

# Utvärdering som instrument för kvalitetsutveckling i nya ST

## Pilotprojekt med två vetenskapliga kurser vid Karolinska universitetssjukhuset

**KRISTINA SUNDBERG**, fil mag, utvärderare/kvalitetsutvecklare, ST-ledningen, Karolinska universitetssjukhuset; forskningsassistent, Medical case centre, institutionen för medicin, Karolinska institutet, Huddinge  
kristina.sundberg@ki.se

**JONAS NORDQUIST**, fil dr, medicinsk pedagog, ST-ledningen,

Karolinska universitetssjukhuset; föreståndare, Medical case centre, institutionen för medicin, Karolinska institutet, Huddinge

**LARS KIHLLSTRÖM**, överläkare, neurokirurgiska kliniken; ST-rektor, FoUU-ledningen, Karolinska universitetssjukhuset, Solna

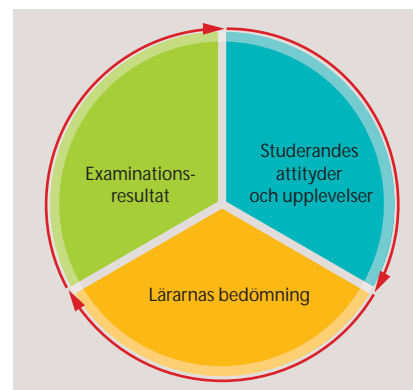
Utvärdering är potentiellt ett utvecklingsverktyg för utbildning. Ambitionen hos ST-ledningen vid Karolinska universitetssjukhuset är därför att använda en modell för utvärdering av utbildning som riktar sig mot utveckling i stället för att enbart fokusera på kursdeltagares reaktioner.

Under november 2009–januari 2010 nyintroducerade ST-ledningen (författare 2 och 3) i ett pilotprojekt två vetenskapliga kurser (kursgivare: Karolinska institutet) som en del av läkarnas nya specialiseringstjänstgöring. Argumentet för att introducera kurserna var att deras lärandemål låg väl i linje med delmål 19 (förmåga till ett medicinskt vetenskapligt syn- och förhållningssätt), som i enlighet med Socialstyrelsens riktlinjer sorterar under kompetensområdet medicinsk vetenskap och kvalitetsarbete [1]. Kurserna vände sig till de ST-läkare som inte tidigare uppnått delmålet som registrerade forskarstuderande. Inför introduktionen av de båda kurserna beslutade ST-ledningen att kurserna skulle utvärderas, vilket var av särskilt stor vikt då kurserna nyintroducerades för ST-läkargruppen. Ambitionen var att utvärderingen skulle bidra till att utveckla kurserna och tydliggöra huruvida kursdeltagarna uppnått aspekter av det lärande som eftersträvats. (Utvärderingen genomfördes av författare 1.)

### Utvärdering för kvalitetsutveckling

Utvärdering är ett underutnyttjat instrument för kvalitetsutveckling av medicinsk utbildning [2]. Trots att det i samband med utbildningsinsatser ofta påtalas att utvärdering sker, så är det oklart om så är fallet. Vad innebär utvärdering? På vilket sätt utvärderas utbildning generellt? Vilka modeller för utvärdering av utbildning är lämpliga att använda?

Definitionen av begreppet utvärdering som använts i samband med det aktuella utvärderingsprojektet är »att beskriva och värdera ett program« [3]. Utvärdering är dock inte likställt med forskning. På ett generellt plan granskar utvärderingen det specifika, medan (den kvantitativa) forskningstraditionen försöker generalisera. Utvärderaren fokuserar på kontexter, intressenter och beslutsprocesser, medan forskaren fokuserar på generaliserbara resultat och teoriskapande. Utifrån några aspekter kan dock utvärderingsområdet sägas ligga närmare den kvalitativa än den kvantitativa forskningstraditionen: exempelvis traditionell generaliserbarhet är inte



Figur 1. Modell för kursutvärdering.

primärt fokus inom vare sig utvärdering eller kvalitativ forskning, och samma metoder används även stundtals för datainsamling och analys [4, 5]. En tydlig skillnad mellan utvärdering och de båda ovan nämnda forsknings-traditionerna är att utvärdering sker inom en »uppdragsram«, medan forskning sker inom en

»teoriram«. Att utvärdering inte kan likställas med forskning innebär dock inte att utvärdering som företeelse saknar forskningstradition [4]. Utvärdering är ett väl beforskat område och kan med fördel användas i utbildningssammanhang för att fastslå hur väl en specifik utbildningsintervention fungerat.

