

Oppdrag for programsensor (oktober 2015)

Evaluering av rutiner, prosedyrer og former for vurdering som brukes i klinisk undervisning i masterprogram i odontologi og bachelorprogram i tannpleie ved UiB

- Hvilke rutiner praktiseres for vurdering i de ulike kliniske kurs?
- Hvilke krav må oppfylles for å oppnå «bestått» i kliniske kurs?
- Bidra til å utarbeide formuleringer som mer nøyaktig beskriver hva som må til for å oppnå bestått i klinisk kurs.
- Evaluere og foreslå forbedringer når det gjelder utforming og bruk av «vurderingsskjema» i klinisk undervisning.

Den 19:e januari, 2016 hölls en serie möten med lärare och studenter på IKO (möteschema bifogas). En Kommentarer och sammanfattning av dessa samtal ges i det nedanstående.

Jag har på institutets hemsida gått igenom samtliga de kliniska ämnenas målbeskrivningar där man kan läsa om förväntat læringsutbyte vad avser Kunskap, Färdighet och Generell kompetens. Målbeskrivningarna anger också i vilka former (Vurderingsformer) ovanstående examineras. Som synes i nedanstående tabell sker examinationen i kliniska ämnen uteslutande genom skriftlig examen:

Sammanställning av vurderingsformer i kliniska ämnen

Ämne	Vurderingsformer	Karakterskala	Emneevaluering
Pedodonti år 5	4 timers skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.
Pedodonti år 4	4 timers skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.
Kariologi 7:e semester	Skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.
Kariologi 10:e semester	Skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.
Periodonti 4:e studieår	4 timers skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.
Endodonti	4 timers skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.
Protetik 4:e studieår	4 timers skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.
Kjeveortopedi	4 timers skriftlig	A-F	Skriftlig evaluering via Mi

	eksamen		side.
Allmenodontologi	6 timers skriftlig eksamen	A-F	Skriftlig evaluering via Mi side.

Det framgår således ingenstans i målbeskrivningarna hur studentens praktiskt kliniska arbete examineras trots att detta kanske utgör uppemot 50 % av curriculum.

I mötet med lärare och studenter framkom ingen tung kritik mot existerande former för bedömning av studenters kliniska arbete. Studentrepresentanterna från Kull 5 t.ex.

"føler seg stort sett komfortable med vurderingsformen"

"ønsker mer individuell tilbakemelding – Ikke bare 'bra' eller 'ok'"

"studenten må kunne ta opp ting på en ordentlig måte"

"det må være lov til spørre. Noen lærere kritiserer i stedet for å svare på spørsmålene"

Studenterna på Kull 5 tycker det är en god idé att först utvärdera sin egen kliniska insats och att instruktören sedan gör sin värdering av studentens arbete.

De synpunkter som framfördes av studenterna i Bergen är på intet sätt unika för detta lärosäte. De kunde lika gärna ha framförts av studenter från någon annan nordisk skola. Uppfattningar och attityder är säkert inte heller desamma hos alla studenter på en kurs. Valda studentrepresentanter är i allmänhet elever som är väl förankrade och vällyckade i utbildningen. Kanske har studenter med sämre resultat varken tid, motivation eller intresse att företräda kursen. Dock är det särskilt för denna grupp av studenter viktigt att kriterier, regelverk, rutiner, rapportering etc. är tydligt kommuniserade, transparenta och tillämpade på ett rättssäkert sätt.

Man får en illustration till ovanstående när man läser vad Kull 4 säger i sina kommentarer:

"Skifte av lærere kort tid før vurderingssamtale."

"Får vurdering fra en lærer som de ikke kjenner"

"Instruktørene har ulike kompetanse, og mener forskjellig om samme arbeid. Spec protetik."

Samma studenter säger vidare:

"Krav til produksjon, problem med å få rette pasienter. Må bruke litt skjønn når det er liten pasient-tilgang"

"Tidlig hjelp til studenter som ligger bak"

”Mulighet for ekstra økter til de som ber om det. Hvis mulig.”

Det är svårare att beskriva krav och kriterier samt att examinera i kliniken jämfört med att mäta teoretisk kunskap. Kanske är det därför det i målbeskrivningarna inte nämns något om klinisk examination.

Det är synnerligen viktigt såväl studenter som kliniska handledare kan arbeta mot en känd och väl beskriven ”standard” – t.ex. en lista med kriterier. Vid institutionen för odontologi, Sahlgrenska akademien i Göteborg finns ett 1-årigt program - Kompletterande utbildning av Utländska Tandläkare (KUT). De sista 15 veckorna av programmet görs i form av VFU i Folktandvårdens kliniker och där samma krav på klinisk färdighet ställs som på sista årets studenter i grundutbildningen. För att underlätta för kliniska handledare att bedöma de utländska tandläkarnas klinik har respektive ämnesområde fått sammanställa en så kallad checklista. I **Bilaga 1** redovisas checklistan för cariologi. Man kan mycket väl ha liknande checklistor i grundutbildningen.

Det är lika viktigt för lärarkollektivet att ha uppställda kriterier efter vilka det kliniska arbetet skall bedrivas. Kanske kan därigenom problem med olika kompetens hos instruktörer som nämns i kritiken från Kull 4 i någon avhjälpas.

Det är också en poäng att studenten gör en egenbedömning av sitt arbete och kommenterar på vilket sätt insatsen ansluter eller inte ansluter till uppställda mål. I **Bilaga 2** redovisas hur egenbedömning tillämpas i den klinikförberedande undervisningen vid institutionen i Göteborg.

Studieresultat skall redovisas tydligt och skriftligt för studenterna. Detta viktigt för att uppnå transparens och en hög grad av rättssäkerhet. I synnerhet gäller detta när mindre goda resultat meddelas. Speciellt i kliniken. Besked om underkända insatser och insatser som kanske får till följd att ett kursavsnitt, eventuellt en hel termin, måste göras om måste meddelas skriftligt och i god tid.

I det nedanstående ges exempel på försök vid andra lärosäten att ge tydlighet åt det kliniska arbetet.

Liftupp är en App och plattform – ett kommersiellt system ursprungligen utvecklat och vid universitetet i Liverpool som på ett drygt år vunnit insteg på 10-talet lärosäten ffa i Storbritanien. Systemet är utformat för att ge stöd för studenter i klinisk undervisning på lärosäten inom tandvård, medicin eller hälsovård. Genom Liftupp får eleverna regelbunden återkoppling som registreras t.ex. via iPads och lagras i en molnbaserad databas. I **Bilaga 3** ges några exempel på hur data registreras. Studenter, handledare och kursledare kan fortlöpande få en bild hur det kliniska arbetet går. Liftupp underlättar bedömningar och skapar transparens, vilket har prisats särskilt av studenterna.

System inneholder en rad ulike verktøy, t.ex. en plattform for OSCE-examination der læreren kan konstruere og gjennomføre stationsbaserte prøv, får stöd for digital retting, återkoppling till studenterna, etc. **Bilagorna 4, 5, 6 och 7** redovisar några publikationer som anknyter till och belyser idéerna bakom Liftnapp.

Hans Sandberg vid Karolinska Institutet (KI) utvecklade för ett antal år sedan den så kallade "StuDentiGroup". Detta är ett digitalt system för att hålla studenterna (och lärarna) fortlöpande informerade om hur ffa det kliniska arbetet fortskrider. **Bilaga 8** redovisar hur man i den kliniska undervisningen på KI kan värdera olika aspekter av en seans när en student utför en fyllning. På liknande sätt har man i Tandhygienistutbildningen i Göteborg ställt upp kriterier för kliniska moment. I det här fallet finns allting på papper och inte som ett digitalt system (Bilaga 9).

Avslutande kommentar

Jag har utvärderat rutiner som används vid IKO för att värdera studieresultat i kliniska kurser. Jag har särskilt tittat på vilka krav som gäller för att uppnå betyget "Godkänd". Sedan några år tillbaka används ett så kallat vurderingsschema framför allt vid kliniska kurser. Vurderingsschemat – som används vid "utvecklingssamtal" med studenterna en gång per termin - undertecknas av både student och lärare och utgör därigenom ett formellt och godkänt protokoll över samtalets innehåll. Vurderingsschemat är ambitiöst och innehåller tämligen många rubriker som skall bedömas enligt nedanstående sammanställning:

Vurdering av ferdighetskurs og klinisk tjeneste - Skjema for samtale mellom student og gruppelærer

Skikkethet/holdninger	Ferdigheter	Kunnskaper
Kommunikasjon	Journalføring	Definisjoner og begreper
Samarbeidsevne	Behandlingsplanlegging	Faktakunnskaper
Pasientrelasjon/ -omsorg	Tidsplanlegging	Innsikt og forståelse av faget
Evne til å vurdere egne prestasjoner	Behandlingsgjennomføring	
Orden	Tekniske ferdigheter	
Hygiene	Selvstendighet	
Punktlighet		

Vid mina samtal med studenter i olika fas och lärare i olika kliniska kurser framkom ingen uttalad kritik mot studentsamtalen eller bruket av vurderingsschemat. Jag bad en av studenterna att skriva några rader om hur vederbörande upplever samtal och schema:

"Generelt i klinikken vurderes en fortløpende mht definerte krav, og holdninger, kunnskaper og ferdigheter vurderes. Basert på dette og egenrapportering (fra studenten) om fremdrift har en studentsamtaler der vurderingsskjema brukes. Vurderingsskjemaene går gjennom på studentsamtaler med hver enkel student ca 1 gang pr semester. Meg bekjent gjøres dette i all hovedsak i alle kliniske fag, og stort sett fra en starter preklinisk og klinisk undervisning. Det er et standardisert skjema som fylles ut av faglærere, og studentene kan kommentere og gi tilbakemeldinger som også noteres. Hvorvidt det ikke brukes gjelder prekliniske fag tror jeg - usikker på dette. Jeg kan i grunn bare svare for min egen del når det gjelder hvorvidt disse skjemaene og

vurderingene fungerer. Det er en ok gjennomgang av hvordan en ligger an som student. Jeg personlig syns det ofte fungerer som en sjekkliste, hvor en kontrollerer at en har "huket av" punktene og ofte står det ikke utfyllende annet en ok, bra eller lignende. Studentsamtalene er ofte hurtig gjennomført, og jeg kunne selv ønske mer konstruktive tilbakemeldinger enn det en får. Jeg vet at medstudenter ofte føler det er vanskelig å ta opp konkrete hendelser og episoder da en stort sett er alene med instruktør og evt sekretær på disse samtalene. Jeg har ikke tilgang på et slikt vurderingsskjema, det kan du evt spørre noen på fakultetet om, så ser du hvilke vurderingsområder som er. Håper dette var til hjelp, jeg har som sagt sagt basert svarene på egen oppfatning."

Utvecklingssamtalen en gång per termin och sättet som dessa genomförs är bra och borgar för en viss "rättssäkerhet". Det är emellertid väl så viktigt att bedömningar i det dagliga arbetet på kliniken sker transparent och i former som ffa studenten är trygg med. I mitt uppdrag ingick att titta på vilka krav som gäller för att uppnå betyget "Godkänd" på kliniska kurser. Jag har letat efter, kriterier, mallar eller instruktioner som lärare och studenter har att hålla sig till för uppnående av godkänt resultat i kliniken. I lärandemålen finns ganska väl beskrivet vad som gäller och hur man examinerar teoretisk kunskap. Men inte en stavelse om hur bedömningar sker i kliniken. Jag kan tänka mig att det finns en outtalad norm eller standard på kliniken och som existerat under många år. Men vad är det som säger en grupp lärare – som inte är kalibrerade – bedömer lika? Och hur skall studenterna som kanske konfronteras med motsägelsefulla besked "kryssa" mellan dessa?

I mitt arbete med denna rapport har jag mött en rad olika nya instrument och sätt att skapa bättre transparens vad gäller bedömningar i kliniken – både vid min egen institution och vid andra lärosäten. Jag har nämnt om detta tidigare i rapporten och jag tror att jag har stöd i modern pedagogik att bedömningar och betygssättning också av kliniska moment bör ske mot givna kriterier. Sådana skapar en norm för både lärare och studenter.

