

Bedöma eller döma?

Språkbedömning och lästest i grundskolans tidigare år

ÅSA WEDIN

Akademien för humaniora, utbildning och samhällsvetenskap
Örebro universitet

Krav på mätbara resultat inom skola och utbildningsväsende ställs allt oftare från politiskt håll inte bara i Sverige och Norden utan över hela västvärlden. Detta gäller i synnerhet färdigheter i läsning och skrivning. Publicering av internationella undersökningar genomförda inom ramen för PISA (Programme for International Student Assessment) och PIRLS (Progress in Reading Literacy Study) följs ofta av nationella krav på högre prestationer, vilka förväntas resultera i bättre utfall vid kommande mätningar. I Sverige har språkundervisningen i grundskolans tidigare år traditionellt fokuserat läsning. Detta fokus har i skolan ofta haft en stark koppling till test och diagnosticering. Sällan har bedömningsmetoderna i sig ifrågasatts eller testen som sådana varit föremål för analys, utan fokus har varit på de elever som lyckats mindre bra i testen och som har uppnått låga resultat.

Vid analys av bedömningar av olika slag vad gäller läsförmåga, är det viktigt att problematisera den syn på läsning och skrivning som ligger inbäddad i dem. Eftersom test och bedömningar har en tendens att reglera vad som händer i klassrummen, genom att lärare ser till att eleverna får öva just det som testas, och eftersom sortering av eleverna utifrån uppfattad läsförmåga ofta görs på grundval av bedömningarna, exempelvis för urval till specialundervisning, är det viktigt att själva testen granskas kritiskt.

Bedömningsexamen sänder signaler inte bara vad gäller vilken typ av kunskap och färdigheter som värderas, utan också om vilken typ av praxis i samband med skrift, litteracitetspraxis, som värderas. På så vis representerar bedömningar ideologiska konstruktioner med en viss syn på kunskap, lärande och samhälle och kan inte studeras som isolerade, neutrala element, utan måste tolkas som inbäddade i olika utbildningsmässiga, sociala, politiska, historiska och ekonomiska sammanhang.

I en artikel om nyinförda nationella litteracitetstest i Danmark poängterar Holm (2009) att analys av bedömningar kan göras på olika nivåer, på själva

testet, på tillämpningen på institutionell nivå och på dess effekter. När det gäller själva testet är det naturligtvis viktigt att analysera vilka kunskaper och färdigheter som i själva verket bedöms, men också vilken kunskapssyn bedömningen baseras på samt vilken typ av litteracitetspraxis som värderas, det vill säga vilka olika sätt att använda skrift och vilka olika perspektiv på skrift som ges värde genom bedömningen.

Läsbedömningar tenderar att värdera den typ av användning av och förståelse av skrift som är vanlig i skolmiljö men att negligera andra typer av praxis som förekommer i övriga situationer eleverna är involverade i och medskapa till. Ett aktuellt sådant exempel är olika typer av litteracitetspraxis som används i elektronisk kommunikation, där skrift ofta kombineras med ljud och bild, både rörlig och stilla, och där användandet ofta samtidigt innefattar både läsande och skrivande, via skärm, mus, tangentbord, penna och papper och där detta inte sällan är kollektiva aktiviteter, långt från det individuella läsande som testen vanligen består av.

När det gäller tillämpningen på institutionell nivå är det viktigt att analysera vilken funktion bedömningen ges i skolan som institution. Exempelvis huruvida bedömningen används för att äska medel eller för att välja ut elever för specialundervisning av olika slag. När det gäller den tredje aspekten, bedömningens effekter, kan analysen enligt Holm (2009) behandla tre nivåer; individnivå, institutionell och social nivå. Han betonar vikten av att aspekter som makt, etik och demokrati beaktas i sådana analyser.

Maktaspekten vad gäller bedömningen belyses särskilt av Shohamy (2001) som hävdar att auktoriteter ofta frestas att använda bedömningen som ett verktyg för att påverka beteende utan att själva riskera att bli utmanade. Enligt Shohamy är bedömning bland annat »an act of demonstrating authority» (Shohamy 2001 s 57). Hon betonar också vikten av att alla inblandade, exempelvis elever och föräldrar i skolsammanhang, ges möjligheter att påverka bedömningarnas utformning. Dessutom framhåller hon att det måste finnas en möjlighet att vägra att bli bedömd och att bedömningens format och innehåll bör vara accepterat av elevernas föräldrar. Hon menar att en dialog mellan de som utför bedömningen och de som blir bedömda är nödvändig av demokratiska och etiska skäl.

Därför kan inte en analys av läsbedömningar behandla endast den tekniska funktionen av bedömningen och frågan om huruvida eleverna lyckas eller inte, utan måste anta ett vidgat perspektiv så att underliggande antaganden i bedömningen och de olika typer av praxis som det resulterar i synliggörs. I denna artikel kommer ett test som används i den svenska grundskolans tidigare år att diskuteras, lästestet »Vilken bild är rätt?» (Lundberg 2001). Testet kommer att jämföras med Skolverkets språkbedömningsverktyg »Nya språket lyfter» (Skolverket 2008). I artikeln kommer testen att relateras till frågor om makt, etik och demokrati. En central fråga kommer att vara: *Vem har makt över bedömningen och med vilken legitimitet?*

TEST OCH BEDÖMNING

I ett historiskt perspektiv kan vi se att bedömningar av den typ vi möter idag med bedömningsskalor och nivåer endast har använts i vårt samhälle drygt hundra år. Guba och Lincoln (1989) beskriver vad de kallar fyra generationer av bedömning. Den första generationen började dyka upp under sent 1800-tal i och med testpsykologins tillkomst. Olika typer av mätningar i skolsammanhang utvecklades och även så kallade intelligenstest, IQ-skalar. Denna typ av bedömning har haft stort inflytande på lästestning och därmed indirekt på läsundervisningen.