### Kursvärdering inte tillräcklig

I utbildningssammanhang används ofta termen utvärdering, och det instrument som återkommande används är kursvärderingsenkäten. Vi hävdar dock att för att uppnå kriterierna i enlighet med vår definition av utvärdering är det inte tillräckligt att enbart granska kursdeltagares attityder till och upplevelser av en kurs. Att sammanställa resultat från kursvärderingsenkäter uppfyller utvärderingskriteriet »beskriva« men inte »värdera«. Utvärdering ska betraktas som ett pedagogiskt redskap som hjälper oss att beskriva och värdera kurser men som också stimulerar till kvalitetsutveckling. För att uppnå detta behöver resultaten från kursvärderingsenkäten kombineras med andra data, som examinationsresultat och lärnarnas bedömningar (Figur 1). På detta sätt kan en kursut-

### ■ sammanfattat

**ST-ledningen** vid Karolinska universitetssjukhuset initierade under höstterminen 2009 en utvärdering av två nyintroducerade vetenskapliga kurser i läkarnas nya specialiseringstjänstgöring. **För att en modell** för utvärdering ska kunna förflytta fokus från »kundnöjdhet« till kvalitetsutveckling av utbildning måste den omfatta utvärde-

rande moment som både beskriver och värderar. **Utvärdering** som ett instrument i utvecklings- och beslutsstypen bör utnyttjas i större omfattning inom medicinsk utbildning. Utvärdering ska ses som en del i en helhet, där kompetensnivån hos den utbildade individen ska utvärderas men också utbildningen per se.

värdering genomförs. Strävan är att inte enbart säkra kvaliteten utan även utveckla den [6]. Kombinationen av olika typer av data ger oss en mer mångfasetterad bild av utbildningsmomentet och större möjligheter till utvecklande insatser. Kursdeltagarnas upplevelse av utbildningens relevans är dock långt ifrån oviktig; upplevelsen av relevans är en av de mest centrala förutsättningarna för att lärande ska kunna utvecklas hos en vuxen individ [7].

## SOLO-taxonomin

Ett genomgående tema i utvärderingen har även varit att värdera i vilken utsträckning utbildningen har påverkat kursdeltagarnas lärandenivåer. Den kunskaptaxonomi (klassificeringsmodell för kunskapsnivåer) vi valt att använda för att belysa och förtydliga deltagarnas nivå av kunskaper är John Biggs SOLO-taxonomi. Utan att gå alltför djupt in i taxonomins konstruktion innebär de tre första stegen ett kvantitativt tillskott av nya kunskaper och informationsbitar samt förmåga att kombinera dem. Ju närmare en person rör sig mot nivå fem, den glidande skalans extrempunkt, desto högre kvalitativ (kognitiv) förståelse har han/hon av ett visst fenomen. Skalan är hierarkisk, vilket betyder att personer som befinner sig där de kan generalisera förståelsen till en ny domän har all den faktakunskap som de föregående fyra nivåerna förutsätter [8] (Figur 2).

## Kirkpatricks pyramid för utvärdering av utbildning

En av de etablerade modeller som i dag används för utvärdering av utbildning är Donald L Kirkpatricks pyramid. Enligt Kirkpatrick är syftet med utvärdering av utbildning att granska om utbildningen bidrar till organisationens mål och möjliggöra utveckling av utbildningen. Kirkpatrick menar att det inom utbildning finns fyra olika nivåer som kan utvärderas: reaktion, lärande, beteende och resultat (Figur 3).