Göteborg den 20 februari 2017
Jan Olsson

Bilaga 1

Sammanställning över mål och kriterier för att uppnå godkänt resultat i ämnet cariologi under en 15 veckors VFU-kurs (TLKVFU) för utländska tandläkare. Kursen ingår i och är en avslutande del av ett 1-årig program – Kompletterande Urbildning för Utländska Tandläkare (KUT) – vid Sahlgrenska akademin, Göteborgs universitet. Under kursen arbetar studenten med patienter på en folktandvårdsklinik under handledning av en utsedd klinisk handledare.



GÖTEBORGS UNIVERSITET

KUT (2016/2017)

VFU - Cariologi

Mål

Efter avslutad VFU-verksamhet förväntas KUT-tandläkaren kunna:

Kunskap och förståelse

- redogöra för den teoretiska kunskapen i förhållande till det praktiska arbetet avseende kariologisk behandling
- redogöra för principer för kariologisk diagnostik och behandlingsplanering
- beskriva principerna för, samt konsekvenserna av, val av diagnostiska metoder, riskbedömningar, terapiplanering, genomförd behandling och prognosbedömning av utförd kariologisk behandling

Färdigheter och förmåga

- självständigt samla in, dokumentera och analysera information erhållen av patienter vid odontologiskt omhändertagande
- självständigt diagnostisera, riskbedöma, terapiplanera och behandla patienter med kariesskador
- visa fördjupad förmåga till lagarbete genom att planera, organisera och leda kariologiska behandlingar i samverkan med andra yrkesgrupper inom tandvården

Värderingsförmåga och förhållningssätt

- uppvisa ett etiskt och moraliskt gott förhållningssätt i tandvårdssituationen
- visa ett professionellt förhållningssätt gentemot patienter, deras anhöriga, kollegor, kliniska handledare och övrig personal
- diskutera kring och reflektera över egen insats vid genomfört arbete i relation till det aktuella behandlingsresultatet

Diagnostik

Undersökning av minst 3 patienter med kariologisk sjukdom

- I undersökningen skall ingå:
 - a) Anamnes
 - b) Cariogram
 - c) Adekvat röntgen
- Sjukdomsgraden kan variera men patienterna måste uppvisa någon form av kariesskador alternativt erosionsskador
- Målet är att säkerställa:
 - a) rätt undersökningsteknik (kliniskt och röntgenologiskt)
 - b) kunskap om kariologiska diagnoser på tand- och individnivå

Se specificerade kriterier.

Terapiplanering

Planering av omhändertagandet av ovanstående 3 patienter som har diagnostiserats med kariologisk sjukdom

- KUT-tandläkaren skall självständigt planera och organisera adekvat omhändertagande
- Terapin måste inte nödvändigtvis utföras av KUT-tandläkaren själv
- Remiss till specialist tandvård (om nödvändigt) accepteras också som adekvat omhändertagande
- KUT-tandläkaren skall visa förmåga till lagarbete i samverkan med andra yrkesgrupper (t ex tandsköterska, tandhygienist)

Bedömningskriterier för kliniska moment

Moment att bedöma	Kriterier för godkänd	Kriterier för underkänd
Självständig kariologisk undersökning	Förmåga att självständigt registrera och diagnostisera initiala och manifesta kariesskador kliniskt och röntgenologiskt. Att kunna ange tidigare utförd restaurativ vård och bedöma dess kvalitet. Registrera och diagnostisera erosionsskador som emalj- eller dentinskada. Även kunna differentialdiagnosticera andra hårdvävnadsdefekter från karies.	Klar över- eller underdiagnostik. Oförmåga att upptäcka och bedöma tidigare utförd restaurativ vård.
Bedöma registrerade skador och övriga defekter samt föreslå behandlingsåtgärder	Kunna diskutera och föreslå behandlingsåtgärder av patientens befintliga skador.	Klar över- och underbehandling av aktuella skador.
Riskbedömning	Självständigt kunna insamla, identifiera och analysera relevanta angrepps- och skyddsfaktorer för karies/erosioner samt värdera den sammanlagda risken.	Oförmåga att identifiera och analysera viktiga risk- och skyddsfaktorer
Kausal behandlingsstrategi	Utifrån framkomna uppgifter om patientens riskfaktorer föreslå orsaksinriktad sjukdomsbehandling i prioriteringsordning med bl.a kostinformation (innehåll och frekvens), fluortillskott (alternativa produkter utifrån patientens möjligheter), oral hygien (teknik, frekvens), salivsstimulerande medel (förslag på metoder och produkter).	Felaktigt föreslagna åtgärder i förhållande till patientens problematik eller felaktig prioriteringsordning.

Moment att bedöma	Kriterier för godkänd	Kriterier för underkänd
Information om kariesprevention	Att ge information om hur karies (erosionsskador) kan motverkas eller förhindras. Denna ska anpassas efter patientens kunskapsnivå, ålder och aktuella risk. Att kunna planera och dokumentera ovanstående.	Oförmåga att anpassa information till aktuell patient och/eller bristande dokumentation.
Kariesexkavering + kavitetutformning	Självständigt kunna exkavera till kariesfrihet. Kunna bedöma djupet av kariesangreppet samt risk för läsion. Kunna bedöma behovet av successiv excavering samt hur denna skall utföras. Optimal utformning av kavitet inklusive hänsyn taget till valt material	Ej kariesfrihet eller överexkavering. Pulpaläsion vid kariesfri yta alternativt pulpaläsion där successiv exkavering varit att föredra. Felaktig utformning av kaviteten vilket påverkar restorationens prognos.
Torrläggning/Matrisläggning	Kunna skapa torrt arbetsfält. Välja matris och kil för att erhålla optimal fyllning.	Ej torrt arbetsfält. Matrisval med risk för avsaknad av approximal kontakt eller cervikal anslutning
Färdig fyllning	Visa upp fyllning med godkänd kvalitet enligt USPHS kriterier: <ul style="list-style-type: none"> • God anatomisk form • Tillfredsställande estetisk färg • Approximala kontakter: Bra kontakt (punktkontakt, ej kontaktyta) till grann tanden med motstånd vid bedömning med 	Avsaknad av fossa och/eller avrundad randvulst. Estetiskt störande färg. Ytkontakt större än 1/3 av den bucko-linguala och ocklusala-cervikala ytan. Placering i fyllningsskarv eller palatinalt/lingualt.

	tandtråd	Avsaknad av kontakt
Kariologisk utvärdering	Efter insatta profylaktiska åtgärder kunna beskriva sina egna och patientens insatser vad som kvarstår.	Avsaknad uppföljande utvärdering.
Prognosbedömning	Kunna sammanfatta patientens nuvarande och framtida kariesrisk och ge motivering till bedömningen	Avsaknad av prognosbedömning och motivering.
Ställningstagande till revisionsintervall	Utifrån kariesrisk och prognosbedömning bestämma lämplig revisionsintervall och eventuella kontroller och ange motivering till detta	Ej individanpassat val av revisionsintervall

Att skicka in

Moment att bedöma	Kriterier för godkänd	Kriterier för underkänd
Patientfall	Se undersökning,diagnostik och terapiplanering av minst 3 patienter	
Material och hantering	<p>Utifrån genomförd kariologisk undersökning och föreslagna behandlingsåtgärder enligt ovan föreslå lämpliga materialval för de tänder som är aktuella för restaurativ behandling Materialvalen på tänderna kan skilja sig åt och skall motiveras utifrån patientens allmänna anamnes och status, kariologiska och övriga orala status. Vid val av komposit skall även adekvat bonding material anges. Hanteringen av de olika material som valts skall beskrivas och motiveras så att hanteringsens betydelse för restorationernas livslängd framgår så långt detta är möjligt. Hållfasthetsmässiga aspekter på materialvalet ska också diskuteras med hänsyn taget till typ av kavitet och lokalisation samt i förekommande fall, betydelsen av preparationens design i förhållande till moderna principer om tandsubstans bevarande preparation. Även de valda materialens biokompatibilitet bör diskuteras.</p>	<p>Uppenbara brister i kunskapen om moderna restaurativa material deras egenskaper, fördelar, nackdelar och hantering vilket påverkar resultatet på ett sådant sätt att patienten påverkas negativt och fyllningarnas livslängd kan anses väsentligt försämrat.</p> <p>Val av material har ej gjorts med hänsyn taget till det aktuella patientfallet/en.</p> <p>Motivering till val av material samt deras hantering saknas eller är uppenbart felaktigt. Över eller underterapi.</p>

Bilaga 2

Exempel på kriterier för egenvärderingen avseende fyllning och puts i klinikförberedande labkurs. Studenten gör först sin egen värdering av utfört moment. Läraren gör därefter sin bedömning.

EGENBEDÖMNING

För varje utförd preparation eller fyllning som avslutas skall en skriftlig egenbedömning efter särskilt formulär göras. När du anser att arbetet är färdigställt gör du en egenbedömning av preparationen eller fyllningen. Särskilda formulär för detta finns. Varje preparation eller fyllning skall värderas ur olika aspekter och dessa bedöms enligt skalan Utmärkt - Acceptabel - Ej acceptabel enligt uppställda kriterier.

Egenbedömningen visar du för din handledare och ni går gemensamt igenom denna. Din handledare gör sin bedömning och ni diskuterar eventuella åsiktsskillnader och justeringar. Målet är givetvis att ditt arbete ur alla aspekter skall vara utmärkt! Då den gemensamma bedömningen visar att preparationen eller fyllningen är tillfyllest signerar handledaren formuläret och du kan lämna detta moment.

Om din preparation eller fyllning på någon punkt får bedömningen Ej acceptabel får du arbeta vidare med preparationen. Om avvikelsen är liten och korrigering är möjlig får du utföra denna och därefter göra en ny egenbedömning. Om avvikelserna från kriterierna är sådana att korrigering inte är möjlig får du byta tanden och göra en ny preparation och egenbedömning. Samtliga egenbedömningar skall sparas i gemensam pärm hos handledarna.

EGENBEDÖMNING och KAMRATBEDÖMNING

De gruppärare som finns på den prekliniska utbildningen står gärna till tjänst att hjälpa er, **men det förutsätter att du själv gjort en analys/egenbedömning (dock ej skriftlig) samt även frågat en kamrat om en bedömning innan handledare tillkallas för bedömning av olika moment.** Enda gången kamratbedömning ej tillämpas är vid diagnostiskt prov liksom de praktiska proven.

Moment	Bra	Acceptabelt	Ej acceptabelt
Förslutning	Kaviteten fullständigt försluten. Ingen sonderbar spalt eller kant	Kaviteten fullständigt försluten. Ingen sonderbar spalt eller kant	Kavitetens ofullständigt försluten. Överskott och/eller underskott kan sonderas
Kontaktområden	Approximala fyllningar har god, punktformig och rätt placerad kontakt till grann tand. God ocklusionskontakt vid Klass I-fyllning	Fyllningar har god ocklusions och/eller grann tandskontakt. Kontaktens ytomfång $\leq 2 \text{ mm}^2$	Fyllningen saknar kontakt till grann tand, är för hög eller låg (Klass I). Kontakten ligger fel i approximalrummet eller omfattar större yta
Anatomisk form	God anatomisk form. Fyllningen följer tandens form mjukt med återgivande av kristor fossor och randvulster	Mindre avvikelser från ideal anatomisk form kan förekomma	Dålig anatomisk form med nivellering av detaljer som kristor och fossor. Approximala ytor vid Klass I och III eller buckala/linguala ytor"
Släthet	Fyllningsytan är slät och jämnt övergående i fyllningens olika anatomiska detaljer	Endast mindre ojämnheter i fyllningsytan	Fyllningsytan är ojämn med övergångar till fyllningens olika anatomiska detaljer. Tandytan har avverkats
Defekter	Inga sonderbara defekter förekommer	Mindre defekter (ej sonderbara med ficksond)	Tydligt sonderbara defekter förekommer
Puts	Fyllningsytan är jämn och utan repor	Mindre repighet och små underskott kan förekomma	Fyllningsytan är repig och ojämn. Sonderbara överskott och /eller tydliga underskott förekommer
Skador på grann tdr; se även info som finns på lab.	Ingen skada på grann tänder och slemhinna	Obetydlig skada på grann tänder och slemhinna	Tydliga skador förekommer på grann tänder och slemhinna

Fyllningen är tillfredsställande utförd då samtliga moment har nått graderingen "acceptabelt"



Examples of dental clinical feedback forms

The screenshot shows a mobile app interface for a dental clinic. A modal window titled 'Procedure Information' is open, displaying a form for 'Direct Restorations'. The form includes fields for 'Quad' (1-8), 'Tooth' (1-8), 'Difficulty' (Basic/Complex), 'Procedure' (Approximal Posterior Restoration), and 'Material' (Amalgam). Below these fields are two scrollable lists: '-- Select Procedure --' and '-- Select Material --'. The background shows a sidebar with a search bar and a list of students, with 'Student 470' selected.

