

Bedömning i svenskämnet

Årskurs 7–9

REDAKTÖRER GUSTAF SKAR & MICHAEL TENGBERG

Svenskläraryrskriften Årsskrift 2014

Svenskläraryrskriften 237

Svensklärföreningen

är en ideell, riksomfattande ämnesförening för den som arbetar med eller är intresserad av språkutveckling och litteraturundervisning från förskola till högskola.

Vi arbetar för

- en genomtänkt syn på språket
- ett aktivt och medvetet förhållningssätt till språk och litteratur
- svenskämnets helhet
- svenskämnets ställning i skola och lärarutbildning
- en levande ämnesdebatt
- fortbildning till både blivande och verksamma svensklärare.

Svensklärföreningen

- ger ut tidskriften Svenskläraren med fyra nummer per år
- ger ut en årsskrift som behandlar ett aktuellt tema
- medverkar vid utgivning av böcker och skrifter med svensksdidaktisk anknytning
- anordnar kompetensutveckling
- skriver remisser och gör uppvaktningar i skol- och utbildningsfrågor
- utgör ett nätverk för språkutvecklare och litteraturvänner i undervisningsvärlden.

Kontakta Svensklärföreningen!

www.svensklaraforeningen.se
info@svensklaraforeningen.se

Medlemsavgift är 250 kr per år
(120 kr för studerande). Plusgiro 89 84-7, Svensklärföreningen

Som ny medlem får du en tidigare utgiven årsbok (värde ca 190 kr).

Du behöver Svensklärföreningen – SLF behöver Dig!

Bedömning i svenskämnet
Svensklärföreningens årsskrift 2014. Svensklärrserien 237

Redaktörer: Gustaf Skar och Michael Tengberg
Omslag: Bitte Granlund/Happy Book AB

© Svensklärföreningen 2015. Distribution Natur & Kultur
Tryckt i Estland 2017
ISBN 978-91-27-13844-5



Innehåll

Medverkande författare	4
Inledning	6
DEL 1 – TEORETISKA UTGÅNGSPUNKTER	23
1. Kunskapsbedömningars grund och syften	24
2. Att kvalitetssäkra bedömning	43
3. Skrivande som grundläggande färdighet: förståelse, bedömning och undervisning	63
4. Att bedöma läsförmåga – teoretiska utgångspunkter	87
5. Uppgifter	109
DEL 2 – BEDÖMNING I PRAKTIKEN	129
6. Skrivbedömning i formativt syfte och återkoppling	130
7. Enkel/utvecklad/välutvecklad elevtext i biologi: Hur bedömer vi det?	153
8. Bedömning av litterär analys i skolan	171
9. Att läsa och analysera noveller i årskurs 9	189
FÖRDJUPNING	211
10. Validitet	212
 Begreppslista	 238

Medverkande författare

KJELL LARS BERGE

är professor i textvetenskap vid Universitetet i Oslo.

GUDRUN ERICKSON

är professor i pedagogik med inriktning mot språk och bedömning vid Göteborgs universitet.

JENNY WIKSTEN FOLKERYD

är universitetslektor i didaktik med inriktning mot svenska vid Uppsala universitet.

ÅSA AF GEIJERSTAM

är universitetslektor i didaktik med inriktning mot svenska vid Uppsala universitet.

MICHAEL T. KANE

innehar The Samuel J. Messick Chair in Test Validity vid Educational Testing Service, USA.

EVA NILSON

är fil. lic. och doktorand i språkdidaktik vid Stockholms universitet.

KARINA PÅLSSON GRÖNDAHL

är fil. lic. i språkdidaktik och arbetar som lärare på Mariaskolan i Stockholm.

GUSTAF SKAR

är fil. dr och leder projektet Læringsstøttende prøver i skrivning
vid Nasjonalt senter for skriveopplæring og skriveforskning,
Høgskolen i Sør-Trøndelag.

SUSANNE STAF

är fil. lic. och doktorand i ämnesdidaktik vid Göteborgs
universitet.

MICHAEL TENGBERG

är docent i pedagogiskt arbete vid Karlstads universitet.

RAGNAR THYGESEN

är professor i specialpedagogik vid Universitetet i Agder.

Inledning

GUSTAF SKAR & MICHAEL TENGBERG

Varför en bok om bedömning i svenskämnet?

Varför behövs en bok om bedömning i svenskämnet? Den ökade uppmärksamheten på frågor om bedömning och utvärdering av elevers kunskaper har gjort att det också i Sverige i dag finns en snabbt växande mängd litteratur om bedömning och betygsättning. Här finns såväl teoretiska perspektiv på relationen mellan bedömning och elevers lärande som praktiska tips om dokumentation och hantering av matriser. Det finns böcker om bedömning i matematik (Björklund Boistrup, 2013), SO-ämnena (Jönsson & Odenstad, 2014) och NO-ämnena (Black & Harrison, 2014). Däremot saknas det teoretiskt grundade beskrivningar av de villkor och överväganden som gäller för klassrumsbedömning i svenskämnet, åtminstone vad gäller aspekterna läsande och skrivande.¹ I den här boken sätts därför frågor om bedömning i svenskämnet i förgrunden. Mer specifikt handlar boken om bedömning av elevers skrivande och läsande, som tillsammans utgör två av ämnets grundläggande lärandeobjekt.

Utgångspunkten för lärares undervisning i svenska återfinns i det centrala innehåll och de kunskapskrav som kursplanen anger. Här beskrivs de aspekter av skrivförmåga och läsförmåga som lära-

¹ Bedömning av elevers muntliga förmågor diskuteras bland annat i Palmér (2010). En praktiskt orienterad redogörelse, som bland annat tar upp matrisbedömning och uppgiftskonstruktioner, finns i Sandin (2009).

re ska utgå från i sin bedömning av elevernas kunskapsutveckling. Men hur gör man egentligen? Hur ska man till exempel förstå själva det fenomen som ska bedömas, det vill säga skrivförmåga och läsförmåga? Och hur lägger man upp undervisningen så att en rättvisande bedömning av elevernas kunskaper blir möjlig?

I den här boken behandlas dessa viktiga frågor om bedömning av läsande och skrivande. Avsikten är dels att beskriva några av de avväganden som lärare måste göra inför bedömningen av läsning och skrivande i klassrummet, dels att problematisera utgångspunkterna för bedömningen och ge förslag på sådana förhållningssätt som kan vara bra att ha med sig som lärare. Vi har därför eftersträvat en kombination av två olika perspektiv, som vi menar måste ingå i bedömningsprocessen: å ena sidan språk- och textteoretiska perspektiv, det vill säga teori om det stoff som ska bedömas, å andra sidan bedömningsteoretiska perspektiv, det vill säga teori om vad bedömning innebär och hur den bör utformas för att bli rättvisande och relevant.

Generellt har frågor om bedömning av elevers kunskaper uppmärksammas i allt högre utsträckning under senare år. Framför allt handlar det om att staten liksom skolans huvudmän söker efter fler och tillförlitligare mått på utbildningens kvalitet. Av det skälet specificeras, med utgångspunkt i kursplanerna, särskilda kriterier för vad som kännetecknar olika kunskapsnivåer. Dessa kriterier återfinns i skolors och enskilda lärares planeringsdokument, men också i de nationella provens bedömningsanvisningar. Förekomsten av den här sortens kriterier implicerar att elevers kunskaper mäts efter en rationell och objektiv standard. Men när elevens prestationer ska ställas mot kriterierna måste det till en bedömare, det vill säga en lärare, som *tolkar* i vad mån prestationerna möter kriterierna. Därmed blir lärarens kunskap om hur kriterierna ska tolkas avgörande för bedömningen. Likaså måste läraren i elevernas prestationer kunna känna igen de kvaliteter som bedömningskriterierna efterfrågar. Är den här texten som eleven skrivit verkligen *sammanhängande* och *begriplig*? Är de exempel

eleven använder från sin egen erfarenhet verkligen passande på ett sätt som *fördjupar/utvidgar resonemanget* i texten? (Exempel lånade från en av bedömningsmatriserna i det nationella skrivprovet i årskurs 9.) Bedömningens rationalitet och objektivitet – eller med en mer precis terminologi, dess *reliabilitet* och *validitet* – beror alltså i hög grad på lärarens kunskaper om vad som ska bedömas, det vill säga vederbörandes ämneskunskaper, liksom lärarens kunskaper om god praxis i bedömning. Det är denna kombination av två avgörande kunskapsområden som den här boken vill lyfta fram.

Bedömning av och för lärande

Bedömning handlar alltså om att samla information om elevens progression i förhållande till uppställda mål. Som begrepp knyts bedömning ofta ihop med betygsättning och utgör säkert i många sinnebild den process som leder fram till betyget, det vill säga som en bedömning *av* lärande. På senare tid har det emellertid också blivit vanligt att se bedömning som ett redskap *för* lärande. Informationen som läraren samlar om elevens kunskaper blir då i första hand framåtsyftande och handlar om att identifiera kommande steg i elevens progression (Gardner, 2012). Bedömning för lärande är något som sker kontinuerligt i klassrummet och som utgör en integrerad del av undervisningen och lärandet. Att lyfta fram begreppet *bedömning för lärande* handlar därför mycket om att rikta ljuset mot de många informella bedömningspraktiker (Leslie & Caldwell, 2009) som ligger till grund exempelvis för lärares återkoppling till elever och för lärares planering av kommande undervisning (Dobson, Eggen & Smith, 2009).

Ibland påtalas att bedömning för lärande, i motsats till betyg och nationella prov, handlar om att jobba med skolans mer komplexa mål, såsom utveckling av elevernas kreativitet och ansvarstagande (Lundahl, 2011). Men urskiljandet av bedömning för lärande kan också ses som ett sätt att understryka ett annat syfte med bedömningen av elevers kunskaper, oavsett vilket kunskapsområde

det gäller. De praktiker som här avses kan vara klassrumssamtal, återkoppling genom rättning av elevers produkter eller kamrat- och självbedömning, men det kan också handla om att använda resultat från standardiserade prov för formativa syften (Black et al., 2003), det vill säga att lära genom att gemensamt bearbeta prov och uppgifter både före och efter själva provtillfället (Black et al., 2010).

Bedömning som kompetensområde och som social process

För att förstå bedömningsprocessen och dess plats i skolan menar vi att bedömning måste ses från två olika perspektiv samtidigt: dels som ett utslag av lärares kunskaper om bedömning, dels som en social process som både styrs av och ger återverkningar på det sammanhang i vilket bedömningen ingår. Ibland sägs det att lärare måste få bättre utbildning i att bedöma olika former av kunskap. Det kan då handla exempelvis om att kunna tolka och konkretisera kunskapskraven i skolans kursplaner. Denna uttolkning förutsätter både djupa ämneskunskaper och djupa kunskaper om vad det vill säga att lära sig olika aspekter av ett visst ämne, till exempel vad det innebär att utvecklas som läsare eller att utvecklas som skribent. I den här boken beskriver vi några av dessa utvecklingssteg och försöker relatera det till vad det innebär att bedöma just läsande och skrivande i förhållande till kunskapskraven för skolåren 7–9.

Bedömningskunskap innefattar också goda principer för hur lärare samlar in relevant information om elevers kunnande, liksom en teoretiskt grundad reflektion över vilka slutsatser som egentligen går att dra baserat på ett visst underlag, vare sig det handlar om ett provresultat eller en elevs uppvisade deltagande i ett klassrumssamtal om en text. Detta kan kallas för kunskaper i bedömningsteori, något som den första delen i den här boken lägger särskild vikt vid.

Samtidigt sker all bedömning inom ramen för en kulturell kontext och tar alltid formen av en social process, som involverar någon som bedömer och som därmed utövar makt över den eller

de som bedöms (Stobart, 2008). Detta maktförhållande kan förstås se olika ut. I synnerhet kan elever göras mer och mindre delaktiga i processen och därigenom bli mer eller mindre medvetna om vad som faktiskt bedöms, hur det går till och vilka syften bedömningen tjänar (Gipps, 1999). En viktig dimension i de senare årens diskussion om bedömning i klassrummet har handlat just om elevers delaktighet och om värdet av självbedömning (se t.ex. Black & Wiliam, 2012; Hattie & Timperley, 2007).

I det långa loppet kommer också bedömningens utformning och innehåll att påverka såväl elever som lärare. Elever anpassar sig så att deras agerande och prestationer i klassrummet ska motsvara det som skolan anser nödvändigt att mäta. På samma vis anpassar lärare sin undervisning så att deras elever ska lyckas så bra som möjligt på nationella prov eller på lokalt standardiserade diagnoser. Detta eftersom resultat på sådana prov i allt högre utsträckning blivit ett sätt att utkräva ansvar inom ramen för utbildningssystemet. När en bedömning av elevers kunskaper får så vittomfattande konsekvenser blir det avgörande att försäkra sig om att bedömningen verkligen inriktas på de långsiktiga och komplexa målsättningarna i ämnet och inte bara på sådant innehåll som lätt låter sig mätas.

Forskning om bedömning i svenskämnet

Givet den betydelse bedömning har för elever, lärare och samhälle kan man lätt tro att det finns en mångfald av svensk forskning om klassrumsbedömning av läsning och skrivande. Så är inte fallet (Skar & Tengberg, 2014; Skar, 2013). En möjlig orsak till det skisseras av Parmenius Swärd (2008), som menar att bedömning uppfattades som något relativt okomplicerat fram till att målrelaterade betyg infördes i mitten av 1990-talet. En annan orsak *kan* vara det att bedömning, till skillnad från undervisning, inte uppfattas tillhöra skolans och lärares kärnverksamhet. Som genomgången ovan visat är det dock en missuppfattning.

Den forskning som trots allt har producerats kan delas upp i

olika kategorier. En möjlig kategorisering är följande: forskning som har med nationella provet att göra, explicit klassrumsanknuten bedömningsforskning och forskning som mer indirekt undersöker bedömning.

Den mesta bedömningsforskningen kopplad till svenskämnet består av undersökningar som tar sig an olika aspekter av det nationella provet. Denna forskning ställer viktiga frågor och ger viktiga svar knutet till betydelsefulla inslag i de bedömningsprocesser som lärare är en del av. Exempel på sådan forskning finns hos Borgström (2010, 2012, 2014) samt Borgström och Yassin (2010), som kritiskt har granskat vilka slutsatser om elevers skrivförmåga som är möjliga att dra på basis av uppgifter i de nationella proven, liksom vilket *konstrukt* (se kapitel 2 samt kapitel 10) som ligger till grund för provet. Tengberg (2014a, 2014b) har ställt liknande frågor kopplat till läsuppgifter. Även projektgruppen bakom det nationella provet har diskuterat och självkritiskt riktat sökljuset mot vad proven egentligen mäter och på vilka teoretiska grunder (Garme & Palmér, 1996; Garme, 2012; Palmér & Östlund-Stjärnegårdh, 2005; Palmér, 2011, 2013). Korp (2006) har bland annat med svenskdelen av nationella proven som utgångspunkt, ställt frågor om hur elever förbereds inför proven och hur proven används. Mickwitz (2011) har undersökt frågor om betyg, vilket ju är ett angränsande område, och fokuserat på vilka grunder betyg sätts och varför. Material från det nationella provet kommer också igen i studier som undersökt vad lärare egentligen fäster vikt vid (Östlund-Stjärnegårdh, 2002) och i undersökningar av bedömersamstämmighet (Borgström & Ledin, 2014; Östlund-Stjärnegårdh, 2009).²

En serie studier som också bör räknas till denna kategori är sådana som handlar om de internationellt jämförande läsförståelseproven. Här finns bland annat Skolverkets (2012) rapport rörande samstämmigheten mellan svenska styrdokument, nationella prov

2 Bedömersamstämmighet har också varit ett tema för undersökningar genomförda av Skolinspektionen (t.ex. 2013).

och PIRLS-provet. af Geijerstam och Folkeryd (2013) analyserar icke godkända elevsvar på PIRLS-provets uppgifter och härleder därifrån de normer som PIRLS etablerar för fenomenet »god läsförmåga«. Vad god läsförmåga innebär och hur den fångas står även i fokus för en senare studie från samma forskargrupp (Liberg, af Geijerstam & Folkeryd, 2014).

Forskning specifikt anknuten till klassrumsbedömning är mer sällsynt, men det finns exempel. Parmenius Swärd (2008) undersökte bland annat elevers upplevelser av skrivbedömning och Brorsson (2007) följde undervisnings- och bedömningsförlopp i högstadiets svenskundervisning. Även Skar (2013) följde skrivundervisning och skrivbedömning, men på gymnasiet, och ur ett validitetsperspektiv. Andra undersökningar (t.ex. Bergman-Claeson, 2003; Kronholm-Cederberg, 2009) har exempelvis inriktat sig på olika aspekter av lärarrespons eller lärarerfarenheter knutna till läsning (Westlund, 2013). Det finns också forskare som mer indirekt tar sig an frågor om bedömning. Exempel är Holmberg (bl.a. Holmberg & Wirdenäs, 2010; Holmberg, 2010), Nordenfors (2011), Bergh Nestlog (2012) och Randahl (2014), som alla studerar skrivpedagogik och som därmed bidrar med värdefull kunskap om vad som blir föremål för bedömning, utan att nödvändigtvis studera själva bedömningsprocessen.

Att svenskämnet i rätt liten utsträckning belysts ur ett bedömningsperspektiv innebär dock inte att det saknas relevant forskning. Det räcker att vi riktar uppmärksamheten mot vårt grannland i väst för att få många exempel på studier av skrivbedömning ur en mängd infallsvinklar. Exemplet på norsk skrivbedömningsforskning är så talrika att utrymmet bara medger att några enstaka projekt nämns. KAL-projektet (Berge, Evensen, Hertzberg, & Vagle, 2005) undersökte ungdomars skrivförmåga ur en rad perspektiv. SKRIV-projektet (t.ex. Smidt, 2010) undersökte skrivundervisning i olika ämnen i olika årskurser, bland annat med bedömningsfokus. NORM-projektet (Berge et al., 2011) är det senaste – och största – och det undersöker bland annat lärares skrivbedömningsnormer,

hur olika skrivuppgifter fungerar och vad som händer med elever som utsätts för systematisk skrivundervisning. Ett annat projekt är det så kallade Läringsstöttande prøver i skrivning, som är ett skrivbedömningsverktyg utformat för formativ bedömning. Verktöget är resultatet av flera års forsknings- och utvecklingsarbete (Berge, 2005; Evensen, 2010, 2014). Den norska forskningen om bedömning av läsning är möjligen inte lika omfattande som forskningen om skrivbedömningen, men här finns exempel på studier av hög vetenskaplig standard som bland annat berör utformningen av testuppgifter i internationella läsförståelseprov (Solheim & Skafstun, 2009), liksom uppgiftsutformningens betydelse för elevers självupplevda förmåga till läsförståelse (Solheim, 2011).

Som genomgången ovan visar finns det mycket forskning som på ett eller annat sätt kan relateras till den bedömningsverksamhet som lärare ägnar sig åt i sin svenskundervisning. Samtidigt visar den rätt långa listan att klassrumsbedömning som process sällan tematiseras i svensk forskning. Så tycks vara fallet även internationellt. Ytterligare en förklaring, vid sidan av de som presenterades inledningsvis, kan vara att andra bedömningsformer och -situationer får en större och mer direkt inverkan på elevers liv. Det gäller exempelvis nationella prov och olika typer av examensprov, som kan leda till stora konsekvenser för individer och grupper. Dessutom verkar det saknas modeller och tankeredskap för att undersöka klassrumsbedömning (jfr Bonner, 2013; se också Newton & Shaw, 2014). I den här boken erbjuds emellertid några sådana modeller och tankeredskap.

Bokens innehåll

I den här boken fokuseras aspekter av bedömning av skrivande och läsande i klassrummet. Avsikten är att på ett ämnesdidaktiskt relevant sätt förklara bedömningsprocessen med utgångspunkt i aktuell bedömningssteori. Boken är uppbyggd av två delar. I den första delen presenteras de allmänna och teoretiska utgångspunkterna för

frågor om bedömning av färdigheter som läsande och skrivande. De kapitel som ingår i denna del behandlar bedömningars olika syften och vad som kännetecknar tillförlitliga bedömningar, men de handlar också mer specifikt om hur uppgifter utformas för att samla information om elevers kunnande liksom om de färdigheter som ska bedömas, det vill säga skrivande och läsande. I bokens andra del tillämpas några av dessa utgångspunkter i fyra empiriska kapitel som beskriver autentiska klassrumssituationer från ett lärarperspektiv. Här diskuteras lärares val av uppgifter för att kunna samla relevant information om elevers skrivande och läsande. Författarna, som alla har erfarenhet av både undervisning i och forskning om skolämnet svenska, diskuterar val av bedömningsformer och former för återkoppling till eleverna med ambitionen att på ett så adekvat sätt som möjligt stödja elevernas fortsatta progression i sitt skrivande och läsande.

Här följer en närmare beskrivning av bokens olika kapitel.

DEL 1: TEORETISKA UTGÅNGSPUNKTER

I det första kapitlet ger *Gudrun Erickson* en introduktion till frågan om bedömningars olika syften i skolan och på vilket sätt dessa relaterar till varandra. Erickson redogör bland annat för de frågor en lärare måste ställa sig inför själva bedömningsprocessen i klassrummet, till exempel *varför* bedömning av elevernas kunskaper genomförs, *vad* det är som ska bedömas och *hur* det ska gå till. Detta är klassiska didaktiska frågeställningar, men till dessa hör också frågor om *när* bedömning ska ske – är det till exempel så att allt eleverna gör i klassrummet ska bedömas – och *vem* som gör vad i bedömningsituationen. Här diskuteras också begreppen *formativ* och *summativ bedömning* liksom hur bedömningsprocessen hänger ihop med undervisningen och elevers lärande.

I det andra kapitlet ger *Gustaf Skar* en introduktion till frågan om validitet och reliabilitet i bedömning. Med utgångspunkt i några konkreta exempel, som alla relaterar till bokens ämnesmässiga

fokus, beskrivs grunderna för den validitetsmodell som utvecklats av Michael T. Kane och som också ligger till grund för den här boken. Att bedöma komplexa färdigheter som skriv- och läsförmåga handlar om att dra slutsatser baserade på underlag i form av vissa elevprestationer, till exempel texter eller lösningar på något slags provuppgifter. Men hur kan vi argumentera för att de slutsatser vi drar är rimliga givet underlaget och syftet med bedömningen? Om detta handlar det andra kapitlet.

Det tredje kapitlet handlar om skrivande som färdighet och har författats av *Ragnar Thygesen* och *Kjell Lars Berge*. Utgångspunkten för kapitlet är ett norskt skolperspektiv på skrivande, i vilket skrivande jämte läsning, räkning och digitala färdigheter betraktas som en grundläggande färdighet, som är både ämnesspecifik och ämnesövergripande. Även om svensk läroplan inte på samma sätt har formulerat skrivande som en grundläggande färdighet i alla ämnen, menar vi att perspektivet är av relevans för svenska förhållanden. Också i Sverige har förväntningarna ökat på att betrakta skrivande och läsande inte bara som ämnesövergripande kompetenser, utan även som ämnesspecifika komponenter, något som i den här boken utvecklas vidare i kapitel sju. I Thygesens och Berges kapitel beskrivs framför allt skrivandets kommunikativa funktioner och författarna presenterar den teoretiska modell över skrivande som färdighet (Skrivhjulet) som bland annat har utvecklats av dem själva. I texten relateras modellen och kriterier för bedömning av skrivande till svenska förhållanden genom kopplingar till kursplanen i svenska.

I det fjärde kapitlet ger *Michael Tengberg* på motsvarande vis en introduktion till den andra grundläggande färdighet som står i fokus för boken, nämligen läsande, eller närmare bestämt läsförmåga. Syftet är i första hand att definiera läsförmåga som fenomen, vilket görs med utgångspunkt dels i kursplanen i svenska, dels i det teoretiska ramverket för PISA:s läsprov. Tengberg argumenterar för att läsförmåga inte kan förstås som en enhetlig kunskapsform utan bör ses som någonting som är sammansatt av flera olika kunskapsdimensioner. För bedömningens vidkommande innebär det

att valida bedömningar av läsförmåga måste grunda sig på bevis som tillvaratar den varians av läsförmåga som exempelvis olika texttyper och lässammanhang möjliggör.

I det femte kapitlet sammanfattar *Gustaf Skar* erfarenheterna från den teoretiska delen av boken under rubriken »Uppgifter«. En uppgift med tillhörande bedömningsanvisningar utgör själva navet i en bedömningsprocess. Uppgifter »operationaliserar« tolkningar av kursplaner och uppfattningar om skrivande och läsande. När en uppgift fungerar optimalt leder den till att elever utför vissa på förhand bestämda handlingar som låter sig tolkas med hjälp av bedömningsverktyg. I kapitlet ges exempel på olika typer av uppgifter och bedömningsredskap och vad som är användbart i förhållande till olika bedömningssyften.

DEL 2: BEDÖMNING I PRAKTIKEN

Det sjätte kapitlet inleder den empiriska delen av boken och är skrivet av *Karina Pålsson Gröndahl*. Här presenteras ett skrivförlopp där elever i årskurs 8 får i uppgift att skriva ett brev till USA:s president Barack Obama. Bedömningsinslaget är i huvudsak formativt och syftar till att identifiera styrkor och svagheter i elevernas texter, samt att genom återkoppling visa på utvecklingsmöjligheter i deras skrivande. Pålsson Gröndahl ger en detaljerad bild av en lärares sätt att resonera i bedömningsituationen, bland annat med utgångspunkt i en bedömningsmatris som syftar till att konkretisera hur kunskapskraven kan komma till uttryck inom ramen för skrivuppgiften. Hon redovisar också exempel på hur återkopplingen formuleras till eleverna och diskuterar i vad mån den fyller sitt syfte givet bokens bedömningsteoretiska utgångspunkter.

I det sjunde kapitlet för *Jenny W. Folkeryd* och *Åsa af Geijerstam* oss utanför svenskämnets omedelbara domän. Här handlar det nämligen om skrivande och framför allt om bedömning av elevtexter i biologiämnet. Ämnesinnehåll, menar Folkeryd och af Geijerstam, skapas genom språket och därför kommer bedömning av elevernas

ämneskunnande med nödvändighet också att innefatta bedömning av hur detta ämneskunnande struktureras och formuleras språkligt. Med utgångspunkt i kursplanen i biologi diskuteras vad skillnaden kan vara mellan »enkla«, »utvecklade« och »välutvecklade« texter inom ämnet. Diskussionen förs med grund i en konkret skoluppgift och autentiska elevtexter, och tillämpar textteoretiska perspektiv på dessa.

Det åttonde kapitlet är skrivet av *Eva Nilson* och handlar om undervisning i och bedömning av litterär läsning, närmare bestämt läsning av en roman. Med utgångspunkt i kursplanens kunskapskrav formulerar Nilson en uppgift där eleverna får välja mellan tre olika ungdomsromaner. Boken läses och analyseras skriftligt med successiv återkoppling från läraren. Den löpande bedömningen av de olika delarna i analysen (personer, miljö, berättarteknik, handling etc.) blir på så vis redskap för att utveckla elevernas analyser och deras förhållningssätt till romanen som litterär form.

I det nionde kapitlet beskriver *Susanne Staf* hur en lärare undervisar om och bedömer elevers analys av en novelltext. Arbetet i klassrummet kretsar kring John Ajvide Lindqvists novell *Pappersväggar* från novellsamlingen med samma namn. Från elevernas analysstexter samlar Staf inte bara information om elevernas kunskaper när det gäller läsförståelse och förmåga att använda gängse litterära analysbegrepp, utan hon visar också på lärarens reflektioner över den egna undervisningen och på vilket sätt formativ bedömning används som ett redskap för att planlägga kommande lektioner i klassen.

I bokens andra del finns också ett kapitel, kapitel 10, för den som vill fördjupa sina kunskaper ytterligare när det gäller validitet i bedömning. Kapitlet är skrivet av *Michael T. Kane*, en av de mest välrenommerade forskarna på området bedömning och mätning av elevers kunskaper. Kapitlet heter kort och gott »Validitet« och beskriver den så kallade argumentbaserade ansatsen till validitet. Denna modell har vunnit stort gehör och är en i det närmaste obligatorisk referens i samtida validitetsforskning. Kapitlet beskriver också validitetsbegreppets historia.

Referenser

- Berge, K.L. (2005). Skrivning som grunnleggende ferdighet og som nasjonal prøve – ideologi og strategier. I A.J. Aasen & S. Nome (red.), *Det nye norskfaget* (s. 1–22). Oslo: Fagbokforlaget.
- Berge, K.L., Evensen, L.S., Fasting, R.B., Matre, S., Solheim, R., Thygesen, R., . . . Solem, H. (2011). *Developing national standards for the teaching and assessment of writing. Rapport fra forprosjekt Utdanning 2020*. Trondheim: Nasjonalt senter for skriveopplæring og skriveforskning.
- Berge, K.L., Evensen, L.S., Hertzberg, F. & Vagle, W. (red.). (2005). *Ungdommers skrivekompetanse (Bind I & II)*. Oslo: Universitetsforlaget.
- Bergh Nestlog, E. (2012). *Var är meningen? Elevtexter och undervisningsspraktiker*. Växjö: Linnéuniversitetet.
- Bergman-Claeson, G. (2003). *Tre lärare – tre världar: lärarkommentarer till elevtexter i tre gymnasieklasser*. Uppsala: Uppsala universitet.
- Björklund Boistrup, L. (2013). *Bedömning i matematik pågår! Återkoppling för elevers engagemang och lärande*. Stockholm: Liber.
- Black, P. & Harrison, C. (2014). *Bedömning för lärande i NO-klassrummet*. Stockholm: Liber.
- Black, P., Harrison, C., Hodgen, J., Marshall, M. & Serret, N. (2010). Validity in teachers' summative assessments, *Assessment in Education*, 17(2), 215–232.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003). *Assessment for learning: putting it into practice*. Buckingham, UK: Open University Press.
- Black, P. & Wiliam, D. (2012). Assessment for learning in the classroom. I J. Gardner (red.), *Assessment and learning* (2 uppl.) (s. 11–32). London: SAGE.
- Bonner, S.M. (2013). Validity in classroom assessment: Purposes, properties and principles. I J.H. McMillan (red.), *SAGE handbook of research on classroom assessment* (s. 87–107). Los Angeles: SAGE Publications, Inc.
- Borgström, E. (2010). Att skriva prov. Om normer och textstruktur i gymnasieskolans skriftliga nationella prov. *Språk och stil*, 20(1), 132–164.

- Borgström, E. (2012). Skrivförmåga på prov. I G. Skar & M. Tengberg (red.), *Svenskämnet i går, i dag, i morgon* (s. 209–223). Stockholm: Svenskläraryöreningen.
- Borgström, E. (2014). Vad räknas som belägg för skrivförmåga? Ett textkulturellt perspektiv på skrivuppgifter i den svenska gymnasieskolans nationella prov. *Sakprosa*, 6(1), 1–34.
- Borgström, E. & Ledin, P. (2014). Bedömarvariation. Balansen mellan teknisk och hermeneutisk rationalitet vid bedömning av skrivprov. *Språk och stil*, 24.
- Borgström, E. & Yassin, D. (2010). Tioåringar skriver nationellt prov. I K. Adelman (red.), *Sjunde nationella konferensen i svenska med didaktisk inriktning* (s. 25–33). Malmö: Nationella nätverket för svenska med didaktisk inriktning.
- Brorsson, B.N. (2007). *Man liksom bara skriver: skrivande och skrivkontexter i grundskolans år 7 och 8*. Örebro: Örebro universitet.
- Dobson, S., Eggen, A.B. & Smith, K. (red.) (2009). *Vurdering, prinsipper og praksis*. Oslo: Gyldendal akademisk.
- Evensen, L.S. (2010). En gyldig vurdering av elevers skrivekompetanse? I J. Smidt (red.), *Rammer for skrivning: om skriveutvikling i skole og yrkesliv* (s. 13–31). Trondheim: Tapir akademisk forlag.
- Evensen, L.S. (2014). Men kan det komme noe godt fra lærerværelset? Om validitet i læreres vurdering. I R. Hvistendahl & A. Roe (red.), *Alle tiders norskdidaktiker* (s. 245–258). Oslo: Novus forlag.
- Gardner, J. (2012). Assessment and learning: introduction. I J. Gardner (red.), *Assessment and learning*. (2 uppl.) (s. 1–9). London: SAGE.
- Garme, B. (2012). Med proven i tiden eller prov i tid och otid! I G. Skar & M. Tengberg (red.), *Svenskämnet i går, i dag, i morgon* (s. 224–233). Stockholm: Svenskläraryöreningen.
- Garme, B. & Palmér, A. (1996). *Samspel och kommunikation: om utvecklandet av nationella kursprov i svenska för gymnasieskolan*. Uppsala: Uppsala universitet.
- af Geijerstam, Å. & J.W. Folkeryd (2013). *Fel men rätt på något sätt? En analys av elevers nollsvaret enligt PIRLS-provet*. I G. Skar & M. Tengberg (red.), *Läsning!* (s. 37–55). Stockholm: Natur & Kultur.
- Gipps, C. (1999). Chapter 10: Socio-cultural aspects of assessment. *Review of Research in Education*, 24(1), 355–392.

- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Holmberg, P. (2010). Genrepedagogik i teori och praktik. Nyretorik och Sydneyskolan i två gymnasielärares klassrum. I C. Falk, A. Nord & R. Palm (red.), *Svenskans beskrivning 30* (s. 123–132). Stockholm: Institutionen för nordiska språk, Stockholms universitet.
- Holmberg, P. & Wirdenäs, K. (2010). Skrivpedagogik i praktiken. *Språk och stil*, 20(1), 105–131.
- Jönsson, A. & Odenstad, C. (2014). *Bedömning i SO: för grundskolan*. Malmö: Gleerups.
- Korp, H. (2006). *Lika chanser i gymnasiet? En studie om betyg, nationella prov och social reproduktion*. Lund: Lunds universitet.
- Kronholm-Cederberg, A. (2009). *Skolans responskultur som skriftpraktik: gymnasieters berättelser om lärarens skriftliga respons på uppsatsen*. Åbo: Åbo Akademi.
- Leslie, L. & Caldwell, J. (2009). Formal and informal measures of reading comprehension. I S.E. Israel, & G.G. Duffy (red.), *Handbook of research on reading comprehension* (s. 403–427). New York, NY: Routledge.
- Liberg, C., Folkeryd, J.W. & af Geijerstam, Å. (2014). God läsförmåga – hur fångas den? I S. Bagga-Gupta, A.-C. Evaldsson, C. Liberg & R. Säljö (red.), *Literacy-praktiker i och utanför skolan* (s. 87–106). Malmö: Gleerups.
- Lundahl, C. (2011). *Bedömning för lärande*. Stockholm: Norstedts.
- Mickwitz, L. (2011). *Rätt betyg för vem? Betygsättning som institutionaliserad praktik*. Stockholm: Stockholms universitet.
- Newton, P.E. & Shaw, S.D. (2014). *Validity in educational and psychological assessment*. Thousand Oaks, CA: SAGE Publications.
- Nordenfors, M. (2011). *Skriftspråksutveckling under högstadiet*. Göteborg: Göteborgs universitet.
- Palmér, A. (2008). *Samspel och solostämmor: om muntlig kommunikation i gymnasieskolan*. Uppsala: Uppsala universitet.
- Palmér, A. (2011). Vad ska bedömas? Gymnasiets nationella prov i svenska och svenska som andraspråk. I M. Ellvin, G. Skar & M. Tengberg (red.), *Svenskämnet i förändring?* (s. 143–155). Stockholm: Svensk lärarföreningen.

- Palmér, A. (2013). Skrivdidaktiska diskurser bakom gymnasieskolans nationella prov: Genre, kommunikationssituation och skrivdidaktiska diskurser. I J.-A. Chrystal & M. Lim Falk (red.), *Tionde nationella konferensen i svenska med didaktisk inriktning, Stockholm 18–19 oktober 2012* (s. 139–148). Stockholm: Nationella nätverket för svenska med didaktisk inriktning.
- Palmér, A. & Östlund-Stjärnegårdh, E. (2005). *Bedömning av elevtext: en modell för analys*. Stockholm: Natur & Kultur.
- Parmenius Swärd, S. (2008). *Skrivande som handling och möte: Gymnasieelever om skrivuppgifter, tidsvillkor och bedömning i svenskämnet*. Lund: Lunds universitet.
- Randahl, A.-C. (2014). *Strategiska skribenter: skrivprocesser i fysik och svenska* (Doktorsavhandling, Örebro universitet).
- Sandin, D. (2009). *Professionell bedömning i svenska*. Malmö: Epago.
- Skar, G. (2013). *Skrivbedömning och validitet. Fallstudier av skrivbedömning i svenskundervisning på gymnasiet*. Stockholm: Stockholms universitet.
- Skar, G. & Tengberg, M. (2014). Vilken forskning bedrivs egentligen inom forskningsfältet svenska med didaktisk inriktning? I: P. Andersson, P. Holmberg, A. Lyngfelt, A. Nordenstam, & O. Widhe (red.), *Mångfaldens möjligheter. Litteratur- och språkdidaktik i Norden* (s. 353–374). Göteborg: Nationella nätverket för svenska med didaktisk inriktning.
- Skolinspektionen (2013). *Olikheterna är för stora, regeringsrapport, Omvärdering av nationella prov i grundskolan och gymnasieskolan, 2013*. Stockholm: Skolinspektionen.
- Skolverket (2012). *Med fokus på läsande. Analys av samstämmigheten mellan svenska styrdokument, ämnesprov i svenska och PIRLS 2011*. Stockholm: Fritzes.
- Smidt, J. (2010). Skrivekulturer og skrivesituasjoner i bevegelse – fra beskrivelser til utvikling. I J. Smidt (red.), *Skrijving i alle fag – innsyn og utspill* (s. 11–35). Trondheim: Tapir akademisk forlag.
- Solheim, O.J. (2011). Impact of reading efficacy on reading comprehension scores in different item formats. *Reading Psychology*, 32, 1–27.
- Solheim, O.J. & Skaftun, A. (2009). The problem of semantic openness and constructed response. *Assessment in Education*, 16, 149–164.
- Stobart, G. (2008). *Testing times: the uses and abuses of assessment*. London: Routledge.

- Tengberg, M. (2014a). Att pröva elevers läsförmåga i årskurs nio. En analys av uppgiftskonstruktioner i det nationella provet i svenska, delprov A: Att läsa och förstå. *Utbildning & Demokrati*, 23(2), 109–132.
- Tengberg, M. (2014b). Konstruktion och bedömning av förmågan att läsa (och förstå) skönlitterär text. Två nedslag i det nationella provets läsförståelsedel, åk 9. *Educare*, 2014:1, 79–97.
- Westlund, B. (2013). *Att bedöma elevers läsförståelse: en jämförelse mellan svenska och kanadensiska bedömningsdiskurser i grundskolans mellanår*. Stockholm: Natur & Kultur.
- Östlund-Stjärnegårdh, E. (2002). *Godkänd i svenska?: Bedömning och analys av gymnasieelevers texter*. Uppsala: Uppsala universitet.
- Östlund-Stjärnegårdh, E. (2009). Bedömaröverensstämmelse – ämnet Svenska. I: *Bedömaröverensstämmelse vid bedömning av nationella prov*. 2009-04-14. Dnr 2008:286. Stockholm: Skolverket.

Del 1

Teoretiska utgångspunkter

1. Kunskapsbedömningars grund och syften

GUDRUN ERICKSON

Att vara lärare innebär att man deltar i, medverkar till och påverkar andra människors utveckling och lärande. Detta är givetvis något mycket positivt, inspirerande och utmanande, men det innebär också ett stort ansvar och kräver gedigna kunskaper av skilda slag. På ett övergripande plan används stundtals begreppen *lärande*, *undervisning* och *bedömning* som sammanfattande beskrivning av de huvudområden som kännetecknar lärares professionella uppdrag. Dessa tre företeelser hör ihop på ett påtagligt sätt och bildar tillsammans en dynamisk pedagogisk och didaktisk helhet.

Temat för den här boken, bedömning, är en av komponenterna i den här helheten. För detta krävs specifika kunskaper, *assessment literacy*, som det ibland rubriceras som, och som en bok av detta slag lämnar ett viktigt bidrag till. Fokus för denna inledande text i boken är en speciell aspekt, nämligen vad som ligger bakom de kunskapsbedömningar som görs i skolan.

Det kommer med andra ord att handla om de olika syften som kan identifieras. Innan vi fördjupar oss i detta ska här först nämnas något om fenomenet bedömning i sig, hur det definieras och uppfattas, och hur det hör ihop med lärande och undervisning. I slutet av kapitlet berörs också kort externa bedömningar i form av nationella prov.

Definitioner och uppfattningar

Bedömning är på samma gång ett både lätt och svårt ord; lätt för att det används på ett så påtagligt och till synes självklart sätt, inte bara i skolans värld utan generellt, men samtidigt svårt på grund av de många tolkningar och associationer som finns och som dessutom inte alltid är vare sig tydliga eller synliga.

Om man letar i ordböcker finner man ett stort antal synonymer till bedömning: *värdering*, *granskning*, *prövning* och *omdöme*, men också *kritik*, *betygssättning* och *dom*. Ett problem är troligen att ordet i sig är besläktat med *döma*. Man brukar tala om ordets makt över tanken, vilket torde gälla i det här fallet, där dessutom bedömning är en synnerligen icke-neutral företeelse som kommer människor riktigt nära och som därför många har synpunkter på. Detta problem är mindre i det engelska språket, där det motsvarande ordet *assessment* inte uppvisar vare sig samma släktskap eller värdeladdade associationer utan har en mera generell och neutral betydelse av uppskattning och (värde-)omdöme. Det är också i denna mera allmänna bemärkelse ordet bedömning kommer att behandlas i den här texten.

Uppfattningarna om bedömning – och dess i vissa sammanhang formella slutpunkt i form av betyg – är många och kan ha såväl personliga som politiska förtecken. De negativa synpunkter som framförs handlar inte sällan om risken för att bedömning tar tid från lärande och undervisning och att mätningen tar överhand, så att det lättast mätbara görs till det viktigaste. Dessutom varnas stundtals för bedömningens potentiellt eller reellt disciplinerande funktion, som ett maktmedel (Bernstein, 1996; Foucault, 1977). Det finns emellertid också uttalat positiva uppfattningar om bedömning som ett viktigt pedagogiskt kommunikationsmedel, som bidrar till elevens successivt ökande insikter om sina egna kunskaper i relation till det som ska läras (Gardner, 2012). Dessutom kan man i somliga politiska kretsar spåra en stark tilltro till bedömning, oftast i form av prov och betyg, som stimulansmedel för elever att lära sig mer och bättre. Den allra vanligaste uppfattningen är dock

sannolikt att bedömning i skolan är något tämligen självklart, ett nödvändigt ont – eller gott – som man som lärare har att hantera på bästa sätt.

Bedömning som del av en helhet

Man hör ibland lärare säga att det inte riktigt var detta de tänkt sig, när de en gång valde att bli lärare, där »detta« står som ett sammanfattande ord för bedömning och dokumentation, samt i vissa fall betygssättning. Yttrandet kan tolkas på olika sätt, bland annat som en suck över att dokumentationskravet i skolan har ökat på ett markant sätt och kanske även att betyg tillkommit där de tidigare inte fanns. Det är svårt att avfärda den här sortens kommentar, även om det finns olika sätt att hantera den beskrivna situationen, såväl individuellt som kollegialt och strukturellt. Men »detta« kan också stå för en mera innehållslig och principiell fråga, nämligen hur bedömning kommer in i och blir en integrerad del i den pedagogiska och didaktiska helheten. Ett sätt att analysera detta är att se på en relation i taget.

Förhållandet mellan lärande och undervisning torde vara relativt okomplicerat, även om det är viktigt att påpeka att kunskap också förvärvas och utvecklas informellt, utanför skolans ram. Det finns sannolikt inte heller särskilt många invändningar mot påståendet att undervisning, förmedlad av kunniga och engagerade lärare i gott samarbete med sina elever, främjar lärande. Detta ska dock inte tolkas som att undervisning automatiskt resulterar i ett oproblematiskt och linjärt lärande; människor lär sig på många olika sätt och utvecklingen går både upp och ned, hit och dit, fram och tillbaka – och står ibland till synes stilla under lång tid, för att därefter fortsätta. Det finns heller inga standardmodeller vare sig för bra lärande eller bra undervisning, men på ett generellt plan kan ändå konstateras att bra undervisning stödjer och stimulerar avsett lärande.

Likaså kan det med lätthet hävdas att bra undervisning förutsät-

ter bra bedömning. För att hjälpa alla elever att komma så långt som möjligt behöver läraren kunna se det som ska ses på många sätt, och med så god skärpa som möjligt. För att kunna kartlägga och beskriva individuella kunskapsprofiler samt göra vägen framåt så bra som möjligt krävs olika bedömningsformer, inklusive vilja och förmåga att involvera eleven själv i analys och planering. Detta kan ses som professionell pedagogisk och didaktisk kompetens, vilket följaktligen innefattar både god undervisnings- och bedömningskunskap.

Undervisning och bedömning är inte identiska företeelser, men de har ett antal gemensamma, ibland identiska, drag (Erickson, 2009). Exempel på detta är relevans och bredd, både visavi underliggande styrdokument och vad gäller metoder. Vid avslutad kurs eller skolform måste hela det aktuella kunskapsinnehållet ha täckts och utvärderats så väl som möjligt. Vidare bör både undervisning och bedömning präglas av tydlighet, lyhördhet, ansvar, rättvisa och respekt. Detta handlar bland annat om en strävan efter så stor transparens som möjligt, både vad gäller mål och metod. Lyhördhet innebär att man funderar lite på och drar slutsatser av det faktum att de flesta människor har två öron och två ögon – men bara en mun; det är väsentligt att lyssna och se, snarare än att snabbt själv leverera lösningar och svar. Dessutom har många av de egenskaper som präglar både bra undervisning och bra bedömning tydliga etiska dimensioner. Dessa har med förhållningssättet till andra och varandra att göra och bör karakterisera den samverkan som finns mellan alla inblandade parter i den pedagogiska processen.

Det som dock tydligast visar att lärande, undervisning och bedömning är ouppslösligt förbundna med varandra kan enkelt uttryckas som att målet är detsamma: både undervisning och bedömning syftar till en så positiv utveckling och ett så bra lärande som möjligt (Erickson & Gustafsson, 2014). Därmed kan vi också fastslå att ett högst väsentligt, övergripande syfte med bedömning är att stödja lärande och undervisning, genom att dels bidra med adekvata iakttagelser av nuläget, dels ge tydlig, konstruktiv och brukbar information vad gäller det fortsatta arbetet.

De männe eviga frågorna

Oavsett vilken typ av bedömning vi talar om finns det ett antal grundläggande frågor som måste ställas inför planering och genomförande av olika typer av åtgärder. Även här kan samma frågor i regel appliceras på helheten och alltså också innefatta lärande och undervisning, även om vi just här fokuserar på bedömning.

Den första frågan är den som står i centrum för det här kapitlet, nämligen *VARFÖR?* Den handlar följaktligen om bedömningsars syften, vilket är av fundamental betydelse såväl för planering och metodval som för analys och användning av resultat. Det går inte att springa ikapp den här frågan, utan den måste både ställas och besvaras innan man sätter igång. Vi återkommer strax till ett längre resonemang kring detta.

Nästa fråga som måste besvaras är *VAD?*, vilket innebär fokus på kunskapsobjektet, som på engelska refereras till som *the construct* (på svenska stundtals även kallat *konstrukt*). Denna fråga är självklart central, eftersom den innebär en konkretisering av det som ska läras, undervisas och därmed bedömas, det vill säga det som styrdokumentet förmedlar. Kunskap är en komplex företeelse och det handlar här både om produkt och process, del och helhet.

HUR? är nästa fråga och den omfattar allt från bedömningsmetoder och uppgiftstyper (se kapitel 5) till beslut kring kravnivåer och återkoppling. Centrala aspekter här är inte minst balansen mellan muntlig, skriftlig och handlingsbaserad bedömning, relationen mellan uppgifter med mera bundna respektive öppna responsformer, holistisk respektive analytisk bedömning, fastställande av krav- och betygsnivåer, återkoppling av resultat samt samverkan i bedömningen med syfte att öka samstämmigheten.

I frågan om *hur* bedömningar ska genomföras ligger även inbäddad en aspekt av tid, det vill säga frågan *NÄR?* Detta innebär beslut om såväl tidpunkter som frekvenser och har delvis även att göra med bedömningens syfte.

Frågan *VEM?* fokuserar främst på subjektfrågan, det vill säga vem som gör vad i olika bedömningsituationer. Man kan

här urskilja flera aktörer: elever, lärare, skolledare samt externa personer och företeelser, de senare bland annat i form av utifrån kommande bedömningsmaterial. I grova drag kan det vara bra att göra en distinktion mellan fyra huvudkategorier av bedömning, nämligen *självbedömning*, *kamratbedömning*, *lärarbedömning* och *extern bedömning*.

Viktigt att säga är dock att dessa olika former inte bara är komplementära utan också stundtals samverkar. En annan ytterst väsentlig aspekt av VEM-frågan handlar om objektsfrågan, nämligen vad det är som bedöms. Bedömning handlar om att så bra som möjligt kartlägga utveckling och kunnande i relation till uppställda mål, inte om att värdera människor eller mänskliga kvaliteter. Det är därför av stor betydelse att vakta på sina ord och betona att det inte är elever som bedöms utan deras kunskaper och kunskapsutveckling. Bachman (1990, s. 20) uttrycker det på följande sätt:

Whatever attributes or abilities we measure, it is important to understand that it is these attributes or abilities and not the persons themselves that we are measuring.

Slutligen är den uppfordrande frågan *OCH...?* av central betydelse, nämligen hur bedömningar och deras resultat används och vilka konsekvenserna av dem blir, såväl för enskilda elever och deras självbild, motivation och arbetsglädje, och i vissa fall deras framtidsmöjligheter, som för lärare, skolor och samhället. Det är här återigen viktigt att betona att bedömningar inte är neutrala företeelser, och att de alltid har effekter. Man kan i detta fall tala om en oundviklig påverkan. Den viktigaste frågan i detta är dock huruvida denna påverkan är positiv eller negativ och, inte minst, vad man som lärare kan göra för att det förstnämnda ska bli fallet, så att de bedömningar som görs främjar lärande och utveckling.

Grundläggande principer och begrepp

Motiven för och sätten att göra kunskapsbedömningar varierar, men det är viktigt att betona att det finns grundläggande principer som gäller för alla typer av kunskapsbedömning. Dessa kan formuleras på lite olika sätt men brukar alltid innefatta begreppen *tydlighet*, *giltighet*, *tillförlitlighet* och *respekt*. Detta finns framskrivet bland annat i de Riktlinjer för god praxis vid prov och bedömning i språk som tagits fram inom EALTA (European Association for Language Testing and Assessment) och som finns översatta till 35 språk och är fritt tillgängliga på nätet (www.ealta.eu.org/guidelines). Begreppen som används torde vara tämligen självförklarande och är alla delaspekter av begreppet *validitet*, som behandlas utförligt av Kane (se kapitel 10). I grunden handlar detta om hur bedömningar utformas, men ännu mer om hur de används, vilka slutsatser som dras, vilka beslut som fattas och vilka åtgärder som vidtas. I detta inkluderas även begreppet *tillförlitlighet*, alltså i vilken utsträckning resultaten går att lita på och bidrar till rättvisa i bedömningen. Detta rubriceras på fackspråk som *reliabilitet*, vars budskap kanske blir ännu tydligare med det norska ordet *rettferdighet*. I anslutning till detta kan vidare en etisk konsekvensaspekt läggas till begreppet *validitet*, nämligen att man i sin planering och användning av bedömning inkluderar de effekter som kan tänkas bli (Kane 2006; Messick, 1989). I detta brukar två huvudsakliga faror nämnas, nämligen att man har med för lite av det som ska bedömas – att man kanske smalnar av sin bedömning så mycket att endast det lätt mätbara kommer med och därmed framstår som det viktigaste – eller att man tvärtom, medvetet eller omedvetet, inkluderar för mycket i sina bedömningar, sådant som inte hör till det man avser att bedöma. På fackspråk rubriceras detta som *underrepresentation av konstruktet* (*construct under-representation*) respektive *konstruktirrelevant varians* (*construct-irrelevant variance*), vilket kan låta tekniskt och avskräckande, men som i själva verket är högst vardagliga företeelser som de flesta lärare är väl medvetna om.

Bedömnings syften

Så tillbaka till den första och mest grundläggande av de tidigare nämnda frågorna, och den som också är i fokus i detta kapitel, nämligen *VARFÖR?* – med vilket syfte – bedömningar görs. Eftersom svaret på frågan kommer att påverka både utformning och genomförande, måste det vara klart innan arbetet kan gå vidare. Skolverket (2011a, s. 7) väljer att besvara frågan med fem mycket vida dimensioner relaterade till skolans uppdrag, nämligen att bedömning syftar till att:

1. Kartlägga kunskaper
2. Värdera kunskaper
3. Återkoppla lärande
4. Synliggöra praktiska kunskaper
5. Utvärdera undervisning

Dessa funktioner säger följaktligen något om den övergripande intentionen utifrån vilken operativa beslut har att fattas; de är alltså ett slags riktninggivare som behöver konkretiseras och som kommer att resultera i praktiska åtgärder, som i sig kan ha olika karaktär.

Det finns även andra sätt att utifrån syften försöka sammanfatta och exemplifiera olika bedömningsformer. Lindblad (1985) nämner följande tio motiv att pröva kunskaper i utbildningssammanhang:

Diagnos, resursanvändning, urval, vägledning, examination, betygssättning, utvärdering, påverkan, stimulans och forskning (författarens övers.).