Den andra generationen bedömningar har uppstått under inflytande av läroplansforskningen och är starkt mål- och kriterierelaterad. I denna typ av bedömning är mätning ett av flera verktyg för att beskriva prestationer. Den tredje generationen bedömningar har uppstått under inflytande av sociologisk teori. Denna typ av bedömning är holistisk och används exempelvis för att ge utlåtande om elevers lämplighet för olika typer av yrken.

I samtliga dessa tre slag av bedömningar används på förhand fastställda gränser och parametrar. Guba och Lincolns (1989) fjärde generationen bedömningar är utvecklad inom ett socialkonstruktivistiskt perspektiv och där tolkning har stor betydelse. Inom denna typ av bedömningar bestäms parametrar och gränser genom förhandling mellan olika intressenter, som exempelvis lärare och elever. Loggböcker och portfolios är exempel på verktyg som används.

När man vill analysera den kunskapssyn olika test och bedömningar bygger på, kan det vara lämpligt att utgå från någon form av kunskapstaxonomi. Aristoteles skilde mellan *techne*, *episteme*, *fronesis* och *noesis* (Linde 2003). *Techne* betecknade praktisk färdighet, *episteme* stod för kunskap och vetande om fakta och relationer, *fronesis* syftade på etisk handlingsförmåga och erfarenhet medan *noesis* var filosofisk reflektion över tidigare vetande och dess grund. Blooms taxonomi (Bloom & Krathwohl 1956), som bland annat ofta används i undervisningssammanhang, skiljer mellan *kunna*, *förstå*, *tillämpa*, *analysera*, *sammanställa* och *bedöma*. Gemensamt för de två taxonomierna är att de första formerna betecknar kunskapsnivåer som betraktas som enkla medan de senare betecknar mer avancerade former av kunskap. Exempel på kunskap på den första nivån är teknisk avkodning och ordagrann förståelse av skriven text medan exempel på de högre nivåerna är att sammanställa, värdera, kritiskt granska och reflektera över kunskap från olika källor.

Många av de vanligt förekommande lästesten tar inte hänsyn till elevernas färdigheter i det språk bedömningen görs på. Detta drabbar främst många andraspråkselever. När de bedöms med test som är utformade för enspråkiga elever, blir bedömningen samtidigt både en bedömning av läsning och språk. Detta medför ett validitetsproblem genom att det blir omöjligt att avgöra om ett visst avvikande eller uteblivet svar är resultatet av läsproblem eller problem med det språk på vilket bedömningen genomförs. Dessa elever riskerar därför att bli bedömda som svagare när det gäller läsning på grund av att deras kompetens endast mäts på detta språk och att deras kompetens i andra språk inte tas med i beräkningen.

För en elev med god läsförmåga på ett annat språk, som ännu inte kommit så långt i svenskan, krävs andra åtgärder än för en elev som inte tidigare kan läsa på ett annat språk, liksom för en elev som även på sitt eller sina tidigare språk har problem med läsning. Men även andra elever med språkliga kunskaper och erfarenheter som skiljer sig från dem som värderas i bedömningen riskerar att uppnå lägre resultat på grund av detta. Det kan gälla elever från hem där andra dialekter eller sociolekter talas och även elever från hem där språk används på sätt som skiljer sig från skolans (Heath 1983, Nauc  r & Boyd 1994).

N  r det g  ller flerspr  kiga elever h  vdar Holm (2009) att det knappast   r etiskt acceptabelt att bed  ma tv  spr  kiga elever efter mallar som utformats f  r enspr  kiga elever. S  rskilt i de fall testet genomf  rs inom tidsramar   r risken stor f  r l  gre resultat f  r andraspr  kselever eftersom just l  sning p   ett andraspr  k f  r de flesta tar l  ngre tid   n l  sning p   ett f  rstaspr  k. Detta riskerar att resultera i or  ttvisa, negativa bed  mningar som kan f   allvarliga effekter f  r flerspr  kiga elevers sj  lvk  nsla men   ven leda till att felaktiga insatser s  tts in.

En av de f   studier som genomf  rts i Sverige, d  r flerspr  kighet kopplats samman med l  s- och skrivutveckling,   r Hedmans (2009) studie av dyslexi och flerspr  kighet. Hon visar bland annat att det inte   r tillr  ckligt att studera fonologisk medvetenhet och avkodning. N  r det g  ller andraspr  kselever som   nnu inte uppn  tt   ldersrelevant niv   i sitt andraspr  k och som konstaterats vara dyslektiker, fann hon att de anv  nde kontexten f  r att   verbrygga bristande ordf  rst  else.

