

Likt, olikt, gemensamt – om god bedömning i språk

Gudrun Erickson

Det övergripande temat för denna text, som bygger på min presentation vid Språklärarnas riksförbunds konferens *Språk så in i Norden* i april 2016, är bedömning av kunskaper – generellt och i språk. Likheter och olikheter mellan olika sammanhang och syften diskuteras, men framförallt betonas de gemensamma, grundläggande drag som kännetecknar all pedagogisk bedömning av god kvalitet.

Från iskallt till glödhet...

Bedömning ses numera i regel som en del av den pedagogiska verksamheten, som något som på olika sätt hör ihop både med individers lärande och lärares undervisning. Att man som elev inte bara vill veta nivån på sina kunskaper vid enstaka tillfällen utan också närmare förstå hur kunskaperna ser ut och utvecklas är tämligen självklart. Som lärare är det svårt att tänka sig att intresset begränsas till undervisning enbart; man vill också få veta om, vad och hur mycket individuella elever lärt sig för att kunna planera vidare på ändamålsenliga sätt. Detta integrerade synsätt, med bedömning som en naturlig del av den pedagogiska processen, är emellertid ingalunda självklart utan påverkas både av kontextuella faktorer och av definitionen av begreppet bedömning.

I Sverige liksom i många andra länder har under senare decennier synen på bedömning genomgått en utveckling från ett i huvudsak tekniskt och mätrelaterat till ett mera pedago-

”I Sverige liksom i många andra länder har under senare decennier synen på bedömning genomgått en utveckling från ett i huvudsak tekniskt och mätrelaterat till ett mera pedagogiskt perspektiv. Denna förändring har varit märkbar både i allmän pedagogisk debatt och i lärarutbildning och fortbildning.”

giskt perspektiv. Denna förändring har varit märkbar både i allmän pedagogisk debatt och i lärarutbildning och fortbildning. Om man jämför 1980-talet med 1990-talet, skulle man kunna säga att intresset för bedömningsfrågor gick från det tämligen iskalla till det glödeta. Detta hänger till betydande del samman med de stora förändringar av skolans styrsystem som genomfördes i samband med införandet av nya läroplaner 1994 och som bland annat innebar att det normrelaterade, ”relativa” systemet ersattes med ett mål- och kunskapsrelaterat bedömnings- och betygssystem. I detta skulle individuella elevers kunskaper relateras till verbala beskrivningar av kunskap och resultaten kommuniceras på ett tydligt och begripligt sätt. Systemet krävde vidare att lärare tolkade kursplaner och kriterier på så likartat sätt

som möjligt och drog samma nivåmässiga slutsatser. Detta ställde givetvis stora krav och medförde att bedömningsfrågor allt mer kom i fokus, liksom det grundläggande kravet på rättvisa och likvärdighet. Olika undersökningar,

såväl från universitet som statliga myndigheter, visade dock snart på betydande problem med tolkningar och konkretiseringar av styrdokument och därmed med likvärdigheten i den svenska skolan. Till följd av detta blev frågor om bedömning och betygsättning mycket snabbt av största intresse, såväl nationellt som lokalt. Detta sammanföll med ett starkt internationellt fokus på bedömningsfrågor, inte minst på formativ bedömning, ibland refererat till som Bedömning för lärande (Black & Wiliam, 1998), och även på feedback (Hattie & Timperley, 2007). Det relativt hastigt påkomna intresset för bedömning var på många sätt positivt, men det hade också en del bekymmersamma effekter, som stundtals ledde till en undran om utvecklingen var på väg från det glödgheta till det något vidbrända.

'Allt är väl inte bedömning...?'

