

Lärares betygsättningspraktik

Alli Klapp Lekholm

Lärares betygsättningspraktik är av största intresse för att förstå vad betygen mäter i termer av elevernas kunskaper och förmågor. Att sätta betyg är en myndighetsutövning utan möjlighet för eleverna att överklaga, och lärarnas beslut om vilket betyg en elev ska få har ofta långtgående konsekvenser för elevens möjlighet till val av vidare studier. I denna artikel diskuteras betygens validitet och att betygen mäter både elevernas ämneskunskaper och personliga egenskaper som till exempel motivation.

I ETT KLASSRUM PÅGÅR en mängd olika skeenden byggda på relationer vilket innebär att läraren måste hantera situationer som kräver snabba och ibland svåra beslut. Det handlar om undervisning, relationer mellan elever och lärare och mellan elever, öppna och dolda konflikter, sociala problem, frånvaro, elevernas attityder till lärande och deras olika motivationsnivå. Det som händer i klassrummet ingår i ett komplext sammanhang där läraren ska få vardagen i klassrummet att fungera; eleverna ska tillgodogöra sig undervisning och utvecklas till demokratiska, kritiskt tänkande medborgare och förhoppningsvis gå ut skolan med godkända betyg. Samtidigt som läraren leder arbetet i klassrummet ska han eller hon bedöma elevernas prestationer, kunskaper och förmågor och denna bedömning ligger sedan till grund för de betyg eleverna får.

Eftersom det är lärarna som bedömer, utvärderar och betygsätter elevernas kunskaper och förmågor

är deras betygsättningspraktik av stort intresse för att förstå vad betygen är ett mått på. Ett grundläggande antagande är att individers kunskaper och förmågor varierar och därför går att mäta. Betyg kan ses som ett summerat mått på elevernas kunskaper och förmågor från många olika mätningar, till exempel klassrumsbedömningar, inlämningsuppgifter och olika typer av prov. Det rådande betygssystemet existerar inom ett visst utbildnings- och bedömningssystem som är konstruerat utifrån vissa grundläggande värderingar, utbildningsteorier och epistemologiska antaganden. Dessa olika antaganden utgör ett fundament från vilka vi kan förstå lärarnas betygsättningspraktik och på vilka grunder betyg sätts.

I Sverige har lärarna en stark autonomi vilket innebär att det är läraren själv (ibland i samarbete med eleverna) som planerar undervisningen, undervisar och i slutändan bedömer och betygsätter

sina elever. Det existerar inte något externt kontrollsystem för lärarnas bedömning och betygsättning av elevernas kunskaper och förmågor utan det är upp till läraren själv att tolka styrdokumentet, bedöma och betygsätta eleverna. Ett långtgående decentraliserat system för bedömning och betygsättning har givetvis vissa fördelar såväl som nackdelar. Inom ett decentraliserat system har lärarna stor makt att bedriva undervisning utifrån egna önskemål. Läraren väljer egna metoder och bedömer och betygsätter eleverna utan standardiserade kontrollinstrument vilket innebär stor frihet att bedriva den undervisning han eller hon önskar. Det som kan vara problematiskt med ett långtgående decentraliserat bedömningssystem handlar om hur vi ska kunna säkra likvärdigheten, att lärarna gör likvärdiga bedömningar av elevernas kunskaper och förmågor och tar likvärdiga beslut när de betygsätter eleverna.

Det finns en stark oro för att vårt bedömningssystem inte lever upp till kraven på likvärdighet och rättvisa när det gäller bedömningar och betygsättning då forskning indikerar att lärarna gör subjektiva bedömningar av elevernas kunskaper och förmågor och att det finns stora skillnader mellan lärare och skolor när det gäller betygsättning (Klapp Lekholm & Cliffordson, 2008; Skolverket, 2005).