Det   r allts   viktigt att eleverna ges denna m  jlighet f  r att en r  ttvis bed  mning ska kunna genomf  ras. Hedman fann   cks   att balansf  rh  llandet mellan elevens olika spr  k hade betydelse f  r testresultatet, n  got som endast kan belysas genom att eleven bed  ms p   sina olika spr  k, inte bara p   andraspr  ket. Hon kunde exempelvis visa att elevers dyslexi kunde p  verka l  sandet i de olika spr  ken p   olika s  tt, bland annat p   grund av skillnader mellan de olika spr  ken n  r det g  ller graden av   verensst  mmelse mellan uttal och stavning. P   s   vis, konstaterade hon, riskerar andraspr  kselever b  de att   ver- respektive underdiagnosticeras n  r det g  ller dyslexi.

BED  MNING AV L  SF  RM  GA I SVERIGE

Mycket av teoribildningen runt l  sning i Sverige har varit fokuserad p   bed  mning av l  sf  rm  ga d  r forskare som Ingvar Lundberg (t ex 1984) och   ke Edfeldt (t ex 1982) spelat betydelsefull roll, s  rskilt inom svensk l  rarutbildning. Under en tid runt 1980-talet fokuserade debatten om undervisning i l  sning och skrivning, b  de bland praktiker och bland forskare, p   undervisningsmetoder; analytiska kontra syntetiska. Fr  n 1990-talets b  rjan har fokus flyttats fr  n undervisningsmetoder till inl  rning och l  rande. Liberg (1990) beskrev l  sning som en process som innefattar b  de analytiska och syntetiska drag.

Man kan dock in i v  ra dagar sk  nja tv   huvuddrag i svensk forskning om barns l  sande och skrivande. Det ena har fokus p   sj  lva l  sprocessen och d  

särskilt dess fonologiska drag, det vill säga kopplingar mellan fonem och bokstäver. Detta perspektiv har främst inriktats på elevers brister och svårigheter och är starkt representerat inom forskning om elever i läs- och skrivsvårigheter och dyslexi (Lundberg 1984, Myhrberg 2007). Det andra perspektivet fokuserar skrift som ett kommunikativt medium och betonar skriftens sociala och kulturella aspekter (Liberg 1990, 2006). Inom detta perspektiv betraktas läsning och skrivning i ett helhetsperspektiv och man studerar människors användning av och förståelse av skrift.

Dessa två perspektiv på skrift återspeglas i de olika typer av lästest som används. Pehrson och Sahlström (1999) presenterar en kartläggning av olika typer av bedömningar som används i den svenska grundskolan för läsning och skrivning. De visar hur tidigare fokus på den enskilda eleven och på urval av elever till specialundervisning där man huvudsakligen arbetade med det eleven misslyckats med, i och med SIA-utredningen 1974 och den nya läroplanen, *Lgr80*, förändrades till fokus på att förändra förutsättningarna för det enskilda barnets deltagande i det ordinarie arbetet i klassen. Man började göra en åtskillnad vad gäller vem som bär problemet, den enskilda eleven eller skolan.

Caroline Liberg har sedan 1990-talets början motsatt sig det traditionella synsättet på läsning och läsförmåga och förespråkar ett perspektiv där skriftspråksutveckling ses i ett större sammanhang. I *Lpo94* och *Lpf96* föreskrivs en holistisk och socialinteraktionalistisk inriktning på läs- och skrivundervisningen i skolan. Samtidigt kan man konstatera att dyslexiförespråkarnas synsätt ofta väljs i skolan, när det gäller bedömning av läsförmåga, framför mer holistiskt inriktade förklaringsmodeller.

Man kan tänka sig att en orsak till att traditionella, standardiserade test väljs kan vara att de lätt förleder lärare och skolledare att tro att man med hjälp av dem på ett enkelt sätt kan få förklaringar till komplexa frågor, så som frågan om hur barn utvecklar sin läsförmåga och varför vissa barn får en mindre gynnsam skriftspråksutveckling. En annan förklaring till deras popularitet kan vara deras fokus på brister hos det individuella barnet i stället för hos skolan. På så vis riskerar ingen enskild lärare att behöva ifrågasättas. Från skolledarhåll kan också ett motiv vara en önskan att styra verksamheten och att få resultat som är enkla att visa upp.

Pehrsson och Sahlström (1999) delar in analysverktygen i normerade och ickenormerade, där de normerade används för bedömning av elevers brister med syfte att exempelvis äska medel och resurser eller för att avgöra vilka elever som ska ha specialundervisning, medan de ickenormerade används för att följa elevernas skriftspråksutveckling och som underlag för undervisningsplanering. De normerade analysverktygen mäter skriftens beståndsdelar utifrån ett grammatikaliskt perspektiv och skiljda från sitt meningsfulla sammanhang.

Ett exempel på detta är det första test som analyseras i denna artikel, där texten förutsätts förstås på *ett* sätt och ha *ett* rätt svar. I denna typ av test är tid en faktor, elevernas förmåga att utföra vissa uppgifter på en viss tid utgör en del i bedömningen. Testen används för att upptäcka brister hos eleverna, brister som ska åtgärdas och som används som argument för utökade resurser och

placering av eleverna i specialundervisning av olika slag. Förslagen till åtgärder består oftast av träning i fonologisk medvetenhet. I något av de fall som ingår i Perhsons och Sahlströms kartläggning uppmanades läraren även att använda själva provet som övningsmaterial när brister upptäcktes.