Fig 1: sample form for determining quad/tooth/procedure/material to give context of the procedure

The screenshot shows the same mobile app interface, but now the 'Direct Restorations' form is displayed as a feedback form. It includes a summary section with the following information: Quad: 3, Tooth: 6, Procedure: Approximal Posterior Restoration, Difficulty: Basic, Material: Amalgam. Below this, there are several rows of feedback questions, each with a rating scale (1-6) and an information icon. The questions are: 'Appropriate management of the pulp', 'Appropriate material selection and handling', 'Appropriate tooth preparation for material', 'Appropriate restoration of tooth contour and anatomy', 'Management of occlusion', 'Iatrogenic damage inc. damage to soft tissues', and 'Procedural Knowledge'. At the bottom, there is an 'Assist only' section with a toggle switch and a speech bubble icon.

Fig 2: sample feedback form for restorative clinical skills

App Store

Mrs Staff 126

Search students

Restorative Clinic

5

X

Absent

Present

Selected

Student 470

15:32

Forms

Patient #64646473

Restorative Clinic Form

Add Alert

Feedback

Direct Restorations

Quad: 3

Add Item

Hide

Edit

Clear

Tooth: 6

Procedure: Approximal Posterior Restoration

Difficulty: Basic

Material: Amalgam

Appropriate management of the pulp

i

X

1

2

3

4

5

6

Appropriate material selection and handling

i

X

1

2

3

4

5

6

Appropriate tooth preparation for material

i

X

1

2

3

4

5

6

Appropriate restoration of tooth contour and anatomy

i

X

1

2

3

4

5

6

Management of occlusion

i

X

1

2

3

4

5

6

Iatrogenic damage inc. damage to soft tissues

i

X

Present

Absent

Procedural Knowledge

i

X

1

2

3

4

5

6

Assist only

i

Add Student

Forms (1/1)

Fig 3: sample feedback form for restorative clinical skills

Commentary

Calling for a re-evaluation of the data required to credibly demonstrate a dental student is safe and ready to practice

Introduction

In the UK, the Francis Report (1) has driven key changes within health care and has focused the need to:

‘Make all those who provide care for patients – properly accountable for what they do... to ensure that the public is protected from those not fit to provide such a service.’ (1)

Irrespective of nationality, this statement underscores the importance of undergraduate education, and its associated assessments, because the best and most cost-effective way to protect the public is to ensure that only the right individuals go on the professional register in the first place. For providers of undergraduate education, this distils down into the problem of how to ensure and demonstrate that our graduates are competent to practice.

In undergraduate dental education, common approaches for demonstration of competency are grounded in the traditions of novice to expert learning (2). In this arena, becoming an ‘expert’ requires ten or more years of experience (3). Consequently, the traditional method for determining competency is the measurement of experience through counting the number and the quality of procedures completed (4). This approach has likely become widely accepted because it appears to have face validity; it is simple to do; progression decisions can easily be defended; it has endured the test of time; and it fulfils a crucial criterion for assessment, namely it is acceptable to stakeholders (5). Data to support the latter statement can be found through reference to the latest round of inspection reports by the UK General Dental Council (GDC) where a focus on, and a drive to increase, the numbers of individual procedures performed by undergraduate learners is still very evident (6).

However, is this traditional approach still the best possible way of measuring competency considering the aforementioned changes in expectation over accountability, combined with advances in our understanding of pedagogy, and available technology?

This paper aims to initiate debate over what should constitute best practice in the assessment of competence. From the evidence-base available we suggest that to truly establish competency sophisticated approaches for data collection, integration, and interpretation are likely to be

needed to meet the demands and expectations of the 21st century. This is because the modern healthcare setting requires its professionals to be responsive and adapt to the ever-changing needs of patients (7). We suggest that in this setting, the important evidence underpinning competency is the longitudinal demonstration of the learner’s ability to independently and simultaneously manage all aspects of the activity being assessed for each patient, over a range of contexts, rather than simply measuring the amount of a specific activity or isolated facets of competency such as communication or professionalism. Furthermore, we will contend that decisions over progression will need to be made on a learner-specific basis through the professional judgement, and consensus of a multidisciplinary expert panel following the objective analysis of large and fully integrated data sets.

What is professional competence?

Professional competence has been defined as:

‘The habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being served.’ (8)

Therefore, the assessment of professional competence is complicated because it requires the daily integration of all data to demonstrate the stability and appropriateness of multiple skills and behaviours over time (habitual). This complex situation is often managed by assimilating the various dimensions of professional competence into a series of outcomes, which are then further organised into a series of domains such as clinical, communication, professionalism and management and leadership. This is the situation that exists in the UK and is described by the GDC in the document ‘Preparing for Practice’ (9).

The ability to convincingly establish competence in each domain is a fundamental requirement for defensible decisions over student progress or graduate registration. Therefore, establishing a suitable approach for the measurement of competence warrants careful consideration.

In 1998, David Chambers published a landmark paper entitled *Competency-based dental education in context* (2). In this publication, he explored the available data in the spectrum of novice-expert learning. He noted that five distinct developmental stages were recognised, novice, beginner, competent, proficient and expert, and that becoming an expert requires ten or more years of experience (3). He

concluded that *five years is only enough time to make a good start* (2) and that graduates would, at best, only be at the level of competent. In this paradigm, 'competent' refers to a developmental stage that is:

'Marked by independence, supported by basic internalized standards and an acceptable repertoire of skills and knowledge.' (2)

Moving forward, the components that describe the developmental stage of 'competent', such as independence, internalised standards, appropriate repertoire of skills and knowledge and understanding, would seem suitable to inform a guiding framework for exploring the measurement of competency at the domain level.

What are the limitations of current approaches to measuring competence?

In dentistry, competency has traditionally been measured through establishing the levels of activity (i.e. numbers of procedures) (4). However, data from multiple studies in medicine support the conclusion that experience does not necessarily predict competency (10–12), and in some cases may be associated with a reduction in competency (11). Moreover, in dentistry, at least one small study suggests that there is no significant difference in outcome between beginners and experts when the focus is on the end product:

'The traditional evaluation criteria in dental education (numbers of tasks completed or their quality defined in objective terms) are probably insufficient to reliably distinguish the level of learning of emerging professionals.' (4)

Whilst there can be no doubt that increased levels of activity broaden experience, it will become evident that it is the type of activity that is important and not the amount of activity that improves competence. The difference in focus between activity type and amount is likely to be decisive because an emphasis on the quantity can lead to learners concentrating on completing tasks, ultimately seeing patients as commodities that are only useful whilst their care contributes to the required skills tally. This is likely to have a deleterious educational impact (13). This is because it is a situation that can only detract from students actively pursuing patient-centred holistic treatment and gaining the required integrated learning approaches.

To facilitate a more robust measurement of competency, we would suggest a move away from the idea of progression through the developmental stages being solely driven by experience, towards the driver being *enhancing performance through deliberate practice* (14). We recognise that many schools are well aware of the limitations of a purely quantitative approach and have already made adjustments to their assessment strategy.

What is the relationship of competence to performance?

The concept of novice to expert learning, within a construct of performance, has been investigated and eloquently described by Anders Ericsson as:

'Nobody becomes an outstanding professional without experience, but extensive experience does not invariably lead people to become experts'... Although everyone in a given domain tends to improve with experience initially, some develop faster than others and continue to improve during the ensuing years. These individuals are eventually recognised as experts and masters. In contrast, most professionals reach a stable, average level of performance within a relatively short time frame and maintain this mediocre status for the rest of their careers.' (14)

Ericsson (14, 15) proposed a model to explain how *professionals reach a stable performance asymptote within a limited time period, whereas the expert performers are able to keep improving their performance for years and decades.*

When a learner is first introduced to a new activity, their primary goal is to reach a level in that domain which is deemed to be acceptable. At this early stage, the learner needs to concentrate hard to avoid mistakes. With more appropriate and focused practice directed by feedback, combined with domain-specific experience, the performance becomes smoother and requires less concentration until eventually it becomes automated. At the stage of automation, the individual loses conscious awareness and therefore is no longer able to make *specific intentional adjustments* without additional external observation and feedback (14). This concept of automation is decisive because once a professional has reached an acceptable skill level, data suggest that more experience does not lead to improved performance (15). Experts, on the other hand, continually and deliberately seek the continued training situation designed to place the desired goal beyond their current level of achievement.

The goal of undergraduate education may not be to create experts but it can certainly utilise the concept of deliberate practice in the development of learners. However, this requires the creation of a powerful learning environment (16) in which a number of key components will need to be brought together:

- Learners are systematically challenged through increasing task difficulty to prevent 'automation'.
- Teachers continually monitor learner performance.
- The provision of multisource feedback from both staff and patients, which is appropriately detailed and timely (17) to enable reflection and subsequent performance modification through *deliberate (focused) practice* (14, 15).

- Continuous opportunities and encouragement for the learner to undertake deliberate practice.

Success in such an environment is predicated around the ability to appropriately measure performance. Without this ability, both the meaningful monitoring of performance and the capability of providing the required levels of feedback become impossible.

How can performance be measured?

In assessment, established wisdom is that any measurement must be in relation to some form of transparently applied criteria or standard. Data from a recent study suggest that measurement scales that are constructively aligned (18) to the level of expertise of the assessor and the developing independence of the learner reduce the levels of disagreement between assessors and thus improve confidence in the assessment outcome (19). Furthermore, the measurement of performance through developing independence is also entirely consistent with the aforementioned required components of competence.

We suggest that a numerical scale anchored to descriptors over the degree of *independence* and *quality* of the learner's performance represents a justifiable approach for measurement that also drives the appropriate educational impact. In addition, with the right longitudinal and triangulated approach, it is a method that could not only be used to inform both the quality and consistency of domain-specific skills, but, by direct inference, also be used to inform both the quality and consistency of the internalised standards being applied by the learner, as well as measuring the learner response to external feedback through analysis of the degree of subsequent change in performance.

The need to assess daily practice, whilst at the same time capturing performance in multiple contexts, implicates the use of Workplace Based Assessment tools (WBAs), either in current or modified forms, linked to the aforementioned numerical scale. WBAs have been shown to have good predictive reliability (20, 21). However, data suggest that great care has to be used in the way they are operationalised and used to make decisions (19, 22). Some of the big challenges are as follows: (i) in the real world, patients, tasks and situations are subject to huge variability, (ii) WBAs are traditionally carried out on a limited number of occasions and designed for a specific task, which gives the learner a task rather than holistic focus and (iii) WBAs are subject to decisions from staff that will be influenced by the context in which they are made and the individuals who are making them. Amongst other things, these issues have resulted in the realisation that *from a psychometric perspective, very large numbers of assessors and cases are required to discriminate reproducibly amongst trainees* (22); and the need for a change in both terminology and focus when considering the

qualitative data from WBAs. It is also of note that it has been suggested that the terms 'credibility' (cf. internal validity) and 'dependability' (cf. reliability) (23) better describe the aims for the trustworthiness of type of data collected through this approach.

