

Abstract: An analysis of item banks for the development of an item-bank framework for Swedish compulsory school

This report compares and analyzes a selection of item banks with regard to the proposed purposes of a future item bank for Swedish compulsory school. Several purposes are presented, but the formative purpose, with its potential to support schools and teachers in their efforts to help the students reach the national goals, is seen as the most important. The item bank can realize this purpose by making the goals clearer and more specific. Furthermore, an item bank can aid the students in reaching the national goals by providing high-quality tasks; assessing students' strengths and weaknesses. An item bank can also contribute to increased agreement between different assessors, with regard to student fulfilment of the national goals, by showing how the national goals can be assessed. This, however, might be restricted to the tasks provided by the item bank. On condition that the proposed item bank includes resources for developing teachers' assessment competence in a more general way, both formative and summative purposes for classroom assessment can be favoured. That the item bank could indeed have this function is shown by the analysis of international experiences.

Anders Jönsson, Doctoral Student, Malmö University
anders.jonsson@mah.se

Gunilla Svingby, Professor Emeritus, Malmö University
gunilla.svingby@mah.se

Förord

Denna rapport är skriven på uppdrag av Skolverket. Uppdraget har inneburit att ta fram ett underlag för att formulera kraven på en svensk nationell provbank för grundskolan, där kraven ska kunna användas för att utveckla provbankens innehåll och dess tekniska lösning. Uppdraget har således varken varit att utforma själva ramverket, eller uppgifterna i provbanken, utan att ta fram underlag för detta.

Föreliggande text presenterar bakgrunden till rapporten (inklusive teoretiska utgångspunkter), en jämförelse av nationella och internationella provbanker samt vilka slutsatser som dras med utgångspunkt från jämförelsen. Däremot presenteras det empiriska underlaget, och konkreta förslag för utformningen av provbanken, endast i begränsad omfattning. Dessa delar återfinns emellertid i originalrapporten, som finns publicerad på Skolverket hemsida (www.skolverket.se).

Anders Jönsson & Gunilla Svingby

Underlag till ramverk för en provbank i grundskolan

Anders Jönsson & Gunilla Svingby

Bedömning är den starkaste faktorn för att påverka inriktningen av elevers ansträngningar (Black, 1998; Eisner, 1985). Bra bedömning utvecklar därför undervisningens kvalitet och elevers lärande. Föreliggande text redovisar resultatet av uppdraget att ge ett forskningsunderlag för det kommande ramverket för en nationell provbank för grundskolan. Regeringen uppdrog 2004 åt Skolverket att ”vidareutveckla det nationella provsystemet och då prioritera att bygga ut provbanken för grundskolan med avseende på ämnen eller ämnesgrupper för såväl de tidigare som de senare åren i grundskolan” (Utbildningsdepartementet, 2004, bilaga 3). Uppdraget gäller således en provbank för hela grundskolan, det vill säga både de skolår då betygs sätts, för nuvarande år 8 och 9, och övriga skolår. Det gäller vidare alla ämnen och inte enbart de fyra som nu omfattas av nationella prov. Innebörden i detta är att provbanken ska vara en resurs för lärare inte bara för betygssättning, utan också för att förbättra elevernas lärande och därmed deras uppnående av målen.

Bakgrunden till regeringens uppdrag till Skolverket kan sökas i förändringen av den politiska styrningen av skolan under det senaste decenniet, vilken bland annat inneburit ökat lokalt ansvar. Genom decentraliseringen av skolan har tidigare styrning med regler ersatts av styrning via mål och resultat. En viktig del i bakgrunden är också det vidgade kunskapsbegrepp, som introduceras i förarbetet till Lpo 94 (och Lpf 94), och som har påverkat läroplanens mål. Det innebär, att målen för elevernas utveckling är mer komplexa och dynamiska än tidigare och dessutom formulerade både som mål att sträva mot och som mål att uppnå.

Av betydelse är vidare Skolverkets slutsats, att lärares betygssättning inte har sådan kvalitet att alla elever kan garanteras en rättvis och likvärdig bedömning och betygssättning (Skolverket, 2004a). Myndigheten för skolutveckling gör i samma fråga bedömningen, att läget visserligen har förbättrats, men att ”skillnaderna i lärares uttolkning och tillämpning (är) alltför stora för att kraven på en likvärdig och rättvis betygssättning ska kunna anses vara uppfyllda” (Myndigheten för skolutveckling 2004, sid. 4). Forskning visar också att lärares professionella repertoar när det gäller bedömning är för begränsad. Den behöver utvecklas och breddas (Svingby, 1998). Regeringen instämmer i Skolverkets slutsats, att ett väl utvecklat provsystem bör kunna stödja skolan och lärarna i ansträngningen att hjälpa eleverna att nå målen (Utbildningsdepartementet, 2004). Uppdraget att utveckla en provbank för grundskolan är i konsekvens med detta.

Syftet med föreliggande artikel

Syftet med föreliggande artikel var att ta fram underlag för att formulera kraven på en svensk nationell provbank för grundskolan med utgångspunkt i: (1) regeringens uppdrag, (2) mål i läroplan och kursplaner, (3) teoretisk och

empirisk prov- och bedömningsforskning samt (4) nationella och internationella erfarenheter av provbanker. Kraven ska kunna användas för att utveckla provbankens innehåll och dess tekniska lösning. Uppdraget var således varken att utforma själva ramverket, eller uppgifterna i provbanken, utan att ta fram underlag för detta.

För att nå detta syfte analyseras internationella och nationella erfarenheter av provbanker, provbankens del i uppfyllandet av provsystemets syften prövas samt olika modeller för den konkreta utformningen av provbanken diskuteras. Det finns därvid en hel rad frågor som måste övervägas, till exempel följande:

- Sambandet provbanken – lärares bedömning
På vilka olika sätt kan lärare komma att använda en provbank?
- Provbanken och lärares kompetens
Vad innebär en provbank för lärares kompetens med avseende på bedömning?
- Provbankens omfattning i förhållande till innehåll och typ av bedömning
Hur ”heltäckande” ska provbanken vara med avseende på innehåll, när den startar – och som målsättning i framtiden? Eftersom målen är generella kan ett visst mål ibland nås via olika ämnesinnehåll. Hur bred ska provbanken vara med avseende på olika sätt att bedöma? Bedömning kan dels variera med avseende på den typ av kunskap som efterfrågas, men också med avseende på om prestationen är individuell eller innebär samarbete, om den är skriftlig, muntlig eller manuell, om prestationen görs vid ett specifikt tillfälle eller gäller under en längre period och så vidare. Ska provbanken fokusera på sådant innehåll och sådana uppgiftstyper, som är svåra och/eller inte vanliga vid bedömning i klassrummet?
- Bedömning av uppgifter och prov
Ska uppgifter/prov förses med bedömningsanvisningar?
- Organisation
Ska provbanken byggas upp och underhållas av ”provbanksexpertis” eller kan lärare/blivande lärare bidra till banken?
Ska innehållet vara statiskt, det vill säga att uppgifter utvecklas vid ett tillfälle och sedan får de ”hålla sig” ett antal år – eller dynamisk, det vill säga att uppgifter tillkommer ständigt?
- Teknik
Vilka tekniska möjligheter ska provbanken erbjuda lärare: söksystem, registrering av uppgifter och resultat och så vidare?

Frågorna handlar om vilka problem som provbanken förhoppningsvis ska lösa, och om vilka problem som provbanken faktiskt kan lösa. Svaren på dessa, och andra, frågor ger en utgångspunkt för att ställa krav på en svensk nationell provbank och de får därmed konsekvenser för utformningen av en sådan.

Det nationella provsystemet för grundskolan

Provbanken ska utvecklas inom ramen för det nationella provsystemet. För närvarande omfattar detta för grundskolan, nationella ämnesprov och diagnostiska material. I Hem- och konsumentkunskap finns dessutom prov i en provbank. Utbildningsdepartementet (2004) anger flera olika syften för provsystemet. Det ska:

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

Det är rimligt att anta att ordningsföljden mellan de olika syften, som formulerats av Utbildningsdepartementet är avsiktlig. De tre första satserna är tydligt inriktade mot provsystemet, som ett stöd till undervisningen och elevernas lärande, vilket är i linje med den generella inriktningen att provsystemet ska stödja skolan och lärarna i ansträngningen att hjälpa eleverna att nå målen.

Det nationella provsystemets första uppgift är således att bidra till måluppfyllelse, det vill säga till att eleverna i ökad omfattning når målen. Den andra uppgiften är att ge elever, föräldrar och lärare information om vad eleven är bra på, samt vad eleven behöver förbättra för att uppfylla målen. I den andra uppgiften ingår att provsystemet ska förtydliga läroplanens och kursplanernas generella målformuleringar. Detta krav ställs också i den tredje satsen med tillägget att konkretisera betygskriterier, vilket rimligen gäller endast för de skolår då betyg sätts. Först den fjärde punkten ställer kravet att provsystemet ska utformas för att stödja bedömning och betygssättning, vilket är rimligt med tanke på att denna uppgift genomförs i en mindre del av grundskolans nio skolår. Den fjärde punkten avser underlag för analys av i vilken utsträckning kunskapsmålen nås på skolnivå, på huvudmannanivå och på nationell nivå, det vill säga underlag för en utvärdering av skolan i förhållande till dess mål.

Det finns vissa oklarheter i formuleringen av syftena. I texten talas till exempel om mål på tre olika sätt: målen, kursmål och kunskapsmålen. Vi uppfattar texten som att de tre uttrycken betecknar samma sak, det vill säga läroplanens och kursplanernas mål. Uttrycket "kunskapsmål" uppfattar vi, i linje med förarbetena till Lpo 94, att det innefattar samtliga aspekter av kunskap (t.ex. förtrogenhet, färdighet, faktakunskap och förståelse) och inte enbart traditionella faktakunskaper.

Uttrycken att förtydliga målen respektive att konkretisera kursmål kan bereda svårigheter. En första tolkning innebär, att de två uttrycken avser samma sak. Rimligen förtydligas målen i och med att de konkretiseras. Det är kanske möjligt att förtydliga målen utan att konkretisera dem. Tolkat så innebär den andra satsen en insnävning i förhållande till den första satsen, till exempel genom att konkretiseringen innehåller ett visst innehåll.

När det gäller nationella provsystemets förmåga att stödja en likvärdig och rättvis betygsättning, analyseras detta i Skolverkets utredning "Det nationella provsystemet - vad, varför och varthän?", som presenterades 2003. Enligt utredningen skulle de nationella ämnesproven kunna fungera betygsstödjande på följande sätt (a.a., sid. 112):

- som mått på den enskilda elevens måluppfyllelse,
- som mått på den egna gruppens grad av måluppfyllelse,
- som förebildligt stödmaterial för lärare och lärarlag.