Den typ av prov som Pehrson och Sahlström (1999) kallar ickenormerande prov, ger däremot underlag för att hitta förutsättningar för att stödja elevernas utveckling och att visa på vägar att förändra undervisningen. I dessa prov ges förslag på pedagogisk planering och de riktar sig både till läraren och till eleven. Den typ av bedömning som utförs via normerade prov kan ses som vad Edfeldt (1982) kallar *reduktionistiska*, man försöker reducera ett komplicerat problem till summan av ett antal mindre sammansatta delförlopp. Normerade läsprov mäter läsning i sin mest begränsade betydelse, som olika specifika beståndsdelar, exempelvis fonologiska aspekter, och endast i viss utsträckning skrift som språk och kommunikation. Endast ickenormerade prov mäter läsning i sin utvidgade betydelse, som en viktig aspekt av lärande i skolan som helhet.

Faran med att använda enkla kvantitativa metoder för att besvara komplicerade frågor har framhållits av Evaldsson och Nilholm (2009). De visar hur en ensidig tro på så kallad evidensbaserad forskning, när det gäller mobbning, kan leda till att formaliserade träningsprogram används i stället för att varje enskild situation behandlas som unik. Riskerna när det gäller lästest som endast mäter en begränsad aspekt av läsning och som inte tar hänsyn till ett bredare spektrum av elevens läsande och skrivande, är att vissa elever onödigtvis utpekas som att ha bristande färdigheter och att planerade insatser inte blir relevanta. Detta visar på faran med att dra alltför långtgående slutsatser av begränsade formella test av den typ som Pehrsson och Sahlström (1999) benämner normerade. Ett aktuellt exempel på ett normerat test är det test som ska analyseras här, testmaterialet »Vilken bild är rätt? En enkel klassdiagnos av läsförståelse» (Lundberg 2001).

VILKEN BILD ÄR RÄTT?

Beskrivning av materialet

Diagnosmaterialet »Vilken bild är rätt? En enkel klassdiagnos av läsförståelse» är av ett slag som var vanligt under svensk grundskolas tidigare år. Materialet har koppling till boken *God läsutveckling: Kartläggning och övningar* (Lundberg & Herrlin 2003). Diagnosen består av ett tiosidigt häfte med 38 uppgifter och två övningsuppgifter. Varje uppgift består av en text, i de flesta fall en mening och i några fall två. Till varje text finns fyra bilder och eleven ska avgöra vilken av bilderna som passar bäst till texten. Endast en av bilderna bedöms som rätt. Provet genomförs på tid, antingen tio minuter eller fem.

Till testet finns en lärarhandledning (Annell, Lundberg & Leven 2003). Denna argumenterar först för diagnostisering av läsförmåga. Här argumenteras för att denna typ av test används för att bedöma »Förmågan att snabbt och säkert läsa meningar och förstå dess innebörder» (s 4). Man motiverar detta

med att denna förmåga är grundläggande och betonar vikten av att »eleven kan läsa med precision, både när det gäller att uppfatta enskilda ord och när det gäller att uppfatta textens grammatiska uppbyggnad» (s 4). Man framhåller att »ett felaktigt svar kan avslöja vilken typ av fel som eleven begått» (s 4) men ingen ledning för de tolkningarna ges.

I handledningen framhåller man att provet genomförs snabbt och lätt. Hela klassen kan genomföra testet på kort tid och det går snabbt att rätta. Man framhåller också att provet är utprovat på ett större antal elever och anger statistiska värden för dessa både i procentvärden och i staninepoäng, för läsning i fem respektive tio minuter. Värdena bygger på antagandet att eleverna följer nominalskalan. Man anger tre vanliga feltyper och ger råd för hur de kan avhjälpas.

Råden innebär i två av fallen att eleven övar att läsa ordlistor med ord som börjar på liknande sätt för att »tvingas in i ordet» (s 6) och i det tredje fallet rekommenderas att »läsa med engagemang och uppmärksamhet» (s 6) och »att läsa texter där man måste följa anvisningar mycket noga» (s 6). Man ger inga råd för hur man hjälper eleverna med att läsa »med engagemang och uppmärksamhet». Handledningen avslutas med åtta råd för hur elever kan stödjas för att uppnå flyt i läsningen.

Provets innehåll

För att ge en bild av provets innehåll kommer tre uppgifter att beskrivas och analyseras. Uppgifterna är tänkta att genomföras i tur och ordning. Varje elev börjar med den första och fortsätter sedan så långt han eller hon hinner inom den angivna tidsramen. Därför är de tidiga uppgifterna avsedda att vara enklast och de sista svårast. Här har jag valt att analysera en av övningsuppgifterna och två av de tidiga uppgifterna, uppgift fem och åtta, eftersom dessa är uppgifter som nästan samtliga elever förväntas genomföra när testet genomförs på tio minuter. Enligt handledningen har vid tidigare prov 96 procent av eleverna klarat fler än fem uppgifter.

Den ena övningsuppgiften lyder: *Sara luktade på en ros*. Bilderna visar en kvinna som luktar på i tur och ordning: (1) fyra rosor, (2) en ros, (3) en blomma av odefinierbar art samt (4) en kvist. Förutom att avläsa texten behöver eleven kunna följande för att lösa uppgiften rätt (om vi antar att eleven inte väljer bild på måfå och har tur att välja »rätt» bild):

- tolka en stiliserad bild, bland annat av en kvinna med svenskt utseende,
- förstå att Sara är en kvinna (inte lika enkelt för barn från kulturer där namnet inte förekommer eller där inte just kvinnonamn slutar på -a),
- veta vad en ros är och hur den ser ut, samt
- förstå att »Sara luktar på en ros» innebär att det bara får finnas *en* ros på bilden, inte fyra.