Irrespective of the terminology used, for WBAs to be employed successfully in the determination of competence, there would seem to be an implicit requirement for the collection, integration and active interpretation, of large continuous and longitudinal data sets. This is because without them, it would not be possible to establish the pattern of different performances across many different contexts within or across the domain(s) of interest.

How often should performance be measured and in what contexts?

Having established a principle for performance measurement, it is necessary to consider the available evidence to inform how often and where such measurements should take place. In other words, what is the acceptable repertoire of skills, what is an appropriate breadth of patients/procedures and what is a sufficient number of occasions to develop the skills?

Medical education is dominated by constructivist views of learning that consider learning as an 'entity' where the context within which the learning occurs may affect its quality, but has little impact on the 'learning' itself (24). A direct consequence of this conventional view is that competence is regarded as a trait, which once achieved is stable irrespective of context. This implies that for any individual skill, the degree of competence can be established, and once acquired is directly transferable to any situation that arises requiring that skill.

However, data strongly suggest that competence is highly context specific (12, 25). Furthermore, modern health care requires its professionals to be responsive to the needs of patients. The ability of an individual to respond to required change has been described as 'capability', which is defined as the *extent to which an individual can adapt to change, generate new knowledge, and continue to improve their performance* (7). Taken together, this means that *we can no longer see competence as 'a state to be achieved'... Competence is not just about acquisition of knowledge and skills, but about the ability to create new knowledge in response to changing work processes* (24). This paradigm shift means that *modern healthcare systems demand that we assess our learners ability to adapt and to flexibly apply and develop knowledge* (24).

Placing these arguments into the arena of dentistry, we would suggest that the true assessment of competence requires a demonstration of the learner's consistency in their ability to simultaneously integrate and appropriately apply and adapt all the relevant domain-specific skills at the

required level of independence, across a range of contexts. Thus, the breadth of experience becomes at least as important as amount. For instance, in Restorative Dentistry, the breadth of context for the placement of direct restorations might include tooth surface, tooth location, material used, difficulty of task (e.g. access, extent of caries and medical history), patient demography (age, gender, ethnicity, disability, anxiety level, etc.) and environment (clinical discipline, in-reach, outreach, etc.). With this approach, the amount of data required to demonstrate competency will be large, variable and learner specific. This is because each learner will see different patients, each contributing a specific set of contexts. Furthermore, each learner will develop at their own rate and will likely have different deliberate practice and feedback needs to stabilise their longitudinal performance.

Knowledge and understanding

The final facet of domain-associated competency to consider is knowledge and understanding. The principles for the objective measurement of knowledge (5, 26, 27), and the appropriate formats within which to do it, are well established and there is no need to elaborate upon them further here.

However, just as with any other component of competence, it cannot be assumed that once the student assimilates knowledge that they will have sufficient understanding to apply it when they come across any relevant situation. This is highlighted by data from a study evaluating the influences of teaching on learning that quote the response of a medical student support this view:

'I found it very difficult to actually study something like 'head injury' without relating it to my own personal knowledge of the clinical situation. . . . I think it is ludicrous to teach something like 'head injury'. . . without having that clinical basis, because (then) you (remember) what you are learning as a series of disconnected facts. . . (just) a very efficient way to pass an exam.' (28)

Long-standing data from work exploring child development provide insight into the problem, as it suggests that the ability to apply knowledge is also highly contextual and requires experience to allow the knowledge to be consolidated and organised in an appropriate way (29–31). This need for knowledge transformation through experience has been well established in medical education (32, 33) and has greatly influenced those studying the development of diagnostic expertise (34), especially in the area of clinical reasoning skills (25, 35).

Overall, data support the hypothesis that the acquisition of knowledge should be developed and concurrently monitored alongside the relevant clinical exposure in real time across contexts of skills application.

The triangulation and aggregation problem

The arguments presented suggest that the data needed for a true demonstration of professional competence are large and complex, with an implicit need for a coherent approach to aggregation and triangulation. Multiple assessment types designed and considered in isolation may lack the required sophistication, a situation that would be true irrespective of how well the various pieces of data were blueprinted together, or how valid and reliable (26) each of the individual assessments were considered to be. To illustrate, a situation that will be familiar to dental academics is one where a student causes concern to experienced clinical faculty. However, the student is pleasant, has managed to undertake the requisite amount of experience and has passed the available WBA's, OSCE's and knowledge examinations. There is probably good cause for the staff concern, but the student's progression is assured because the available data, although spanning domains, are considered in self-contained 'assessment packets', that is they are barely passing in several areas but the outcome is nevertheless a pass. It should be considered that someone in this situation is probably not competent overall, but the available data and the way it is integrated lack sufficient sophistication and resolution to reflect the legitimate concerns of the experienced teachers. We would contend that in the situation of a dental programme, a true measure of competency cannot be established from isolated assessments even when they are triangulated together, be they OSCEs, WBAs or written tests, especially where the focus of aggregation is the assessment instrument rather than the domain or skill. We propose that an enhanced measurement process that ensures the right outcomes for learners, patients and stakeholders is required. It is a process underpinned by the full integration and triangulation of data from all domains and contexts combined with an understanding of the performance within them. Crucially, within this paradigm, data from simulation, objective assessment and patients should be viewed as different contexts, which through appropriate assessment design strategies involving a coherent approach can be integrated to demonstrate competence. Clearly, it will be necessary to identify where triangulation is appropriate, and work in postgraduate medicine developing 'Entrustable Professional Activities' where data are required to be integrated and triangulated from many competencies spanning multiple domains to holistically demonstrate a real-world skill, may be a good model (36). If clinical academics were able to evaluate each and every clinical episode in terms of a cross-domain data set, where any outcome falling below the required level of independence highlighted an insufficiency, then this would allow them to reflect the overall ability of the learner to holistically manage the patient on that occasion. The data

derived from individual episodes of patient treatment would be integrated on a longitudinal basis and interpreted to determine patterns of consistency, which when further triangulated across contexts and with other assessment data would give a closer reflection of the true competency of the learner. Further benefits of such an approach is that it could (i) with the right management of WBA data allow for identification and moderation of staff who were not giving the learners appropriate feedback, in essence 'failing to fail', a known issue in dentistry (37) and (ii) serve to enhance the utility of assessment because its purpose is not just to identify who passes or fails, but rather to make the assessment process part of everyday learning and reflection (38).

Crucially, the aggregation, triangulation and interpretation of this personalised and complex data derived from a variety of contexts will not be straightforward or, lend itself to a purely quantitative approach. This will necessitate a move from individual disciplines behaving as independent entities when making progress decisions towards an integrated approach where a multidisciplinary panel functions as an *interpretive community* (24) to establish the ability of an individual to practise dentistry because:

"Truth" is a matter of consensus among assessors who have to arrive at judgments on performance that are as informed and sophisticated as can be at a particular point in time' (24).

Recommendations for data to inform a decision over competency

Overall, the available data strongly suggest that the demonstration of competence requires a coherent approach to the longitudinal aggregation and triangulation of data. Based on our analysis of the available evidence, the following five broad principles are suggested to inform credible, dependable and trustworthy decisions over learner progression:

- Consistency, demonstrated through the longitudinal measurement of performance is a key parameter to establish competence. Measurement of performance should be grounded in the developing independence of the learner.
- Both the *breadth* (i.e. the different contexts) and the consistency (number of *occasions* at the appropriate level) of performance are key drivers in developing and demonstrating competence. A number of parameters in relation to each assessed performance should be recorded as an indicator of context and enable the triangulation of data between and across contexts.
- In a dental programme, single assessments are not the best way of establishing or developing student competence, as these do not provide a sufficient breadth of contexts, an appropriate educational impact or longitudinal insight. Sophisticated methods of assessment data collection, integration and triangulation both within and across domains are required.
- It is essential that knowledge be linked to real-world patient encounters in multiple contexts, as well as from appropriately aligned theoretical and simulated situations that require the learner to process information and make relevant clinical decisions in a highly aligned and contextual manner.
- Progress decisions are best reached through the judgement of a multidisciplinary interpretive community informed by comprehensive data and a sophisticated approach to interpretation as discussed.

References

- 1 Staffordshire NM Foundation Trust Public Inquiry. Report of the Mid Staffordshire NHS foundation trust public inquiry executive summary. The Stationery Office, 2013.
- 2 Chambers DW. Competency-based dental education in context. *Eur J Dent Educ* 1998; 2: 8–13; Blackwell Publishing Ltd.
- 3 Chi M, Glaser R, Farr MJ. The nature of expertise. Farr MJ, ed. Hillsdale, NJ: Erlbaum, 1998.
- 4 Chambers DW, Geissberger M. Toward a competency analysis of operative dentistry technique skills. *J Dent Educ* 1997; 61: 795–803.
- 5 Norcini JJ, Shea JA. The credibility and comparability of standards. *Appl Measur Educ* 1997; 10: 39–59.
- 6 General Dental Council. Inspection reports [Internet]. [cited 2015 Nov 6]. Available from: <https://www.gdc-uk.org/Dentalprofessionals/Education/Pages/Dentist-qualifications.aspx>.
- 7 Fraser SW, Greenhalgh T. Coping with complexity: educating for capability. *BMJ* 2001; 323: 799–803.
- 8 Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002; 287: 226–235.
- 9 General Dental Council: Preparing for Practice. London: GDC, 2011.
- 10 Dawes R. House of cards: psychology and psychotherapy built on myth. New York, NY: Free Press, 1996.
- 11 Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. *Ann Intern Med* 2005; 142: 260–273.
- 12 Butterworth JS, Reppert EH. Auscultatory acumen in the general medical population. *JAMA* 1960; 174: 32–34.
- 13 Van Der Vleuten CP. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract* 1996; 1: 41–67.
- 14 Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004; 79 (Suppl. 10): S70–S81.
- 15 Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. *Med Educ* 2007; 41: 1124–1130.
- 16 Vermetten YJ, Vermunt JD, Lodewijks HG. Powerful learning environments? How university students differ in their response to instructional measures. *Learn Instr* 2002; 12: 263–284.
- 17 Nicol DJ, Dick DM. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Stud High Educ* 2006; 31: 199–218.
- 18 Biggs J. Enhancing teaching through constructive alignment. *High Educ* 1996; 32: 347–364.

- 19 Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales. *Med Educ* 2011; 45: 560–569.
- 20 Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): a preliminary investigation. *Ann Intern Med* 1995; 123: 795–799.
- 21 Prescott L, Hurst Y, Rennie JS. Comprehensive validation of competencies for dental vocational training and general professional training. *Eur J Dent Educ* 2003; 7: 154–159.
- 22 Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. *Med Educ* 2012; 46: 28–37.
- 23 Driessen E, Van Der Vleuten C, Schuwirth L, van Tartwijk J, Vermunt J. The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study. *Med Educ* 2005; 39: 214–220.
- 24 Govaerts M, van der Vleuten CPM. Validity in work-based assessment: expanding our horizons. *Med Educ* 2013; 47: 1164–1174.
- 25 Feltovich PJ, Barrows HS. Issues of generality in medical problem solving. In: *Tutorials in problem-based learning*. Assen: Van Gorcum, 1984: 128–142.
- 26 Downing SM. Validity: on the meaningful interpretation of assessment data. *Med Educ* 2003; 37: 830–837.
- 27 Norcini JJ. Setting standards on educational tests. *Med Educ* 2003; 37: 464–469.
- 28 Entwistle A, Entwistle N. Experiences of understanding in revising for degree examinations. *Learn Instr* 1992; 2: 1–22.
- 29 Vosniadou S, Brewer WF. Theories of knowledge restructuring in development. *Rev Educ Res* 1987; 57: 51–67.
- 30 Piaget J. *The child's conception of physical causality*. New York, NY: Harcourt, Brace, 1930.
- 31 Piaget J. *The child's conception of the world*. Totowa, NJ: Littlefield, Adams, 1929.
- 32 Lesgold AM, Feltovich PJ, Glaser R, Wang Y. *The acquisition of perceptual diagnostic skill in radiology* (Tech. Rep. No. PDS-1). Pittsburgh, PA: University of Pittsburgh, Learning Research & Development Center, 1981.
- 33 Boshuizen HP, Schmidt HG, Custers E. Knowledge development and restructuring in the domain of medicine: the role of theory and practice. *Learn Instr* 1995; 5: 269–289.
- 34 Boshuizen HP. Does practice make perfect? In: Boshuizen HP, Bromme R, Gruber H, eds. *Professional learning: gaps and transitions on the way from novice to expert*. Dordrecht: Springer, 2004: 73–95.
- 35 Eberhard J, Klomp HJ, Föge M, Hedderich J, Schmidt HG. The intermediate effect and the diagnostic accuracy in clinical case recall of students and experts in dental medicine. *Eur J Dent Educ* 2009; 13: 128–134.
- 36 ten Cate O. Nuts and bolts of entrustable professional activities. *J Grad Med Educ* 2013; 5: 157–158.
- 37 Bush HM, Schreiber RS, Oliver SJ. Failing to fail: clinicians' experience of assessing underperforming dental students. *Eur J Dent Educ* 2013; 17: 198–207.
- 38 Black P, William D. *Assessment and classroom learning*. Assess Educ 1998; 5: 7.