På nivå 1 – reaktion – kan vi mäta deltagarnas reaktioner på utbildningen. Det är på denna nivå som resultat från traditionella kursvärderingsenkäter återfinns. I gränslandet mot nivå 2 – lärande – återfinns även resultat från kursutvärderingar. Lärande, dvs utbildningens inverkan på deltagarnas utveckling av kunskap, attityder och färdigheter, är en förutsättning för att vi på nivå 3 – beteende – ska kunna utvärdera om utbildningen bidragit till att förändra deltagarnas reella beteenden. Nivå 4 – resultat – är den mest komplicerade att utvärdera eftersom den till stora delar berör kostnadseffekter och måluppfyllelse på makronivå i verksamheten [9].

Utvärderingen av de två vetenskapliga kurserna genomfördes på Kirkpatricks nivå 1 (reaktion) och nivå 2 (lärande); nivå 1 för kvalitetssäkring och kvalitetsutveckling av kurserna och nivå 2 för att fastslå specifika önskade lärandenivåer i pilotprojektet – kvalitetssäkring.

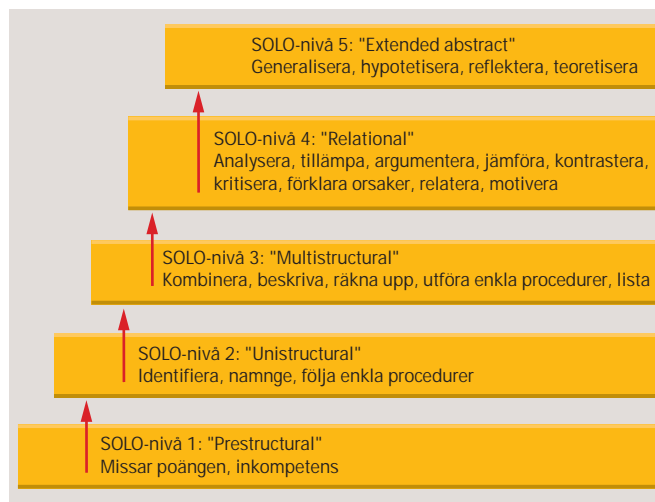
Syftet med artikeln är att beskriva hur användningen av en beskrivande och värderande kursutvärderingsmodell i ett pilotkursprojekt kan kompletteras med en utvärdering av lärandemål för att både kvalitetssäkra och kvalitetsutveckla kurser för ST-läkare.

## METOD

### Scientific writing and publishing

Utvärderingsunderlaget för kursen »scientific writing and publishing« på Kirkpatricks nivå 1 var kursutvärdering, öppen deltagande observation vid kursgenomförandet och en uppföljningsfråga till kursdeltagarna via mejl ett år efter avslutad kurs. Kursvärderingsenkäten fokuserade främst på områdena lärandemål, lärandeaktiviteter och examination.

Kursdeltagarna fick i enkäten ta ställning till en rad olika



Figur 2. SOLO-taxonomin enligt Biggs.



Figur 3. Kirkpatricks modell för utvärdering av utbildning.

påståenden och ange i vilken grad de instämde i dessa på en skala från 1 till 5. Den öppna deltagande observationen genomfördes av utvärderaren genom att observera lektions- och seminarietillfällen och genom att under kursens gång samtala med kursdeltagarna om deras upplevelser. Observationens syfte delgavs deltagarna i samband med kursstart, och löpande observationsanteckningar fördes. Dessa användes främst i utvärderingens analyskede och då som ett verktyg för att förtydliga uttalanden från kursdeltagare och kursledare som framkommit i andra delar av utvärderingsmodellen.