Lindblad betonar att prov som har olika syften i regel inte går att byta ut mot varandra, vilket han menar är något som ofta förbises. Samma resonemang kan med lätthet föras även kring andra bedömningsformer. Några kommentarer kan vara på sin plats vad gäller de angivna syftena. Det är till exempel viktigt att komma ihåg att en

diagnos syftar till att få fram information som leder till ändamålsenliga och effektiva åtgärder. Detta innebär förvisso inte bara att söka efter brister för att kunna påverka och avhjälpa dessa, utan också om att visa på styrkeområden i elevers enskilda kunskapsprofiler. För att man under sitt lärande ska kunna, vilja och orka ta itu med det svåra och mindre självklara, behöver man också kunna få känna sig kapabel och ges möjlighet att flyga ännu högre inom områden där styrkorna finns. Detta är gammal, praktisk lärarkunskap som tyvärr ofta glöms bort – inte minst i samband med diagnostisk bedömning. Samma handfasta konstaterande kan göras vad gäller *påverkan*, där man inte bör underskatta den starka signaleffekt bedömningar har vad gäller elevers uppfattning om lärandet, inte minst ur perspektivet vad eleverna tror att läraren anser viktigt, eller allra viktigast. Det finns inte särskilt mycket forskning kring detta fenomen, och den som finns avser ofta påverkan av utifrån kommande prov. Dock är det klokt att som lärare inse att det man ofta fokuserar på i undervisningen och tydligt bedömer, också kommer att uppfattas som det viktigaste. Så stor transparens och så god kommunikation som möjligt kring de centrala frågorna som tidigare nämnts – varför, vad, hur osv. – torde därför vara det bästa sättet att undvika missförstånd, men det är likafullt viktigt att inse att elever med stor sannolikhet tror mera på det de ser göras än på det de hör sägas. Det räcker alltså inte bara att säga att till exempel muntligt kommunicerande eller val av adekvata lässtrategier är centralt, detta måste också utvärderas på ett tydligt sätt, jämförbart med andra aspekter av kunnandet inom ämnet.

Ytterligare ett sätt att beskriva olika avsikter med bedömning är att använda begreppen *formativ* respektive *summativ*, som kan sägas representera två olika aspekter av bedömningsuppdraget, vilket inte minst har att göra med hur gjorda iakttagelser används. Dessa begrepp har även kommit att dominera diskussionen om bedömning i skolan under senare tid, ofta under benämningen »Bedömning *för* lärande« (BFL) respektive »Bedömning *av* lärande«.

Formativ och summativ bedömning

Begreppen formativ och summativ bedömning myntades i slutet av 1960-talet (Scriven, 1967) och avsåg ursprungligen utvärdering i högre utbildning. De kom emellertid snabbt att vidareutvecklas och användas i undervisningsrelaterade sammanhang med fokus på stödjande av elevers lärande (Black & Wiliam, 1998; Sadler, 1989), men även med visst fokus på lärares professionella utveckling. Formativ bedömning avser den typ av bedömning som syftar till att påverka eller forma lärandet, men också undervisningen, medan summativ bedömning står för det sammanfattande eller summerande, vilket kan innefatta mycket lite eller väldigt mycket beroende på planering och sammanhang.

Formativ bedömning refereras, som tidigare nämnts, också till som »Bedömning för lärande« och definieras ofta utifrån tre frågor som härstammar från Sadler (1989) och som har vidareutvecklats av den så kallade Assessment Reform Group, som bedrev intensivt arbete kring bedömningsfrågor mellan 1996 och 2010 (www.nuffieldfoundation.org/assessment-reform-group). De tre frågorna är följande:

1. Var är eleverna i sitt lärande?
2. Vart är de på väg?
3. Hur ska de på bästa sätt ta sig dit?

Ovanstående frågor kan och bör emellertid också ställas i första person, det vill säga med eleverna som subjekt: »Var är jag nu i mitt lärande?«, »Vart är jag på väg?« och »Hur ska jag på bästa sätt ta mig dit?« – detta för att stimulera eleverna till reflektion både kring den egna kunskapsprofilen och kring det som ska läras, det vill säga målen för ämnet. En ytterligare god effekt är att svaren och kommentarerna kring frågorna ger läraren värdefull, inte sällan unik, information från den som lär sig, det vill säga den verkliga huvudpersonen i processen. Oavsett subjekt kan vi konstatera att formativ bedömning till mycket stor del handlar om kommuni-

kation mellan elev och lärare, men även mellan elev och elev, till ömsesidig nytta för att öka lärandet i sig och kunskapen om hur lärandet kan utvecklas till att bli alltmer positivt och effektivt.

Summativ bedömning avser avstämning av kunskapsläget och refereras, som sagt, ofta till som Bedömning av lärande. Det handlar här om att vid vissa bestämda tidpunkter ta reda på vad eleven har lärt sig. Denna form av bedömning kan tyckas vara ett tämligen självklart inslag i all undervisning. Det framstår därför som besynnerligt att summativ bedömning stundtals framställs som en stor och ödesmättad slutbedömning, där väldigt mycket står på spel, där alla åtgärder är för sena och kunskaperna dessutom tenderar att bedömas på ett ytligt och föga pedagogiskt sätt. Visserligen förekommer större, sammanfattande bedömningar i svensk skola, dock inte i form av examensprov och dessutom vid ytterst få tillfällen. I själva verket torde den allra mesta summativa bedömningen försiggå som en del av undervisningen och utgöra en typ av »mikro-summativ« iakttagelse, som används dels för att bidra till ändamålsenlig fortsatt undervisning och lärande, dels för att tillsammans med många andra likartade observationer bilda underlag för summerande omdömen. Ett exempel på en studie av denna typ av bedömning i svensk kontext och inom ramen för ämnet svenska är Skar (2013), som analyserar fallstudier av skrivbedömning i gymnasiet.

RELATIONEN MELLAN DET FORMATIVA OCH SUMMATIVA

Det finns många beskrivningar av relationen mellan formativ och summativ bedömning. Ett kortfattat sätt är att säga att den formativa bedömningen försiggår kontinuerligt med syfte att stödja lärandet, medan den summativa görs vid definierade tillfällen för att ta reda på vad och hur mycket eleven lärt sig. En annan, något mera spektakulär karakteristik, ges av Stake (citerad i Scriven, 1991), som konstaterar att »when the cook tastes the soup, that's formative; when the guests taste the soup, that's summative«. Det handlar enligt Stake följaktligen om att under processens gång

optimera produkten för att resultatet ska bli så gott som möjligt. Hur illustrativ och slagkraftig denna metafor än må vara, bidrar den emellertid till synen på summativ bedömning som något slutgiltigt och tämligen ödesmättat, vilket inte riktigt stämmer med verkligheten. Flera forskare, till exempel Taras (2005) och Harlen (2012), har också varnat för en dikotomisering av formativ och summativ bedömning samt påpekat vikten av att se dem som komplement till varandra. Det finns i detta sammanhang också anledning att påminna om de grundprinciper som gäller för all typ av bedömning, nämligen att den ska vara präglad av tydlighet, giltighet, tillförlitlighet och respekt.

Konstaterandet att olika bedömningsformer vilar på samma grundprinciper betyder inte att alla metoder passar för allt, eller att samma bedömning kan ha hur många syften som helst och generera resultat med samma giltighet för alla dessa. Tvärtom är det väsentligt att vid planeringen av olika bedömningar definiera vad som är huvudsyftet och låta detta styra såväl utformning som användning av olika typer av material och metoder. Vidare är det av grundläggande betydelse att inför sammanfattande beslut kring kunskapsnivåer hämta belägg från bedömningar av olika slag.

Allt är väl inte bedömning?

En följd av att kunskapsbedömning kommit att fokuseras allt mer, är att språkbruket kring företeelsen behöver diskuteras. Frågan i rubriken till detta avsnitt, ställd av en lärarstudent, kan illustrera detta. Den aktualiserar inte minst vikten av att skapa rimlig tydlighet kring *när* bedömning sker, för att undvika det som i »Gemensam europeisk referensram för språk: lärande, undervisning och bedömning« (Council of Europe/Skolverket 2009; s. 183) beskrivs som risken att »göra livet till ett enda oändligt prov för eleven och till en byråkratisk mardröm för läraren«. Skolverket har i »Allmänna råd för grundskolan« (2011b, s. 25) uttryckt den första risken på följande sätt:

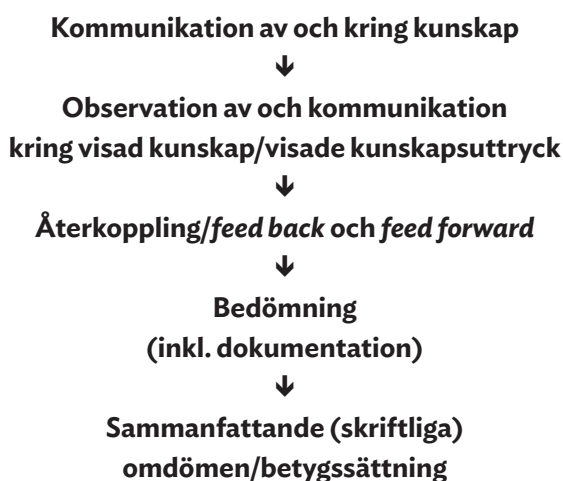
Det ska dock påpekas att det utifrån ett pedagogiskt perspektiv är viktigt att eleverna inte upplever att allt de gör bedöms hela tiden. Det är därför viktigt att eleverna är införstådda med när och hur läraren tänker följa upp deras kunskaper och vad som fokuseras vid olika tillfällen.

Detta kan sägas vara en relativt stor kontrast till det somliga lärare, i uppenbar välmening, säger till sina elever, nämligen att de inte behöver oroa sig för proven, eftersom läraren inte ger sådana utan »bedömer hela tiden«. Yttranden av det slaget kan möjligen lugna vissa elever, men säkerligen oroa många andra som kommer att undra om de i det ständigt granskande läge som i så fall råder bör hålla inne med undringar och frågor som kan visa på bristande insikt i olika skeden av lärandet. Det vore mycket olyckligt om ett ord och uttryckssätt från en lärare på detta sätt skulle ge signalen att elevernas självklara och nödvändiga, ibland korta, ibland långa, sökande efter förståelse skulle värderas och bedömas i sig. Det finns emellertid en risk med den relativa tydlighet rörande *när* bedömningar görs, som förordas av Skolverket, om deras budskap uppfattas som att enstaka prov räcker som underlag för lärares sammantagna bedömning. Enligt läroplanerna ska läraren utnyttja all tillgänglig information vid betygssättning och får alltså inte förlita sig på ett för smalt underlag. Som i så många andra pedagogiska och didaktiska sammanhang gäller det att med förstånd och i kommunikation med eleverna skapa en god balans mellan olika företeelser i verksamheten.

Frågan om dokumentation – det som i Gemensam europeisk referensram för språk (se ovan) refereras till som något som kan bli en »byråkratisk mardröm för läraren« – är i någon mån kopplad till diskussionen om bedömning och dess definition. Det förefaller därför väsentligt att diskutera vad olika beteckningar står för och kanske framför allt var gränsen går mellan det vi refererar till som undervisning respektive bedömning. En hel del av det som i dagens retorik refereras till som bedömning skulle faktiskt snarare kunna betecknas som god, lyhörd och tydlig undervisning, som har som

naturligt inslag kommunikation med elever enskilt och i grupp kring kunskap och kunskapsuttryck. Det finns mycket och många som talar för det och det är en diskussion som säkert kommer att växa i omfattning.

Nedanstående enkla modell (Figur 1.1), något vidareutvecklad från Erickson och Gustafsson (2014), över relationen mellan kommunikation och bedömning kan fungera som ett underlag för diskussion av den här frågan, som givetvis har betydelse långt bortom frågan om ordval:



Figur 1.1. Relationen mellan kommunikation och bedömning.

En grundläggande aspekt av lärarprofessionen utgörs av kommunikation av och kring kunskap. I detta ligger naturligt nog också observation av och kommunikation kring den kunskap eleverna visar – de kunskapsuttryck som framträder. Denna kommunikation innebär även successiv, mer eller mindre riktad och precis, återkoppling, som inte enbart kopplar tillbaka, det vill säga reagerar på och kommenterar det som har sagts och gjorts i dåtid eller försiggår i realtid (*feedback*), utan den bidrar även med framåtsyftande kommentarer, det som av Sadler (1989) rubriceras

som *feed forward*. I detta ingår även elevernas egna reflektioner kring lärandet, individuellt och tillsammans med andra.

I något skede övergår dessa resonerande aktiviteter mellan lärare och elever i något som innebär mer än kommunikation, nämligen en något mera systematiserad, i vissa delar standardiserad och av läraren dokumenterad, värdering i relation till uppsatta mål; kommunikationen och återkopplingen kan därmed sägas ha övergått i det som vanligen refereras till som bedömning. Slutligen kan iakttagelser från olika bedömningstillfällen summeras i muntliga och skriftliga omdömen eller i betygssättning, vilket följaktligen innebär en formell summering av iakttagelser, vilka också har syftet att ge framåtsyftande information.

Nationella prov

Den absolut största andelen kunskapsbedömningar görs lokalt, i klassrum runt om i landet – och i världen. Det rör sig om olika typer av bedömningar, med olika syften och därmed olika tillvägagångssätt. Det finns emellertid också bedömningar som kommer utifrån och som har delvis andra syften. Vi ska här mycket kort beröra en sådan företeelse, nämligen de svenska nationella proven.

Sverige har en lång tradition vad gäller centralt tillhandahållna prov med syfte att stödja lärares bedömning och betygssättning inom det rådande systemet (Marklund, 1987). De konkreta syftena med de nationella proven har varierat mellan det breda och det smala. Efter införandet av det målrelaterade styrsystemet på 1990-talet skulle proven under ett antal år, förutom att ge stöd för rättvisa, likvärdighet och uppföljning, också bidra till konkretisering och implementering av kursplanernas kunskapssyn samt till enskilda elevers måluppfyllelse. Detta visade sig dock resultera i alltför stor variabilitet i utformning, tolkning och användning, varför syftena i dag har begränsats till att vara mera uttalat summativa. De nationella proven ska:

- Stödja en likvärdig och rättvis bedömning och betygssättning, och
- Ge underlag för en analys av i vilken utsträckning kunskapskraven uppfylls på skolnivå, på huvudmannanivå och på nationell nivå.

Den formativa ambitionen lever emellertid delvis kvar i Skolverkets konstaterande av att de nationella proven också kan bidra till:

- Att konkretisera kursplanerna och ämnesplanerna
- Att ge en ökad måluppfyllelse för eleverna (se vidare www.skolverket.se/bedomning/nationella-prov)

Det primära syftet med de nationella proven är följaktligen att bidra till rättvisa och likvärdighet. Detta innebär bland annat att frågan om samstämmigheten mellan bedömare måste ägnas kraft och uppmärksamhet, något som understryks av resultat från studier genomförda av de nationella skolmyndigheterna (Skolinspektionen, 2013; Skolverket, 2009).

Det finns här en hel del kritik, inte minst mot Skolinspektionens undersökningar (Gustafsson & Erickson, 2013), men det innebär inte att saken som sådan kan ignoreras. Frågan är emellertid komplex och innefattar en mängd yttre ramfaktorer och traditioner liksom aspekter av kunskapsdefinitioner och provutformning, det senare diskuterat i en nyligen framlagd doktorsavhandling med fokus på de nationella proven i svenska (Borgström, 2014).

Bedömning för lärande och likvärdighet

Synen på skolans kunskapsuppdrag och bedömningens roll i detta kan diskuteras och definieras på olika sätt. Generellt kan dock sägas att bedömning har två distinkta icke utbytbara funktioner att fylla; dels att stödja lärande och (kunskaps)utveckling, dels att säkra kvalitet och därmed befrämja rättvisa och likvärdighet på såväl

individuell som strukturell nivå. Behovet av det senare accentueras ytterligare av svenska lärares tämligen unika ansvar för betygssättning, med allt vad det innebär av påverkan på elevers självuppfattning, utbildningsval och livschanser. Det är i detta sammanhang viktigt att återigen betona de gemensamma grundläggande principer som bör finnas för alla typer av bedömningar, oavsett syfte och sammanhang, nämligen de kvaliteter som inte minst innebär krav på tydlighet, giltighet, tillförlitlighet och respekt. Detta har i grunden att göra med den etik som vi gemensamt måste kämpa för ska prägla hela samhället, och därmed i allra högsta grad även skolan och utbildningen.

Referenser

- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bernstein, B. (1996). *Pedagogy, symbolic control and identity*. London: Taylor & Francis.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1).
- Borgström, E. (2014). *Skrivbedömning: om uppgifter, texter och bedömningsanvisningar i svenskämnets nationella prov*. Örebro: Örebro universitet: Institutionen för humaniora, utbildnings- och samhällsvetenskap.
- Council of Europe/Skolverket (2009). *Gemensam europeisk referensram för språk: lärande, undervisning, bedömning*. Stockholm: Skolverket.
- Erickson, G. (2009). ”Att bäras åt”: om den goda bedömningens flerdighet och ömsesidighet. I U. Tornberg (red.), *Språkdidaktiska perspektiv*. Stockholm: Liber.
- Erickson, G. & Gustafsson, J.-E. (2014). Bedömningens dubbla funktion – för lärande och likvärdighet. I U. P. Lundgren, R. Säljö & C. Liberg (red.), *Lärande, skola, bildning* (3 uppl.) (s. 559–594). Stockholm: Natur & Kultur.
- Foucault, M. (1977). *Discipline and punishment* (A. Sheridan, övers.). London: Allen Lane.
- Gardner, J. (red.). (2012). *Assessment and learning* (2 uppl.). London: Sage Publications.
- Gustafsson, J.-E. & Erickson, G. (2013). To trust or not to trust? Contrasting findings on teachers’ assessments. *Educational Assessment, Evaluation and Accountability*, 25(1), 69–87.
- Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. I J. Gardner (red.), *Assessment and learning* (2 uppl.) (s. 87–102). London: Sage Publications.
- Kane, M. (2006). Validation. I R.L. Brennan (red.), *Educational Measurement* (4 uppl.) (s. 17–64). Westport, CT: American Council on Education/Praeger Publishers.

- Lindblad, T. (1985). On the whys and hows of language testing. I V. Kohonen & A.J. Pitkänen (red.), *Language testing in school: AFinLA yearbook 1985*. Tammerfors: Tammerfors universitet.
- Marklund, S. (1987). *Skolsverige. Del 5: Läroplaner*. Stockholm: Liber.
- Messick, S.A. (1989). Validity. I R.L. Linn (red.), *Educational Measurement* (3 uppl.) (s. 13–103). New York: American Council on Education/Macmillan.
- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science* 18, 119–144.
- Scriven, M. (1967). The methodology of evaluation. I R.W. Tyler, R.M. Gagne & M. Scriven (red.), *Perspectives of curriculum evaluation*. Chicago, IL: Rand McNally.
- Scriven, M. (1991). Beyond formative and summative evaluation. I M.W. McLaughlin & E.D.C. Phillips (red.), *Evaluation and education: A quarter century*. Chicago: University of Chicago Press.
- Skar, G. (2013). *Skrivbedömning och validitet – Fallstudier av skrivbedömning i svenskundervisning på gymnasiet*. Stockholms universitet: Institutionen för språkdidaktik.
- Skolinspektionen (2013). *Olikheterna är för stora om rättning av nationella prov i grundskolan och gymnasieskolan, 2013*. Stockholm: Skolinspektionen.
- Skolverket (2009). *Redovisning av uppdrag om skillnaden mellan betygsresultat på nationella prov och ämnesbetyg i svenska, matematik och engelska i årskurs 9*. Stockholm: Skolverket (Dnr 2008:3789).
- Skolverket (2011a). *Kunskapsbedömning i skolan – praxis, begrepp, problem och möjligheter*. Stockholm: Skolverket.
- Skolverket (2011b). *Allmänna råd för planering och genomförande av undervisningen*. Stockholm: Skolverket.
- Taras, M. (2005). Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466–478.

2. Att kvalitetssäkra bedömning

GUSTAF SKAR

All bedömning får konsekvenser. Ju mer omfattande dessa konsekvenser är, desto säkrare vill vi vara på att bedömningen håller god kvalitet. Inom bedömningsverksamhet brukar man ofta tala om kvalitativ bedömning som valid och reliabel. I det här kapitlet ska vi se närmare på begreppen *validitet* och *reliabilitet*.

Kapitlet har emellertid ytterligare ett syfte, nämligen att fungera som en introduktion till en av grundstenarna i denna bok – Michael T. Kanes validitetskapitel, som redovisar Kanes validitetsmodell. Kane är en av vår tids mest inflytelserika validitetsteoretiker. Alltsedan en banbrytande artikel i början av 1990-talet (Kane, 1992)¹ har han influerat vårt sätt att förstå och undersöka validitet, eller *giltighet* som det också ibland benämns. Vi kommer i den här introduktionen att skissera konturerna på Kanes validitetsmodell, som har sitt ursprung i mätteoretiska (psykometriska) sammanhang.

Kapitlet är upplagt på följande sätt. Inledningsvis kommer jag att presentera ett fiktivt fall av bedömning. Fallet kommer att användas som exempel i texten igenom. Därefter kommer jag att redogöra för den så kallade argumentbaserade ansatsen till validitet. Kane själv (se kapitel 10) presenterar en både utmärkt

1 Se också t.ex. Kane (2006, 2011, 2013) och två flitigt citerade artiklar som författats med andra (Crooks, Kane, & Cohen, 1996; Kane, Crooks, & Cohen, 1999).

och komprimerad version av validitetsbegreppets historia. Denna begreppsdefinition bildar sedan utgångspunkt för ett resonemang om hur det fiktiva fallet kan förstås med utgångspunkt i begreppen validitet och reliabilitet. Allra först ska vi dock se närmre på vad bedömning är.

Något om språkbedömning

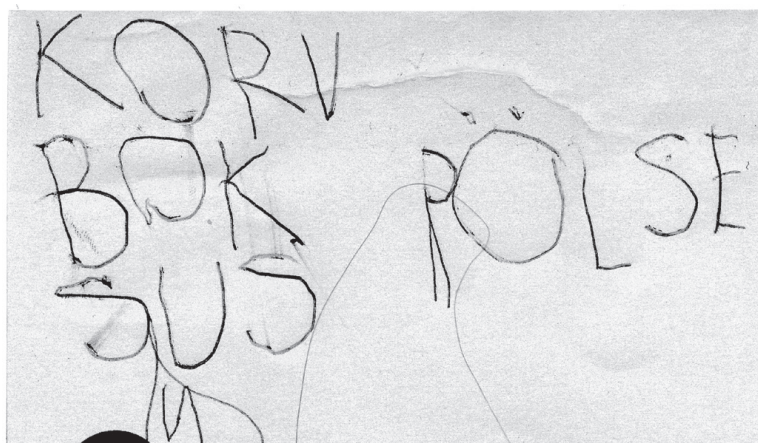
Innan vi tar oss an vårt fiktiva bedömningsfall kan det vara värt att närmare fundera över vad det är vi gör när vi tänker oss bedöma någons skriv- och/eller läsförmågor. I typfallet kommer vi att med utgångspunkt i ett eller annat empiriskt underlag – exempelvis en text eller ett muntligt svar – försöka dra slutsatser om sådant vi *inte* ser eller hör.

I Figur 2.1 här intill återges en text som är författad av en 5-årig flicka. Texten beskriver olika varor som ska inhandlas. Fastän vi bara har fyra ord att tillgå – *korv*, *bok*, *snus*, *pölse* – vore det inte konstigt att dra slutsatser om vad flickan kan och inte kan göra med skriften som verktyg.

Vi kan se att flickan, författaren av lappen, har förstått något om texters materiella aspekt, eftersom hon avpassat skriftens storlek till arkets yta. Dessutom är många ord rättstavade, vilket tyder på en begynnande medvetenhet om fonem-grafem-relationer. Vi kan emellertid också, till exempel, se att bokstäverna verkar vara producerade med viss möda, vilket indikerar att flickan ännu inte behärskar att koda med finmotorik. Vidare att ett ord (*snus*) initialt är felstavat och därefter korrigerat, något som antyder vissa utmaningar hos henne avseende stavning av ett egentligen ljudrätt ord. Många av oss skulle säkert utan större tveksamhet mena att författaren av denna inköpslista antagligen inte kan skriva en längre diskursiv text.

Poängen med dessa exempel på vilka slutsatser vi drar, är att bedömning är en komplex akt. Får vi i uppgift att bedöma om flickan kan stava till *korv*, *bok*, *snus* och *pölse* svarar vi på ett sätt,

medan uppgiften att bedöma om flickan kan stava över huvud taget skulle kunna ge ett annat svar. Vi skulle då tala om *förmåga att stava*, och förmågor är något som inte syns, varken på personer eller på enstaka papperslappar. När vi vill använda en text som den i figuren, eller en annan text, för att göra mer än att beskriva just den enskilda texten, drar vi slutsatser om sådant vi bara kan anta.



Figur 2.1. En inköpslista författad av en 5-årig flicka.

Att dra slutsatser av detta slag är ett givet inslag i en lärarvardag. Som lärare är vi ofta mindre intresserade av själva det empiriska underlaget (korv, bok, snus, pölse), än vad det faktiskt kan berätta om en elevs förmåga att utföra liknande och andra handlingar vid andra tillfällen. Att konstatera att en elev kan skriva fyra specifika ord hjälper varken till i undervisningsplaneringen eller för att dra slutsatser i förhållande till abstrakta målformuleringar. Av den anledningen är de bedömningar vi gör sällan bara beskrivningar eller bedömningar av kvaliteten i en produkt, utan av kvaliteten tolkad som ett uttryck för någons förmåga. Med andra ord gör vi inferenser från det synliga och materiella (t.ex. texten eller svaret) till det osynliga och immateriella.

Validitet

ETT FIKTIVT FALL

Låt säga att vi vill ta reda på hur väl eleverna i klass 9A kan skriva utredande uppsats. Vi börjar med att konkludera vilket slags bedömning vi vill göra och varför, och kommer fram till att vi vill kunna dra slutsatser om elevernas förmåga i mer generella ordalag. Vi vill, när bedömning är gjord, kunna påstå något om givna elevers förmåga att skriva utredande uppsatser även vid andra tillfällen. Syftet med bedömningen är att planera den kommande undervisningen, eftersom eleverna kommer att placeras i olika undervisningsspår beroende på resultatet på uppsatsprovet. Ett undervisningsspår innebär i det här fallet ett tidsbegränsat anpassat upplägg.

VAD ÄR VALIDITET?

För många kan begrepp som validitet och reliabilitet uppfattas som tekniska. Faktum är dock att information, det vill säga slutsatser, från bedömningar bara är intressant om vi kan lita på den, det vill säga om den är valid. När vi i slutet av en kurs ska samla våra slutsatser om elevernas kunskaper och förmågor vill både vi, eleverna, deras föräldrar och andra intressenter, att både slutsatser och beslut ska vara valida.

En någorlunda vedertagen definition av validitet är följande: »Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests« (American Educational Research Association, American Psychological Association & National Council on Educational Measurement, 2014, s. 11). I klartext betyder det att validitet handlar om i vilken utsträckning det finns bevis som stödjer rimligheten i att tolka ett provresultat på ett visst sätt, givet ett visst syfte (t.ex. betygsättning, undervisningsplanering etc.).

Många av oss har emellertid en gång fått lära att »validitet föreligger om instrumentet mäter det som instrument ska mäta« (jfr t.ex. Svenska Akademiens ordbok). Med en sådan definition är

validitet nästan en inneboende egenskap i mätinstrumentet, vilket skulle innebära att exempelvis ett läsförståelsetest skulle vara valitt för att det mäter läsförståelse. Inom psykometrin har man dock länge fört fram att själva mätinstrumentet i sig inte är föremål för validering.² Om syftet är att undersöka elevers matematikförmåga kommer ett »giltigt« läsförståelsetest inte att kunna göra jobbet. Därför är validitet en abstraktion och något vi kan påstå »finns«, »föreligger« eller »är etablerat« när syftet med en mätning harmoniserar med vårt instrument.

Andra kapitel i den här boken (se kapitel 1, samt kapitel 5) diskuterar ingående frågan om relationen mellan bedömning och syfte, och hur den avgör vilka bedömningsverktyg som passar. Med en validitetsmodell får vi ett instrument för att strukturera undersökningar om hur syften och instrument harmoniserar. Inbakat i en validitetsmodell är också vetenskapsteoretiska uppfattningar om vad som krävs för att vi ska kunna påstå saker, till exempel att någon kan skriva eller läsa eller att någon kan mycket om andra världskriget.

Begreppet validitet har varit och fortsätter att vara ett begrepp »i rörelse«. För det första finns det en rörelse i tid. Den definition av begreppet som ovan presenterades som vedertagen är frukten av intensiva meningsutbyten över lång tid om hur man fångar fenomenet validitet; olika vetenskapsteoretiska paradig har inverkat på konceptualiseringar av begreppet. För det andra är det egentligen felaktigt att tala om vedertagna uppfattningar av vad validitet är. I en nyligen publicerad genomlysning av begreppets historia och utveckling pekar Newton och Shaw (2014) på att litteraturen på området föreslår åtminstone 150+ »typer« av validitet. Michael T. Kanes modell får förvisso anses vara allmänt accepterad, och

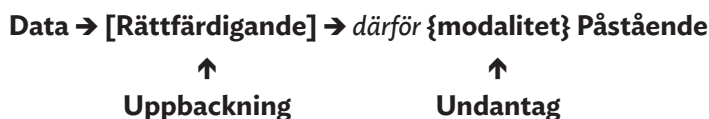
2 Det kan i detta sammanhang vara intressant att peka på en formulering i det verk som i det närmaste fungerar som en »officiell« definition av validitet, *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 2014). I denna står följande: »It is incorrect to use the unqualified phrase 'the validity of the test'.« (s. 11).

dessutom något som man i allmänhet förhåller sig till, men läsaren bör känna till att den alltså inte är allenarådande.

KANES MODELL – EN ARGUMENTBASERAD ANSATS³

Kanes modell för validering bygger på en så kallad argumentbaserad ansats. En sådan ansats tar sikte på validitet som abstraktion och utgår i Kanes fall från en specifik modell för argumentationsanalys (Toulmin, 2003).

Toulmins modell (se Figur 2.2) innebär en struktur för att skärskåda belägg och rättfärdiganden som stödjer ett påstående. Grunden i modellen är enkel. Låt säga att vi får syn på någon med ett bandage runt huvudet. Vi konstaterar då att den här personen har skadat sitt huvud. Egentligen vet vi inte det, men vi gör en inferens från ett belägg (bandagerat huvud) till ett påstående om huvudskada (jfr också inferensen från inköpslistan ovan, t.ex. stavningsförmåga). Inferensen möjliggörs av ett rättfärdigande, nämligen vår omvärldskunskap som säger oss att personer som skadar huvudet kan få bandage. (Skulle vi se det bandagerade huvudet på en Halloweenfest är det dock inte säkert att vi skulle ha gjort samma inferens.)



Figur 2.2. Toulmins modell för argumentationsanalys; från data till ett påstående, stött av ett rättfärdigande (som stöds av uppbackning). Modalitet avser de reservationer som är nödvändiga.

Kanes validitetsmodell bygger på en argumentbaserad ansats och utgår från att validitet inte är en fast egenskap hos till exempel ett prov. I stället är det rimligare att tänka sig att man kan påstå att en

³ Delar av de resonemang som förs i avsnittet förekommer också i Skar (2014).

slutsats om en elevs kunskap, eller en åtgärd baserad på ett provresultat (t.ex. att placera någon i en särskild studiegrupp), är mer eller mindre valid. När vi kommer med detta påstående (vårt argument för validitet) bygger det på att vi klagjort en lång rad inferenser i vad Kane kallar ett tolknings- och användningsargument.

Ett tolknings- och användningsargument specificerar de bedömningsbaserade påståenden vi gör och de handlingar vi utför, liksom på vilka grunder vi påstår saker eller handlar som vi gör. Med hjälp av ett explicitgjort tolknings- och användningsargument redovisar vi alltså hur påstående, handling och bakomliggande skäl hänger samman. Låt oss ta ett exempel:

Låt säga att vi har bestämt oss för att diagnostisera en klass explicita grammatikkunskaper i syfte att anpassa undervisningen efter klassens behov. Vi kommer för det första behöva undersöka den allmänna grammatikkunskapen genom något slags bedömning, och därför kommer vi att utföra en tolkning av ett eller annat provresultat. För det andra kommer vi att vilja kunna fatta beslut efter genomförd diagnostisering, det vill säga vi kommer att *använda* provresultaten för att fatta ett eller annat beslut. I ett *tolknings- och användningsargument*, det vill säga i ett »argument« för att det vi gör är rimligt, behöver vi därför åtminstone specificera:

1. Hur säkra vi kan vara på att våra »bedömningsregler« (t.ex. bedömningsmatriser) gör att vi bedömer det vi vill bedöma. De många olika möjliga bedömningarna av inköpslistan i Figur 2.1 visar betydelsen av att specificera vad vi ska leta efter.
2. Hur säkra vi kan vara på att en enskild observation av grammatikkunskaper skulle stå sig om vi gjorde om provet. Det vill säga, är resultaten konsistenta (reliabla) eller kommer de att variera på grund av tillfälligheter?
3. Hur säkra vi kan vara på att just det här provet lånar sig till slutsatser om annat än hur eleverna presterar på provet. Medan punkt 2 handlar om möjligheten att generalisera

över provtillfällena handlar denna punkt om att dra en mer långtgående slutsats. Kan vi, med andra ord, vara förvissade om att provet säger något om den grammatikkunskap vi vill dra slutsatser om?

4. Hur säkra vi kan vara på att våra beslutsregler är i linje med våra bedömningssyften (som ju är att möjliggöra en god skrivundervisning). Är vi säkra på att vi planerar undervisningen på rätt sätt med utgångspunkt i resultaten från det diagnostiska provet? Kan vi vara övertygade om att konsekvenserna av vårt agerande är i enlighet med vad vi önskat oss?

Om vi på ovanstående frågor kan svara att vi är mycket säkra, ja då har vi just »argumenterat« för bedömningsprocessens validitet.

DEN ARGUMENTBASERADE ANSATSEN I PRAKTIKEN

Låt oss nu återgå till klass 9A för att se hur en argumentbaserad ansats till validitet skulle kunna ta sig uttryck i ett klassrumsbedömnings-sammanhang.

För att bedöma kvaliteten i vår bedömning av elevers förmåga att skriva utredande uppsats ställer vi upp ett tolknings- och användningsargument. Vi specificerar vad ett provresultat är tänkt att betyda, nämligen en indikation på elevers förmåga att skriva utredande uppsats, samt vilka bevis som måste föreligga för att vi ska kunna »ta steget« från vårt empiriska underlag till vårt påstående. Vi specificerar också hur vi kommer att använda provresultatet och hur ett sådant användande kan rättfärdigas.

Vi kan tänka på provresultatet som ett påstående: »Jag påstår att eleven är duktig på att skriva utredande uppsats.« Påståendet kan ta sig uttryck som ett uppgiftsbetyg (t.ex. 10 poäng; B; ***** 😊) eller en kommentar, till exempel: »Du skriver bra och det framgår tydligt vad du vill utreda.« Både betyget och kommentaren antyder dels ett »provresultat«, dels en innebörd; resultatet är högt vilket

med stor sannolikhet betyder att skribenten skrev bra på denna uppgift, eller, kanske, att skribenten kan skriva bra på denna typ av uppgifter i allmänhet.

För att kunna göra ett dylikt påstående måste vi åtminstone på goda grunder kunna hävda att våra bedömningsregler är av god kvalitet, att resultaten är generaliserbara i ordets vidaste bemärkelse och att användningsregeln gynnar bedömningens syfte.

Först måste vi dock ha klart för oss vad utredande uppsats är; vi måste definiera den skrivhandling vi tänker oss bedöma. Utan att veta vad utredande uppsatsskrivande går ut på får vi svårt att hävda att någon kan eller inte kan skriva utredande uppsats. Även om våra kursplaner bygger på teoretiska utgångspunkter om vad det innebär att använda språket för att kommunicera, erbjuder de inga explicita teoretiska beskrivningar av till exempel skrivande, vilket gör att sådana får sökas från annat håll (jfr kapitel 3).

I bedömningsteori talar man om att mäta konstrukt (jfr kapitel 1), det vill säga osynliga mer eller mindre fixa egenskaper (t.ex. aggressivitet, ångest) eller förmågor (t.ex. skrivförmåga, läsförmåga). Skrivförmåga är just en osynlig förmåga och det är vår förståelse av den som ligger till grund för våra konklusioner om kvaliteten hos bedömningsreglerna, om möjligheten till generalisering och så vidare.

BEDÖMNINGSREGLER

Vi måste ha bedömningsregler som kan hjälpa oss att ta oss från en elevtext som »data« till ett resultat som »påstående« (se Figur 2.2). Vår konstruktdefinition, det vill säga den definition av skrivförmåga vi utgår från, ger oss då uppbackning för vårt rättfärdigande.

Bedömningsregler när det gäller skrivbedömning innehåller oftast anvisningar om vad vi ska leta efter i texten, till exempel retorisk slagkraftighet, stavning eller styckesmarkeringar. På den högsta nivån kanske elevtexten innehåller en analys som styrks av källor, medan texter på den lägre nivån kan innehålla mindre

underbyggda analyser. Vanligen ordnas beskrivningar av detta slag i matriser, som kan ställas upp hur som helst (så länge uppställningen är användbar). Beskrivningarna brukar motsvara en bokstav (t.ex. D) som i sin tur antas motsvara en viss betygsnivå. Ibland vill man dock kanske bedöma varje aspekt för sig (vilket ger en bokstav för respektive aspekt), något som kallas *analytisk bedömning*. I andra sammanhang ger man kanske i stället ett helhetsomdöme (bokstaven representerar då texten som helhetsprodukt), vilket kallas *holistisk bedömning*. Valet mellan analytisk eller holistisk bedömning har oftast med bedömningens syfte att göra. Vill vi diagnostisera elever är det fördelaktigt med analytisk bedömning eftersom vi då kan få nyanserade (eller »taggiga«) profiler.

GENERALISERBARHET – RELIABILITET

Vi måste också vara övertygade om att det resultat vi får fram efter bedömningen inte i någon överdriven grad beror på slumpen. För att försäkra sig om att så inte är fallet brukar man undersöka en bedömnings reliabilitet.⁴

I klassrumssammanhang finns det åtminstone två typer av reliabilitet som är viktiga att ta hänsyn till:

1. Bedömarreliabilitet
2. Uppgiftsreliabilitet

Med bedömarreliabilitet menas att olika bedömare bedömer en och samma prestation på samma vis, medan uppgiftsreliabilitet innebär att olika uppgifter, som antas mäta samma förmåga, verkligen leder till likartade prestationer från eleven. Oavsett om bedömningen ska användas som underlag för kommande undervisning eller som underlag för kursbetyg måste värderingen, som sker med hjälp av bedömningskriterierna, vara reliabel, det vill

4 I Kanes modell är sålunda reliabiliteten en av förutsättningarna för validitet.

säga så fri från slumpmässighet som möjligt. Om så inte är fallet är det svårt att exempelvis veta om eleverna verkligen behöver träna mer på stavning eller om det är slumpmässigheten i bedömningen som fått oss att tro det. Förutom att två olika bedömare av samma elevtext bör göra likartade bedömningar är det också viktigt att den enskilda läraren vet att hon eller han gör samma bedömning av samma elevtext oavsett tillfälle, vilket kräver välformulerade och konkretiserade bedömningskriterier och en hel del träning.

Ett enkelt sätt att skatta bedömarreliabiliteten är att antingen vara flera kollegor som *oberoende* av varandra läser samma texter, eller att läraren själv läser samma texter med en tids mellanrum. Genom att sedan sätta upp olika lärares bedömningar i en tabell kan vi få mycket information om hur överens vi är och hur relativt stränga eller generösa vi är. I Tabell 2.1 nedan återges fiktiv data som illustrerar hur en sådan tabell kan se ut.

		Gustaf					Totalt
		E	D	C	B	A	
Michael	E	10	2				12
	D	2	5		5		12
	C			7	2		9
	B				13	4	17
	A	1				1	2
	Total	13	7	7	20	5	52

Tabell 2.1. Illustration av samstämmighetsundersökning.

Vi kan utläsa en rad intressanta saker från tabellen. För det första ser vi att det verkar finnas en tämligen stor procentuell fullständig överensstämmelse. Alla de skuggade fälten är sådana där bedömar-

na är överens. Totalt rör det sig om 36 av 52 fall (dvs. i 69 procent av fallen). Räknar vi i tillägg sådana omdömen som bara är ett steg från varandra är siffran 88 procent (46/52). Dock ser vi också att det finns fall då Michael har bedömt en text motsvara A, medan Gustaf har bedömt den motsvara E, vilket är ett problem. Vidare har Gustaf i fem av fallen satt B, då Michael satt D, även detta är ett problem. Ett grovt förenklat sätt att uttrycka detta problem på är att det i en viss andel av fallen är slumpen som avgör vilket betyg eleverna får, eller att betyget i en viss del av fallen beror på läsaren (läraren) och inte på textens kvalitet.

Vi kan också utläsa något annat ur tabellen, nämligen hur stränga och generösa Gustaf respektive Michael är. Låt oss se närmare på Tabell 2.2 som illustrerar detta.

	E	D	C	B	A	Totalt
Gustaf	13	7	7	20	5	52
Michael	12	12	9	17	2	52
Totalt	25	19	16	37	7	104

Tabell 2.2. Bedömarstränghet.

Redan vid en snabb okulärbesiktning ser vi att det finns en viss skillnad i stränghet mellan de två bedömarna. Medan 48 procent av Gustafs bedömningar motsvarar B och A (20 + 5), har Michael använt de två översta betygsstegen i 36 procent av fallen (17 + 2).

En reliabilitetsundersökning som denna ger inget svar på frågan om *varför* det till viss del råder oenighet mellan Gustafs och Michaels bedömningar, även om stränghetsundersökningen ger några indikationer på detta. Det viktiga i sammanhanget är att resultaten inverkar på vårt slutgiltiga mål, det vill säga att påstå något

om elevers förmåga att skriva utredande uppsats, för att sedan göra bruk av detta påstående. Vi kan inte vara 100 procent övertygade om att slutsatser om elevers skrivförmåga bara beror på skrivförmågan, eller på vem som har bedömt uppsatsen. Emellertid kan vi, om vi vill, säga att en 70-procentig fullständig samstämmighet är tillräckligt god för att generaliseringsmöjligheterna inte ska vara alltför försvagade.

Uppgiftsreliabilitet berör flera aspekter av själva uppgiftslösningen. Den första aspekten handlar om sådant som naturligt varierar, som humör, motivation, hunger, trötthet och liknande. När vi försätts i en testsituation kommer våra chanser att lösa en uppgift på ett tillfredsställande sätt delvis bero på vår dagsform. Allt annat lika betyder det att om eleverna skriver samma prov många gånger så kommer sådana faktorer att »ta ut varandra«. Medelvärden av resultaten kommer då med större säkerhet att avspegla elevernas faktiska förmåganivå, än vad som varit fallet om resultat bara hade räknats från ett testtillfälle.

Den andra – för klassrumssammanhang viktiga – aspekten berör det som ofta kallas standardisering. Formerna för uppgiftslösning kan variera stort. Elever kan ibland vara fria att lösa uppgifter hemma, medan de vid andra tillfällen måste lösa dem under övervakning i skolans datorsal. Ibland kan också ett och samma testtillfälle erbjuda variation – någon elev får extra handledning, en annan får det inte. En mindre skillnad, men likafullt viktig, kan vara att vissa elever har snabba datorer, medan andra får arbeta med datorer som ständigt måste startas om. Ju fler saker som skiljer förutsättningarna för testdeltagande åt, desto lägre är graden av standardisering. Om man mäter samma sak flera gånger och de enskilda elevernas resultat varierar kraftigt kan låg grad av standardisering vara orsaken till det.

Varför är då reliabilitet viktigt? Det korta svaret är att vi behöver kunna lita på att provresultat inte är ett uttryck för tillfälligheter. Om vi tillåter stora mängder osystematiskt »brus« i våra provresultat omöjliggör vi generalisering, och utan generalisering – att

samma elev skulle kunna producera en liknande text av liknande kvalitet vid ett annat tillfälle – sitter vi åter med data som inte medger sådana slutsatser som vi är intresserade av. För att kunna dra slutsatser om exempelvis elevers stavningsförmåga (jfr Figur 2.1) eller förmåga att skriva utredande uppsats, måste vi kunna anta att de skulle kunna vara i stånd att skriva på ett liknande sätt i ytterligare en provsituation. Låg reliabilitet hindrar dock sådana antaganden.

EXTRAPOLERING

Som det nyss har redogjorts för vill vi ofta något mer än att enbart ta reda på hur en elev kan svara på en given uppgift. Vi vill även *extrapolera* resultat från ett sammanhang (provsammanhanget) till ett annat sammanhang (sammanhang utanför provsituationen). Elevers förmåga att skriva utredande uppsatser om ett givet ämne under mer eller mindre givna förhållanden är av mindre intresse för oss än deras förmåga att skriva utredande uppsatser i stort.

Att extrapolera ett resultat betyder att dra slutsatser om sådant vi inte kan se. Det betyder att vi låter ett provresultat indikera hur väl en elev skulle klara av något utanför ett provsammanhang. En extrapolering är möjlig om man: 1) har definierat den förmåga man vill mäta, 2) har en föreställning om hur den förmågan tar sig uttryck i »världen utanför provet« och 3) kan belägga att skillnaden mellan sättet att utföra en uppgift på ett prov och i världen utanför provet, inte är alltför omfattande.

Provuppgifter speglar på ett idealt sätt vår konstruktdefinition, men ofta tvingas vi mäta på ett sådant sätt att uppgifterna bara till viss del speglar situationen utanför provet. Ibland kan iveren att få »objektiva« och jämförbara provresultat leda till att man mäter något mycket smalare än vad man hade tänkt uttala sig om. Det är till exempel inte ovanligt att skrivförmåga mäts genom enklare grammatikprov eller lucktest. Vinsten är förstås att man eliminerar den osäkerhet det innebär att ha mänskliga bedömare, medan för-

lusten är att vi får betydande svårigheter när vi ska hävda något om relationen mellan resultatet på lucktestet och elevernas förmåga att skriva diskursiva texter.⁵

Även standardiseringsbehovet kan innebära svårigheter att extrapolera. Säg att vi vill undanröja alla tvivel om plagiat eller annat fusk och låter våra elever skriva sin utredande uppsats utan tillgång till internet och utan möjlighet att låta andra läsa och kommentera texten. Då har vi effektivt reducerat möjligheten att påstå något om hur eleven kommer att kunna klara sig i den verkliga världen (där ju sådana hjälpmedel finns tillgängliga). Ett sådant pris kan dock vara väl värt att betala, beroende på vad som står på spel.

BESLUT

När vi har ett påstående, eller en fastlagd tolkning av innebörden av ett provresultat, vill vi oftast fatta någon typ av beslut. Låt oss fortsätta med exemplet från klass 9A.

Syftet med provet är att möjliggöra en placering av elever i olika undervisningsspår. Provet är tänkt att få goda konsekvenser i så måtto att alla elever, oaktat förkunskaper, ska kunna få en anpassad och adekvat undervisning. Ansvariga lärare på skolan har bestämt att elever som får 25 poäng på provet ska hamna i spår A, medan elever med mindre än 25 poäng ska hamna i studiegrupp B.

För att beslutet ska vara valitt krävs det att provet mäter elevernas förmåga att skriva utredande uppsats. Vid sidan av detta måste också beslutsregeln vara gynnsam visavi det syfte som specificerats. Lärarna måste på goda grunder veta att just 25 poäng är en rimlig kravgräns. Skulle det visa sig att så inte är fallet kan elever felaktigt klassificeras som lämpliga för studiegrupp A eller studiegrupp B. Vidare måste lärarna vara övertygade om att beslutsregeln resulterar

5 Att mäta »för lite« leder till underrepresentation av konstruktet (*construct under-representation*) och att mäta »för mycket« (t.ex. också mäta läsförmåga när man bara vill mäta skrivförmåga) leder till konstruktirrelevant varians (*construct-irrelevant variance*) (Messick, 1996, se även kapitel 1).

i just de goda konsekvenser man tänkt sig (avpassad undervisning till alla). Men även om vi – för exemplet skull – antar att det blir så (att var och en får delta i den undervisning hon eller han bäst behöver), kan vi fortfarande inte hävda att beslutet är rimligt. Vi måste också kontrollera för eventuella negativa konsekvenser.

Det går knappast att uttömmande presentera alla potentiellt negativa konsekvenser som kan uppkomma, men vi skulle här – åter för exemplet skull – kunna tänka oss att det blir stigmatiserande att tillhöra studiegrupp B, alltså den grupp som de elever hamnar i som får färre än 25 poäng. Lärarna upptäcker kanske då en bit in på terminen att eleverna i den gruppen blir tråkade av de andra, vilket kan resultera i att flera av dem, även om de presterar väl i undervisningen, aldrig mer frivilligt vill ägna sig åt skrivande och skrivundervisning. Lärarna måste då väga in den skada beslutet innebär, det vill säga de negativa konsekvenserna, mot dess nytta, det vill säga de positiva konsekvenserna. Om lärarna då menar att stigmatiseringen av eleverna i studiegrupp B är ett allvarligt problem som bör undvikas till varje pris, ja då ska beslutsregeln ändras (jfr Kane, 2013).

En annan potentiellt negativ konsekvens har med värderingar att göra. Som lärare bör man fundera på relationen mellan en konstruktetikett (t.ex. skrivförmåga) och uppgifters innehåll i termer av vems kunskap som är giltig. I Parmenius Swärd (2008) finns exempel på elever som utanför skolan både skriver mycket och får uppskattning för sitt skrivande, men som i skolan inte bedöms vara goda skribenter och därigenom tappar självförtroende, intresse och motivation. Att det då troligen rör sig om olika konstrukt kan ha undgått både elever och lärare.

När vi ska sätta slutbetyg eller uppgiftsbetyg, som används som underlag för slutbetyg, måste vi dessutom veta att eleverna haft möjlighet att lära sig just det skrivande som vi bedömer. Annars avspeglar betyget vad eleverna kan *utan* undervisning, och inte *efter* undervisning.

För att komplicera det hela måste avslutningsvis följande

kommenteras. Ett betyg i svenska under vårterminen i årskurs 9 är på samma gång ett påstående och ett beslut. Det är ett påstående om elevkunskaper så till vida att det implicerar en viss måluppfyllelse, där till exempel ett E är tänkt att betyda att en elev kan vissa saker. Det är också ett beslut eftersom det i egenskap av att vara myndighetsutövning markerar om, och i sådana fall på vilken nivå, en elev har klarat en given kurs. De enskilda kursbetygen utgör i sin tur underlag för ett »betygssnitt« som anger om en elev har en fullständig eller ofullständig examen. Denna är, i nästa led, ett certifikat som används av eleven för att till exempel söka gymnasieutbildning.

Avslutning

Som nämndes inledningsvis får all bedömning konsekvenser. Eftersom dessa konsekvenser kan vara omfattande måste vi vara säkra på att bedömningen håller hög kvalitet.

Det här kapitlet har introducerat begreppen validitet och reliabilitet som redskap för att tala om mätningars eller bedömningars kvalitet. Det har också skissartat och ytligt redovisat konturerna av en av vår samtids mest inflytelserika validitetsmodeller, nämligen den av Michael T. Kane. Jag har avsiktligt avstått från att ägna kapitlet åt en taxonomisk framställning till förmån för en framställning som förhoppningsvis röjer något av sättet på vilket man *tänker med hjälp av en validitetsmodell*.

I många länder erbjuds både lärare och elever externt producerade och administrerade prov. I Sverige har man valt en annan väg, vilket lägger ett stort ansvar på läraren. I ett annat sammanhang har jag sammanfattat det ansvaret så här:

Givet situationen i Sverige är kraven på lärare därmed mycket höga; förutom att vara experter på lärande, kunskap och ett eller flera skolämnen, måste lärare också vara väl insatta i former för och konsekvenser av bedömning och betygsättning. Dessutom

måste de dela uppfattningar om lärande och kunskap och bedömning inom sitt ämnesområde med andra lärare. Åtminstone om elever runtom i landet skall garanteras en likvärdig skola och om betygen skall vara likvärdiga i den bemärkelsen att de faktiskt representerar likvärdiga kunskaper oaktat var i landet en elev genomgått sin utbildning. (Skar, 2013, s. 1).

Att ha något slags validitetsmodell som tankestötta i den situation som beskrivs ovan kommer att gynna inte bara läraren, utan även andra intressenter, som elever, föräldrar och allmänheten.

Referenser

- American Educational Research Association, American Psychological Association & National Council on Educational Measurement (red.). (2014). *Standards for educational and psychological testing*. Washington: American Educational Research Association.
- Crooks, T.J., Kane, M.T., & Cohen, A.S. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, 3(3), 265–282. doi:10.1080/0969594960030302.
- Kane, M.T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527–535. doi:10.1037/0033-2909.112.3.527.
- Kane, M.T. (2006). Validation. I R.L. Brennan (red.), *Educational measurement* (4 uppl.) (s. 17–64). Westport, CT: Praeger Publishers.
- Kane, M.T. (2011). Validating score interpretations and uses. *Language Testing*, 29(1), 3–17. doi:10.1177/0265532211417210.
- Kane, M.T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. doi:10.1111/jedm.12000.
- Kane, M.T., Crooks, T.J. & Cohen, A.S. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18(2), 5–17. doi:10.1111/j.1745-3992.1999.tb00010.x.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. doi:10.1177/026553229601300302.
- Newton, P.E. & Shaw, S.D. (2014). *Validity in educational and psychological assessment*. Thousand Oaks, CA: SAGE Publications.
- Parmenius Swärd, S. (2008). *Skrivande som handling och möte: gymnasieelever om skrivuppgifter, tidsvillkor och bedömning i svenskämnet [Writing as action and interaction – upper secondary school students on writing tasks, time conditions and assessment in the subject of Swedish]* (Doktorsavhandling, Lunds universitet/Malmö högskola, Lärarutbildningen). Tillgänglig: <http://hdl.handle.net/2043/6562>.
- Skar, G. (2013). *Skrivbedömning och validitet. Fallstudier av skrivbedömning i svenskundervisning på gymnasiet. [Validity and classroom-based writing assessment: case studies of writing instruction and assessment in upper secondary school.]* (Doktorsavhandling, Stockholms universitet). Tillgänglig: <http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-88035>.

- Skar, G. (2014). *Potentiella dilemman i svensklärares klassrumsbedömning*.
Manus antaget för publicering.
- Toulmin, S. (2003). *Uses of argument* (2 uppl.). Cambridge: Cambridge
University Press.