L. J. Dawson¹, B. G. Mason¹, V. Bissell² and C. Youngson¹

¹University of Liverpool School of Dentistry, Liverpool, UK,

²University of Glasgow School of Dentistry, Glasgow, UK
email: ldawson@liv.ac.uk

Validity in work-based assessment: expanding our horizons

Marjan Govaerts & Cees PM van der Vleuten

CONTEXT Although work-based assessments (WBA) may come closest to assessing habitual performance, their use for summative purposes is not undisputed. Most criticism of WBA stems from approaches to validity consistent with the quantitative psychometric framework. However, there is increasing research evidence that indicates that the assumptions underlying the predictive, deterministic framework of psychometrics may no longer hold. In this discussion paper we argue that meaningfulness and appropriateness of current validity evidence can be called into question and that we need alternative strategies to assessment and validity inquiry that build on current theories of learning and performance in complex and dynamic workplace settings.

METHODS Drawing from research in various professional fields we outline key issues within the mechanisms of learning, competence and performance in the context of complex social environments and illustrate their relevance to WBA. In reviewing recent socio-cultural learning theory and research on performance and performance interpretations in work settings,

we demonstrate that learning, competence (as inferred from performance) as well as performance interpretations are to be seen as inherently contextualised, and can only be understood '*in situ*'. Assessment in the context of work settings may, therefore, be more usefully viewed as a socially situated interpretive act.

DISCUSSION We propose constructivist–interpretivist approaches towards WBA in order to capture and understand contextualised learning and performance in work settings. Theoretical assumptions underlying interpretivist assessment approaches call for a validity theory that provides the theoretical framework and conceptual tools to guide the validation process in the qualitative assessment inquiry. Basic principles of rigour specific to qualitative research have been established, and they can and should be used to determine validity in interpretivist assessment approaches. If used properly, these strategies generate trustworthy evidence that is needed to develop the validity argument in WBA, allowing for in-depth and meaningful information about professional competence.

Medical Education 2013; **47**: 1164–1174
doi:10.1111/medu.12289

Discuss ideas arising from the article at
[www.mededuc.com 'discuss'](http://www.mededuc.com/discuss)



Educational Development and Research, Maastricht University,
Maastricht, the Netherlands

Correspondence: Marjan Govaerts, Educational Development and Research, Maastricht University, PO Box 616, Maastricht 6200 MD, the Netherlands. Tel: 00 31 433 885 746;
E-mail: marjan.govaerts@maastrichtuniversity.nl

INTRODUCTION

Work-based assessment (WBA) is potentially the best way of assessing professional competence, i.e. the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, judgement, emotions, values and reflection in day-to-day practice.¹ Work-based assessments include assessment tools such as mini-clinical evaluation exercise, direct observation of practical skill, professionalism mini-evaluation exercise, multi-source feedback as well as in-training evaluation reports that typically require clinical assessors to convert trainee performance into a numerical score, according to predefined rules and criteria, to obtain accurate and easily communicable descriptions of a trainee's ability. However, although WBA may come closest to assessing habitual performance, research findings raise serious concerns about utility of WBA for summative assessment purposes. First, assessment tasks in the real world are unpredictable and inherently unstandardised and they will not be equivalent over different administrations. From a psychometric perspective, this poses serious threats to reliability and validity of assessment. Second, as professional judgement is inherent in WBA, serious concerns are raised about the subjectivity of assessments. Raters are generally considered to be major sources of measurement error.^{2,3} Performance ratings are considered to be unacceptably biased, suffering from halo and leniency effects, and intra- and inter-rater reliability of performance ratings are often found to be substandard.⁴⁻⁶ Weaknesses in the quality of measurement on top of problems in the implementation of WBA instruments have even resulted in widespread cynicism about WBA in the profession.⁷

As is apparent from a focus on quantifiable measures of assessment quality, most criticisms of WBA stem from approaches to validity and validation consistent with the quantitative framework of psychometrics. In essence, validity refers to the degree to which the proposed interpretations and the uses of assessment outcomes (e.g. performance ratings or test scores) in terms of decisions and actions are adequate and appropriate, as justified by evidence or theoretical rationales.^{8,9} Validation can then be defined as 'developing a scientifically sound validity argument to support the intended interpretation of test scores and their relevance to the proposed use'¹⁰ through accumulation and integration of different kinds of evidence from different sources. Or, as stated by Koch and DeLuca¹¹: '...validation should be a generative process that promotes continuous

inquiry into assessment practice'. What is rarely addressed explicitly, though, is that our approaches to WBA – reflected in the way we design and evaluate assessment practices – are inextricably linked to our implicit theories of learning, performance and competence. In this article, it is our intent to illustrate that an exclusive focus on traditional psychometric approaches to validity and validation in WBA may no longer be appropriate by their disregard for key issues with respect to competence development, performance and assessment in complex and dynamic workplace settings.

Within the predictive, deterministic framework of psychometrics, assessment typically aims for generalisable explanations or predictions.^{9,12} Central to the psychometric discourse in current assessment are its almost exclusive focus on the inference of a true score representing true performance; its pursuit of a specified level of consistency that is assumed to be conditional on technically sound measurement (reliability) and the assumption of error (noise that needs to be eliminated) when repeated measurements fail to yield consistent results. The almost exclusive use of psychometric tools in validation of WBA, that is the way we develop the validity argument in WBA, reflects theoretical assumptions underlying our interpretations and uses of assessment outcomes that conceptualise assessment as a scientific measurement of abstract, latent and stable dispositions within individuals. In current approaches to WBA and validation of WBA, three assumptions in particular seem to stand out:

- 1 Learning (professional development) is a deterministic, linear process that can be identified and specified in advance; task performance and learning (as represented by assessment scores) are typically abstracted and interpreted independent of context;
- 2 Competence, as inferred from performance, is a fixed, permanent and decontextualised attribute, i.e. an inherent trait or ability of health care workers (or trainees), and
- 3 Performance can be 'objectified' and assessors, if they were only capable to do so, would be able to rate and observe some true level of performance.

There is, however, increasing and compelling research evidence that challenges the assumptions underlying our approaches to WBA. For instance, findings from research in industrial and organisational psychology show that job performance lacks temporal stability, especially in highly complex

jobs.^{13,14} True intra-individual variation in job performance may result from changes in the individual (e.g. due to motivation, fatigue, changing levels of competence) as well as changes in the job environment.¹⁴ Similarly, research findings in medical education indicate that context (i.e. task environment or work environment) critically influences behaviours in practising doctors. Durning and colleagues,¹⁵ for instance, reported that contextual factors affected clinical reasoning performance by experts (board certified internists) in ways that were very specific to the situation and were influenced by participants in the encounter (patient and doctor), their goals and the setting. So, although some aspects of job performance can be expected to be relatively stable over time (cognitive ability, perhaps), variability in performance ratings in WBA may very well reflect true performance variability within individuals. Similarly, increasing evidence from industrial and organisational psychology, as well as medical education, supports contentions that rater effects in WBA do not represent (mere) rater biases, but rather represent alternative and complementary valid perspectives on trainee performance,¹⁶ challenging our interpretations of between-rater differences in WBA.

Recent research findings and growing understanding of learning in complex social environments therefore suggest that meaningfulness and appropriateness of current validity evidence in WBA can be called into question, and common validity theory, which is framed in psychometrics, may no longer hold: we may be operating on faulty assumptions. In the following, we will discuss changing conceptions of learning and performance in work-based settings and will present research findings to substantiate the need for expanded conceptions of validation and validity theory. Drawing from research in various professional fields we will discuss the assumptions underlying the psychometric approaches to WBA and will propose alternative strategies to assessment and validity inquiry that are embedded in qualitative research paradigms and built on current theories of workplace learning and contextual performance.

WBA AND PREDICTABILITY OF LEARNING

In medical education, perspectives originating from behaviourist, cognitivist and constructivist learning theories have long dominated developments in instruction and assessment. These learning theories have in common that they focus on individual

learners, that they stress cognitive aspects of performance (i.e. thinking and reflection) and that learning is treated as a 'thing' or product located in the mind of the learner. Although these theories acknowledge that context influences quality of learning processes and thus how well learning occurs, their view is that the nature of what is learned or is to be learned, is relatively independent of context.¹⁷ They generally treat workplace learning as a linear process, akin to formal learning, through which a learner develops from incompetent to competent, largely neglecting the role of social, cultural and organisational factors in shaping learning and performance development. During the past decades, however, more robust theories of workplace learning have emerged, expanding the limiting assumptions underlying the theories described above.

Especially the group of socio-cultural theories of workplace learning seem to offer more powerful frameworks for understanding learning in workplace settings (See Glossary Table for definition of terms used). Socio-cultural learning theories claim that learning and learning outcomes emerge through active participation in activities of a community and interaction with the complex and dynamic systems of the work environment.¹⁸ Socio-cultural learning theories therefore consider learning and expertise development to be inextricably linked to features of the context in which the learning occurs; learning processes as well as learning outcomes change as contexts change.^{17,19} What, how and why trainees learn is shaped by unique experiences and the meaning or consequences that trainees and co-participants (e.g. supervisors, assessors, co-workers and patients in a clinical context) attach to these experiences.⁹ Socio-cultural learning theories, with their focus on knowledge produced by social interaction, are particularly useful for thinking about learning in clinical training and health care settings. In these settings, learning is produced by a trainee's engagement in non-standardised and unpredictable tasks of authentic health care practices and the ongoing social interaction around authentic tasks, shaped by (unique) physical, social and organisational contexts.²⁰ Learning in clinical work settings then inevitably becomes a dynamic, non-linear and non-deterministic process. The increasing complexity of health care as well as its ever-changing context furthermore demand that we move beyond predictability of individual learning and competence towards conceptualisations of competence as a collective, situated and dynamically produced through interaction and learning in functional clinical groups.²⁰

Not only is team-based care rapidly becoming the norm in our health care systems (requiring a shift in focus from individual competence to team competence), the complex and dynamic nature of health care systems also implies that we can no longer see competence as 'a state to be achieved'. Rather, nowadays, notions of work-based learning and competence should include the ability to continuously adapt to change. Competence it is not just about acquisition of knowledge and skills, but about the ability to create new knowledge in response to changing work processes.²¹ From this perspective, learning involves learning things 'that aren't there yet', through exchange and interactions in social networks and collaborative processes in communities of practice that adapt to continuously evolving circumstances.^{22,23} Complex and dynamic interactive processes between the learners and their environment then 'mutually reconstruct both the learner and the environment'. Learning is 'expansive'²² and can be conceptualised as 'an increasing (collective) capacity for acting in flexible, constructive and innovative ways appropriate to the challenges of ever changing circumstances'.¹⁷ Learning for future practice thus implies that learning is an ongoing process without a clear endpoint; learning is never complete. This is directly opposed to traditional approaches in medical education where learning focuses on planned, formal events with well-defined and stable learning outcomes.²⁴ Very recent theories of workplace learning therefore explicitly question whether predictable and decidable systems of workplace learning can be designed and implemented. These theories, some of which build on complexity theory, emphasise the view that learning is an ongoing *creative* process, emergent from its context in unpredictable and unanticipated ways.¹⁷