Efter en storföreläsning för blivande lärare fick jag en gång en fråga, som var en nyttig påminnelse dels om hur viktigt det är med ord, dels om att ord ibland kan få makt över tanken. Personen konstaterade först att bedömning ju verkade vara ett spännande område men sade sedan med viss desperation i rösten att 'allt väl ändå inte är bedömning'... Uppenbarligen hade jag inte varit tydlig nog kring definitioner och med det viktiga påpekandet att undervisningen självklart måste ge rejält utrymme för lärande utan att elever känner sig granskade och värderade i sin kunskapsutveckling. Vikten av att skapa rimlig tydlighet kring *när* bedömning sker och *vad* som bedöms berörs i den gemensamma europeiska referensramen för språk/GERS (Europarådet/Skolverket, 2009), där det varnas för att "göra livet till ett enda oändligt prov för eleven och till en byråkratisk mardröm för läraren" (sid. 183). Detta kan relateras till något vissa lärare, i uppenbar välmening, ibland säger till sina elever, nämligen att de inte behöver vara oroliga för prov, för läraren

ger inga, eller ytterst få, sådana utan bedömer 'hela tiden'. Detta är säkerligen avsett som en upplysning om att man visst kan visa att man kan saker utanför de mera formella bedömningstillfällena, men de flesta elever torde också höra ett annat budskap, nämligen att även tecken på att man inte förstått och inte kan komma att dokumenteras. Yttranden av detta slag kan möjligen lugna vissa men säkerligen oro många andra, som kommer att undra om man i det ständigt granskande läge som råder bör hålla tyst med undringar och frågor som kan visa på bristande insikt. Det vore mycket olyckligt om ord och uttrycks sätt på detta sätt skulle ge signalen att det självklara och nödvändiga, ibland korta, ibland långa, sökandet efter förståelse skulle värderas och bedömas i sig. Samtidigt som detta konstateras, är det naturligtvis också viktigt att inte ge intryck av att bedömningen reduceras till enstaka provtillfällen där väldigt mycket står på spel. Detta understryker hur viktigt det är att ha en öppen och kontinuerlig dialog både med elever och kolleger kring frågor om lärande, undervisning och bedömning och att tillsammans söka efter goda balanser och lösningar. En central fråga här är hur man kan och bör distingera mellan kommunikation kring kunskap, återkoppling i samband med detta samt bedömning som leder fram till ett sammanfattande omdöme, vare sig detta är i betygsform eller ej (Erickson & Gustafsson, 2014).

Det dubbla uppdraget

Bedömning har – och har alltid haft – två uppenbara grundfunktioner, nämligen att stödja lärande och att stödja rättvisa och likvärdighet. Detta innebär i det första fallet att, i samverkan med eleverna, kartlägga hur individuella kunskapsprofiler ser ut, klargöra vart man är på väg och komma fram till hur man ska ta sig dit på bästa sätt (Assessment Reform Group, 2002). Denna kontinuerliga process har också som

syfte att bidra till elevers ökade medvetenhet om och medagerande/agens i det egna lärandet. Bedömning innebär också att vid olika tillfällen summera hur långt lärandet kommit och hur det gestaltar sig samt att göra detta på så tillförlitliga och stabila sätt som möjligt. Man kan referera till detta som formativ respektive summativ bedömning eller som bedömning *för* respektive *av* lärande, men det är viktigt att betona att bakom båda ligger vissa gemensamma och grundläggande principer. Dessa kan diskuteras och preciseras på olika sätt, men det torde råda enighet kring att de har att göra med de vida begreppen *tydlighet, giltighet, tillförlitlighet* och *respekt*. Elever ska inte behöva gissa vad kunskapen handlar om eller hur den värderas, det som bedöms måste ha tydlig och god förankring i den ämnes- och kunskapssyn som uttrycks i styrdokumentet, bedömningar måste präglas av stringens och konsekvens, den etiska aspekten måste vara tydlig och respekt genomsyra hela processen. Sett i ljuset av detta är all dikotomisering av de båda bedömningsfunktionerna – att den ena skulle utesluta den andra – både osaklig och negativ, vilket också påpekats av forskare inom området (Harlen, 2012; Taras, 2005).