I vårt bedömningssystem sker lärarnas bedömningar av elevernas kunskaper och förmågor främst i undervisningssituationen i klassrummet. Klassrumsbedömningar innebär att läraren bedömer elevernas prestationer, kunskaper och förmågor i ett socialt sammanhang och att läraren anpassar undervisningen efter elevernas behov. Detta innebär att variationen mellan olika klassrum är stor och att en standardiserad form av undervisning och bedömningsmetoder inte är relevant. Lärarens beslut om elevernas betyg grundade på bedömningar i klassrummet kan få långtgående conse-

kvenser för eleven eftersom betyg är vårt främsta instrument för antagning till nästa nivå i utbildningssystemet och därmed av stor betydelse för elevens möjligheter till vidare utbildning. Att lärarna gör likvärdiga, rättvisa och jämförbara bedömningar av elevernas kunskaper och förmågor är av stor betydelse för betygens legitimitet, för elevernas rättsäkerhet och för hela vårt utbildningssystem.

Enligt styrdokumentet är det enbart elevernas kunskaper som ska bedömas och ligga till grund för betygsättning, det vill säga att betygen förutsätts vara endimensionella. Detta innebär att betygen ska vara likvärdiga mellan lärare, skolor och över tid och därmed legitimeras deras roll som urvalsinstrument och som utvärderingsinstrument för utbildningssystemet. Lärarna tolkar styrdokumentet, till exempel läroplaner och kursplaner, och då finns det givetvis skillnader mellan lärare och mellan skolor i hur dessa tolkas. Dock tillåter vårt system att undervisningens innehåll och metod kan variera mellan lärare och skolor medan bedömningarna och betygsättningen av elevernas kunskaper och förmågor däremot inte får variera. Detta innebär att lärarna ska ha ett gemensamt förhållningssätt till bedömning och tolka elevernas kunskaper och förmågor på ett likvärdigt sätt när de betygsätter eleverna.

Samhället förutsätter alltså att betyg är ett objektivi- tätt mått på ämneskunskaper. En mängd forskning visar dock att lärarna bedömer elever utifrån olika grunder och personliga föreställningar (Brookhart, 1993; Hidi, Renningen & Krapp, 2004). Lärare inom olika ämnen väger till exempel kunskaper och förmågor olika. Lärare inom språk tycks exempelvis väga in både ämneskunskaper och elevernas personliga egenskaper i betyget medan lärare i matematik tar hänsyn till elevernas personliga egenskaper först när eleverna befinner sig på gränsen

till ett högre betyg (Pilcher, 1994). Det objektiva förhållningssättet förutsätter att läraren kan bedöma eleverna utan att ta in övrig information om eleven vilket kan innebära en viss motsättning då läraren dels ska ha en god relation till eleven och stötta elevens lärande och samtidigt kunna ta beslut om betyg som kan ha långtgående negativa konsekvenser för eleven. Att arbeta som lärare innebär att han eller hon utvecklar ett personligt, nära och subjektivt förhållningssätt till sina elever medan bedömningsprocessen ställer krav på att läraren intar en objektiv och distanserad hållning gentemot sina elever. Black och Wiliam (1998) menar att bedömningsprocesser är sociala processer som utspelas i sociala sammanhang, utförda av, för och med individer som aktörer.

För att öka rättssäkerheten för eleverna när det handlar om betygsättning bör lärarnas bedömnings- och betygsättningspraktik göras så transparent som möjligt. Lärare, elever, föräldrar och samhället i övrigt bör veta på vilka grunder betygen sätts och hur lärarna bedömer elevernas kunskaper och förmågor och kunna utvärdera att lärarnas bedömningar och beslut sker på ett rättvist och likvärdigt sätt.

Lärarnas bedömningspraktik

McMillan (2003) har utvecklat en teori kring lärarnas bedömningspraktik som kontrasterar de traditionella bedömningsprinciperna med inlärnings-teorier. Traditionella mättekniker, som används för att utföra storskaliga, objektiva prov, kan vara svåra att applicera på bedömningar som sker i klassrummet. I klassrummet är det många olika faktorer som lärarna måste ta hänsyn till när de bedömer elevernas kunskaper och förmågor och det handlar även om lärarens egna värderingar kring vad kunskap är, lärarens förhållningssätt till lärande och vilka kunskaper läraren har om bedömning och

hur bedömning påverkar elevernas lärande.