3. Skrivande som grundläggande färdighet: förståelse, bedömning och undervisning

RAGNAR THYGESEN & KJELL LARS BERGE

Ämnesspecifikt och ämnesövergripande lärande

I skolsammanhang anses några färdigheter vara grundläggande.¹ Det innebär att de anses utgöra en förutsättning för lärande i alla ämnen och för ett ämnesövergripande lärande. I den norska skolan räknas skrivande, muntlighet, läsning, räkning och digitala färdigheter som grundläggande, enligt den läroplan som togs i bruk 2006. Denna uppfattning är principiellt viktig när det gäller förståelsen för vad det innebär att lära och hur lärande går till. Att förstå, lära och vara verksam inom ett ämne kan inte ses som oberoende av vad det vill säga att skapa mening med språket eller med hjälp av andra meningsskapande teckensystem. Eleverna måste förberedas för att skriva, läsa och tala på ett ämnesrelevant sätt. Drar vi paralleller till vuxenlivet, innebär yrkeskompetens i så kallade praktiska yrken, till exempel att kunna bygga ett hus eller att reparera en bil, dessutom att vi måste kunna tala med varandra på ett relevant och precist sätt, att vi kan ta till oss relevant kunskap

1 Författarna till detta kapitel tillhör en forskargrupp som gemensamt har utvecklat de teoretiska modeller som presenteras i kapitlet, liksom de bedömningskriterier som redovisas. Gruppen består i alfabetisk ordning av följande forskare: Kjell Lars Berge (Universitetet i Oslo), Lars S. Evensen (Norges teknisk-naturvitenskapelige universitet), Rolf B. Fasting (Høgskolen i Oslo og Akershus), Trine Gedde-Dahl (Høgskolen i Oslo og Akershus), Synnøve Matre (Høgskolen i Sør-Trøndelag), Hildegunn Otnes (Norges teknisk-naturvitenskapelige universitet), Randi Solheim (Høgskolen i Sør-Trøndelag), Ragnar Thygesen (Universitetet i Agder) och Wenche Vagle (Universitetet i Oslo).

genom läsning och att vi kan skriva för att dokumentera det som görs. Ett ämne är alltså inte bara något som sitter i huvudet, och det är heller inte uteslutande tyst kunskap som vi praktiserar genom och med kroppen. Ett ämne är också något som vi »utövar« genom att själva och tillsammans med andra artikulera oss genom språket och andra meningsskapande resurser i texter. Att vara kompetent i ett ämne innebär med andra ord att kunna tala, läsa och skriva på ett ämnesrelevant sätt.

Det är emellertid viktigt att understryka att skrivande som grundläggande färdighet också ska ta hand om uppgifter som inte är direkt ämnesspecifika. Varken vuxna eller elever skriver enbart inom tydligt definierade ämnen. Att kunna skriva på ett relevant sätt är också ett villkor för att kunna delta i politiska eller andra frivilliga organisationer. Dessutom är förmågan att skriva viktig för alla människors personliga utveckling. Som många texthistoriskt orienterade forskare har visat (t.ex. Lotman, 1990; Luhmann, 1986) utvecklades den moderna skönlitteraturen genom skriftliggörande av en persons självkommunikation. Det är frestande att, som exempel, hänvisa till Karl Ove Knausgårds romanverk *Min kamp*, där ett karaktäristiskt drag är just det självutforskande. Liknande självkommunicerande strategier i texter kunde påvisas i den så kallade KAL-undersökningen av 15-16-åringars skrivande vid examensprovet i norska för grundskolan (Berge, Evensen, Hertzberg & Vagle, 2005). I KAL-undersökningen dokumenterades att många av de duktigaste elevskribenterna använde berättande som egenkommunikation om upplevelser. Eleverna hade sig själva som modelläsare (Berge, 2005). Följaktligen måste skrivande som grundläggande färdighet förstås både som en ämnesövergripande färdighet och som en färdighet i varje ämne.

Det faktum att ämnesskrivande har blivit en viktig uppgift för skolan innebär att det ställs mer omfattande krav på skrivundervisningen, samt på bredden och kvaliteten i elevernas texter. Denna utvidgning av skrivundervisningen kan lätt utgöra grund för en konflikt mellan å ena sidan skrivpedagoger som är orienterade

mot formalistisk kunskap, och å andra sidan skrivpedagoger som är orienterade mot eleven som en skapande författare av texter. Så länge det sistnämnda perspektivet dominerar kommer elevernas intresse för vad de vill skriva om och hur de vill skriva till stor del att vara avgörande för vilket slags skrivundervisning som erbjuds. Många elever kommer då att specialisera sig på att skriva berättelser med utgångspunkt i egna erfarenheter och privata teman. När ämnesspecifikt skrivande blir en viktig uppgift kan skolan emellertid inte längre tillåta att elever bara skriver utifrån egna intressen och preferenser. Deras intressen kommer då också att utmanas genom att det ställs tydligare formella krav på hur deras texter ska utformas. Det senare gäller inte minst kravet på språkbruk, stavning, interpunktion och dessutom användning av textmediet.

En sådan här utvidgning av skrivundervisningen gör att den tillsynes knyts närmare den australiska genreskolans skrivpedagogiska sätt att tänka (Berge, Coppock, & Maagerø, 1998). I den betonas särskilt explicit undervisning av formella färdigheter. Detta till trots är det, som Hertzberg (2001) har pekat på, inte nödvändigtvis någon motsättning mellan elevorienterad undervisning och sådan formalistiskt orienterad undervisning som genreskolan föreskriver. Också en formorienterad och genreorienterad skrivundervisning tar sin utgångspunkt i elevernas erfarenheter, nyfikenhet och intressen, men grunddragen kommer att vara annorlunda. En formorienterad och genreorienterad skrivundervisning tvingar eleverna att lära sig att handskas med ämneskunskaper och egna erfarenheter genom en rad olika sätt att skriva som de måste tillägna sig i skolan. Muntligt berättande är i en sådan undervisning inte längre ett relevant rättesnöre för elevers skriftliga texter, utan även ganska unga elever konfronteras här på ett tydligt vis med och utmanas i att skriva på ett vetenskapligt sätt.

Vad är skrivande?

När skrivande uppfattas som grundläggande för både ett ämnesspecifikt och ämnesövergripande lärande, kan skrivundervisningen inte längre innebära att elever specialiserar sig på att skriva berättelser med utgångspunkt i privata upplevelser. Eleverna måste skriva texter som både har ett relevant ämnesinnehåll, och som byggs upp i enlighet med strukturen för texter i det aktuella ämnet. Detta innebär att följande fråga blir relevant att ställa: Hur ska skrivundervisningen planeras för att skapa intresse hos elever att skriva goda ämnestexter, det vill säga texter som inte bara tar utgångspunkt i egna erfarenheter och intressen? Eller för att formulera frågan på ett annat sätt: Vilka förutsättningar krävs för att man ska lyckas med undervisningen i ämnesspecifikt skrivande? Dessa frågor är ämnade att inspirera lärare att för sig själva och för sina elever tydliggöra hur de önskar att använda skrivande på ett relevant sätt i ett visst ämne. Vad innebär det att skriva i NO-ämnen? Vad betyder relevant skrivförmåga i geografi? Vilka slags texter ska skrivas i livskunskap? Lärarna måste alltså reflektera över vad de använder skrivande till i sina respektive ämnen. I tillägg måste de prioritera olika skrivaktiviteter i de olika årskurserna.

Vi kommer i det nedanstående att rikta uppmärksamheten mot några förutsättningar som vi menar är centrala för en ämnesrelevant undervisning där skrivande ses som grundläggande för lärande.

Grundläggande förutsättningar

Skrivande och skrivförmåga måste förstås som ett komplext fenomen, som inte kan generaliseras utifrån en viss typ av skrivande. Att skriva innebär att kunna delta i många olika aktiviteter, vilka i sin tur kräver olika strategier. Att man har utvecklat en förmåga att skriva fiktionsberättelser innebär inte att man nödvändigtvis också kan skriva goda debattartiklar. Så när skrivundervisningen ska utvecklas i och mellan ämnen, hur komplex ska skrivförmågan definieras då, och vilka urval ska göras i relation till denna komplexitet?

Om det inte finns någon generell modell över hur skrivförmåga utvecklas, individuellt eller på gruppnivå, måste skrivförmåga då hållas för att vara något som först och främst går att förstå kulturellt och situationellt? För den norska forskargrupp som har utvecklat modellerna för undervisning av skrivande som en förutsättning för ämnesspecifikt och ämnesövergripande lärande, har en utgångspunkt för arbetet varit att lösningarna på utmaningarna måste bygga på å ena sidan en kombination av forskningsbaserade insikter, och å andra sidan på kunskap om duktiga och engagerade lärares praktiska erfarenheter av skrivundervisning. Eftersom det finns så lite skrivforskning att ta sin utgångspunkt i, har det därför varit helt nödvändigt att bygga på lärares praxisbaserade insikter.

Under sitt arbete med modellerna utformade forskargruppen, i samarbete med representanter för Utdanningsdirektoratet², en definition av skrivande som grundläggande färdighet:

Skrivande har flera grundläggande funktioner – med skrift kan vi kommunicera med människor som inte är närvarande samtidigt som oss; vi kan tillvarata våra tankar och vi kan sortera dem, och vi kan föreställa oss dagens och gårdagens verklighet, men också det som ännu inte varit.

Grundläggande skrivförmåga är förmågan att använda dessa funktioner genom skriftligt meningsskapande och skriftligt hantverk. I skriftligt meningsskapande kommunicerar vi med olika syften i olika situationer – genom text, komposition, meningar, ord, stavning, interpunktion, teckningar och andra symboler. Genom det skriftliga hantverket samordnar vi mentala, motoriska och formella delförmågor som till exempel ger oss en läslig handstil och en god språklig framställning.

Skrivsituationer är ofta kopplade till verksamhet i ämnen.

2 Utdanningsdirektoratet är en norsk statlig myndighet under Kunnskapsdepartementet. Utdanningsdirektoratet ansvarar för implementering av politiska beslut på utbildningsområdet i Norge, samt för tillsyn och kvalitetsutveckling av förskola och skola. (Jfr Skolverket.) (Red. anm.)

Sättet att skriva i olika ämnen återspeglar därför ämnenas egenart, inre logik och arbetssätt. I några ämnen är tekniska beskrivningar av apparatur och framkomliga arbetssätt viktiga; i andra ämnen är det förmågan att skildra känslor som är viktig. Något som är speciellt för skrivförmåga är därmed att den är en förutsättning för de flesta ämnen, samtidigt som den är särskilt central i modersmålsämnet. Det nya är att betona betydelsen av skrivundervisning i alla ämnen.

På den mest elementära nivån innebär skrivförmåga att kunna vända sig till andra genom enkla meddelanden som blir lästa och besvarade. Utvecklingen av skrivförmåga från en elementär nivå sker när barn erfar att skrivande är något de behöver på successivt fler områden och lär sig sätt att skriva på som ger dem de redskap de behöver till det. Förmågan att använda sådana redskap stärks genom relevant träning under hela skolgången – både manuellt och elektroniskt – liksom över de gränser som skolämnena sätter. Till att börja med är utvecklingen förbunden med elevens vardagsnära erfarenheter; senare blir den också mer samhällsorienterad och ämnesanknuten. Genom läroplanens kompetensmål utvecklas skrivförmåga i alla ämnen. (<http://www.udir.no/Upload/lareplaner/veiledning/grf.pdf>)

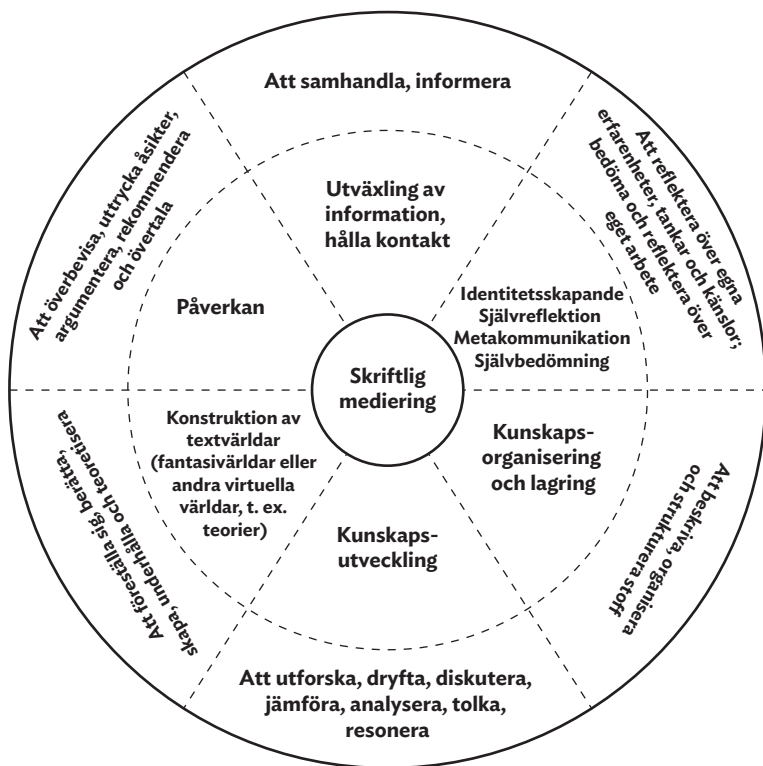
Som texten visar, tillfredsställer definitionen förväntningar om att skrivförståelse ska knytas till ämnena och deras kompetensmål.³ Strama krav som dessa kan emellertid ha den olyckliga konsekvensen att skrivande i skolan uteslutande förstås som ämnesskrivande. Det är därför viktigt att understryka, som vi varit inne på tidigare i detta kapitel, att skrivförmåga är mer än skrivande i ett ämne och utifrån det ämnets premisser. Detta är en uppfattning som kommer klart till uttryck i inledningen till den norska skolreformen:

³ Kompetensmålen i den norska läroplanen uttrycker vad eleverna ska kunna göra (jfr de svenska kunskapskraven), men inte på en graderad skala. (*Red. anm.*)

De grundläggande färdigheterna läsning, skrivande, muntlighet, räkning och digitala färdigheter ingår på ett relevant sätt i alla ämnen. De grundläggande färdigheterna är inte ämnen i sig, utan de är viktiga för att kunna agera i ämnet. Samtidigt har de grundläggande färdigheterna ett egenvärde som går utöver gränserna för och mellan ämnena. De grundläggande färdigheterna utgör resurser för bemästrande av komplexa lärande- och samhandlingsituationer i utbildning, i yrkesliv och i samhälle och hem. Att bemästra de grundläggande färdigheterna skapar förutsättningar för den enskilda människans självtillit och myndiggörande. På så sätt utgör de grundläggande färdigheterna resurser för skolans bildningsarbete. (<http://www.kunnskapsloftet.no>)

Att skriva innebär att utföra många olika handlingar i olika situationer. Med hjälp av skrift förmedlar vi kunskap som redan finns, vi skapar ny kunskap, vi utvecklar teorier och föreställer oss alternativa världar, och vi använder skrift för att förhålla oss till varandra och för att upprätthålla mellanmännsliga kontakter. Att skriva på dessa olika sätt innebär också att vi skapar mening på olika sätt. Våra skrivhandlingar är dessutom representativa både för etablerade kulturkontexter och för friare situationskontexter. I skolan innebär detta att eleverna både ska kunna skriva på ett ämnesrelevant sätt (t.ex. en vetenskaplig rapport i NO) och behärska en situation där det rör sig om att övertyga läsaren om en viss uppfattning (t.ex. en insändare i en tidning eller en tidskrift).

SKRIVHJULET – EN MODELL FÖR KOMMUNIKATIVA,
STRUKTURELLA OCH TEKNOLOGISKA SIDOR AV SKRIVANDE
Det så kallade Skrivhjulet illustrerar skrivande som en grundläggande färdighet i skolan (och i samhället i övrigt). Modellen består av två huvuddelar: *skrivandets kommunikativa funktioner* och *den semiotiska medieringen* som skriften och språket innebär, det vill säga strukturella och teknologiska sidor av skrivande.

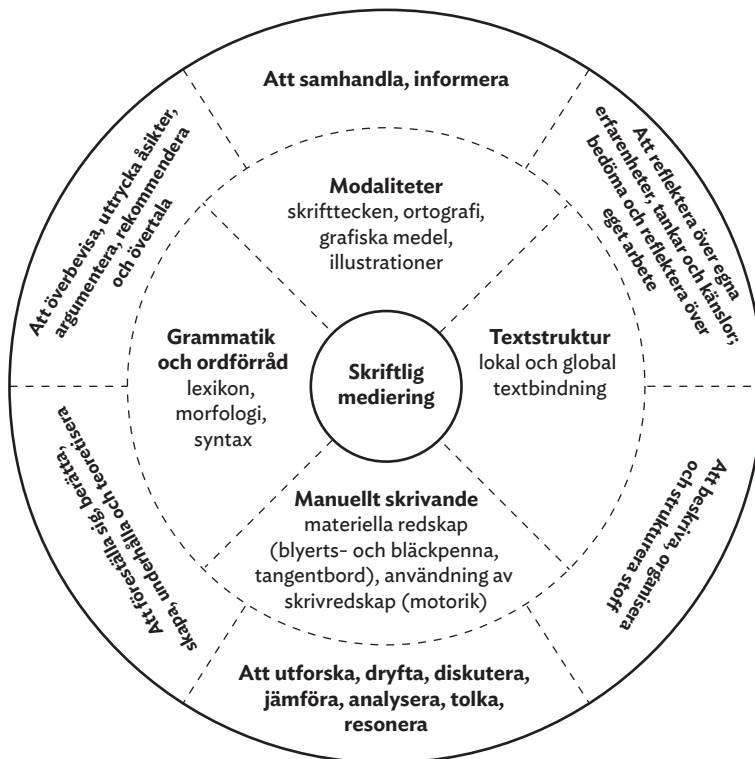


Figur 3.1. Skrivhjulet med fokus på funktionella aspekter av skrivande.

Den ena delen av skrivcirkeln, skrivandets kommunikativa funktioner, består av flera cirklar, som kan vridas i förhållande till varandra och därmed peka på nya komplexa sammanhang (jfr ett kullager). Den innersta cirkeln omfattar följande grundläggande syften som skrivande har i vår kultur: *identitetsutveckling och självreflektion, metakognition och självbedömning; kunskapsorganisering, -lagring och -strukturering; kunskapsutveckling; konstruktion av textvärldar samt påverkan.*

Den yttersta cirkeln omfattar i sin tur följande situationsbestämda skrivhandlingar som svarar mot dessa funktioner: *att samhandla; att reflektera över; att beskriva; att utforska; att föreställa sig samt att övertalning.*

Skrivhjulet är menat att vara en generalisering av allt som vi kan använda skrivande till. Det är inte begränsat till särskilda genrer och relationen mellan *handlingar* och *syften* är inte låst, utan flexibel och delvis överlappande. Vi kan med andra ord utan vidare berätta för att skapa en egen fiktiv värld där möjligheten att skapa en underhållande historia är ett mål i sig självt (*konstruktion av textvärldar – att föreställa sig*). Men vi kan också berätta en historia för att föra fram och hävda åsikter där berättelsen fungerar som ett exempel, som Astrid Lindgrens artikel »Pomperipossa i Monismannien« eller Aisopos fabler (*påverkan – att föreställa sig*). På så sätt kan de två cirkelarna vridas och ge upphov till flera olika kombinationer av syften och skrivhandlingar.



Figur 3.2. Skrivhjulet med fokus på semiotiska medieringsresurser.

I den andra delen av skrivcirkeln, som fokuserar den semiotiska medieringen, är skriftspråket en semiotisk resurs på samma sätt som det muntliga språket.⁴ Därför är också redskapen som utgör fundamentet för att skriva inlemmade i modellen. En kompetent skribent måste ha ett skriftligt ordförråd och kunskap om grammatiska normer och behärska skriftens modaliteter, inklusive regler för interpunktion och stavning. Vidare måste hon eller han ha kunskap om olika sätt att strukturera text på samt ha utvecklat vissa motoriska färdigheter i att använda blyerts- och bläckpenna, tangentbord och digitala skrivprogram. Avsikten med de funktioner som är knutna till skriftens redskap är att presentera det skrivna på ett sätt så att texten kommunicerar med en tänkt eller given läsare.

Skrivhjulet kan med andra ord användas för att på ett systematiskt sätt urskilja allt vi använder skrivande till och hjälpa oss att etablera fokus i skrivuppgifter. Om vi i undervisningen till exempel önskar att eleverna ska beskriva fotosyntesen är det organisering av kunskap vi önskar bedöma.

KRITERIER FÖR FORMULERING AV RELEVANTA SKRIVUPPGIFTER

Skrivuppgifter ska självklart vara förankrade i Skrivhulets sätt att definiera skrivande på. Dessutom måste konstruktionen av dem vara nära förbunden med de lärandemål som sätts för ämnet eller det aktuella arbetsområdet, vilket betyder att lärandemålet måste preciseras tydligt. Vi vill med anledning av detta referera till några undersökningar som understryker betydelsen av att lärandemål tydliggörs.

Doherty och Hilberg (2007) genomförde en kvasiexperimentell undersökning där 23 lärare och 394 elever i årskurs 3 och 5 i två ame-

4 Den teoretiska inspirationen till hjulet kommer bland annat från funktionellt inriktad semiotik, som Pragskolan (Gammelgaard, 2003), talhandlingsteori och socialemiotik (Berge, Coppock, & Maagerø, 1998). Det är också inspirerat av teorier om semiotisk mediering (Mertz & Parmentier, 1985).

rikanska grundskolor deltog. Resultaten visar att lärarnas användning av planerad, målinriktad undervisning predicerade elevernas prestationer i språkförståelse, läsning, stavning och vokabulär på ett standardiserat prov i slutet av skolåret. Det framgår emellertid att den målinriktade undervisningen bara utgjorde en del av ett mer omfattande pedagogiskt upplägg, som även karaktäriserades av en dialogisk undervisningsform som innebar att lärare och elever hade gemensamt fokus på målen för lärandet. Integreringen av målorientering i en sammanhängande pedagogisk kontext tycks alltså vara mycket betydelsefull.

Copeland med flera (2002) rapporterade att eleverna i deras undersökning vann på att förhålla sig till specifika målformuleringar och kontinuerlig vägledning medan de arbetade mot uppsatta mål. Det är intressant att notera att elever som mottog utvidgad vägledning i att själva formulera mål för eget arbete oftare lyckades nå de uppsatta målen.

Hattie och Timperley (2007) har betonat att lärare bör göra elever förtroga med ämnets överordnade mål. Dessutom kan lärarna avtala individuella mål och göra eleverna medvetna om kriterier för måluppfyllelse. Även dessa nyzeeländska forskare framhäver betydelsen av att involvera elever i utformningen av konkreta mål för skolarbetet. Frågan som elever bör ställa sig själv är: *Vart är jag på väg?* Målen måste då relateras till specifika prestationer eller en specifik förståelse. Väl formulerade mål innehåller både en utmaning och en prestation, eller en förståelse. Att lärarna sörjer för utmanande uppgifter resulterar både i att eleverna blir mer engagerade och i bättre prestationer. Mål är också mer verkningfulla när eleverna känner en gemensam förpliktelse att nå dem. Sådana förpliktelser kan skapas på olika sätt; ett effektivt sätt är att göra det med hjälp av förebilder. Till exempel kan lärare låta elever se en film om en högpresterande elev som utför en uppgift i en klassrumssituation. Detta resulterar ofta i att eleverna efteråt uppvisar en förpliktelse att formulera svåra mål. Specifika och konkreta mål leder till större engagemang medan hierarkiskt or-

ganiserade mål ger vägledning både på kort och lång sikt. Mål som utmanar eleverna just på den nivå där de upplever svårigheter, mål som eleverna har formulerat själva och mål som är satta av läraren och som eleverna uppfattar som förnuftiga, motiverar eleverna mer än lättare och mer tillfälliga mål gör, enligt Zimmerman (2008).

I nästa steg, efter att målen har formulerats, måste uppgifterna preciseras med utgångspunkt i Skrivhjulet, det vill säga innehålla ett klart fokus på det specifika syftet med texten som eleven ska producera, samt information om vilken skrivhandling som är aktuell eller relevant att använda.

Som exempel kan vi utgå från kursplanen i svenska där ett viktigt tema, förstås, är *språket*. I svenskämnet syftesformulering står det att: »Genom undervisningen ska eleverna ges möjlighet att utveckla sina kunskaper om svenska språket, dess normer, uppbyggnad, historia och utveckling samt om hur språkbruk varierar beroende på sociala sammanhang och medier« (<http://www.skolverket.se/laroplaner-amnen-och-kurser/grundskoleutbildning/grundskola/svenska>).

Specifika mål som relaterar till detta är att »anpassa språket efter olika syften, mottagare och sammanhang« och »urskilja språkliga strukturer och följa språkliga normer« (ibid.). Med utgångspunkt i kursplanens centrala innehåll finns sedan möjlighet att formulera mer konkreta lärandemål kopplade till detta. En sådan utgångspunkt skulle kunna vara följande centrala innehåll: »Beskrivande, förklarande, utredande, instruerande och argumenterande texter, till exempel tidningsartiklar, vetenskapliga texter, arbetsbeskrivningar och blogginslag. Texternas syften, innehåll, uppbyggnad och språkliga drag.« (ibid.).

På vilket sätt kan eller bör eleverna då arbeta mot detta lärandemål i sitt skriftliga arbete? En genomgång av aktuella syften och möjliga skrivhandlingar i Skrivhjulet pekar mot syftet *kunskapsorganisering, -lagring och -strukturering* och mot skrivhandlingen *att reflektera*. Skrivuppgiften till eleverna kan då formuleras så här: »Skriv en text där du reflekterar över varför olika situationer,

sammanhang och mottagare innebär att du talar och skriver på olika sätt. Ta utgångspunkt i några konkreta exempel.« Läraren kan i arbetet med att formulera kriterierna (se nästa avsnitt) då specificera hur man i en sådan text kan se att »eleven kan skriva [...] med [...] språklig variation, [enkel/utvecklad/välutvecklad] textbindning samt [...] fungerande anpassning till texttyp, språkliga normer och strukturer« (ibid.).

Andra exempel på hur kriterier för utformning av uppgifter kan användas återges nedan. Uppgifterna är konstruerade med utgångspunkt i skrivhandlingarna *att beskriva*, *att föreställa sig* och *att överbevisa*, och är i de två första fallen formulerade för elever i årskurs 5 och i det sista fallet för elever i årskurs 8.⁵ Nedan anges handling, tema och uppgift:

- *Att beskriva* Den digitala vardagen: Beskriv vad du använder din dator till en vanlig eftermiddag eller kväll i ett svar till mormor och morfar/farmor och farfar.⁶
- *Att föreställa sig* Livet i familjen: Föreställ dig att du byter roll med en vuxen för en dag. Skriv om denna dag.
- *Att överbevisa* Den digitala vardagen: Skriv till rektor för att överbevisa henne eller honom om att också din skola ska starta ett projekt där eleverna får egna bärbara datorer.

Skrivhjulet lägger på det här sättet grunden för en varierad undervisning i skrivande som ett redskap för såväl ämnesspecifikt som ämnesövergripande lärande. Genom att introducera modellen utmanas elevernas etablerade förståelse av vad skrivande i skolan innebär och de kan motiveras till att utforska de många möjligheter som hjulet representerar.

5 Alla uppgiftsexempel är hämtade från Utdanningsdirektoratets urvalsprov (»utvalgsprøve«) i skrivande som grundläggande färdighet år 2012. Urvalsproven är en central del av den norska, nationella bedömningen av elevers skrivande.

6 I det norska originalet används begreppen »bestemor« och »bestefar«, vilka kan syfta på såväl far- som morföräldrar. (*Red. anm.*)

BEDÖMNING AV ELEVTEXTERS KVALITET

Som en kommentar till det nationella provsystemet på Nya Zeeland har den mångåriga ledaren för NEMP, Terry Crooks (2001), betonat att en helt nödvändig förutsättning för lärande är *motivation*. Och en av de viktigaste påverkansfaktorerna för motivation är just bedömning. Det är därför av fundamental betydelse att genom bedömning optimera den motiverande effekt som en återkoppling innebär. En viktig aspekt av bedömning är hur eleverna uppfattar ett kriterierelaterat bedömningssystem (t.ex. det svenska) när det gäller deras egen motivation för lärande. Kriterierelaterad bedömning kan höja elevernas motivation om de inser att bedömningen baserar sig på individuella prestationer i förhållande till mål och inte i förhållande till hur andra elever presterar (s.k. normrelaterad bedömning).

Den forskargrupp som utvecklade Skrivhjulet och utformade de norska nationella urvalsproven i skrivande har även utvecklat riktlinjer för hur elevtexter ska bedömas. Till grund för deras riktlinjer ligger en så kallad primärdragsmodell. *Primärdragsbedömning* placerar sig mellan holistiska och analytiska bedömningsformer. Denna tredje väg för bedömning riktar uppmärksamhet mot centrala drag i olika skolämnena på ett relevant sätt.

I bedömningen av en elevtext är skrivhandlingen (jfr Skrivhjulet) då primärdraget. Med hjälp av primärdragsmodellen bedöms områdena *kommunikation*, *innehåll*, *komposition*, *språkbruk*, *stavning*, *interpunktion* och *bruk av skriftmediet* i ljuset av den skrivhandling som framgår av uppgiften (Thygesen et al., 2007). Listan över bedömningsområden är inte särskilt originell eftersom samma områden återkommer i de flesta bedömningsmatriser som används när skrivande ska bedömas (Berge, 2005).

Kommunikation relaterar både till elevens förmåga att skriva på ett sätt som är relevant i förhållande till ett specifikt syfte och specifika skrivhandlingar, men också till förmågan att få läsaren med på det perspektiv och den förståelse som förs fram i texten.

Innehåll relaterar till elevens förmåga att använda den kunskap

som hon eller han behöver för att framställa texten på ett relevant sätt. Elevens kunskap och/eller erfarenhetsbaserade insikt kan aldrig helt skiljas från hur den uttrycks. Därför är kvaliteten i texter beroende av skribentens kunskaps- och erfarenhetsnivå. En svår avvägning när det gäller innehållet kan vara i vilken grad kunskaperna som eleven redogör för är korrekta. Det är dock självklart att läraren i sin bedömning måste ta hänsyn till om upplysningar i texten är korrekta eller inte. En aldrig så välskriven text om exempelvis ett författarskap kan inte innehålla sakfel och ändå godkännas, lika lite som en NO-text som hävdar att jorden är platt kan godkännas.

Komposition relaterar till textstrukturell kvalitet. Eleven ska kunna skapa en enhetlig text med en yttre form som ger läsaren en uppfattning om vilket slags text det handlar om. Den ska ha en utvecklad struktur, något som omfattar både global och lokal komposition. Texten kan också byggas upp runt olika texttyper, som argumentation, kausalitet, narrativitet och deskriptivitet. Hur dessa kommer till uttryck och är utvecklade i texten i förhållande till varandra bidrar också till att avgöra textens kvaliteter.

Språkbruk handlar om elevernas förmåga att kunna använda språkets resurser, ord och syntax på sätt som är relevanta för skrivhandlingen och situationen de skapar texten i. På så sätt har kravet om språkbruk med stil och register att göra. Studier av utvecklingen av elevspråk visar att det är krävande att gå från ett sätt att använda språket på, som har en förlaga i det muntliga registret, till att använda språket i specialiserade register som är anpassade till skriftspråkliga normer. Man kan helt enkelt inte skriva som man talar. Den stora norska elevtextundersökningen visade emellertid att till och med tämligen svaga elever inte utan vidare skriver på ett muntligt sätt när de skriver berättelser. De anpassar språket till situation och medium och försöker skriva så att man kan tolka textens mening på ett relevant sätt också i situationer där man inte känner till vem det är som har skrivit den. Denna situationssensitivitet verkar nästan intuitiv (Karlsson, 1997).

Kraven om *stavning* och *interpunktion* är oproblematiska. Det

är viktigt att alla elever utvecklar automatiserade färdigheter på dessa områden. Många lärare i dag är bekymrade för nivån på stavnings- och interpunktionsförmågan i skolan. Det har hävdats att dessa områden nedprioriteras av språkpolitiska skäl. Den norska forskargruppen önskar dock tydliggöra att kravet på stavning och interpunktion måste gälla.

Av de områden som bedöms enligt primärdragsmodellen är det mest originella, och kanske det enda kontroversiella, att bedömningsområdet *bruk av skriftmediet* ses som ett multimodalt fenomen. Detta beror på att skolans skriftkultur, särskilt i modersmålsämnet (dvs. norsk- och svenskämnen), har varit utpräglat monomodal. I det som definieras som skrift ingår både traditionella verbal-språkliga och grafiska utformningar, som layout och illustrationer. Därför ingår det i förståelsen av skrivförmåga att texterna ska se ut på ett sätt som bjuder in läsaren till att bry sig om dem. Texternas visuella och estetiska kvaliteter ska också bedömas. Följaktligen är det viktigt att eleverna genom texten kan visa att de har en funktionell och läslig handstil. Alla som har läst elevtexter vet att många elever har en handstil som gör det svårt att avgöra vad som faktiskt står i texten. Emellertid har den omfattande användningen av datorer och olika typer av programvara bidragit till att människor i allmänhet i större utsträckning även fokuserar på texters estetiska dimensioner. Eftersom skrivande ses i ett multimodalt perspektiv måste det i bedömningen därför också ingå andra former för meningsskapande resurser än verbalspråket, till exempel format, layout, illustrationer, tabeller, figurer och grafer.⁷ Dessa bedömningsområden är desamma från en text till en annan och från en skrivhandling till en annan. Likväl måste kriterierna anpassas efter den enskilda skrivhandlingens särdrag. Det kommer nämligen inte att vara samma krav på kommunikation, innehåll, komposition, språkbruk, stavning, interpunktion och bruk av skriftmediet i ett

7 Utdanningsdirektoratet önskade inte att inkludera denna bedömningsdimension i 2012 års urvalsprov (jfr not 5).

personligt brev som i en text där eleven ska förklara hur en historisk händelse, som beslutet om ny grundlag i Norge den 17 maj 1814, var möjlig.

De ovan beskrivna bedömningsområdena kan med fördel konkretiseras med hjälp av »stöttande frågor«, vars uppgift är att leda bedömningen mot centrala drag inom respektive område. Detta kan exemplifieras med utgångspunkt i den ovannämnda uppgiften om varför olika situationer innebär olika sätt att tala och skriva. För bedömningsområdet *kommunikation*, som alltså gäller hur väl elevens text kommunicerar, kan relevanta stöttande frågor vara: Har texten en rubrik som berättar vad som kommer att hända, eller som gör läsaren nyfiken? Används kommunikationsfrämjande frågor, uttryck eller tilltal? För området *innehåll*, som gäller hur pass relevant textens innehåll är, kan en stöttande fråga vara: Handlar texten om hur olika situationer inverkar på sättet att tala och skriva?

De bedömningar av elevtexten som läraren gör kan sedan översättas till profiler över elevens prestationsnivå på de olika områdena (kommunikation, innehåll, komposition osv.). Det är en fördel att skapa profiler för varje skrivuppgift som eleven genomför.

Bedömning främjar lärande när den inkluderar lärandemål som är förstådda och som är gemensamma för lärare och elever, när den hjälper elever att förstå och acceptera fastlagda normer och när den gör elever delaktiga i självbedömning. En sådan bedömning innebär en återkoppling som hjälper eleverna att ta nästa steg och som bygger upp en tillit hos dem så att de kan göra arbetet bättre (Crooks, 2001).

Pedagogisk uppföljning baserat på formativ bedömning

Fokus på specificerade mål, formulering av utmanande uppgifter och grundlig analys av olika kvaliteter i elevens text är meningslöst om inte arbetet följs upp pedagogiskt av läraren genom bedömning som stödjer lärandet. Återkopplingen på elevtexten ska bidra till att

eleven upplever glädje i att känna att hon eller han kan bemästra skrivande.

Vad menar vi då med *återkoppling* i pedagogiska sammanhang? Återkoppling förstås som information från en lärare till en elev som gäller aspekter av elevens agerande eller förståelse. Återkopplingen måste ses som en konsekvens av elevens agerande.

Återkopplingen bör ge information som direkt angår uppgiften eller lärandeprocessen. Detta kan ske på många olika sätt, till exempel genom affektiva processer, såsom ökad insats, motivation eller engagemang hos eleven. Alternativt kan läraren initiera olika kognitiva processer hos eleven – i vilka det ingår att restrukturerat förståelsen, att bekräfta för eleven att hon eller han förstått korrekt eller inte, att indikera att ytterligare information är tillgänglig eller nödvändig, att visa vilken riktning eleven ska följa i det vidare arbetet och/eller indikera alternativa sätt att förstå given information.

Det är av avgörande betydelse att återkopplingen innehåller uppgiftsrelevant information; dess effekt är ett resultat av i vad mån återkopplingen inriktas mot hur väl eleven lyckas med uppgiften. Däremot ger yttre bedömningsmekanismer, som betyg, liten vinst eftersom de lätt kan undergräva elevens vilja att ta ansvar för att motivera sig själv och reglera sitt eget lärande. Återkopplingen har störst verkan när läraren ger eleven stor respons på korrekta svar och mindre på felaktiga svar, och när den tar sin utgångspunkt i en förändring i förhållande till tidigare försök.

Effekten influeras också av svårighetsgraden hos målen och uppgifterna; den har störst verkan när målen är specifika och utmanande och när uppgiftskomplexiteten är låg. Återkopplingen fungerar också bäst när eleven inte uppfattar den som ett hot mot självbilden; en icke-hotande situation ger utrymme för fokus på själva återkopplingen.

Vi använder Hattie och Timperleys (2007) modell för återkoppling som ram för att förstå varför speciella former för återkoppling främjar lärande, och som hjälp för att identifiera betingelser som enligt all sannolikhet leder till bättre läranderesultat. Huvudsyftet

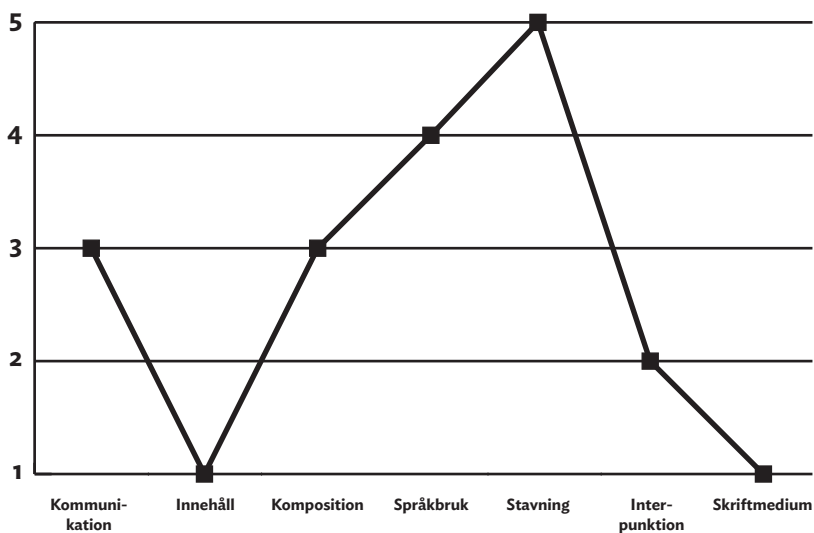
med återkoppling är att reducera avståndet mellan elevens nuvarande kunskap, förståelse och lärandebeteende och ett uppsatt mål. De strategier som både läraren och eleven använder för att reducera detta avstånd kan vara mer eller mindre verkningsfulla när det gäller att öka lärandet, därför är det viktigt att förstå vilka resultat som olika betingelser leder till.

Verkningsfull återkoppling måste svara på tre frågor, som antingen kan ställas av läraren eller eleven själv: *Vart är jag på väg?* (Vad är målet/målen?), *Var är jag nu?* (Hur väl har jag lyckats nå målen?) och *Vilket är nästa steg?* (Hur bör jag arbeta för att öka framgången?)

Hur verkningsfulla svaren på dessa frågor är för att minska avståndet är delvis beroende av mot vilket område återkopplingen inriktar sig. Dessa områden är: *mängden av arbete eleven ska lägga ner på att lösa uppgiften, sättet på vilket eleven förstår hur uppgiften ska lösas, självreglering och självuppfattning*. Återkoppling får olika verkan beroende på vilket område som avses.

Hur fungerar då återkoppling när syftet är att minska avståndet mellan befintlig och önskad förståelse? Det finns flera sätt på vilka läraren kan hjälpa till med detta, som alla inkluderar att ställa lagom utmanande och specifika mål enligt hur vi beskrivit det ovan. Specifikt formulerade mål är mer verkningsfulla än generella mål, först och främst eftersom de i större grad fångar elevens uppmärksamhet, men också för att de kan göra att återkopplingen blir mer fokuserad. Specifika mål med tillhörande återkoppling kan också lättare innehålla information om kriterierna för att nå målen än mer generella mål kan.

Genom återkoppling kan läraren också hjälpa till med att förklara målen, öka engagemanget eller stärka eleven i dennes ansträngningar för att nå dem. Målen kan också göras mer hanterliga genom att snäva in ett antal rimliga förväntningar på måluppfyllelsen. Mer generellt kan läraren skapa en lärandemiljö där eleven utvecklar en förmåga att styra sitt eget lärande och förmåga att upptäcka fel. Detta resulterar i att eleven ger sig själv återkoppling om hur målet



- 1: Mycket låg nivå
- 2: Låg nivå
- 3: Genomsnittsnivå
- 4: Hög nivå
- 5: Mycket hög nivå

Figur 3.3. Resultatprofil.

ska nås. I vad mån återkoppling bidrar till dessa processer beror dock till stor del på vad som fokuseras i återkopplingen och på det område som återkopplingen inriktas mot.

Verkningsfull undervisning inkluderar att läraren bedömer hur väl eleverna förstår det som hon eller han har undervisat om, så att nästa undervisningssteg kan anpassas efter elevernas aktuella nivå. Därefter följer återkopplingsdelen, som fokuserar på de tre ovannämnda frågorna. En ideal lärandemiljö uppstår när både lärare och elever söker svar på dessa frågor.

Att svara på frågan *Var är jag nu?* innebär att läraren ger eleven information relaterad till de uppsatta målen, ofta kopplad till ett slags standard och/eller i förhållande till hur väl eller dåligt eleven har utfört en specifik del av en arbetsuppgift (jfr bedömningsområdena återgivna ovan). Återkopplingen bidrar till, när den

innehåller information om framgång och/eller hur eleven bör fortsätta arbetet, att stärka elevens motivation för att lära.

Läraren kan vidare hjälpa eleven att bedöma sitt eget arbete och sitt sätt att lära genom att påvisa styrkor och svagheter i uppnådd kunskap och förmågor, något som medvetandegör eleven om sitt ämneskunnande. Konstruktiva återkopplingar kan ges i form av både formella och informella samtal och genom muntliga och skriftliga kommentarer. Genom dokumentation och konkretisering får eleven på så sätt svar på frågan om hur hennes eller hans situation är just nu.

Den positiva effekten av återkoppling kan också användas för att svara på frågan *Vilket är nästa steg?* Det gör läraren genom att ge eleverna information som resulterar i ökade möjligheter att lära. Eleven bör få hjälp av läraren att tänka målrelaterat och reflektera över vad hon eller han ska lära sig, vad hon eller han faktiskt har lärt sig, vad som återstår och hur arbetet bör se ut framåt. Den framåtsyftning som uttrycks i frågan kan visa sig ha stor betydelse för lärandet.

Läraren och eleven kan tillsammans lägga upp arbetsplaner med utgångspunkt i elevens individuella förutsättningar. Läraren kan med fördel presentera de förväntningar som finns på eleven. Det kan exempelvis röra sig om större utmaningar, högre grad av självständighet i arbetet eller bättre grepp om vad eleven förstått och inte förstått. De kan samtala om olika lärandestrategier och komma överens om vilka lärandestrategier som ska användas.

Återkoppling kan också röra frågan om självreglering, som inkluderar bättre förmåga att göra självbedömningar och tillförsikt inför vidare engagemang i till exempel en skrivuppgift. Ett exempel: »Du är bekant med hur ett argument inleds. Se om du fått med dig huvuddragen i det första avsnittet av din text.« Denna typ av återkoppling kan ha stor inverkan på elevens tro på sin egen förmåga att lyckas, förmåga att kontrollera och bedöma sitt agerande i lärandesituationer samt självförtroendet i lärandesammanhanget. Återkoppling på denna nivå rör hur eleven övervakar, styr och

reglerar sitt eget arbete mot lärandemålen. Självregering innefattar autonomi, självkontroll, självbestämmande och självdisciplin.

Till sist kan återkoppling vara personlig på ett sådant sätt att den rör eleven som person. Forskning visar dock att denna typ av återkoppling har lite att göra med hur stor insats eleven lägger ned i arbetet med en uppgift. Beröm är ineffektivt eftersom det innehåller knapphändig information i förhållande till de tre frågorna (se ovan) och eftersom det riskerar att flytta fokus från uppgiften. Det är emellertid viktigt att skilja mellan allmänt beröm och beröm som direkt relaterar till arbetsinsatsen, till elevens självreglering, till engagemanget eller till arbetsprocessen. Den sortens beröm kan stärka elevens tro på sig själv och i nästa steg leda till att eleven engagerar sig än mer i uppgiften.

Referenser

- Berge, K.L. (2005). Studie 6: Tekstkulturer og tekstkvaliteter. I K.L. Berge, L.S. Evensen, F. Hertzberg & W. Vagle (red.), *Ungdommers skrivekompetanse. Bind II: norskeksamen som tekst* (s. 11–190). Oslo: Universitetsforlaget.
- Berge, K.L., Coppock, P. & Maagerø, E. (1998). *Å skape mening med språk*. CAF: Oslo.
- Berge, K.L., Evensen, L.S., Hertzberg, F. & Vagle, W. (red.). (2005). *Ungdommers skrivekompetanse* (Bind I & II). Oslo: Universitetsforlaget.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Copeland, S.R., McCall, J., Williams, C.R., Guth, C., Carter, E.W., Fowler, S.E., Presley, J.A. & Hughes, C. (2002). High school peer buddies: A win-win situation. *Teaching Exceptional Children*, 35(1), 16–21.
- Crooks, T. (2001). *The validity of formative assessments*. Paper presented to the British Educational Research Association annual conference, University of Leeds, 13–15 September 2001.
- Doherty, R.W. & Hilberg, R.S. (2007). Standards for effective pedagogy, classroom organization, English proficiency, and student achievement. *The Journal of Educational Research*, 101(1), 24–35.
- Gammelgaard, K. (2003). *Tekstens mening: en introduktion til Pragerskolen*. Frederiksberg: Roskilde Universitetsforlag.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hertzberg, F. (2001). Tusenbenets vakre dans. *Rhetorica Scandinavica*, 18, 92–105.
- Karlsson, A.-M. (1997). Textnormer i och utanför skolan – att skriva insändare på riktigt och på låtsas. I: G. Håkansson (red.) *Svenskarnas beskrivning*, 22. Lund: Lund UP.
- Lotman, J. (1990). *Universe of the mind: A semiotic theory of culture*. London: I.B. Tauris.
- Luhmann, N. (1986). *Love as passion: The codification of intimacy*. London: Polity Press.

- Mertz, E. & Parmentier, R.J. (1985). *Semiotic mediation: Sociocultural and psychological perspectives*. Orlando: Academic Press.
- Thygesen, R., Evensen, L.S., Berge, K.L., Fasting, R.B., Vagle, W. & Haanæs, I.R. (2007). *Nasjonale prøver i skriving som grunnleggende ferdighet*. Sluttrapport. Nasjonalt senter for leseopplæring og leseforskning, Universitetet i Stavanger.
- Zimmerman, B.J. (2008). Goal setting: A key proactive source of academic self-regulation. I: D.H. Schunk, & B.J. Zimmerman (red.), *Motivation and self-regulated learning* (s. 267–296). Lawrence Erlbaum Associates.

Hemsidor

- Grundläggande färdigheter: <http://www.udir.no/Upload/Lareplaner/veiledning/grf.pdf>
- Kunnskapsløftet: <http://www.kunnskapsloftet.no>
- Svenskämnet: <http://www.skolverket.se/laroplaner-amnen-och-kurser/grundskoleutbildning/grundskola/svenska>

4. Att bedöma läsförmåga – teoretiska utgångspunkter

MICHAEL TENGBERG

Vad är egentligen läsförmåga?

I det här kapitlet ges en introduktion till begreppet *läsförmåga*. Meningen är att bidra med kunskaper som kan hjälpa lärare att göra bedömning av elevers läsförmåga till en självklar och integrerad del av undervisningen i svenskämnet. Varje form av bedömning måste ta sin utgångspunkt i en klar definition av *vad* det är som ska bedömas. Det här kapitlet syftar därför till att besvara frågan om vad som ska bedömas när det gäller elevers läsande och handlar om att ringa in själva fenomenet läsförmåga med utgångspunkt framför allt i den rådande kursplanen i svenska. Kapitlet kommer också att handla om några av de generella villkor som är förenade med bedömning av läsförmåga, utan att för den skull gå in på specifika bedömningspraktiker. Eftersom boken riktar sig mot bedömning i årskurserna 7–9 handlar kapitlet särskilt om de analyserande och tolkande läspraktiker som blir aktuella i högstadiet. När elever kommer till högstadiet förväntas de exempelvis vara förtrogna med genrer som sakprosatext och skönlitteratur. De förväntas också kunna göra enklare sammanfattningar av såväl huvudsakligt innehåll i sakprosatext som av handlingsförlopp i berättande texter. Detta är mål som framgår av kursplanen i svenska och som uttrycker den grundläggande läsförmåga som högstadiets läsundervisning tar sin utgångspunkt i.

För att kunna göra en relevant planering av läsundervisningen

behöver den lärare som övertar elever i årskurs 7 ganska snart bilda sig en uppfattning om hur det är ställt med elevernas läsförmåga. Många lärare använder sig kanske av diagnosmaterial, för vilket det ofta finns både bedömningsanvisningar och förslag på uppföljning i form av textmaterial och aktiviteter. Men de kunskaper eller förmågor som synliggörs i en diagnos måste vägas mot andra bevis, eller belägg, för hur eleverna läser, förstår och förhåller sig till olika slags texter. Därför är det nödvändigt att man som lärare själv har relevanta kunskaper om läsförmåga och läsprocesser så att man kan dra rimliga slutsatser om vad eleverna kan och inte kan.

Som redan framgått används här begreppet läsförmåga i stället för exempelvis »läsning« eller »läsförståelse«. Med läsning avses ju inte endast en kunskap eller kunskapsform, utan också en aktivitet, som inte nödvändigtvis alltid bör bli föremål för bedömning. När det gäller begreppet *läsförståelse* kan det betraktas som snävare än läsförmåga. Inte minst framstår läsförståelse som begränsande i relation till skönlitteraturen, eftersom dess krav på läsaren inte enbart går att beskriva i termer av *förståelse*. Begreppet läsförmåga är också avsett som en översättning av det engelska begreppet *reading literacy*, som används i bland annat PIRLS- och PISA-undersökningarna. Med *reading literacy* menas en individs förmåga att »förstå, använda, reflektera över och engagera sig i skrivna texter, i syfte att uppnå sina mål, utveckla sina kunskaper och sin potential och delta i samhället« (OECD, 2009, s. 14). Läsaren betraktas som någon som själv aktivt konstruerar mening med utgångspunkt i sina tidigare kunskaper och erfarenheter. Läsprocessen involverar därmed informationssökande och personligt meningsbyggande simultant, även om de uppgifter som prövar läsförmågan och som ska reflektera olika sätt att använda texter brukar koncentrera sig på någon av de olika delprocesserna.

Läsförmåga – ett sammansatt konstrukt

En slutsats av detta blir alltså att läsförmåga inte utgör något enhetligt konstrukt¹ utan att det är ett komplext fenomen som består av en rad olika delaspekter snarare än att vara en enhetlig egenskap hos en individ (se t.ex. Duke, 2005). De flesta tester, och andra slags bedömningsmaterial, används i avsikt att kunna fånga läsförmågan med ett enda mått. När det gäller betygsättning är detta förvisso också en av förutsättningarna för verksamheten. I betyget i ämnet svenska ska måttet på läsförmågan dessutom vägas samman med andra konstrukt, såsom skrivförmåga och muntlig förmåga. Men när syftet med bedömningen är formativt finns det anledning att understryka betydelsen av att uppmärksamma läsförmåga som ett fenomen uppbyggt av olika aspekter, såsom avkodning, läsflyt, ordkunskap, inferensförmåga på lokal och global nivå, kritisk reflektion, identifikation av textstruktur och textstrategier och så vidare. Vidare, menar Leslie och Caldwell (2009), måste lärarens planering av läsundervisningen utgå från den repertoar av texter och teman som eleverna förväntas behärska. Detta sammantaget innebär att den mångfald av aspekter som gemensamt sorterar under det vi kallar läsförmåga måste bli en del av lärares professionella medvetande och språk om de ska kunna användas i bedömningen. Detta föranleder en diskussion om vilka faktorer som i så fall påverkar elevers läsförmåga liksom vilka olika delar som är relevanta vid bedömning. När det gäller faktorer som påverkar läsförmågan kommer det här kapitlet främst att handla om betydelsen av vilket slags text som läses liksom vilka aktiviteter och vilka syften som omgärdar läsningen. När det gäller att definiera de olika delarna av konstruktet kommer jag utgå dels från hur läsförmåga beskrivs i kursplanen i svenska, dels från hur det teoretiska ramverket för PISA-undersökningen preciserar olika aspekter av läsförmågan.

1 Med konstrukt avses helt enkelt den egenskap eller förmåga hos en individ som man avser bedöma (se vidare kapitel 1).

Läsförmågan i kursplanen i svenska

Kursplanen i svenska för grundskolan inleds genom att definiera språk som »människans främsta redskap för att tänka, kommunicera och lära« (Skolverket, 2011a). Uttrycket att äga ett språk används inte sällan i betydelsen att ha förmåga att uttrycka sig i tal och skrift, det vill säga att *producera* språk, medan läsande i detta sammanhang beskrivs som en *receptiv* färdighet. Men läsande är självfallet också en aktiv språkhandling, som innebär att språkliga symboler på flera olika nivåer integreras och tolkas som information eller mänskligt uttryck i någon form. Att avkodningsprocesser såsom igenkänning av ord, ljud och meningsbyggnad ställer konkreta språkliga krav är lätt att inse. Men även de processer som brukar hänföras till *läsförståelse* (aktivera bakgrundskunskaper, integrera textavsnitt, producera inferenser, värdera textinnehåll etc.) utgår från ett oerhört komplicerat bemästrande av språkliga symboler och segment. Att läsa är således inte någon passiv språkhandling. Tvärtom. Detta underförstås i kursplanen när den använder språkutveckling som ett begrepp som (om än outtalat) inkluderar läsande, skrivande, lyssnande och talande.

Den förväntade utvecklingen av läsförmåga i årskurserna sju till nio definieras i kursplanen genom det för undervisningen angivna centrala innehållet tillsammans med kunskapskraven för de olika betygsnivåerna. Definitionen rymmer två olika dimensioner. Dels specificeras de färdigheter eller förmågor i umgänget med text som eleverna ska utveckla: »[l]ässtrategier för att förstå, tolka och analysera texter från olika medier. Att urskilja texters budskap, tema och motiv samt deras syften, avsändare och sammanhang.« (Skolverket, 2011a) Dels specificeras de texttyper och textstrukturer som elever ska bli förtrogna med:

- Skönlitteratur för ungdomar och vuxna från olika tider, från Sverige, Norden och övriga världen. Skönlitteratur som belyser människors villkor och identitets- och livsfrågor. Lyrik, dramatik, sagor och myter.