Grundläggande frågeställningar och begrepp

Likaväl som det finns vissa grundläggande principer för bedömning, kan det hävdas att vissa generella basfrågor med fördel kan användas dels för att ringa in och beskriva pedagogisk bedömning som fenomen, dels för att konkret planera bedömningar av skilda slag (Takala, Erickson, Figueras & Gustafsson, 2016). Frågan *varför* handlar om syften, *vad* om den

kunskap som ska bedömas, *hur* om olika tillvägagångssätt vid konstruktion av bedömningsmaterial och situationer, liksom vid analys och återkoppling av resultat; *när* berör tidpunkter, frekvens och sekvensering, *vem* aktörer och agens, och slutligen fokuserar frågan *och...?* användning och konsekvens. Samtliga frågor, och i synnerhet den sista, hänger tydligt samman med validitet, eller giltighet, som utgör en grundbult i kvalitetssäkringen av bedömning. Enligt den traditionella definitionen handlar validitet om att verkligheten bedöma det som ska bedömas, men begreppet har kommit att få en utvidgad betydelse, där slutsatser, användning och effekter är i fokus. I synnerhet två företeelser brukar anges som allvarliga hot mot validitet, nämligen att man har med för lite av den kunskap som avses respektive tvärtom, att man tar med sådant som inte riktigt hör dit, i sina bedömningar. Dessa två hot, som refereras till som *construct underrepresentation* och *construct irrelevant variance* (Messick, 1996), är ytterst påtagliga och förekommer både i mera vardagliga bedömningssituationer och i stora internationella studier. Ett tydligt exempel på underrepresentation är att muntlig språkfärdighet inte ingick i den europeiska språkstudien

”Elever ska inte behöva gissa vad kunskapen handlar om eller hur den värderas, det som bedöms måste ha tydlig och god förankring i den ämnes- och kunskapssyn som uttrycks i styrdokumentet, bedömningar måste präglas av stringens och konsekvens, den etiska aspekten måste vara tydlig och respekt genomsyra hela processen.”

2011 (Skolverket, 2012), vilket givetvis hade konsekvenser för resultaten. Vidare kan konstateras att om elevers förståelse av det de läst och hört enbart ska visas genom långa, skriftliga svar, finns risk att man på ett irrelevant sätt bedömer skrivande i stället för förståelse.

En viktig aspekt av validitet är tillförlitlighet, eller reliabilitet, dvs. att resultaten är så fria som

möjligt från slump och därmed går att lita på som indikatorer på en persons förmåga inom ett visst område. Detta kräver att uppgifter ska vara relevanta, klara och entydiga i sin utformning, men det handlar också om att bedömare är rimligt överens i sina bedömningar. Strävan efter samstämmighet är givetvis central men har ibland lett till oönskade effekter. Ett exempel är att man i vissa sammanhang undviker uppgifter med produktiva svar och skapar prov med bara flervalsfrågor, vilket blir besynnerligt givet en kommunikativ språksyn. Skolinspektionen har vidare i sina undersökningar konstaterat att bedömning av skriven text – mest i svenska, men även i engelska – genererar variationer i lärares bedömningar som anses för stora och därför förordad att skrivmomentet tas bort från nationella prov. Detta förslag har dock ej accepterats vare sig från vetenskapligt eller politiskt håll.

Begrepp som validitet och reliabilitet uppfattas ofta som rent teoretiska, med relevans för akademiska analyser av skilda slag men föga brukbara i konkreta pedagogiska sammanhang. Detta är att snäva in betydelsen av något som i själva verket fyller en högst användbar och konkret funktion, inte minst i planering och analys av arbetet med bedömning. En manifestation av detta finns i de *Guidelines for Good Practice in Language Testing and Assessment*, som tagits fram inom EALTA (*European Association for Language Testing and Assessment*). Detta dokument kan fritt laddas ner (www.ealta.eu.org/guidelines.htm) och innehåller dels en inledning kring det som kon-

stituerar god bedömning på ett generellt plan, dels riktade avsnitt till olika kategorier av användare, inte minst lärare. Företeelser som nämns kring analysen av den egna praktiken är ordnade under huvudrubrikerna *Bedömnings syfte(n) och specificering*, *Bedömningsmetoder* och *Konsekvenser*. Riktlinjerna, som finns tillgängliga på 35 språk, har visat sig fungera väl och används runt om i Europa och övriga delar av världen som stöd för utveckling av goda metoder kring språklig bedömning.