Ytterligare faktorer som har inverkan på bedömning och betygsättning av elever är olika externa faktorer såsom skol- och kommunalpolitiska åtgärder och föräldrars inverkan. Likaså kan olika typer av externa tester påverka lärarnas undervisning, bedömningar och betygsättning. Vardagen i klassrummet inbegriper alla dessa olika faktorer och påverkar arbetet i klassrummet. Miljön i klassrummet, elevernas bakgrund, elevernas attityder till lärande och deras uppförande interagerar med lärarens värderingar och mål med undervisningen.

Dessa olika interagerande och konkurrerande faktorer påverkar lärarnas bedömningar av elevernas prestationer och speglar lärarnas egna värderingar kring elevernas lärande och motivation, den praktiska situationen i klassrummet och de externa faktorerna som till exempel nationella prov. Lärarna kan också tänkas ta beslut kring bedömningar som utgår från andra grunder, till exempel att lärarna vill att eleverna ska lyckas med sitt skolarbete och därför modifierar och anpassar bedömnings-situationerna så att eleverna har möjlighet att prestera. Enligt McMillan (2003) ändrar lärarna bedömningarna i enlighet med olika karaktärsdrag hos eleverna och grundar sin bedömning och betygsättning så att elevernas motivation och engagemang för lärande ska bibehållas och stärkas. Att läraren tar hänsyn till dessa faktorer i bedömnings- och betygsättningsprocessen gör dock betygsättningsprocessen implicit och mindre transparent.

Ett flertal forskare menar att bedömnings- och betygsättningsprocessen måste omförhandlas för att bättre anpassas till vår nuvarande syn på lärande och undervisning (Brookhart & Durkin, 2003; McMillan, 2003). Det är inte lämpligt att använda traditionella mätprinciper för att bedöma elevernas kunskaper i klassrummet. Att samla, tolka och utvärdera information är en pågående process och

bedömningar som sker i klassrummet och lärares betygsättningspraktik är en subjektiv och intuitiv process där lärarna behöver kompetens och förmåga att utvärdera sina tolkningar och effekterna av dessa. Det är nödvändigt att fokusera på hur vi kan validera klassrumsbedömningar och lärarnas betygsättning genom att använda en bredare definition av validitet som erkänner att lärarnas beslut är just subjektiva och intuitiva och där konsekvenserna av deras beslut också är en del i valideringsprocessen. Därför blir lärarnas förmåga att analysera sin egen bedömnings- och beslutsprocess av största vikt för att deras beslut ska uppfattas som rättvisa, likvärdiga och legitima (Shepard, 2000).

Ett annat perspektiv för att tydliggöra den motsättning som lärarna hanterar är genom en distinktion mellan en kvantitativ och en kvalitativ modell för bedömning. Den kvantitativa modellen för bedömning innebär att läraren försöker eliminera influenser från elevens sociala bakgrund, den fysiska kontexten i klassrummet såväl som elevernas personliga egenskaper när bedömningar och betygsättning görs. Den kvalitativa modellen för bedömning innebär däremot att läraren använder all information om eleven och elevens prestationer, kunskaper och förmågor och personliga egenskaper vilka integreras i bedömningen för att läraren ska kunna få en säker bild av elevens kunskaper och förmågor (Kane, 2006).

Betygens dimensionalitet och effekter av elevernas personliga egenskaper

Ett flertal studier har visat att betyg inte enbart mäter elevernas ämneskunskaper utan även olika personliga egenskaper såsom ansträngning, intresse, motivation och uppförande (Klapp Lekholm & Cliffordson, 2009; McMillan, Myran & Workman, 2002).

För att kunna förstå vad betygen egentligen är ett mått på och vilka kunskaper, förmågor och egenskaper som lärarna betygsätter har betygens dimensionalitet undersökts i tre olika studier (Klapp Lekholm & Cliffordson, 2008; Klapp Lekholm & Cliffordson, 2009; Klapp Lekholm, 2008). Ett huvudintresse har varit att undersöka om betygen mäter annat än ämneskunskaper, till exempel elevernas personliga egenskaper såsom intresse och motivation. Vidare har studierna undersökt hur olika elev- och skolkaraktäristika påverkar betygen och hur olika systematiska skillnader i betygen beror på olika skolkaraktäristika, till exempel skillnader mellan kommunala och fristående skolor. Utifrån resultaten är det möjligt att se att betygen är flerdimensionella, det vill säga att betygen mäter flera dimensioner av elevernas kunskaper, förmågor och egenskaper.