Utredningsgruppen framhåller problem med samtliga tre funktioner och man drar slutsatsen, att de nationella ämnesproven inte är bra verktyg för att avgöra om enskilda elever har uppnått målen. Bland skälen nämns att provmiljön inte ger alla individer samma förutsättningar att visa sitt kunnande, det otydliga sätt som målen formuleras i styrdokumentet, att proven inte förmår fånga väsentliga egenskaper hos eleverna, till exempel att på ett kreativt sätt utnyttja olika hjälpmedel (och inte bara den tillåtna formelsamlingen eller lexikonet), eller att med hjälp av andra människor söka sig fram till svar (a.a., sid. 140).

Provens betygsstödjande funktion är oklar och svårförståelig för lärarna dels eftersom de nationella ämnesproven av en rad skäl, till exempel begränsningar i tid och format, lokala innehållsliga val och omständigheter vid provtillfället, inte kan omfatta samtliga aspekter av målen. Det innebär att proven inte till fullo kan uppfylla syftet att konkretisera kursmål och betygsriterier (a.a., sid. 16-18). Ett annat problem med avseende på förebildlighet är att varken prov, elevlösningar eller bedömningsanvisningar på grund av sekretessreglerna kan publiceras. Det innebär, att lärare och elever inte kan få de kommenterade autentiska elevlösningar, som skulle kunna fungera som stöd, trots att detta är ett av de nationella ämnesprovets syften.

Skolverkets utredning nämner inte att bidra till ökad måluppfyllelse för eleverna, som ett syfte för provsystemet, vilket kan förklaras av att utredningen framför allt förhåller sig till provdelen av provsystemet, det vill säga de nationella ämnesproven för grundskolan respektive kursproven för gymnasieskolan. För grundskolan gäller utredningen därmed i stor utsträckning främst de ämnen som omfattas av ämnesprov och de betygsgivande åren i grundskolan, medan övriga skolår behandlas i liten omfattning.

Med stöd i utredningen kan konstateras, att den planerade nationella provbanken för grundskolan gäller för samtliga ämnen och för alla skolår. Den måste således genom hela grundskolan inriktas mot att stödja lärarnas ansträngningar att hjälpa eleverna att nå målen, bland annat genom att beskriva hur elevernas kunskaper utvecklas i förhållande till målen. Det innefattar kravet, att provbanken på ett mer fullständigt sätt konkretiserar och exemplifierar målen med hjälp av ett brett utbud av olika slags bedömningsuppgifter, och till dessa hörande bedömningsanvisningar (inklusive exempel på elevsvar). För de ämnen som inte har nationella ämnesprov krävs också, att provbanken visar hur betygsriterierna kan operationaliseras och bedömas. Dessa olika krav är inte antagonistiska. En provbank, som visar på olika sätt att bedöma ett mål, kan samtidigt stödja elevernas lärande och vara till stöd vid betygsättning genom att konkretisera

läroplansmålen via exempel på elevsvar och genom olika slag av provuppgifter. Självfallet kan en väl utformad provbank, som fokuserar på elevens kontinuerliga utveckling och prestationer, också stödja bedömningen i de skolår där betygssättning inte sker.

Teoretiska utgångspunkter

Syftet med detta avsnitt är att formulera utgångspunkter för vårt urval av provbanker och för analysen av de valda provbankerna. Bedömning av elever kan göras med olika syften. Det kan handla om att kontrollera elevers kunskaper i förhållande till uppsatta mål, till exempel för att ange att eleven har klarat en gräns eller för att rangordna elever sinsemellan. Bedömning används också för att ”disciplinera” elever och för att få dem att anstränga sig. Med bedömning kan man vidare rikta in elevers ansträngningar. Effekterna av bedömning för dessa syften beror på hur viktigt det är för eleven att klara målet eller att få bra betyg. Studier visar, att för en del av eleverna är dessa syften inte viktiga, vilket betyder att bedömning med denna typ av inriktning inte fungerar särskilt väl (Svingby, 1998).

Syftet med ”konventionella prov” (Korp, 2003) är oftast just att kontrollera i vilken utsträckning eleven har vissa bestämda kunskaper i slutet av ett kursavsnitt. Detta kan betecknas som summativ kontroll. En sådan bedömning utförs inte i första hand för att bidra till att eleven utvecklas mot målen, utan istället för att kontrollera i vilken utsträckning eleven faktiskt har uppnått dessa. Kännetecknande för många summativa prov är att de resulterar i någon form av meritvärde, till exempel en poängsumma eller ett betyg. Dessa är oftast inte möjliga att direkt relatera till de kunskaper eller kunskapsbrister, som eleven uppvisat. Bristen på relevant feedback gör det svårt för eleven att lära från provresultatet. Summativa prov i skolan kan vara lärarkonstruerade eller externa (som ämnesproven inom provsystemet). I det senare fallet dikteras villkoren för proven av provkonstruktören. Bedömningen görs oftast av läraren, vilken vid externa prov tilldelas någon form av bedömnings-anvisningar.

Användningen av konventionella, summativa prov – ofta externt utformade – har ökat internationellt. Aktuell forskning kring bedömning har studerat hur den utbredda användningen av sådana prov har påverkat undervisningen och elevernas lärandestrategier och kunskapssyn. Man har med stor enighet konstaterat, att denna provtradition dels fört med sig en likriktning av undervisningen och dels lett till ytinriktat lärande och reproducerande lärandestrategier hos eleverna. Andra forskare har visat att den här typen av prov missgynnar vissa grupper av elever. Resultaten har lett till en omfattande kritik mot konventionella prov och andra traditionella bedömningar, dels eftersom flera kvaliteter som värderas högt i läroplaner och kursplaner inte kunnat mätas med proven (t.ex. metakognitiva processer, handling i autentiska situationer, divergent tänkande), dels för att proven ger upphov till stress samt påverkar självförtroende och motivation negativt hos elever som ofta misslyckas (Menucha Birenbaum m.fl., 2006; Caroline Gipps, 2001; Korp, 2003; Shepard, 2002).

En förändrad kunskapssyn

Parallellt med kritiken av den traditionella bedömningstraditionen har intresset tilltagit för alternativa sätt att se på prov och bedömning (Shepard, 2002). Ett skäl är förändringen i sättet att betrakta kunskap och lärande som inträffat under senare år, och som ofta benämns "social-konstruktivistisk" eller "sociokulturell" syn på kunskap och lärande (Wells, 1999).

Denna kunskapssyn, vilken bland annat uttrycks i skolans styrdokument, beskriver kunskap som sociala konstruktioner, vilka hjälper oss att förstå vår omvärld och att skapa mening i denna (Skolverket, 1996). Termen "social-konstruktivistisk" syftar på att all mänsklig kunskap är konstruerad i den meningen att den varken finns "där ute" för oss att upptäcka, eller inom oss i en opåverkad och objektivt tillägnad form. Personliga, såväl som vetenskapliga, kunskaper ses istället som sociala och kulturella konstruktioner. Till skillnad från synen på kunskap som något objektivt och direkt överförbart, från exempelvis läraren till en mer eller mindre passiv elev, är lärande en menings- och kunskapsskapande process; ett aktivt förlopp där den lärande måste vara involverad tillsammans med andra och/eller tillsammans med kulturella verktyg och där tidigare erfarenheter spelar en stor roll för hur och vad vi lär oss.

Utifrån kunskapssynen, kan man härleda konsekvenser för bedömningspraktiken; både för varför och vad man bedömer samt vilka metoder man använder, vilket presenteras nedan.

Formativ bedömning

Syftet med en formativ bedömning är inte att i första hand att kontrollera i vilken grad eleven nått fram till målet, utan på vilket sätt eleven klarar av olika delar av målet i fråga. Detta för att kunna klargöra för läraren och för eleven på vilka punkter, eller inom vilka områden, eleven har sin styrka, samt var hon/han behöver förbättra sig. Den formativa bedömningen kan också helt inriktas mot att hjälpa eleven att utveckla kunskap om sitt eget lärande. Enligt detta synsätt är bedömningen ett verktyg, bland flera, i elevens utveckling mot målen och ett viktigt hjälpmedel för elevens lärande.

Formativa bedömningar kan se ut på olika sätt. De kan till exempel konstrueras och genomföras på samma sätt som summativa bedömningar, det vill säga som ett konventionellt prov - och naturligtvis rikta sig mot samma mål. Skillnaden i förhållande till den summativa bedömningen är att eleven får relevant feedback utifrån sina prestationer.

Formativ bedömning kan också skilja sig från konventionella prov. Elevens arbete kan till exempel samlas in under en längre period och utvecklingen analyseras i förhållande till målet (s.k. portföljbedömning). Det är inte heller nödvändigt att uppgiften som bedöms genomförs av en enda elev, utan den kan även utföras av en grupp elever, tillsammans med en vuxen eller med stöd av olika typer av hjälpmedel (s.k. dynamisk bedömning), eller under en rad andra betingelser (se t.ex. Black & Dylan Wiliam, 1998, Wynne Harlen & Ruth Deakin Crick, 2003; James W. Popham 2002; Shepard, 1991). Bedömningens syfte är då inte att se vad eleven kan utföra ensam och utan stöd, utan att visa vad hon/han kan åstadkomma under mer optimala förhållanden. Man ansluter då till Lev Vygotskijs (1978) klassiska teori om bedömning som ett sätt att utveckla

eleven (zonen för proximal utveckling). Bedömningssituationen kan därmed utnyttjas som ett tillfälle att lära (Gipps, 2001; Wells, 1999; 2001).

Den formativa bedömningen är effektiv i detta avseende. I flera studier av effekten av formativ bedömning förbättras elevernas genomsnittliga studieresultat generellt och den största förbättringen gäller de svagpresterande eleverna (Black & William, 1998).

Feedback

Den sekretess som omgärdar många summativa bedömningar innebär, att eleverna inte ges möjlighet att få relevant feedback utifrån sina prestationer, och feedback är en av de mest centrala komponenterna i formativ bedömning (Black & William, 1998; Sadler, 1989; Wiggins, 1998). För att feedback ska vara effektiv för eleven krävs att hon/han: (1) vet vilka målen och bedömningskriterierna är, (2) har förmågan att jämföra sin egen prestation med målen samt (3) har möjligheten att så att säga "minska avståndet" mellan målen och sin prestation (Sadler, 1989).

Feedback kan vara av olika slag. Pat Tunstall och Caroline Gipps (1996) delar, med utgångspunkt i en studie av lärares återkoppling till eleverna, in feedback i två kategorier: värderande och beskrivande. Medan "värderande feedback" är generellt dömande, till exempel som allmänt beröm eller allmänt negativa utlåtanden, är "beskrivande feedback" tydligt specifik och uppgiftsrelaterad. Denna feedback kallas ibland just "uppgiftsrelaterad feedback".

Distinktionen är viktig, eftersom flera empiriska studier har visat att "värderande feedback" antingen inte har någon effekt alls på lärandet eller till och med ger negativa effekter. "Beskrivande eller uppgiftsrelaterad feedback" tenderar däremot att ge positiva effekter på lärandet (Black & William, 1998; Avraham N. Kluger & Angelo DeNisi, 1996).