Utvärderingen på Kirkpatricks nivå 2 genomfördes med hjälp av ett individuellt skriftligt kunskapsprov för deltagarna (SOLO-nivå 2 och 3) före och efter genomförd kurs; samma prov användes vid båda tillfällena. Kunskapsprovet innehöll frågor om kursens innehåll på temat innehåll, form och process för vetenskapligt författande; alla prov rättades efter facit av samma person. Syftet med utvärderingen på Kirkpatricks nivå 2 var främst att, utöver utifrån examinationsresultaten, säkerställa om deltagarna uppnått de lärandenivåer som introduktionen av kursen avsett.

## Epidemiology I: introduction to epidemiology

Utvärderingsunderlaget för kursen »epidemiology I: introduction to epidemiology« på Kirkpatricks nivå 1 innefattade kursutvärdering, deltagande observation vid kursgenomförandet och en uppföljningsfråga till kursdeltagarna via mejl ett år efter avslutad kurs.

Kursvärderingsenkäten fokuserade främst på områdena lärandemål, lärandeaktiviteter och examination. Kursdeltagarna fick i enkäten ta ställning till en rad olika påståenden och ange i vilken grad de instämde i dessa på en skala från 1 till 5. Den öppna deltagande observationen genomfördes av ut-

värderaren genom att observera lektions- och seminarietillfällena samt samtala med kursdeltagarna om deras upplevelser under kursens gång. Observationens syfte delgavs deltagarna i samband med kursstart, och löpande observationsanteckningar fördes. Dessa användes främst i analyskedet av utvärderingen och då som ett verktyg för att förtydliga de uttalanden från kursdeltagare och kursledare som framkommit i andra delar av utvärderingsmodellen.

Utvärderingen på Kirkpatrick's nivå 2 genomfördes med hjälp av en gruppövning/formativt examinerande moment (concept mapping) som utfördes av kursdeltagarna vid kursens slut. Kursdeltagarna genomförde gruppövningen i två grupper om sex respektive sju personer. I övningen identifierade respektive grupp en rad olika epidemiologiska termer och grupperade dem sedan efter upplevda relationsmönster. Det fanns inget exakt facit för gruppövningens önskade slutresultat utan syftet var att gruppmedlemmarna under övningsprocessen skulle kunna identifiera, kombinera och relatera en rad olika epidemiologiska begrepp (SOLO-nivå 3 och 4).

En kontrollgrupp bestående av sju ST-läkare som inte gått kursen och som inte var registrerade som forskarstuderande eller hade doktorerat genomförde samma gruppövning i september 2010. Syftet med utvärderingen på denna nivå var främst att, utöver utifrån examinationsresultaten, säkerställa om deltagarna uppnått de lärandenivåer som introduktionen av kursen avsett.

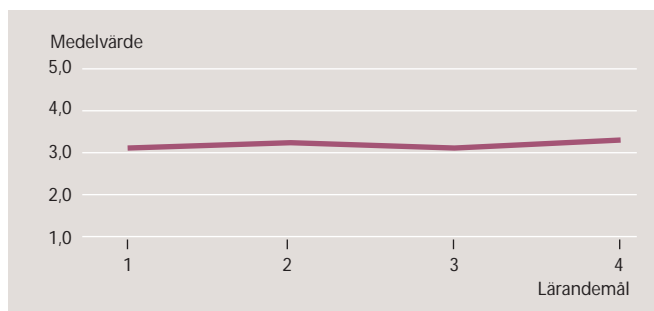
## RESULTAT

Kursen »scientific writing and publishing« är i sin ursprungliga form anpassad för målgruppen forskarstuderande vid Karolinska institutet och genomförs vid institutionen för folkhälsovetenskap. Smärre anpassningar till ST-läkargruppen gjordes därför före kursstart. Föreläsare med klinisk bakgrund tillkom, men en större förberedande uppgift och en uppgift i form av en inledning till en artikel utgick på grund av ST-läkarnas arbetsbelastning. Den deltagande observationen visade dock att kursdeltagarna trots dessa justeringar delvis upplevde en lägre nivå av relevans vad gäller lärandeaktiviteternas koppling till gruppens specifika bakgrund.