Betygen innehåller dels en ämnesrelaterad dimension där elevernas ämneskunskaper finns med, dels en dimension som är gemensam över flera betyg och som inte är relaterad till elevernas ämneskunskaper. Den gemensamma dimensionen i betygen mäter främst elevernas personliga egenskaper. Betygen mäter till allra största del elevernas ämneskunskaper men ungefär tre till fem procent av betygen innehåller annat än ämneskunskaper vilket innebär att lärarna tar hänsyn till annat än elevernas ämneskunskaper när de betygsätter sina elever. Det kan synas vara en låg procentsats som förklaras av elevernas personliga egenskaper men i och med att denna dimension existerar för flera ämnen blir effekten stor när ämnesbetygen summeras i jämförelsetalet (det summerade betyget som används som antagningsinstrument) i avgångsbetyget.

Att elever anser att de är kompetenta i ett visst ämne, att de trivs i skolan, att deras föräldrar är engagerade i deras skolarbete och att de samarbetar

med andra har en positiv påverkan på den ämnesrelaterade dimensionen i betygen. Elever som tycker att de klarar av skolarbetet bra och som har en positiv tro på sin egen förmåga i olika ämnena når högre betyg (Klapp Lekholm & Cliffordson, 2009). Det verkar även som om föräldrarnas engagemang i elevens skolarbete har större betydelse för betygen i svenska än i engelska och matematik.

Elevernas specifika intresse av att lära sig mer inom ett visst ämne påverkar betyget positivt medan elevernas generella intresse att lära sig mer i skolan påverkar den gemensamma dimension som främst handlar om elevernas personliga egenskaper. Betydelsen av elevernas personliga egenskaper för betygen är större för de elever som upplever att de inte klarar av skolarbetet så bra. Detta kan tolkas som att lärarna använder sig av en kompensatorisk betygsättningspraktik där lärarna kompenserar de lågpresterande eleverna med hjälp av deras intresse och motivation att lära sig mer i skolan.

Detta stöds också av Stiggins et al. (1989) som fann att elever som är högpresterande betygsätts utifrån kunskaper, medan lågpresterande elever betygsätts utifrån både kunskaper och personliga egenskaper. Likaså är dessa resultat i linje med tidigare forskning som visar att lärare utvecklar olika rationaliteter för betygsättning som inkluderar lärarnas egna moraliska ställningstaganden, till exempel att lärarna tar hänsyn till individuella skillnader hos eleverna och strävar efter att underlätta och stötta elevernas lärande. Att betygen mäter elevernas personliga egenskaper som till exempel motivation, innebär att den kompensatoriska betygsättningspraktiken, där lärarna anpassar betygsättningspraktiken efter elevernas kunskaper, förmågor och personliga egenskaper även innebär att lärarna tar hänsyn till de konsekvenser betyget får för eleverna, lärarna och skolan.

Resultatet av studierna visar även att flickor får högre betyg i svenska och engelska och att deras personliga egenskaper till större del räknas in i betygen i jämförelse med pojkarna. En viktig mekanism bakom flickornas högre betyg verkar vara att flickor har ett större intresse och motivation för skolarbetet (Klapp Lekholm & Cliffordson, 2009). Murphy (2000) menar att flickornas bättre betyg beror på att flickorna kommer till skolan bättre förberedda för skolarbete och att de utvecklat attityder till lärande och utbildning som ligger väl i linje med skolans arbete och syn på lärande och därför blir flickorna bättre belönade i betygen. Flickornas motivation blir då en egenskap som indirekt påverkar inlärning och kan då ses som relevant för betygen.