- Språkliga drag, uppbyggnad och berättarperspektiv i skönlitteratur för ungdomar och vuxna. Parallellhandling, tillbakablickar, miljö- och personbeskrivningar, inre och yttre dialoger.
- Några skönlitterära genrer och hur de stilistiskt och innehållsligt skiljer sig ifrån varandra.
- Några skönlitterärt betydelsefulla ungdoms- och vuxenboksförfattare från Sverige, Norden och övriga världen och deras verk, samt de historiska och kulturella sammanhang som verken har tillkommit i.
- Beskrivande, förklarande, utredande, instruerande och argumenterande texter, till exempel tidningsartiklar, vetenskapliga texter, arbetsbeskrivningar och blogginlägg. Texternas syften, innehåll, uppbyggnad och språkliga drag.
- Texter som kombinerar ord, bild och ljud, och deras språkliga och dramaturgiska komponenter. Hur uttrycken kan samspela med varandra, till exempel i tv-serier, teaterföreställningar och webbtexter.
- Kombinationer av olika texttyper till nya texter, till exempel informerande texter med inslag av argumentation.

Man kan tänka sig att de lässtrategier som benämns och som också finns ytterligare förtydligade i kommentarmaterialet (Skolverket, 2011b) ska kunna tillämpas på alla dessa texttyper. I kunskapskraven för betyg i slutet av årskurs 9 finns också en nivåindelning som definierar vad som menas med »grundläggande läsförståelse« (för betyget E), »god läsförståelse« (för betyget C) respektive »mycket god läsförståelse« (för betyget A).

Sättet att skilja på de olika nivåerna är inte alldeles genomskinligt, men det utgår från en tänkt progression av vad det vill säga att förstå text på ett kvalificerat sätt. Medan eleven för betyget E exempelvis ska kunna »göra enkla sammanfattningar av olika texters innehåll«, ska eleven för betyget C kunna »göra utvecklade sammanfattningar« och för betyget A »välutvecklade sammanfatt-

ningar«. En central aspekt av bedömningen av elevers läsförmåga utifrån kunskapskraven handlar således om att lärare tillsammans definierar vad som kännetecknar enkla, utvecklade respektive välutvecklade sammanfattningar av olika texters innehåll. På motsvarande vis bör man försöka hitta en samsyn för att likvärdigt kunna bedöma läsflyt, förmågan att sätta texter i sammanhang, förmågan att göra analyser och att identifiera budskap, det vill säga alla de övriga aspekter av läsandet som preciseras i kunskapskraven.

Sammantaget framtonar i kursplanen en ganska komplex definition av läsförmågan. Å ena sidan ger den en relativt tydlig vägledning till läraren ifråga om vad för slags texter och textstrukturer som eleverna bör få bekanta sig med under högstadietiden (här nämns genrer såsom argumenterande texter, lyrik och dramatik, liksom strukturella komponenter såsom berättarperspektiv, parallellhandling och tillbakablick). Å andra sidan kan det alltså vara svårt att avgöra vad som skiljer exempelvis utvecklade från välutvecklade sammanfattningar av olika texters innehåll. Till denna komplexitet fogas också kursplanens skrivningar om språkbruk och språklig medvetenhet, som också utgör en viktig dimension av läsförmågan. Inte minst ställer skönlitteraturen många gånger krav på uppmärksamhet inför och medvetenhet om språkliga nyanser, till exempel i form av ords värdeladdning och associationsfält, eftersom också till synes små stilistiska medel kan få avgörande betydelse för texttolkningen.

Det finns tyvärr inget alexanderhugg som löser den här komplexiteten och förvandlar fenomenet läsförmåga till något mycket enklare som en lärare kan ta till i undervisning och bedömning. Tvärtom måste varje lärares analys av sina elevers kunnande uppmärksamma och synliggöra mångsidigheten i fenomenet. Sedan läroplansreformen 1994 har vi vant oss vid kursplaner såsom påfallande kortfattade dokument, i vilka ämnesstoffet inte ges mer än en översiktlig och partiell definition. Kursplanen i svenska kan till exempel inte sägas implicera någon underliggande teori om läsning och läsförmåga. För att tolka kursplanens innehåll och kravnivåer

behöver lärare därför ha tillgång till en mer gedigen teoretisk grund, som låter förstå hur satserna i kursplanen kan konkretiseras i form av textval, klassrumsaktiviteter och kriterier för bedömning. En viktig del av den här bakgrunden består av gedigna kunskaper i litteraturhistoria och litteraturvetenskap, liksom goda kunskaper om språkvetenskapliga och ämnesdidaktiska teorier och modeller. Ett annan viktig del representeras av de teoretiska ramverken för PIRLS- och PISA-undersökningarna.

Läsförmågan i PISA

De teoretiska ramverk som ligger till grund för PIRLS- och PISA-undersökningarna innehåller schematiska beskrivningar av vilka olika komponenter som kan sägas ingå i begreppet läsförmåga (Mullis et al., 2009; OECD, 2009). Dess syfte är att ligga till grund för utformningen av provuppgifter så att man dels kan säkerställa att provet verkligen mäter det som det är tänkt att mäta, dels försäkra sig om att provets innehållsfokus förblir konstant över tid. Båda ramverken bygger på de senaste decenniernas läsforskning och kan därför ses som högst relevanta utgångspunkter för en diskussion om klassrumsbedömning av elevers läsande. I det här kapitlet används i första hand PISA:s ramverk, som är mer komplext än det ramverk som används i PIRLS. Betygsättning måste naturligtvis utgå från kursplanens kunskapskrav, men i den formativa bedömningen av elevers läsutveckling kan PISA:s beskrivning av olika komponenter i läsförmågan vara till god hjälp.

I ramverket beskrivs tre olika dimensioner utifrån vilka de deltagande elevernas läsförmåga bedöms: *text*, *läsprocess* och *läs-sammanhang*. Med avseende på var och en av dessa tre dimensioner väljer PISA ut texter och uppgifter som ska användas för att synliggöra och pröva läsförmåga i vid mening.

TEXT

När det gäller den första dimensionen, text, skiljer man först och främst mellan tryckta och elektroniska texter. Att PISA skiljer mellan tryckta och elektroniska texter betyder inte att man menar att förmåga att läsa tryckta texter är något helt annat än förmåga att läsa elektroniska texter. Det betyder i stället att om man vill skapa sig en sammansatt bild av en elevs läsförmåga bör man pröva den i relation till såväl tryckta som elektroniska texter. När det gäller elektroniska texter åsyftas inte enbart skärmbaserad läsning utan även läsning av *hypertexter*, det vill säga icke-sekventiella texter som förutsätter någon form av aktiv navigering av läsaren med hjälp av de navigeringsverktyg som texten och exempelvis en dator eller en läsplatta erbjuder. Vad gäller text skiljer PISA också på löpande texter, som är organiserade i meningar och stycken, och icke-löpande texter, som utgörs av listor, tabeller, diagram, reklambilder etc. Inte sällan blandas dessa två textformat, exempelvis i dagstidningar. Samtidigt kan det innebära något helt annat att orientera sig och finna rätt information i en löpande text än att finna rätt information i ett diagram. En grundlig bedömning av elevers läsförmåga förutsätter därför att de tränas och prövas i båda dessa avseenden. Slutligen skiljer PISA på vad man kallar olika texttyper. I Sverige är vi bekanta med uppdelningar som den mellan sakprosa och skönlitterär text. I PISA menar man att den här gränsen är grumlig, inte minst eftersom litterära texter inte sällan lånar typiska drag från sakprosatekst och vice versa. Därför skiljer man inte på sakprosa och skönlitteratur utan på beskrivande, berättande, utredande, argumenterande, instruerande och transaktiva texter. Att här göra en utläggning om skiljelinjerna mellan dessa olika texttyper skulle dock kräva alltför mycket utrymme. Liknande termer används dessutom i den svenska kursplanen. Däremot är det intressant att notera hur den traditionella genreavgränsningen utifrån konventioner för språkliga, stilistiska och strukturella drag i texter, som antas reglera läsarens förväntningar på texten, i PISA:s ramverk bytts mot en typologisering utifrån de språkhandlingar

(beskriva, berätta, utreda etc.) som antas vara bärande i texten (se vidare kapitel 3).

LÄSPROCESS

Den andra dimensionen framstår åtminstone i en svensk kontext som ett viktigt komplement till förståelsen av vad läsförmåga är. Här specificeras nämligen de olika läsprocesser genom vilka en läsare interagerar med en text. Läsprocesser ska förstås som sätt att definiera den kognitiva komplexitetsgrad som kännetecknar de olika uppgifter en läsare tar sig an antingen i förhållande till en viss aspekt i texten eller i förhållande till texten som helhet. Enligt Afflerbach och Cho (2011), som gjort åtskilliga studier på området, måste bedömning av elevers läsande ta sin utgångspunkt i en bred förståelse av de kognitiva processer som ingår. I PISA-undersökningen skiljer man på tre olika läsprocesser: *söka och inhämta information; sammanföra och tolka*; samt *reflektera och utvärdera* (OECD, 2009, s. 34ff).

Att söka och inhämta information handlar i detta sammanhang om att dels lokalisera dels använda uttalad information i en text. Det är viktigt att komma ihåg att förmågan att lokalisera och välja rätt information är beroende av hur texten är organiserad, liksom av hur mycket distraherande information texten rymmer. I en berättelse där det vimlar av karaktärer kan det vara svårt att hålla isär relationerna mellan olika karaktärer medan detta ofta är lättare i en berättelse med få karaktärer.

Att sammanföra och tolka utgör en mer sammansatt process, som naturligtvis förutsätter att läsaren kan söka och inhämta information. Med sammanföra menas att läsaren integrerar innehåll från olika delar av en text, exempelvis mellan meningar som följer på varandra eller mellan information som tillhandahålls i olika styckena i en text. Med tolka menas att läsaren drar en slutsats som inte finns explicit utskrivet i texten. Den här sortens meningsskapande skiljer sig alltså från sökläsandet och är nödvändigt för att läsaren ska kunna skaffa sig en bredare förståelse av textens ärende och

mening. Läsaren förväntas till exempel kunna göra jämförelser och visa på kontraster mellan olika karaktärer eller jämföra presenterad information i sakprosatexter.

Med reflektera och utvärdera, slutligen, menas att läsaren bygger vidare på sina tidigare kunskaper och hämtar tankegångar och perspektiv utanför texten för att exempelvis sätta det lästa i ett sammanhang och utvärdera representationer eller argument som används i texten. I någon mening förutsätter detta slags processande att läsaren ställer sig utanför texten och analyserar den som produkt. Hon förväntas exempelvis kunna analysera och kommentera språk- och stilval i texter, texters genretillhörighet liksom att kunna identifiera motiv för handlingar och representationer i berättelser.

De här tre läsprocesserna har en hel del gemensamt med de fem faser i den litterära läsprocessen som Langer (2011) beskriver och som många svensklärare är bekanta med. Man kan ur detta perspektiv se PISA:s läsprocessbegrepp dels som mer anpassade även till annan typ av läsning än den skönlitterära, och dels som mer anpassad till en traditionell kunskapstaxonomi (jfr Anderson & Krathwohl, 2001) i det att den går från enkla till mera sammansatta processer. Denna beskrivning av olika läsprocesser utgör därför ett användbart verktyg för den lärare som vill planera för en allsidig textbearbetning i klassrummet och därigenom ett allsidigt stöd till elevers läsutveckling.

LÄSSAMMANHANG

Den tredje dimensionen handlar om lässammanhanget och blir för PISA delvis ett sätt att definiera olika texter: texter för *privat* respektive *offentligt* bruk samt texter i *utbildnings-* respektive *yrkessammanhang*. Att dessa sammanhang överlappar varandra är givet. Texter som skrivs för eller antas fylla privata eller personliga syften kan läsas exempelvis i utbildnings- eller yrkessammanhang och där fylla en ny funktion. Detta gör det svårt att använda texttyper för att definiera lässammanhang i provsituationer likt PISA.

Övergripande måste man givetvis anta att alla texter där först och främst läses i ett *provsammanhang*, vilket för med sig speciella betingelser för meningsskapandet. När det gäller klassrumsbedömning av elevers läsning kan det dock vara angeläget att fundera över i vad mån en läsares uppmärksamhet inför och förmåga att tolka exempelvis privat eller personlig kommunikation kan antas representera samma läsares förmåga att läsa offentlig kommunikation. Det finns till exempel skäl att anta att romanläsning involverar andra processer om den läses självvalt på fritiden än om den läses som en del av en skoluppgift som ska redovisas och bedömas (Olin-Scheller, 2006).

HUR FÖRHÅLLER SIG KURSPLANENS OCH PISA:S DEFINITIONER AV LÄSFÖRMÅGA TILL VARANDRA?

Som lärare bör man naturligtvis fundera noggrant över hur de olika definitionerna av läsförmåga som ges i kursplanen respektive i PISA förhåller sig till varandra. Skolverket har i några olika rapporter angett att det skulle vara god överensstämmelse mellan den läsförmåga som prövas i PISA och den som definieras av kursplanen i svenska (se t.ex. Skolverket, 2010, s. 25ff). Denna slutsats bygger emellertid inte på någon djupare analys av relationen och bör betraktas med viss skepsis. Ett skäl till detta är att PISA definierar och prövar läsförmåga som en ämnesövergripande kompetens, medan kursplanen definierar det som en i huvudsak ämnesspecifik kompetens. Det visar sig exempelvis i att PISA lägger större vikt än den svenska kursplanen vid förmågan att läsa icke-kontinuerliga texter, som exempelvis diagram och tabeller. På motsvarande sätt finns det i kursplanen en större betoning av skönlitteraturens betydelse och av förmågan att läsa skönlitterära texter och att sätta in dem i ett historiskt och kulturellt sammanhang, något som av förklarliga skäl är svårt att pröva i en internationellt jämförande studie som PISA. Jämför man läsförståelsedelen i det nationella provet i svenska med PISA-provet kan man också se att betoningen av olika läsprocesser

skiljer sig åt och att sättet man prövar läsförmågan på följer olika typer av logik (Tengberg, 2014).

Detta gör som sagt inte PISA:s ramverk ovidkommande för svenska förhållanden. Tvärtom är parallellerna många och i flera fall bidrar PISA:s ramverk till att förtydliga hur komplext fenomenet läsförmåga är och vilka olika hänsyn som behöver tas i samband med bedömning. Man bör därför se det som ett komplement till kursplanen och ett stöd i att konkretisera vad läsförmåga faktiskt handlar om. Två av de aspekter som delvis lyfts i PISA:s ramverk, men som förtjänar att belysas ytterligare, är dels betydelsen av textens uppbyggnad och karaktär, dels betydelsen av de klassrumsaktiviteter och syften som omgärdar läsningen i skolan. Detta diskuteras i de två följande avsnitten.

Textens betydelse

Textens uppbyggnad och karaktär har förstås en avgörande betydelse för möjligheten till förståelse och meningsskapande i läsningen. Detta får i sin tur en omedelbar inverkan på bedömningen av elevers läsförmåga, eftersom en lärare behöver samla ingående information om vilka texttyper och textstrukturer som skapar särskilda problem för just hennes elever. Även om skolans »traditionella textgenrer« (berättande, beskrivande och argumenterande texter) behållit en viktig plats inom svenskämnet har diversiteten av texter och uttrycksformer ökat avsevärt.

I dag förväntas de texter som läses i skolan inte bara representera en större social och kulturell mångfald än tidigare, dessutom har själva textbegreppet vidgats, framför allt i takt med samhällets digitalisering. Multimediala uttrycksformer med nya sorters verktyg för interaktion och meningsskapande ingår på ett naturligt sätt i elevers läsande vardag och ska också ingå i undervisningen. Detta komplicerar analysen av hur läsförmåga ska förstås som fenomen, liksom på vilket sätt den ska prövas och bedömas i skolan och därigenom bilda underlag för lärarens fortsatta undervisning.

Forskning visar att förhållandet mellan läsning av olika sorters texter är komplicerat. En grundläggande skillnad brukar framhållas mellan läsning av skönlitteratur och läsning av sakprosa. Medan det ofta finns en återkommande struktur i berättelser, som elever tidigt lär sig känna igen, tenderar sakprosatexter att integrera flera olika strukturer, något som elever som regel är betydligt mindre förtrogna med och vilket därför gör sakprosaläsningen till en större utmaning (Pearson & Hamm, 2005; Williams, 2011). Samtidigt visar studier att duktiga läsare ofta är duktiga på att läsa såväl skönlitteratur som sakprosa, medan de mindre duktiga läsarna har problem med båda texttyperna. I en amerikansk studie kunde man bland annat visa att den fördel som elever har vid läsning av skönlitteratur uppträder på global textnivå (dvs. för läsförståelseuppgifter som kräver sammanföring och tolkning av större textpartier), medan ingen skillnad observerades på lokal textnivå (dvs. för uppgifter som kräver sökläsning och där svaret återfinns explicit på ett enskilt ställe i texten; Best et al., 2006). Studien visade också att skillnader i bakgrundskunskaper hade större betydelse vid läsning av sakprosatext än vid läsning av skönlitteratur. Detta är intressant och får betydelse i synnerhet för formativ bedömning av elevers läsförmåga. Resultatet antyder nämligen att nivån på elevers läsförståelse kan komma att variera i högre grad för läsning av sakprosa än för läsning av skönlitteratur, eftersom förståelse av sakprosatext är mer beroende av läsarens förkunskaper. För att få en allsidig bild och kunna dra generaliserbara slutsatser (se kapitel 10) om elevens läsförmåga behöver en lärare därför pröva elevens läsning på ett flertal olika texter och i olika sammanhang. En annan implikation är att det blir extra viktigt att träna elever i att dra nytta av sina bakgrundskunskaper under läsningens gång.

När det gäller digital läsning (läsning på skärm) och multimedial läsning (läsning som kombinerar bild och text eller text och navigeringsverktyg) vet vi ännu ganska lite om hur detta slags meningsskapande skiljer sig från läsning av tryckt text på papper. Undervisning och bedömning av digital och multimodal läsning

kan också utgå från de tre läsprocesserna som beskrivits ovan. Men att lokalisera eller att tolka information på nätsidor, eller att reflektera över och utvärdera dess utformning, ställer likväl annorlunda krav på läsaren. Från jämförande forskning vet vi dock att läsförståelsen försämras signifikant när elever läser en text som en pdf på skärm jämfört med när samma text läses på papper (Mangen et al., 2013; Wästlund et al., 2005). Orsaken till detta är inte helt känd, men forskare har bland annat föreslagit att läsarens överblick över texten som en helhet försämras när hon måste skrolla fram och tillbaka i den. Detta får betydelse eftersom förståelse bygger på minne av texten och detta minne är till stor del visuellt organiserat på så sätt att läsare minns och förstår text delvis genom att komma ihåg var på sidan en viss typ av information finns.

Det är i sammanhanget också intressant att notera att experimentella studier har visat att lärare som ska bedöma uppsatser har svårare för att komma ihåg innehåll i texterna och uppger att förståelse och bedömning försvåras när de läser uppsatserna på skärm jämfört med när de läser dem utskrivna på papper (Johnson & Nádas, 2009). Huruvida dessa skillnader handlar om en generationsvis successivt övergående ovana vid skärmen som medium eller om det faktiskt handlar om biologiska, kognitiva begränsningar hos människan kan vi inte svara på ännu. Men den lärare som samlar information om sina elevers läsförmåga bör alltså ta hänsyn till mediets betydelse. Detta kan vara särskilt viktigt på skolor som satsar på egna datorer eller läsplattor till varje elev och där mycket läsning blir just skärmbaserad.

Klassrumsaktiviteter och syften – läsningens situerade natur

Den kontinuerliga bedömning i klassrummet som görs av lärare, speciallärare eller specialpedagog syftar givetvis till att utvärdera elevers kunskapsutveckling i förhållande till ämnets kunskapskrav. Till skillnad från standardiserade tester, som förutom att vara vägledande vid betygsättning även syftar till utvärdering av kommuners

och skolors prestationsnivåer, används klassrumsbedömning också för att anpassa undervisningen så att den ger bästa möjliga förutsättningar för att utmana och stödja elevers kunskapsutveckling där de för närvarande befinner sig. Det gör att klassrumsbedömning blir avhängig de specifika klassrumsaktiviteter i vilka läsningen försiggår och består ofta i att läraren observerar elevernas deltagande i aktiviteter som bygger på läsning snarare än att konstruera specifika provuppgifter för att *mäta* elevernas läsning (Leslie & Caldwell, 2009). Elevernas progression prövas således i samband med textsamtal i grupp eller i helklass, genom löpande loggboksskrivande under läsningen exempelvis av en roman eller genom längre enskilda skrivuppgifter (ibland rapporter eller recensioner).

Alla vet att en lärare inte kan tränga in i huvudet på eleven för att observera den enskilda läsprocessen. I stället måste läsförmåga och läsoplevelser medieras i tal eller skrift. En viktig aspekt av bedömningen blir därför att reflektera över i vad mån förmågan att hitta och tolka information i texter eller förmågan att utvärdera texter samtidigt är uttryck för elevens förmåga att *verbalisera* läsprocessen. Låt säga att vi har en elev som är en duktig skribent och som visar god förmåga att ge uttryck för nyanserade tolkningar av texter, och en elev som har svårare för att uttrycka sig i skrift. Det är då inte självklart att den senare saknar förmåga att tolka texten på samma vis som den första eleven. Ändå ger det konkreta underlaget för bedömning oss skäl att tro att dessa båda elevers läsförmåga skiljer sig åt på ett signifikant vis.

Från studier av läsförståelseprov vet vi att uppgiftens utformning påverkar elever på olika sätt. Exempelvis tenderar uppgifter som innebär att elever ska lämna kortare eller längre skriftliga svar att gynna flickor framför pojkar, något som tros hänga samman med flickors generellt högre prestationsnivåer i språkfärdighet (Lafontaine & Monseur, 2006; Lissitz & Hou, 2012). Lafontaine och Monseur visar med sin studie att skillnaden i läsprestation mellan flickor och pojkar minskar när proven innehåller en högre andel flervalsfrågor. Även om detta förstås handlar om statistiska

skillnader, som inte kan antas gälla i varje enskild elevgrupp, så blir det viktigt att läraren i klassrummet uppmärksammar skrivandets betydelse för att uttrycka läsförmåga och ger förutsättningar även för andra sätt att visa den.

Sett till bedömningens konstruktvaliditet kan man argumentera för att skrivförmågan i så liten grad som möjligt borde få inverka på bedömningen när det är läsförmåga vi är specifikt intresserade av. En generell mätteoretisk princip säger att om vi vill få detaljerad information om elevers kunskaper inom ett givet område bör deras prestationer inte bero av faktorer som ligger utanför det vi vill mäta (Koretz, 2008). Å andra sidan kan man med hänsyn till frågan om *ekologisk validitet* hävda att om utbildningen i svenskämnet ytterst syftar till att förbereda elever för att delta i samhällsliv och kulturliv genom att kommunicera kunskaper och erfarenheter med människor, så existerar en sådan sak som läsförmåga inte oberoende av individens förmåga att uttrycka denna läsförmåga antingen i tal eller i skrift. Därmed ska man kanske förstå kontextualiserade prövningar av läsförmågan – det vill säga prövningar som ingår i en pågående dialog om texter, en dialog som eleverna uppfattar som meningsfull och stöttande – som ett mer autentiskt sätt att mäta och bedöma.

När det gäller underlag som samlas in som bevis på elevers läsförmåga bör man också ta hänsyn till under vilka omständigheter elever formulerar sina svar. I provsituationer, som exempelvis det nationella delprovet i läsning, har eleverna hela tiden tillgång till den text de ska svara på och kan gå tillbaka till den för att finna stöd för olika tolkningar. Å andra sidan är tidspressen en faktor som begränsar många elever och ger liten tid för någon egentlig eftertanke. I skrivuppgifter som bearbetas i klassrummet och/eller hemma över en längre period finns både tillgång till texten och gott om tid. I helklassamtalet däremot får eleven sällan möjlighet att gå tillbaka till texten för att finna stöd för sina svar (textinnehållet måste i stället huvudsakligen minnas) och tidspressen är klart påtaglig. I gengäld rymmer samtalet en responsiv karaktär i det att

man inom dialogen kan resonera om olika tolkningar, upprepa en fråga, lägga till ett andra argument om det första inte övertygar etc. Den möjligheten finns inte i den sortens upprepad envägskommunikation som provkonstruktion och provtagande utgör. Studier har också visat att mått på läsförståelse blir signifikant högre när elever har tillgång till de texter de prövas på medan de besvarar frågorna, vilket implicerar att bedömningsunderlag från situationer där eleverna inte haft texten tillgänglig tenderar att underskatta elevernas läsförmåga (Leslie & Caldwell, 2009). Alternativt kan man säga att sådana situationer prövar en annan aspekt av läsförmågan, det vill säga en aspekt som i hög grad beror av vad eleven *minns* av texten, inte enbart av vad hon eller han *förstår*.

Avslutande kommunikation – bedömning av läsförmåga

Mot bakgrund av den mångsidiga natur som läsförmågan utgör bör man dra slutsatsen att valida bedömningar av elevers läsförmåga måste grunda sig på bevis som belyser dessa många olika aspekter av läsandet. När vi bedömer, det vill säga när vi avgör vad en elev kan i förhållande till givna kunskapskrav, gör vi en konkretisering av kursplanen och låter enskilda uppgifter representera kursplanens abstrakta konstruktdefinition. För att kunna dra slutsatser om vad elever kan, baserat på deras prestationer på enskilda uppgifter, måste vi därför dels se till att de olika uppgifterna ger en så god representation som möjligt av den läsförmåga som kursplanen stipulerar, dels förvissa oss om att eleverna har fått flera chanser att visa sin kunskap så att slutsatsen från bedömningen inte blir beroende av dagsform, den valda texten, vilken diskussionsgrupp man hamnade i den dagen etc.

William (2010) påpekar att *hur* man mäter något ytterst är en fråga om vad man mäter. Om vi väljer att pröva elevers läsförmåga genom att låta dem formulera en längre text där de beskriver exempelvis sin analys av en novell eller sina kontinuerliga iakttagelser under läsningen av en roman, så är detta inte enbart ett *sätt* att pröva

deras läsförmåga, utan också en konkretisering av *vad* det är som prövas. Om vi i stället skulle be eleverna att muntligen beskriva sin analys av texten skulle det med all säkerhet innebära att andra aspekter av läsprocessen belystes. Vi skulle få syn på andra sidor av hur eleven tagit sig an texten och vilka slutsatser hon kommit fram till baserat på sin läsning. Valet av metod får följaktligen betydelse för vilket ämnesinnehåll som synliggörs. Detta blir alltså viktigt att tänka på när man som lärare funderar över hur elevernas progression i läsningen ska bedömas i samband med ett visst arbetsområde eller inför betygsättning.

En annan aspekt som avslutningsvis förtjänar att nämnas är att när det gäller läsning och bearbetning av texter, så är det minst lika viktigt att undersöka (och pröva) elevers förmåga att *ställa* frågor, inte bara att besvara dem. Av tradition är vi oerhört vana vid att textbearbetning i skolan innebär att läraren ställer frågor om texten och elever svarar för att visa vad de förstått. Men eftersom en väsentlig del av såväl litterär analys som av argumentationsanalys och faktaanalys består i att formulera prövande, kritiska frågor till en text är det avgörande att läraren i undervisning och bedömning också uppmärksammar vikten av att få elever att ställa frågor. En rad studier visar att förmågan att generera frågor, konstruktiva, öppna och prövande frågor, under läsningens gång korrelerar starkt med olika mått på hög läsförmåga (se t.ex. Janssen, 2002). Att träna elever i att formulera frågor om sådant de upplever som förvånande, svårt, konstigt, upprörande osv. i texter är därför en viktig del i läsundervisningen, en av de viktigaste lässtrategierna faktiskt, för att återvända till ett begrepp som lyftes inledningsvis i kapitlet. Eftersom ett successivt frågeställande till texten också är att betrakta som en aspekt av god läsförmåga bör den också uppmärksammas i lärarens bedömning, det vill säga den bör användas för att värdera elevens kunskaper, för att återkoppla kunskapsbedömning till eleven och för att planera den fortsatta undervisningen.

Det här kapitlet har handlat om vad läsförmåga är för något, eller åtminstone hur det kan beskrivas som utgångspunkt för

undervisning och bedömning i svenska i årskurserna 7–9. Jag har medvetet hållit mig till en ganska allmän definition i syfte att täcka in hela konstruktet. Av det skälet har jag inte haft utrymme för några mer ingående beskrivningar av de komponenter man kan förvänta sig ska ingå exempelvis i bedömning av en litterär analys, en argumentationsanalys eller läsning av annan sakprosatext. I bokens empiriska del finns dock två kapitel där några av dessa aspekter konkretiseras och sätts in i ett klassrumssammanhang.

Ett av kapitlen, kapitel 8, handlar om hur elevers skriftliga analyser med successiv återkoppling från läraren kan användas för att bedöma och utveckla elevers romanläsning. Det andra kapitlet, kapitel 9, handlar om hur en längre skrivuppgift läggs till grund för bedömning av elevers novellläsning.

Referenser

- Afflerbach, P.P. & Cho, B. (2011). The classroom assessment of reading. I M.L. Kamil, P.D. Pearson, E. Birr Moje & P.P. Afflerbach (red.), *Handbook of reading research, vol. IV*. New York: Routledge.
- Anderson, L.W. & Krathwohl, D.R. (2001). *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman.
- Best, R., Ozuru, Y., Floyd, R.G. & McNamara, D.S. (2006). Children's text comprehension: effects of genre, knowledge, and text cohesion. I S. Barab, K.E. Hay & D.T. Hickey (red.), *ICLS '06 Proceedings of the 7th international conference on learning sciences* (s. 37-42). Mahwah, NJ: Lawrence Erlbaum Ass.
- Duke N.K. (2005). Comprehension of what for what: Comprehension as a nonunitary construct. I S.G. Paris & S.A. Stahl (red.), *Children's reading comprehension and assessment*. Mahwah, NY: Lawrence Erlbaum Ass.
- Janssen, T. (2002). Instruction in self-questioning as a literary reading strategy: An exploration o empirical research. *L1 – Educational Studies in Language and Literature*, 2, 95-120.
- Johnson, M. & Nádas, R. (2009). Marginalised behaviour: Digital annotations, spatial encoding and the implications for reading comprehension. *Learning, Media and Technology*, 34(4), 323-336.
- Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard U.P.
- Lafontaine, D. & Monseur, C. (2006). *Impact of Test Characteristics on Gender Equity Indicators in the Assessment of Reading Comprehension*. Liège: Universitetet i Liège.
- Langer, J.A. (2011). *Envisioning literature: Literary understanding and literature instruction* (2 uppl.). New York: Teachers College Press.
- Leslie, L. & Caldwell, J. (2009). Formal and informal measures of reading comprehension. I S.E. Israel & G.G. Duffy (red.), *Handbook of research on reading comprehension*. New York, NY: Routledge.
- Lissitz, R.W. & Hou, X. (2012). The contribution of constructed response items to large scale assessment: Measuring and understanding their impact. *Journal of Applied Testing Technology*, 13(3),

- 1–50. Hämtad 2013-05-27, från <http://www.testpublishers.org/assets/volume%2013%20issue3%20constructed%20response.pdf>.
- Mangen, A., Walgermo, B.R. & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research*, 58, 61–68.
- Mullis, I.V.S., Martin, M.O., Kennedy, A.M., Trong, K.L. & Sainsbury, M. (2009). *PIRLS 2011. Assessment framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.
- OECD (2009). *PISA 2009. Assessment framework: Key competencies in reading, mathematics and science*. <http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/pisa2009assessmentframework-keycompetenciesinreadingmathematicsandscience.htm> Hämtad 2013-05-23.
- Olin-Scheller, C. (2006). *Mellan Dante och Big Brother: en studie om gymnasieelevers textvärldar* (Doktorsavhandling). Karlstad: Karlstad University Press.
- Pearson, P.D. & Hamm, D.N. (2005). The assessment of reading comprehension: A review of practices: Past, present, and future. I S.G. Paris & S.A. Stahl (red.), *Children's reading comprehension and assessment* (s. 13–70). Mahwah, New Jersey: Lawrence Erlbaum Ass.
- Skolverket (2010). *Rustad att möta framtiden? PISA 2009 om 15-åringars läsförståelse och kunskaper i matematik och naturvetenskap*. Rapport 352. Stockholm: Fritzes.
- Skolverket (2011a). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011*. Stockholm: Fritzes.
- Skolverket (2011b). *Kommentarmaterial till kursplanen i svenska*. Stockholm: Fritzes.
- Tengberg, M. (2014). Att pröva elevers läsförmåga i årskurs nio: en analys av uppgiftskonstruktioner i det nationella provet i svenska, delprov A: Att läsa och förstå. *Utbildning & Demokrati*, 23(2), 109–132.
- Wiliam, D. (2010). What counts as evidence of educational achievement? The role of constructs in the pursuit of equity in assessment. *Review of Research in Education* 34, 254–284.
- Williams, J. (2011). Teaching comprehension of narrative and expository text. I R. E. O'Connor & P. F. Vadasy (red.), *Handbook of reading interventions* (s. 254–278). New York/London: The Guildford Press.

Wästlund, E., Reinikka, H., Norlander, T. & Archer, T. (2005). Effects of VDT and paper presentation on consumption and production of information: Psychological and physiological factors. *Computers in Human Behavior*, 21, 377-394.

5. Uppgifter

GUSTAF SKAR

Att vara elev innebär till stor del att vara någon som löser uppgifter. Även om lärare inte alltid tänker på sin verksamhet i sådana termer relaterar såväl gruppdiskussioner, att svara på frågor vid genomgång av läxan som att skriva något i läsloggen, till att lösa uppgifter. Till skillnad från den vardagsspråkliga betydelsen av att lösa en uppgift måste en elev dock inte nå ett visst resultat för att en uppgift ska vara »löst«; som lärare är vi lika intresserade av sådant som indikerar att en elev *inte* kan svara på vissa frågor. Men trots att både elever och lärare ägnar mycket tid åt uppgifter saknar vi ibland ett gemensamt språk för att tala om hur uppgifter konstrueras och vad vi kan använda dem till.

I det här kapitlet behandlas uppgifter som förekommer i svenskundervisning. Vi kommer att ta upp frågan om vilken information som olika uppgiftstyper kan leda till, men också titta närmare på läsuppgifter, skrivuppgifter och integrerade uppgifter. Först kommer vi dock att undersöka vad en uppgift egentligen är och hur uppgifter relaterar till bedömningars syfte.

Uppgifter, uppgiftsform och uppgiftskvalitet

Utgångspunkten i det här kapitlet är att alla handlingar en lärare ber eleverna att genomföra *i syfte att skaffa underlag i en kunskapsbe-*

dömningsprocess kan kallas uppgifter. Det innebär dels att uppgifter definieras med utgångspunkt i sin funktion, dels att uppgifter är ett begrepp med bredd. Allt är dock inte uppgifter i den bemärkelse som avses här; vissa handlingsiniterande uppmaningar leder inte till underlag i en bedömningsprocess, till exempel att be eleverna ställa upp stolarna efter avslutad lektion och liknande. I det här avsnittet ska vi se närmare på hur uppgifter och dess yttre form kan förstås i relation till bedömnings syften.

Som tidigare konstaterats genomförs klassrumsbedömning med en rad olika syften (se kapitel 1). Hill och McNamara (2012) erbjuder en definition av klassrumsbedömning som tar hänsyn till en lärares skiftande behov av att göra många olika typer av bedömningar. Definitionen belyser också komplexiteten i lärares bedömningsarbete. Enligt dem är klassrumsbedömning:

... varje värdering lärare (och/eller elever) gör av kvaliteten i elevarbete (eller flera elevers arbete), och lärares (eller elevers) användning av information från denna värdering i undervisnings-, återkopplings-, rapporterings-, hanterings- eller socialiserings-syften (Hill & McNamara, 2012, s. 396, författarens övers.).

Det finns flera aspekter av denna definition som är värda att kommentera. För det första kan vi se att vi kan urskilja två syften som vi känner igen som primära, dels att bedöma för att skaffa underlag inför den kommande undervisningen, dels att bedöma för att skaffa underlag inför rapportering (t.ex. betygsättning). För det andra kan vi se att lärare bedömer elevers arbete också i andra syften än att dra slutsatser om deras kunskaper. Den typen av bedömning (som t.ex. går ut på att förstärka visst beteende) lämnar vi dock därhän, även om den också bygger på uppgifter. För det tredje kan vi se att formerna för de handlingar som inbegrips när vi talar om bedömning varierar när vi tar ett brett grepp. Eftersom bedömning avser varje värdering av elevers arbete relaterat till undervisning och/eller rapportering, måste vi inkludera såväl markerade och

omarkerade bedömningar som formella och informella bedömningar. Den inre och yttre utformningen på en uppgift kan förstås i relation till de tre dimensionerna, *grad av »markeradhet«, formalitet och syfte.*

En uppgift kan vara markerad eller omarkerad. En markerad uppgift innebär att elever känner till att den handling de utför är en del av en bedömningssituation. Ett exempel är ett skriftligt läxförhör. En uppgift är omarkerad när elever inte vet eller inte är säkra på att den handling som utförs är en del av en bedömningssituation. En omarkerad uppgift kan till exempel äga rum när en lärare ber elever läsa högt ur en lärobok och utan att ha meddelat det bedömer deras avkodningsförmåga och läshastighet, vilket kan ligga till grund för undervisningsplanering. Graden av markering är viktig för elever och särskilt när uppgiftsresultatet ska ligga till grund för viktiga beslut (t.ex. betyg). Många lärare känner säkert igen frågan »kommer du att bedöma detta?» som till exempel kan ställas när elever i sin loggbok ombeds »reflektera« eller »diskutera« en läst text. Frågan är mer än befogad. För det första finns etiska och rättssäkerhetsmässiga hänsyn att ta – i många fall har elever rätt att få veta att deras handlingar är föremål för bedömning. För det andra kommer eleverna, precis som vilka arbetstagare som helst skulle ha gjort, att angripa uppgifterna på olika sätt beroende på om mycket eller lite står på spel. I en reflektionstext som inte bedöms kan eleverna tillåta sig att pröva lite olika grepp, men i en text som bedöms vill de ofta agera på ett sätt som resulterar i en för dem gynnsam bedömning. I fall där bedömningen får långtgående konsekvenser måste därför eleven tillåtas att agera informerat.

En uppgift kan till sitt yttre också vara formell eller informell. Informella uppgifter förekommer i ganska stor omfattning i klassrumsbedömning. Vi gör en »allmän koll« av hur eleverna förstått läxan genom att ställa frågor till mer eller mindre slumpvis valda elever. Vi har kanske en pop quiz som får lösas lite hur som helst.

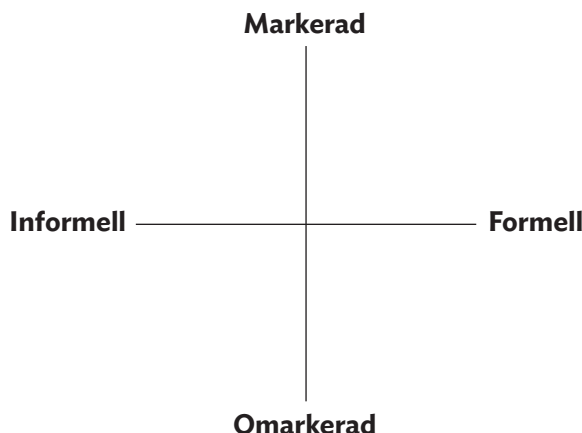
En formell uppgift omgärdas däremot av vissa ramar och inte

sällan är den en del av ett prov.¹ Ett prov kan bestå av flera delprov (så som Högskoleprovet), som i sin tur består av flera uppgifter. I storskalig testning föredrar man prov med många uppgifter som fokuserar på ett smalare område, eftersom det ger provtagaren flera möjligheter att visa sin kunskap på, vilket är gynnsamt i termer av reliabilitet (jfr kapitel 10). Exempel på prov som kan ha riklig uppgiftsmängd i svenskundervisning är grammatikprov.

Många gånger kanske vi omtalar bedömningar med olika formalitetsgrad, som prov. Här är begreppet dock reserverat för formella, markerade bedömningar som genomförs under specificerade omständigheter (jfr Cizek, 2012). Detta beror på att graden av formalitet inverkar på vilken typ av information som bedömningen möjliggör. Ett prov måste alltid administreras, vilket innebär att det måste hanteras (förvaras, delas ut, samlas in) och att provtagarna i förekommande fall måste övervakas. För till exempel det nationella provet i svenska finns angivna riktlinjer för vilken tidsmängd en provtagare (elev) har på sig att skriva provet och vilka hjälpmedel som ska finnas tillgängliga. Dessa riktlinjer relaterar till olika grader av standardisering (jfr kapitel 2). Ju mer en provlösning har standardiserats, desto bättre är chanserna att generalisera resultatet till andra liknande provsituationer och att jämföra olika provresultat.

I Figur 5.1 syns fyra fält som kan användas för att typologisera olika uppgifter. Även om en dylik typologisering inte tar sikte på uppgiftens innehåll kan det ändå vara klokt att göra en sådan ibland. Om jag som lärare främst använder mig av information från omarkerade, informella uppgifter kommer jag att ha ett kvalitativt annorlunda underlag än om jag också hade haft tillgång till information från markerade och formella uppgifter. Vilken typ av informationsunderlag som är nödvändigt är avhängigt bedöm-

¹ Wikström (2013, s. 71f) kallar uppgifter som ingår i prov för »frågor«, vilket skulle kunna betraktas som den gängse termen i bedömningslitteraturen. I detta kapitel är dock syftet att synliggöra uppgiften som ett generiskt redskap som lärare använder sig av i olika sammanhang. Dessutom är inte alla uppgifter syntaktiska frågor.



Figur 5.1. Två av tre dimensioner relaterade till uppgifters form.

ningssyftet. För att uppnå ett särskilt syfte med en bedömning är en uppgifts innehållsliga utformning och utförande avgörande. Utformningen kan därför sägas möjliggöra eller omöjliggöra vissa bedömningssyften. Ett exempel är bedömning i diagnostiserande syfte som möjliggörs i högre grad om bedömningsmatriser utformas så att det går att göra tämligen nyanserade bedömningar av till exempel olika dimensioner av skriv- eller läsförmåga (se kapitel 3 samt kapitel 4). Ett annat exempel på hur uppgiftens utformning bidrar till att möjliggöra ett bedömningssyfte är när bedömningsresultatet ska användas för att jämföra provtagare (bl.a. så vill vi ju att betyget A på ett prov ska betyda samma sak oavsett vilken elev som får det). Då spelar till exempel själva instruktionen stor roll; vi vill vara förvissade om att alla elever har tolkat instruktionen på samma sätt, annars omöjliggörs jämförelsen!

Bedömningsprocesser ser olika ut och kräver i regel olika typer av uppgifter. Med det menas att en uppgift i sig varken kommer att vara bra eller dålig. I förhållande till validitet (se kapitel 10) är det uppgiftens yttre och inre utformning i relation till bedömningssyftet (dvs. det vi vill påstå om elever och det vi vill göra med våra påståenden) som är av intresse. Av detta följer att en »fyller-i-övning« inte alltid är avgjort dålig, lika lite som en komplex hemtentamen

alltid är bra. En bra uppgift är den som gör jobbet. I resterande del av detta kapitel ska vi definiera olika uppgifter och resonera om deras kvalitet i förhållande till olika bedömningsavsikter.

Vad innehåller en uppgift?

Gemensamt för alla uppgifter, vare sig de är markerade/omarkerade eller formella/informella, är att de i huvudsak har tre beståndsdelar (Haladyna & Rodriguez, 2013, s. 44). Dessa beståndsdelar kommer att ta sig olika uttryck beroende på uppgiftens karaktär och bedömningens syfte. De tre beståndsdelarna är:

1. *En fråga/en uppmaning till provtagare/elev*, till exempel »När verkade Voltaire?» eller »Skriv en argumenterande uppsats om varför skoldagen bör förkortas.«
2. *Villkor för uppgiftslösning*, som i klassrumssammanhang kan vara mer eller mindre explicitgjorda. Villkoren specificerar hur uppgiften ska lösas och kan till exempel innebära att en grupp elever får lektionen på sig att skriva den argumenterande uppsatsen efter eget tycke. Ibland är villkoren för uppgiftslösning inte allmänt kända eller efterlevda, även om de är förutsatta. Det gäller till exempel undervisning med små barn, som inte alltid svarar först efter att de har nominerats som talare. När en elev sålunda svarar utan att »räcka upp handen« efterlevs inte uppgiftslösningens villkor.
3. *Bedömningsregler*. Traditionellt har både lärare och provkonstruktörer talat om *rättning* för att benämna aktiviteten att värdera, poängsätta eller på annat sätt sedermera meddela resultat efter genomförd uppgift (t.ex. så kallar Haladyna & Rodriguez, 2013, denna beståndsdel för *scoring procedure*). Rättning är dock klokt att reservera för mer eller mindre okomplicerad »klerikal« bedömning (ibid.) av elevgensvar, som att ange 1 poäng när eleven valt rätt alternativ på en flervalsfråga och 0 poäng när eleven valt fel alternativ.

Bedömningsregler kan i klassrumssammanhang vara uttalade eller på förhand uppställda. Exempel på uttalade regler får vi när vi återvänder till helklassgenomgången av läxan. Läraren har antagligen inte verbaliserat olika nivåer av kunnande relaterat till läxan och har inte heller något explicitgjort system för att förhålla de svar som inkommer till uppsättningar av kriterier. Troligare är att läraren efter en stunds genomgång har en känsla för om klassen gjort läxan eller inte. Inför mer avgörande bedömningar formuleras dock ofta bedömningsregler på förhand. Både i storskaliga testsammanhang och i klassrummet bestäms vilka kriterier till exempel en uppsats måste möta för att vara godkänd (jfr kapitel 3).

Inte sällan anges dessutom olika nivåer av godkänt, något som får till följd att kriterier formuleras även för dessa. Själva regeln är därefter ganska enkel; om vi kan identifiera till exempel en lämplig stilnivå i texten möter texten kriteriet om lämplig stilnivå, vilket kan rendera ett poäng eller matchning mot en verbal deskriptor. Till skillnad från när vi *rättar* något kan vi dock inte se stilnivån med blotta ögat – slutsatsen om en texts stilnivå (som vi mycket väl kan argumentera gott för och visa manifestationer av) bygger med andra ord på en tolkning.

I nästa avsnitt ska vi närmare granska uppgifters beståndsdelar och relatera dem till olika bedömningssyften.

Grundläggande klassifikation av uppgifter

Hur uppgifter kan klassificeras är en öppen fråga. Mest avgörande är vem som gör klassifikationen. Det finns exempel på skrivforskare som intresserat sig för vilket skrivande som bedöms givet en viss uppgift (Hillocks, 2002), liksom det finns läsforskare och litteraturvetare som undersöker vilken typ av läsning olika uppgifter

innebär (t.ex. Tengberg, 2014). Sådana (innehålls)klassifikationer är avgörande för att bättre få klart för sig mot vilka kunskapsdomäner olika uppgifter förhåller sig. Det finns också kunskapsteoretiska klassificeringar av uppgifter som till exempel relaterar de uppgifter/göromål elever ägnar sig åt i skolan till den typ av handlingar och verksamheter som eleven ska ägna sig åt utanför skolan. Det finns till exempel gott om teoretiker som ur ett kulturhistoriskt eller sociokulturellt perspektiv analyserat vad det innebär att vara elev och hur detta är något annat än det som ett provpoäng är tänkt att indikera (se Greeno, 1997; Lave & Wenger, 1991). Som elev lär man sig mer och annat än vad svar på uppgifter förmår att visa, till exempel att sitta på ett speciellt sätt, att räcka upp handen och att vara tyst när det är prov.

Man kan också ta ett bedömningsteoretiskt perspektiv på olika uppgiftstyper. Ett sådant bidrar till förståelse av att frågan om vad som ska bedömas aldrig kan frikopplas från frågan om hur en uppgift bedöms; ur ett bedömningsteoretiskt perspektiv är vad- och hur-frågorna alltid sammankopplade.

Grundläggande för den som väljer uppgiftstyp är graden av direkthet. I många samtida texter förordas »autentiska« och »direkta« uppgifter och ett begrepp som är särskilt populärt är *performance assessment*. Bakom autenticitetssträvanden ligger en önskan om att låta uppgifter vara så med verkligheten överensstämmande som möjligt för att steget från vad en provtagare gör till påstående om vad hon eller han kan göra i en måldomän ska bli minimalt (jfr eng. *high fidelity*). Det kan dock vara klokt att komma ihåg att bedömning i sig innebär att en handling filtreras i ljuset av de kriterier som används för att bedöma den (McNamara, 1996; Messick, 1996); vi kan nästan aldrig konstruera helt autentiska uppgifter.

Mot bakgrund av sådana autenticitetssträvanden har olika tekniker olika fördelar och kan också vara behäftade med olika problem. Att till exempel mäta skrivförmåga med flervalfrågor är fördelaktigt i det perspektivet att inverkan från en enskild bedöma-

re blir minimal, men negativ eftersom uppgiften inte innebär att elever skriver, vilket medför allvarliga problem i extrapoleringsledet (Kane, Crooks & Cohen, 1999).

Ett exempel på problematiken som varje uppgiftskonstruktör står inför är hur uppgiftens verklighetsanknytning ska viktas mot slutsatsernas säkerhet. Om elever skriver hemma kan man å ena sidan inte vara säker på vem som har skrivit, men man har å andra sidan skapat en uppgiftstyp som liknar skrivande i världen utanför uppgiftssituationen. Ett annat exempel på denna konflikt har vi kunnat se när Skolinspektionen (t.ex. 2012, 2013) efter genomförda omrättningar av de nationella proven föreslagit att faktiskt skrivande bör strykas som uppgift. Denna konflikt mellan hur och vad har av forskare kommenterats i en rapport:

Problemet med reliabilitetsbrister i bedömning av längre uppsatser har uppmärksamats av Skolinspektionen vid alla genomförda omrättningsomgångar. Skolinspektionens huvudförslag har varit att uppsatsskrivning ska avlägsnas från de nationella proven. Problemet med denna lösning är att förmåga att skriva på svenska och engelska är centrala mål i läroplanerna för både grundskola och gymnasium. Det finns inte heller någon annan känd metod för att pröva förmågan att skriva uppsats än just genom uppsatsskrivning (Gustafsson, Cliffordson, & Erickson, 2014, s. 41).

Motivet bakom Skolinspektionens förslag är alltså att hög reliabilitet medger generalisering. Som Gustafsson med kolleger (2014) påpekar är generalisering dock inte liktydigt med möjlighet till extrapolering. Dessutom kan man tänka sig att ett nationellt prov utan en skrivdel skulle förlora en del av sin legitimitet. Att prov *uppfattas* som rimliga är i teknisk validitetsmening inte avgörande för kvaliteten, men ändå av största betydelse – risken är annars stor att ingen litar på resultatet.

Huvudtyper av uppgifter

Grovt sett kan man dela in uppgifter i tre huvudtyper. Nedanstående klassifikation, liksom begrepp för undergrupper av uppgifter, bygger på Haladyna och Rodriguez (2013):

1. Uppgifter med bundna svarsalternativ (*selected response*), som rättas objektivt av människa eller dator.
2. Öppna uppgifter (*constructed response*), som rättas objektivt av människa.
3. Öppna uppgifter som bedöms genom att en produkt tolkas med utgångspunkt i kriterier.

UPPGIFTER MED BUNDNA SVARSALTERNATIV

Uppgifter med bundna svarsalternativ innebär att den som löser uppgiften väljer bland färdigformulerade alternativ. Enligt Haladyna och Rodriguez (2013, kap. 5) finns det åtminstone nio huvudkategorier av uppgiftstyper med bundna svarsalternativ och därtill en rad underkategorier. Den mest kända av uppgiftstyperna är sannolikt flervalsfrågan, som består av en stam (*stem*) och fyra alternativ. Av alternativen är ett korrekt och de tre andra felaktiga. De felaktiga alternativen kallas distraktorer.

Fördelar med flervalsfrågor och andra uppgifter med bundna svarsalternativ är att de kan mäta kunskap eller till exempel läsförståelse på ett avancerat och reliabelt sätt utan att den som löser uppgiften måste utföra andra komplexa handlingar (t.ex. att skriva). Det går till exempel att formulera uppgifter för att undersöka om elever som läst en text kan identifiera sådant som finns på textens yta och sådant som kräver att eleven gör inferenser. En annan och avgörande fördel är att bedömningen går snabbt att genomföra och är säker.

Det finns dock även några uppenbara nackdelar med bundna svarsalternativ. Sådana uppgifter medger inte bedömning av handlingar som att *resonera* eller *sammanfatta*, vilka alla kan vara

relevanta i relation till just läsning. En annan klar nackdel är att det finns en viss risk för att den som löser uppgiften svarar rätt genom gissning. Ju färre alternativ, desto högre är sannolikheten. Ytterligare en nackdel är att uppgifter av detta slag är mycket svåra att ta fram; för att förvissa oss om att tänkta innebörder i provresultaten ska vara möjliga krävs välskrivna uppgifter som i regel tar lång tid att utveckla. I storskalig testverksamhet kostar utvecklingen tusentals kronor per uppgift (Gavin Brown, personlig kommunikation, 19 mars 2014; Haladyna & Rodriguez, 2013, s. 45).

OBJEKTIVT RÄTTADE ÖPPNA UPPGIFTER

Öppna uppgifter kan enligt Haladyna och Rodriguez (2013) delas in i sådana som rättas objektivt och sådana som bedöms subjektivt. Exempel på den förra typen kan vara följande två uppgifter som följer på ett längre textavsnitt om en pojke som får en hund: 1) Vad fick pojken på gott humör? 2) Hur mycket kostar hundmaten?

Eftersom eleven här själv producerar ett svar är uppgiftstypen fri från krav på att på förhand konstruera rimliga alternativ. Detta innebär också eliminering av chansen att gissa samtidigt som en välformulerad bedömningsregel anger exakt hur ett svar ska utformas för att godkännas. Den här uppgiftstypen lämpar sig väl för tillfällen då det är nödvändigt att veta att testtagaren till exempel faktiskt kan dra vissa slutsatser eller på annat sätt kan använda läsning i något syfte.