Bedömningsanvisningar

Väl utformade bedömningsanvisningar ger förutsättningar för formativ bedömning bland annat genom att underlätta för läraren att ge feedback, samtidigt som sådana anvisningar kan understödja likvärdig bedömning (Jönsson & Svingby, 2008). Bedömningsanvisningar kan dock se mycket olika ut och utformningen beror till stor del på vilken typ av kunskaper som ska bedömas. En bedömningsanvisning som består i ett facit till räkne- eller flervalsuppgifter, vilka bara har ett rätt svar, behöver inte vara särskilt utförlig. Men om man ska bedöma komplexa uppgifter och kunskapskvaliteter, behövs mer utförliga anvisningar, till exempel en bedömningsmatris. En bedömningsmatris (eng. "scoring rubric") är en uppsättning kriterier för vad som ska bedömas tillsammans med en kvalitetsskala för respektive kriterium (Arter & McTighe, 2001).

En viktig poäng med bedömningsmatriser är, att de bidrar till öppenhet och tydlighet i bedömningen. Med matrisen tvingas läraren bli explicit med vad som ska bedömas, vilket dels ger eleverna möjlighet att koncentrera sig på det som anses viktigt, men också möjlighet att lära sig hur bedömningen går till och i förlängningen att bedöma sin egen förmåga. För att elever ska kunna förstå och använda matrisen krävs att den är relativt enkel och tydlig.

Kriterierna i en bedömningsmatris ska gälla en handling eller en produkt, som visar att eleven behärskar det som ska bedömas enligt målet. Kriterier som endast gäller abstrakt förståelse eller kunskap räcker inte.

Att dela upp bedömningsmatrisen i flera separata aspekter som bedöms separat (t.ex. innehåll, stavning, grammatik i en uppsats) ger en så kallad analytisk matris (Arter & McTighe, 2001). Analytiska matriser skiljer sig från holistiska matriser, vilka ger ett samlingsomdöme för prestationen som helhet.

Forskning visar, att analytiska matriser oftast ger en bättre tillförlitlighet än holistiska. Tillförlitligheten vid användning av både holistiska och analytiska matriser kan ökas genom träning av bedömarna samt om dessa har tillgång till exempel på bedömda elevlösningar. Ett annat sätt att förbättra tillförlitligheten är genom ökad detaljspecificering (Jönsson & Svingby, 2008). Nackdelarna med en allt för detaljerad matris är emellertid påtagliga. För många kriterier och för detaljerad beskrivning av dessa minskar användbarheten. Om man till exempel bedömer elever medan de agerar i klassrummet, är det svårt att hantera en allt för komplex matris.

En annan invändning mot detaljspecificerande bedömningsanvisningar har förts fram av Eisner (1985; 1991), som menar att sådana hämmar kreativiteten hos både elever och lärare. Det gäller i särskilt hög grad när elevernas prestationer bedöms utifrån på förhand specificerade och noggrant definierade mål eller kvantitativa nivåbeskrivningar (eng. "standards"). Vid denna typ av bedömningsanvisningar läggs stor vikt vid jämförbarhet, objektivitet och rättvisa och att eleverna får exakt information från början om vad de ska lära sig. Eisner kritiserar detta och menar, att kunskapen behandlas instrumentellt och riskerar leda till en mekanisk undervisning, som begränsar möjligheterna till improvisation och kreativitet. Dessa eventuella negativa effekter har självfallet att göra med hur bedömningsanvisningarna utformas. Risken för alltför detaljerad styrning blir mindre, när matrisen utformas för att bedöma övergripande kunskapskvaliteter istället för detaljer eller kvantitativa mått (Wiggins, 1998). Försök med lärarstuderande antyder att möjligheterna för elever att själva medverka i utformningen av matrisen förefaller ha påtagligt positiva effekter både på elevernas acceptans av kriterierna och på deras resultat (Dysthe & Engelsen, 2003; Engelsen, 2006).

Komplexa uppgifter och processbedömning

För att fånga det breda spektrum av kunskaper och färdigheter, som uttrycks i skolans styrdokument, däribland elevernas förmåga att lösa komplexa problem, resonera kritiskt och kunna använda sina kunskaper i autentiska situationer, krävs att arsenalen av bedömningsinstrument breddas. Detta gäller både summativ och formativ bedömning. Framför allt måste metoder utvecklas och användas för att bedöma komplexa och mångfacetterade kunskapskvaliteter. I summativa sammanhang är så kallade "slutna" uppgifter vanliga. Det handlar om, att inte bara problemet (eller frågan) redan är formulerat, utan också om att det bara finns "ett rätt svar" och ofta också bara "en rätt väg" att komma fram till lösningen. Vi kan tala om endimensionella uppgifter och uppgifter "utan frihetsgrader" (Eisner, 1985). Många forskare argumenterar för användningen av "öppna" uppgifter (Wiggins, 1998). Med "öppna" menas att det finns stor frihet i hur eleven kan lösa uppgiften, men också möjligheter för eleven att själv formulera

problemet och mer än ett bra och riktigt svar. Samuel Messick (1996) beskriver en gradient från starkt strukturerade bedömningar, som flervalfrågor, till mer ”öppna” bedömningsformer som muntliga presentationer eller portföljbedömningar, vilka i engelskspråkig litteratur går under beteckningen ”performance assessments”. ”Performance assessment” innebär, enligt Messick, att eleven måste utföra eller skapa något under så pass lång tid att antingen processen, produkten eller båda kan bedömas. För att bedömningen ska ha kunna stödja elevernas lärande är det av vikt att bedömningen omfattar även själva processen och inte bara slutprodukten (Engelsen, 2006; Shepard, 2002).

Shepard (2002) ger exempel på några ”alternativa” (d.v.s. icke-konventionella) bedömningsmetoder för bedömning andra kunskapsformer än fakta. Hon förespråkar olika observationsbaserade metoder, som kan användas i undervisningen och intervjuer för att förstå hur eleven resonerar och vilka strategier hon/han använder sig av. Intervjun kan också lätt göras ”dynamisk”, vilket innebär att eleven får hjälp av läraren att visa vad hon/han kan under mer gynnsamma betingelser. Portföljbedömning kan användas för att på ett naturligt sätt dokumentera vad eleven presterat, samtidigt som metoden erbjuder ett underlag för gemensam granskning av elevens prestationer och självreflektion. Denna bedömningsmetod kan således vara ett viktigt instrument för att utveckla metakognitiva förmågor.

En slutsats är, att lärare som regel bör använda flera olika metoder och vid varje tillfälle välja den metod, som bäst fångar de kunskaper man avser bedöma. Ur rättvisesynpunkt borde eleverna också ges möjlighet att demonstrera sina kunskaper på det sätt som de själva anser bäst visar vad de kan (Engelsen, 2006; Shepard, 2002).

Kunskapssyn och krav på provbanken

Läroplanens kunskapssyn är tillsammans med uppdragstexten en viktig utgångspunkt för uppbyggnaden av ett ramverk för provbanken. Skolans mål innehåller för grundskolans ämnen sammantaget ett spektrum av kunskaper inom relativt breda innehållsliga områden. Ett gemensamt mål för alla ämnen är förmåga att fungera i autentiska och komplexa sammanhang. Grundskolans ramar innebär vidare att skolan lokalt har uppdraget att precisera och anpassa målen till lokala omständigheter och enskilda elever. Organisationen innebär ett stort ansvar för skolans lärare och ledning att formulera lokala mål och till dessa anpassad bedömning. En grundläggande utgångspunkt är att bedömningen ska göras med utgångspunkt i läroplanens mål och utifrån dess kunskapssyn.

Skolverkets kvalitetsgranskning från 2000 gav en antydning om brister i förhållande mellan vad som undervisas och vad som bedöms. Man har också visat, att en del lärare inte uppfattat skillnaden i kunskapssyn mellan tidigare och nuvarande betygssystem (Selghed, 2004), men också att lärare, som säger sig vilja mäta förståelse ändå i realiteten bedömer relativt enkla minneskunskaper (Svingby, 1998). Resultaten antyder en brist både på verktyg och kompetens för att bedöma mer komplexa kunskapskvaliteter och elevkompetenser. Bakom ligger även en motsättning mellan de psykometriskt orienterade ideal, som förknippades med det tidigare relativa betygssystemet, och den kvalitativa kunskapssyn med det målrelaterade betygssystem, som dagens läroplan företräder. En ovilja att acceptera en

mindre mätinriktad bedömning av elevernas kunskaper, som framkommer i studier kring de nationella proven i grundskolan, kan vara uttryck för en sådan konflikt. Flera lärare uttrycker kritik mot de muntliga delproven, i vilka bedömningen anses vara alltför subjektiv (Skolverket, 2004a).

Bedömningen i klassrummet har brister, när det gäller systematik, medvetenhet, öppenhet och tydlighet med avseende på vad bedömningen gäller, och hur olika prestationer värderas. Möjligheterna för eleven att förstå vad bedömningen gäller, och vad som värderas, är också otillräcklig (t.ex. Cederberg, 2006). Bristerna gäller framförallt bredden i bedömningsformer. För lite av den moderna bedömningsrepertoaren används. Problemet tycks också gälla hur bedömningen motsvarar läroplanens och kursplanernas mål. Många av grundskolans mål är komplexa och svåra att bedöma, och även om man kan se innovativa exempel på bedömning, saknas bra modeller för bedömning av dessa mål i klassrummet. Den svenska forskningen inom området har vidare en blygsam omfattning och undervisningen på lärarhögskolorna outvecklad, vilket innebär att också lärarnas kompetens är dåligt utvecklad (Högskoleverket, 2005). Samtidigt har det i anslutning till de nationella utvärderingarna för samtliga grundskolans ämnen utvecklats både en kompetens och en reservoar av uppgifter för år 9 och i någon mån också för det femte skolåret. Denna kompetens är spridd över landet och kopplad till lärarutbildningen och borde självfallet kunna användas både vid uppbyggnaden av och underhållet av provbanken.

På provbanken kan ställas kravet, att den ska bidra till att lösa de problem och svårigheter avseende bedömning i klassrummet, som visats i forskningsrapporter och på annat sätt.

Exempel på provbanker

I det följande presenteras några exempel på prov- och uppgiftsbanker. Dessa exempel har valts för att de representerar delvis olika lösningar på ”problemet” provbank. Vi presenterar således inte en översikt över fältet. Provbankerna har valts från ett större antal, som vi företrädesvis har funnit genom sökning via Internet, huvudsakligen med hjälp av sökmotorn Google. Som söktermer har använts olika kombinationer av ”item banks”, ”assessment/test delivery systems” och ”assessment/test design systems”. Sökningen innefattar endast engelskspråkig litteratur och sökresultatet har givit en tydlig överrepresentation av källor från USA. Potentiellt intressanta provbanker, som publicerats på andra språk, finns således inte representerade. Andra sökmetoder krävs för att finna dem. Skolverkets provbanksmaterial för gymnasieskolan, och provbanken för hem- och konsumentkunskap, kompletterar de internationella exemplen.