Även om elevernas personliga egenskaper kan tyckas vara irrelevanta för betygen och inte bör ingå i dem är det också möjligt att diskutera detta utifrån ytterligare ett perspektiv. I läroplanen ingår dels specifika mål i kursplanerna för varje enskilt ämne där eleverna ska uppnå vissa mål för ämnen, dels ingår flera övergripande mål, giltiga för alla ämnen, som bland annat handlar om att eleverna ska utveckla en positiv attityd till lärande och skolan och utveckla motivation för lärande. Att betygen innehåller en dimension som är gemensam för flera betyg innebär att betygen innehåller eller mäter förmågor och egenskaper hos eleverna som är viktiga för alla dessa betyg. En förklaring kan vara att vissa av elevernas förmågor och egenskaper som är viktiga för prestation i de flesta ämnen, till exempel motivation, inte fångas i den ämnesrelaterade dimensionen. En viktig fråga att diskutera är huruvida elevernas motivation är en relevant eller irrelevant variabel i betygen.

Inkapslingsteorin kan bidra till förståelsen av vad betygen utgör ett mått på (Gustafsson & Carlstedt, 2006). Denna teori menar att människor föds med

en generell förmåga som har en övergripande påverkan på individers utveckling av kunskaper och förmågor, som utvecklas genom de erfarenheter som barn gör under sin uppväxt. Omsatt till skolan innebär detta att en elevs generella förmåga finns med i alla typer av situationer och undervisningssammanhang och påverkar utvecklingen av elevens specifika kunskaper och förmågor i till exempel ett visst skolämne. För att testa individers generella förmåga används test som främst mäter primära kognitiva förmågor, till exempel intelligens-test av olika slag. Skolans nationella prov däremot mäter elevernas kunskaper och förmågor i relation till läroplanen och mäter både elevernas kognitiva och icke-kognitiva förmågor. Likaså reflekterar betygen ett brett spektrum av elevens kognitiva såväl som icke-kognitiva förmågor, till exempel motivation, självdisciplin och intresse.

Det har visat sig att betygen har en god förmåga att predicera studieframgång. Detta förklaras med att elevernas samlade specifika kunskaper och förmågor inkluderar både elevens generella kognitiva förmåga och individuella egenskaper vilka är av betydelse för lärande och studieframgång. Betygen reflekterar elevernas kunskaper och förmågor och inkluderar både ämnesspecifika kunskaper och personliga egenskaper, till exempel motivation för lärande, vilket innebär att just de personliga egenskaperna kan avgöra elevens framgång inom skolsystemet och resultat i högre utbildning (Cliffordson, 2004, 2008).

Betygens validitet

Betygen är av stor betydelse för eleverna då de fungerar som det primära antagningsinstrumentet för nästa steg i utbildningssystemet. Konsekvenserna för dem som blir bedömda är allvarliga och påverkar ofta individernas möjligheter till vidare utbildning. I en perfekt värld skulle bedömningar

av elevernas kunskaper och förmågor vara desamma oavsett vem som bedömer, när bedömningen görs och var någonstans det sker. Kvaliteten på lärarnas bedömningar, tolkningar och beslut är av stor betydelse eftersom de har långtgående konsekvenser för eleverna men även för samhället i stort.

I ett bedömningssystem där beslut om ett visst betyg grundar sig på många olika bedömningar vid olika tillfällen och där flera lärare har gjort bedömningar höjs rimligtvis kvaliteten och reliabiliteten på betygen jämfört med om beslut om betyg tas utifrån något enstaka bedömningstillfälle där till exempel eleverna tillfälligt presterar över eller under sin egentliga förmåga. Betyg är alltså ett summativt mått grundat på ett stort antal bedömningssituationer, och denna bedömningsprocess utgör en grund för validering av lärarnas betygsbeslut där varje informationskälla bidrar till att lärarna kan göra en rimlig bedömning och ta beslut om vilket betyg en elev ska få. Det är främst lärarens slutsatser och tolkning av elevernas resultat som måste hålla hög kvalitet, inte själva bedömningsmomentet i sig. Validering innebär ett sätt för läraren att samla bevis för eller emot en viss tolkning och beslut, vilket innebär att det inte är bedömningsproceduren eller provpoängen i sig som valideras utan lärarnas tolkningar och slutsatser.