Här kan man tänka att risken är stor att de barn, till exempel vissa flerspråkiga barn som ännu inte utvecklat samma kompetens i svenska som sina klasskamrater och som kanske inte förstår ordet ros eller som kanske inte vet hur en ros ser ut, inte frågar utan hellre sitter tysta och hoppas förstå bättre längre fram. Handledningen ger inte heller läraren något råd om att verkligen undersöka

om eleverna har förstått. I stället kan man där läsa en ordagrann förlaga för lärarens instruktioner: »Vilken bild passar bäst? Peka! Ja, det är rätt! Det kan ju inte vara bilden med många rosor, eller bilden med en annan blomma, eller bilden med en kvist» (s 9). Detta är ett tydligt exempel på hur man som lärare kan missa olikheter i elevernas förkunskaper genom att förutsätta att alla har liknande förkunskaper. Risken är stor att elever som av en eller annan orsak inte förstår de inledande övningsuppgifterna snarare försöker dölja det än be om ytterligare förklaring.

Den femte uppgiften lyder: »Lisa har gjort en lykta av snöbollar. Den lyser så fint.» Bilderna visar en flicka som i tur och ordning: (1) står på knä och lägger bollar på varandra (ringar på bilden som antas tolkas som snöbollar), (2) står och ser på en pyramidformad hög av bollar med taggar ritade runt som antas tolkas som att den lyser, (3) står på knä och lägger bollar på varandra med taggar runt (liknande de på bild två), och slutligen (4) står på knä och lägger en fjärde boll på tre andra med två liggande bredvid. För att förstå att det är den andra bilden som avses behöver eleven kunna följande:

- förstå att Lisa representeras av det barn som visas på bilden,
- förstå vad »en lykta av snöbollar» är (Detta prov genomfördes i oktober i utprovningsen. Det är inte bara andraspråkselever som kan ha svårt att förstå vad en snölykta är i oktober. I stora delar av landet är det sällan möjligt att göra sådana, än mindre självklart att man har haft tillfälle till att göra det.),
- förstå att cirklar symboliserar snöbollar,
- förstå att en pyramid av snöbollar betyder snölykta (den som byggt en snölykta vet att man knappast kan bygga den pyramidformad utan i stället måste ge den en kupad form för att inte snöbollarna ska ramla ner, i synnerhet om man vill tända ett ljus i den),
- förstå att ritade taggar runt ett föremål betyder att det ska visa att det lyser,
- förstå att man här inte ska tänka på att man, när man gör en snölykta, måste sätta i ljuset och tända detta innan de sista bollarna läggs dit och att alltså inte bild tre är rätt trots att den stämmer med hur man tänder en snölykta, samt
- förstå att »har gjort» innebär att görandet är avslutat och att det är därför som just bild två betraktas som rätt

I detta exempel finns ännu större risker att själva bildtolkningen blir ett problem. För många barn är snölyktor inte något som hör till deras vardag, särskilt inte i oktober och november när genomförande av provet bör genomföras för att kunna jämföras med den statistik som finns i handledningen. Bildtolkningen i detta fall ställer krav på att kunna tolka inte bara en stiliserad bild av en snölykta utan också av de taggar som på två bilder ritats utanför snölyktan för att visa att den lyser.

Den åttonde uppgiften, också den inom ramen för vad 80 procent av eleverna hann med i år två på tio minuter enligt den presenterade statistiken, lyder: »Siri sitter på golvet och läser. Hon äter också på ett äpple.» Bilderna visar en flicka som i tur och ordning: (1) sitter på golvet och tittar i en bok med ett äpple bredvid, (2) sitter på en pall, håller i boken, blundar och biter i ett äpple,

(3) sitter på golvet, med boken hopslagen, blundar och biter i äpplet, och slutligen (4) sitter på golvet, med boken uppslagen, blundar och biter i äpplet. Här gäller det, för att eleven ska välja bild fyra, att eleven förstår att huruvida flickan blundar eller inte är oväsentligt för denna uppgift och att det inte betraktas som att hon äter ett äpple när det ligger bredvid.

Sammantaget kan man säga att själva bildhanteringen innebär risker för fel-tolkning och på så sätt utgör en viktig del i vad som testas. För att kunna lösa dessa uppgifter »rätt», det vill säga så att de ger poäng, måste eleverna inte bara avläsa och förstå meningarna utan också kunna avläsa och tolka bilderna på avsett sätt. De måste dessutom behärska den typ av litteracitetspraxis som testet representerar, att avläsa en mening, avskild från sitt sammanhang, tolka på det sätt som förväntas, koppla den till rätt bild samt i vissa fall bortse från tidigare egna erfarenheter.

Detta innebär att använda en lässtrategi som normalt inte används i de texter eleverna läser och skapar utanför skolan. De överensstämmer inte heller särskilt väl med de typer av lässtrategier som de kommer att behöva utveckla senare under skoltiden, med bland annat kritiskt tänkande och selektivt läsande, och inte heller med de strategier de använder vid elektroniskt läsande och skrivande av olika slag.