En svårighet med uppgiftstypen kan vara att formulera bedömningsregeln. Inte sällan dyker nya oväntade svar upp, svar som också äger sin rimlighet. Om det är en bedömningsituation där mycket står på spel kan det därför vara idé att testa uppgiften på kollegor innan den sjösätts i undervisning.

ÖVRIGA ÖPPNA UPPGIFTER

Den sista uppgiftstypen är en som svensklärare är väl bekanta med. En öppen uppgift med subjektiv bedömning innebär många fördelar, särskilt i termer av autenticitet och direktet. En uppgift av detta slag kan vara mycket omfattande («skriv en uppsats på valfri plats, om valfritt ämne») eller begränsad («skriv en argumenterande text under 20 minuter») beroende på mot vilken kunskaps- eller färdighetsdomän uppgiften förhåller sig. Den här typen av uppgift kan utformas så att det är lätt att argumentera för extrapoleringsledet, vilket dock brukar inverka på generaliserbarheten (Kane et al., 1999). En potentiell svaghet med det här uppgiftsformatet är att det kräver just subjektiv bedömning (vilket, förvisso, i sig är något naturligt och verklighetstroget; texter utanför skolan bedöms i regel av människor) som kan innebära icke-jämförbara resultat. Inom skrivbedömningsfältet har man länge uppmärksammat problemet att läsare sällan är överens om hur en text ska värderas (Berge, 1996; Björnsson, 1960; Edgeworth, 1890; Jølle, 2014).

För att förstå relationen mellan en uppgift och ett resultat kan modellen nedan (Figur 5.2) vara behjälplig. Figuren illustrerar en situation med en öppen uppgift som ska bedömas subjektivt. Den som löser uppgiften tolkar instruktionen (t.ex. *skriv ett brev till rektor med x ord på y minuter och med z hjälpmedel på ett visst ställe*) och producerar sedan en lösning på den. Lösningen bedöms av en bedömare som tolkar en bedömningsanvisning för att närma sig gensvaret. För uppgifter med bundna svarsalternativ finns ingen tolkande bedömare. Vad modellen indikerar, vilket är viktigt att komma ihåg, är att det aldrig finns någon direkt relation mellan en elevs läs- eller skrivförmåga och ett provresultat – alla uppgifter är i någon mening indirekta (Messick, 1996) och resultatet är alltid en produkt av interaktion.



Figur 5.2. Relationen provtagare–uppgift–bedömning–resultat.
Bearbetning av figur i McNamara (1996).

Uppgiftstyper i svenskundervisning

LÄSNING

Även om det finns hävdvunna tekniker för att mäta läsförmåga är det en avgjort komplicerad och tämligen osäker verksamhet. Förutom att välja vilken *typ* av läsförmåga (se kapitel 4) vi är intresserade av att mäta måste vi också besluta oss för vilken eller vilka texter som ska ingå i uppgiften och om vi ska utforma uppgifterna med bundna svarsalternativ eller som öppna uppgifter. Att läsa och förstå något tar ju sig inte alltid produktiva uttryck, vilket bör innebära avvägningar i valet av format (Hughes, 2003). Vi måste också bestämma hur lång tid en provtagare ska få på sig att lösa uppgifterna. Slutligen kvarstår att bestämma hur svaren ska rättas eller bedömas.

Alderson (2000) har i en omfattande genomgång redovisat ett flertal tekniker, till exempel:

- *Cloze test*, som innebär att vart n -te (t.ex. vart femte) ord tas bort, vilket resulterar i att uppgiftslösaren ska fylla tomrummet. Till exempel: Genom att ta bort _____ femte ord mäts, enligt _____ förespråkare, förmågan att ta _____ mening ur en

text. (Genom att ta bort *valt* femte ord mäts, enligt *teknikens* förespråkare, förmågan att ta *huvudsaklig* mening ur en text.)

- *Matchningstest*, som innebär att man till exempel parar ihop rubriker med textstycken.
- *Ordningssuppgifter*, i vilka det gäller att sätta stycken i rätt ordning.

En del av dessa tekniker kan förekomma i läromedel och i externt producerade tester. För den lärare som ska använda sig av resultatet gäller det att känna sig säker på vilken typ av läsförmåga uppgiftstypen medger mätning av.

För svensklärare torde de vanligaste uppgifterna vara endera sådana med bundna svarsalternativ (konventionella flervalsfrågor eller rätt-fel-uppgifter) eller öppna frågor med både objektiv rättning och subjektiv bedömning. Som vi tidigare varit inne på finns en rad fördelar med respektive teknik, men också fallgropar. Hughes (2003, s. 154f) har formulerat ett antal goda råd för den som själv ska konstruera läsuppgifter. Några av dessa följer här:

- Formulera uppgiften så att rätt svar inte går att hitta i texten utan att förstå den. Fenomenet kan uppstå när ordalydelsen i uppgiften är nära nog identisk med den text som provtagaren ska identifiera.
- Formulera uppgiften så att den inte kan lösas med hjälp av omvärldskunskap. Ett exempel är:
»Sveriges huvudstad är: _____.« Här spelar både val av text och uppgiftsformulering in.
- Gör uppgifterna oberoende av varandra. Oberoende innebär i detta sammanhang att det inte ska vara möjligt att svara på en fråga genom att svara på en annan.

Ytterligare två råd som förekommer frekvent i litteraturen rör formulering av alternativ i samband med flervalsfrågor och utformning av bedömningsregler. Alternativ bör formuleras så att även de

felaktiga är rimliga (och i fallet med ja–nej-frågor att själva frågan är rimlig). Anledningen är att en provtagare annars kan utesluta alternativ utan att nödvändigtvis ha förstått texten. Vad gäller bedömningsregler bör man i förväg fundera mycket noggrant över hur öppna frågor ska rättas. Om provet innebär *high stakes* (dvs. får stora konsekvenser för eleven) bör man i rättvisans namn och för jämförbarhetens skull bara godkänna sådana svar som passar den på förhand fastställda rättningsmallen, alternativt rätta om samtliga svar om man justerat bedömningsregeln i efterhand.

SKRIVANDE

En bra skrivuppgift kräver att en provtagare skriver något. Om så inte är fallet testas något annat, till exempel redigeringsförmåga som när uppgiften består i att rätta en text, eller grammatiska kunskaper som när uppgiften består i att välja vilket alternativ som innehåller en korrekt formulering.

Liksom i fallet med läsförmågeuppgifter kräver skrivuppgifter noggranna övervägningar av vilket skrivande som ska mätas. Med terminologin i denna bok (se kapitel 3) kommer vi här att rikta intresset mot att undersöka elevers förmåga att utföra olika skrivhandlingar. Vi behöver då en uppgift som »uppmanar« (*prompts*) elever att skriva i ett särskilt syfte. Beroende på hur skrivförmåga har definierats kan uppmaningen komma att innehålla information om skrivandets syfte, textens tänkta publik (och därmed, ibland implicit, publiceringsforum) och omfattning. Uppmaningen kan också specificera ett tänkt innehåll. Ett exempel på en skrivuppmaning är följande:

Du är elevrådets ordförande och har fått i uppdrag att skriva en insändare till lokaltidningen Lokalbladet om varför skolmaten bör bli bättre. En insändare får inte överstiga 250 ord.

Den här uppgiften uppmanar provtagaren att argumentera och

säger något om såväl den tänkta publiken (allmänheten) som avsändaren (representant för elevrådet). Den anger också innehåll och omfattning.

Ett annat exempel på en uppgift som uppmanar elever att skriva, men som potentiellt kommer att vålla bekymmer är följande:

Skriv ett blogginlägg om valfritt ämne.

Det enda som specificerats i denna uppmaning är textens format. Vem som ska skriva vad, i vilket syfte och till vem har lämnats därhän, vilket kan skapa problem när texten ska bedömas. Ett mycket kort blogginlägg (»Jag gillar blommor!«) är lika mycket ett blogginlägg som ett långt blogginlägg som utreder det militärindustriella komplexet. Sätillvida vi är intresserade av att veta mer om provtagaren än att hon eller han kan foga samman bokstäver och ord, behöver en uppmaning ta sikte på skrivandets funktion för att en bedömare ska kunna avgöra om en text är bra eller dålig.

Som tidigare nämnts är bedömning av texter hämtade från skrivuppgifter svår. Som lärare gäller det då att på förhand specificera vilka dimensioner av skrivande som är relevanta att fokusera på och hur resultatet ska kommuniceras.

Trots att det inte alltid är samma skrivteori som ligger bakom olika provkonstruktioner ser bedömningsregler för skrivande förvånansvärt lika ut runtom i världen. Det är vanligt att man inkluderar dimensioner som:

- *»Röst«/kommunikation*, det vill säga skribentens förmåga att ikläda sig den angivna skribentrollen och vända sig till en angiven mottagare
- *Innehåll*, som relaterar till förmågan att skriva sakligt korrekt och/eller hålla sig till ett givet ämne
- *Struktur*, som relaterar till förmågan att ordna och binda ihop texten på ett för sammanhanget logiskt/ändamålsenligt sätt
- *Stil*, som relaterar till ändamålsenlig ton och vokabulär

- *Olika aspekter av formalia* (grammatik, rättstavning, interpunktion m.m.).

De olika dimensionerna kan bedömas var för sig, vilket kallas *analytisk bedömning*. För att sådan bedömning ska ha effekt (t.ex. att peka ut problematiskt område) bör separata resultat för varje dimension redovisas. Dimensionerna kan också bedömas med helhetspoäng, vilket kallas *holistisk bedömning*. Bedömningen avspeglar då vanligen hur väl texten, överlag, »gör sitt jobb«.

Bedömningens syfte avgör vilken av bedömningsmetoderna som är lämpligast. (Parentetiskt kan nämnas att nationella proven i svenska kombinerar de båda typerna genom att avkräva dels en helhetsbedömning, dels en bedömning av olika dimensioner.)

INTEGRERADE UPPGIFTER

De flesta svensklärare som administrerat ett nationellt prov är bekanta med det som brukar kallas *integrerade uppgifter* (Davies et al., 1999). En integrerad uppgift innebär ett försök från uppgiftskonstruktörens sida att bedöma flera förmågor (t.ex. läsning och skrivande) samtidigt. Ett typexempel är en skrivuppgift som bygger på mer eller mindre omfattande läsning. De texter som används kan till exempel finnas i ett läshäfte som delas ut före eller i samband med att en uppgift ska genomföras. En annan vanlig integrerad uppgift är bokrecensionen, som i idealfallet bygger på läsning av en bok och därefter författande av en recension. Vad som skiljer en integrerad uppgift från en läs- eller skrivuppgift som innehåller skrivande respektive läsande är att bedömningsregeln tar hänsyn till båda förmågorna.

Det finns många och klara fördelar med integrerade uppgifter. Den mest uppenbara är att de till stor del liknar uppgifter/handlingar som genomförs utanför skolan. Varken i arbets- eller privatlivet behöver vi särskilt ofta delta i glosförhör, skriva grammatiktentor eller rätta avsiktligt ogrammatiska meningar eller me-

ningsfragment. I stället använder vi både receptiva och produktiva färdigheter i olika kommunikativa syften.

Dessvärre finns också nackdelar med integrerade uppgifter. Dels kräver uppgiftstypen avancerade bedömningsregler som förmår inkludera aspekter av alla förmågor som testas. Dels är uppgiftstypen problematisk när bedömningssyftet är diagnostiskt. Att avgöra varför en elevlösning inte är tillfyllest kan vara komplicerat när eleven behövt läsa längre texter och därefter skriva. En dålig förståelse av det lästa kan leda till en dålig skriftlig utformning, och omvänt kan en utmärkt förståelse för en ovan skribent vara svår att kommunicera. En annan problematik med integrerade uppgifter är att de på grund av att de ofta är omfattande är svåra att genomföra vid upprepade tillfällen. Detta inverkar på möjligheten att generalisera resultaten.

Summering och lästips

Det här kapitlet har fokuserat på att definiera uppgifter som förekommer mer eller mindre frekvent i svenskundervisning. Förutom att presentera vanliga begrepp i uppgiftskonstruktionssammanhang har kapitlet också fokuserat på relationen mellan vad som ska bedömas och hur det bedöms. Den viktigaste konklusionen är den att ingen uppgift i sig varken är bra eller dålig. Uppgifters kvalitet avgörs i relationen mellan dess utformning och de påståenden om elevers läs- eller skrivförmåga vi vill göra.

Som har gjorts uppenbart är uppgiftstyper och i synnerhet uppgiftskonstruktion ett fält i sig. I dags dato finns få skolrelevanta beskrivningar av uppgiftskonstruktioner på svenska. En nyligen utkommen bok är dock Wikströms *Konsten att göra bra prov* (2013). På engelska finns fler titlar. Två standardverk som relaterar till läs- och skrivbedömning är Aldersons *Assessing reading* (2000) och Weigles *Assessing writing* (2002). Andra titlar som kan vara av intresse är Haladyna och Rodriguez (2013) om utveckling av uppgifter, Hughes (2003) om språkbedömning och Nitko och Brookhart (2011) om kunskapsbedömning i vid bemärkelse.

Referenser

- Alderson, J.C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Berge, K.L. (1996). *Norsksensorenes tekstnormer og doxa: en kultursemiotisk og sosiotekstologisk analyse* (Doktorsavhandling). Trondheim: Norges teknisk-naturvitenskapelige universitet.
- Björnsson, C.H. (1960). *Uppsatsbedömning och uppsatsskrivning*. Stockholm: Almqvist & Wiksell.
- Cizek, G.J. (2012). An introduction to contemporary standard setting. I G. J. Cizek (red.), *Setting performance standards: concepts, methods, and perspectives* (2 uppl.) (s. 3–14). New York: Routledge.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. & McNamara, T.F. (1999). *Dictionary of language testing*. Cambridge: Cambridge University Press.
- Edgeworth, F.Y. (1890). The element of chance in competitive examinations. *Journal of the Royal Statistical Society*, 53(3), 460–475, 644–663.
- Greeno, J.G. (1997). Response: On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5–17. Retrieved from <http://www.jstor.org/stable/1176867>
- Gustafsson, J.-E., Cliffordson, C. & Erickson, G. (2014). *Likvärdig kunskapsbedömning i och av den svenska skolan – problem och möjligheter*. Stockholm: SNS Förlag.
- Haladyna, T.M. & Rodriguez, M.C. (2013). *Developing and validating test items*. New York och London: Routledge.
- Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395–420. doi:10.1177/0265532211428317.
- Hillocks, G. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.
- Hughes, A. (2003). *Testing for language teachers* (2 uppl.). Cambridge: Cambridge University Press.
- Jølle, L. (2014). Pair assessment of pupil writing: A dialogic approach for studying the development of rater competence. *Assessing Writing*, 20, 37–52. doi:10.1016/j.asw.2014.01.002.

- Kane, M.T., Crooks, T.J. & Cohen, A.S. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18(2), 5–17. doi:10.1111/j.1745-3992.1999.tb00010.x.
- Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. doi:10.1177/026553229601300302.
- McNamara, T.F. (1996). *Measuring second language performance*. New York: Longman.
- Nitko, A.J. & Brookhart, S.M. (2011). *Educational assessment of students*. Upper Saddle River, N.J.: Pearson Merrill Prentice Hall.
- Skolinspektionen (2012). *Lika för alla? Omvärldsmätning av nationella prov i grundskolan och gymnasieskolan under tre år, 2012*. Stockholm: Skolinspektionen.
- Skolinspektionen (2013). *Olikheterna är för stora, regeringsrapport: omvärldsmätning av nationella prov i grundskolan och gymnasieskolan, 2013*. Stockholm: Skolinspektionen.
- Tengberg, M. (2014). Att pröva elevers läsförmåga i årskurs nio. En analys av uppgiftskonstruktioner i det nationella provet i svenska, delprov A: Att läsa och förstå. *Utbildning & Demokrati*, 23(2), 109–132.
- Weigle, S.C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Wikström, C. (2013). *Konsten att göra bra prov: vad lärare behöver veta om kunskapsmätning*. Stockholm: Natur & Kultur.

Del 2

Bedömning i praktiken

6. Skrivbedömning i formativt syfte och återkoppling

KARINA PÅLSSON GRÖNDAHL

Det finns olika anledningar till att vi ger återkoppling på våra elevers texter. Den främsta är säkerligen att vi vill försöka få våra elever att utvecklas i sitt skrivande. En annan kan vara att vi vill försöka få eleverna att förstå att skrivandet är en process och att en text kan förbättras genom att man bearbetar den. Kanske ger vi också återkoppling i hopp om att våra elever ska tycka och fortsätta tycka att det är roligt att skriva. Men tänk om vår återkoppling varken stöttar eller utvecklar våra elevers skrivande? Eller får dem att vilja lägga ner tid på att bearbeta sina texter. Tänk om det är så att eleverna helt enkelt tappar sugen och aldrig mer vill plita ner ett endaste ord till. Vilka blir följderna då för våra elever på lång sikt?

Bedömningens syfte och konsekvens

Bedömningens syfte framträder först när lärare tar beslut om hur bedömningsinformation ska användas. Bedömningsinformation som används för att sätta till exempel ett slutbetyg i årskurs 9 sägs användas i *summativt syfte*, medan information som används för att göra en bedömning av hur väl eleverna verkar tillägna sig den pågående undervisningen, eller för att ge en individuell återkoppling till varje elev, har ett *formativt syfte* (bedömning för lärande). Oavsett om bedömningens syfte är summativt, formativt eller både och, är det

viktigt att vi försöker stärka *validiteten* i våra bedömningar. Validitet handlar om att vi ska sträva efter att göra våra bedömningar *relevanta* och *trovärdiga*. Relevanta utifrån att vi faktiskt bedömer det vi säger att vi ska bedöma och trovärdiga utifrån att bedömningar görs på ett likvärdigt, rättssäkert och professionellt sätt.

Likvärdighet i bedömning är en viktig aspekt. Men när bedömningen syftar till att stötta och utveckla elevers skrivförmåga handlar likvärdighet inte om att eleverna ska få exakt samma återkoppling av läraren. I stället bör återkopplingen fokusera olika aspekter av elevers skrivande, beroende på vad varje elev behöver arbeta med för att utvecklas. Därför bör likvärdighet i stället ses i relation till återkopplingens kvalitet (Stobart, 2012, s. 234). Om återkopplingens kvalitet är hög är sannolikheten större att elever ges möjlighet att förstå och använda återkopplingen. Om återkopplingen i stället är av bristande kvalitet minskar elevers möjlighet till förståelse och användning av den återkoppling som ges.

Bedömning som syftar till att stötta och utveckla behöver kanske framför allt diskuteras i relation till sitt syfte. Om syftet med att ge återkoppling på elevers texter är att texterna ska förbättras bör vi som lärare vara intresserade av att undersöka om elevernas texter verkligen blir bättre. Och om syftet är att elever ska motiveras till att fortsätta skriva bör vi vara intresserade av att deras motivation till skrivande ökar. Det här är självklart en svår och tidskrävande utmaning, men ett intresse riktat mot vad vår återkoppling gör eller inte gör kan bidra till en ökad medvetenhet. En ökad medvetenhet kan då, så småningom, resultera i att de avsikter vi har som lärare kan uppnås i högre grad.

De konsekvenser som uppstår efter bedömning av en elevs förmåga har olika verkan. Ofta brukar bedömning i summativt syfte anses vara förknippad med *high-stakes assessment*, det vill säga bedömning förknippad med stora eller avgörande konsekvenser för eleven, till skillnad från bedömning i formativt syfte som anses vara *low-stakes assessment*, det vill säga bedömning förknippad med mindre avgörande konsekvenser för eleven. För elever står

det alltså mer på spel när bedömning används i summativt syfte än i formativt syfte. Om en lärare ger ett felaktigt betyg till en elev som gör att den eleven inte kommer in på till exempel önskat gymnasieprogram, anses konsekvensen vara mer allvarlig än om till exempel återkopplingen varken förstås eller används av eleven. Men stämmer verkligen detta? Vad får det för konsekvenser om vi lärare kontinuerligt ger återkoppling som är av låg kvalitet? Får inte även dessa tolkningars användning negativa konsekvenser för elevernas möjlighet att lära det som ska bedömas, och i slutändan också för elevernas betyg?

Kapitlets upplägg

Det här kapitlet fokuserar bedömning i formativt syfte. Eftersom det belyser en av flera alternativa bedömningsformer kommer det att inledas med en kort introduktion om bedömningens förändrade roll för att förklara varför alternativa bedömningsformer fått mer uppmärksamhet. Därefter får vi följa en lärares tankar och handlingar under tre faser av ett skrivförlopp: en planerings-, en genomförande- och en återkopplingsfas/bedömningsfas. Precis som i bokens övriga kapitel kommer Kanes validitetsmodell (2006) (se även kapitel 10) att användas för att belysa, och diskutera, vilka konsekvenser lärarens beslut får. Eftersom kapitlet fokuserar bedömning i formativt syfte blir det även viktigt att i det undervisningsexempel som används belysa vilka potentiella svårigheter som kan uppstå i den här specifika bedömningsformen. Framför allt fokuseras en av de viktigaste strategierna i formativ bedömning, nämligen återkoppling.

Kapitlet avslutas med en reflektion av undervisningsexemplet samt en kort introduktion till bedömningsformen, *dynamic assessment* (Poehner, 2008). Dynamisk bedömning erbjuder ett annat förhållningssätt till återkoppling. Ett förhållningssätt som kanske kan bidra till att utmana oss i hur vi tänker kring relationen mellan undervisning och bedömning.

Bedömningens förändrade roll

Bedömningens roll har förändrats under loppet av de senaste decennierna. Från att tidigare enbart ha använts som ett sorteringsinstrument används bedömning i dag även som ett pedagogiskt verktyg. Bedömning som pedagogiskt verktyg kan hjälpa oss att både stärka våra elevers lärande och att utveckla vår undervisning. Framför allt kan bedömningens förändrade roll ses som en strävan att öka elevers delaktighet i bedömningsprocessen (Gipps, 1999), en strävan som även uttrycks i grundskolans läroplan (Skolverket, 2011a, s. 15).

På grund av bedömningens förändrade roll har alternativa bedömningsformer, som till exempel själv- och kamratbedömning, fått mer uppmärksamhet. Inte minst har formativ bedömning med sina fem nyckelstrategier fått stort gehör. En av dessa nyckelstrategier, återkoppling, har visat sig ha avgörande betydelse för elevers möjlighet att lära och för deras motivation (Black & Wiliam, 1998). Tanken att återkoppling ska fungera som en bro mellan bedömning och undervisning, liksom mellan lärare och elever, är tilltalande. Ännu mer tilltalande blir tanken när det visat sig att det finns en särskild form av återkoppling som verkar ha stor potential att påverka elevers möjlighet att lära, nämligen den uppgiftsrelaterade (task-based feedback; Hattie & Timperley, 2007).

För många lärare som undervisar i skrivande är återkoppling ingen ny företeelse. De förvånas heller inte över att återkoppling, som riktar sig mot en specifik uppgift, kan ha störst påverkan på elevers möjlighet att lära. Lärare som undervisar i skrivande har sedan länge, innan återkopplingens nyintåg, ägnat tid åt att rätta och ge respons på elevers texter (Sommers, 1982). Det som dock är förvånande är att återkoppling fortfarande verkar vara en svår konst att bemästra. Återkoppling har inte bara visat sig vara svår för lärare att ge (Ferris, 2003), utan även för elever att ta emot, förstå och använda. Det här kapitlet avser att belysa några av svårigheterna med återkoppling och förhoppningsvis sätta igång tankar hos lärare som till vardags arbetar med bedömning av skrivande.

Läraren

Läraren som beskrivs i det här kapitlet undervisar en grupp elever i årskurs 8 i svenska. Vi kommer att få följa honom när han planerar och genomför undervisningen och ger återkoppling till eleverna. Syftet med hans bedömning är att samla in elevers utkast för att identifiera styrkor och svagheter i elevernas texter, för att sedan använda informationen för att visa varje elev vilka utvecklingsmöjligheter texten har. Återkopplingen som läraren kommer att ge är skriftlig och varje elev kommer att få en individuell skriftlig kommentar.

LÄRAREN PLANERAR UTIFRÅN LÄROPLANEN

Läraren inleder arbetet med att ta fram läroplanen, närmare bestämt kursplanen i ämnet svenska. Här anges inledningsvis de fem förmågor (se Skolverket 2011a, s. 222) som undervisningen ska ge eleverna möjlighet att utveckla:

- Formulera sig och kommunicera i tal och skrift
- Läsa och analysera skönlitteratur och andra texter för olika syften
- Anpassa språket efter olika syften, mottagare och sammanhang
- Urskilja språkliga strukturer och följa språkliga normer
- Söka information från olika källor och värdera dessa.

Läraren bestämmer sig för att eleverna ska ges möjlighet att utveckla den första förmågan, med betoning på att formulera sig och kommunicera i skrift. Han studerar också kunskapskraven som anger bedömningens inriktning, och då särskilt kraven för bedömning av den skriftliga förmågan. Läraren läser kunskapskraven för betyget A eftersom han, även om alla elever inte har A i svenska, vill försäkra sig om att planeringen gör det möjligt för alla elever att arbeta mot de kvaliteter som betonas för det högsta betygssteget.

Eleven kan skriva olika slags texter med god språklig variation, välutvecklad textbindning samt väl fungerande anpassning till texttyp, språkliga normer och strukturer. De berättande texter eleven skriver innehåller välutvecklade gestaltande beskrivningar och berättargrepp samt dramaturgi med komplex uppbyggnad. Eleven kan söka, välja ut och sammanställa information från ett varierat urval av källor och för då välutvecklade och väl underbyggda resonemang om informationens och källornas trovärdighet och relevans.

Sammanställningarna innehåller välutvecklade och nyanserade beskrivningar och förklaringar, välutvecklat ämnesrelaterat språk samt väl fungerande struktur, citat och källhänvisningar. Genom att kombinera olika texttyper, estetiska uttryck och medier så att de olika delarna samspelar på ett ändamålsenligt och effektivt sätt kan eleven förstärka och levandegöra sina texters budskap. Dessutom kan eleven ge välutvecklade och nyanserade omdömen om texters innehåll och uppbyggnad och utifrån respons bearbeta texter mot ökad tydlighet, kvalitet och uttrycksfullhet på ett väl fungerande sätt (Skolverket 2011a, s. 232).

Genom att läsa igenom kunskapskraven försäkras läraren sig också om att han vet vad som ska bedömas, innan han går vidare och planerar för undervisningens utformning. Läraren inser att han inte kan täcka in samtliga kunskapskrav genom att enbart låta eleverna genomföra en skrivuppgift. I stället måste han välja ut några av kunskapskraven och försäkra sig om att skrivuppgiften som han planerar för gör det möjligt att bedöma de aspekter som framträder i de kunskapskrav han har valt. Eftersom det centrala innehållet anger vad undervisningen i ämnet ska behandla för årskurserna 7–9, läser läraren även igenom detta avsnitt.

Nu börjar en idé växa fram hos läraren om att låta eleverna skriva ett brev. Läraren tänker att de flesta elever någon gång har skrivit ett brev och borde vara bekanta med hur ett brev kan se ut, både till form och till innehåll. Läraren har läst om genrepedagogik och vet

att brev ofta innehåller vissa element, som till exempel användandet av en hälsningsfras och avskedsfras, men är också medveten om att brev kan skilja sig åt beroende på i vilket syfte brevet skrivs. Det är ju en sak att skriva ett brev till en kompis och berätta vad man har gjort under sommaren, tänker han, och en helt annan att skriva ett brev till Lady Gaga och be henne att spela på en välgörenhetsgala. Det faktum att brevet kan varieras i många avseenden anser läraren vara något positivt, något som gör det möjligt att bedöma hur väl eleverna anpassar sitt skrivande till »... texttyp, språkliga normer och strukturer.« (Skolverket 2011a, s. 237). Därför funderar läraren vidare kring vem eleverna ska skriva till, vad eleverna ska skriva om och varför de ska skriva brevet; eftersom brevets vem, vad och varför kommer att påverka hur det skrivs gäller det att hitta en spännande person!

För just den här uppgiften bestämmer läraren att mottagaren av brevet ska vara Barack Obama. Barack Obama har vid tiden för lärarens arbetsplanering nyligen varit på besök i Sverige¹ och läraren tror att det kan bidra till att uppgiften känns aktuell och att eleverna kanske i högre grad känner sig motiverade inför uppgiften. Vidare tänker han att syftet med brevet ska vara att försöka övertyga Obama att genomföra något slags förändring efter sitt Sverigebesök. Därför blir det kanske bra, tänker läraren, om eleverna får skriva om ett aktuellt ämne – arbetslöshet – som gör det möjligt för dem att jämföra svenska och amerikanska förhållanden, för att sedan kunna argumentera för den förändring som de ska föreslå. Läraren bestämmer därför att jämförelserna ska grunda sig på såväl fakta som på egna erfarenheter.

Läraren funderar sedan över upplägget för skrivuppgiften och tänker sig följa en skrivprocessmodell (Strömquist, 2007). Läraren vet att det finns flera skrivprocessmodeller, varav de flesta innehåller fyra faser i ett skrivförlopp: planering, utkast, återkoppling/respons och bearbetning. Ett sådant upplägg gör det möjligt, re-

¹ Barack Obama besökte Stockholm i september 2013.

sonerar läraren, att ge eleverna återkoppling/respons och låta dem bearbeta responsen, som också är en förmåga som ska bedömas enligt kunskapskraven.

Läraren vet att det finns lässvaga elever i gruppen, liksom det finns elever som har svårt att välja texter på egen hand. Därför väljer han själv ut två kortare texter om arbetslöshet: en text som handlar om arbetslöshet i USA och en text om arbetslöshet i Sverige. I och med ett sådant upplägg försäkras läraren om att samtliga elever får tillgång till och förståelse för ett ämnesinnehåll som de kan använda, om de vill, i sitt brev. Därefter tänker han att eleverna ska skriva ett första utkast och lämna in till honom. Han planerar sedan att läsa igenom elevernas utkast och förse dessa med rättningar och individuella kommentarer som ska syfta till att hjälpa varje elev att förbättra sitt brev. Eleverna ska i nästa steg få bearbeta sina brev med hjälp av lärarens kommentarer och lämna in det bearbetade brevet för en slutgiltig bedömning.

Läraren anser att skrivuppgiften, såsom den nu är planerad, även täcker ytterligare en förmåga som eleverna ska ges förutsättning att utveckla i svenskundervisningen, nämligen att:

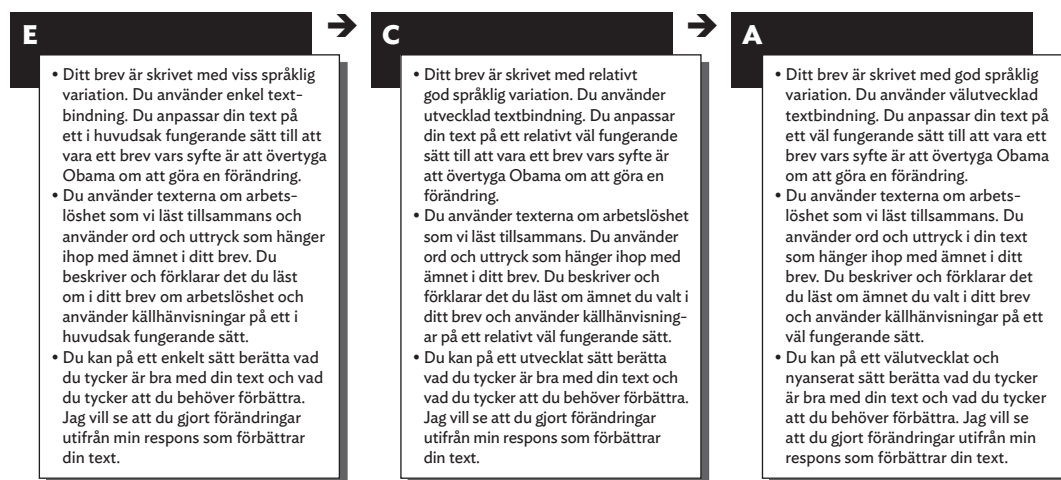
- Formulera sig och kommunicera i tal och *i skrift*
- Anpassa språket efter olika syften, mottagare och sammanhang.

De kunskapskrav som läraren tänker rikta sin bedömning mot blir följande (de angivna värdeorden är skrivna enligt betygssteg E/C/A):

1. Eleven kan skriva olika slags texter med viss/relativt god/god språklig variation, enkel/utvecklad/välutvecklad textbindning samt i huvudsak/relativt väl/väl fungerande anpassning till texttyp, språkliga normer och strukturer.
2. Dessutom kan eleven ge enkla/utvecklade/välutvecklade och nyanserade omdömen om texters innehåll och uppbyggnad

och utifrån respons bearbeta texter mot ökad tydlighet, kvalitet och uttrycksfullhet på ett i huvudsak/fungerande/väl fungerande sätt.

Läraren utformar sedan en bedömningsmatris (Figur 6.1) som han ska dela ut till eleverna där han försöker konkretisera hur aspekterna i kunskapskraven kan komma till uttryck och bedömas i den här specifika skrivuppgiften.



Figur 6.1. Bedömningsmatris för den specifika uppgiften.

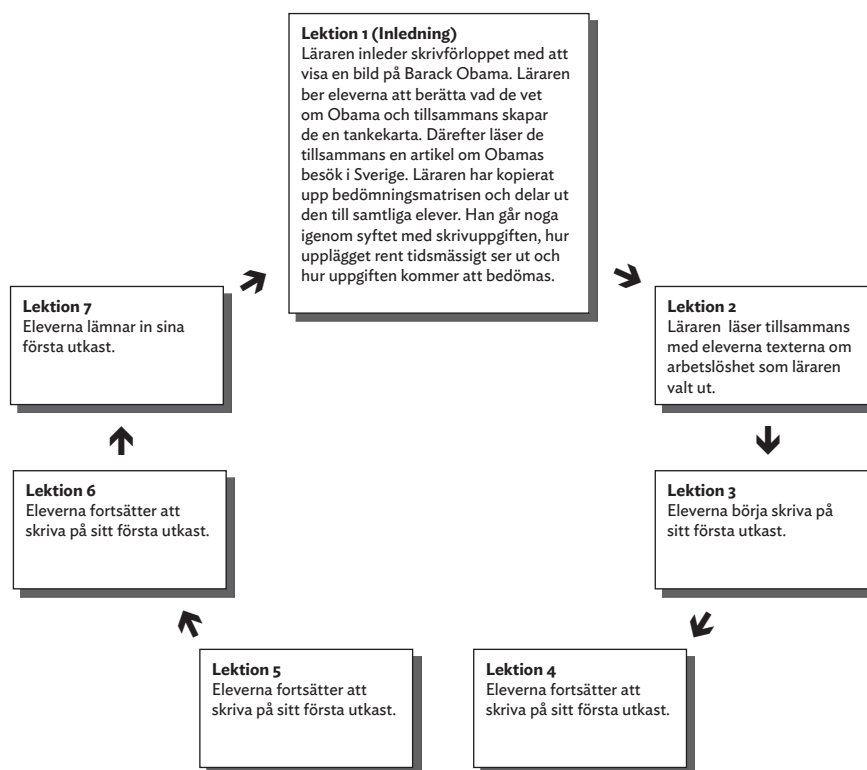
Läraren känner sig nöjd med bedömningsmatrisen. Dock är den sista meningen som lyder: »Jag vill se att du gjort förändringar utifrån min respons som förbättrar din text«, egentligen inte en kvalitetsaspekt som berör själva texten, utan den handlar snarare om en kvalitetsaspekt som riktar in sig på en process. Det är en aspekt som läraren inte riktigt har beaktat i bedömningsmatrisens utformning.

Bedömningsmatrisen planerar läraren att använda när han ger eleverna återkoppling, för att försäkra sig om att den återkoppling

som ges utgår från det han sedan tänker bedöma. Innehållet i matrisen, tänker han, gör det också möjligt för honom att avgöra var eleven befinner sig och vart eleven är på väg. Återkopplingen kan då bli framåtsyftande. Hur eleven ska ta sig dit kommer läraren tydliggöra skriftligt i den individuella kommentar som varje elev ska få ta del av.

LÄRAREN GENOMFÖR LEKTIONERNA

Nästa steg är att läraren genomför lektionerna såsom illustreras av Figur 6.2 nedan:



Figur 6.2. Lektionsplanering.

Efter sju lektioners arbete är det dags för eleverna att lämna in sina utkast. Läraren samlar in texterna och tar med sig dem hem över helgen för att kunna lämna tillbaka utkastet med skriftlig återkoppling på måndagen.

LÄRAREN ÅTERKOPPLAR/BEDÖMER ELEVERNAS UTKAST

Läraren sätter sig med den hög med utkast som han tagit med sig hem och plockar fram den första texten. Han läser igenom den och börjar sedan stryka under stav- och grammatikfel. Där han tycker att han inte riktigt förstår skriver han ett frågetecken i marginalen. I slutet av varje text skriver läraren »Bra!«.

När han läser bedömningsanvisningarna går det upp för honom att han inte riktigt vet hur han ska bedöma elevernas förmåga att bearbeta sina texter enligt följande »... mot ökad tydlighet, kvalitet och uttrycksfullhet på ett i huvudsak/fungerande/väl fungerande sätt«. Läraren ser framför sig ytterligare en helg där han ska sitta och jämföra elevernas ursprungliga texter med deras bearbetade versioner. Han bestämmer sig för att be eleverna att lämna in en kort text där de helt enkelt får beskriva vilka förändringar de gjort och motivera sina förändringar. Han går vidare med att skriva den individuella kommentaren till det första utkastet som eleven Karin har skrivit (Figur 6.3).

Hej Karin!

Tack för att jag fick läsa din text. Jag tycker din text är bra men det finns självklart sådant som kan förbättras.

Jag tycker att du skulle kunna ha en bättre inledning.

Variera din meningsbyggnad och ditt ordval.

Använd styckeindelning så blir din text lättare att läsa.

Figur 6.3. Exempel på individuell kommentar.

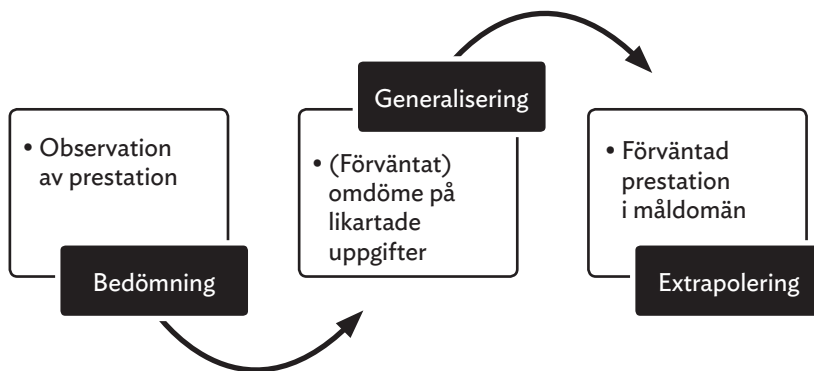
Efter flera timmars läsning och skrivande av individuella kommentarer är läraren till slut klar med hela klassens texter. Han lägger högen med texter i sin väska för att kunna dela ut de kommenterade utkasterna under måndagslektionen.

När läraren på lektionen sedan lämnar tillbaka texterna till eleverna berättar han att de har två lektioner på sig att bearbeta sina texter. Några elever suckar. En elev räcker upp handen och frågar om man får skriva hemma. Läraren svarar att han vill att eleverna ska skriva under lektionerna. En annan elev frågar om man får renskriva på datorn, och läraren tycker att eleverna får göra det om de vill.

De kommande lektionerna arbetar eleverna vidare med att bearbeta sina texter och de flesta lämnar in efter två lektioner. Några elever vill ha mer tid och läraren ber dem då att lämna in så fort det bara går. Efter ytterligare två veckor har samtliga elever lämnat in och läraren läser igenom texterna för att göra en slutgiltig bedömning.

Kanes validitetsmodell

Om man som lärare vill dra en slutsats om att en elev har kunskaper om någonting i mer generell mening, krävs det att läraren gör en trovärdig tolkning av vad resultatet från bedömningen representerar. Läraren i exemplet ovan vill för undervisningens skull kunna dra slutsatser om elevernas generella skrivförmåga, och inte bara om deras förmåga att skriva ett brev i syfte att övertyga Barack Obama om att genomföra någon slags förändring kopplat till arbetslöshet. Kanes modell (2006) (se kapitel 10) illustrerar att tolkningar måste vara trovärdiga i tre led, för att slutsatsen som helhet ska kunna betraktas som giltig (Figur 6.4).



Figur 6.4. Inferenser i tre led utifrån Kane (2006).

OBSERVATION AV PRESTATION

I det första ledet i modellen observeras elevens prestation, som i vårt exempel består av två prestationer: det av eleven skrivna utkastet till brevet och det av eleven bearbetade brevet. Trovärdigheten i tolkningen stärks om läraren bedömer det han har sagt att han ska bedöma enligt den bedömningsmatris som han skrivit och delat ut till eleverna. När läraren börjar läsa igenom elevernas utkast stryker han bara under stav- och grammatikfel i texterna; han fokuserar alltså sin återkoppling/bedömning på de bevis på språklig korrekthet som finns i elevernas texter.

I lärarens kommentarer till Karins text (Figur 6.3) kan vi se att han riktar in sin återkoppling på: brevets inledning, variationen av meningar och ordval samt styckeindelningen, vilket även är något som han skrivit att han ska bedöma. Däremot riktar hans återkoppling/bedömning enbart in sig på vissa aspekter av det första kunskapskravet och lämnar andra aspekter därhän, som till exempel hur väl brevet är anpassat till brevets mottagare. Om läraren gjort sin bedömning utifrån sin bedömningsmatris och gett återkoppling på samtliga aspekter i kunskapskraven skulle bedömningens trovärdighet ha ökat.

Att som lärare göra en urskilning av de aspekter som ska

bedömas kan vara svårt, men en konsekvent tillämpning av kunskapskraven bidrar till en giltigare tolkning. Under bedömningens gång kan det visa sig att aspekter synliggörs som läraren kanske inte riktigt uppfattat som betydelsefulla för uppgiften. Det behöver i och för sig inte vara problematiskt, så länge som läraren bedömer det han säger att han ska bedöma.

OMDÖME PÅ LIKARTADE UPPGIFTER

I det andra ledet i Kanes modell används omdömet på den enskilda uppgiften till att även gälla andra likartade uppgifter. Men om skrivuppgiften i exemplet byts ut mot en annan likvärdig uppgift, kommer eleven då att uppvisa ungefär samma resultat i den nya uppgiften? Trovärdigheten stärks förstås om en elev presterar likadant på flera jämförbara uppgifter, men försvagas om elevens resultat visar sig vara uppgiftsberoende. Att generalisera från en typ av uppgift är svårt. Ett sätt att ändå försöka öka trovärdigheten i generaliseringsledet blir att låta eleverna genomföra flera jämförbara skrivuppgifter. Följdfrågan blir då naturligtvis: Hur många jämförbara texter ska elever skriva? Det är förstås svårt att svara på. Å ena sidan kan läraren rent praktiskt fundera över hur många skrivuppgifter som hinner läggas in under terminen eller läsåret, när det finns andra förmågor som elever också ska ges möjlighet att utveckla. Det kan vara svårt att inom ramen för en termin, och till och med inom ramen för ett läsår, göra flera liknande uppgifter, vilket krävs om resultaten ska gå att generalisera. Å andra sidan bör läraren fundera över vad det får för konsekvenser om elever bara genomför en enda skrivuppgift. Vissa elever kanske redan är förtrogna med att skriva brev och/eller argumenterande texter. Andra elever kanske är förtrogna med problem som finns i den amerikanska och svenska kontexten, och har på så vis lättare för att göra jämförelser. Det innebär kanske att vissa elever kommer att prestera jämförelsevis bättre på just den här uppgiften än de skulle gjort på en annan uppgift som ägnades åt att pröva samma förmåga.

FÖRVÄNTAD PRESTATION I MÅLDOMÄNEN

I det tredje och sista ledet utgår modellen i vårt exempel från att elevernas resultat på brevpuppgiften i den specifika undervisningskontexten även gäller i andra kontexter, till exempel utanför skolan eller i andra domäner, såsom svenskämnet i stort. Det innebär att en elev som bedöms klara av att skriva ett brev i syfte att övertyga, även ska klara av att göra det i en annan situation. Trovärdigheten är beroende av relationen mellan den enskilda bedömningen och hur bedömningsverktygen överensstämmer med lärarens kunskap om måldomänen, det vill säga den domän av kunskaper och färdigheter som man genom en viss bedömning vill kunna uttala sig om. Om vi tittar på lärarens bedömningsmatris (Figur 6.1) så kan vi se att det finns en viss överensstämmelse mellan det som ska bedömas och målbeskrivningarna.

Det som kan stå i vägen för bedömning i formativt syfte

Vad vet vi då om det som hindrar respektive det som stöttar bedömning i formativt syfte? Stöttar läroplanen användandet av den här bedömningsformen? Hindrar elevernas inställning till och erfarenhet av den här bedömningsformen dem från att lägga ner tid och energi på att bearbeta sina texter? En sak är säker, bedömning som används i syfte att stötta och utveckla lärande blir svår att implementera och använda, om man inte först är medveten om att det finns många faktorer som kan påverka (Stobart, 2012). Nedan följer några exempel på *yttre* och *inre faktorer*, som kan påverka det som sker i klassrummet, den så kallade inre kontexten.

YTTRE FAKTORERS PÅVERKAN PÅ DEN INRE KONTEXTEN

Utbildningens roll och status i ett samhälle påverkar elevers inställning till skolan och till den undervisning som sker där. Den kan således antingen ha en stöttande eller hindrande funktion för den bedömning som används i formativt syfte. Utbildningens roll

konkretiseras i skolans styrdokument. I läroplanens övergripande mål och riktlinjer för bedömning och betyg (Skolverket, 2011a, s. 18) har skolan två mål:

1. Varje elev ska utveckla ett allt större ansvar för sina studier
2. Varje elev ska utveckla förmågan att själv bedöma sina resultat och ställa egna och andras bedömning i relation till de egna arbetsprestationerna och förutsättningarna

Ett större ansvar hos eleverna för sina studier inbegriper i den första punkten en ökad delaktighet i de processer som de är en del av i skolan. Just bedömningsprocessen är en sådan process där ökad delaktighet kan uppnås, genom alternativa bedömningsformer.

Den andra punkten i riktlinjerna lyfter sådana alternativa bedömningsformer som redan berörts i kapitlets inledning, som till exempel själv- och kamratbedömning. Dessa former för bedömning kan samspela i undervisningen med lärarens bedömning för att uppnå en ökad delaktighet, och därmed ge möjlighet för elever att utveckla större ansvar för sina studier. I Skolverkets »Allmänna råd« (Skolverket, 2011b) finns också stöd för bedömning i formativt syfte. Sammanfattningsvis kan man säga att skolans styrdokument är utformade för att främja alternativa bedömningsformer.

INRE FAKTORERS PÅVERKAN PÅ DEN INRE KONTEXTEN

Även den inre kontexten, klassrumskontexten, måste främja användandet av bedömning i formativt syfte. För att till exempel den återkoppling som ges i klassrummet ska bli produktiv har Stobart (2012) identifierat några betydande inre faktorer:

- Tillit och motivation
- En delad förståelse för vad som ska läras
- Det formativa i ett summativt klimat

Tillit och motivation är viktiga förutsättningar för att elever ska våga och vilja i klassrummet, samt för att de ska engagera sig i klassrumsarbete. De måste förstå vad arbetet i klassrummet syftar till och att diskussioner kring vad som är svårt är nödvändiga för att de ska kunna komma runt sina svårigheter. Elever måste också vilja dela med sig av sina texter så att olika typer av svårigheter får träda fram och bemötas. Läraren är den som tillsammans med eleverna skapar tillit och motivation. En lärare som ser elevers svårigheter som något som måste utforskas vidare visar respekt och intresse för elevernas läroprocess.

En av strategierna i formativ bedömning är att försöka tydliggöra för eleverna »vart de är på väg«. Det innebär att eleverna måste få veta vad som ska läras och hur en god prestation ser ut. Antagandet bakom detta är att man har lättare att lära sig om man får se ett exempel på hur lärandet ser ut/kommer till uttryck. Om eleverna får se ett exempel på ett brev som är väl anpassat till mottagaren, är det kanske lättare för dem att åstadkomma en liknande anpassning i sina brev.

Om tydlighet saknas i bedömningen är det ett hot mot validiteten (Stobart, 2012, s. 237). Kunskapskraven behöver därför tolkas och förklaras av läraren genom konkreta exempel. Däremot är det svårt att se till att de exempel som visas inte bara blir »mallar« som eleverna kopierar. Risken är då att det blir ett lärande som ligger på ytan och som faktiskt inte gör att elever utvecklar en djupare förståelse för vad kvalitet i ett visst arbete innebär. Kunskapskraven ska heller inte tydliggöras så att de av eleverna uppfattas som en checklista på vad som ska läras. Torrance (2007) kallar det här för *criteria compliance* och menar att om kunskapskrav överspecificeras leder det till *assessment as learning*, i stället för *assessment for learning*. Om eleverna inte förstår vad lärandet syftar till blir det ju svårt för dem att förstå varför återkoppling ges, vad den syftar till och hur återkopplingen ska användas.

Den sista faktorn som Stobart (2012, s. 238f) pekar ut som ett hot/hinder är den summativa bedömningens inverkan. I länder där

till exempel slutprov står för en betydande del av läsårsbetyget, står naturligtvis de här proven högt i kurs. Det blir således svårt att i en sådan kontext försöka argumentera för att andra bedömningsformer är viktiga och användbara.

UPPNÅ KVALITET I ÅTERKOPPLING

Likvärdighet i återkoppling handlar, som tidigare nämnts, inte om att alla elever ska få samma återkoppling, utan om återkopplingens kvalitet. Parr och Timperley (2010) har identifierat några aspekter av återkopplingens kvalitet, som kan vara användbara i lärares arbete: läraren måste veta vad kvalitet är i förhållande till den specifika uppgiften, hur den uppgift som bedöms ligger till i förhållande till den kvalitet som eftersöks, vilka svårigheter den bedömda texten uppvisar i förhållande till den kvalitet som eftersöks, och slutligen måste läraren kunna beskriva/artikulera en lösning till eleven.

Att som lärare kunna beskriva/artikulera en lösning är inte helt enkelt. Om vi går tillbaka och tittar på den skriftliga kommentar som Karin fick (Figur 6.3), så kan vi fundera över om lärarens kommentar ger Karin någon vägledning i hur hon kan förbättra sin text.

Hälsningsfrasen och lärarens tack (»Hej Karin! Tack för att jag fick läsa din text.«) accentuerar det faktum att återkoppling har en känslomässig påverkan, vilket är en viktig funktion att beakta, även om den inte direkt hjälper Karin att förbättra sin text. Den andra delen i andra stycket (»Jag tycker din text är bra«) är väldigt generell, eftersom den inte ger en specifik beskrivning av vad som är bra. En tydlig positiv återkoppling kan trigga igång förbättringar i texten, och är självklart viktig för att motivera Karin till fortsatt arbete. De tre uppmaningar som följer pekar på vad som kan förbättras: en bättre inledning, en variation av meningar och ord samt styckeindelning. Däremot står det ingenting om hur Karin ska gå tillväga för att åstadkomma detta.

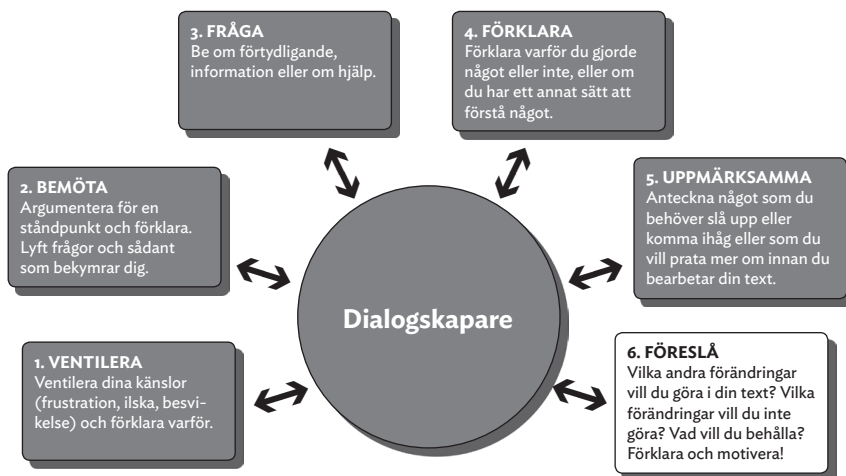
Behövs det mer personlig information i Karins inledning till brevet eller framgår det inte varför hon skrivit brevet? Har Karin

kanske använt hälsningsfrasen »Tja Obama!«, vilket inte riktigt passar för brevets mottagare och syfte? Det står inte heller något om hur variation i meningsbyggnaden kan uppnås. Är det textbindning som åsyftas? Kanske hade några konkreta förslag från läraren på hur textbindning kan utvecklas i texten varit behjälpligt. Den tredje och sista uppmaningen förklarar i och för sig att styckeindelning kommer att göra brevet mer läsvänligt. Dock förutsätter den att Karin vet vad ett stycke är och när ett nytt stycke ska göras.

Oavsett hur vi lärare väljer att formulera våra kommentarer måste vi ta hänsyn till att alla elever inte har, eller inte har haft, möjlighet att tillägna sig ett språk om språket, ett *metaspråk*. Alla elever vet inte vad ett stycke är eller vad meningsbyggnad är, men det är något som alla elever kan lära sig om de får möjlighet att göra det. Vi måste också förstå och ta hänsyn till att vår återkoppling inte alltid ses som en stöttning. Vissa elever tycker kanske att vår återkoppling försämrar, eller rättare sagt, förändrar deras texter. Därför kanske de väljer att bortse från den. Då handlar det inte längre om huruvida eleverna förstår eller inte, utan om att texten är personlig och att eleven vill behålla sitt personliga uttryck.

SÄKERSTÄLLA ÅTERKOPPLINGENS/BEDÖMNINGENS SYFTE

Om läraren i vårt exempel skulle vilja försöka säkerställa att återkopplingen fyller sin funktion/syfte, skulle han kunna skapa en dialog om återkopplingen genom att på ett medvetet sätt få eleverna att reagera på den, och agera utifrån den. Gay (1998, s. 1) föreslår fem användbara dialogskapare som läraren skulle kunna ha hjälp av: *ventilera, bemöta, fråga, förklara* och *uppmärksamma*. Ett sjätte förslag till dialogskapare, som också är Gays men som inte beaktas i hennes ursprungliga modell, är *föreslå*, som jag tror kan bidra till att ge elever utrymme att lyfta och motivera egna tänkta förändringar eller icke-förändringar, vilket inte beaktas i Gays ursprungliga modell (1998) (Figur 6.5).