Information som lärare samlar från varje elev för att kunna ta beslut om ett visst betyg är tänkt att fungera som argument för eller emot ett visst betyg. Irrelevant varians i betygen innebär att betygen innehåller sådant som inte är relevant för betyget. Det kan till exempel handla om att elevens könstillhörighet påverkar betyget eller att minoritetselever blir felaktigt bedömda och får ett felaktigt betyg på grund av deras grupptillhörighet. Det kan också handla om att beslut om ett visst betyg grundar sig på bedömningar som mäter irrelevanta

kunskaper för just det ämnet som är i fokus, till exempel att specifika språkkunskaper krävs för att kunna lösa uppgifter i matematik.

Ett annat hot mot betygens validitet handlar om att de bedömningssituationer som lärarnas beslut grundar sig på inte innehåller vissa viktiga ämnesområden som bör vara med. Till exempel att lärare bedömer elever och tar beslut om betyg utifrån bedömningar som inte hanterar hela den domän eller ämnesområde som är i fokus och som betyget ska avspegla. Ytterligare en aspekt som också bör diskuteras i detta sammanhang är vilken typ av kunskaper och förmågor som är i fokus. I vårt utbildningssystem vill vi att eleverna utvecklar en viss typ av komplexa kunskaper och förmågor såsom analytiska och reflekterande förmågor, kritiskt tänkande och utvecklande av demokratiskt tankesätt och förhållningssätt till kunskap. Dessa typer av kunskaper tangerar elevernas personliga egenskaper och förhållningssätt till lärande.

Gustafsson och Balke (1993) och Andersson (1998) hittade en generell skolprestationsdimension i betygen. Gustafsson och Balke visade att det existerar en relativt stor del varians i betygen som inte är relaterad till kognitiv förmåga. Andersson (1998) antog att den generella skolprestationsfaktorn kunde vara relaterad till elevernas personliga egenskaper som till exempel motivation. Att betygen är flerdimensionella ligger också väl i linje med inkapslingsteorin som framhåller att betygen är ett mått som innehåller kulturellt värderad kunskap (Gustafsson & Carlstedt, 2006). Betygen ses som ett brett mått som reflekterar en mängd kognitiva såväl som icke-kognitiva förmågor hos eleverna. Den gemensamma dimensionen kan alltså fånga ett flertal olika egenskaper hos eleverna som värderas som något positivt av lärarna. Elever som vill nå goda resultat i skolan måste prestera utifrån både de formella och de informella krav som ställs

på dem från skolans sida. Till exempel tycks motivation för att lära sig mer vara en grundläggande egenskap och inställning hos eleven, som är nödvändig för att nå goda resultat i skolarbetet men som inte finns med som ett explicit kriterium i kursplanerna för ämnesbetygen (Klapp Lekholm & Cliffordson, 2009). Kursplanerna innehåller explicita kriterier för varje ämnesbetyg och betygsnivåerna, men också förgivettaganden kring de övergripande attityder som eleverna bör utveckla. Kriterierna för de olika ämnesbetygen är relaterade till elevernas kunskaper och förmågor i varje enskilt ämne (Skolverket, 2008).

Frågan om huruvida elevernas motivation för lärande är relevant för betygen är relaterad till hur vi tolkar läroplanen och kriterierna för varje enskilt ämne. I svenska, engelska och matematik är inte motivation en explicit förmåga som ska finnas med i betygen och kan därför ses som något som skapar betygsskillnader mellan eleverna. Dock är ett av de övergripande syftena med bedömnings- och betygssystemet att utveckla och stärka elevernas motivation för lärande. Om motivation är en underliggande egenskap som är av stor betydelse för elevernas inlärningsprocesser, prestationer och för att lyckas inom utbildningssystemet, då är motivation av stor vikt och relevans för betygen och därför något som alla elever bör vara medvetna om.