Eftersom bildtolkning är en kulturell fråga kan man anta att just elever med annan sociokulturell bakgrund än den dominerande, så som många flerspråkiga elever och elever som kommer från socioekonomiskt svaga miljöer, missgynnas av detta. Eleven behöver alltså ha utvecklat kunskap om hur man löser uppgifter av just detta slag, vad vi skulle kunna kalla skoltestkunskap. I detta ingår bland annat att veta vad man ska fokusera på, så som vissa detaljer, utan att låta annat så som tidigare erfarenheter eller egen fantasi och inlevelseförmåga spela in. Eftersom provet genomförs på tid behöver eleven också förstå vikten av att genomföra uppgifterna snabbt och utan att tänka efter eller att försöka sig på att tolka det lästa.

Att testet inte alls är utvecklat med tanke på att en ökande andel av de svenska skoleleverna har en flerspråkig bakgrund eller en mångkulturell uppväxtmiljö, framgår av exemplet med snölyktan, en av de uppgifter som de flesta elever förväntas lösa. Enligt Skolverkets (2010 Tabell 8A) statistik är det 19,4 procent av eleverna som är berättigade till modersmålsundervisning, det vill säga som antingen talar ett annat språk än svenska hemma eller ett andra språk, eller talar svenska plus ett annat eller fler andra språk.

Att det i detta fall blev en extra svårighet även för de elever som är erfarna byggare av snölyktor ökar inte uppgiftens lämplighet. Man nämner att även elever med annat modersmål än svenska deltagit i utprovningen men att man valt att inte markera dem eftersom »en mångkulturell och flerspråkig bakgrund är ju allt vanligare i svenska skolor» (s 5). Att inte hänsyn tagits till elevernas språkliga bakgrund blir extra problematiskt med tanke på den låga andelen lärare med utbildning i svenska som andraspråk och modersmålsämnets låga status (jfr Skolverket 2008).

Detta test, av det normerade slaget med på förhand angivna gränser, är utformat på det sätt som Guba och Lincoln (1989) ser som första generationens bedömningar, alltså av det slag som började utvecklas i slutet på 1800-

talet. I denna testform förutsätts att alla barn följer samma utvecklingsgång, i samma ordning och samtidigt. Testet mäter läsförmåga i sin allra mest begränsade form, som en teknisk färdighet. Att det dessutom både är kulturokänsligt och etnocentriskt framgick av analysen av de tre uppgifterna.

Den typ av kunskap som bedöms är av det slag som betraktas som enklast i Aristoteles (Linde 2003) och Blooms (Bloom & Krathwohl 1956) taxonomier. Att inte ens frågor som elevers slarv, chanstagande, associationer och okoncentration behandlas i lärarhandledningen visar att läsning betraktas som en ren teknisk färdighet, avskild från ett socialt och kulturellt sammanhang, något som kan genomföras *rätt* eller *fel*. Det blir då intressant att se på ett bedömningsmaterial av det slag som Pehrson och Sahlström (1999) ser som ickenormerat, ett material för språkbedömning utarbetat av svenska skolmyndigheter, »Nya språket lyfter» (Skolverket 2008), en vidareutveckling av det tidigare »Språket lyfter» (Skolverket 2002)

NYA SPRÅKET LYFTER

»Nya språket lyfter» är ett diagnosmaterial utarbetat av Skolverket (2008) för bedömning av elevers måluppfyllelse i svenska och svenska som andraspråk. Läsning och läsförmåga är i fokus men även andra aspekter av språk behandlas. I förordet sägs att materialet ska användas av lärare i de båda svenskämnen i samarbete med lärare i modersmål. Bedömningsmaterialet inleds med en kort genomgång av forskning om barns språkutveckling, med särskilt fokus på skrift. Flerspråkiga barns språkutveckling lyfts fram som en positiv möjlighet.

Diagnoserna bygger på lärarens observation, vilket motiveras i en inledande presentation av materialets innehåll och uppläggning. Man ger följande exempel på dokumentation av elevernas utveckling: foton, elevernas produkter, inläsningar på band samt översikt av det slag som materialets observations-schema utgör. Observationsschemat är uppbyggt med olika observationspunkter, som i huvudsak behandlar läsning och skrivning, men även tal och samtal finns med som särskilda observationsspunkter.

Observationsspunkterna presenteras tillsammans med kortare teorigenomgångar och förklaringar och samlas till avstämningar. Den tredje av de fyra avstämningarna är avsedd att motsvara mål att uppnå för år tre. För varje observationspunkt och avstämning ges förslag på vad som bör noteras vid bedömningen och förslag på vad undervisningen kan innehålla för att stimulera denna utveckling. Denna bedömning representerar vad Guba och Lincoln (1989) ser som den fjärde generationen av test.

Den stora skillnaden mellan de två bedömningsmaterialen ligger i synen på vad som ska bedömas och vilket material som krävs för att göra detta. En annan skillnad är att medan man i »Vilken bild är rätt?» framhäver hur enkelt det är att genomföra och tolka testet, visar »Nya språket lyfter» på den komplexitet som utveckling av språk och läsförmåga innebär, genom att översiktliga teorigenomgångar finns inlagda där olika frågor som bedöms vara av vikt vid bedömningen behandlas.