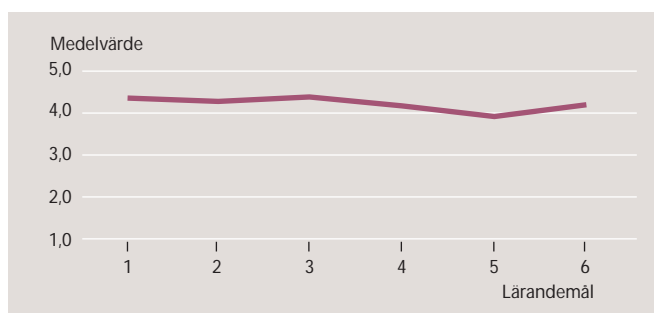
Svarsfrekvensen på kursvärderingsenkäten var 93 procent (n=15), och deltagarnas upplevelse var att de överlag uppnådde kursens lärandemål »till viss del« (svarsalternativ 3) (Figur 4). Kursdeltagarna ansåg att tydligare information om inlämningsuppgifter, tydligare kommunicerat syfte, handledning i samband med gruppuppgifter, förbättrad logistik kring kursen och en minskad total arbetsbörda under veckan skulle kunna ha bidragit till att man uppnått kursens lärandemål i större utsträckning.

Kursutvärderingen visade dock att samtliga kursdeltagare blev godkända på de examinerande momenten vid första examinationstillfället, och den kursansvariges bild av kursdeltagarnas uppnådda lärandenivå efter avslutad kurs var att den utvecklats under kursens gång samt att kursdeltagarna uppnått lärandemålen på ett tillfredsställande sätt. Den individuella uppföljningsfrågan ett år efter avslutad kurs besvarades av 9 av 15 deltagare. Av de svarande deltagarna ansåg 4 att de som ett resultat av kursen bla har kunnat tolka artiklar, strukturera manuskript och planera forskning på ett bättre sätt än tidigare. 5 ansåg att kursens format och/eller ST-utbildningens karaktär (huvudsakligen klinisk inriktning) bidragit till att de hittills inte kunnat dra nytta av kursens innehåll i någon större praktisk utsträckning.

I en jämförelse med resultaten från det individuella skriftliga kunskapsprovet, som genomfördes före kursstart och efter



Figur 4. Kursen »scientific writing and publishing«: medelvärde för nivå av uppfyllelse av lärandemål 1–4.



Figur 5. Kursen »epidemiology I: introduction to epidemiology«: medelvärde för nivå av uppfyllelse av lärandemål 1–6.

avslutad kurs, visade dock ett t-test att deltagarna efter genomförd kurs hade höjt sin kunskapsnivå inom området ( $P=0,005$ ).

Kursen »epidemiology I: introduction to epidemiology« är i sin ursprungliga form anpassad för forskarstuderande vid Karolinska institutet och genomförs vid Institutet för miljömedicin. Intentionen med den aktuella kursomgången var att genom tät och fortlöpande kommunikation mellan kursansvarig och kursdeltagare kunna se till att ST-läkargruppens behov och förkunskaper skulle få inflytande över kursens utformning. Den deltagande observationen visade att den avsedda kommunikationen fungerade väl och att kursdeltagarna överlag upplevde en hög nivå av relevans vad gäller lärandeaktiviteterna i kursen.

Svarsfrekvensen på kursvärderingsenkäten var 92 procent (n=12). Deltagarnas upplevelse var att de överlag uppnådde kursens lärandemål »i högre grad« (svarsalternativ 4) och/eller »helt och hållet« (svarsalternativ 5) samt att kursens lärandeaktiviteter och formativa examinerande gruppuppgifter hade bidragit till detta (Figur 5). Kursutvärderingen visade att samtliga kursdeltagare blev godkända på de examinerande momenten vid första examinationstillfället. Den kursansvarige var mycket nöjd med gruppens insats och ansåg att kursdeltagarna uppnått lärandemålen på ett utmärkt sätt. Ett år efter avslutad kurs besvarades den individuella skriftliga uppföljningsfrågan av 8 av 13 deltagare. Av dem som svarade var samtliga positiva till den upplevda nyttan av kursen och ansåg att de som ett resultat av kursen bla kunnat diskutera forskningsrön på kliniken, kritiskt granska vetenskapliga artiklar och arbeta med ST-projektet på ett bättre sätt än tidigare.