I samband med urvalet har även andra begränsningar gjorts. Så har till exempel endast internetbaserade lösningar valts. Skälet är att kapaciteten hos dagens datorer, och tillgången på höghastighetsuppkoppling mot Internet, ger goda möjligheter att skapa prov- och uppgiftsbanker, som kan lagra, hantera och distribuera komplexa data på ett enkelt sätt (Quellmalz & Haertel, 2004). Ett undantag är provsystemet GSCE, som är helt pappersbaserat.

En annan avgränsning är att de exempel, som presenteras, uteslutande är hämtade från icke-kommersiella projekt. Även om det finns en stor mängd kommersiella provbanker, ger dessa en ytterst begränsad insyn i hur banken

är uppbyggd, samt vilket forsknings- eller utvecklingsarbete som bedrivits. Detta begränsar möjligheterna att diskutera respektive provbanks för- och nackdelar.

Det är värt att notera att majoriteten av de provbanker som presenteras, gäller naturvetenskap och matematik. Detta grundar sig inte i ett medvetet urval, utan är en följd av utbudet. Det är den inriktning, som förefaller ha prioriterats, när provbanker konstruerats.

I det följande görs korta presentationer av provbankerna, och kommentarer ges med avseende på deras möjlighet att uppfylla de syften som formulerats för den svenska provbanken. För mer utförliga beskrivningar, inklusive exempel på uppgifter och bedömningsanvisningar, finns i originalrapporten till Skolverket (Skolverket, 2007).

Principled Assessment Designs for Inquiry – PADI

PADI är en tekniskt avancerad provbank, som väver samman modern teori för pedagogiska mätningar med naturvetenskapligt didaktisk forskning. Ett gemensamt teoretiskt ramverk garanterar likvärdig bedömning genom att det innehåll bedömningen riktar sig mot är noga specificerat, samtidigt som bedömningen görs efter en gemensam och jämförbar skala (Cromley & Mislevy, 2004; Hamel & Schank, 2005; Mislevy, Haertel & Murphy, 2005; Quellmalz & Haertel, 2004). Det faktum att ämnesdidaktisk forskning inarbetats i uppgiftsmallarna, i form av exempelvis elevers vardagsföreställningar kring naturvetenskapliga fenomen, torde ge goda förutsättningar för diagnosticering av elevernas kunskaper. Genom att använda provuppgifter från ett internationellt utbildningsprojekt, där eleverna gör vetenskapliga mätningar i vatten, luft och jord (GLOBE), visar man att PADI-systemet, trots hög grad av standardisering, inte är begränsat till lättbedömda uppgifter, utan att systemet kan hantera även mer autentiska uppgifter (Haydel DeBarger, Yumoto & Quellmalz, 2005).

Ytterligare ett exempel på samverkan mellan PADI och andra projekt, är samarbetet med Full Option Science System (FOSS). Här visas PADI-systemets potential för vad som ofta kallas ”computerized adaptive testing”, eller CAT. Det innebär att systemet skattar elevens förmåga utifrån ett förtest och anpassar efterföljande uppgifter efter detta. CAT kan dels spara tid, men också reducera den frustration som det innebär att behöva svara på frågor som ligger antingen för högt eller för lågt i förhållande till den egna förmågan (Kennedy, Timms, Long, Ketchner & Bluestein, 2005; Timms & Kennedy, 2005).

Ett problem med CAT är i allmänhet att eleven inte har någon möjlighet att gå tillbaka och förbättra sina svar (Urbina, 2004). FOSS-modellen har löst detta problem genom att dels innehålla ett bedömningsläge, där eleven själv kan välja när hon/han känner sig redo att testa sina kunskaper och dels erbjuda ett övningsläge, där eleven kan få textbaserad feedback. Själva bedömningsarbetet är dock dolt för både elev och lärare. Elevens svar skickas till en rättningsserver, som i sin tur skickar tillbaka resultat som en form av kunskapshistorik (Kennedy m.fl., 2005). Det innebär emellertid att läraren inte har någon möjlighet att se och förstå det bedömningsarbete som utförs.

Secondary Mathematics Assessment and Resource Database (SMARD)

Syftet med SMARD, som lades ner i februari 2006, var att underlätta för lärarna att ta del av varandras uppgifter. Det fanns inga ambitioner för att bedömningen av uppgifterna i banken skulle vara likvärdig. Uppgifterna varierade också med avseende på om de åtföljdes av bedömningsanvisningar. Även då sådana anvisningar förekom var de av högst skiftande utformning. Det fanns heller inga kvalitetskrav på de uppgifter, som togs in i databasen (SMARD hemsida, 2006-01-20). Erfarenheterna från SMARD pekar på problemen med en provbank som helt och hållet bygger på lärares bidrag.

Performance Assessment Links in Science (PALS)

Att det finns ett samband mellan mål och bedömning är en förutsättning för måluppfyllelse. PALS tydliggör detta samband genom så kallade "assessment charts", en sorts matriser som läraren kan använda för att kontrollera överensstämmelsen mellan mål och de uppgifter hon/han valt för sin bedömning. PALS innehåller dessutom förebildliga uppgifter, som åtföljs av tydliga bedömningsanvisningar och mättekniska data (Quellmalz m.fl., 2000). Sammantaget ger PALS-provbanken goda förutsättningar för en likvärdig bedömning och ökad måluppfyllelse, men även möjligheter att fånga elevernas starka och svaga sidor med stöd av bedömningsanvisningarna.

PALS har lagt stor vikt vid användarvänlighet och ger lärare stöd för att anpassa uppgifterna efter lokala behov. Man erbjuder också resurser för att användarna ska kunna lära sig mer om kritiska moment i konstruktionen av bedömningsuppgifter av hög kvalitet (Quellmalz m.fl., 2000), vilket torde vara essentiellt för att få till stånd positiva spridningseffekter utöver de bedömningsuppgifter, som hämtas eller modifieras via databasen.

The Assessment Design and Delivery System (ADDS)

ADDS främsta syfte är att lära lärarna mer om bedömning, för att på så sätt generellt höja kvaliteten på lärarnas bedömningar. Avsikten är således att överskrida effekten av en kvalitetshöjning, som är begränsad till de uppgifter som erbjuds via provbanken. Man har därvid skapat en provbank, där bedömningens "tools of the trade" är inbyggda i systemet. Användarna lär sig genom att de måste ta ställning till moment, som anses vara kritiska för att skapa bedömningsuppgifter av hög kvalitet. Sålunda måste användaren specificera vilka kunskaper bedömningen avser innan själva uppgiften skapas. På samma sätt måste användaren först formulera bedömningsanvisningar innan han/hon kan skriva instruktionen till uppgiften (Vendlinski, Niemi & Wang, i.å.). Detta fokus på mål, kunskapsformer och bedömningsanvisningar torde ge möjligheter till ökad likvärdighet, måluppfyllelse samt diagnostisering av elevernas kunskaper.

Genom att databasen innehåller en viss mängd expertkonstruerade uppgifter, har man skapat förutsättningar för användaren att utnyttja förebildliga exempel. Genom att ge användarna möjlighet att bidra till provbanken (inklusive någon form av kvalitetsgranskning; Vendlinski m.fl., i.å.), finns dessutom förutsättningar för att lärarna känner ett ägandeskap och

att databasen byggs upp på ett mer dynamiskt sätt jämfört med en provbank som skapas enbart av experter.

Regional Examination Boards for 13 – 16 GCSE

Syftet med GCSE-systemet var för det första att utveckla bra prov, som konkretiserade ämnesmålen både med avseende på innehåll och olika typer av prestationer. I detta inkluderades syftet att med hjälp av proven bedöma elever på ett, inom regionen, likvärdigt sätt. Intressanta aspekter av detta är dels den mycket stora täckningen av innehållet i kurserna, och dels möjligheten för eleven att själv välja innehåll att visa sina kunskaper på. Systemet syftade vidare till att höja lärarnas bedömningskompetens genom förebildlighet och direkta utbildningsinsatser. De utförliga bedömningsanvisningarna med kommenterade elevsvar var ett sätt att förverkliga detta syfte, liksom de många kringmaterialen, läroböckerna och det kontinuerliga samarbetet mellan lärare och olika slags experter vid konstruktion och utprovning av uppgifter. Systemet producerade en oerhört stor mängd uppgifter. Även om målförankringen och bedömningskriterierna är kvalitetshöjande, skiftar kvaliteten självfallet.

Frågetecken kan säkerligen också ställas i förhållande till den mät-tekniska kvaliteten med avseende på likvärdighet. Systemet är ett av få, som också innehåller prov inom de samhällsorienterande ämnena.

Grundskolans provbank i Hem- och konsumentkunskap

Provbanken för Hem- och konsumentkunskap innehåller få, men förebildliga uppgifter, som konkretiserar målen och är tydligt i linje med den kunskaps-syn, som uttrycks i styrdokumentet. Uppgifterna är autentiska, processinriktade och förankrade i aktuella vardags- och samhällsfrågor. De inbjuder dessutom till ett aktivt deltagande i bedömningen från elevens sida. Ett uttalat formativt syfte genomsyrar upplägget. Uppgifterna skulle med fördel kunna användas i fler skolår än i år 9, som anvisningarna anger (de Ron & Feldt, 2006; Provbanken i hem- och konsumentkunskap, 2006). Sammantaget motsvarar provbanken för Hem- och konsumentkunskap flera av de krav, som framhålls i detta utkast till ramverk.

Provbankens tre uppgifter är vidare intressanta genom att de så tydligt är inriktade mot bedömning av elevers handlande. Kriterierna riktar in bedömningen på processer, där läraren ska iaktta både elevers handlande och deras samtal om sitt handlande – både individuellt och tillsammans andra. Kriterierna är relativt vaga, vilket förutsätter kreativitet från lärarens sida. För att bedömningen ska kunna användas som reflektionsunderlag för eleverna, krävs då också att bedömningsunderlaget från lärarens sida görs explicit. Ett problem med kriterierna är, att de inte bildar en gradient för de olika nivåerna utan är relativt fristående. Ett annat problem är det faktum att bedömningsanvisningarna inte ges till eleverna förrän efter det att de avslutat uppgiften (Provbanken i hem- och konsumentkunskap, 2006), vilket försvårar för eleverna att lära sig av denna.

Vad gäller syftet att stötta lärarnas betygssättning i ämnet kan man möjligen ställa sig mer tveksam, då det inte förekommer något uttalat stöd

för hur sammanvägningsprocessen ska gå till¹⁴. Kriterierna i bedömningsanvisningarna innehåller inga nivåbeskrivningar och det erbjuds heller inte några bedömda elevsvar, som kan underlätta en likvärdig bedömning. Provbanken ger dock goda möjligheter för ökad måluppfyllelse, samt även möjligheter att fånga elevernas starka och svaga sidor med stöd av bedömningsanvisningarna.