Bland annat behandlas barns språkutveckling och råd ges för hur lärare kan bedöma och dokumentera elevernas språkutveckling, med särskilt fokus på skrift. Bedömningen grundas på olika typer av information och knyts till aktuell forskning. Medan »Vilken bild är rätt?» sätter fokus på elevens brister riktas intresset i »Nya språket lyfter» mot elevens kunskaper och färdigheter, det som eleven kan. I detta material framhåller man att elever och föräldrar ska involveras bland annat genom att elevens eget analyschema som knyts till den individuella utbildningsplanen (IUP) ska utgöra grunden för samtal mellan lärare, elev och förälder när det gäller elevens kunskapsutveckling. På så vis involveras i detta bedömningsmaterial både eleverna själva och deras föräldrar i själva bedömningen även om huvudansvaret vilar på läraren. Medan »Vilken bild är rätt» riktar fokus på elever som lyckas sämre, ger »Nya språket lyfter» råd för utvecklande av undervisningen för att stimulera elevernas utveckling i vid bemärkelse.

DISKUSSION

Det är viktigt att betona att alla former av test och bedömning, när det gäller läsförmåga, representerar en viss syn på kunskap, lärande och språk liksom ett visst perspektiv på språkanvändning och litteracitetspraxis. Därför är det av grundläggande betydelse att de olika antaganden och värderingar som olika bedömningsredskap bygger på, synliggörs. Detta ställer krav på en god bedömningskompetens hos lärare, inte bara i skolans senare år utan även bland lärare i förskola och grundskolans tidigare år. När det gäller andraspråkselever är det särskilt viktigt att bedömningar av deras språk genomförs av personer med kompetens i andraspråkutveckling.

Den typ av test som vi såg i det första exemplet, »Vilken bild är rätt?» är ett exempel på test där läsning behandlas som teknik och inte som den kommunikativa och sociala aktivitet det är. Lärarna reduceras till administratörer av testen i stället för att få stöd i sitt arbete och med att ställa krav på ökade resurser för lärarutbildning, fortbildning, skolläda och forskning. Vi såg också att det var kulturokänsligt och etnocentriskt.

Lästest av det slaget ger fel signaler till elever och föräldrar (och lärare) om vilka typer av läsande som är viktiga både för skolframgång, för nuvarande liv utanför skolan och för kommande arbetsliv. Testen stämmer särskilt illa överens med de typer av läsande och språkanvändning som är betydelsefulla vid elektronisk kommunikation; framför datorn eller med mobilen. De bygger på en syn på språk, skrift och läsning som inte stämmer med modern forskning och inte med de kunskapsbehov som finns. I synnerhet för bedömning av andraspråkselevs läsande är den typen av test inte relevant.

Samtidigt är det förståeligt att lärare väljer att använda dessa typer av test och att skolledare beordrar sina lärare att göra det. Enkelheten i handhavandet, den korta tid som krävs och de klara nivågraderingarna är förförande när det gäller detta slags test. I en skolsituation där nya arbetsuppgifter för lärare tillkommer, med rapporter och utlåtanden som ska skrivas, är valet av denna typ av test frestande. Den typ av helhetsbedömningar som »Nya språket lyfter» utgör exempel på, ställer höga krav både på lärares kompetens, på deras

användning av sin arbetstid och på deras organisering av undervisningen. För att lärare ska kunna förväntas genomföra bedömningar av det slag som presenteras i »Nya språket lyfter» krävs att skolan tillförs nödvändiga resurser.

Lärarna behöver ha god kompetens när det gäller barns språkutveckling, både vad avser tal och skrift och enspråkiga och flerspråkiga barn. Detta kräver inte minst att lärarutbildningarna, och anknytande forskningsämnen, får tillräckliga resurser för kompetensutveckling och forskning. För att modersmålslärarna ska kunna involveras i arbetet, som avsett, krävs att de erbjuds relevant utbildning och att modersmålsämnets status uppvärderas. Placering av undervisningen i modersmål efter skoltid och centraliserad till vissa skolor, omöjliggör i många fall den typ av samarbete mellan klasslärare, lärare i svenska som andraspråk och modersmålslärare som förespråkas i materialet. För att vi ska få se den typ av bedömningar som »Nya språket lyfter» representerar mer allmänt genomförda i skolan, krävs en förändring vad gäller attityd och ökad kunskap, bland lärare och skolledare men även bland utbildningspolitiker.

När de här aktuella proven relateras till maktfrågan kan man konstatera att båda bedömningsverktygen kan betraktas som medel för maktutövning. I det första fallet, »Vilken bild är rätt?» är de som utsätts i huvudsak elever som påverkas i sin syn på var fokus i läsutvecklingen bör ligga och i synnerhet de elever som uppnår låga resultat och som riskerar att bli särbehandlade med sänkt självförtroende till följd. Även de elever som får en missvisande bedömning, och därmed ickerelevanta åtgärder, drabbas.

Man kan också hävda att lärarna i viss mån blir utsatta, eftersom de får en roll som inte värderar deras professionella kunskap. När det gäller »Nya språket lyfter» ser jag i stället en risk för att lärarna blir utsatta på ett annat sätt genom att de påförs utökade och i viss mån nya arbetsuppgifter. Det är rimligt att anta att Skolverket utformat detta material i syfte att påverka och förändra undervisningen i skolan. Detta syfte kan vara vällovligt men bra skolundervisning ställer krav på resurser.