Gruppövningen »concept mapping« resulterade i att båda deltagargrupperna identifierade majoriteten av de 52 inkluderade begreppen och relaterade begrepp, samtidigt som alla

gruppledmedlemmar deltog aktivt i övningen. Kontrollgruppens genomförande av samma gruppövning visade att deras gemensamma kunskaps- och färdighetsnivå inte var lika hög som kursdeltagarnas; övningen uppnådde aldrig SOLO-nivå 4. Majoriteten av de epidemiologiska begreppen identifierades dock av kontrollgruppen, och kartläggningen av begreppen genomfördes på SOLO 3-nivå, men endast 3 av 7 gruppmedlemmar var aktiva i processen. Kursdeltagarna uppnådde således överlag en högre lärandenivå inom det aktuella området och hade en högre nivå av deltagaraktivitet än kontrollgruppen.

## DISKUSSION

Deltagarnas upplevelse av att de dels uppnått lärandemålen i något lägre utsträckning, dels haft mindre praktisk nytta av kursen »scientific writing and publishing« beror troligen mer på utformningen av kursen än på dess faktiska innehåll. Den deltagande observationen visade exempelvis att kursdeltagarna upplevde en lägre nivå av koppling mellan lärandeaktiviteterna och deras bakgrund som ST-läkare. Betydelsen av deltagarnas upplevelse av lärandeaktiviteter som relevanta och anpassade till deras specifika behov blir här central. Även det faktum att deltagarna upplevde en bristfällig kommunikation vad gäller syfte med lärandeaktiviteter samt ändringar i kursens schema är en trolig bidragande orsak.

Utvecklingsarbetet med formerna för främst »scientific writing and publishing« fortsätter, och inför genomförandet av kursen för ST-läkare i november 2010 kom fokus att läggas på tydligare information till deltagarna om syftet med olika moment i kursen och eliminering av logistiska problem. Att få den specifika gruppen deltagare, ST-läkare, att känna att kursen och dess lärandeaktiviteter känns relevanta för just dem blir av högsta prioritet.

Begränsningar att beakta i samband med utformningen av lärandeutvärdering på Kirkpatrick's nivå 2 är den risk för »testeffekt« som kan uppstå då ett kunskapstest genomförs vid två tillfällen nära inpå varandra. Testeffekten innebär att den täta upprepningen av provmomentet medför en förhöjd kunskapsnivå i sig, oavsett utformningen av det utbildningsmoment vars effekt man avsett att följa upp [10]. Testeffekten kan således ha spelat in vad gäller resultaten av det skriftliga eftertestet som genomfördes i samband med kursen »scientific writing and publishing«.

Den uppföljningsfråga som kursdeltagarna fick via mejl ett år efter avslutad kurs angående deras upplevelse av kursen och den nytta kursen givit i deras dagliga arbetsliv hade även med fördel kunnat ställas i samband med en fokusgruppsintervju för ett mer uttömmande svar. Då det av schematekniska skäl inte var möjligt att samla alla berörda ST-läkare vid ett och samma tillfälle gick frågan i stället ut via mejl.

Då de två kurser som utvärderats i detta projekt nyintroducerades på ST-läkarutbildningen vid Karolinska universitets-

sjukhuset och således hade karaktären av ett pilotprojekt, valde vi att komplettera vår kursutvärderingsmodell (Kirkpatrick's nivå 1 – se Figur 1) med en utvärdering av deltagarnas lärandenivåer (Kirkpatrick nivå 2). Kombinationen av dessa två nivåer gjorde att vi fick en mer mångfasetterad utvärdering. Genomförande av en utvärdering av lärande på Kirkpatrick's nivå 2 innebär dock flera metodologiska utmaningar (provkonstruktion, införskaffande av kontrollgrupper etc) och är vid användning av etablerade examinationsmetoder i en kurs inget som behöver genomföras efter varje kurstillfälle. I stället visar våra fleråriga erfarenheter av praktiskt bruk av kursutvärderingsmodellen (Figur 1) i professionsutbildning att den med engagemang från alla inblandade intressenter är ett tillräckligt instrument i sig för att skapa en levande kvalitetsutvecklingskultur.