Although social learning theory is increasingly being used in medical education,¹⁹ much of current theorising still seeks to understand and explain workplace learning so that conditions that uniformly support and enhance quality learning can be identified and implemented. In fact, a lot of current efforts to improve work-based learning and assessment seem to aim for the design of clinical training that steers trainees' learning in predictable ways, through development of the 'right' theories of professional development, better analyses of task environments and the technology to model them,¹² as well as specifying standards for competent performance that have to be achieved at predefined stages in the learning process (e.g. milestones project).²⁵ In other words: if it would only be possible to predict what, when and how people learn, it would

also be possible to design assessments using predetermined correct responses or models of performance.¹² Such (law-like) predictability is necessary to make models of assessment, learning and performance compatible with the psychometric framework. However, conceptualisations of learning as inherently situated, collaborative, transformational and expansive (i.e. focusing upon knowledge production rather than reproduction) challenge assumptions of predictability and uniformity in what is learned and what is to be learned. Assessment that focuses on predefined and specified learning outcomes then necessarily becomes an oversimplification of an arbitrary stage in the process of professional development.²⁶

WBA AND COMPETENCE AS A FIXED ATTRIBUTE

Although context specificity or performance variability from one case or task to the next is a well-known phenomenon in medical education,²⁷ current approaches to assessment and its validation build on assumptions that there must be some level of true performance that can be 'measured': variability of an individual's performance over time or across tasks and work settings is typically viewed as measurement error. Competence is conceptualised as a stable trait, to be inferred from performance sampling within the professional domain, and expertise, once developed and established is considered to be portable and transferable from one context to another. In fact, most licensure and certification procedures seem to build on exactly this assumption.

There is an increasing body of research that challenges these conceptualisations of competence and professional performance. Within-person variation in performance is substantial and can be as large as between-person differences.^{28–30} Obviously, performance of learners changes during training, as they learn and develop through participation in professional practice. Indeed, the focus of current WBA is ongoing evaluation and provision of feedback to improve performance and expertise development.³¹ It would seem self-evident that conceptions of performance stability no longer hold within a context that intentionally aims for performance changes. We also readily accept that learners and professionals are not always performing at their best, and that performance varies from day to day or even within the same day. Especially in highly complex jobs, performance lacks temporal stability.^{13,14} Reasons may be motivational (e.g. changes in performance goals

and effort due to conflicting tasks), physiological (e.g. fatigue) or any other unstable cause affecting individual performance, such as mood or emotional experiences.³²

More importantly, however, there is an increasing body of research indicating that the dynamic nature of performance in work settings is caused by environmental factors, i.e. opportunities and constraints in the work setting, even in experts and talented performers. Research findings in industrial and organisational psychology and human resource management suggest that talented performance is not directly portable from one company to another, thereby challenging one of the foundational assumptions underlying human resource practices in organisations, namely that talent can be bought. In general, research findings indicate that performance is contextual and that 'talent won't transfer unless it maps to the challenges of the new environment'.³³ For instance, 'star' investment analysts on Wall Street showed significant short- and long-term performance decline after moving to another firm and the drop in performance persisted for up to 5 years.³⁴ Research findings suggested that specific features of the new role and work setting influenced the drop in performance. The contextual and situated nature of job performance was affirmed by findings that stars who moved with a group of colleagues performed better than those who moved solo. A study on the portability of leadership also showed that highly talented chief executive officers who were recruited by other firms did not always deliver; whether skills and experience proved valuable in the new job depended on specific characteristics of their new work environment.³³ Similarly, research on intra-individual performance variation in football players showed that a significant portion of variance could be explained by constraining actions of others, including teammates. Moreover, susceptibility to environmental constraints varied across players and job complexity, suggesting that performance is determined by the interaction between person, task and environment.³⁰ These findings are consistent with the notion of performance and competence being the product of cultural and social circumstances and of ongoing interaction with individuals and groups (teams) in a specific work setting.

Recent research in medical education equally challenges naïve assumptions about performance stability and generic transferability of knowledge and skilful practice. In their study on family practitioners' performance, Wenghofer and colleagues,³⁵ for

instance, found that the doctor's work setting as well as systemic (community-related) factors significantly impacted performance, with varying effects across different performance dimensions. The study furthermore showed that, although doctor factors significantly influenced performance, they were not nearly as important as previously assumed. The critical influence of context on doctor behaviour was also illustrated in a study by Ginsburg and colleagues,³⁶ who reported that practising internists' approaches to professional dilemmas were malleable and dependent on individual patient characteristics, the doctor's affective response and relationship with the patient, the nature of the diagnosis as well as the doctor's relationships with co-workers in the health care system. They concluded that a doctor's performance was subject to 'multiple interdependent, idiosyncratic forces unique to each situation'.

Despite powerful research evidence, however, the notion that performance genuinely fluctuates over (short) periods of time and cannot be defined independently of its context has not really affected assessment researchers yet. If we want to capture the complex and multifaceted construct of professional competence we need to focus on aspects that go beyond the technical and context-free aspects of performance. On the contrary, unique and continually changing work contexts in modern health care systems demand that we assess our learners' and doctors' ability to adapt and to flexibly apply and develop knowledge and skills in the face of evolving circumstances. In line with this approach, performance variability resulting from interaction with contextual factors should not be dismissed as 'measurement error', but considered as potentially valuable and meaningful information in the appreciation of an individual's professional competence.³⁷

WBA AND OBJECTIFICATION OF PERFORMANCE

From a socio-cultural perspective, performance is socially constructed and determined by each person's perception of and interaction with situational characteristics of the task at hand. When this framework is applied to the assessment of performance in work settings, a picture emerges of performance that can never be 'objective', but is always conceptualised and constructed according to the perspectives and values of an individual assessor, influenced by his or her unique experiences and the social structures in the assessment task and its context.³⁸

In fact, research findings in industrial and organisational psychology indicate that assessors' judgements of performance in work settings can only be understood *in situ*: assessor behaviours are framed within the context in which assessment takes place. In WBA, assessors are engaged in complex and unpredictable tasks, more often than not in a context of time pressures and conflicting as well as ill-defined goals.^{39,40} Assessors' behaviours and assessment outcomes are furthermore influenced by a broad range of other factors in the work context, such as interpersonal relationships (with the learner as well as with co-workers), political, emotional and cultural factors.^{41,42} Central to constructivist, socio-cultural approaches to assessment is the view that assessors can no longer be seen as passive measurement instruments, but as active information processors who interpret and construct their own personal reality of the assessment context. Or, as stated by Delandshere and Petrosky⁴³: 'Judges' values, experiences, and interests are what makes them capable of interpreting complex performances, but it will never be possible to eliminate those attributes that make them different, even with extensive training and "calibration".' This implies that there can be honest disagreement within and across communities of practice: a specific supervisor-assessor's conception of appropriate performance in, for instance, a patient encounter may be different from that of co-workers, the trainee or the patient. Differences in an assessor's interpretation and scoring of performance-related behaviours may then be viewed as 'distinct views of a common individual's job performance that may be equally valid'⁴⁴ or 'meaningful differences in..... behavior across sources, especially when each source rates... behavior in different situations'.¹⁶

Recent research in medical education^{45,46} confirms findings from industrial and organisational psychology. A study by Govaerts *et al.*⁴⁶ for instance, explored the use of performance theories by experienced and trained assessor-supervisors in general practice. Findings showed that, when observing and evaluating trainee performance, assessors interactively used general as well as task-specific performance theory and person schemas to arrive at judgements and decisions about performance effectiveness. Between-assessor differences in the performance dimensions used in the assessment of performance were substantial, though, reflecting assessor idiosyncrasy in the interpretation of task performance as a result of differing personal experiences, beliefs and professional values. These findings provide support for socio-cultural approaches

to WBA, in which assessors are to be seen as 'social perceivers' who construct and reconstruct their own performance theories and conceptualisations of competence through training, socialisation and task experience. Consequently, assessors in work settings are inherently idiosyncratic, and multiple assessors will have multiple constructed realities. Assessment that is framed in socio-cultural, constructivist theories thus challenges the assumption, underlying psychometric assessment theory, of the existence of a single true score.

IMPLICATIONS FOR WBA AND VALIDATION

What emerges from learning theories as described above and research evidence about performance and performance interpretations being inherently contextualised is the need to reconsider assumptions underlying common WBA practices.

On the basis of the research and insights presented in this paper, we want to argue that assessment in work settings is a socially situated interpretive act, which is inherently value laden. It reflects the experiences, the meanings, intentions and interpretations of individuals involved in the assessment process ('the interpretive community').⁴⁷ Conceptions of learning and performance based in socio-cultural theory call for assessment that does not just focus on learning outcomes, but also (and perhaps even more so) on the processes underlying learning, performance and performance interpretations in dynamic, complex workplace settings. This implies that the purpose of assessment is not to 'objectively' and 'accurately' quantify learning or learning outcomes, but to understand what, how and why trainees and doctors are learning. This entails understanding and explicating context, i.e. the relationship between learners, the learning environment and the larger social systems within which learning is occurring.⁹ Assessment questions need to address learners' experiences, the activities that they are engaged in as well as the social, cultural and ethical issues that shape learning, learning outcomes and performance interpretations.¹² Assessment questions, in other words, need to be grounded in inquiry traditions that offer rich, situated accounts of contextualised learning, performance and assessor judgements in order to capture, understand and evaluate multiple, diverse instances and interpretations of learning and performance in complex social systems. Inquiry systems that are situated within qualitative research paradigms (e.g. constructivist-interpretive) seem to be well suited for this task.

During the past decades, 'interpretivist approaches' to assessment have been proposed, in line with social-constructivist and socio-cultural theories of learning and performance.^{9,11,12,48,49} A central feature of these approaches is that performance assessments are seen as social constructions or interpretations, rather than absolute, objective truths⁴⁹; there is no single 'true' score or 'objective' rating of performance. Rather, 'truth' is a matter of consensus among assessors who have to arrive at judgements on performance that are as informed and sophisticated as can be at a particular point in time. Various methodological approaches in interpretivist assessment have been described. Kuper *et al.*⁵⁰ for instance, suggested an ethnographic approach and use of interviews and focus groups to capture a broad range of interpersonal behaviours in specific contexts and to generate rich, meaningful assessments of doctor competence. In the setting of teacher education, case study approaches have been adopted to develop an assessment scheme for the purpose of teacher certification.⁴³ Although each approach has its own origin and nuances, key characteristics of interpretivist assessment approaches could be summarised as follows^{43,48,49,51}:

- 1 In WBA assessment, tasks are not interchangeable, but make unique contributions to learning and assessment. As assessments in work settings are 'socially constructed' between assessors and the person who is being assessed, learners typically prepare a paper or portfolio documenting their learning and assessment activities to capture situated assessment processes. Assessment asks learners to describe the contexts in which they work (and learn), to document their learning experiences, learning goals and learning plans as well as assessment activities (work sampling, for instance) and performance evaluations. Knowing how a learner perceives the demands of any particular assessment task is considered critical information in performance interpretations. Therefore, the learner's point of view is typically incorporated in the assessment process, as are intermittent feedback cycles with critical analyses and reflection on learning and task performance;
- 2 Assessments rely on narratives rather than numerical scores: assessments seek to purposefully generate elaborate, written evaluative statements about performance by expert judges – those who are most knowledgeable about the context in which assessment occurs, intentionally capturing and accounting for context-specific aspects of performance. As scores have

little intrinsic meaning, assessment instruments challenge assessors to provide narrative comments that are useful in guiding the learner's competence development as well as meaningful in decision making about competence achievement;