Bergenhenegouwen (1987) menar att de implicita kraven från lärarnas och skolans sida delvis är omedvetna och att eleverna därför inte får vetskap om hur de kan möta informella och implicita krav. Om inte de informella och implicita krav och normer som existerar i utbildningssystemet görs explicita i undervisningen blir resultatet ett utbildningssystem utan transparens, som är orättvist och utan likvärdighet och med stora skillnader mellan lärare och skolor. Att bedömnings- och betygsättningspraktiken är transparent och öppen är en grundlägg-

gande förutsättning om betygen ska kunna utvärderas och valideras så att de blir ett rättvist mått på elevernas kunskaper och förmågor.

Referenser

- Andersson, A. (1998). The dimensionality of the leaving certificate in the Swedish Compulsory school. *Scandinavian Journal of Educational Research*, 42(1), 25–40.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35–36.
- Brookhart, S. M. (1993). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 30(2), 123–142.
- Brookhart, S. M. & Dunkin, D. T. (2003). Classroom assessment, student motivation and achievement in high school social studies classes. *Applied Measurement in Education*, 16(1), 27–54.
- Bergenhengouwen, G. (1987). Hidden curriculum in the university. *Higher Education*, 16, 535–543.
- Cizek, G. J., Fitzgerald, S. & Rachor, R. E. (1995). Teachers' assessment practices: Preparation, isolation and the kitchen sink. *Educational assessment*, 3(2), 159–179.
- Cliffordson, C. (2004). Betygsinflation i de målrelaterade gymnasiebetygen. *Pedagogisk forskning i Sverige*, 9(1), 1–14.
- Cliffordson, C. (2008). Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education. *Educational Assessment*, 13(1), 56–75.
- Gustafsson, J.-E. & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28(4), 407–434.
- Gustafsson, J.-E. & Carlstedt, B. (2006, augusti). Abilities and grades as predictors of achievement: The encapsulation theory of intelligence: New evidence. Paper presented at the symposium "The investment theory of intelligence: New evidence, new challenges" at the annual meeting of the American Psychological Associations, New Orleans.
- Hidi, S., Renninger, K. A. & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (Eds.), *Motivation, emotion and cognition: Integrative perspectives on intellectual functioning and development* (89–115). Mahwah, NJ: Lawrence Erlbaum Association, Inc.
- Kane, M. (2006). Validation. In R.L Brennan (Ed.), *Educational Measurement* (17–64). Westport, CT: Praeger Publishers.
- Klapp Lekholm, A. (2008). Grade and grade assignment: effects of student and school characteristics. *Göteborg studies in educational science*, Vol. 269. Göteborg: Acta Universitatis Gothoburgensis.
- Klapp Lekholm, A. & Cliffordson, C. (2008). Discrepancies between school grades and test scores at individual and school level: effects of gender and family background. *Educational Research and Evaluation*, 14(2), 181–199.
- Klapp Lekholm, A. & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, 15(1), 1–23.
- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice.

- Educational Measurement: Issues and practice*.
Winter 2003, 34–43.
- McMillan, J. H., Myran, S. & Workman, D.
(2002). Elementary teachers' Classroom
Assessment and Grading Practices. *The Journal
of Educational Research*, 95(4), 203–213.
- Murphy, P. (2000). Equity, assessment and gen-
der. In J. Salisbury & S. Riddell (Eds.), *Gender,
Policy & Educational change* (134–152). London:
Routledge.
- Pilcher, J. K. (1994). The value-driven meaning of
grades. *Educational Assessment*, 2(1), 69–88.
- Skolverket (2005). National assessment and gra-
ding in the Swedish school system. Hämtad 4
augusti, 2008 från [http://www.skolverket.se/
publikationer?id=1524](http://www.skolverket.se/publikationer?id=1524).
- Skolverket. (2008). Ämnesprovet 2007 i grund-
skolans årskurs nio. Hämtad 4 augusti 2008
från [http://www.skolverket.se/
publikationer?id=1779](http://www.skolverket.se/publikationer?id=1779).
- Shepard, L. A. (1993). Evaluating test validity. In L.
Darling-Hammond (Ed.), *Review of research in
education* (405–450). Washington, DC:
American Educational Research Associations.
- Stiggins, R. J., Frisbie, D. A. & Griswold, P. A.
(1989). Inside high school grading practices:
Building a research agenda. *Educational
Measurement: Issues and Practice*, 8(2), 5–14.