## KONKLUSION

Utifrån två exempel på utvärderingar av vetenskapliga kurser inom läkarnas specialiseringstjänstgöring vid Karolinska universitetssjukhuset menar vi att för att utvärdering av utbildning ska kunna uppnå sin fulla potential som kvalitetsutvecklingsinstrument måste den modell för utvärdering som används innefatta både beskrivande och värderande moment.

En traditionell kursvärdering fokuserar enbart på kursdeltagarnas reaktioner på en kurs; en beskrivning av kursdeltagarnas upplevelser kan göras. För att kunna värdera denna beskrivning behöver dock en kursutvärdering genomföras; reaktionerna måste vägas samman med examinationsresultat samt lärarnas bedömningar. På detta sätt kan utvärderingsresultaten användas för kvalitetsutveckling; utveckling av kvalitet kan inte bygga på enbart reaktioner. Kursutvärderingen kan med fördel även kombineras med andra moment såsom observationer och uppföljningsfrågor. Är syftet med utvärderingen att dessutom få en uppfattning om huruvida kursdeltagarna som ett resultat av kursen höjt sina lärandenivåer inom det aktuella området måste utvärderingsmodellen innefatta utvärderingsmoment på Kirkpatrick's nivå 2.

Utvärdering är ett instrument som har varit centralt i beslutet att fortsätta att erbjuda de vetenskapliga kurserna »scientific writing and publishing« och »epidemiology I« för Karolinska universitetssjukhusets ST-läkare. Vid ST-ledningen anser vi att nivån av kunskap kring utvärdering av medicinsk utbildning bör höjas och att utvärdering som ett instrument i utvecklings- och beslutssyfte i utbildningssammanhang bör utnyttjas i större omfattning. Därvidlag ska utvärderingen ses som en del i en helhet, där kompetensnivån hos den utbildade individen ska utvärderas men också utbildningen per se.

■ *Potentiella bindningar eller jävsförhållanden: Inga uppgivna.*

**Kommentera** denna artikel på [Lakartidningen.se](http://Lakartidningen.se)

## REFERENSER

1. Socialstyrelsens föreskrifter och allmänna råd om läkarnas specialiseringstjänstgöring. SOSF 2008:17.
2. Tavakol M, Gruppen LD, Torabi S. Evaluation: Using evaluation research to improve medical education. *The Clinical Teacher*. 2010;7(3):192-6.
3. Jerkedahl Å. Utvärdering – steg för steg. Om projekt- och programbedömning. Stockholm: Norstedts Juridik AB; 2005.
4. Karlsson O. Utvärdering – mer än en metod. Tankar och synsätt i utvärderingsforskning. Stockholm: Kommentus förlag; 1999.
5. Kuper A, Lingard L, Levinson W. Critically appraising qualitative research. *BMJ*. 2008;337:687-9.
6. Nilsson KA. Kursutvärdering – för studenternas medverkan i kvalitetsarbetet. Lund: Lunds universitet; 1998. Rapport nr 98:202.
7. Knowles M, Holton EF III, Swanson R. *The adult learner*. London: Elsevier; 2005.
8. Biggs J, Tang C. *Teaching for quality learning at university*. Maidenhead: Open University Press; 2007.
9. Kirkpatrick DL, Kirkpatrick JD. *Evaluating training programs. The four levels*. San Francisco: Berrett-Koehler Publ Inc; 2006.
10. Karpicke JD, Roediger HL. The critical importance of retrieval for learning. *Science*. 2008;15(5865):966-8.