Materialet ställer höga krav på lärarna, på hur de ska organisera undervisningen liksom på att de ska ha tillräckliga kunskaper för att kunna genomföra bedömningen. Om skolan tillförs nödvändiga resurser, både vad gäller personal för att kunna genomföra arbetet och vad gäller nödvändig fortbildning, ser jag att resultatet av att skolorna följer en bedömningsplan av den typ som »Nya språket lyfter» representerar, kan innebära skapande av möjligheter för elever att uppnå ökade kunskaper och i slutändan högre resultat på bedömningar av olika slag. Om inte finns en risk för att lärares känsla av otillräcklighet ökar, vilket knappast gagnar undervisningen.

I de nya direktiven för lärarutbildningen utökas glädjande nog studietiden för lärare för de tidiga åren i skolan till fyra år men tyvärr minskas samtidigt möjligheterna för denna lärargrupp till fördjupade kunskaper. För att lärare ska kunna göra goda bedömningar av elevernas språkutveckling och kunna planera verksamheten väl, måste det finnas utrymme för vissa lärare att skaffa sig specialkunskaper om språkutveckling och bedömning.

NOT

Tack riktas till Örebro universitet som möjliggjort detta arbete samt till Catharina Schmidt och Lars Holm för värdefulla kommentarer. Ett särskilt tack riktas till Lars-Erik Widmark som uppmärksammade mig på flera problem som lärare och elever utsätts för av lästest av olika slag.

LITTERATUR

- Annell, B., Lundberg, I. & Leven, E. 2003: *Kom och läs! Vilken bild är rätt? En enkel klassdiagnos av läsförståelse*. Stockholm: Natur och Kultur.
- Bloom, B.S. & Krathwohl, D.R. 1956: Taxonomy of educational objectives. The classification of educational goals, by a committee of college and university examiners. I *Handbook I: Cognitive domain*. New York: Longmans, Green.
- Edfeldt, Å.W. 1982: *LUS: Läsprocessen grundbok om läsforskning*. Stockholm: Liber.
- Evaldsson, A-C. & Nilholm, C. 2009: Evidensbaserat skolarbete och demokrati: Mobbing som exempel. *Pedagogisk forskning i Sverige*, 14(1), 65–82.
- Guba, E.G. & Lincoln, Y.S. 1989: *Fourth Generation Evaluation*. London: Sage.
- Heath, S.B. 1983: *Ways with words: Language life and work in communities and classrooms*. Cambridge: Cambridge University Press.
- Hedman, C. 2009: *Dyslexi på två språk: En multipel fallstudie av spansk-svensktalande ungdomar med läs- och skrivsvårigheter* Stockholm: Stockholms universitet, Centrum för tvåspråkighetsforskning.
- Holm, L. 2009: Who evaluates the evaluation? An analysis of computer-adaptive testing using literacy as an example. *Nordisk Pedagogik*, 29(2), 137–148.
- Liberg, C. 1990: *Learning to read and write*. (RUUL, 20) Uppsala: Uppsala University, Department of Linguistics.
- Liberg, C. 2006: *Hur barn lär sig läsa och skriva*. Lund: Studentlitteratur.
- Linde, G. 2003: *Kunskap och betyg*. Lund: Studentlitteratur.
- Lundberg, I. 1984: *Språk och läsning*. Malmö: Liber.
- Lundberg, I. 2001: *Vilken bild är rätt? En enkel klassdiagnos av läsförståelse*. Stockholm: Natur & Kultur.
- Lundberg, I. & Herrlin, K. 2003: *God läsutveckling: Kartläggning och övningar*. Stockholm: Natur & Kultur.
- Myhrberg, M. 2007: Läs- och skrivsvårigheter. I *Att läsa och skriva: Forskning och beprövad erfarenhet*. Stockholm: Myndigheten för skolutveckling.
- Naclér, K. & Boyd, S. 1994: Turkish and Swedish pre-school children's language socialisation. Implications for literacy development. I L. Huss (red): *Många vägar till tvåspråkighet: Föredrag från ett forskarsymposium vid Göteborgs Universitet den 21–22 oktober 1999*. Uppsala: Uppsala University, Centre for Multi-ethnic Research.
- Pehrsson, A. & Sahlström, F. 1999: *Kartläggning av läsning och skrivning ur ett deltagarperspektiv: Analysverktyg för alla*. (Specialpedagogiska rapporter, nr 14) Göteborg: Göteborgs universitet, Institutionen för specialpedagogik.
- Shohamy, E. 2001: *The power of test: A critical perspective on the uses of language tests*. Essex: Pearson Education.
- Skolverket, 2002: *Språket lyfter: Lärarhandledning*. Stockholm: Skolverket.
- Skolverket, 2008a: *Med annat modersmål: Elever i grundskolan och skolans verksamhet*. Stockholm: Fritzes.
- Skolverket, 2008b: *Nya språket lyfter: Diagnosmaterial i svenska och svenska som andraspråk för grundskolans årskurs 1–5*. Stockholm: Skolverket.
- Skolverket, 2010: *Skolor och elever i grundskolan läsåret 2009/10*. <http://www.skolverket.se/sb/d/1638#paragraphAnchorO> (2010-05-17)