Skolverkets provbank för gymnasieskolan

Utformningen av prov och uppgifter inom provbanken för ämneskurser påminner om de nationella proven. Uppgifterna är konstruerade av experter och håller hög kvalitet. Till alla uppgifter finns antingen facit eller mer utförliga bedömningsanvisningar, och i vissa fall även bedömningsmatriser och bedömda elevlösningar, där anvisningarna generellt sett ger goda förutsättningar för en likvärdig bedömning. I de fall det finns mer utförliga och beskrivande anvisningar, kan dessa användas i formativt syfte, men de flesta uppgifterna och bedömningsanvisningarna är utformade för summativ bedömning. Större delen av provmaterialet gäller skriftliga uttrycksformer, men det förekommer också experimentella uppgifter (fysik) samt uppgifter där eleverna ska förhålla sig muntligt till inspelat tal (moderna språk) (Provbanken i fysik, 2006; Provbanken i moderna språk, 2006).

Möjligheten för provbanken *att fungera förebildligt* handlar om hur väl proven representerar målen innehållsligt och med avseende på olika kunskapsstyper. Syftet att öka elevernas möjligheter att nå målen – ökad måluppfyllelse – kräver att provbanken erbjuder verktyg för att uppmärksamma och pröva ett bredare spektrum av målen. I och med att fysikprovbanken innehåller anvisningar för bedömning av laborationer (Provbanken i fysik, 2006), kan man hävda att den representerar kursplanens hela bredd.

Det stöd som provbanken erbjuder för lärares egen provkonstruktion gäller främst förändringar av provets sammansättning, ordningsföljd eller poängsättning, medan det saknas stöd för att modifiera själva uppgifterna eller bedömningsanvisningarna. Provbanken i fysik ger till exempel instruktioner till lärarna om hur provbankens klassificeringssystem kan användas för egen provkonstruktion. Även om denna resurs innehåller en del generella överväganden, som överensstämmelsen mellan mål och provets syfte, är den i första hand inriktad mot hur användaren kan utnyttja uppgiftsdatabasen för att skapa välkomponerade skriftliga prov. Med ”välkomponerade” avses att det finns en balans i exempelvis antalet uppgifter, uppgifter av olika karaktär och öppenhet, samt att uppgifterna riktar sig mot olika kognitiva nivåer (Provbanken i fysik, 2006; Provbanken i matematik, 2006).

Provbanken för yrkesprogrammen avviker påtagligt från övriga provbanker för gymnasieskolan. Istället för att bedöma elevernas kunskaper med avseende på ämnesmål, fokuserar provbanken på elevernas demonstrerade yrkeskompetens. Det betyder, att provbanken innehåller uppgifter som ställer eleven inför autentiska, simulerade yrkessituationer och kräver att eleven ska uppfatta problemet i situationen, olika handlings-

¹⁴ Därmed inte sagt att ett sådant stöd är önskvärt, då det bland annat kan ge signaler om att betyg kan sättas med utgångspunkt från enstaka prov. Proven kan då få rollen av slutprov, vilket inte är förenligt med tanken om lokalt friutrymme och deltagande målstyrning (Skolverket, 2003).

möjligheter och välja den bästa. I provmaterialet för yrkesprogrammen läggs fokus i bedömningen främst på elevens förmåga att i komplexa situationer kunna välja ut och använda adekvata kunskaper från olika ämneskurser och inte på bedömning av de enskilda kunskaperna inom respektive ämne (Provbanksprojektet, 2006). Typen av bedömning tillsammans med bedömningens inriktning mot kompetenser och programmål innebär att kvaliteter, som sällan utsätts för bedömning, uppmärksammas och prövas (Skolverket, 2001).

En gemensam kompetensprofil, som kopplas till program- och provspecifika bedömningsmatriser, ger förutsättningar för en *likvärdig bedömning*. Detsamma gäller syftet *diagnos* av elevernas kunskaper. Sambandet mellan mål och bedömning, som är en förutsättning för *måluppfyllelse*, tydliggörs i denna provbank genom sambanden mellan mål, bedömningsmatriser och kompetensprofiler (Provbanksprojektet, 2006). Motsvarande tydlighet vad gäller relationen till de nationella målen finns inte i alla delar av provbanken för ämneskurser.

Provbanken fyller syftet *förebildlighet* genom att konkretisera programmålen. Uppgifterna är utformade för att kunna anpassas till olika önskemål beroende på kursens utformning och lärarens tillämpning av denna (Skolverket, 2005). Det är emellertid, i likhet med provbanken för ämneskurser, osäkert om uppgifterna kan användas som förebilder i den meningen, att lärarna på egen hand kan konstruera likartade uppgifter. Bedömningsmatriserna och kompetensprofilerna kan visserligen användas även i andra bedömningssituationer, men systemet innehåller ingen information om hur man som lärare kan göra för att modifiera befintliga, eller skapa egna, uppgifter eller matriser. Syftet att stödja likvärdighet, måluppfyllelse och diagnos blir därför i princip begränsade till de uppgifter som tillhandahålls i provbanken.

Tekniskt är Skolverkets provbank för yrkesprogrammen en avancerad provbank, som ger stora möjligheter att erbjuda meningsfulla och autentiska provuppgifter. Innehållet i provbanken är indelat efter respektive program och innehåller inga sökfunktioner, vilket dock inte erbjuder några problem för användaren eftersom antalet prov är förhållandevis litet (Provbanksprojektet, 2006).

Samtidigt ställer tekniken krav på att användarna måste ha tillgång till datorer med vissa prestanda och vissa program installerade, vilket i sin tur innebär att lärarna måste ha tillgång till datorsalar, eller motsvarande, vid provtillfället. Eftersom proven kräver tillgång till datorer ställs det större krav på planering samt krav på datorerna vad gäller prestanda och mjukvara (Gerrevall, 2001).

Vid en jämförelse mellan de delar av provbanken, som riktar sig mot ämneskurser, och de delar som riktar sig mot yrkesprogrammen, är det tydligt att provmaterialet för ämneskurser har en mycket starkare betoning på stöd vid betygssättning än materialet för yrkesprogrammen. Det omvända förhållandet gäller för syftet att visa på elevers starka och svaga sidor, där provmaterialet för yrkesprogrammen innehåller kompetensprofiler och bedömningsmatriser, som kan utnyttjas för feedback, men också för att förmedla till eleverna vad som förväntas av dem. Sålunda är matriserna för yrkesämnena uppdelade i separata kriterier med nivåbeskrivningar utifrån betygsstegen, vilket även finns i matematik, men inte i Moderna språk och endast i begränsad omfattning i naturvetenskapliga ämnen. Vidare är

matriserna för yrkesämnena inte knutna till specifika uppgifter, utan snarare till vissa ämnesområden, vilket gör det möjligt för eleverna att själva använda matriserna för att få en uppfattning om vad som förväntas av dem och vad som anses viktigt i mer generell mening än lösningen på enstaka uppgifter. Den här typen av generella kriterier finns även för skriftlig produktion i Moderna språk samt för Biologi A, men det är bara provbanken för yrkesämnen som kombinerar kvalitativa nivåbeskrivningar med generella kriterier (Provbanken i biologi, 2006; Provbanken i fysik, 2006; Provbanken i matematik, 2006; Provbanken i moderna språk, 2006; Provbanksprojektet, 2006).

Provbanksmaterialet för yrkesämnena betonar således, i motsats till de nationella proven, utöver summativa också bedömningens formativa funktioner, det vill säga uppgiften att hjälpa eleverna att upptäcka, och förbättra, sina starka och svaga sidor. Att ge förebilder för lärares konstruktion av egna prov beskrivs som en angelägen uppgift för alla delarna i provsystemet för gymnasieskolan.

Den svenska provbanken för gymnasieskolan tycks skilja sig från de övriga provbanker, som presenterats här, genom att i högre grad än dessa lägga vikt vid likvärdighet och i lägre grad betona stödet till lärare. Till skillnad från flera av de internationella exempel som presenterats, saknar den svenska provbanken för gymnasieskolan således som uttalat syfte att stödja lärares kompetensutveckling.

En jämförelse mellan provbanken för gymnasieskolans ämneskurser och den planerade provbanken för grundskolan visar på viktiga skillnader. När det gäller att täcka hela bredden av innehåll och kunskapsformer tycks uppgiften avsevärt svårare för grundskolans ämnen. Varje ämnes mål är så påtagligt bredare i grundskolan än i gymnasieskolan, att problemet att tillhandahålla uppgifter för alla lokala preciseringar knappast är möjlig. Som jämförelse kan vi se på fysikämnet. Kursplanerna i Fysik A på gymnasiet innehåller främst formuleringar av typen ”ha kunskap om” (vilket visserligen kan diskuteras vad detta innebär) och ”ha kännedom om”, medan grundskolans fysikämne även har formuleringar av typen ”kunna använda argument” och ”kunna föra diskussioner” (Skolverkets kursinformations-system, 2006). Man kan se det som att måloppfyllelse för gymnasiefysiken mer handlar om att täcka hela bredden av ämnesinnehållet, medan det i grundskolan i större utsträckning krävs att man också täcker bredden av olika kunskaps- och uttrycksformer.

En slutsats blir, att eftersom provbanken för grundskolan i flertalet ämnen knappast kan tillhandahålla tillräckligt antal uppgifter för alla lokala behov (jämför grundskolans provbank för Hem- och konsumentkunskap), behövs ett system med stöd för lärares egen provkonstruktion.

Jämförelse av provbankerna

I jämförelsen av provbankerna tas de frågor som ställdes i inledningen upp, men också frågor som har aktualiserats av genomgången av provbankerna. Inledningsvis ställdes frågan om provbankens omfattning i förhållande till innehåll och mål. Att det finns en koppling mellan mål och bedömning, är en grundläggande förutsättning för måloppfyllelse. Det är också en förutsättning för rättvis bedömning och för ett rättvist betyg (Skolverket, 2004c).

Vad gäller de exempel som givits från USA, har man i flera av dessa provbanker begränsat sig till vissa kunskapsformer och uttryckssätt, och därmed inte gjort anspråk på att företräda hela den bredd av målformuleringar, som kan tänkas finnas lokalt. PADI och PALS inriktar sig således i första hand på uppgifter med anknytning till naturvetenskapligt arbetssätt, medan uppgifterna i ADDS främst hanterar centrala teorier inom naturvetenskapen. Men även med denna begränsning har man inom exempelvis PALS haft svårigheter att representera allt det ämnesinnehåll som ingått i intentionerna, till exempel naturvetenskapens karaktär (Quellmalz m.fl., 2000).