”En viktig aspekt av validitet är tillförlitlighet, eller reliabilitet, dvs. att resultaten är så fria som möjligt från slump och därmed går att lita på som indikatorer på en persons förmåga inom ett visst område. Detta kräver att uppgifter ska vara relevanta, klara och entydiga i sin utformning, men det handlar också om att bedömare är rimligt överens i sina bedömningar.”

Och bedömning i språk då...?

Det mesta som sagts hittills i denna text handlar om bedömning på ett allmänt och ämnesövergripande plan, om än med hänvisningar till språk. Detta är medvetet för att visa att många av de teorier och förhållningssätt man behöver för att hantera bedömning på goda och flexibla sätt är generaliserbara över ämnen och kontexter, vilket i sig ger möjlighet till synergier och samverkan i vidare bemärkelse än vanligt.

Dock finns det självklart också ett behov att konkretisera med ämnesrelaterad kunskap. Det finns här många goda referenser till litteratur, vars innehåll och referenslistor fördjupar frågan. Att göra ett urval är däremot hart när omöjligt. Nämnas kan dock den ”*Assessing-serie*” som utgivits av Cambridge University Press, med fokus på olika områden – receptiva såväl som produktiva och interaktiva färdigheter, språkliga *elements* som ordförråd och grammatik, bedömning för särskilda syften, datorbaserad bedömning, bedömning av yngre elevers språkfärdighet osv. Vidare finns ett stort antal

referenser till material och litteratur kring GERS och den europeiska språkportfolion på Europarådets hemsida (www.coe.int/t/dg4/linguistic/cadre1_en.asp). Bland tidskrifter med fokus på bedömning i språk kan i synnerhet två nämnas, nämligen *Language Testing* och *Language Assessment Quarterly*. Det är också vanligt att mera allmänna tidskrifter med språkdidaktisk inriktning innehåller artiklar kring bedömningsrelaterade frågor, då ofta med fokus på mera kontinuerliga former.

Inom ramen för denna begränsade text kommer jag i det följande enbart att ge några korta, språkrelaterade kommentarer på basis av två av de grundfrågor som nämnts tidigare. En helt fundamental sådan är givetvis *Vad?*, dvs. den som avser själva kunskapsobjektet – det som ska läras, undervisas och bedömas, ofta refererat till som *the construct*. Detta finns först och främst beskrivet i skolans kurs- och ämnesplaner, men eftersom dessa är tydligt relaterade till GERS finns även där kunskap och inspiration att hämta. I dessa dokument preciseras den kommunikativa språksyn, som sedan 1980-talet varit tydlig i de svenska styrdokumenterna och som i GERS refereras till som en handlingsorienterad språksyn. Språkens användning för olika syften och sammanhang står följaktligen i fokus på ett tydligt sätt, såväl undervisnings- som bedömningsmässigt. Tolkningen av detta vållar ibland problem, i synnerhet vad gäller synen på mål och medel – vad som behövs för att språket ska fungera. Det kan här vara på sin plats att ta till en traditionell bild av relationen mellan del och helhet och likna utvecklingen av ett brukbart språk vid bygget av ett välfungerande och vackert hus. Grunden skulle då kunna ses som en stabil och genomtänkt, kommunikativ språksyn. För att bygga huset behövs givetvis ett antal beståndsdelar, som väljs och används strategiskt med anpassning

till syftet, mottagaren och situationen. Dessa ”element” är inte minst brädor eller tegelsten – ord och fraser – spik eller murbruk för att få det hela att hänga ihop – grammatik och syntax – samt färg och puts för att skapa ett gott ytskikt, som fyller en funktion både för hållbarhet och skönhet – uttal och stavning. Utan beståndsdelar av god kvalitet blir det inget hållbart hus att leva och trivas i. Detta torde knappast vara möjligt att ifrågasätta, men det är också väsentligt att fastslå att beståndsdelarna *inte* är huset... Metaforen må ha sina luckor och kan kritiseras på olika sätt, men den har visat sig brukbar i diskussionen av hur man kan se på utvecklingen av språk, att det inte råder något motsatsförhållande mellan det som ibland refereras till som *language elements* och *language skills*, men att det i slutändan givetvis är huset i sig som utvärderas.