Figur 6.5. Dialogskapare i återkoppling till elever.

Läraren hade i samtliga faser också varit behjälpt av en teoretisk skrivmodell som Skrivhjulet (se kapitel 3). I modellen beskrivs skrivande som en komplex förmåga att utföra olika skrivhandlingar för olika syften, i stället för att begränsa skrivandet utifrån en viss texttyp. Enligt modellen skriver vi i syfte att *förhandla, reflektera, beskriva, utforska, föreställa oss* och *argumentera*.

Fördelen med att utgå från modellen är att det blir tydligt att till exempel ett brev kan inbegripa flera syften. Ytterligare en fördel är att den tydliggör den tekniska aspekten av skrivande. Med andra ord vad eleverna behöver kunna och behärska för att uppnå syftet med att skriva sitt brev till Obama. Belysandet av den tekniska aspekten gör det också möjligt att säkerställa att undervisningen behandlat det som sedan ska bedömas.

Ett annat sätt att se på återkoppling

Sammanfattningsvis kan det konstateras att det inte är lätt att bedöma elevers skrivande. Inte heller är det lätt att ge återkoppling, vilket framstår ganska tydligt i exemplet. Det finns många faktorer

som påverkar hur återkoppling förstås och används av eleverna. Å ena sidan kan man som lärare bli medveten om dessa faktorer och försöka utveckla kvaliteten på sin återkoppling. Å andra sidan tål det att fundera på vilka andra alternativa bedömningsformer som finns för att stötta och utveckla elevers skrivande.

Ett intressant alternativ, som presenteras av Poehner (2008), är det som kallas *dynamisk bedömning*. Dynamisk bedömning skiljer inte undervisning och bedömning åt, vilket anses vara bedömningsformens främsta styrka. Återkoppling ges i stället av läraren i form av ledtrådar, där ledtrådarna successivt ökar mer och mer i sin stöttande funktion i att lösa de problem som eleven har. På det här sättet kan läraren få reda på hur mycket stöttning en elev behöver få, givet ett specifikt problem i texten, samtidigt som eleven lär sig att använda ledtrådarna (Alavi & Taghizadeh, 2014).

Referenser

- Alavi, M.S., & Taghizadeh, M. (2014). Dynamic assessment of writing: The impact of implicit/explicit mediations on L2 learners' internalization of writing skills and strategies. *Educational Assessment*, 19 (1), 1–16.
- Black, P. & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: GL Assessment.
- Ferris, D. (2003). Responding to writing. I B. Kroll (red.), *Exploring the dynamics of second language writing*, (s. 119–140). Cambridge: Cambridge University Press.
- Gay, P. (1998). Dialogizing response in the writing classroom: Students answer back, *Journal of Basic Writing*, 17 (1), 3–17.
- Gipps, C. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24, 355–392.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77 (1), 81–112.
- Kane, M.T. (2006). Validation. I R. L. Brennan (red.), *Educational measurement* (4 uppl.) (s. 17–64). Westport, CT: Praeger Publishers.
- Parr, M.J., & Timperley, H. (2010) Feedback to writing, assessment for teaching and learning and students progress. *Assessing Writing*, 15, 68–85.
- Poehner, M. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. New York: Springer.
- Skolverket. (2011a). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011*. Stockholm: Fritzes.
- Skolverket. (2011b). *Allmänna råd med kommentarer om planering och genomförande av undervisningen – för grundskola, grundsärskola, specialskolan och sameskolan*. Stockholm: Fritzes.
- Sommers, N. (1982). Responding to Student Writing, *College Composition and Communication*, 33(29), 148–156.
- Stobart, G. (2012). Validity in Formative Assessment. I J. Gardner (red.), *Assessment and learning* (2 uppl.) (s. 233–242). London: Sage Publications.
- Strömquist, S. (2007). *Skrivprocessen: teori och tillämpning* (3 uppl.). Lund: Studentlitteratur.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14 (3), 281–294.

7. Enkel/utvecklad/ välutvecklad elevtext i biologi: Hur bedömer vi det?

JENNY WIKSTEN FOLKERYD & ÅSA AF GEIJERSTAM

Att skriva olika typer av texter är en återkommande aktivitet i många, för att inte säga alla, olika skolämnen. Skrivandet kan gälla allt från korta svar på frågor till långa självbärande texter, det kan genomföras spontant och under en kortare period eller väl planerat under en längre tid (jfr kapitel 3). Traditionellt har bedömningen av elevers skrivande framför allt skett i svenskämnen och i andra språkämnen, medan elevers texter i övriga ämnen i första hand har betraktats som nedteknad kunskap, eller »fakta«, och bedömts utan att ta hänsyn till hur innehållet skapas genom språket. Med den kunskap som finns från forskning om språk i olika verksamheter och ämnen, och med de styrdokument som gäller i Sverige i dag, behöver vi dock omvärdera hur elevtexter i andra ämnen än de språkliga såväl undervisas som bedöms. Det räcker inte med att bedöma ämnesinnehåll som om det vore skilt från hur det språkligt är uttryckt. I detta kapitel kommer vi därför att närmare diskutera frågan om bedömning av elevtexter i andra ämnen än svenskämnen med exempel från elevtexter i biologi i årskurs 8.

I internationell forskning har olika aspekter av det så kallade ämnesspråket diskuterats under de senaste årtiondena. Här har man bland annat konstaterat att en elev för att lära sig ett ämnesområde även behöver lära sig detta områdes sätt att använda språket (se t.ex. Schleppegrell, 2004). Även på nationell mark har det senaste

årtiondet inneburit en hel del forskning om ämnesspråk. Språkliga aspekter i ett flertal olika skolämnen (som geografi, samhällskunskap, biologi, kemi och slöjd) har studerats (Eriksson, 2010; af Geijerstam, 2006; Lindqvist, kommande; Nygård Larsson, 2011; Sellgren, 2011). Dessutom har såväl Skolverket som kommuner satsat på fortbildning inom områden som gått under benämningar som »språk i alla ämnen«, »ämnesspråk« och »språkutvecklande arbetssätt i alla ämnen«. Grundtankarna i dessa fortbildningsinsatser är att språket är centralt i alla ämnen, och att alla lärare därmed behöver vara bekanta med det sätt på vilket språket används i det egna ämnet. Det handlar då om mer än enbart vilka ord som används (de ämnesspråkliga begreppen) och omfattar också sådant som till exempel texters övergripande struktur, hur man samtalar i ämnet och på vilket sätt man tänker sig att texter inom ämnet ska läsas och förstås.

Det råder alltså ingen tvekan om att frågor om språkanvändning är centrala i alla skolämnen. Det innehåll som vi vill att eleverna ska tillgodogöra sig både kan och ska alltså enligt läroplanen bedömas genom sådant som elevers skriftliga och muntliga produktion. Det betyder naturligtvis inte att språklig produktion är det *enda* sättet att dokumentera och bedöma kunskap. I många ämnen står även annat görande i centrum (t.ex. slöjdandet i textilslöjd eller den fysiska aktiviteten i idrott och hälsa). Det vi kan konstatera mot bakgrund av forskning och styrdokument är dock att språklig produktion är en central aspekt för bedömning i alla skolämnen.

I kunskapskraven som formuleras i Lgr 11 för de olika betygen återkommer i flera ämnen liknande formuleringar som skiljer de olika betygsstegen åt. Formuleringar i kunskapskraven i biologi som »enkla (betyg E)/utvecklade (betyg C)/välutvecklade (betyg A) texter och andra framställningar med viss/relativt god/god anpassning till syfte och målgrupp« finns även i motsvarande formuleringar i andra ämnen. Dessa generella definitioner kan vara svåra för eleven att konkretisera i sitt eget arbete (jfr kapitel 8). Det kan även vara svårt för läraren som ska bedöma elevens texter

och andra framställningar, samt som förväntas ge återkoppling till eleverna om vad som kan utvecklas för att nå ett högre betyg kommande terminer. Återkopplingen till eleverna kan naturligtvis delvis ske genom att visa exempel på texter som motsvarar de olika betygen. För att i än högre grad konkretisera vad som kan vara skillnaden mellan exempelvis en enkel, en utvecklad respektive en välutvecklad text inom ett visst ämne vill vi dock gå längre i konkretiseringen, och ge förslag på språkvetenskapliga verktyg som på ett relativt enkelt sätt kan användas för att bedöma elevers texter i dessa avseenden. Eftersom det centrala innehållet i ett ämne som biologi är tydligt innehållsligt relaterat (det handlar till exempel om att »identifiera, sortera och gruppera organismer efter deras släktskap«) är det också centralt att de språkvetenskapliga verktyg som används för bedömning i första hand fångar innehållsliga aspekter i texterna.

Vi vill understryka att denna text bör läsas som ett förslag på hur man kan bedöma skrivuppgifter i många olika skolämnen. Biologi fungerar här som ett exempel, men resonemanget kan enkelt överföras även till andra ämnen. Överförbarheten påverkas inte minst av att formuleringarna i kunskapskraven återkommer mellan olika ämnen. Vi vill dock påpeka att även om modellen för att bedöma texter kan överföras mellan olika ämnen så innebär detta inte att elevernas förmåga att skriva dessa texter enkelt transporteras från ett ämne till ett annat, eller ens från en skrivsituation till en annan. Varje text är situerad i ett visst sammanhang, skrivs med ett visst syfte och med ett visst innehåll.

Bedömningens syfte

I kursplanen för biologi i grundskolan presenteras inledningsvis ämnets övergripande syfte (Skolverket, 2010). Sammanfattningsvis ska eleverna ges förutsättningar för att utveckla sin förmåga att:

- Använda kunskaper i biologi för att granska information,

kommunicera och ta ställning i frågor som rör hälsa, naturbruk och ekologisk hållbarhet

- Genomföra systematiska undersökningar i biologi
- Använda biologins begrepp, modeller och teorier för att beskriva och förklara biologiska samband i människokroppen, naturen och samhället.

Språkliga aspekter är väsentliga för samtliga syften som formuleras inom ämnet biologi (såväl som inom andra ämnen). Vi har dock valt att exemplifiera vårt resonemang genom den tredje syftespunkten ovan och med hjälp av språkvetenskapliga verktyg visa hur förmågan att använda *begrepp* samt att *beskriva* och *förklara* kan bedömas.

Syftespunkterna i kursplanen konkretiseras i det centrala innehållet för ämnet vilket sammanfattas under fyra huvudområden, eller måldomäner: *natur och samhälle, kropp och hälsa, biologin och världsbilden* samt *biologins metoder och arbetsätt*.

Här har vi valt att närmare undersöka det sistnämnda området och mer specifikt hur syftet att *använda biologins begrepp, modeller och teorier för att beskriva och förklara biologiska samband i människokroppen, naturen och samhället* kan bedömas just när det gäller subdomänen som handlar om hur organismer identifieras, sorteras och grupperas utifrån släktskap och utveckling. De exempel som används i kapitlet är autentiska texter skrivna av elever i årskurs 8. Som ingång till uppgiften gjordes en gemensam genomgång av blötdjur, då läraren också antecknade på tavlan. Denna taveltext om blötdjur användes också som inledning i alla elevers texter. Uppgiften bestod sedan i att läsa om olika typer av blötdjur och självständigt, och med biologins begrepp, redovisa informationen i skrift. Någon mer detaljerad information om hur texten skulle skrivas gavs varken muntligt eller skriftligt. Två av de texter som skrevs används här för att visa vilket bedömningsresonemang som kan föras utifrån dessa. Texterna skiljer sig åt på flera sätt och är därför intressanta att diskutera utifrån ett bedömningsperspektiv.

ELEVTEXT 1.

Snäckor

Dom kan leva i vatten och på land.

Vid västkusten är strandsnäcken vanligast.

Vanligast i sötvatten är skivsnäkan och dammsnäcken.

Vinbärgessnäcken är ätlig.

Musslor

Vissa musslor är ätliga.

Det kan bildas pärlor på skalen av Sandkorn.

Musslor lever på smådjur, bakterier.

Bläckfisk

Det finns två sorters bläckfiskar

8armad och 10armad.

Den 10armade bläckfisken har världens största ögon

ELEVTEXT 2.

Musslor: Musslorna är skyddade av två skinnhalvor som hålls ihop av ett gångjärn och av väldigt starka muskler. På några musslor täcks insidan av pärlemor. Där inne kan en pärla bildas av till exempel ett sandkorn. I Japan odlas pärlor på konstgjord väg.

Snäckor: Dom har ofta ett spiralvridet skal, som stängs med ett lock. Snäckorna var troligen bland dom första organismer som började ta sig upp på land. Med sin taggiga tunga raspar dom i sig växtdelar.

Bläckfiskar: Det finns 8armad och 10armad bläckfisk. På armarna sitter det sugploppar.

Armarna kan bli 15 långa.

Det atlantiska jättebläckfisken är 10armad och den kan bli 17 meter lång och väga två ton. Den har världens största ögon som kan bli 40 cm tvärsöver.

I det följande kommer dessa elevtexter att kommenteras utifrån språkliga aspekter som i forskning visat sig vara speciellt viktiga för att visa hur information kan presenteras så att en läsare lättast tar till sig av den. Det har också visat sig att elever som på ett tydligt sätt kan presentera information själva i högre grad har förstått innehållet (af Geijerstam 2006; Pretorius, 1996). De kunskapskrav i biologi som vi speciellt fokuserar är följande:

- Hur eleven använder informationen på ett i huvudsak/relativt väl/väl fungerande sätt i diskussioner och för att skapa enkla/utvecklade/välutvecklade texter och andra framställningar med viss/relativt god/god anpassning till syfte och målgrupp.

Dimensioner av bedömning

En av de vanligaste aspekter som diskuteras då det gäller språk-användning i de naturvetenskapliga ämnena är *ordförrådet*. Att lära sig förstå och använda ämnesspecifika begrepp betraktas som en central del av att lära sig ett naturvetenskapligt ämne. En stor del av den forskning som behandlat elevers språk-användning i naturvetenskap har också handlat just om begrepps-bildning, och man har till exempel kontrasterat elevers vardagliga förståelse av begrepp med den vetenskapliga betydelsen av samma begrepp (se t.ex. Andersson & Bach, 2005; Wallin, 2004; Wickman & Östman, 2002). Det råder heller ingen tvekan om att ordförrådets storlek är avgörande för hur väl en läsare kan förstå och ta till sig innehållet i till exempel en läromedelstext. Internationell forskning visar till exempel att åtminstone 95 procent av orden behöver vara kända för en läsare för att hon eller han ska kunna ta till sig innehållet i en text (Nation, 2001). Det finns alltså starkt stöd för att låta elevernas ämnesspecifika begreppskunskap vara en central del av innehållet i undervisningen och därmed också av bedömningen i ett ämne. Ett sätt att bedöma begreppskunskapen är att granska elevers texter med avseende på deras ordförråd. Genom att undersöka elevernas

egen produktion snarare än att undersöka förståelsen genom någon form av »begreppstest« (liknande ett glosförhör), får man som lärare en uppfattning om det aktiva ordförrådet hos eleven snarare än om det passiva. För att kunna bedöma en text med avseende på dess ordförråd behöver man dock ta ställning till textens syfte och målgrupp. I till exempel en argumenterande text som är skriven som en insändare i en ungdomstidning (vi kan tänka oss en text för eller emot byggandet av vindkraftverk i närheten av skolan) kan man förvänta sig färre ämnesspecifika ord än i en beskrivande text där eleven redogör för hur vindkraft fungerar. Därför måste bedömningen av förekomsten av sådant som vardagliga eller ämnesspecifika ord i en text alltid vara relaterad till textens syfte och målgrupp.

En annan aspekt av texten som är central för att avgöra såväl dess grad av utveckling som dess anpassning till syfte och målgrupp är *textens genrestuktur*, alltså hur den hänger samman i sina olika delar. Även här måste textens struktur sättas i relation till dess syfte. För att återgå till exemplet om vindkraft så behöver insändaren som argumenterar för eller emot att bygga vindkraftverk i närheten av skolan ha en annan struktur än elevens redogörelse för hur vindkraftverket fungerar. Som lärare är det här alltså viktigt att vara klar över vilken slags text man förväntar sig att eleverna ska skriva, för att kunna bedöma resultatet.

För att hitta stöd för hur texter i naturvetenskapliga ämnen kan betraktas och bedömas kan inspiration hämtas från australiensisk genreforskning. Här har man bland annat diskuterat en utveckling från att initialt kunna hantera texter som används för handlingar och göranden i naturvetenskapen, till att använda språket för att förklara och dokumentera på olika sätt och slutligen till att kunna hantera språket för att utmana vetenskapen genom diskussioner och argumentationer (Veel, 1997). Här blir det intressant att relatera detta synsätt på språkanvändning i naturvetenskap till den svenska kursplanens syfte att »beskriva och förklara biologiska samband«. Kursplanens syften kan här sägas motsvara Veels genrer

som innebär att organisera vetenskaplig information samt att ge vetenskapliga förklaringar till olika händelser. Som vi ser det kan man för att bedöma detta ha stor hjälp av den kunskap som inom språkvetenskapen finns om beskrivande och förklarande texter.

För att kunna avgöra hur utvecklad en text är, vilket är en av de dimensioner som skiljer betygsstegen åt, måste man också ta hänsyn till hur de innehållsliga teman som diskuteras är strukturerade och organiserade i förhållande till varandra. Detta handlar alltså om textens innehållsliga struktur och kan diskuteras i termer av makro- och mikroteman. Med makrotema menas här huvudämnet för hela eller delar av texten och med mikrotema menas ett undertema eller en speciell aspekt av huvudämnet (Folkeryd, 2014; Hellspong & Ledin, 1997). En text som innehåller många mikroteman som inte kopplas till huvudämnet kan uppfattas som fragmentarisk och svår att förstå. Sådana texter blir ofta mer listliknande, där mycket ny information av olika slag presenteras utan att fördjupas. En text där olika mikroteman tydligt kopplas till huvudämnet skapar däremot en tydlig röd tråd för läsaren att följa. När det gäller att avgöra hur utvecklad en text är utgör sådana aspekter av bedömning därför viktiga inslag.

För att en läsare ska kunna bygga upp en sammanhängande förståelse för innehållet i texten är det också viktigt att innehållet binds samman inom och mellan meningar. Det handlar exempelvis om att användningen av ord som *och*, *då* eller *därför* gör det lättare för den som läser att förstå hur idéer och information i texten hänger ihop. Genom att olika typer av sådana *utvidgande språkliga markörer* används kan man också säga att innehållet fördjupas. Hur väl en text är utvecklad kan man därför bedöma genom att exempelvis ta hänsyn till:

- Hur två enskilda men relaterade idéer fogas samman (*tillägg* med hjälp av t.ex. orden *och*, *men*, *utan*, *fast*, *som*)
- Hur informationen specificeras genom angivelser av tid, plats och anledning (*specificering* med hjälp av t.ex. orden *för*, *eftersom*, *så*, *om*, *då*, *när*)

- Hur idéer kan definieras eller klargöras (*utveckling* med hjälp t.ex. av orden *till exempel, alltså*) (af Geijerstam, 2006; Halliday & Matthiessen, 2004; Holmberg & Karlsson, 2006).

Sammanfattningsvis bedöms alltså texterna utifrån följande dimensioner:

- a. Ordförråd
- b. Genrestruktur
- c. Innehållslig struktur
- d. Utvidgande språkliga markörer

Dimensioner av bedömning – textexempel

De språkliga aspekter som vi har presenterat ovan kan alltså sammanfattas som fyra olika dimensioner som är viktiga för bedömning av innehåll. I detta avsnitt visar vi hur de två elevtexter som presenterades tidigare i kapitlet kan bedömas med avseende på dessa dimensioner.

A. ORDFÖRRÅD

Vad gäller ordförrådet kan vi se att elevtext 1 innehåller i stort sett uteslutande ämnesspecifika ord för biologi respektive ord som oftast används i skolans sammanhang. Texten domineras helt av artnamn (*strandsnäcka, dammsnäcka, musslor, 10armad bläcfisk* etc.). Till viss del finns även mer allmänt skolspråkliga uttryck som *att leva i vatten och på land, ätlig* och *smådjur*. Några bärande uttryck som är hämtade utanför skolan återfinns inte i texten. Elevtext 2 innehåller också många ämnesspecifika begrepp, med relativt många artnamn. Här finns dock en variation i vad de ämnesspecifika orden handlar om, och vi ser ord som *skinnhalvor, spiralvridet skal* och *växtdelar*. Förutom detta finns ett antal mer allmänt skolspråkliga ord som *bildas* och *odla pärlor*. Det som framför allt skiljer text 1 och 2 åt vad

gäller ordförrådet är att text 2 även innefattar ett par mer vardagligt präglade ord, som *gångjärn*, *lock* och *tvärsöver*.

Här ställs man som lärare inför att avgöra vad som är önskvärt i denna typ av text. Är en text som innefattar en variation av ord »bättre« än en text som endast innefattar ämnes- och skolspråkliga ord? Hur mycket vardagsspråk är i så fall lagom? Ett direkt svar på detta är naturligtvis omöjligt att ge. Det man med grund i tidigare forskning kan säga är däremot att en elev som skriver en text där hon eller han varierar mellan läromedlens eller taveltextens begreppsanvändning och den egna vardagsspråkliga vokabulären, troligen har tagit till sig innehållet i högre grad än en elev som endast använder begrepp direkt hämtade ur läromedlen eller lärarens tavelanteckningar. I tidigare studier har man nämligen kunnat se att elevtexter som har en mycket hög grad av tekniska termer (till exempel på grund av många ämnesspråkliga ord) också är svår att förstå för den elev som skrivit den (af Geijerstam, 2006).

En annan aspekt av ordförrådet som är väsentlig att ta hänsyn till i bedömningssammanhang handlar om hur väl det är anpassat till syfte och målgrupp. Elevernas ålder är i detta fall en viktig faktor. Dessa texter är skrivna av elever i årskurs 8, som alltså är cirka 14 år gamla. Även om målgruppen för denna textuppgift inte tydliggjordes som »andra 14-åringar«, så är det rimligt att relationen mellan typerna av begrepp förändras med elevernas ökande ålder. Andelen ämnesspråkliga och vardagsspråkliga begrepp i en biologitext skriven av en 10-åring kan inte förväntas vara densamma som i en text skriven av en elev i gymnasiet, som snart ska gå ut det naturvetenskapliga programmet. Med detta resonemang som grund är det alltså inte orimligt att en beskrivande text om blötdjur i årskurs 8 innefattar en viss andel vardagsspråkliga ord.

B. GENRESTRUKTUR

Elevtexterna som vi använder som exempel i detta kapitel är tacksamma att analysera vad gäller deras struktur. De är relativt

enkla på så sätt att de är typiska »beskrivande rapporter« i Veels terminologi (Veel, 1997), där en art först nämns och sedan beskrivs på olika sätt. Alla texter som skrivs i skolan är inte så enkla i sin uppbyggnad, och de är därmed inte heller lika enkla att bedöma vad gäller övergripande struktur.

Strukturen i våra båda texter är relativt liknande. Båda inleder varje nytt stycke med en rubrik: »Musslor«, »Snäckor« respektive »Bläckfiskar«. Varje rubrik fungerar som ett generellt påstående som följs av en beskrivning av olika typiska drag för den art som ska beskrivas. En skillnad mellan texterna är dock att de beskrivande stegen i text 1 är relativt kortfattade, medan de i text 2 är mer utbyggda.

C. INNEHÅLLSLIG STRUKTUR

När det gäller den innehållsliga strukturen i elevtext 1, så ser vi att det huvudämne som uttrycks i rubriken (Snäckor) kommenteras i tre mikroteman. Vi får reda på någonting om boplats (*Dom kan leva i vatten och på land*), förekomst (*Vid västkusten är strandsnäckan vanligast. Vanligast i sötvatten är skivsnäkan och dammsnäckan.*) och ätbarhet (*Vinbärgessnäckan är ätlig.*). Endast när det gäller ett av mikroteman får vi utökad information, i detta fall när det gäller vilken typ av snäcka som är vanligast förekommande i olika sammanhang.

Det andra huvudämnet behandlar musslor och också här möter vi informationen organiserad i tre mikroteman. Det handlar om ätbarhet (*Vissa musslor är ätliga.*), om pärlbildning (*Det kan bildas pärlor på skalen av sandkorn.*) och om födan för musslor (*Musslor lever på smådjur, bakterier.*). Till skillnad från det första makrotemat om snäckor, ser vi här endast ett påstående per undertema. Informationen beskrivs eller förklaras inte vidare vilket snarast ger intrycket av en listliknande presentation om musslor.

Även det tredje huvudämnet uttrycks i rubriken, där ämnet bläckfiskar pekats ut. I avsnittet innehåller båda meningarna in-

formation som berör ett undertema om olika typer av bläckfiskar. Men eftersom makroteman och mikroteman är relativa begrepp, skulle här den första meningen (*Det finns två sorters bläckfiskar 8armad och 10armad.*) kunna sägas skapa ett nytt makrotema om just den 10armade bläckfisken. Mikrotemat till detta nya makrotema utgörs då av den andra meningen (*Den 10armade bläckfisken har världens största ögon.*). Avsnittet om bläckfiskar är på så sätt mer innehållsligt sammanhängande än avsnitten om snäckor och musslor. Det är dock inte många olika saker som vi får reda på om bläckfisken.

När det gäller den innehållsliga organisationen i elevtext 2 så ser vi att det första temat handlar om musslor. I huvudsak beskrivs sedan musslor utifrån två mikroteman som handlar om musslans konstruktion (*Musslorna är skyddade av två skinnhalvor som hålls ihop av ett gångjärn och av väldigt starka muskler.*) och pärlbildning i musslor (*På några musslor täcks insidan av pärlemor. Där inne kan en pärla bildas av till exempel ett sandkorn. I Japan odlas pärlor på konstgjord väg.*). Den största skillnaden i jämförelse med elevtext 1 är att textens innehåll här i större utsträckning fördjupas genom att mer information ges om ett av textens mikroteman.

Det andra huvudämnet handlar om snäckor och här organiseras informationen precis som i elevtext 1 i tre underteman. Vi får reda på något om skalet (*Dom har ofta ett spiralvridet skal, som stängs med ett lock.*), något om snäckans historik (*Snäckorna var troligen bland dom första organismer som började ta sig upp på land.*) och något om snäckans föda (*Med sin taggiga tunga raspar dom i sig växtdelar.*)

Det tredje huvudämnet behandlar bläckfiskar och precis som i elevtext 1 är samtliga meningar ordnade under ett undertema, i detta fall typer av bläckfiskar. Till skillnad från elevtext 1 får vi dock något mer information om bläckfiskar och det skapas också ett nytt makrotema (*10armad bläckfisk*) som i högre grad byggs ut genom tilläggsinformation (*Det atlantiska jättebläckfisken är 10armad och den kan bli 17 meter lång och väga två ton. Den har världens största ögon som kan bli 40 cm tvärsöver.*).

D. UTVIDGANDE SPRÅKLIGA MARKÖRER

En skillnad mellan de två elevtexterna finner vi dock inte bara när det gäller hur informationen ordnas i makroteman och mikroteman, utan också när det gäller på vilka sätt informationen byggs ut och fördjupas. I elevtext 1 ser vi i de flesta meningarna inga utvidgningar. När sådana förekommer handlar det uteslutande om tillägg av information med hjälp av ordet »och« (*Dom kan leva i vatten och på land.*). I något fall är det också så att den explicita sammanbindningsmarkören utelämnas (*Musslor lever på smådjur, bakterier*). Detta bidrar ytterligare till det mer listliknande intrycket av den texten.

I elevtext 2 sammanlänkas informationen i betydligt större utsträckning. Även här handlar det i de flesta fallen om utvidgning genom tillägg av information. De sammanlänkande orden är kursiverade i följande exempelmeningar:

- Musslorna är skyddade av två skinnhalvor *som* hålls ihop av ett gångjärn *och* av väldigt starka muskler.
- Dom har ofta ett spiralvridet skal, *som* stängs med ett lock.
- Snäckorna var troligen bland dom första organismer *som* började ta sig upp på land.
- Det atlantiska jättebläckfisken är 10armad *och* den kan bli 17 meter lång *och* väga två ton.
- Den har världens största ögon *som* kan bli 40 cm tvärsöver.

I dessa fall ser vi hur en ny sats lägger till ny information till föregående sats, vilket innebär en rikare beskrivning av mikrotemat än vad vi ser i exempeltext 1.

Sammanfattningsvis kan en bedömning av elevtext 1 i relation till kunskapskraven visa att den är en enkel text med viss anpassning till syfte och målgrupp, eftersom:

- a. Texten i hög grad innehåller skolspråkliga och ämnesspecifika ord. Variationen mellan vad orden hänvisar till är dock liten,

och ingen uppblandning med vardagsspråkliga ord finns, vilket man hade kunnat förvänta sig i denna typ av text i denna åldersgrupp.

- b. Texten har en struktur som lämpar sig för dess syfte. Den inleds med en rubrik som talar om vad som ska beskrivas, och efter det följer beskrivningar av just detta. Dock utvecklas dessa beskrivningar i låg grad.
- c. Alla innehållsteman berör på något sätt textens huvudämne (exempelvis om musslor). Det är dock endast vid ett par tillfällen som mikrotemat fördjupas eller breddas genom flera påståenden eller beskrivningar av huvudämnet.
- d. De påståenden eller beskrivningar som ges i texten oftast inte utvidgas, för att på så sätt skapa sammanhang och bidra till fördjupad information. När det finns utvidgningar är de i form av tilläggsinformation som uttrycks med hjälp av ordet »och«.

Elevtext 2 är i stället en utvecklad text med relativt god anpassning till syfte och målgrupp eftersom:

- a. Texten innehåller såväl skolspråkliga, ämnesspecifika och vardagsspråkliga ord. Med de vardagsspråkliga begreppen som *gångjärn*, *lock* och *sugploppar* lyckas texten bra med att åskådliggöra sin beskrivning av blötdjuren.
- b. Texten har en struktur som lämpar sig för dess syfte. Varje stycke inleds med en rubrik som sedan följs av beskrivningar av det som ska beskrivas. De beskrivande delarna i texten är relativt utvecklade.
- c. Alla innehållsteman berör textens huvudämne och är tydligt ordnad i mikroteman. Den innehållsliga strukturen fördjupas också i och med att flera påståenden ges om ett mikrotema.
- d. De påståenden eller beskrivningar som ges i texten ofta utvidgas, och därmed fördjupar den information som ges om huvudämnet. Viss variation finns när det gäller vilken typ av

utvidgning som görs, men ofta handlar det om sammanlänkade satser genom användning av orden *och* samt *som*.

För att texterna ska nå ett ännu högre betyg behöver ordförrådet omfatta såväl vardagsspråkliga som mer ämnesspecifika ord, där variationen skulle ge en rikare beskrivande text anpassad efter syfte och mottagare. Genrestrukturen i båda texterna kan stärkas genom att de olika leden i texten utvecklas mer. Det kan göras dels genom att varje innehållsdel får en mer utförlig introduktion (nu är det endast rubriken som introducerar varje ny typ av blötdjur), dels genom att de drag som beskrivs får en något längre beskrivning. På så sätt kan man säga att textens olika genresteg utvecklas.

När det gäller innehållslig struktur och användningen av utvidgande språkliga markörer skulle ett högre betygssteg ha kunnat nås om varje huvudämne hade utvecklats med hjälp av fler fördjupade mikroteman. En mer välutvecklad text skulle med fördel också ha innehållit fler utvidgade satser och framför allt satser som på olika sätt byggt ut informationen. Med det menas exempelvis att informationen inte bara byggs ut med tillägg utan också med specificeringar (exempelvis med hjälp av *för att*, *eftersom*) och utvecklingar (exempelvis med hjälp av definitioner och omformuleringar).

Ett generaliserat omdöme

Det vi har exemplifierat ovan handlar om en skrivuppgift i biologi. Vi har också visat på en mall som kan användas så att lärare som bedömer på ett systematiskt sätt ska kunna ta ställning till aspekter som visat sig väsentliga för bedömning i biologi. I enlighet med Kanes modell bör läraren dessutom kunna argumentera för att bedömningen på den enskilda skrivuppgiften är generaliserbar (jfr kapitel 10). För att kunna dra den sortens slutsatser måste läraren ta hänsyn till liknande skrivuppgifter som bedömts vid andra tillfällen eller som bedömts av annan lärare. Ett generaliserat skrivresultat kan sedan användas för att säga något om elevens förmåga

i måldomänen. Måldomänen utgörs i det här fallet av förmågan att skapa texter och andra framställningar, med anpassning till syfte och målgrupp när det gäller att använda biologins begrepp, modeller och teorier, för att beskriva och förklara biologiska samband i människokroppen, naturen och samhället.

Förutom att omdömen om elevernas förmåga inom en domän exempelvis kan leda till ett terminsbetyg, kan de också ligga till grund för andra beslut. I fallet med elevtexterna som analyserats i detta kapitel såg vi exempelvis att ingen av texterna bedömdes vara välutvecklad. När samtliga texter i klassen analyserades sågs samma resultat. Ett tänkbart beslut från lärarens sida skulle då kunna vara att ändra sin undervisning. Inför skrivuppgiften fick eleverna en omfattande muntlig genomgång av ämnet, men inga specifika instruktioner gavs om hur informationen i just denna typ av text skulle kunna presenteras, struktureras och fördjupas med hjälp av olika språkliga medel.

För att elever ska kunna skriva texter med de kvaliteter som formuleras i kunskapskraven behöver förutsättningar för detta ges i undervisningen. Eleverna måste få hjälp med hur texter skrivs inom olika ämnesområden, exempelvis med hjälp av modelltexter och annan explicit undervisning om skrivande i ämnet. De måste få möta det ämnesspecifika ordförrådet, hur texter byggs upp i innehåll och form och hur information kan fördjupas och byggas ut. Samtliga av dessa aspekter måste också relateras till det övergripande syftet med uppgiften och vem som är mottagare.

Vidare måste lärarna, för att kunna bedöma elevernas texter, få tillgång till konkreta verktyg som kan hjälpa dem att avgöra om texterna till exempel är enkla/utvecklade eller välutvecklade. Med tanke på att liknande formuleringar återkommer i kunskapskraven i många ämnen kan de språkvetenskapliga verktyg som här presenterats fungera inte bara inom biologiämnet utan också inom andra ämnen. Ämnesinnehållet, ett ämnesspecifikt ordförråd, uppgifternas syfte och texternas struktur varierar mellan olika ämnen, men de verktyg som presenterats i detta kapitel kan användas oavsett skolämne.

Referenser

- Andersson, B. & Bach, F. (2005). On designing and evaluating teaching sequences taking geometrical optics as an example. *Science Education*, 89(2), 196–218.
- Eriksson, I. (2010). *Innehållet i fokus – kemiundervisning i finlandssvenska klassrum* (vol. 8). Stockholm: Stockholms universitets förlag.
- Folkeryd, J.W. (2014). Hjärtan, hjul och hävstånger – Innehåll i elevernas sakprosatexter på lågstadiet. I *Mångfaldens möjligheter – litteratur och språkdidaktik i Norden*. Elfte nationella konferensen i svenska med didaktisk inriktning. Göteborg.
- af Geijerstam, Å. (2006). *Att skriva i naturorienterande ämnen i skolan*. Uppsala: Uppsala universitet.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2004). *An introduction to functional grammar* (3 uppl.). London: Arnold.
- Hellspong, L. & Ledin, P. (1997). *Vägar genom texten: handbok i brukstextanalys*. Lund: Studentlitteratur.
- Holmberg, P. & Karlsson, A.-M. (2006). *Grammatik med betydelse: en introduktion till funktionell grammatik*. Uppsala: Hallgren & Fallgren.
- Lindqvist, E. (2014). *Litteracitetsaktiviteter i slöjd och samhällskunskap – Fallstudier av flerspråkiga elevers textanvändning*. Stockholm: Institutionen för språkdidaktik.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nygård Larsson, P. (2011). *Biologiämnets texter. Text, språk och lärande i en språkligt heterogen gymnasieklass*. Lund: Lunds universitet.
- Pretorius, E.J. (1996). A profile of causal development amongst ten-year-olds: Implications for reading and writing. *Reading and Writing: An interdisciplinary Journal*, 8, 385–406.
- Schleppegrell, M.J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, N.J. Routledge.
- Sellgren, M. (2011). *Den dubbla uppgiften. Tvåspråkiga elever i skolans mellanår arbetar med förklarande genre i SO*. Stockholm: Stockholms universitet.
- Skolverket (2010). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011*. Stockholm: Skolverket.

- Veel, R. (1997). Learning how to mean – scientifically speaking: Apprenticeship into scientific discourse in the secondary school. I F. Christie & J. Martin (red.), *Genre and institutions: Social processes in the workplace and school* (s. 161–196). New York: Continuum Intl Pub Group.
- Wallin, A. (2004). *Evolutionsteorin i klassrummet. På väg mot en ämnesdidaktisk teori för undervisning i biologisk evolution*. Göteborg studies in Educational Sciences 212. Göteborg: Acta Universitatis Gothoburgensis.
- Wickman, P.-O. & Östman, L. (2002). Learning as Discourse Change: A Sociocultural Mechanism. *Science Education*, 86(5), 601–623.

8. Bedömning av litterär analys i skolan

EVA NILSON

Bedömningens syfte och bakgrund

Att läsa litterära texter och att skriva om sitt bokläsande är en vanlig aktivitet i skolan, och en av grundstenarna i ämnet svenska. Bokrapporten, och bokrecensionen, är återkommande uppgifter både i grundskolan och på gymnasiet. Man blir tilldelad en bok i skolan som man läser under ett antal veckor hemma, för att sedan skriva en rapport om den i skolan. Den uppgift, eller undervisningsprocess, som här ska presenteras, utgår ifrån att läsa och skriva om det lästa, men på ett lite mer stöttande sätt än så. Det finns flera syften med den uppgift jag här ska beskriva, och jag vill därför ta upp några av de didaktiska implikationer som jag hoppas kunna lyfta med detta undervisningsupplägg.

Att läsa en hel roman tar mycket tid i anspråk och är därför svårt att hinna med på lektionstid. En vanlig didaktisk situation är att vi som lärare introducerar en skönlitterär bok för att sedan förutsätta att eleven läser densamma hemma under de kommande veckorna. Boken förväntas vara utläst vid ett specifikt datum då själva redovisningen, ofta i form av en rapport, samtal eller en recension ska ske. Vi lämnar alltså eleven med en bok att läsa utanför skolan, men examinerar läsförståelsen i klassrummet. I litteraturutredningen *Läsandets kultur* (SOU 2012, s. 89), uppmärksammas just hur elever hänvisas att läsa hemma, men att lässvaga elever inte alltid gör det.

Ett annat didaktiskt problem som berör litteraturläsandet, och

som också har uppmärksammats inom läsforskningen, är att det vid examination av litterär läsning är vanligt att elevernas samtal och reflektioner hamnar långt ifrån de texter som ska diskuteras. Omvärldskunskapen, i stället för den litterära texten, kommer i fokus (Mossberg-Schüllerqvist & Olin-Scheller, 2011). Detta kan ha många orsaker, men en kan vara att eleven inte har märkt att att läsa hela texten och försöker leda samtalet ut från texten, i stället för in i den.

Många elever har också svårigheter att läsa en skönlitterär text utifrån ett estetiskt perspektiv, och läser den i stället som en direkt återspeglning av verkligheten. Att elever ifrågasätter personer och händelser i skönlitterärt berättande som om de vore verkliga, i stället för fiktiva, är ett vanligt inslag i litteraturundervisning. Den litterära texten läses *faktivt* i stället för *fiktivt*.¹ Också för läraren kan det ibland vara oklart, vid examination av läsförståelse och förmågan att analysera litteratur, vad det är vi examinerar. Det ibland otydliga lärandeobjektet vid läsning av just den litterära texten gör att dialogen om vad som examineras blir otydlig för eleven, och ibland även för läraren (Tengberg, 2011).

Ser vi till de kunskapskrav som lärare ska tydliggöra och konkretisera för eleven vid betygssättning i årskurs 9 är dessa formulerade i ganska vaga ordval. Ska kunskapskrav bli begripliga för eleven måste krav och begrepp bli mer konkreta och relaterade till enskilda uppgifter. Att göra betygssystemet så transparent som möjligt, med kriterierelaterade kunskapskrav, var också själva idén med det betygssystem som infördes med Lpo 94, och som Lgr 11 bygger vidare på.

Ovanstående är alltså områden som vi som lärare har att utmana och förtydliga för eleven genom de litteraturdidaktiska val vi gör utifrån läroplanen. Min utgångspunkt för detta kapitel blir att visa hur man kan arbeta för ökad förståelse av den litterära texten,

¹ För en mer fördjupad diskussion om faktiv och fiktiv läsning se Bo Steffensen, *När barn läser fiktion* (2005).

i det här fallet en hel roman, där innehåll och form kan ses som varandras förutsättningar i tolkningen av texten.

Syfte, mål, centralt innehåll, kunskapskrav

Utgångspunkten för det undervisningsupplägg som presenteras i det här kapitlet är således att ge exempel på hur man kan undervisa i litterär läsning och hur detta sedan kan examineras på ett för eleven så transparent sätt som möjligt. Eleverna i exemplet har i uppgift att läsa en roman som de, under tiden de läser, ska skriva en reflekterande och analyserande text om. I kursplanen i svenska står under rubriken »Syfte« att eleven ska »ges förutsättningar att utveckla sin förmåga att 'formulera sig och kommunicera i tal och skrift' samt 'läsa och analysera skönlitteratur och andra texter för olika syften'« (Skolverket, 2011, s. 222).

Den tänkta uppgiften som beskrivs i det här kapitlet utgår ifrån dessa två syften i kursplanen. I dess centrala innehåll listas även ett antal komponenter i den litterära texten som eleven ska uppmärksamma i sin litteraturläsning. Man lyfter fram att eleven ska kunna urskilja texters budskap, teman och motiv, samt se hur personer och miljö är beskrivna i litteraturen. Även språkliga drag, uppbyggnad och berättarperspektiv ska lyftas fram, samt berättelsemönster, och hur både inre och yttre dialoger fungerar. De litterära texterna ska belysa »människors villkor och identitets- och livsfrågor«, vilket blir ett sätt att få in omvärldskunskap och att kunna knyta denna till den egna erfarenheten. Litteraturen ska också kunna knytas till en historisk och kulturell miljö. Detta är aspekter som blir viktiga vid utformningen av undervisningen.

Eftersom vi som lärare vill att eleverna ska komma så nära texten som möjligt i sina analyser, kommer jag att ta med den del i kursplanens centrala innehåll där citatteknik och källhänvisningar kommer in. I det centrala innehållet i ämnet svenska lyfts det också fram att eleven ska kunna ta emot respons på sitt skrivande, vilket innebär att eleven ska bearbeta sin analys efter lärarens återkopp-

ling så att arbetet växer fram del för del till en hel analys av romanen i fråga.

I kunskapskraven för olika betygsnivåer i årskurs 9 presenteras en tänkt kunskapsmässig utveckling, om än genom generella definitioner. Genom den uppgift som presenteras i det här kapitlet blir flera olika kunskapskrav aktuella för bedömning. För att till exempel visa läsförståelse ska eleven kunna göra sammanfattningar av texters innehåll som ska kopplas till tidsaspekter och orsakssamband i berättelser. Eleven ska utifrån egna erfarenheter och kunskaper om omvärlden kunna föra resonemang om texters budskap. I sammanfattningar av verk ska eleven dessutom kunna använda ett ämnesrelaterat språk och ha en fungerande struktur i sina texter där citat- och källhänvisningar finns med. Eleven ska vidare kunna ta emot respons och bearbeta sina texter för att förtydliga innehållet i dem.

För att kunna gradera kunskapskraven används definitioner som »enkla«, »till viss del«, »i huvudsak fungerande« för betyget E, till »utvecklade«, »relativt utvecklade« för betyget C till »välutvecklade« för högsta betyg. De resonemang eleven ska kunna föra om de litterära verken kan på E-nivå handla om »tydligt framträdande budskap«, för betyget C om »tydligt framträdande och budskap som kan läsas mellan raderna« och för betyget A om »tydligt framträdande och budskap som kan läsas mellan raderna eller är dolda«.

De olika betygsstegens logik handlar om att gradera kvaliteter i elevers läsande. Här går man från »enkla« till »välutvecklade« sammanfattningar, och även det ämnesrelaterade språket kan vara »enkelt« till »välutvecklat«. När eleven använder sig av citatteknik kan denna vara »fungerande« till att bli »effektiv«. De olika betygsstegen, och kunskapskraven, kan dock bli alltför abstrakta för att eleven konkret ska kunna förstå vad som behöver förändras/förbättras för ett högre betyg. Även om eleven får en matris för att se vilka skillnader som betygsstegen står för bör läraren också arbeta konkret med gensvar där eleven får tillfälle att arbeta fram sina analyser genom den återkoppling som läraren ger.

Att arbeta både med en matris och exemplifiera och konkretisera genom elevens eget arbete kan vara ett sätt att lära ut det som sedan ska bedömas. Konsten att skriva en litterär analys är svår att lära ut endast teoretiskt, och undervisningen kan därför bli alltför abstrakt för att eleven ska kunna ta den till sig. Det handlar i stället om att stötta eleven i hans eller hennes egen bearbetning av sin text genom att dialogiskt ge kommentarer och frågor till analysen av texten.

Analysuppgiften

Den specifika uppgift som valts för att arbeta med ovannämnda centrala innehåll och kunskapskrav är att eleverna ska skriva ett antal reflektioner och analyser till en skönlitterär bok, här en roman.² Detta är en uppgift som kan läggas upp ganska fritt efter vad läraren vill att eleverna ska ta fasta på i sin analys och den går lätt att modifiera efter klass och nivå. Det finns en pedagogisk poäng med att eleverna ska läsa en längre text som underlag till sin litterära analys. Den långa texten kräver ett tålamod i läsandet som kan vara svårt att hitta själv för den läsovana eleven. Tiden det tar att läsa kan då samtidigt ge tid för eftertanke och reflektion över texten. Att låta arbetet med romanläsningen och analysarbetet ta ett antal veckor kan göra att undervisningen och läsandet kan upplevas som mindre fragmentariskt än vad som ibland blir fallet med korta texter och korta skrivuppgifter till det lästa. Till den här uppgiften får eleverna välja mellan tre olika ungdomsböcker, nämligen Johanna Nilssons *Hon går genom tavlan, ut ur bilden* (1996), Peter Nilssons *Hjärtans fröjd* (1992) och Sara Kadefors *Sandor slash Ida* (2001). Såväl *Hjärtans fröjd* som *Sandor slash Ida* har både flickor och pojkar som

² Detta är en uppgift som jag på intet sätt har hittat på själv. Flera varianter av uppgiften kan hittas på olika nätforum. Mitt bidrag är att använda den till att visa hur man kan konkretisera bedömningen av något så svårt som analys av litterär text i grundskolan. Uppgiften kan göras på många olika nivåer och kan med lätthet anpassas efter litteratur och svårighetsgrader för kursen. I min presentation har jag utgått ifrån det centrala innehållet och de kunskapskrav som finns för årskurs 7–9 i Lgr 11.

huvudpersoner och alla tre har ett lite ovanligt berättelsemönster med parallellhandlingar, tillbakablickar och framtidsscenarioer som förhoppningsvis kan visa eleverna på litteraturens möjligheter att ge perspektiv på upplevelsen av tid och rum. Texterna har några år på nacken, men är fortfarande flitigt använda både i grundskolans senare år och i gymnasiet första kurs eftersom tematiken i romanerna är fortsatt aktuell.

Hon går genom tavlan, ut ur bilden handlar om Hanna som är tio år och som blir mer och mer utfryst i klassen som hon går i. Den ensamhet och utsatthet som Hanna känner, liksom den gruppdynamik som presenteras i klassen, visar kamratgemenskap och skoltillvaro på en metanivå som kan passa bra att läsa i de övre klasserna i grundskolan eller i början på gymnasiet. Berättelsen om den unga flickan Hanna bryts av med ett antal nedslag där vi möter Hanna under en sjukhusvistelse ett antal år senare. Just skiftet mellan nutid och dåtid kan vålla problem för läsaren samtidigt som det ger en ytterligare dimension till läsningen och förståelse av bokens tema.

Hjärtans fröjd handlar om ung pojkes första möte med kärleken. Trots ett fragmentariskt berättande med en uppbruten kronologi är boken relativt lättläst. Den starka berättelsen om kärlekens kraft och utsattheten i att vara förälskad varvas med den oerfarna pojkens tafatthet och blir på så sätt en berättelse de flesta kan känna igen sig i.

Sandor slash Ida handlar om två ungdomar som möts på ett chattforum på nätet och inleder en vänskapsrelation baserad på den anonymitet de upplever genom sitt textutbyte. Berättelsemönstret utgår från att vi växelvis får följa Sandor och Idas vardag, och därigenom får vi lära känna två till det yttre mycket olika ungdomar som möts kring gemensamma livsfrågor. Här kan just skiftet mellan de två huvudpersonerna underlätta läsandet, för trots bokens lite större sidantal upplevs den ofta som lättläst.

Arbetet med litteraturanalysen läggs sedan upp på följande sätt: eleverna får utifrån muntliga och skriftliga instruktioner ett

antal sidor att läsa per vecka, för att därefter under lektionstid skriva en text om specifika delar i boken. Vid det nästkommande lektionstillfället får eleverna tillbaka sin text för bearbetning efter lärarens kommentarer. Samtidigt får de ett nytt läsbeting och så fortsätter det tills man har uppnått de delar i analysen som det centrala innehållet har lyft fram, och som uppgiften syftar till att behandla. När varje deltext är klar lämnas den till läraren som kommenterar både analysens innehåll och språket i texten. För att eleven ska kunna göra en litterär analys behöver han eller hon ha ett ämnesrelaterat språk. Här handlar det om att eleven både ska kunna, och våga, använda sig av de litterära termer som kan hjälpa till att ge en så precis och exakt formulering som möjligt av det eleven vill lyfta fram i texten. Veckan, eller lektionen, därpå går eleven tillbaka till sin text och utvecklar analysen utifrån lärarens kommentarer. Eleven får därefter ett nytt läsbeting på x antal nya sidor, beroende på hur tjock boken är och hur lång tid man har lagt upp arbetet på.

Uppgiften startar med att eleven läser cirka 20–30 sidor i sin roman och därefter ombeds att skriva om bokens inledning utifrån en rad frågor: Vad händer på bokens allra första sida? Vill man läsa vidare? I så fall varför, varför inte? Vad händer under de första sidorna i boken? Vad tror du kommer att hända längre fram? Verkar boken lättläst, tung, rolig, spännande etcetera? Att komma in i en bok alldeles i början kan för många läsare vara svårt, och dessa frågor är tänkta att ge läsaren en insikt i vad en inledning i en roman är, och vad den pekar fram emot. Vad en inledning har för betydelse för en läsoplevelse av en bok, och vad det kan finnas för motstånd inför att börja läsa en lång text, är frågor som eleven kan fundera på och som kan vara en del i en analys där eleven sedan arbetar sig igenom boken. I slutet av analysen kan de första frågorna sedan ställas i ett annat ljus och kanske i sig göra att nästa bok eleven börjar läsa blir lättare, just för att eleven då är medveten om inledningens betydelse för historien, men också om den ibland upplevda svårigheten med att få grepp om berättelsen på de första sidorna.

Vid nästa lektionstillfälle får eleven i uppgift att skriva om miljön. I berättelser förlagda i samtidsmiljöer kan det vara svårt att fästa berättelsen vid en speciell miljö, just för att den kan vara alltför bekant för läsaren. Genom att se hur den historiska tiden eller den sociala och/eller geografiska miljön framträder i berättelsen kan det vara intressant att reflektera över vad denna har för betydelse. Och framför allt, om den inte är framskriven, att eleven reflekterar över varför den inte är det. Har tid och rum, eller social miljö, någon betydelse eller är det underordnat berättelsen? Ungdomsböckerna som används i den här analysuppgiften har liknande historisk och geografisk miljö, men vid läsning av böcker från andra tider och andra platser kan de olika miljöer som skildras i dessa vara av betydelse. Här kan det vara av speciell vikt att se hur miljön har betydelse, respektive inte har betydelse för tolkning och upplevelse. Vad är likt och vad är olikt? Vad får det för betydelse för min läsning? Analysdelar att ta upp i nästa steg i uppgiften är personbeskrivningar, som i sig kan se väldigt olika ut. Hur får vi reda på vilka personerna i berättelsen är? Varför har författaren gjort dem på det sätt som de framträder i romanen? Får vi reda på något om dem genom yttre eller inre dialog, beskrivningar eller händelser? Är de trovärdiga? Kan man känna igen sig i dem?

Därefter ska eleven arbeta med att resonera om hur berättelsemönstret i den bok de läser ser ut. Som nämndes ovan har just dessa tre romaner ett speciellt berättelsemönster, vilket har betydelse för hur innehållet presenteras. Ibland kan just ett icke-kronologiskt berättande komplicera förståelsen av texten och det är därför viktigt att eleven blir medveten om den litterära formens betydelse för att förstå vad denna kan tillföra innehållet i berättelsen. Vad tillför den? Hur tänker eleven/läsaren om detta? Här är det också viktigt att analysera vilket perspektiv som berättaren intar i boken. En vanlig tanke hos många elever är att författaren är jaget i boken. En sådan typisk faktiv läsning, där allt i böckerna förväntas avspegla en reell verklighet, gör att det kan vara svårt att skilja mellan berättare och författare, speciellt i fall där dessa skulle kunna vara en och

samma person, vilket ofta tas för givet om författaren och berättaren t.ex. har samma kön. För att få fram en litterär läsning behövs därför en distans i analysen mellan författaren och berättaren, och kanske också huvudpersonerna, där berättarperspektivet lyfts fram som ett val av författaren. Också språket kan analyseras under denna del. Med vilken språkdräkt presenteras personerna, och hur låter författaren dem uttrycka sig själva i texten? Vad kan språkliga uttryck säga oss om vilka personerna i texten är?

I analys av litteratur kan man lätt hamna i ett slags färdighetsträning där boken ska dissekteras som ett ting i sig, för att få fatt i de olika delarna. Det är dock viktigt att inte se litteraturen som icke-kontextbunden, det vill säga som varken knuten till läsaren eller till läsarens värld. Frågor att ställa sig kan vara: Vilken igenkänning upplever eleverna och vilka identitets- och livsfrågor lyfter boken fram? Vad säger litteraturen mig som läsare och vad ser jag för budskap i det jag läser? Att det är eleven som läsare som ska få komma fram är förstås oerhört viktigt. Analysen ska vara där för att lyfta fram läsningen för eleven, inte som en exercis där delar ska dissekteras för sin egen skull, utan som ett sätt för eleven att få fatt i vad litterär text är och kan vara.