Detta är ett centralt problem för en provbank för grundskolan, då det i många fall finns en större bredd av kunskapsformer och uttrycksformer representerade i målen för grundskolan jämfört med exempelvis gymnasieskolan. Om bedömningen inte stämmer med målen, eller bara med en del av målen, blir bedömningen inte rättvis, men bristande överensstämmelse får också andra konsekvenser. Eftersom bedömning styr (vissa) elevers uppmärksamhet och studiestrategier, kan en bedömning som riktas mot minneskunskaper få som konsekvens att elever lär sig utantill, även om målen framhåller mer komplexa förmågor. PALS-projektet erbjuder flera fördelar i detta sammanhang. Den mest uppenbara är att banken tydliggör kopplingen mellan mål och bedömning.

PADI-systemet, med dess inriktning mot komplexa uppgifter av undersökande och problemlösande karaktär, har också potentialen att konkretisera kursmål och betygskriterier samt visa på starka och svaga sidor hos eleverna. Den senare funktionen stärks dessutom av, att man bygger in resultat från didaktisk forskning i sina så kallade ”uppgiftsmallar”. Till skillnad från de övriga system som beskrivits, har uppgifterna i PADI en gemensam stomme för hur uppgifterna utformas och bedöms, vilket dels innebär att de kan modifieras, lagras och distribueras per automatik, dels att bedömningen mellan olika uppgifter och olika bedömare blir jämförbar. Dessa uppgifter kan således ingå även i standardiserade sammanhang. En uppgiftsbank uppbyggd med hjälp av ett sådant system skulle därför sannolikt kunna kopplas till såväl nationell utvärdering som kontroll av elevernas måloppfyllelse. PADI-systemets komplexitet är också dess svaghet, när det gäller förebildlighet. Om lärarna inte förstår hur likvärdigheten åstadkommit eller vad den innebär, kan man inte få några spridningseffekter till lärares egen provkonstruktion.

Vid utformningen av den planerade provbanken behöver denna punkt uppmärksammas.

Urval och täckning

En provbank kommer troligen inte att kunna täcka alla de ämnesområden, som omfattas av grundskolans undervisning. PALS-projektet, som enbart gällde naturvetenskapliga ämnen, hade således cirka 200 uppgifter i sin databas efter tre år. Trots ambitioner om motsatsen hade projektets uppgifter en tyngdpunkt i traditionellt väl tillgodosedda områden, medan exempelvis uppgifter med anknytning till personliga eller sociala perspektiv förblev underrepresenterade (Quellmalz m.fl., 2000).

Om urvalet av uppgifter i en statligt sanktionerad provbank är begränsat, kan det på gott och ont ge signaler om vilka mål och kriterier, eller inne-

hållsliga områden, som bör prioriteras i undervisningen. Då det skett en förskjutning från faktakunskaper och enkla färdigheter mot mer övergripande förmågor och förhållningssätt i skolans styrdokument (Lars Lindström, 2005), borde det vara rimligt att komplexa uppgifter kommer att finnas i den planerade provbanken för den svenska grundskolan. I en bank, där uppgifterna inte är avsedda att användas i standardiserade prov, finns heller inga hinder för att inkludera komplexa uppgifter. (Därmed inte sagt att det vore omöjligt att standardisera komplexa uppgifter, vilket visas bland annat genom PADI-projektet). De internationella provbanker som beskrivs här innehåller också en stor andel uppgifter av denna typ.

En nationell provbank kan även ha som funktion att komplettera försummade eller eftersatta områden. Detta gäller exempelvis de övergripande målen, vilka tycks bli eftersatta till förmån för uppnåendemål och betygskriterier i samband med bedömning (Skolverket, 2001). Med urvalet av uppgifter kan man lyfta fram mål och innehållsområden samt demonstrera variationer i uppgiftstyper. Provbanken kan på så sätt konkretisera läroplanens mål på liknande sätt som provmaterialet för yrkesprogrammen konkretiserar övergripande program mål.

ADDS systemet är intressant i detta perspektiv. Systemet har medvetet utvecklats med utgångspunkt i forskning kring formativ bedömning och djupinläring, och erbjuder därför stöd åt lärarna för att konstruera uppgifter, som inte följer traditionella bedömningsmönster.

Förebild för lärares egna prov (spridningseffekter)

PALS-projektet satsar på att ge exempel på bra prov och provuppgifter, samt att inspirera lärare att använda uppgifterna som förebilder när de utvecklar egna prov/uppgifter. Lärarna erbjuds således möjligheten att modifiera uppgifterna i provbanken, så att de passar deras syften, alternativt att konstruera egna uppgifter med hjälp av provbanken. PALS erbjuder dessutom användarna möjligheter att lära sig mer om bedömning och om effektiv bedömningsdesign. Lärarna får se exempel på bra respektive dåliga prov och uppgifter, vilka kommenteras. De får också se bra respektive dåliga bedömningsanvisningar och får referenser till artiklar, böcker och rapporter som kan ge ytterligare information och stöd.

Det visade sig i utvärderingar, att lärarna faktiskt använde exemplen för att konstruera egna uppgifter och anvisningar. Lärarna lärde sig också, hur man använder bedömningsmatriser för att förbättra sin bedömning, och erfarenheterna från PALS integrerades i den egna bedömningspraxisen (Quellmalz m.fl., 2000). Upplägget hade således positiva spridningseffekter utöver de uppgifter, som fanns att hämta i själva projektet. Också ADDS riktar in sig på att påverka lärarnas reguljära bedömningar. Även inom detta projekt har man påvisat positiva resultat av provbanken, med avseende på vilka kunskapsformer som lärare prövar i sin egen undervisning, och när det gäller lärares användning av bedömningsmatriser (Lee m fl, 2005).

Bedömningsmatriser

Vid bedömning är det inte bara utformningen av uppgiften som är viktig. Minst lika viktig är bedömningen av elevernas prestationer. Om det finns behov av bättre utformade uppgifter, finns det ännu större behov av bättre,

mer explicita och mer systematiskt utformade bedömningar. Bedömningsmatriser spelar här en central roll. Användningen av matriser är viktigt när det gäller att stödja en likvärdig bedömning. Tidigare forskning har visat att likvärdigheten oftast är låg vid bedömning av komplexa uppgifter (Dunbar, Koretz & Hoover, 1991), men också att olika former av strukturerade bedömningsanvisningar ger en ökad överensstämmelse i bedömningen (Jönsson & Svingby, 2008). Väl genomarbetade bedömnings-anvisningar har även andra förtjänster. Ett genomgående tema i forskning av bedömning med hjälp av bedömningsmatriser är både glädjen av, som nyttan med, öppenhet (eng. "transparency") i bedömningen. Bra bedömningsanvisningar visar eleven vad som är viktigt i undervisningen, och gör därmed bedömningen tydlig, vilket både elever och lärare upplever som positivt. Bedömningsmatriser kan även förbättra lärarnas möjligheter att ge kontinuerlig feedback och förslag på förbättringar av elevernas prestationer (Jönsson & Svingby, 2008). Väl genomarbetade bedömnings-anvisningar kan därför ses som en central komponent vid målrelaterad bedömning, och utformningen av sådana är avgörande för uppfyllandet av de syften, som formulerats för provbanken.

Flera av de internationella systemen använder bedömningsanvisningar. Inom ADDS-systemet läggs stor vikt vid konstruktionen av bedömningsanvisningar, vilket tycks ge gott resultat. En jämförande studie visade till exempel att lärare som arbetat med ADDS, själva utvecklar bedömningsanvisningar för sina prov (Lee m.fl., 2005). ADDS-systemet tycks fylla både syftet att, via ett urval färdiga uppgifter och prov, konkretisera kursmål och betygskriterier, samt syftet att utbilda lärare att själva utveckla uppgifter. Uppgifter som i sin tur leder både till likvärdig bedömning och till möjligheter att diagnostisera eleverna. Också PALS ger stöd för lärares utveckling av bedömningsanvisningar. En annan styrka med PALS är den genomtänkta anpassning av lärares egna uppgifter, som görs i förhållande till hela systemet. Uppgifterna indexerats i förhållande till mål i kursplanen. Man utarbetar bedömningsanvisningar och övningshäften, prövar ut uppgifterna samt förser dem med information om exempelvis interbedömarreliabilitet.

Involvera lärarna

Det finns en rad skäl till att involvera lärarna i provbanken och att på så sätt öka deras bedömningskompetens. Det är troligt att det, för att kunna acceptera och reguljärt använda en nationell provbank för att öka måluppfyllelsen, krävs kompetensutveckling av lärarna (Gerrevall, 2001). Även om lärare överlag är positiva till uppgifterna i grundskolans nationella prov och till gymnasieskolans provbanksmaterial, reagerar vissa lärare negativt på öppna uppgifter, som de anser gör bedömningen alltför subjektiv (Skolverket, 2004a, b). Tidigare studier har också gett en bild av bedömningspraxis i den svenska skolan som tydligt traditionell (Lindqvist, 2003; Skolverket, 2000; Svingby, 1998; Törnvall, 2002).

Ett sätt att engagera lärare i arbetet med en provbank är att låta dem bygga upp denna med hjälp av egna uppgifter, så som gjordes i SMARD. Det finns åtskilliga fördelar med en databas som konstrueras av lärarna själva. En fördel är att risken minskar för att det ska uppstå en klyfta mellan innehållet i provuppgifterna och lärarnas undervisning. Om lärarna själva bygger upp banken krävs inga insatser från experter, vilket gör uppbyggnaden avsevärt billigare. Om syftet inte inbegriper användning av uppgifterna i standar-

diserade prov, krävs heller inte någon omfattande indexering eller utprovning av uppgifterna. Lärarna kategoriserar själva uppgifterna i samband med att de skickar in dem till databasen.

Ett problem med uppgiftsbankar som bygger på lärarnas bidrag är emellertid uppgifternas kvalitet. Det är uppenbart svårt att konstruera bra uppgifter, vilket enligt Kristian Ramstedt (1992) ofta underskattas. Ramstedt konstaterar, att en stor del av de uppgifter som erfarna och intresserade lärare konstruerade till de centrala proven i fysik, kunde inte användas på grund av bristande kvalitet. Erfarenheterna från det svenska provsystemet pekar på vikten av att uppgifterna i en prov- eller uppgiftsbank håller hög kvalitet. En lärarkonstruerad prov- eller uppgiftsbank skulle visserligen vara billigare än en bank konstruerad av experter, men skulle samtidigt riskera låg kvalitet och låg legitimitet hos lärarna.

Användarvänlighet

De internationella provbanker, som granskats är med få undantag tekniskt avancerade. De erbjuder lärarna möjligheter att på ett enkelt sätt söka uppgifter inom de ämnesområden de vill pröva. De erbjuder också möjligheter för läraren att ta en uppgift från banken och ändra den så att den passar hans/hennes speciella situation. Möjligheter att förbättra egna uppgifter och förse dem med bedömningsmatriser finns i vissa banker. Några banker erbjuder läraren möjligheten att utifrån det mål för provet, som läraren vill pröva, kombinera lämpliga uppgifter. Flera banker erbjuder möjligheten för läraren att spara de uppgifter hon/han valt för att på ett enkelt sätt kunna ta fram dem vid ett senare tillfälle. De flesta av systemen förefaller ha utvecklat användarvänlig teknik.