- 3 All stakeholders in the assessment process are thus continuously challenged and required to document their performance interpretations as well as to articulate underlying values and assumptions;
- 4 Written performance evaluations are collected across a broad range of tasks, contexts and assessors, in order to gain in-depth understanding of a person's performance repertoire and capability to adapt to various task requirements, and
- 5 Inferences about professional competence are based on critical review of all available performance evidence, through open deliberative and critical dialogue among stakeholders in the assessment process. An interpretive approach does not imply that interpretations are bound to single assessment occasions or to single performance documentations. Meaningful interpretations can, and should be, constructed across assessment occasions and performance evaluations. Data from multiple sources are to be triangulated, reviewed and discussed to identify patterns of performance across tasks and contexts as well as any outlying aspects of performance. Interpretations are repeatedly tested against all available evidence, until a coherent interpretation or an integrative judgement on an overall level of performance can be accounted for^{43,48}. If necessary, decisions involve inquiry strategies for additional information gathering about specific aspects of performance. This does not mean that 'anything goes'; essentially, final decision making requires professional judgements that should be corroborated, motivated and substantiated in such a way that the judgement is defensible and credible. To guide the performance evaluation, interpretive categories or dimensions can be developed through collective discussion of values and standards. The critical review of the evidence, the questioning of the different interpretations and assumptions as well as the documentation of the decision-making process are all essential and contribute to the validity and fairness of the final decision. Part of the strength of interpretive approaches to assessment is its traceability, through documentation of rich, meaningful information and

articulation of values and standards. External evaluators may then assume an auditing role to ensure that the process is equitable, reflects professional standards and is sufficiently rigorous to protect the public from incompetent professionals. In this respect, interpretive assessment may be more trustworthy than assessments relying on a set of scores that mask assessors' thinking.⁵¹

These views on assessment are fundamentally different from prevailing psychometric-based, reductionist (positivist-oriented) approaches to assessment. What both the psychometric-based and constructivist-interpretivist assessment approaches have in common, though, is that inferences about professional competence need to be credible and defensible, based on trustworthy evidence. Within both frameworks, assessment validation comprises the 'development of a series of inferences and assumptions leading from the observed performances to conclusions and decisions...' and 'evaluation of the plausibility of these inferences and assumptions using appropriate evidence'.⁵² Clearly, traditional notions of reliability and validity related to quantitative evaluation of assessment practices have limited usefulness in the evaluation of situated performance interpretations. The theoretical assumptions underlying interpretivist assessment approaches, as described above, call for validity theory that provides the theoretical framework and the conceptual tools to guide the validation process in qualitative assessment inquiry. Although we acknowledge that there is considerable debate about the value and legitimacy of alternative sets of criteria and standards to assess qualitative inquiry, basic principles of rigour specific to qualitative inquiry have been put forward over the past decades, and we argue that they can and should be used to determine 'validity' (i.e., trustworthiness, credibility and defensibility) of the qualitative inquiry in interpretivist assessment approaches. Criteria and standards that can be used to judge the adequacy of constructivist-interpretivist assessment have been suggested by Lincoln and Guba^{53,54} in their classical work on evaluation. They suggest the use of criteria such as trustworthiness (i.e. credibility, transferability, dependability and confirmability) and authenticity (i.e. fairness, openness, negotiation and shared understanding) to evaluate assessment quality. They furthermore propose the use of various techniques or methodological strategies to bring rigour to the qualitative inquiry. These techniques include: prolonged engagement in the assessment process; peer debriefing; analysis of disconfirming evidence (i.e. actively seeking counterexamples that

challenge emerging interpretations), member checks and progressive subjectivity (to achieve credibility) as well as thick, rich description (to achieve transferability) and the audit trail, external audit and documentation of the assessment decision processes (to achieve dependability and confirmability). Some strategies need to be addressed in the assessment design stage, whereas others are applied during data collection and interpretation or after interpretation of performance data (similar to the application of techniques and strategies to ensure validity in standardised assessments).⁵⁵ Examples of these approaches to assessment validation have been described in typically context-bound assessments of portfolios.^{49,56–58} If used properly, methodological approaches as described above generate trustworthy evidence that is needed to develop the validity argument in interpretivist assessment approaches. In conclusion, similar to the positivist approach to validation, interpretivist assessment has the intent to construct generalising interpretations about a learner and his performance. However, the strategies to arrive at these interpretations and to provide evidence on the strength of these generalisations rest on different approaches.

CONCLUDING REMARKS

Based on contemporary learning theories and research evidence illuminating the context specificity of performance and performance interpretations, we argue that we need to expand our approaches to assessment inquiry in work settings and validity theory underlying validation processes.

We do not want to claim that contextualised perspectives on assessment can only be covered by the constructivist-interpretivist assessment framework. Alternative frameworks, such as Brunswik's Probabilistic Functionalism and Lens Model, also describe ecological perspectives on judgement and decision making.⁵⁹ Our argument, however, is that when building on specific frameworks in (evaluation of) assessments, one has to be very clear about assumptions underlying its use. On the basis of socio-cultural learning theories we propose approaches towards WBA that are grounded in qualitative (constructivist-interpretivist) research paradigms, to generate in-depth understanding of and meaningful information about critical aspects of professional competence. Rich, narrative evaluations of performance as well as articulation of underlying performance theories and values not only enhance the formative function of the assessment system to

maximise learning,⁵⁸ but are indispensable for trustworthy decision making in summative assessments. Our constructivist-interpretivist approach to WBA seems to cater to the growing awareness in the literature that an exclusive focus on the psychometric discourse may no longer be helpful in facing assessment challenges in modern health care practices and education.^{60,61}

We do not want to pretend that approaches as described in this paper provide solutions to all problems in WBA. Nor do we want to build an argument against the use of quantitative performance data in assessment of professional competence. Numerical ratings as well as standardised assessments are valuable elements in programmatic approaches to competence assessment.⁶² Rather, we should aim for careful balancing of quantitative and qualitative approaches in our assessment programmes, justifying our choices on the basis of assessment purposes as well as conceptualisations of learning and performance/competence.

Implications of interpretivist approaches to WBA include a shift from numbers to words in performance assessment as well as assessors who are willing and able to create an 'interpretive community'. This means that assessors must be able to demonstrate commitment to articulation of their own values and assumptions underlying judgements; they must be willing to engage in critical dialogue and meaningful negotiation, offer criticisms to others and be open for change in the light of the negotiation. The biggest challenge may very well be to make the necessary commitments of time and energy that are required to achieve trustworthiness in the assessment process. However, we feel that expanding our assessment repertoire with constructivist-interpretivist approaches may support new and much-needed directions in assessment and professional accountability. Engagement in discussion about performance values by communities of practice may furthermore fuel the debate about what constitutes excellence in professional competence and how assessment systems may contribute to improving the quality of patient care.

Finally, we think that conceptualisations of assessment and validity as described in this paper apply to all kinds of unstandardised assessments – in a range of (school-based) educational contexts. Changes in assessment towards assessment for learning, as well as acknowledgement that current measurement practices in educational assessment are not in line with current theories of learning and cognition, increas-

ingly call for reconsideration of conventional notions of assessment and assessment validity. In medical education, research into questions raised by interpretivist assessment approaches is badly needed.

Contributors: MG and CvdV worked collaboratively to develop the primary content of this paper. MG wrote the initial draft of the manuscript. Both MG and CvdV contributed to revisions of the initial draft for intellectual content and clarity. Both authors approved the final manuscript for publication.

Acknowledgements: The authors want to thank Kevin Eva, Tim Wilkinson, Cathy Haigh and an anonymous reviewer for their valuable comments, which helped improve the contents of this paper.

Funding: None.

Conflicts of interest: None.

Ethical approval: Not applicable.

GLOSSARY TABLE

Social/socio-cultural learning theories emphasise learning through active participation in social (authentic, professional activities). Learners develop by actively engaging in ongoing processes of workplaces. The learning processes as well as learning outcomes (performance) are determined by social, organisational, cultural and other contextual factors. However, socio-cultural learning theories also reject the idea that the individual learner should be the exclusive focus of analysis: learning can be either individual or social (collective).¹⁷

Constructivist-interpretivist assessment approaches view assessment to be value laden and socially constructed. Assessors are social beings who construct the assessment according to their own values, beliefs and perceptions. Performance can therefore never be objective. The interpretive approach focuses on participants' own perspectives in conceptualising and reconstructing their experiences, expectations, interpretations and assumptions.³⁸

Trustworthiness of qualitative assessment inquiry is important to evaluate its worth. Trustworthiness involves establishing⁵⁵:

Credibility, or confidence in the 'truth' of the findings;

Transferability, or showing that findings have applicability in other contexts;

Dependability, or showing that findings are consistent and could be repeated;

Confirmability, or the degree of 'neutrality' (findings not shaped by investigator bias, motivation or interest).

Specific strategies can be used for establishing each of these criteria in qualitative assessment inquiry.⁵⁸

REFERENCES

- 1 Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002;**287** (2):226–35.

- 2 Albanese MA. Challenges in using rater judgments in medical education. *J Eval Clin Pract* 2000;**6** (3):305–19.
- 3 Downing SM. Threats to the validity of clinical teaching assessments: what about rater error? *Med Educ* 2005;**39**:353–5.
- 4 Kreiter CD, Ferguson KJ. Examining the generalizability of ratings across clerkships using a clinical evaluation form. *Eval Health Prof* 2001;**24**:36–46.
- 5 van Barneveld C. The dependability of medical students' performance ratings as documented on in-training evaluations. *Acad Med* 2005;**80** (3):309–12.
- 6 Cook DA, Beckman TJ, Mandrekar JN, Pankratz VS. Internal structure of mini-CEX scores for internal medicine residents: factor analysis and generalizability. *Adv Health Sci Educ Theory Pract* 2010;**15** (5):633–45.
- 7 Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. *Med Educ* 2012;**46** (1):28–37.
- 8 Kane M. Validation. In: Brennan RL, ed. *Educational Measurement*. Westport, CT: American Council on Education/Praeger 2006; 621–94.
- 9 Moss PA, Girard BJ, Haniford LC. Validity in educational measurement. *Rev Res Educ* 2006;**30**: 109–62.
- 10 American Educational Research Association, American Psychological Association & National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association 1999.
- 11 Koch MJ, DeLuca C. Rethinking validation in complex high-stakes assessment contexts. *Assess Educ Princ Pol Pract* 2012;**19** (1):99–116.
- 12 Delandshere G. Assessment as inquiry. *Teach Coll Rec* 2002;**104** (7):1461–84.
- 13 Fisher CD. What if we took within-person performance variability seriously? *Ind Organ Psychol* 2008;**1**:185–9.
- 14 Sturman MC, Cashen LH, Cheramie RA. The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. *J Appl Psychol* 2005;**90** (2):269–83.
- 15 Durning SJ, Artino AR, Boulet JR, Dorrance K, Van der Vleuten CPM, Schuwirth LWT. The impact of selected contextual factors on experts' clinical reasoning performance (does context impact clinical reasoning performance in experts?). *Adv Health Sci Educ Theory Pract* 2012;**17** (1):65–79.
- 16 Lance CE, Hoffman BJ, Gentry WA, Baranik LE. Rater source factors represent important subcomponents of the criterion construct space, not rater bias. *Hum Resource Manag Rev* 2008;**18**:223–32.
- 17 Hager P. Theories of workplace learning. In: Malloch M, Cairns L, Evans K, O'Connor BN, eds. *The SAGE Handbook of Workplace Learning*. Los Angeles, CA: Sage Publications 2011; 17–32.
- 18 Bleakley A. Broadening conceptions of learning in medical education: the message from teamworking. *Med Educ* 2006;**40**:150–7.
- 19 Mann KV. Theoretical perspectives in medical education: past experience and future possibilities. *Med Educ* 2011;**45**:60–8.
- 20 Lingard L. Rethinking competence in the context of teamwork. In: Hodges BD, Lingard L, eds. *The Question of Competence: Reconsidering Medical Education in the Twenty-first Century*. London: Cornell University Press 2012; 42–70.
- 21 Fraser SW, Greenhalgh T. Coping with complexity: educating for capability. *BMJ* 2001;**323**:799–803.
- 22 Engeström Y, Sannino A. Studies of expansive learning: foundations, findings and future challenges. *Educ Res Rev* 2010;**5** (1):1–24.
- 23 Mennin S. Self-organisation, integration and curriculum in the complex world of medical education. *Med Educ* 2010;**44**:20–30.
- 24 Bleakley A. Blunting Occam's razor: aligning medical education with studies of complexity. *J Eval Clin Pract* 2010;**16**:849–55.
- 25 Ten Cate O, Scheele F. Competency-based postgraduate training: can we bridge the gap between theory and clinical practice? *Acad Med* 2007;**82**:542–7.
- 26 Hager P, Hodgkinson P. Moving beyond the metaphor of transfer of learning. *Br Educ Res J* 2009;**35** (4): 619–38.
- 27 Eva KW. On the generality of specificity. *Med Educ* 2003;**37** (7):587–8.
- 28 Fisher CD, Noble CS. A within-person examination of correlates of performance and emotions while working. *Hum Perf* 2004;**17**:145–68.
- 29 Deadrick D, Bennett N, Russell C. Using hierarchical linear modeling to examine dynamic performance criteria over time. *J Manage* 1997;**23**:745–57.
- 30 Stewart GL, Nandkeolyar AK. Exploring how constraints created by other people influence intraindividual variation in objective performance measures. *J Appl Psychol* 2007;**92** (4):1149–58.
- 31 Norcini J, Burch V. Workplace-based assessment as an educational tool. AMEE Guide no. 31. *Med Teach* 2007;**29** (9):855–71.
- 32 Beal DJ, Weiss HM, Barros E, MacDermid SM. An episodic process model of affective influences on performance. *J Appl Psychol* 2005;**90** (6):1054–68.
- 33 Groysberg B, McLean AN, Nohria N. Are leaders portable? *Harv Bus Rev* 2006;**84** (5):92–101.
- 34 Groysberg B, Lee L-E. Hiring stars and their colleagues: exploration and exploitation in professional service firms. *Organ Sci* 2009;**20** (4):740–58.
- 35 Wenghofer E, Williams AP, Klass DJ. Factors affecting physician performance: implications for performance improvement and governance. *Health Policy* 2009;**5** (2):e141–60.
- 36 Ginsburg S, Bernabeo E, Ross KM, Holmboe ES. "It depends": results of a qualitative study investigating how practicing internists approach professional dilemmas. *Acad Med* 2012;**87** (12):1–9.