Den sista delen av analysen i den här uppgiften måste därför vara elevens eget omdöme om boken. För att detta inte ska bli endast »bra«, »godkänd« eller »intressant« är det viktigt att läraren lyfter fram att omdömet ska kopplas till det som eleven har fokuserat på tidigare i sin analys. Kanske är berättelsen inte trovärdig för att dialogerna inte fungerar, miljön känns för avlägsen, personernas agerande är för avvikande etcetera. Kanske är den bra för att man känner igen sig i de inre dialogerna och att boken på så sätt får en funktion för läsaren. Kanske tar boken upp en problematik som gör att eleven växer i sin kunskap om något speciellt område. En litterär text kan upplevas som meningsfull eller inte av många olika anledningar, och här handlar det om att få eleven att formulera dessa anledningar. Lärarens konstruktiva kritik blir ett sätt att få till en kritisk läsning hos eleven, där läsningens roll i kommunikationen mellan text och läsare kommer fram.

Eleven ombeds att läsa den valda romanen hemma, men att skriva i skolan. Läraren kan ge väldigt korta läsbe-ting, som i sig bearbetas genom elevens skriftliga analys varje vecka. Tanken är att eleven inte ska lämnas ensam med hela boken i ett svep utan att han eller hon ska känna att det han eller hon precis har läst utgör underlaget för den kommande lektionen. Med en snabb återkoppling på den skriftliga analysen, som då heller inte blir så mastig för läraren vecka för vecka, får eleven på så sätt en fortlöpande kommunikation med läraren om sina tankar om texten.

Genom att dela upp den litterära analysen i olika delar arbetar eleven fram en text som läraren kontinuerligt ger återkoppling på. De olika analysdelarna, såsom bokens inledning, personer, miljö, berättarteknik, handling och budskap, samt elevens egna erfarenheter eller annan omvärldskunskap som relateras till litteraturen, bildar till slut en hel elevtext. På så sätt lotsas eleven fram till att hitta kriterier och ett ämnesrelaterat språk att använda för sitt omdöme. Denna återkoppling, som här fungerar som undervisningsmetod, utgår ifrån elevens text samtidigt som läraren kan hjälpa varje elev med dennes analys och reflektion. Återkoppling kan se olika ut i det pedagogiska arbetet, men utifrån egen erfarenhet förespråkar jag här att man som lärare arbetar så dialogiskt som möjligt kring texten. Just att arbeta med en mer dialogisk och processororienterad återkoppling på elevers litterära analyser gör att de bearbetar sina egna reflektioner på ett mer fördjupat sätt. En sådan återkoppling ställs emot en mer tillrättavisande och värderande återkoppling där eleven ändrar sina analyser efter lärarens markeringar (Newell, 1994). Dialogisk och processororienterad återkoppling utgår ifrån att läraren hjälper eleven i dennes löpande bearbetning av sin text (Hattie & Timperley, 2007). Lärarens kommentarer handlar därför om hur eleven ska förtydliga, fördjupa, bredda etcetera det som denne själv har lyft fram i texten.

Den här sortens kommentarer utgör alltså en formativ bedömning och bildar underlag för den kommande, planerade undervisningen. Visar eleverna att de inte har förstått berättarperspektiv,

inte fördjupar sina personbeskrivningar, inte kopplar sitt omdöme om texten till sin egen analys etcetera, blir detta i sig ett tecken för läraren om att bli än tydligare kring dessa delar i sin undervisning. I slutändan, när den summativa bedömningen ska göras, blir det också tydligt vilken återkoppling läraren har gett och undervisningen och bedömningen blir på så sätt tydligt ihopkopplade i elevens egen text, både för läraren och för eleven.

Formativ och summativ bedömning

I bedömningen av elevens text bör kravet vara att de olika delarna i den analys som beskrivits i uppgiften ovan finns med och att dessa är åtskilda ifrån varandra. Själva det pedagogiska upplägget med denna uppgift är att eleven, genom att dekonstruera en text, kan få fatt i hur just den skönlitterära texten fungerar. Vilka delar består den av och vilken mening har dessa i gestaltandet av ett budskap etcetera?

För att undervisning, lösning av uppgiften och bedömning ska stå i önskvärd relation bör den återkoppling som eleven får hela tiden peka på hur analysen kan utvecklas, vilka krav den skriftliga kommunikationen har och hur väl citaten speglar det eleven lyfter fram i sin analys. Om detta ska bli synligt för eleven bör den formativa återkopplingen vara tydlig och explicit. Den litterära analys som har beskrivits i den här uppgiften färdigställer eleven genom en process med hjälp av handledning av läraren. Det är alltså inte ett prov som ska skrivas utifrån memorerade kunskaper, utan något som utvecklas utifrån vad eleven ser i texten. Slutprodukten betygsätts utifrån på hur välutvecklad analysen är och hur väl eleven kopplar den egna texten till den lästa romanen primärt, men då läraren även vill att annan kunskap som eleven har ska kunna appliceras på texten kan inte ens den smalaste analysuppgift förneka kunskap om omvärlden, eller i alla fall att eleven relaterar texten till någon slags omvärld.

Jag har här visat hur en undervisningsprocess kan kopplas till

det centrala innehållet i Lgr 11, och hur dess innehåll i sin tur kan relateras till de olika betygskriterierna som är framskrivna för betygen E, C och A. Det bedömningen av en uppgift i litterär analys framför allt ställer krav på är att läraren har en definition av vad en välutvecklad analys på den här nivån egentligen är. Det blir dock nästan omöjligt att göra detta på ett begripligt sätt för eleven, utan att gå via den litterära texten och elevens eget arbete. Jag menar här att man måste arbeta processinriktat för att återkoppling och betygssteg ska bli begripliga för eleven. Det mål som finns med uppgiften är att eleven själv ska ha resonerat om textens olika delar utifrån sin egen läsning, men också förstås genom den kunskap som läraren har förmedlat. Här blir det omöjligt att på förhand tala om vad eleven ska skriva i sin analys, eftersom ett färdigt svar från läraren på hur litteratur kan upplevas och vad den kan säga oss omintetgör syftet med själva uppgiften. Bedömningen måste i stället gå via den stöttning läraren ger eleven och hur responsen tas upp i elevens egen bearbetning av texten. Det blir i relationen mellan den litterära texten, elevens analys och lärarens återkoppling som bedömningen måste ske. Eftersom uppgiften innebär att det är elevens analys av texten, och att texten skrivs av eleven i skolan, bör läraren kunna dra slutsatsen att det litterära kunnande som eleven delar med sig av i sina analyser också speglar förmågan att analysera det verk han eller hon skriver om.

Kan läraren då utifrån uppgiften bedöma elevens förmåga att analysera text generellt? Läsförmågan som här examineras handlar om litterär text, att läsa en roman, och denna kan förstås inte automatiskt överföras till all annan text. Denna uppgift tar inte heller upp alla delar som är uppräknade i måldomänen över litterär kunskap i kursplanens centrala innehåll och kunskapskrav, men den täcker flera av dem.

När det gäller läsförståelse och analys av litteratur utifrån de områden som tas upp i det centrala innehållet, det vill säga citat- och källhänvisning (i alla fall till viss del, då eleven här endast hänvisar till ett verk), bearbetning av egen text och ämnesrelaterat

språk, menar jag ändå att de resultat som kommer fram i uppgiften visar elevens förmåga i måldomänen i relation till dessa.

Uppgiften byggs upp av elevens eget skrivande, men med tydlig återkoppling av läraren. Hur självständig den litterära analysen är kan i och för sig diskuteras. Att skriva analytisk text utan någon återkoppling före färdigställandet av texten är dock en ovanlig skrivsituation, och kan därför inte vara ett mål i sig. De litterära analyser som elever förväntas skriva på gymnasiet, och senare i universitets- och högskoleutbildning, kommer att kommenteras av lärare och handledare och utgör alltså en vanlig lärandesituation i fortsatta studier. Hur långt eleven når i sin litterära läsning och analys av litteraturen med stöd i undervisningssituation, har bland annat att göra med var eleven börjar sin kunskapsresa, men framför allt om hur lärandeprocessen ser ut.

Med stöd i Kanes validitetsmodell (se kapitel 10) kan vi precisera de olika led i bedömningen som krävs för att en tolkning av elevens kunskaper ska vara giltig. Till att börja med gör läraren en observation av elevens prestation. Elevens delanalyser, samt de kommentarer som läraren ger eleven, observeras kontinuerligt för att se hur eleven hanterar dessa. När hela texten är skriven får den ett omdöme som utgår ifrån de kunskapskriterier som finns för årskursen. Detta omdöme på den enskilda uppgiften antas därmed motsvara den nivå som eleven skulle prestera på andra likartade uppgifter och utgör således det första steget mot ett mer generellt omdöme av elevens kunskapsnivå. Tillsammans med andra uppgifter ska betygsomdömet på uppgiften alltså kunna leda fram till ett omdöme om elevens kunskaper i måldomänen. Vad som menas med likartade uppgifter tas inte upp i denna text, men den litterära analysförmågan och kunskapen om litterär analys som efterfrågas i årskurs 9 både kan och bör testas även vid andra tillfällen och med andra texter. För att kunna sätta ett slutgiltigt betyg för årskurs 9 bör eleven kort sagt ha fått flera möjligheter att göra uppgifter som testar samma sak.

De resultat eleven visar genom den uppgift som beskrivs i

det här kapitlet menar jag möjliggör några slutsatser om elevens kunskapsnivå i måldomänen (extrapolering), men det finns naturligtvis också komplikationer kring uppgiften som kan stå i vägen för den reliabla och valida bedömning som läraren gör av elevens kunskaper.

Validitetshinder

Ett exempel på vad som kan försvåra extrapolering till måldomänen är val av litteratur, vilken kan skapa olika mycket engagemang för undervisningsprocessen hos olika elever. Val av litteratur kan alltid stå i vägen för läsandet och därför naturligtvis också för den analys som eleven sedan gör. Nyfikenheten, som blir fördjupningens drivkraft, kan variera om eleven tycker boken är spännande eller på annat sätt engagerande. Med en annan roman skulle läraren således kunna få ett annat resultat från samma elev.

Eftersom det i uppgiften är viktigt att läraren kan hjälpa eleven framåt i sin analys måste läraren ha läst samma bok som eleven. Det är också viktigt att böckerna man som lärare väljer har liknande form, om det nu är det man vill lägga fokus på, men också att det finns variation, till exempel gällande längden på böckerna, så att elever med olika förutsättningar kan få en uppgift som passar dem i deras läsutveckling. Här är vi i skolan nästan alltid begränsade av det litteraturbestånd som finns i klassuppsättningar och de böcker vi som lärare är väl inlästa på. I en klass med 20–30 elever kan det förstås bli en omöjlig uppgift, även för den väl belästa läraren, att sätta sig in i ett lika stort antal skönlitterära texter inför ett läs- och skrivmoment, som ovan skisserats.

Den analysuppgift som har presenterats här baseras på att elevens läsning ska examineras genom skriftlig framställning. Detta arbetssätt har både för- och nackdelar. En fördel som forskning har visat är att just skrivandet av den här typen av uppgifter, där eleven skriver under tiden de läser sina böcker, gör att de kan fördjupa sina resonemang, kan få fler idéer att skriva om och hjälp att klargöra för

sig själva hur de tänker om litteraturen – men också att den kunskap som de får ut av uppgiften stannar kvar på ett annat sätt (Wong et al., 2002). Vad som dock går förlorat i en sådan här uppgift, som till största delen bygger på ett individuellt textbyggande – om än med hjälp av handledning – är dialogen med klasskamraterna. De nya perspektiv som själva samtalet kan leda till genom att eleverna diskuterar och ventilerar det de själv tänker och tycker kommer inte fram i denna uppgift på samma sätt. Ibland kan samtalet i grupp visa hur man kan tänka om en text samtidigt som de svårigheter som klasskamraterna kan lyfta fram i läsandet kan avlasta den elev och läsare som har fastnat i något läsmoment. Här gäller det att, trots att uppgiften är skriftlig, se till att lektionstid ändå används till tal om texten, inte bara mellan lärare och elev, utan också elever emellan. Detta underlättas om eleverna läser ett begränsat antal olika böcker.

Att examinera litterär kunskap genom skriftlig framställning kan också innebära en stor nackdel då eleven kan ha svårt att formulera sig i skrift, men samtidigt ha en god litterär analysförmåga. I bedömning av läsning, genom skriftlig framställning, är det därför viktigt att se hur läsande och skrivande vägs emot varandra, och hur detta för elevens del kan stå emot vartannat. Om skrivandet som medierande redskap står i vägen för eleven och utgör ett problem i själva formulerandet av kunskapen om litteraturen, hindras eleven att uttrycka just de läsförmågor och analysförmågor hon eller han besitter eller behöver hjälp med att utveckla. Om skrivandet blir för tungt och eleven känner att det är svårt att få fram det han eller hon vill säga, kan även den elev som gärna läser och som har mycket tankar om en text slås ned och prestera sämre. Det är därför viktigt att just analytisk förmåga av litterär text (och naturligtvis också annan text) även examineras på andra sätt. Den skriftliga framställningen får inte stå i vägen för vare sig läsglädje eller analysförmåga. I bedömningen av den skriftliga analysen kan läraren dessutom få svårigheter med en text som är svårläst. Det kan vara svårt att få fatt i det som eleven vill säga med sin text om inte själva skriftspråket

fungerar fullt ut. Skrivandet, som i sig kan göra att eleven fördjupar sina kunskaper, kan alltså stå i vägen både för elevens möjligheter att visa sina kunskaper men också för lärarens möjligheter att se dessa i elevens text. Bedömningen kan på så sätt bli icke-valid.

Såsom den här undervisningsprocessen har presenterats tänker jag mig, vilket också har framgått tidigare, att eleven läser hemma men skriver i skolan. Att just skriva i skolan har både för- och nackdelar. Fördelar kan vara att det skapar möjlighet för läraren att kunna hjälpa eleven vid skrivtillfället, och att läraren vet att det är elevens egen text som examineras. Nackdelar, å andra sidan, kan vara att inspiration och koncentration inte alltid infinner sig vid skrivtillfället när man måste skriva under en viss tid och på en viss plats, och liknar inte riktigt en autentisk skrivsituation. Den standardisering av skrivsituationen som eftersträvas i den här undervisningsprocessen kan därför försvåra extrapoleringen.

Vill man, som läraren i den här uppgiften, att eleven hela tiden ska revidera sin text och arbeta med respons, förutsätter det nästan att eleven skriver på en dator. Det är dock inte alltid så att dator-tätheten i skolan kan leva upp till detta och en sådan situation kan förstås försvåra analysuppgiftens genomförande.

Sammanfattning

I den undervisningsprocess som ovan har skisserats utgår undervisningen och bedömningen av densamma ifrån de problem som kan uppstå både för elev och lärare vid läsning och bedömning av den litterära analysförmågan hos eleven. Långa läsbeting, otydliga lärandeobjekt och svårigheter med att läsa litteratur som estetisk verksamhet är några orsaker till detta. Kunskapskraven som är framskrivna i Lgr 11 är definierade bland annat utifrån värderande adjektiv som »enkla«, »utvecklade« eller »välutvecklade«. I en litteraturdidaktisk undervisningssituation där elevens egen läs- och analysförmåga ska examineras går det dock inte att på förhand säga vad en välutvecklad läsning är, utan denna bedömning måste utgå

ifrån det eleven själv tar fram och sedan kan utveckla med hjälp av återkoppling från läraren. Litteratur som konstform kan inte examineras utefter givna kriterier, då själva upptäckten i litteraturen och läsningen måste vara individens egen. Hur långt eleven kommer i sin litterära analys, där skolan efterfrågar ett antal nedslag som kan analyseras och förhoppningsvis fördjupas, blir därmed det som i slutändan examineras.

Problemen med att extrapolera de resultat som kommer fram i elevens text hänger bland annat samman med hur läraren i undervisningssituationen hanterar valet av litteratur, möjligheten för eleven att visa sin läsförmåga genom skrivförmåga samt med kravet på eleven att prestera vid en viss plats eller på en viss tid, vilket sällan liknar en autentisk skrivsituation utanför skolan. De begränsningar som själva skolsituationen kan ge hoppas jag ändå ska kunna överträffas av den potential till lärande och valid bedömning som ovan nämnda undervisningsscenario kan erbjuda.

Referenser

- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112.
- Kadefors, S. (2001). *Sandor slash Ida*. Stockholm: Bonnier Carlsén.
- Mossberg Schüllerqvist, I. & Olin-Scheller, C. (2011). *Fiktionsförståelse i skolan: svensklärare omvandlar teori till praktik*. Lund. Studentlitteratur.
- Newell, George E. (1994). The effects of written between-draft responses on students' writing and reasoning about literature. *Written Communication*, 11(3), 311–347.
- Nilsson, J. (1996). *Hon går genom tavlan, ut ur bilden*. Stockholm. Wahlström & Widstrand.
- Nilsson, P. (1992). *Hjärtans fröjd*. Stockholm. Rabén & Sjögren.
- Skolverket (2011). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011*. Stockholm: Fritzes.
- SOU 2012:65. *Läsandets kultur. Slutbetänkande av Litteraturutredningen*. Stockholm: Fritzes.
- Steffensen, B. (2005). *Når børn læser fiktion*. Köpenhamn: Akademisk Forlag.
- Tengberg, M. (2011). *Samtalets möjligheter: om litteratursamtal och litteraturreception i skolan*. Stockholm: Brutus Östlings Bokförlag Symposium.
- Wong, B.Y.L., Kuperis, S., Jamieson, D., Keller, L. & Cull-Hewitt, R. (2002). Effects of guided journal writing on students' story understanding. *The Journal of Educational Research*, 95(3), 179–191.

9. Att läsa och analysera noveller i årskurs 9

SUSANNE STAF

Vad ska bedömas och varför?

Att läsa och analysera litterära texter tillhör svenskämnets centrala förmågor, förmågor som får allt större betydelse ju högre upp i skolsystemet eleven befinner sig. Det syns i svenskämnets kursplaner och nationella prov, men också i engelska och moderna språk, där litteraturen är en del av ämnesinnehållet. Av det skälet, men också av andra skäl, behöver svensklärare bjuda in elever till läsning och skrivande som öppnar texterna och avmystifierar begrepp som *tolkning* och *analys*. Ett vanligt och ofta fruktbart sätt att göra det på är att i samtal och gemensam läsning visa på mönster och begrepp som leder bortom den alltför bokstavliga, eller *faktiva* läsningen (Mossberg Schüllerqvist & Olin-Scheller, 2011, s. 53–54). Eftersom läsutveckling är en långsiktig process och texternas komplexitet ökar med elevernas ålder, måste det dessutom göras om och om igen. Sist men inte minst måste elever få möjlighet att lära sig hur tolkningar kan verbaliseras i tal och skrift.

Det här kapitlet, och mitt bedömningsresonemang, utgår från en klass i årskurs 9 som läser och analyserar John Ajvide Lindqvists novell *Pappersväggar*, ur samlingen med samma namn från 2007. Eftersom uppgiften är tänkt som inledning till ett längre novelltema är den kort, förhållandevis enkel och görs på lektionstid. Eleverna får viss stöttning av högläsning och en skrivinstruktion med definierade analysbegrepp, men måste i övrigt lösa uppgiften på egen hand.

Noveller är en tacksam genre för litteraturundervisning på högstadiet. De är berättartekniskt och språkligt intressanta, lätta att hitta i klassuppsättning och ofta så korta att alla elever kan läsa och diskutera hela verk. Det koncentrerade formatet ger dessutom utrymme för fler gemensamma texter som kan vidga erfarenheterna och användas som referenspunkter för fortsatt läsning och diskussion.

Den novelluppgift som beskrivs i kapitlet har flera syften. Ett syfte är att utvärdera resultatet av tidigare undervisning genom att översiktligt bedöma den novellanalys eleverna skriver. Ett annat är att konkretisera de aktuella kursplanemålen genom att använda resultatet som grund för en genomgång i formativt syfte (Black & Wiliam, 1998). Det tredje syftet, som redan nämnts, är att inleda ett novelltema som ska utmynna i en bedömning av elevernas kunskaper på området. Som all undervisning förutsätter detta att både lärare och elever är klara över vad som ska bedömas, och på vilka grunder det ska ske.

Vilken del av det centrala innehållet riktas uppgiften mot?

I kursplanens avsnitt »centralt innehåll« anges huvudinnehållet i svenskundervisningen för respektive stadium. Till högstadiets centrala innehåll hör strategier för att förstå, tolka och analysera texter, liksom förmåga att urskilja budskap, tema, motiv, syften och sammanhang. Undervisningen i svenska ska ge eleverna möjlighet att utveckla läs- och skrivstrategier utifrån texternas »typiska uppbyggnad och språkliga drag«. Att strategierna varierar med textens syfte nämns däremot inte i kursplanen. De skönlitterära textslag som namnges är lyrik, dramatik, sagor och myter, medan epiken, i form av romaner och noveller, får tänkas ingå i gruppen »några skönlitterära genrer«, som återfinns i samma avsnitt. Vidare ska de litterära texterna representera olika tider och geografiska områden och belysa människors villkor, det vill säga livs- och identitetsfrågor. Svenskundervisningen ska ge eleven möjlighet

att utveckla kunskap om texternas språkliga drag, uppbyggnad, berättarperspektiv, parallellhandlingar, tillbakablickar, miljö- och personbeskrivningar samt inre och yttre dialoger. Kursplanen ger alltså starkt stöd för en undervisning som arbetar med att utveckla elevernas läs- och analysförmågor.

Min uppgift är tänkt för elever i årskurs 9. Mot slutet av grundskoletiden har elever kunskap om de vanligaste skolgenrerna, men det är ofta en tyst kunskap. Jag behöver därför undervisa på ett sätt som aktiverar och utvecklar den tysta genrekunskapen. Novellanalysen i mitt exempel bygger på två genrer: novellen och den litterära analysen. Undervisningen behöver uppmärksamma att texterna har skilda syften, men också hjälpa eleverna att utveckla de språkliga resurser som behövs för att klara uppgiften.

Novellen

Den analysuppgift som inleder temat innebär att eleverna läser och kommenterar en samtida svensk novell. Eftersom klassen går i årskurs 9 har jag en bild av elevernas läsvanor och förmågor, men vet också att en avgränsad uppgift ofta fungerar väl vid starten av det sista läsåret, då föresatserna är som störst. Eleverna får tillfälle att repetera tidigare kunskaper och jag får en möjlighet att utvärdera resultatet av tidigare undervisning i formativt syfte. Novelluppgiften är alltså tänkt som avstamp för det fortsatta arbetet med att utveckla elevernas läs- och analysförmågor.

Novellen *Pappersväggar* är skriven av John Ajvide Lindqvist, vars böcker kategoriseras som vuxenlitteratur men tilltalar många tonåringar. Den ingår i en samling med tio noveller med likartade teman, men med varierande längd och svårighetsgrad. Det är också ett av skälen till valet, eftersom jag vill återanvända samma novell i en jämförande uppgift längre fram i arbetet. För att den här bokens läsare ska få en uppfattning om handlingen följer här en kort sammanfattning:

Pappersväggar handlar om en nioårig pojke som får en kartong

av sin pappa. För pojken blir kartongen ett hus och en tidsmaskin med vars hjälp han förflyttar sig i tid och rum. Under det knappa dygn som berättelsen utspelar sig baxar pojken kartongen till en närbelägen glänta, där han tidigare haft en koja. Vid middagen samma kväll ber han pappan att få övernatta i kartonghuset. Pappan svarar ja men berättaren noterar att han verkar ledsen. När pojken återkommer till gläntan är det kväll och fantasilekarna känns överflödiga. Han läser en stund och somnar till, men vaknar snart av stegljud från något han först tror är ett djur. När stegen närmar sig kartongen hör pojken att ljudet kommer från en tvåbent varelse. Snart är den så nära kartongen att han kan känna dess varma andedräkt genom väggen. Trots sin rädsla kan pojken inte motstå frestelsen att sträcka handen mot värmen. I samma ögonblick känner han en smekning längs kartongväggen och strax därefter försvinner varelsen bort. Novellen slutar med att pojken tänker att varelsen lämnat honom.

Pappersväggar är relativt lättläst men innehåller också analytiska utmaningar. Eftersom intrigen är enkel och sammanhangen vardagliga räknar jag med att även klassens ovana läsare kan ta sig in i texten utan problem. Att huvudpersonen, John, är nio år gammal erbjuder en möjlighet till identifikation, samtidigt som den åldersmässiga skillnaden mellan nioåringen och niondeklasserna skapar distans. Historien berättas kronologiskt och i första person, vilket är vanligt i barn- och ungdomslitteratur och därför välbekant för eleverna. Vana läsare märker snart att berättarrösten tillhör den vuxne pojken, en synvinkel som skapar ett tomrum i texten som utmanar elevernas tolkningsförmåga, men också öppnar för fler tolkningar. Liknande tomrum är vanliga i koncentrerade texter som noveller, och därför viktiga att adressera i undervisningen. Slutligen är texten kort, vilket är en förutsättning för att den ska kunna läsas och kommenteras inom ramen för en 80-minuterslek-tion.

GENOMFÖRANDET

För att uppgiften ska kunna genomföras på så begränsad tid krävs förberedelser, och rimliga förväntningar på resultatet. Syftet med den snäva tidsramen är att eleverna ska ha möjlighet att arbeta koncentrerat, och under så likartade förhållanden som möjligt. Inget hindrar att uppgiften görs på mer än en lektion, men att påbörja och avsluta vid samma tillfälle kan vara befriande i en skolvardag med ständiga omstarter. Eftersom snäva tidsramar också riskerar att stressa oroliga själar, behöver eleverna vara införstådda med att resultatet ska användas formativt. I idealfallet är detta både presenterat och diskuterat innan uppgiften genomförs.

När lektionen startar ligger novelltexten på elevernas bänkar. Jag beskriver förutsättningarna för lektionen, som inleds med att jag högläser novellen och instruktionen och fortsätter med att eleverna skriver en kort analys. Innan läsningen presenterar jag författaren och texten. Syftet är dels att skapa förståelse och nyfikenhet inför läsningen, dels att rikta klassens uppmärksamhet mot hur händelserna berättas, snarare än vad som händer. Jag ber dem att markera de avsnitt eller meningar i novellen som de uppfattar som särskilt intressanta eller anmärkningsvärda. Till sist påminner jag om att vi tidigare diskuterat noveller med hjälp av de begrepp som ingår i uppgiften. Tider och arbetsgång står på tavlan, allt för att procedurfrågor inte ska bryta arbetsron.

Högläsningen fyller flera funktioner. Lektionen inleds gemensamt med en stunds lyssnande som ökar koncentrationen och samlar tankarna. De betoningar och tempoväxlingar som tillkommer förstärker texten och hjälper de ovana läsarna att tolka den. Att lästakten är gemensam innebär också att alla får samma skrivtid. En tänkbar nackdel är att högläsningen styr elevernas tolkningar, men det bedömer jag som mindre viktigt än att alla snabbt kommer in i texten och kan påbörja uppgiften. Under skrivandet behöver eleverna också läsa om delar av novellen och kan då upptäcka fler nyanser. Slutligen finns inga förväntningar på att analyserna ska vara originella, även om de gärna får vara personliga så långt uppgiften och tiden medger.

När högläsningen är avslutad ber jag eleverna titta tillbaka på sina markeringar medan jag fångar upp eventuella frågor. Så snart alla är införstådda med uppgiften startar skrivtiden. Vid det laget har cirka 20 av lektionens 80 minuter gått och en timme återstår. Jag uppmanar eleverna att använda hela tiden och tystnaden sänker sig över klassrummet.

Den instruktion som eleverna har framför sig ser ut som följer:

NOVELLANALYS

Kommentera novellen som du uppfattat den efter dagens läsning. Din analys ska innehålla tre delar: en inledande sammanfattning, en analysdel och en avslutning. För att texten ska kunna läsas av någon som inte läst *Pappersväggar* ska du styrka dina påståenden med exempel.

1. Sammanfatta först novellens händelseförlopp. Vad handlar texten om – och vad leder det till? Syftet med inledningen är att ge läsaren en grund för att förstå sammanhanget och din analys. Ett eller högst två stycken är tillräckligt.
2. Analysera sedan novellen med hjälp av de fyra begreppen i listan nedanför. Detta är textens huvuddel där du diskuterar novellen med stöd av citat och/eller exempel.
3. Avsluta med att sammanfatta vad du kommit fram till. Vad uppfattar du som textens grundtanke? Hur tolkar du titeln? Vad i texten uppskattade du – och vad undrar du fortfarande över? På vilka punkter liknar eller skiljer sig *Pappersväggar* från andra noveller du läst?

BEGREPP

- **Tid** När utspelar sig händelserna? Hur lång tid förflyter? Är händelserna berättade i kronologisk ordning (tidsordning) eller på annat sätt? Innehåller texten tidsförkortningar,

tillbakablickar eller framåtblickar? Vad gör det för din bild av det som sker? Sammanfatta och ge exempel.

- **Personer** Vilka personer förekommer i texten? Vem eller vilka uppfattar du som viktigast – och varför? Vilka övriga personer finns det och vilken roll spelar de? Framställer författaren dem genom att beskriva dem, genom deras handlingar, vad de säger eller kanske allt på en gång? Sammanfatta med exempel ur texten.
- **Synvinkel** Vem berättar, det vill säga med vems ögon ser vi det som sker? Är det någon som själv deltar i handlingen eller en utomstående berättare? Jag-berättelse eller tredjeperson? Använder författaren samma synvinkel i hela novellen, eller har han valt att växla? Vilken effekt har synvinkeln på dig som läsare? Sammanfatta och ge exempel.
- **Språket** Novellförfattare måste kunna berätta mycket med få ord. Hur skulle du beskriva språket i *Pappersväggar*? Är det vardagligt eller mer litterärt? Hur använder författaren litterära grepp som kontraster, upprepningar eller bildspråk? Finns det meningar eller avsnitt som känns särskilt viktiga – och i så fall på vilket sätt? Sammanfatta med stöd av citat/exempel.

När du är klar: Läs igenom texten och kontrollera att läsaren kan följa dina tankar. Numrera sidorna och skriv namn på samtliga papper.

När lektionen är slut samlar jag in analyserna och gör en första genomläsning. Jag skriver inga kommentarer i elevernas texter, men noterar tendenser och exempel som kan användas som underlag för kommande lektioner. Jag antecknar också lite om varje elevs text, men avstår från individuella kommentarer tills vi arbetat med fler noveller. På det sättet undviker vi att lägga dyrbar lektionstid på individuell återkoppling utifrån ett litet bedömningsunderlag, något som lätt blir fallet om eleverna vant sig vid att varje liten uppgift kommenteras och betygssätts. Däremot försöker jag upp-

märksamma de elever som överträffat sina tidigare resultat genom att uppmärksamma exempel ur deras texter. En förutsättning för hela arbetet är förstås att klassen är införstådd med bedömningsrutinerna.

VAD BEHÖVER JAG VETA – OCH HUR RELATERAR

UPPGIFTEN TILL DE SLUTSATSER JAG VILL KUNNA DRA?

Novelluppgiften bygger på två av de fem övergripande förmågor som svenskundervisningen ska utveckla: »att läsa och analysera skönlitteratur och andra texter för olika syften« och »att formulera sig och kommunicera i tal och skrift«. Jag vill att uppgiften ska ge en aktuell bild av elevernas förmåga att med visst stöd, men huvudsakligen på egen hand, läsa och analysera skönlitterär text. Med andra ord vill jag identifiera eventuella luckor i kunskapsdomänen litteraturanlys och bedöma om klassen är redo för mer utmanande texter och uppgifter.

Vad är det då jag mer specifikt behöver veta? För det första vill jag veta hur väl eleverna förstår händelser och orsakssamband i novellen. Eftersom uppgiften görs på tid och saknar utrymme för diskussion, har jag valt en novell som utspelar sig i vardagliga miljöer. Händelser och orsakssamband är därför relativt okomplicerade och den sammanfattning som inleder analysen är framför allt tänkt som en hjälp in i skrivandet och i textens föreställningsvärld (Langer, 2005).

För det andra vill jag se om och hur eleverna använder de litterära begrepp som ingår i uppgiften. Det förutsätter i sin tur att eleverna förstår och haft möjlighet att använda begreppen i gemensamma tolkningssammanhang, något bland andra Langer (2005) poängterat. I den instruktion eleverna får ingår definitionerna, vilket kan verka överflödigt, men frigör tid och tankekraft till begreppsanvändningen, som är det jag vill bedöma.

Bedömningen måste också relateras till den text som ligger till grund för analysen. Som nämnts är *Pappersväggar* en kort novell

där begrepp som *handling* och *personer* är relativt lätta att se och kommentera. Begrepp som *tid*, *synvinkel* och *språk* kräver däremot större läsvana och ett mer nyanserat språk. Ovana läsare kan också ha svårt att se att de litterära begreppen beskriver en helhet. Detta underlättas inte heller av att uppgiften, som ber dem kommentera ett begrepp i taget. Jag kan alltså inte räkna med att eleverna upptäcker samband mellan begreppen, men notera om någon gör det. Eftersom uppgiften är skriftlig kan jag också notera om språket övertygar läsaren om slutsatserna i analysen. I instruktionen finns kravet att texten ska vara självbärande, vilket är välbekant för klassen, men värt att uppmärksamma eftersom instruktionen består av frågor. Sist, men inte minst, ska texterna visa om eleverna klarat att underbygga sin analys med relevanta exempel eller citat.

Att allt detta ska fungera inom ramen för en koncentrerad uppgift är knappast realistiskt. Eleverna behöver därför veta att det är en bearbetad version av analysen som ska bedömas mot slutet av temat. Bedömningsunderlaget består av flera analysuppgifter som ska visa i vilken mån, och med vilken kvalitet, eleven kan:

- Sammanfatta texters huvudinnehåll med angivna orsaks- och tidssamband
- Använda litterära begrepp på ett adekvat sätt
- Stödja sina iakttagelser med textexempel som visar att de har förstått text och begrepp
- Resonera kring texters huvudtanke/huvudtankar
- Relatera textinnehållet till andra texter och egna erfarenheter.

Novelltemat berör också andra kunskapskrav i svenska, exempelvis förmågan att välja lässtrategier utifrån texters särdrag, men eftersom det är svårt att bedöma utifrån få och likartade texter ingår det inte i bedömningsunderlaget.

Återkoppling i helklass

När jag läst och bedömt analyserna av *Pappersväggar* återkopplas resultatet till klassen. Genomgången tar en lektion och bygger på anonymiserade exempel ur elevernas texter. Syftet är att konkretisera kursplanemålen genom att knyta dem till citat ur elevtexterna, men också att visa eleverna vad de kan och behöver utveckla. Genomgången följer instruktionen och börjar med inledningen, som ska sammanfatta handlingen och inleda analysen. Samtliga elevtexter innehåller någon slags sammanfattning, men både den och inledningsfunktionen behöver ofta utvecklas. För att åskådliggöra det visar jag klassen tre exempel:

1. Novellen handlar om en pojke som får en enorm kartong utav sin pappa. Han gör den till sitt eget 'hus' och ska t.o.m. sova i den. Han tar ut den i skogen men mitt i natten hör han ett ljud, det är någon/något utanför.

2. Den här novellen handlar om John som berättar om sommaren då han var nio år. Han berättar om dagen då han var nio år. Han berättar om dagen som hans pappa tog hem en stor kartong till honom. John älskar kartongen och vill tillbringa all sin tid i den. Han vill till och med sova i den, men det blir inte riktigt som han tänkte.

3. John är en liten pojke som tycker om att fly från verkligheten och ofta skapar sina egna världar i fantasin. Han har en pappa som jobbar på sågen. Han vet att sonen gillar just sådana små äventyr så en dag tar han med sig en stor pappkartong. Den är så stor att John kan ligga raklång och stänga till helt runt om sig, förutom en centimeter vid öppningen. Hela världen runt om slutar existera och tomrummet i lådan fylls av hans fantasi.

Nästa dag har daggen gjort den släta ytan alldeles bucklig. Den passar inte längre in i trädgårdens atmosfär, så John bestämmer

sig för att flytta den till skogen där hans gamla koja låg. Det blir perfekt. Så perfekt att han vill spendera natten där, och det gör han också. Han pratar med pappa, packar och säger hejdå till pappan som verkar orolig men ändå ler tillbaka och vinkar hejdå. Så på natten känner John lugnet av ensamheten, tills stegen av en okänd gestalt närmar sig lådan.

Därefter får eleverna parvis diskutera två frågor:

1. Vad får läsaren veta om novellen?
2. Vad signalerar att detta är inledningen till en novellanalys?

Syftet med frågorna är att alla ska få syn på att sammanfattningen fyller dubbla funktioner och också måste fungera som inledning till analysen. Jag poängterar att alla texter innehåller en sammanfattning, men att både den och inledningsfunktionen kan utvecklas och bli tydligare. Längre fram i temat kommer vi att diskutera hur delar av analysen kan presenteras redan i inledningen, men tills vidare får detta räcka.

I uppgiftens huvuddel har eleverna kommenterat *Pappersväggar* med hjälp av de litterära begrepp som ingår i uppgiften. Det första är tid och instruktionen ger förslag som att kommentera novellens tid, den berättade tiden och de berättartekniska knep författaren använder för att visa tidens gång. Huvudfrågan är vad detta betyder för läsarens bild av texten, men eftersom många elever uppfattat delfrågorna som jämbördiga består många analyser av en rad svar på detaljnivå. Med detta i åtanke visar jag två exempel:

1. Jag tror att novellen utspelar sig under sommaren därför på sidan 279 står det 'genom de tunna väggarna hörde jag sjungande fåglar, havet som lapade mot strandkanten, svaga röster, båt-motorer.' Detta tycker jag känns väldigt somrigt, speciellt med båtmotorer och sjungande fåglar. Den utspelar sig under två dagar och i kronologisk ordning. Det finns även tidsförkortningar som

t. ex denna på sidan 279 'När pappa hade åkt till jobbet nästa dag gick jag ut till kartongen' och även på sidan 281 när John har varit ute i skogen och nästa stycke börjar med att hans pappa frågar 'Vad har hänt med kartongen?'

2. Novellen handlar om ett minne som huvudpersonen John berättar. En tillbakablick till 'en sommar då han var 9 år'. Det kanske utspelar sig på 70- eller 80-talet. Berättelsen handlar om ungefär 1,5 dygn, från en eftermiddag till kvällen dagen efter. Först handlar det om när pojken får kartongen på en eftermiddag. Sen hoppar författaren fram till morgonen efter, då pojken flyttar sin kartong. Tredje delen är när pojken frågar om han får sova i kartongen som han flyttar och hur han packar. Fjärde delen är när pojken är i kartongen på kvällen och ska sova.

Jag uppmärksammar att eleverna resonerar, eller har påbörjat ett resonemang, att det är ett av bedömningskriterierna och att liknande exempel finns i samtliga analyser. Trots att verb som resonera och analysera ofta förekommer i skolinstruktioner tyder resultatet på att några elever är osäkra på vad det innebär. Ett viktigt mål för kommande lektioner blir därför att visa hur man resonerar om en litterär text.

Nästa begrepp är personer. Vilka personer som förekommer i *Pappersväggar* är inte svårt att upptäcka, men många elever har avstått från att nämna varelsen, vilket jag vill uppmärksamma dem på. Innan vi läser några exempel ber jag dem därför att tänka efter vilka personer de kommenterade i sin analys:

1. De personer som förekommer i texten är John och hans pappa. Man kan även räkna varelsen i slutet som en individ. Jag uppfattar John som den viktigaste personen och det är honom allt fokus ligger kring. Hans pappa har en mycket betydelsefull roll också då det är han som kom hem med lådan från första början. John är en fantasifull pojke. Han framkommer som väldigt modig och

självständig. Man får reda på att Johns pappa jobbar inom byggindustrin. Han verkar också väldigt lugn och oansvarsfull då de flesta pappor inte skulle låta sin unga son sova ensam i en kartong i skogen. Några vuxna nämns också i texten, som John går förbi när han är på väg till skogen, man får dock inte veta något om dem bara att de vet hans namn.

2. De personer som förekommer i texten är John och pappan. Jag uppfattar John som viktigast eftersom han har huvudrollen, medan pappan endast har en biroll. Det finns även några anonyma vuxna som John träffar på stigen när han kör kartongen i skogen i en barnvagn.

Det första exemplet innehåller ett utvecklat resonemang som inkluderar den mystiska varelsen. Jag kommenterar styrkorna men ger också några exempel på hur språket och dispositionen skulle kunna bli mer övertygande. Exempel två är betydligt kortare, men resonerande och mer skriftspråkligt än det första exemplet. Även här uppmärksammar jag styrkorna och möjligheten att vidareutveckla texten med citat och begrepp som huvud- och biperson.

Efter tid och personer kommer synvinkel, ett komplext begrepp som resulterat i flera intressanta kommentarer. Jag visar tre exempel:

1. Nu när jag analyserar texten känns vissa saker lite mer klara. På något sätt och vis känns det som om detta kan vara minnen från författarens liv eftersom han skrev 'Jag kommer alltid att minnas honom så'. Känns bara väldigt äkta helt enkelt.

2. Det är för det mesta en jagberättelse men det sker andra skrivsätt som beskriver hans tankar osv. Det känns som att allt återberättas och att man själv är ett barn när man läser det.

3. Man läser novellen genom 9-åringens ögon och den är i jag-form. När man läser 'på den tiden arbetade pappa på sågen' får man en känsla av att texten är tillbakablickande.

När vi läst ber jag klassen fundera på vad i respektive exempel som visar att det är en analys. Eleverna uppmärksammar att citaten innehåller exempel, att skribenterna förklarar hur och på vilka grunder de tolkat texten och ordval som jagberättelse och tillbakablickande. Jag kompletterar med sambandet mellan tid och synvinkel, som syns i samtliga exempel.

Nästa begrepp är språket. Flera elever har skrivit att de är osäkra på vilka grunder språket kan beskrivas som litterärt eller vardagligt, kunskaper som instruktionen förutsatte att de hade. Eftersom de flesta ändå kommenterat novellspråket som vardagligt eller litterärt ber jag dem lägga märke till vilka kriterier som används i tre exempel:

1. Språket är lätt att förstå eftersom det inte är några gamla ord utan ganska vardagligt. På vissa ställen är det lite mer invecklat, eftersom han vill beskriva exakt hur det var med hjälp av dofter, färger och känslor, t. ex 'De vanliga buskarna och träden blev nya och främmande som om jag tittat på film'. Jämför också ofta så vi ska förstå ännu bättre, kartongen doftade gott som alldeles nya leksaker på födelsedagen.
2. Språket i texten är litterärt. Det finns stilfigurer som liknelser t ex 'som alldeles nya leksaker' eller 'som om jag tittat på film'.
3. Jag tyckte att språket var lagom avancerat. Det fanns ett ord som jag inte förstod, 'pterodaktyler'. Jag tog reda på vad det betydde och fick reda på att det betydde flygödla. Annars var det enkla ord som man lätt förstod. Jag tyckte författaren använde ett målände språk. Jag som läsare fick både känna dofter, ljud och känslor. Jag anser att språket är litterärt, men ändå vardagligt så att man förstår vad författaren vill få ut av novellen.

Utifrån exemplen listar vi kriterier på litterärt respektive vardagligt språk, och fyller på, med ytterligare exempel. Listan blir ett bra underlag för att beskriva de meningar som eleverna citerat i sina analyser, men eftersom lektionstiden är begränsad sparar jag uppgiften till en kommande lektion.

I textens avslutande del ska eleverna sammanfatta sin analys och novellens budskap. Det är den minst utvecklade delen i många av texterna, vilket är väntat med tanke på den begränsade skrivtiden. Kommentarer får därför vänta till kommande lektioner, när klassen haft möjlighet att diskutera *Pappersväggar*. I stället visar jag några exempel som kan väcka nyfikenhet och frågor inför textsamtalet kommande lektion:

1. Textens grundtanke är jag osäker på. Det är en liten pojke som leker att han har en helt annan värld med sin kartong och så i slutet kommer det en varelse. Textens betydelse är kanske bara att berätta en berättelse som författaren varit med om, han återberättar en händelse när han var liten och det är en ganska spännande och intressant text.

2. Det jag har kommit fram till är att novellen visar att man kan skapa sin egen värld vart man än är. Grundtanken är nog att vi har så mycket för oss i vår värld så att man bara ska ge sig in i sin egen värld för en tid. Titeln är ju kopplat till kartongens väggar och att det är ett lätt material som våra drömmar.

3. Varelsen som kommer till John på natten behöver inte vara hans pappa. Det kan vara en metafor för hans barndom som tar ett farväl, när den stryker med handen mot kartongen får den en sista titt på barnet John. Barndomen ser att John är rädd för vad som finns där ute men han klarar det. Barndomen försvinner 'stegen försvann bort i gräset'. Novellen sista mening lyder så här: 'Det hade lämnat mig'. I den här stunden vet John att barnet inom honom är borta.

Därmed är genomgången klar. Om tiden medger ber jag eleverna sammanfatta vad de lärt sig och formulera eventuella frågor. Är lektionen slut gör vi det i inledningen av nästa lektion, så att jag som lärare får en bild av vad eleverna uppfattat och har med sig in i fortsättningen av temat.

Fortsättningen och det samlade bedömningsunderlaget

I följande avsnitt presenteras huvuddragen i fortsättningen av novelltemat. Eftersom det omfattar flera veckor ligger tonvikten mer på didaktiska överväganden och mindre på beskrivningar av undervisningen.

Det fortsatta arbetet inleds med ett textsamtal, som ger eleverna möjlighet att diskutera och eventuellt ompröva sin första tolkning. Samtalet bygger på de frågor som klassen formulerat i sina första analyser, men också på några väl valda utdrag ur novellen. Att litteratursamtal har potential att utveckla elevers tolkningar visar avhandlingen *Samtalets möjligheter* (Tengberg, 2011), som studerar lärarledda novellsamtal i högstadiet. Hultin (2006) visar att litteratursamtal har skiftande syften och identifierar sammanlagt fyra samtalstyper: det undervisande förhöret, det informella boksamtalet, det kultur- och normdiskuterande samtalet och det texttolkande samtalet. Samtalet i mitt novelltema kan närmast beskrivas som texttolkande, men syftar också till att stärka bilden av tolkningar som förhandlingsbara, vilket inte är självklart för alla. Efter samtalet skriver eleverna en kort kommentar om vad de vill utveckla i nästa version av analysen.

Under resten av temat varvas enskild och gemensam läsning med samtal och skrivande. Ett resultat av den formativa bedömningen är att jag fått syn på svårigheterna i att skriva en sammanhängande analys utifrån frågor. Instruktionen erbjöd ett visst stöd för läsning och tolkning men gav mindre vägledning i hur man faktiskt skriver en analys. Den formativa bedömningen visade också att många elever nogsamt hade besvarat frågorna men inte

klarat att formulera de samlade tolkningarna som jag förväntade mig. Ett viktigt inslag i kommande undervisning blir därför att läsa och diskutera andras novellanalyser.

Om elevernas analyser ska övertyga potentiella läsare och motsvara kunskapskraven behöver de bli mer skriftspråkliga. Skrivtiden för novelluppgiften gav inte utrymme för omarbetningar, men ordval och svajiga prepositioner tyder också på att många texter har muntliga förebilder. Ett mål med undervisningen är att eleverna ska kunna formulera tolkningar på ett språk som gör rättvisa åt deras tankar. Även om det tar längre tid än temat varar, är målet att alla utvecklar sitt skriftspråk. Lektionerna innehåller därför övningar där eleverna parvis kommenterar valda avsnitt i de noveller eller analyser som vi läser. Av samma skäl inleds många lektioner med exempel som visar att de gör framsteg.

Novelltemat ger också tillfälle att använda de litterära begrepp som är en viktig del av svenskämnets metaspråk. Kursplanens krav på adekvat begreppsanvändning innebär dels att kunna använda begreppen som redskap för att göra analysen, dels att de ska kunna redovisa resultatet i muntliga eller skriftliga textkommentarer. Analysen av *Pappersväggar* visar att klassen behöver övning på bägge dessa områden, liksom på att definiera och använda begrepp som kontraster och bildspråk. Att svenskundervisningen på högstadiet behöver utveckla ett mer specialiserat språk för textanalys visar också Tengberg (2011).

Det samlade bedömningsunderlaget för novelltemat innehåller tre till fyra analyser, varav en jämförande. Texter och uppgifter erbjuder en viss progression så att arbetet bygger på likartade förutsättningar utan att bli monotont. Mitt exempel inleds med en kort samtidsnovell och fortsätter med två eller tre noveller från andra kulturer eller tider. Avslutningsvis jämför eleverna *Pappersväggar* med en längre novell på samma tema och av samma författare.

Ett avslutande resonemang om uppgifternas möjligheter och begränsningar

Ett huvudsyfte med den analysuppgift som jag beskrivit var att bedöma elevtexterna med de kriterier som gäller för den avslutande bedömningen. Resultatet visade att så gott som samtliga analyser innehöll exempel på de fyra första kriterierna. Eleverna sammanfattade novellen, använde några litterära begrepp, försökte underbygga sina iakttagelser med exempel och resonerade kring novellens budskap och titel. I några texter fanns också exempel som relaterade *Pappersväggar* till egna erfarenheter eller andra texter, vilket var det femte och sista bedömningskriteriet, men bara utgjorde en liten del av uppgiften. Bedömningen visade således att eleverna, med stöd av högläsning, riktade frågor och definierade begrepp, kunde läsa och analysera en kort samtida novell för vuxna.

Resultatet visade samtidigt att det fanns en stor variation i hur eleverna uppfattat novellen och inte minst i hur pass bokstavligt, eller mimetiskt, de läst den. Påfallande många elever verkade till exempel säkra på att den John som är novellens huvudperson var samma John som författat den. Det var också stora skillnader i hur iakttagelserna formulerades i allt från välutvecklade resonemang till korta okommenterade påståenden. Om eleverna fått mer tid hade några texter sannolikt varit mer utvecklade, men tidsramarna resulterade också i att alla slutförde uppgiften. En slutsats av bedömningen var att den fortsatta undervisningen behövde innehålla gemensamt arbete med exempelanalyser som kunde utveckla det metaspråk som bland andra Tengberg (2011) efterlyser.

Punktlistan med de fem kriterierna fungerade alltså som underlag för en bedömning, som konkretiserade målen för kommande undervisning. Som underlag för en bedömning av individuella kunskaper är listan mer tveksam, eftersom den riskerar att reducera en komplex bedömningsuppgift till en besiktning, där läraren bockar av enstaka förmågor i den takt bevisen dyker upp. Eleverna måste förvisso uppfylla de kriterier som ingår i kunskapskraven för respektive betyg, men det som i slutänden avgör kvaliteten är inte

enskilda besiktningspunkter, utan den helhet de resulterar i. Att förmedla detta till eleverna uppfattar jag som den svåraste delen i bedömningsuppdraget.

I det betygsunderlag jag gett exempel på är samtliga analyser skriftliga, vilket kan tyckas ensidigt, men förklaras av att skriftliga analyser spelar en avgörande roll i såväl svenska, engelska och moderna språk. Under det arbete som leder fram till dessa skriftliga analyser finns däremot åtskilliga möjligheter till samtal, tankeskrivande och diskussioner, som kan ge kunskaper och fler infallsvinklar.

Enligt Kanes modell (Kane, Crooks, & Cohen, 1999; samt kapitel 10) ska en giltig eller valid bedömning vara trovärdig i tre led. I det första ledet ska läraren observera elevens prestation, vilket i mitt exempel motsvaras av den inledande novellanalysen som bedömdes kollektivt och i formativt syfte. Som nämnts antecknade jag lite om varje elevtext, vilket gav ett bredare underlag för bedömningen av den jämförande novellanalys som avslutade novelltemat.

I det andra ledet ska resultatet på en enskild uppgift vara likartat för andra, jämförbara uppgifter. I det bedömningsunderlag jag skisserat ingår analyser av flera noveller, varav den sista är jämförande och därför mer komplex. Novelltemat som helhet innehåller alltså en progression som inleds med kortare noveller i vardagliga miljöer, fortsätter med noveller från andra språkområden, kulturer eller epoker och avslutas med en uppgift där eleven analyserar och jämför två noveller. I praktiken är det en viss skillnad på att läsa och analysera samtida texter i kända miljöer och sådana som kräver större inlevelse och andra bakgrundskunskaper, men eftersom variationen ingår i kursplanen måste den återspeglas i undervisning och bedömning. Som lärare behöver jag däremot vara medveten om skillnaderna och se till att eleverna får förutsättningar för att klara de delvis olika uppgifterna.

För att uppfylla det tredje ledet i Kanes modell ska resultatet vara giltigt även i andra kontexter. Detta betyder i första hand andra läsuppgifter i svenskämnet. Kunskaper om läsning och litteratura-

analys är också relevanta i språkämnen som modersmål, engelska eller moderna språk, men även i samhällskunskap eller historia kan fiktionstexter användas som underlag för betygsgrundande uppgifter. Utanför skolkontexten kan elever tänkas kommentera litteratur, film eller spel i bloggar eller andra publika sammanhang där analyskunskap spelar roll för dialogen med andra läsare eller mediekonsumenter.

Avslutningsvis kan jag konstatera att svårigheterna med att bedöma elevers kunskaper om läsning och analys inte försvinner med lång yrkeserfarenhet. Att det finns goda skäl att ägna svenskundervisningen åt att utveckla sådana kunskaper blir jag däremot alltmer övertygad om.

Referenser

- Ajvide Lindqvist, J. (2007). *Pappersväggar: tio berättelser*. Stockholm: Ordfront.
- Black, P. & Wiliam, D. (1998). *Inside the black box: raising standards through classroom assessment*. London: GL Assessment.
- Hultin, E. (2006). *Samtalsgenrer i gymnasieskolans litteraturundervisning: en ämnesdidaktisk studie*. Örebro: Örebro universitet.
- Kane, M.T., Crooks, T.J. & Cohen, A.S. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18(2), 5-17.
- Langer, J.A. (2005). *Litterära föreställningsvärldar: litteraturundervisning och litterär förståelse*. Göteborg: Daidalos.
- Mossberg Schüllerqvist, I. & Olin-Scheller, C. (2011). *Fiktionsförståelse i skolan: svensklärare omvandlar teori till praktik*. Lund: Studentlitteratur.
- Tengberg, M. (2011). *Samtalets möjligheter: om litteratursamtal och litteraturreception i skolan*. Stockholm/Stehag: Brutus Östlings Bokförlag Symposion.