En del provbanker erbjuder både möjligheten att eleverna gör provet elektroniskt eller utanför datorn. Denna flexibilitet uppskattas av användarna. Några av systemen erbjuder ett varierat utbud av bilder, simuleringar, filmer med mera som kontext till uppgifterna. Även om tekniken är relativt lätt att hantera, kräver de flesta provbanker en arbetsinsats från lärarens sida. Detta gäller i särskilt hög grad de system, som har ambitionen att utveckla lärares kompetens.

I en undersökning om provbanken för gymnasieskolan, framkommer att många lärare inte utnyttjar resursen, vilket bland annat beror på tidsbrist (Skolverket, 2005). De internationella erfarenheterna erbjuder inga lösningar på detta problem. De konstaterar tvärtom, att det tar tid för användare att lära sig hantera ett nytt system. Det tar förstås olika lång tid att lära sig hantera ett system, beroende på vilken grad av användarvänlighet det har. Väsentligt är att tidsinsatsen för att lära sig funktionerna måste vara värd investeringen. Detta illustreras av skillnaderna mellan PALS- och PADI-systemen.

Medan PALS har en tydlig inriktning mot användarna, har man inom PADI prioriterat de mättekniska aspekterna, vilket gör systemet svårt att använda för vanliga lärare. I PALS-systemet ges användarna möjlighet att söka uppgifter i databasen både utifrån målformuleringar och utifrån ämnesområde. Sökresultaten visas sedan tillsammans med en tydlig relevansmarkering. Vidare kan användarna recensera de uppgifter de använt, så att andra kan få stöd när de ska välja uppgifter. Om uppgifterna inte passar de lokala målen, eller behöver modifieras, finns det ett interaktivt stöd, som underlättar anpassningen med bevarad kvalitet. De personliga ”assessment

charts” (d.v.s. matriser som indikerar överensstämmelse mellan mål och bedömningsuppgift), som en gång skapats, kan dessutom sparas i en databas. Allt detta gör att användaren på ett enkelt sätt kan utnyttja systemets resurser. Inom PADI-projektet har man istället valt att i så kallade uppgiftsmallar (vilka finns för alla uppgifter) integrera såväl ett teoretiskt ramverk för bedömning, som resultat från didaktisk forskning. Genom att utnyttja dessa mallar kan användaren skapa egna bedömningsuppgifter av hög kvalitet och med hög mätteknisk standard. Problemet är att systemet baseras på ett komplext teoretiskt och tekniskt fundament, vilket i princip är otillgängligt för läraren. Det innebär, att användaren måste ta hjälp av starkt strukturerade ”guider” för att kunna konstruera uppgifter med hjälp av systemet, men även att det förmodligen inte finns några möjligheter för lärare att överblicka eller förstå vad man gör. I det samarbetsprojekt med FOSS som beskrevs tidigare, skickas till exempel elevernas prestationer till en rättningsserver, där underlaget tolkas med utgångspunkt från de statistiska algoritmer som finns inbyggda i uppgiftsmallen, varefter resultatet skickas tillbaka till elev och lärare. Läraren har ingen insyn i hur processen går till, utan får automatiskt färdiga tolkningar/resultat att förhålla sig till. Ett sådant system involverar inte lärarna, utan riskerar att fasa ut läraren från bedömningsarbetet, vilket helt överlämnas åt experter.

Frågan om hur provbanken ska byggas upp och underhållas är relevant här. Samtliga presenterade provbanker, utom SMARD, har byggts upp av experter i nära samarbete med forskare, men i flera av dem finns ändå en involvering av lärare, antingen som i ADDS där lärarna själva utvecklar uppgifter med hjälp av mallar som systemet tillhandahåller, eller som i GCSE där verksamma lärare kontinuerligt samarbetar med olika slag av experter för att ta fram och bedöma uppgifter. Poängen med att kombinera forskare/experters med verksamma lärare är uppenbar. Genom forskarnas medverkan kan provbanken ha hög ”teknisk” kvalitet och genom lärarnas medverkan garanteras användbarhet i klassrummet och intresse för provbanken från lärarnas sida. Ett annat kännetecken för merparten av de studerade provbankerna är deras samtida teoretiska förankring och dynamiska karaktär. Det innebär att med bakgrund i en utvecklad teoretisk ram, gällande bedömning av kunskap generellt och i det specifika ämnet, utvecklas kontinuerligt nya uppgifter och uppgiftsformer. Ett problem för alla provbankerna utom GCSE-systemet, är att alldeles för få uppgifter har utvecklats för att täcka ämnets olika delar. GCSE-systemets modell med flera parallella centra för utveckling av uppgifter utifrån en gemensam teoretisk bas är intressant, eftersom den uppenbarligen resulterade i en mycket stor mängd uppgifter av hög kvalitet.

Slutsatser för utformning av en provbank

Utifrån analysen av provbanker kan man formulera de krav som provbanken ska möta. Kraven utgår dels från den syn på mål, kunskap och bedömning, som fastslås i nationella styrdokument och dels från de problem, som förts fram via forskning om och erfarenheter av bedömningen i skolan. Helt grundläggande är provbankens fokus. Även om olika syften presenteras, framstår syftet att stödja skolan och lärarna i ansträngningen att hjälpa eleverna att nå målen, som det viktigaste syftet för provbanken. Som

genomgången har visat kan detta syfte realiseras på olika sätt, men centralt är att banken förtydligar och i och med detta konkretiserar målen. Under förutsättningen att undervisningen stöder eleverna, kan provbanken även bidra till ökad måluppfyllelse genom att visa på elevernas starka och svaga sidor.

På liknande sätt som för måluppfyllelsen kan provbanken, via förebildlighet, bidra till ökad likvärdighet. Uppgifter som exemplifierar målen och som visar på olika sätt att uttrycka och bedöma dessa, bör leda till större likhet mellan lärares bedömning av om/hur väl elever uppnår målen. Detta gäller troligen endast för de uppgifter som ingår i banken. Under förutsättning att provbanken också satsar på att utveckla lärares bedömningskompetens, kan emellertid ökad måluppfyllelse och likvärdighet gynnas mer generellt, samtidigt som en kvalitetshöjning kan ske av lärarnas bedömningar. Att ett sådant utökad syfte är en möjlig uppgift för en provbank framgår av de internationella erfarenheterna.

I linje med ovanstående resonemang, föreslås utformningen av provbanken utgå från följande huvudsyften:

1. Resursen ska erbjuda förebildliga exempel på hur målen i läroplan och kursplaner kan konkretiseras och bedömas.
2. Resursen ska tydliggöra elevers starka och svaga sidor, med ett tydligt fokus på bedömningens formativa funktioner i förhållande till läroplanens och kursplanens mål.
3. Resursen ska stödja likvärdig och rättvis bedömning i förhållande till målen i läroplanen och respektive kursplan.
4. Resursen ska bidra till utvecklingen av lärarnas bedömningskompetens.

Avslutningsvis ges ett urval av slutsatser för utformningen av provbanken. Ett mer utförligt förslag presenteras i originalrapporten till Skolverket (Skolverket, 2007).

Uppgifternas innehåll och utformning

En ökad måluppfyllelse kräver både att det som bedöms är relevant i förhållande till målen, samt att bedömningen gäller hela den bredd av mål som uttrycks i styrdokumentet. Uppgifterna i bedömningsresursen bör därför erbjuda en stor variation av bedömningsformer. Konkretisering av mål (och i förekommande fall betygskriterier) med hjälp av uppgifter med stor variation, leder till större möjligheter för lärare att bedöma om eleverna har nått målen, och om de är på väg att nå dessa.

Också innehållsligt bör bedömningsresursen eftersträva bredd. Det är inte troligt, att resursen kommer att kunna erbjuda uppgifter som representerar alla tänkbara operationaliseringar av målen. Det är emellertid rimligt, att resursen ger exempel på bedömning inom skilda innehåll förankrade i läroplan och kursplaner. Det är vidare rimligt, att bedömningsresursen lägger stor vikt vid sådant innehåll, som antas vara svårt och tidskrävande för lärare att bedöma, liksom kunnande som sällan behandlas och bedöms.

Bedömningsanvisningar

För samtliga ovan angivna syften är bedömningsanvisningar av mycket stor betydelse. Banken bör därför innehålla bedömningsanvisningar med explicit förankring i mål tillsammans med kommenterade elevsvar.

Organisation

För att skapa en bedömningsresurs med hög kvalitet och som har legitimitet hos användarna, bör resursen i någon mening vara expertkonstruerad. Samtidigt är det viktigt att användarna har någon form av inflytande över uppgifterna i databasen, både för att skapa känsla av ägandeskap, men också för att gynna en mer dynamisk uppbyggnad av resursen. En uppgiftsbank för grundskolans alla ämnen kan svårigen omfatta alla tillämpningar av målen. Istället bör databasen byggas ut successivt för att omfatta fler och fler aspekter på målen, där de uppgifter som finns i banken, ses som exemplifieringar av hur målen kan konkretiseras på skilda innehåll och i olika sammanhang.

Användning

För att ge möjlighet till flexibel användning utan att i allt för stor utsträckning styra format eller ämnesinnehåll, bör bedömningsresursens databas bestå av enskilda uppgifter och inte av färdiga prov. Samtidigt måste uppgiftsbanken vara användarvänlig för att inbjuda till användning, trots den tidsinsats som krävs för att lära sig systemet.

Kategorisering av uppgifter

För att lärare ska kunna söka efter uppgifter i databasen, måste uppgifterna kategoriseras enligt några på förhand definierade kriterier. Dessa kategorier bör till en del vara bekanta för användaren, som ett led i resursens användarvänlighet, men kan även kompletteras med andra kategorier, som bedöms relevanta för att stödja provbankens utbildande syfte.

Utbildningsresurser

För att stötta bedömningsresursens utbildande funktion föreslås resurser av två slag: dels de som likt ADDS är "inbyggda" i systemet och därmed är mer eller mindre osynliga och tvingande för användarna att utnyttja, dels de utbildningsresurser som likt PALS erbjuds i form av valbara alternativ. Vissa grundläggande faktorer, som behövs för att garantera bedömningens kvalitet, byggs in i systemet, medan fördjupning och ytterligare exempel erbjuds som valbara alternativ.

Tekniska data/Kvalitetskriterier

Ett sätt att rapportera kvaliteten för en uppgift i en uppgiftsbank, är så kallad interbedömarreliabilitet. Ibland rapporteras även någon form av svarsfrekvenser, till exempel hur stor andel som klarade respektive uppgift vid utprovningen.