Vad gäller frågan *Vem?* finns mycket att säga, om subjekt och objekt, ansvarsfördelning, samverkan och respekt. Först och främst torde det vara så, att om man som lärare känner att bedömning enbart är ens eget ansvar blir bördan blytung; man måste ”bäras åt”, som en mig närstående två-och-ett-halvtåring uttryckte det en gång (Erickson, 2009). Detta innebär att man på genomtänkta och välunderbyggda sätt använder sig av elevers självbedömning och kamratbedömning, liksom av kollegial samverkan och extern bedömning. Det finns ett stort värde i att fler ögon tittar, öron lyssnar och röster hörs i arbetet med lärande, undervisning och bedömning. Detta har att göra med de grundläggande principer som tidigare nämnts: tydlighet, giltighet, tillförlitlighet och respekt, och det ökar möjligheterna att göra adekvata iakttagelser och använda dem väl. Vad gäller elevers agens finns även god inspiration att hämta i GERS och den europeiska språkportfolion, som på ett systematiskt sätt fokuserar elevers språkliga medveten-

het och ansvar för det egna lärandet (Little & Erickson, 2015). Angående frågan om *Vem?* är det slutligen väsentligt att låta den innefatta både subjekt och objekt. Bedömning handlar om att så väl som möjligt kartlägga utveckling och kunnskap 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 detta 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.

Detta citat får bilda avslutning på en text om ett ämne som har en central roll i den språkpedagogiska och didaktiska processen och som mål har att bidra såväl till gott, effektivt och njutbart lärande som till rättvisa och likvärdighet.

Referenser:

Assessment Reform Group (2002). *Research-based principles of assessment for learning to guide classroom practice*. http://methodenpool.uni-koeln.de/benotung/assessment_basis.pdf

Bachman, L. F. (1990). *Fundamental considerations in Language Testing*. Oxford: Oxford University Press.

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.

Erickson, G. (2009). "Att bäras åt" – Om den goda bedömningens flerfaldighet och ömsesidighet. I Tornberg, U. m. fl. *Språkidaktiska perspektiv. Om lärande och undervisning i främmande språk* (s. 159-174). Stockholm: Liber.

Erickson, G. & Gustafsson, J-E. (2014). Bedömning för Lärande och Likvärdighet. I Lundgren, U.P., Sälljö, R & Liberg, C. (red.), 3 utg., *Lärande, skola, bildning* (s. 559-594). Stockholm: Natur & Kultur.

Europarådet/Skolverket (2009). *Gemensam europeisk referensram för språk: lärande, undervisning, bedömning*. Stockholm: Skolverket.

Harlen, W. (2012). On the Relationship between Assessment for Formative and Summative Purposes. I J. Gardner (ed.), *Assessment and Learning*, Second edition; s. 87-102. London: Sage Publications.

Hattie, J. & Timperley, H. (2007) The power of feedback. *Review of Educational Research*, 77(1), 81-112.

Little, D. & Erickson, G. (2015). Learner Identity, Learner Agency, and the Assessment of Language Proficiency: Some Reflections Prompted by the Common European Framework of Reference for Languages. *Annual Review of Applied Linguistics*, 120-139.

Messick, S. A. (1996). Validity and Washback in Language Testing. *Language Testing* 13, 241-256.

Skolverket (2012). *Internationella språkstudien 2011. Rapport 375*. Stockholm: Skolverket.

Taras, M. (2005). Assessment – Summative and Formative – Some Theoretical Reflections. *British Journal of Educational Studies* 53(4), 466-478.

Takala, S., Erickson, G., Figueras, N. & Gustafsson, J-E. (2016). Future Prospects and Challenges in Language Assessments. I Tzagari, D. & Banerjee, J. (eds.), *Contemporary Second Language Assessment, Contemporary Applied Linguistics* (pp. 299-315). London: Bloomsbury Academic.



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

Hon har lång erfarenhet av nationellt och internationellt arbete kring lärande, undervisning och bedömning, är projektledare för utvecklingen av de nationella prov- och bedömningsmaterialen i engelska och moderna språk samt var 2013-2016 president i EALTA. Gudrun är sedan 2006 hedersmedlem i Språklärarnas riksförbund.