Fördjupning

10. Validitet

MICHAEL T. KANE

Sammanfattning

Provresultat används för att dra slutsatser om provtagare liksom för att fatta beslut baserade på dessa slutsatser. Att validera en tolkning eller användning av provresultat är att bedöma rimligheten hos de slutsatser som dras. I en argumentbaserad modell för validering specificeras de föreslagna tolkningarna och användningarna av provresultat i termer av inferenserna och antagandena i ett tolknings-/användningsargument (interpretation/use argument, IUA). Vidare bedöms validiteten i tolkningen och användningen genom att man bedömer tolknings-/användningsargumentets koherens och rimligheten hos dess inferenser och antaganden. De bevis som behövs för validering är de som krävs för att bedöma tolknings-/användningsargumentet. Tolkningar och användningar av provresultat som är sammanhängande och godtagbara och som stöds av lämpliga bevis betraktas som valida, medan tolkningar och användningar som inte stöds av de tillgängliga bevisen inte är valida.

Vår förståelse av validitet utvecklades gradvis under 1900-talet i takt med att frågor uppstod om vad provresultat betyder och hur de skulle kunna användas (Angoff, 1988). I samband med att nya metoder har introducerats för att hantera nya frågor har validitetsteorin utvecklats till att bli mer komplicerad, men de grundläggande principerna är tämligen enkla. En föreslagen tolkning eller användning av provresultat är valid i den utsträckning som de påståenden som

framförs är rimliga. Om tolkningen eller användningen är enkel kan valideringen vara enkel; om tolkningen eller användningen i stället är anspråksfull kan valideringen vara krävande.

I början av 1900-talet innehöll prov av mentala förmågor uppgifter som ansågs kräva den förmåga som var av intresse, och tolkningen av resultaten inbegrep en generalisering utifrån observerade provprestationer till en slutsats rörande förmågan.³ Den här formen av tolkning tillämpas fortfarande i stor omfattning inom utbildningsväsendet; elevresultat på ett urval uppgifter som förknippas med en viss kunskapsdomän tolkas i termer av förväntad prestation inom den kunskapsdomänen. Om ett provresultat är baserat på ett relativt stort och representativt urval av prestationer inom ett visst aktivitetsområde, kan det vara rimligt att acceptera resultaten som ett mått på prestationsnivå inom det området. Bevis som stödjer tolkningar i termer av akademiska innehållsdomäner eller förmågor som förknippas med sådana domäner har refererats till som *innehållsrelaterade validitetsbevis* (Angoff, 1988).

Vid tiden omkring år 1920 användes provresultat även för att förutsäga framtida prestationer, som en grundval för selektion och placering i akademiska sammanhang och i arbetslivssammanhang. Med ett kriteriemått som kan antas avspegla värdet av en utfallsvariabel kan validitet bedömas i termer av hur pass bra provresultat förutsäger kriteriepoäng. Även den här formen av tolkning tillämpas fortfarande i stor omfattning och stöds av *kriterierelaterade bevis*. Till exempel kan korrelationer mellan provresultat (t.ex. resultat på ett högskoleprov) och kriteriepoäng (t.ex. genomsnittsbetyg för första året på college) användas för att berättiga prediktioner om kriteriet på grundval av provresultaten (Cureton, 1951).

På 1950-talet bekymrade sig psykologer och andra forskare över hur man skulle validera tolkningar i termer av psykologiska egenskaper (t.ex. ångest och aggressivitet), vilka inte definierades

³ Det vill säga att slutsatser om en osynlig förmåga drogs på basis av provresultat. (Red. anm.)

såsom innehållsdomäner och för vilka inga adekvata kriteriemått fanns att tillgå. Cronbach och Meehl (1955) lade fram *konstruktvaliditet* som ett ramverk för validering av resultattolkningar i termer av sådana psykologiska konstrukt, eller egenskaper. Denna konstruktvaliditetsmodell baserades på den positivistiska vetenskapsfilosofin som var vitt accepterad vid den här tiden. Genom konstruktvaliditetsmodellen definierades ett teoretiskt konstrukt implicit genom sin roll i en teori (eller ett nätverk av hypotetiserade förhållanden), och den konstruktbaseade resultatolckningen och teorin bedömdes sedan tillsammans. Om implikationerna av teorin stöddes av empiriska data skulle teorin och tolkningen gällande provresultat i termer av teoretiska konstrukt vara understödda. Om implikationerna av teorin däremot stod i strid med dessa data skulle teorin eller provet betraktas som tvivelaktiga (Cronbach, 1971). Bevis för att provresultaten är relaterade till en annan variabel på sätt som teorin uppger att de ska vara relaterade, understödjer till exempel både teorin och tolkningen gällande provresultaten som mått på konstruktet, såsom detta definieras av teorin.

Det ursprungliga förslaget gällande konstruktvaliditet (Cronbach & Meehl, 1955) formulerades i termer av teoretiska konstrukt införlivade i högt utvecklade vetenskapliga teorier (t.ex. konceptet kraft i Newtons rörelselagar). Som ett resultat av detta var den ursprungliga versionen av konstruktvaliditet inte omedelbart tillämpbar på de flesta fenomen som mäts inom pedagogiken och sociologin. De flesta provresultat tolkas snarare i termer av innehålls- eller prestationsdomäner, och/eller i termer av löst sammanhållna uppsättningar av antaganden rörande fenomenet och dess förhållande till andra variabler, och följaktligen är den ursprungliga, strikta versionen av konstruktvaliditet inte tillämpbar i större utsträckning.

Emellertid betraktades den begreppsmässigt rika modellen som lades fram av Cronbach och Meehl (1955) som ett användbart generellt ramverk för validering av alla typer av tolkningar av provresultat. Vissa av de formella kraven i den ursprungliga

modellen genomgick gradvis en lättnad eller togs bort för att göra konstruktmodellen mer tillämpbar inom utbildning och sociologisk forskning. Kriterierelaterade bevis för att resultat på ett fallenhetsprov (*aptitude test*) med säkerhet var relaterade till senare prestationer i ett utbildningsprogram ansågs till exempel stödja det allmänna antagandet att den formen av fallenhet behövdes i den formen av utbildningsprogram.

Messicks (1989) *sammansatta modell för konstruktvalidering* bredade den allmänna uppfattningen om konstruktvaliditet till att omfatta alla legitima tolkningar och användningar av provresultat genom att bedöma alla de förhållanden och antaganden som är förknippade med fenomenet ifråga. Provresultaten bedömdes i termer av hur väl de överensstämde med den allmänna uppfattningen av fenomenet eller konstruktet. Validiteten i provresultaten som mått på konstruktet skulle bedömas i termer av sin begreppsmässiga och empiriska relevans för konstruktet och i termer av sin mottaglighet för källor till konstruktirrelevant varians.

Messicks (1989) sammansatta modell sammanförde de olika valideringsmodellerna inom ett gemensamt ramverk, men den tillhandahöll inte några begränsningar gällande vilka typer av bevis som var relevanta för validering och tenderade därför att vara relativt öppen. Utan en väldefinierad konstruktteori tenderar konstruktvalidering att vara väldigt öppen och kan resultera i »en utdragen, till och med oändlig process« (Cronbach, 1989, s. 151).

Den argumentbaserade modellen för validering

Den argumentbaserade modellen för validering utformades för att bibehålla den inneboende generaliteten i den sammansatta konstruktmodellen (Messick 1989), samtidigt som den skulle tillhandahålla en enklare och mer systematisk approach till validering (Chapelle, Enright, & Jamieson, 2008; Cronbach, 1988; Kane, 2006, 2013; Shepard, 1993).

Min version av den argumentbaserade modellen (Kane, 2006,

2013) tillämpar två typer av argument, vilka tjänar två viktiga syften. *Tolknings-/användningsargumentet* (*interpretation/use argument, IUA*) specificerar påståendena (dvs. de inferenser och antaganden som inbegrips i de föreslagna tolkningarna och användningarna), och *validitetsargumentet* ger en bedömning av den föreslagna tolkningen och användningen av provresultaten.

Den föreslagna tolkningen och användningen kan betraktas som valid om tolknings-/användningsargumentet är tydligt, koherent och fullständigt, samt om dess inferenser är skäliga och dess antaganden rimliga. De typer av bevis som krävs för validering avgörs av de påståenden som framställs, där mer anspråksfulla påståenden kräver fler bevis än mindre anspråksfulla påståenden. Den allmänna tanken är att noggrant analysera vad som påstås, vilket sker genom att den föreslagna tolkningen och användningen av resultaten bryts ner i ett nätverk av inferenser och antaganden, som sedan kan bedömas på ett systematiskt sätt.

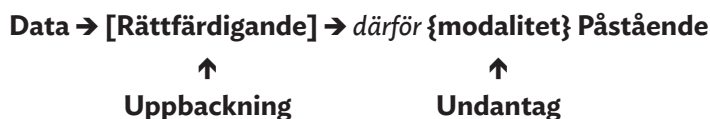
TOLKNINGS-/ANVÄNDNINGSGRUNDEN

Tolknings-/användningsargumentet specificerar det resonerande som ingår i användningen av provresultaten för att dra slutsatser och fatta beslut. Avsikten är att analysera den logiska grunden för tolkningen och användningen av provresultaten och samla detta till ett argument som sedan kan bedömas i termer av följdriktighet och rimlighet.

Tolknings-/användningsargumentet spelar samma roll som en vetenskaplig teori spelar i Cronbachs och Meehls (1955) redogörelse för konstruktvaliditet, genom att specificera de påståenden som framställs. Dock ger den argumentbaserade modellen utrymme för en uppsättning av tolkningar och användningar som sträcker sig från enkla tolkningar i termer av en enkel färdighet, till konstrukt-tolkningar baserade på en teori av någon form och i nästa steg till användningar av provresultaten.

Inferenserna i tolknings-/användningsargument kan ta sig

många uttryck, men de tenderar att vara *presumptiva* i den meningen att de fastställer en förmodan till fördel för den föreslagna tolkningen eller användningen, men fastställer den inte utom allt tvivel. Toulmin (1958) föreslog en modell för presumptiva argument som antar att varje inferens i argumentet utgår från ett givet underlag (*datum*)⁴ av någon form och framställer ett *påstående* av något slag. Det givna underlaget kan bestå av observationer (t.ex. provprestationer), eller av påståendet från en tidigare inferens. De presumptiva inferenserna kan sammankopplas på olika sätt för att bilda presumptiva argument, i vilka påståendet från en inferens kan tjäna som givet underlag för nästa inferens.



Figur 10.1. Toulmins modell för argumentationsanalys; från data till ett påstående, stött av ett rättfärdigande (som stöds av uppbackning). Modalitet avser de reservationer som är nödvändiga.

En inferens är beroende av ett *rättfärdigande*, vilket är en allmän regel med vars hjälp man ska kunna sluta sig till ett påstående utifrån ett givet underlag. I provsammenhang medför till exempel bedömningsregeln ett rättfärdigande av rörelsen från en observerad provprestation till ett provresultat. Rättfärdigandet kräver i allmänhet stöd, vilket betecknas som *uppbackningen* av rättfärdigandet. Vissa rättfärdiganden kan vara mycket rimliga a priori och behöver därför inte ytterligare uppbackning, men de flesta rättfärdiganden kräver någon form av uppbackning, och rättfärdiganden som bemyndigar starka påståenden kräver i allmänhet omfattande uppbackning.

Rättfärdigandena kan åtföljas av en *modifierare*, som indikerar

4 Kane använder singularformen av *data* i betydelsen »givet underlag«. (*Red. anm.*)

vår tilltro till inferensen. Modifieraren kan vara kvalitativ (t.ex. genom ord som »ibland« och »nästan alltid«) eller kvantitativ. Generaliseringsinferenser baserade på statistiska modeller kan till exempel modifieras av mätningars standardfel eller konfidensintervall, och en regressionsbaserad prediktion kan modifieras av standardfelet i en uppskattning eller konfidensintervall (Haertel, 2006).⁵

Ett presumtvt argument ger stöd för de påståenden som framställs, men Toulmin (1958) lämnar utrymme för undantag, eller *invändningsvillkor*, som indikerar villkor under vilka rättfärdigandet inte skulle gälla. Till exempel kan standardutformningen av ett tolknings-/användningsargument behöva modifieras för att passa provtagare med handikapp, för vilka provets standardutformningar eventuellt inte är lämpliga. Sådana anpassningar förändrar inte nödvändigtvis resultatens betydelse eller de sätt på vilka de kan användas; att man ger elever med nedsatt syn eller dyslexi extratid på ett prov som inte ska avslutas inom en begränsad tid (dvs. att de flesta elever avslutar provet utan att få ont om tid eller behöver skynda sig mot slutet) skulle till exempel inte förändra provresultatets betydelse och skulle inte ha några implikationer för de flesta provanvändningar. I sammanhang där hastighet är ett betydande inslag (t.ex. i certifiering av akutläkare) skulle däremot det faktum att vissa kandidater får mycket extratid kunna komma att störa den föreslagna tolkningen och användningen av resultaten.

Tolknings-/användningsargumentet består av en följd eller ett nätverk av presumtiva inferenser, som leder från de observerade prestationerna till den föreslagna tolkningen och användningen, och det tillämpas varje gång provprestationer bedöms, tolkas och

⁵ En mätning är alltid förbunden med osäkerhet och det finns en stor uppsättning statistiska modeller för att skatta osäkerheten i mätningar. Konfidensintervall presenteras ibland i samband med opinionsundersökningar och kan exempelvis ange inom vilket spann (t.ex. mellan 34,2 och 37,3 procent) som den faktiska andelen presumtiva röstare ligger. Konfidensintervallet är en säkrare skattning än en enskild procentsats, eftersom det tar hänsyn till att det finns faktorer (t.ex. antalet medverkande i undersökningen) som spelar roll för vad som är möjligt att tolka från utfallet av en mätning. (*Red. anm.*)

används. När tolknings-/användningsargumentet har utvecklats och validerats kan det användas för att tolka varje enskild provtagares prestation, såvida inte speciella omständigheter leder till ett undantag.

Tolknings-/användningsargumentet erbjuder ett ramverk för validering genom att ange huvuddragen för de inferenser och antaganden som ska bedömas. I takt med att vi breddar tolkningen genom att tillföra inferenser tenderar vi att göra den föreslagna tolkningen mer informativ och kanske mer användbar, men vi behöver även fler bevis för validering. Tolknings-/användningsargumentet bör uttryckas tillräckligt tydligt och detaljerat för att det ska vara möjligt att identifiera de inferenser och antaganden som är inbegripna i den föreslagna tolkningen och användningen, i synnerhet alla eventuella tvivelaktiga antaganden. En brist i vilken del som helst del av tolknings-/användningsargumentet undergräver argumentet som helhet, även om resten av argumentet är väl underbyggt.

VALIDITETSARGUMENTET

Validitetsargumentet erbjuder å sin sida en bedömning av tolknings-/användningsargumentet. Föreslagna tolkningar och användningar kan betraktas som valida i den utsträckning som tolknings-/användningsargumentet är fullständigt i den meningen att det till fullo representerar den föreslagna tolkningen och användningen av provresultaten, att det är koherent i den meningen att det är inbördes konsekvent och att dess inferenser och antaganden är antingen inherent rimliga eller stöds av lämpliga bevis.

Olika former av inferenser kommer att kräva olika former av bevis för sin bedömning. Till exempel förlitar man sig när det gäller bedömningsregler i allmänhet på expertutlåtanden om kriterierna som används vid bedömning och på empiriska bevis för att kriterierna är konsekvent tillämpade. Provprogram som inbegriper prediktioner om framtida prestationer kräver i allmänhet empi-

riska studier för att fastställa länken mellan provresultaten och den framtida prestation som är av intresse. Provresultatsbaserade beslutsrutiner kräver bevis för att rutinen ska kunna uppnå sina mål utan oacceptabla negativa konsekvenser.

Tolknings-/användningsargumentet erbjuder ett ramverk för identifiering av de former av data och analyser som behövs för validering. De påståenden som framställs varierar mellan en viss tolkning/användning och en annan, och därför varierar de bevis som krävs för att underbygga påståendena. Dock innefattar validering alltid en bedömning av de föreslagna tolkningarna och användningarna av resultaten. För att argumentera till fördel för de föreslagna tolkningarna och användningarna av resultaten är det nödvändigt att tillhandahålla uppbackning för alla inferenserna i tolknings-/användningsargumentet, och att utesluta alla rimliga ifrågasättanden av tolknings-/användningsargumentet. Det är viktigt att specificera tolknings-/användningsargumentet så pass detaljerat att dess svagaste del kan identifieras och bedömas.

Observera att det är tolkningen och användningen av provresultaten som valideras, snarare än provet i sig. Det skulle inte vara möjligt att bedöma validiteten hos provresultat utan att, explicit eller implicit, tillämpa någon form av föreslagen tolkning eller användning.

Man kan tänka sig den kedja av inferenser som leder från provtagares prestationer på ett prov till en slutsats, eller till ett beslut om provtagaren, som en serie broar som man måste passera för att ta sig från provresultatet till tolkningen eller användningen. Chapelle, Enright och Jamieson (2008, s. 9) har utvidgat denna metafor:

... man kan tänka sig att ett giltigt tecken krävs för att man ska kunna korsa var och en av broarna och att giltiga tecken erhålls när adekvat uppbackning tillhandahålls för de antaganden som är förknippade med var och en av inferenserna.

Och om en av broarna saknas fallerar tolkningsargumentet i sin helhet.

I likhet med vetenskapliga teorier kan tolkningsargument bestridas på flera olika sätt, men en särskilt effektiv form av bestridande utgår från en rimlig alternativ tolkning. Bedömningen av rimliga konkurrerande tolkningar är därför en viktig komponent i bedömningen av alla föreslagna tolkningsargument.

Några ofta använda inferenser

I det här avsnittet undersöker jag några ofta använda inferenser i tolknings-/användningsargument för provprogram i utbildningssammanhang. Dock är de inferenser som undersöks här inte avsedda att betraktas som en fullständig checklista. Tolknings-/användningsargumentet för en föreslagen tolkning och användning av provresultat bör utvecklas för att avspegla denna tolknings och användnings inneboende påståenden. Ett enskilt tolknings-/användningsargument skulle inte nödvändigtvis inbegripa alla de inferenser som diskuteras nedan, men skulle kunna inbegripa ytterligare inferenser.

BEDÖMNING⁶

Bedömningsinferensen för oss från provprestationerna (data) till resultatet (påståendet), och rättfärdigandet för denna inferens är en bedömningsregel eller matris som specificerar kriterier för tillmätning av olika resultat. Uppbackningen för bedömnings-

6 Kane använder engelska scoring, som dock har en vidare betydelse än att tillfoga poäng och som skulle kunna översättas med »generering av provresultat«. Ett provresultat kan vara en siffra (t.ex. »9 poäng«) eller en bokstav (t.ex. »F«) eller en kommentar (t.ex. »du klarade provet bra«). Därför ligger den smalare betydelsen av begreppet bedömning ganska nära till hands; vi bedömer om svaret är rätt eller fel eller om texten har konventionell eller okonventionell struktur. Man bör dock vara klar över att bedömning på svenska i sin tur ibland har ett vidare användningsområde där det används för att tala om både processer och produkter. Vi kan tala om bedömningsarbete (provkonstruktion, rättning, återkoppling m.m.) och resultat. (Red. anm.)

inferensen skulle i det typiska fallet bygga på expertbedömning. Genom granskning kan bevis erhållas för att bedömningsreglerna implementeras på ett konsekvent och korrekt sätt, och i de fall där bedömare ingår i bedömningen erhålls information om bedömar-samstämmighet genom interreliabilitetsundersökningar.⁷

INFERENSER RELATERADE TILL KONVERTERADE PROVRESULTAT

Olika modeller för konvertering av provresultat kan användas för att omvandla observerade resultat till konverterade provresultat (*scaled scores*), som berikar tolkningen genom att ge ytterligare information om vad resultaten betyder (Kolen, 2006).⁸ Till exempel kan observerade resultat ges en normrelaterad tolkning i termer av de procenttal av en viss population (t.ex. tredjeklassare inom ett visst skolsystem) som har resultat vid eller under varje observerat resultat. Rättfärdigandet för denna inferens skulle sätta varje enskilt provresultat i förbindelse med motsvarande procenttal av populationen, och uppbackningen för detta rättfärdigande skulle då härledas ur data rörande poängdistribution i populationen. För att den normrelaterade tolkningen ska vara så användbar som möjligt är det nödvändigt att poängdistributionen baseras på lämpliga urvalsplaner, att urvalen är stora och att datasamlingen är relativt ny.

Kriterierelaterade tolkningar av olika slag kan också berika tolknings-/användningsargumentet genom att indikera vad provtagare med olika observerade resultat kan förväntas ha förmåga att göra. Till exempel kan provtagare med diverse resultat beskrivas

7 Kane refererar här till det faktum att ganska mycket bedömningsarbete, till exempel rättande av högskoleprov, sker automatiserat med hjälp av datorer. I andra fall kan arbetet förvisso utföras manuellt, men utan att den som rättar behöver tolka elevgensvaret i förhållande till kvalitetsbeskrivningar (exempelvis betygskriterier) (se också kapitel 5 av Skar för utförlig diskussion). (*Red. anm.*)

8 Konverterade provresultat uppstår när så kallade råpoäng konverteras till ny poängskala. Ett exempel är högskoleprovet där antalet rätt (= råpoängen) konverteras till ny poäng på skalan 0–2, vilket innebär att exempelvis 49 råpoäng kan motsvara 1,3 »konverterade poäng«. (*Red. anm.*)

i termer av specifika uppgifter som de kan utföra, eller i termer av lärandepressioner som indikerar att provtagare med vissa resultat kan prestera på en viss nivå i en progression av stigande färdighetsnivåer (Kolen, 2006). Att koppla läs- och skrivresultat till svårighetsgrad hos texter kan särskilt vara till hjälp när resultaten används i utbildningssammanhang (t.ex. för att identifiera texter med en för eleven lämplig nivå).

GENERALISERING

Generaliseringsinferensen för oss från ett observerat resultat på ett visst prov till det förväntade resultatet vid upprepningar av provet. Generalisering förändrar inte resultatets värde, men utvidgar tolkningen från en bedömning av provtagarens observerade prestationer till det förväntade värdet av provtagarens resultat vid alla tidpunkter som är möjliga i provproceduren. Detta inbegriper således olika provuppgifter, olika provplatser och olika provtillfällen samt (i fall där bedömare är inbegripna) olika bedömare.

När vi tolkar provresultat generaliserar vi vanligtvis vad gäller ett antal observationsvillkor. Vi säger något i stil med »Mary gjorde bra ifrån sig på läsförståelseprovet«, snarare än »Mary gjorde bra ifrån sig på formulär 08B av läsförståelseprovet, som mrs Smith höll med klassen förra torsdagen«. Med andra ord generaliserar vi vad gäller provformer, tillfällen, kontexter och vem som administrerar provet, och vi antar att resultatet inte skulle variera i någon större utsträckning om dessa aspekter av provsituationen hade varit till viss del annorlunda.

Observationsvillkoren (t.ex. uppgiftstyper, kontexter, provvakter, tillfällen, bedömare) som medges inom definitionen för provproceduren definierar ett universum av möjliga observationer, *generaliseringsuniversumet* (*universe of generalization*), och medelvärdet inom generaliseringsuniversumet för en provtagare är det *generaliserade resultatet* (*universe score*) för provtagaren (Brennan, 2001; Haertel, 2006).

Inom generaliserbarhetsteori (*generalizability theory*) betecknas dessa olika typer av observationsvillkor (t.ex. uppgifter, kontexter, bedömare) som *aspekter (facets)*.⁹

Vi kan till exempel ha en uppgiftsaspekt och en tillfällesaspekt, där olika uppgifter skulle vara villkor för uppgiftsaspekten och olika tillfällen skulle vara villkor för tillfällesaspekten. En enskild observation från generaliseringsuniversumet skulle då inbegripa något villkor av varje aspekt (t.ex. uppgift, tillfälle, kontext eller bedömare).

Generaliseringsinferensen för oss från en provtagares observerade resultat till en uppskattning av provtagarens generaliserade resultat. Tolkningen av resultatet breddas från en bedömning av prestationer som verkligen har bedömts till ett påstående om provtagarens förväntade prestation inom ett universum av möjliga provprestationer. Det observerade resultatet är det givna underlaget (*datum*) och det generaliserade resultatet är påståendet.

Bevis, eller uppbackning, för generaliserbarhet tenderar att utvecklas under utvecklingen av prov. Provmaterial och rutiner utvecklas och generaliserbarheten hos de resulterande provresultaten bedöms. Om resultaten inte är tillräckligt generaliserbara inom en viss aspekt kommer anpassningar att göras för att förbättra generaliserbarheten. Om resultaten exempelvis inte visar någon större konsistens i förhållande till ett urval uppgifter kan antalet uppgifter i provet utökas, eller så kan uppgiftsspecifikationerna stramas upp. Om resultaten på uppsatser inte är generaliserbara vad gäller bedömare kan bedömningsmatriserna eller utbildningen av bedömare eventuellt behöva förbättras (eller också kan antalet bedömare utökas). Alla sådana förändringar medför kompromisser, och revideringsprocessen tenderar därmed att vara komplicerad.

Statistisk urvalsteori erbjuder ett användbart ramverk för analyser av generaliseringsinferenser. Den stödjer sig på antaganden

⁹ *Generalizability theory* är en psykometrisk modell för beräkning av i vilken utsträckning olika så kallade aspekter (facets, t.ex. bedömares stränghet eller uppgiftens svårighet) bidrar till provresultat. (*Red. anm.*)

om urval av prestationer (t.ex. slumpvis urval) från generaliseringsuniversumet. Dessa urvalsantaganden förverkligas dock nästan aldrig i praktiken; provuppgifter och provkontexter skapas av provutvecklarna, bedömare rekryteras och utbildas och provtillfällen styrs av schemamässiga faktorer. Belägg måste alltså påvisas för att urvalet, som inte är slumpmässigt, ska vara tillräckligt representativt för att generaliseringsinferensen ska fungera som avsett, även om observationsurvalet inte är slumpmässigt i strikt bemärkelse. I det typiska fallet stöds representativiteten hos urvalet av två former av bevis. Först görs en ansträngning för att göra urvalet av uppgifter och observationsvillkor så representativt för generaliseringsuniversumet som möjligt, och därefter görs en ansträngning för att identifiera och eliminera alla effekter som skulle göra urvalet avsevärt icke-representativt för generaliseringsuniversumet och därmed innebära statistisk bias.

Om provet till exempel ska användas för att bedöma läs- och skrivförmåga skulle det vara olämpligt att låta alla avsnitt handla om sport; ett sådant skevt urval skulle i statistiskt avseende innebära bias och i det mer allmänna avseendet att det skulle sätta vissa elever i en mindre gynnsam position. Om provutvecklare verkligen har ansträngt sig för att göra urvalet representativt för generaliseringsuniversumet och ingen allvarlig källa till systematiska urvalsfel kan identifieras skulle det vara rimligt att anta att urvalet är representativt.

Statistisk urvalsteori erbjuder även en modifierare för generaliseringsinferensen i form av en uppskattning av det slumpmässiga urvalsfel som är förknippat med den. Modifieringen eller osäkerheten bestäms genom att man med hjälp av statistiska metoder skapar ett förväntat värde i generaliseringsuniversumet genom att utgå från det slumpmässiga urvalsfelet.

Observera att om tolkningen inte utgår från att resultaten är generaliserbara för vissa observationsvillkor (t.ex. kontexter eller tillfällen) skulle variabiliteten för dessa villkor inte inverka menligt på den föreslagna tolkningens validitet. Om tolkningen till exem-

pel är knuten till en specifik period (t.ex. ett förhör som hålls varje vecka) skulle variabilitet under förlängda perioder (t.ex. över flera veckor) tolkas som verkliga förändringar snarare än som slumpfel.

Reliabilitet som ett nödvändigt villkor för validitet?

I klassisk provteori och i senare psykometriska modeller anses reliabilitet eller generaliserbarhet vara ett nödvändigt, men inte tillräckligt, villkor för validitet. Diskussionen som presenteras ovan indikerar varför detta antagande i allmänhet framstår som godtagbart, men den indikerar även vissa förbehåll som bör tas med i beräkningen när detta antagande görs.

Inom den argumentbaserade modellen för validering är generaliserbarhet ett nödvändigt villkor för validiteten hos nästan alla provresultattolkningar, eftersom nästan alla av dessa tolkningar innefattar generalisering inom ramen för något generaliseringsuniversum som sträcker sig bortom de observationer som faktiskt görs. Om vi antar att vi har för avsikt att generalisera till en viss uppsättning tillfällen är det nödvändigt med en viss nivå av generaliserbarhet gällande tillfällen, och om vi antar att vi kommer att generalisera till ett visst uppgiftsuniversum så är det nödvändigt med en viss nivå av generaliserbarhet gällande uppgifter (Brennan, 2001). Inom ramen för klassiska diskussioner om reliabilitet bedöms generaliserbarhet ifråga om tillfällen i termer av *stabilitetskoefficienter*, och generaliserbarhet ifråga om uppgifter bedöms i termer av *parallellmetoden* (*parallel forms reliability*), eller mätningar av *intern homogenitet* (t.ex. alfakoefficienten).¹⁰

Generaliserbarhet/reliabilitet är dock inte ett tillräckligt villkor för validiteten hos de flesta resultattolkningar, eftersom dessa tolkningar inbegriper ytterligare inferenser (t.ex. extrapolering)

¹⁰ Begreppen relaterar till olika statistiska metoder för att undersöka den förväntade stabiliteten i resultat. Alfakoefficienten, till exempel, kan användas för att uttrycka i vilken utsträckning det är rimligt att anta att en viss elev presterar som förväntat på uppgifter som liknar varandra. (*Red. anm.*)

som inte behandlas av reliabilitets- eller generaliserbarhetsbevis, och för att etablera validiteten hos en tolkning behöver vi adekvat uppbackning till alla inferenserna i tolkningen.

Extrapolering

En *extrapoleringsinferens* utsträcker tolkningen från förväntad prestation inom ramen för generaliseringsuniversumet för provet – det generaliserade resultatet – till den förväntade prestationen inom ramen för en *måldomän* som innefattar hela skalan av prestationer av intresse, utifrån den föreslagna tolkningen av resultatet. *Resultatet i måldomänen* (*target score*) är det förväntade värdet inom ramen för måldomänen, och det indikerar hur pass bra provtagaren kan prestera inom måldomänen.

Av praktiska skäl är det i allmänhet inte möjligt för provprogram att tillämpa slumpvisa eller representativa urval från måldomänen, och det skulle heller inte vara legitimt att helt enkelt generalisera utifrån det observerade provresultatet till resultatet i måldomänen. Istället kräver tolkningen av observerade resultat i termer av resultatet i måldomänen en separat extrapoleringsinferens.

Rättfärdigandena för extrapoleringsinferenser kan bedömas empiriskt genom en undersökning av förhållandet mellan provresultat och resultat baserade på *kriteriebedömningar* (*criterion assessment*), som direkt bedömer prestationen i måldomänen. Till exempel kan ett språkprov bestå av korta avsnitt och kortvariga konversationer följt av objektiva frågor, medan kriteriebedömningen skulle kunna innefatta konversationer på tu man hand på detta språk.

Extrapoleringsinferensen kan även bedömas genom att man analyserar förhållandet mellan provprestationerna och prestationerna man vill uttala sig om (prestationer i måldomänen). I den utsträckning som provuppgifterna täcker merparten av måldomänen, eller de färdigheter som krävs av uppgifterna i måldomänen, kan extrapoleringsinferensen anses vara rimlig utan empiriskt stöd

i någon högre grad. I den utsträckning som provprestationerna skiljer sig väldigt mycket från de prestationer man vill uttala sig om skulle extrapoleringsinferensen behöva åtminstone något empiriskt stöd (Messick, 1989).

I vissa fall kan provresultat användas för att förutsäga prestation inom ett annat område och/eller vid en viss tidpunkt i framtiden. Rättfärdigandet för sådana prediktionsinferenser skulle kunna vara en regressionsekvation som sätter ett kriteriemått av framtida prestation i relation till provresultaten, med uppbackning som härletts ur en empirisk undersökning där regressionsekvationen uppskattas. Det uppskattade standardfelet gällande uppskattning för regressionsekvationen ger oss då en modifierare för prediktionerna.

Teoribaserade inferenser

Teoribaserade inferenser utsträcker tolkningen från resultat i måldomänen till teoretiska konstrukt som definieras i termer av sina roller inom en teori (Cronbach & Meehl, 1955). Rättfärdigandet för inferenser från resultatet i måldomänen (*target score*), som används som en indikator för ett konstrukt, till en uppskattning av konstruktpoängen, har resultatet i måldomänen som det givna underlaget (*datum*) och det uppskattade konstruktvärdet som påståendet. Bevisen som stödjer teorin och analyserna av indikatorns relevans för sitt konstrukt ger uppbackning för rättfärdigandet. Såväl teorin som de föreslagna konstrukttolkningarna av provresultaten stöds i den utsträckning som en teoris prediktioner stöds. En teori som överlever en mängd empiriska utmaningar kan accepteras, åtminstone på sannolika grunder.

Om vi har fler än en indikator för ett konstrukt bör de stämma överens (*konvergent bevis*). Om en teori ponerar att två konstrukt är oberoende av varandra bör deras indikatorer också vara oberoende (*divergent bevis*). Om en teori gör gällande att två konstrukt är förbundna med varandra på ett visst sätt (t.ex. funktionellt, i omvänt

förhållande eller genom en modell över latent egenskaper) bör de observerade empiriska förhållandena mellan de motsvarande indikatorerna vara i linje med teorins prediktioner (Messick, 1989).

Användningar av resultat

Beslutsinferenser för oss från en tolkning i termer av en måldomän eller ett konstrukt till ett beslut om provtagaren, där det till exempel preciseras att elever med vissa resultat behöver extrahjälp inom något område, eller att en elev kommer att placeras i en viss studiegrupp inom en kurs.

Tolknings-/användningsargumentet för resultatbaserade beslut inbegriper i allmänhet en kedja av inferenser som leder från observerade prestationer till en resultattolkning, och sedan vidare till ett beslut som åtminstone till en del är grundat på de uppskattade värdena av domän- eller konstruktpoängen. Beslutsinferensen har som sitt givna underlag (*datum*) ett tolkat resultat och som påstående ett beslut.

Vi kan använda provresultat för att fatta olika sorters beslut (Bachman & Palmer, 2010), och när vi bedömer lämpligheten hos besluten förlitar vi oss i allmänhet på de bevis som indikerar att beslutsprogrammets mål uppnås (avsedda, positiva konsekvenser), samt att oönskade bieffekter (icke-avsedda, negativa konsekvenser) begränsas till ett minimum. Beslutsprogram som i allmänhet uppnår sina mål och som inte har allvarliga negativa konsekvenser betraktas som acceptabla, eller valida, och de som inte uppnår sina avsedda mål eller har allvarliga negativa konsekvenser betraktas som oacceptabla, eller icke-valida. Konsekvenserna (eller s.k. wash-back) kan vara positiva eller negativa.

Tre typer av konsekvenser har identifierats som särskilt relevanta för bedömningen av provprogram: de avsedda utfallen, påverkan på skyddade grupper och systemiska effekter (t.ex. påverkan på kursplaner). Bedömningen av dessa konsekvenser inbegriper i allmänhet kompromisser (Kane, 2013). Ett provprogram som

medför betydelsefulla utfall kan betraktas som acceptabelt, även om det har vissa negativa konsekvenser, men ett program med allvarliga negativa konsekvenser kommer sannolikt att förkastas även om det uppnår sina mål.

Sådana konsekvenser som är unika hos en enskild provtagare (t.ex. missnöje med resultaten) skulle vanligtvis inte tas med i beräkningen när beslutsregeln bedöms, såvida de inte indikerar ett allmänt problem med provet eller tolknings-/användningsargumentet. Sådana unika utfall kan behöva studeras från fall till fall, men de är inte nödvändigtvis relevanta för bedömningen av provprogrammet som helhet.

Ett exempel: Kompetensprov

Inom utbildningsväsendet används prov i regel för att bedöma elevers framgång inom de kunskaps- eller färdighetsområden som lärs ut. Ett prov i läsning kan till exempel användas för att bedöma en elevs färdighetsnivå när det gäller att läsa prosa på ett visst språk. En kompetens som läsförmåga (jfr kapitel 4) kan definieras såsom en *måldomän* som specificerar de typer av uppgifter och de uppsättningar av kontexter och observationsvillkor som förknippas med den kompetensen. Måldomänen för läsförmåga skulle kunna definieras i termer av en viss uppsättning av tryckta texter, däribland böcker, tidskrifter, tidningar, instruktioner, tidtabeller, mejl etc., såväl som specifikationerna av de villkor under vilka texturvalen ska läsas.

Måldomänen för en kompetens som läsförmåga skulle i det typiska fallet inbegripa en varierad samling »autentiska« uppgifter och prestationer inom en stor uppsättning kontexter; de är således inte begränsade till provuppgifter i standardiserade provsituationer. Värdet som tillskrivs en provtagares kompetens inom en viss domän är i grunden ett påstående om hur väl provtagaren kan utföra de uppgifter som förknippas med denna kompetens.

Vissa kompetenser har en bred definition (t.ex. färdighetsnivå

inom läsförmåga) medan andra har en snävare definition (t.ex. ordförråds kunskap eller förmåga att använda interpunktion). Uppsättningen av uppgifter som inkluderas i måldomänen kommer i allmänhet att regleras av praktiska frågor; domänen kan vara grundad på kursplanemål eller på de krav som ställs på sökande till ett visst jobb. Till exempel skulle en arbetsgivare som använder resultaten från en bedömning av kommunikativ kompetens i ett främmande språk på arbetsplatsen (t.ex. i kommunikation med kunder) ta i bruk en måldomän som täcker de former av arbetsplatsprestationer som är vanliga och viktiga.

I de flesta fall antas kompetenser avspegla en viss egenskap, färdighet eller samling färdigheter hos provtagaren, och prestationerna som inbegrips i måldomänen kan vara specificerade för att avspegla dessa färdigheter. Emellertid kan kompetensen även definieras i termer av en måldomän utan att man för den skall specificerar alla de färdigheter som utgör prestationerna. Teorier av någon form kan ha kapacitet att förklara prestationerna i måldomänen, däremot behövs inte en exakt förklaring av hur uppgifterna utförs (t.ex. hur en text avkodas och tolkas) för att provresultat ska kunna tolkas som mått på en kompetens som läsförmåga. Kompetensen definieras i grunden av sin måldomän.

Kompetensprov

Uppsättningen av observationer som inkluderas i måldomänen bör vara i linje med den avsedda tolkningen av kompetensen, och måldomänen för en brett definierad kompetens kan därmed inbegripa en stor uppsättning uppgifter i en stor uppsättning situationer. Till exempel skulle måldomänen för prestationer inom läsförmåga förmodligen inbegripa läsning av en tidning eller en roman, men det skulle i allmänhet inte vara genomförbart att inkludera uppgifter av denna längd och komplexitet i ett prov. I stället skulle provet förmodligen inbegripa ett urval av korta avsnitt hämtade ur längre artiklar eller böcker. Provuppgifterna och provvillkoren

skulle ha standardiserats på flera olika sätt och de observationer som inkluderats i provet skulle hämtas från en del av måldomänen.

Det skulle i allmänhet inte vara skäligen att betrakta prestationerna i ett kompetensprov som ett slumpmässigt eller representativt urval från måldomänen för den specifika kompetensen, och det skulle heller inte vara legitimt att generalisera utifrån det observerade resultatet till resultatet i måldomänen. I stället skulle tolkningen av provresultaten i termer av en brett definierad kompetens kräva minst tre huvudinferenser: en bedömningsinferens, en generaliseringsinferens och en extrapolering från det generaliserade resultatet till resultatet i måldomänen.

Att validera tolkningar i termer av kompetenser

Tolknings-/användningsargumentet för en kompetens inbegriper i det typiska fallet bedömnings-, generaliserings- och extrapoleringsinferenser. I bedömningsfasen bedöms den observerade prestationen och utmynnar i ett observerat resultat (t.ex. ett råpoäng eller ett konverterat provresultat av någon form), det observerade resultatet generaliseras sedan till det generaliserade resultatet, som i nästa steg extrapoleras till ett resultat i måldomänen.

Rättfärdigandet för bedömningsinferensen, som i det typiska fallet utvecklas i samband med provutveckling, tillskriver ett resultat (påståendet) till en provtagares observerade provprestation. Uppbackningen för bedömningsinferensen inbegriper oftast ett utslag från paneler med experter som utvecklar och granskar bedömningskriterierna, men även bevisen för att bedömningsrutinerna tillämpas konsekvent och korrekt.

Generaliseringsinferensen behandlar de observerade resultaten som en uppskattning av det generaliserade resultatet (medelvärdet inom generaliseringsuniversumet). Tolkningen utvidgas från ett provresultat baserat på de observerade provprestationerna (*datum*) till ett påstående om förväntad prestation inom generaliseringsuniversumet.

Extrapoleringsinferensen utvidgar sedan tolkningen från generaliseringsuniversumet till måldomänen. Resultatet förändras i allmänhet inte, men tolkningen breddas för att täcka måldomänen av »autentiska« prestationer som är förknippade med kompetensen. Extrapoleringsinferensen inbegriper inte en enkel statistisk generalisering, utan snarare ett mer anspråksfullt språng från påståenden om provprestationer till påståenden om hela prestationsskalan i måldomänen, inklusive prestationer utanför prov och provsammanhang.

Vår tilltro till extrapoleringsinferensen beror åtminstone till en del på förhållandet mellan generaliseringsuniversumet och måldomänen för kompetensen. I den utsträckning som provet täcker merparten av måldomänen är det troligt att extrapoleringen kommer att anses rimlig, och extrapoleringsinferensen behöver då eventuellt inte fullt så mycket empiriskt stöd. Om provet å andra sidan inbegriper uppgifter eller svarssätt som skiljer sig avsevärt från uppgifter eller svarssätt i måldomänen, är möjligheten att varians på grund av *irrelevant metod* (*irrelevant method variance*) kommer att begränsa den föreslagna tolkningens validitet, vilket är en befogad farhåga (Messick, 1989).

»Tänka högt-protokoll«, där provtagare redogör för hur de tar sig an provuppgifter, kan ge mer detaljerade analyser av processerna som tillämpas vid provuppgifter (Cronbach, 1971). Om processerna som tillämpas vid provuppgifter liknar de processer som tillämpas vid andra uppgifter i måldomänen stöds extrapoleringen. Om processerna i stället är avsevärt annorlunda får extrapoleringen mindre analytiskt stöd och skulle behöva mer empiriskt stöd.

Som tidigare påpekades definieras inte måldomäner godtyckligt, utan kan till exempel avspegla innehållet i undervisningen, i färdigheter som antas vara nödvändiga i vissa sammanhang (t.ex. ett jobb eller ett utbildningsprogram) eller i vissa antaganden om underliggande egenskaper. Provprestation kan till exempel förknippas med en egenskap eller också förklaras genom en samling av specifika kunskaper, färdigheter och bedömningsförmåga, som tas i anspråk i olika kombinationer av provuppgifterna.

Slutord

Den argumentbaserade modellen för validering specificerar ingen särskild form av tolkning för resultat, men den kräver att de påståenden som görs läggs fram tydligt och med adekvat stöd. Valideringen av en föreslagen tolkning och användning inbegriper utvecklingen av ett tolknings-/användningsargument, en bedömning av dess koherens och fullständighet och en bedömning av rimligheten hos dess inferenser och antaganden.

Mer anspråksfulla tolkningar kräver mer uppbackning, men en tydlig specifikation av den föreslagna tolkningen och användningen begränsar även omfattningen av de bevis som krävs för validering. Som en följd av detta behöver validering inte vara speciellt mödosamt. Resultattolkningar som framlägger mer blygsamma påståenden (t.ex. prestationsbaserade tolkningar) kan behöva färre bevis för validering, medan mer anspråksfulla tolkningar och användningar kommer att behöva fler bevis för sin validering.

Vissa inferenser och antaganden (t.ex. bedömning, generalisering, konvertering, extrapolering och beslut) är vanligt förekommande i tolknings-/användningsargument och vi har stor erfarenhet av att bedöma dessa inferenser och antaganden. Validering är visserligen inte lätt, men det är i allmänhet möjligt att göra en hyfsat bra valideringsinsats genom lämplig grad av ansträngning (beroende på den föreslagna tolkningen och användningen av resultaten).

Om den avsedda tolkningen av provresultat inte är tydlig är det lätt att bortse från implicita antaganden eller att göra onödiga antaganden. Ett tolknings-/användningsargument som utelämnar inneboende antaganden i tolkningen och användningen kan exempelvis komma att leda till en överdrivet optimistisk slutsats om den föreslagna tolkningens och användningens validitet. Ett tolknings-/användningsargument som inbegriper inferenser eller antaganden som inte behövs för den avsedda tolkningen eller användningen kan i stället komma att leda till en onödigt pessimistisk slutsats om validiteten.

Tolknings-/användningsargumentet ger även en grundval för bedömning av validitetsargumentets riktighet. Om tolknings-/användningsargumentet är koherent och representerar de föreslagna tolkningarna och användningarna på ett precist sätt och dess inferenser och antaganden är rimliga kan den föreslagna tolkningen och användningen betraktas som valida. Som en följd av detta kan de bevis som krävs för validering vara omfattande, men trots det avgränsade och relativt väl definierade.

Ett explicit framlagt tolkningsargument har tre avgörande funktioner: det ger vägledning i provutveckling, det ger ett ramverk för validering och det ger en grundval för bedömning av de föreslagna tolkningarnas och användningarnas validitet. I den utsträckning som tolkningsargumentet är koherent och dess inferenser och antaganden rimliga är de föreslagna tolkningarna och användningarna rättfärdigade, eller valida.

Referenser

- Angoff, W.H. (1988). Validity: An evolving concept. I H. Wainer & H. Braun (red.), *Test validity* (s. 9–13). Hillsdale, NJ: Erlbaum.
- Bachman, L. & Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford, UK: Oxford University Press.
- Brennan, R. (2001). *Generalizability theory*. New York, NY: Springer-Verlag.
- Chapelle, C.A., Enright, M.K. & Jamieson, J. (2008). *Building a validity argument for the test of English as a foreign language*. New York, NY: Routledge.
- Cronbach, L.J. (1971). Test validation. I R. L. Thorndike (red.), *Educational measurement* (2 uppl.) (s. 443–507). Washington, DC: American Council on Education.
- Cronbach, L.J. (1988). Five perspectives on validity argument. I H. Wainer & H. Braun (red.), *Test validity* (s. 3–17). Hillsdale, NJ: Erlbaum.
- Cronbach, L.J. (1989). Construct validation after thirty years. I R. E. Linn (red.), *Intelligence: Measurement, theory, and public policy*, (s. 147–171). Urbana, IL: University of Illinois Press.
- Cronbach, L.J. & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Cureton, E.E. (1951). Validity. I E.F. Lindquist (red.), *Educational measurement* (s. 621–694). Washington, DC: American Council on Education.
- Haertel, E.H. (2006). Reliability. I R. Brennan (red.), *Educational measurement* (4 uppl.) (s. 65–110). Westport, CT: American Council on Education and Praeger.
- Kane, M.T. (2006). Validation. I R. Brennan (red.), *Educational measurement* (4 uppl.) (s. 17–64), Westport, CT: American Council on Education and Praeger.
- Kane, M.T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1–73.

- Kolen, M. (2006). Scaling and norming. I R. Brennan (red.), *Educational measurement* (4 uppl.) (s. 155–220). Westport, CT: American Council on Education and Praeger.
- Messick, S. (1989). Validity. I R.L. Linn (red.), *Educational measurement*, (3 uppl.) (s. 13–103.) New York, N.Y.: American Council on Education and Macmillan.
- Shepard, L. (1993). Evaluating test validity. I L. Darling-Hammond (red.), *Review of research in education* (s. 405–450). Washington, DC: American Educational Research Association.
- Toulmin, S. (1958). *The uses of argument*. Cambridge, U.K.: Cambridge University Press.

Begreppslista

Analytisk bedömning Analytisk bedömning avser ett tillvägagångssätt i vilket flera aspekter eller dimensioner av något bedöms. Strikt analytisk bedömning innebär att det som är föremål för bedömning, exempelvis en text, får flera resultat. Så kan en text erhålla två separata resultat för till exempel språkbruk och stavning. Se särskilt kapitel 2.

Bedömersamstämmighet Bedömersamstämmighet avser i vilken utsträckning två eller flera bedömare når liknande slutsatser. Samstämmighet kan mätas på olika sätt, till exempel i vilken utsträckning två bedömare ger exakt samma betyg eller i vilken utsträckning två bedömare rangordnar elevsvar på samma sätt.

Bedömningsregel En bedömningsregel specificerar hur en elevlösning på en given uppgift ska rättas eller bedömas. Ett exempel är kortsvarsuppgifter i läspröv, där bedömningsreglerna anger vilka svar som accepteras som korrekta lösningar på uppgiften.

Extrapolering Extrapolering avser en slutsats från ett generaliserat provresultat (se generalisering) till kunskaper i en måldomän, exempelvis skrivförmåga. Det betyder att ett provresultat används som en indikation på elevens skrivförmåga även utanför provsammanhanget. Se kapitel 2.

Generalisering Generalisering avser ett antagande om att ett provresultat skulle vara identiskt eller nära identiskt om samma eller liknande uppgifter genomfördes vid ett annat tillfälle, eller bedömdes av andra lärare. Se kapitel 2.

Holistisk bedömning Till skillnad från analytisk bedömning innebär holistisk bedömning att en elevprestation ges ett sam-

manfattande omdöme. En elevtext kan exempelvis få betyget »B« baserat på att den uppfyller kriterier hämtade från kunskapskraven. Det betyder att bedömningen av olika dimensioner av texten (t.ex. interpunktion, språkbruk, innehåll) har vägts samman till ett resultat.

Konstrukt Konstrukt avser den förmåga som är föremål för bedömning. Ett exempel på konstrukt är skrivförmåga. Definitionen av konstruktet avgör hur uppgifterna som ligger till grund för bedömningen bör utformas för att konstruktrelevanta slutsatser ska vara möjliga att dra.

Konstruktirrelevant varians Konstruktirrelevant varians innebär att ett provresultat påverkas även av andra förmågor än de som man avsett att mäta. Ett vanligt exempel rör läsförmåga som ofta provas genom att elever i skrift återger en tolkning eller förståelse av en text. Det innebär att inte bara läsförmågan utan också skrivförmågan blir avgörande för resultatet.

Korrelation Anger samband mellan två variabler, exempelvis samband mellan socioekonomisk bakgrund och skolprestationer. Korrelation anges som ett värde mellan 1 och -1 , där värden nära 1 betyder att sambandet är starkt positivt, det vill säga höga värden på den ena variabeln sammanfaller med höga värden på den andra variabeln. Värden nära -1 indikerar starkt negativt samband, det vill säga att höga värden på den ena variabeln sammanfaller med låga värden på den andra variabeln. Värden omkring 0 betyder att samband saknas.

Läsförmåga Läsförmåga inbegriper alltifrån avkodning och läsflyt till förmåga att söka och föra samman information i text, kunna dra slutsatser på lokal och global nivå samt kritiskt reflektera över det lästa. Se kapitel 4.

Läsprocesser Med läsprocesser avses aspekter av läsandet som relaterar till de olika typer av kognitiv utmaning som läsaren ställs inför. I PISA-provet definieras tre läsprocessnivåer: söka och inhämta information; sammanföra och tolka det lästa; samt reflektera och utvärdera språk och innehåll i texten. Se kapitel 4.

Måldomän Ett konstrukt kan definieras i termer av en måldomän, som då avser alla de kunskaper och färdigheter som man genom en viss bedömning vill kunna uttala sig om. Måldomänen för konstruktet läsförmåga kan exempelvis antas innefatta alla de texter och aspekter av läsande som vi förväntar oss att elever ska behärska med utgångspunkt i kursplanen i svenska.

Reliabilitet Med reliabilitet avses bedömningens tillförlitlighet, det vill säga den grad med vilken man kan hävda att bedömningen inte beror av slumpmässiga fel eller mätfel. För att undersöka reliabiliteten i en bedömning är det vanligt att granska dels i vad mån upprepade bedömningar ger samstämmiga resultat, dels om olika bedömare värderar elevernas prestationer och insatser på samma sätt.

Skrivförmåga Skrivförmåga innebär att utifrån specifika syften kunna utföra givna skrivhandlingar genom att ta lämpliga skriftliga redskap i bruk. Se kapitel 3.

Tolknings-/användningsargument Tolknings-/användningsargument specificerar den kedja av inferenser som går från ett visst provresultat till de slutsatser man önskar kunna dra och den användning man vill göra utifrån provresultatet. Det bygger således på en serie argument, som i tur och ordning stödjer de antaganden om elevernas kunskaper och färdigheter som görs baserat på en bedömning. Se kapitel 10.

Underrepresentation av konstrukt Det att man i bedömningen inkluderat för lite av det som egentligen ska bedömas, det vill säga att innehållet i bedömningen inte helt och hållet representerar konstruktet. En bedömning av allmän skrivförmåga bör exempelvis inte enbart basera sig på uppgiften att skriva en novell, utan bör även inkludera andra sorters skrivande som vi innefattar i begreppet allmän skrivförmåga.

Uppgift Uppgifter avser de handlingar som en lärare ber elever att genomföra i syfte att skaffa underlag för bedömning. Det innebär att vi med uppgifter avser såväl formella situationer, till exempel när elever skriver ett nationellt prov, som mer informella situationer, exempelvis när elever ombeds läsa högt ur en bok.

Validitet Validitet handlar om giltigheten i en bedömning. En bedömning är att betrakta som valid i den mån det finns goda argument som stödjer de tolkningar och/eller beslut som görs baserat på den eller de prestationer som bedömts. Två vanliga hot mot validiteten utgörs av konstruktirrelevant varians (dvs. att bedömningen påverkas av sådant som inte är avsett att bedömas) och underrepresentation av konstruktet (dvs. att det som inkluderas i bedömningen inte representerar hela konstruktet).

Validitetsargument Validitetsargumentet utgör en bedömning av de föreslagna tolkningarna och användningarna av en viss bedömning. Dessa tolkningar och användningar är att betrakta som valida om varje steg i tolknings-/användningsargumentet är tydligt definierat och kan stödjas av goda argument. Se kapitel 10.

Ämnesspråk Refererar till det sätt på vilket språket används inom ramen för ett specifikt ämnesområde. Ämnesspråk handlar inte enbart om specifika begrepp och uttryck utan också om förväntade strukturer för hur man samtalar inom ämnet liksom hur texter skrivs och läses.