Olika kvalitetskriterier kan emellertid vara olika viktiga beroende på bedömningens syfte. Interbedömarreliabilitet är exempelvis framför allt av vikt när resultaten har ett stort inflytande på individernas liv (s.k. "high-stakes assessment") och där slumpen därmed inte kan tillåtas ha så stort inflytande, medan svarsfrekvenser är av intresse när man behöver sortera individer för någon form av urval. När huvudsyftet med bedömningen är att stötta lärandet, kan dock andra kriterier än reliabilitet och spridning vara till större nytta (t.ex. upplevd autencitet eller meningsfullhet; se Baartman,

Bastiaens, Kirschner & van der Vleuten, 2006). Andra kriterier, som att eleverna vet vad som förväntas av dem, är däremot viktiga oavsett syftet med bedömningen.

Referenslista

- Arter, Judith & McTighe, Jay (2001). *Scoring rubrics in the classroom*. Thousand Oaks: Corwin Press, Inc.
- Baartman, Liesbeth K. J., Bastiaens, Theo J., Kirschner, Paul A. & Van der Vleuten, Cees P. M. (2006). The wheel of competency assessment: Presenting quality criteria for Competency Assessment Programmes. *Studies in Educational Evaluation*, 33, 153-177.
- Birenbaum, Menucha, Breuer, Klaus, Cascallar, Eduardo, Dochy, Filip, Dori, Yehudit, Ridgway, Jim, Wiesemes, Rolf & Nickmans, Goele (2006). A learning integrated assessment system. *Educational Research Review*, 1, 61-67.
- Black, Paul (1998). *Testing: Friend or foe?* London: Falmer Press.
- Black, Paul & Wiliam, Dylan (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7-74.
- Cederberg, Meta (2006). *Utifrån sett - inifrån upplevt: Några unga kvinnor som kom till Sverige i tonåren och deras möte med den svenska skolan*. Doktorsavhandling, Malmö högskola.
- Cromley, Jennifer G. & Mislevy, Robert J. (2004). *Task templates based on misconception research (No. 646)*. Los Angeles: Center for the Study of Evaluation (CSE).
- de Ron, Lena & Feldt, Maria (2006). *Bedöma och lära i hem- och konsumentkunskap*. Stockholm: PRIM-gruppen.
- Dunbar, Stephen B., Koretz, Daniel M. & Hoover, H. D. (1991). Quality control in the development and use of performance assessments. *Applied measurement in education*, 4, 289-303.
- Dysthe, Olga & Engelsen, Knut Steinar (Red.) (2003). *Mapper som pedagogisk redskap. Perspektiver og erfaringer*. Oslo: Abstrakt Forlag.
- Eisner, Elliot (1985). *The Art of Educational Evaluation: A personal View*. London & Philadelphia: The Falmer Press.
- Eisner, Elliot (1991). Taking a Second Look: Educational Connoisseurship Revisited. I Milbrey W. McLaughlin & D. C. Phillips (Red.), *Evaluation and education at quarter century* (ss. 169-187). Chicago: The National Society for the Study of Education.
- Engelsen, Knut Steinar (2006). *Gjennom fokustrengsel. Laererutdanning i møte med IKT og nye vurderingsfomer*. Doktorsavhandling, Universitetet i Bergen.
- Gerrevall, Per (2001). Generella kompetenser bedöms via gymnasieskolans nya yrkesprov. I Skolverket *Villkor och vägar för grundläggande yrkesutbildning* (ss. 187-204). Stockholm: Skolverket.
- Gipps, Caroline (2001). Sociocultural aspects of assessment. I Gunilla Svingby & Sofia Svingby (Red.), *Bedömning av kunskap och kompetens* (ss. 15-67). Stockholm: Lärarhögskolan i Stockholm, PRIM-gruppen.
- Hamel, Lawrence & Schank, Patricia (2005). A wizard for padi assessment design, *AERA Annual Conference*. Montreal, Canada.
- Harlen, Wynne & Deakin Crick, Ruth (2003). Testing and Motivation for Learning. *Assessment in Education: Principles, Policy & Practice*, 10, 169-207.

- Haydel DeBarger, Angela, Yumoto, Futoshi & Quellmalz, Edys (2005). Using PADI templates as an alternative structure for specifying GLOBE investigation strategies, *AERA Annual Conference*. Montreal, Canada.
- Högskoleverket (2005). *Utvärdering av den nya lärarutbildningen vid svenska universitet och högskolor*. Stockholm: Högskoleverket.
- Jönsson, Anders & Svingby, Gunilla (2008). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144.
- Kennedy, Cathleen, Timms, Michael, Long, Kathy, Ketchner, Susan & Bluestein, Dan (2005). Using padi to develop an online self-assessment system that supports student science learning in foss, *AERA Annual Conference*. Montreal, Canada.
- Kluger, Avraham N. & DeNisi, Angelo (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254-284.
- Korp, Helena (2003). *Kunskapsbedömning – hur, vad och varför*. Stockholm: Myndigheten för skolutveckling.
- Lee, John J., Vendlinski, Terry & Niemi, David (2005). Computer-guided assessment design, *AERA Annual Conference*. Montreal, Canada.
- Lindqvist, Susanna (2003). *Elevers uppfattningar och upplevelser av bedömning i matematik i skolår 5*. Stockholm: PRIM-gruppen.
- Lindström, Lars (2005). Pedagogisk bedömning. I Lars Lindström & Viveca Lindberg (Red.), *Pedagogisk bedömning: Om att dokumentera, bedöma och utveckla kunskap* (ss. 11-27). Stockholm: HLS förlag.
- Messick, Samuel (1996). Validity of performance assessments. I Gary W. Phillips (Red.), *Technical issues in large-scale performance assessment* (ss. 1-18). Washington, DC: National Center for Education Statistics.
- Mislevy, Robert J., Haertel, Geneva D. & Murphy, Robert (2005). Overview of the PADI Design System, *AERA Annual Conference*. Montreal, Canada.
- Myndigheten för skolutveckling (2004). *Kompetensutveckling av lärare om bedömning och betygssättning*. Stockholm: Myndigheten för skolutveckling.
- Popham, James W. (2002). *Classroom assessment: what teachers need to know*. Boston: Allyn and Bacon.
- Provbanken i biologi. Umeå universitet, 2006-05-22 från <http://www.umu.se/edmeas/provbank/pbbi/index.html>.
- Provbanken i fysik. Umeå universitet, 2006-05-22 från <http://pb-fy.edmeas.napb.se/>
- Provbanken i matematik. Umeå universitet, 2006-05-22 från <http://pb-ma.edmeas.napb.se/>
- Provbanken i moderna språk. Göteborgs universitet, 2006-05-22 från <http://www.ped.gu.se/sol/nafs/nafs.htm>.
- Provbanken i hem- och konsumentkunskap. Skolverket, 2006-10-13 från <https://provbanken.skolverket.se/pbhc/site/index.html>
- Provbanksprojektet. Skolverket, 2006-03-23 från <http://194.47.65.140/>.
- Quellmalz, Edys & Haertel, Geneva D. (2004). *Use of technology-supported tools for large-scale science assessment: Implications for assessment practice and policy at the state level*. Committee on Test Design for K-12 Science Achievement, Center for Education, National Research Council.

- Quellmalz, Edys, Hinojosa, Thomas, Hinojosa, L. & Schank, Patricia (2000). *Performance assessment links in science (pals): An on-line resource library*. Arlington: National Science Foundation.
- Ramstedt, Kristian (1992). *Något om uppgiftsbanker*. Umeå: Avdelningen för pedagogiska mätningar.
- Sadler, Royce D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Selghed, Bengt (2004). *Ännu icke godkänt. Lärares sätt att erfara betygssystemet och dess tillämpning i yrkesutövningen*. Doktorsavhandling, Malmö högskola.
- Shepard, Lorrie A. (1991). Will National Tests Improve Student Learning? *Phi Delta Kappan*, 73, 232-239.
- Shepard, Lorrie A. (2002). The role of classroom assessment in teaching and learning. I Virginia Richardson (Red.), *Handbook of Research on Teaching*, 4th ed. (ss. 1066-1101). Washington DC: American Educational Research Association.
- Skolverket (1996). *Bildning och kunskap. Särtryck ur Skola för bildning*. Stockholm: Skolverket.
- Skolverket (2000). *Nationella kvalitetsgranskningar 2000: Rapport nr 190*. Stockholm: Skolverket.
- Skolverket (2001). *Att organisera kunskap – om skolans kunskapsuppdrag i teorin, i praktiken och i framtiden*. Stockholm: Skolverket / Liber.
- Skolverket (2003). *Det nationella provsystemet – vad, varför och varthän?* Stockholm: Skolverket.
- Skolverket (2004a). *Skolverkets lägesbedömning 2004 av förskoleverksamhet, skolbarnsomsorg, skola och vuxenutbildning: Rapport 249*. Stockholm: Skolverket.
- Skolverket (2004b). *Prövostenar i praktiken: Grundskolans nationella provsystem i ljuset av användares synpunkter*. Stockholm: Skolverket.
- Skolverket (2004c). *Handlingsplan för en rättssäker och likvärdig betygssättning*. Stockholm: Skolverket.
- Skolverket. (2005). *Skolverkets provbank. Hur används den och vad tycker användarna?* Stockholm: Skolverket.
- Skolverket (2007). *Underlag till ramverk för en provbank/bedömningsresurs i grundskolan*. Stockholm: Skolverket.
- Skolverkets kursinformationssystem. Skolverket, 2006-03-23 från <http://www3.skolverket.se/ki03/front.aspx>.
- SMARD hemsida. Central Queensland University, 2006-01-20 från <http://smard.cqu.edu.au/>.
- Svingby, Gunilla (1998). *Proven, kunskapen och undervisningen: Rapport nr 138*. Stockholm: Skolverket.
- Timms, Michael & Kennedy, Cathleen (2005). FOSS: Using PADI to Develop an Online Self-Assessment System that Supports Student Science Learning, *AERA Annual Conference*. Montreal, Canada.
- Tunstall, Pat & Gipps, Caroline (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal*, 22, 389-405.
- Törnvall, Maj (2002). *Uppfatningar och upplevelser av bedömning i grundskolan*. Licentiatavhandling, Malmö Högskola.
- Urbina, Susana (2004). *Essentials of psychological testing*. New Jersey: Wiley.

Utbildningsdepartementet (2004). *Uppdrag till statens skolverk avseende det nationella provsystemet: Regeringsbeslut U2004/5293/S*. Stockholm: Utbildningsdepartementet.

Vendlinski, Terry, Niemi, David, & Wang, Jia (inget årtal). *Learning assessment by designing assessments: An on-line formative assessment design tool*. University of California, Los Angeles, 2006-01-20 från <http://adds.cse.ucla.edu/reports.htm>.

Vygotsky, Lev S. (1978). *Mind in society*. Cambridge: Harvard University Press.

Wells, Gordon (1999). *Dialogic Inquiry: Towards a sociocultural practice and theory of education*. New York: Cambridge University Press.

Wells, Gordon (Red.) (2001). *Action, talk and text: learning and teaching through inquiry*. New York: Teachers College Press.

Wiggins, Grant (1998). *Educative assessment*. San Francisco: Jossey-Bass.