- 37 Schuwirth LWT, van der Vleuten CPM. A plea for new psychometric models in educational assessment. *Med Educ* 2006;**40**:296–300.
- 38 Gipps C. Socio-cultural aspects of assessment. *Rev Res Educ* 1999;**24**:355–92.
- 39 Murphy KR, Cleveland JN. *Understanding Performance Appraisal. Social, Organizational and Goal-based Perspectives*. Thousand Oaks, CA: Sage Publications 1995.
- 40 Levy PE, Williams JR. The social context of performance appraisal: a review and framework for the future. *J Manag* 2004;**30**:881–905.
- 41 Tziner A, Murphy KR, Cleveland JN. Contextual and rater factors affecting rating behaviour. *Group Organ Manag* 2005;**30** (1):89–98.
- 42 Ferris GR, Munyon TP, Basik K, Buckley MR. The performance evaluation context: social, emotional, cognitive, political and relationship components. *Hum Resource Manag Rev* 2008;**18**:146–63.
- 43 Delandshere G, Petrosky A. Capturing teachers' knowledge: performance assessment and post-structuralism. *Educ Res* 1994;**23** (5):11–8.
- 44 Landy FJ, Farr JL. Performance rating. *Psychol Bull* 1980;**87** (1):72–107.
- 45 Kogan JR, Conforti L, Bernabeo E, Lobst W, Holmboe E. Opening the black box of clinical skills assessment via observation: a conceptual model. *Med Educ* 2011;**45**:1048–60.
- 46 Govaerts MJB, Van de Wiel MWJ, Schuwirth LWT, Van der Vleuten CPM, Muijtjens AMM. Workplace-based assessment: raters' performance theories and constructs. *Adv Health Sci Educ Theory Pract* 2013;**18** (3):375–96.
- 47 Shay SB. The assessment of complex performance: a socially situated interpretive act. *Harv Educ Rev* 2004;**74** (3):307–29.
- 48 Moss PA. Enlarging the dialogue in educational measurement: voices from interpretive research traditions. *Educ Res* 1996;**25** (1):20–8. 43.
- 49 Johnston B. Summative assessment of portfolios: an examination of different approaches to agreement over outcomes. *Stud High Educ* 2004;**29** (3):395–412.
- 50 Kuper A, Reeves S, Albert M, Hodges BD. Assessment: do we need to broaden our methodological horizons? *Med Educ* 2007;**41**:1121–3.
- 51 Delandshere G, Petrosky AR. Assessment of complex performances: limitations of key measurement assumptions. *Educ Res* 1998;**27** (2):14–24.
- 52 Kane MT. Terminology, emphasis, and utility in validation. *Educ Res* 2008;**37** (2):76–82.
- 53 Guba E, Lincoln Y. *Fourth Generation Evaluation*. London: Sage Publications 1989.
- 54 Lincoln YS, Guba EG. *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications 1985.
- 55 Krefting L. Rigor in qualitative research: the assessment of trustworthiness. *Am J Occup Ther* 1991;**45** (3):214–22.
- 56 Driessen E, van der Vleuten CPM, Schuwirth L, van Tartwijk J, Vermunt J. The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study. *Med Educ* 2005;**39**:214–20.
- 57 Tigelaar DEH, Dolmans DHJM, Wolfhagen IHAP, van der Vleuten CPM. Quality issues in judging portfolios: implications for organizing teaching portfolio assessment procedures. *Stud High Educ* 2005;**30** (5):595–610.
- 58 van der Vleuten CPM, Schuwirth LWT, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Cl Ob* 2010;**24** (6):703–19.
- 59 Wigton RS. What do the theories of Egon Brunswik have to say to medical education? *Adv Health Sci Educ Theory Pract* 2008;**13**:109–21.
- 60 Hodges B. Assessment in the post-psychometric era: learning to love the subjective and collective. *Med Teach* 2013;**35**(7):564–8.
- 61 Schuwirth L, Ash J. Assessing tomorrow's learners: in competency-based education only a radically different holistic method of assessment will work. *Six things we could forget*. *Med Teach* 2013;**35**(7):555–9.
- 62 Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach* 2011;**33**:478–85.

Received 6 April 2013; editorial comments to author 8 May 2013; accepted for publication 14 June 2013

Formative assessment and self-regulated learning: A model and seven principles of good feedback practice.

DAVID J. NICOL¹
University of Strathclyde

DEBRA MACFARLANE-DICK
University of Glasgow

¹ Author for Correspondence; Dr David J. Nicol, Centre for Academic Practice, Graham Hills Building, University of Strathclyde, 50 George St., Glasgow, G1 1QE.
e-mail: d.j.nicol@strath.ac.uk

Formative assessment and self-regulated learning: A model and seven principles of good feedback practice.

Abstract

The research on formative assessment and feedback is re-interpreted to show how these processes can help students take control of their own learning – i.e. become self-regulated learners. This reformulation is used to identify seven principles of good feedback practice that support self-regulation. A key argument is that students are already assessing their own work and generating their own feedback and that higher education should build on this ability. The research underpinning each feedback principle is presented and some examples of easy-to-implement feedback strategies are briefly described. This shift in focus, whereby students are seen as having a proactive rather than a reactive role in generating and using feedback, has profound implications for the way in which teachers organise assessments and support learning.

Introduction

This paper positions the research on formative assessment and feedback within a model of self-regulated learning. Formative assessment refers to assessment that is specifically intended to generate feedback on performance to improve and accelerate learning (Sadler, 1998). A central argument is that, in higher education, formative assessment and feedback should be used to empower students as self-regulated learners. The construct of self-regulation refers to the degree to which students can regulate aspects of their thinking, motivation and behaviour during learning (Pintrich and Zusho, 2002). In practice, self-regulation is manifested in the active monitoring and regulation of a number of different learning processes: e.g. the setting of, and orientation towards, learning goals; the strategies used to achieve goals; the management of resources; the effort exerted; reactions to external feedback; the products produced.

Intelligent self-regulation requires that the student has in mind some goals to be achieved against which performance can be compared and assessed. In academic settings, specific targets, criteria, standards and other external reference points (e.g. exemplars) help define goals. Feedback is information about how the student's present state (of learning and performance) relates to these goals and standards. Students generate internal feedback as they monitor their engagement with learning activities and tasks and assess progress towards goals. Those more effective at self-regulation, however, produce better feedback or are more able to use the feedback they generate to achieve their desired goals (Butler and Winne, 1995). Self-regulated learners also actively interpret external feedback, for example, from teachers and other students, in relation to their internal goals. Although research shows that students can learn to be more self-regulated (see Pintrich, 1995; Zimmerman & Schunk, 2001), how to enhance feedback (both self-generated and external) in support of self-regulation has not been fully explored in the current literature. This paper helps address this gap by proposing seven principles of good feedback practice in relation to the development of self-regulation.

The rationale for re-thinking formative assessment and feedback

Over the last two decades, there has been a shift in the way teachers and researchers write about student learning in higher education. Instead of characterising it as a simple acquisition process based on teacher transmission, learning is now more commonly conceptualised as a process whereby students actively construct their own knowledge and skills (Barr and Tagg,

1995; De Corte, 1996; Nicol, 1997). Students interact with subject content transforming and discussing it with others in order to internalise meaning and make connections with what is already known. Terms like 'student-centred learning', which have entered the lexicon of higher education, are one reflection of this new way of thinking. Even though there is disagreement over the precise definition of student-centred learning, the core assumptions are active engagement in learning and learner responsibility for the management of learning (Lea, Stephenson and Troy, 2003).

Despite this shift in conceptions of teaching and learning, a parallel shift in relation to formative assessment and feedback has been slower to emerge. In HE, formative assessment and feedback are still largely controlled by and seen as the responsibility of teachers; and feedback is still generally conceptualised as a transmission process even though some influential researchers have recently challenged this viewpoint (Yorke, 2003; Boud, 2000, Sadler, 1998). Teachers 'transmit' feedback messages to students about what is right and wrong in their academic work, about its strengths and weaknesses, and students use this information to make subsequent improvements.

There are a number of problems with this transmission view when applied to formative assessment and feedback. Firstly, if formative assessment is exclusively in the hands of teachers, then it is difficult to see how students can become empowered and develop the self-regulation skills needed to prepare them for learning outside university and throughout life (Boud, 2000). Secondly, there is an assumption that when teachers transmit feedback information to students these messages are easily decoded and translated into action. Yet, there is strong evidence that feedback messages are invariably complex and difficult to decipher and that students require opportunities to construct actively an understanding of them (e.g. through discussion) before they can be used to regulate performance (Higgins, Hartley and Skelton, 2001; Ivanic, Clark and Rimmershaw, 2000). Thirdly, viewing feedback as a cognitive process involving only transfer of information ignores the way feedback interacts with motivation and beliefs. Research shows that feedback both regulates and is regulated by motivational beliefs. External feedback has been shown to influence how students feel about themselves (positively or negatively) and what and how they learn (Dweck, 1999). Research also shows (Garcia, 1995) that beliefs can regulate the effects of feedback messages (e.g. perceptions of self-efficacy might be maintained by re-interpreting the causes of failure). Fourthly, as a result of this transmission view of feedback, the workload of teachers in HE increases year by year as student numbers and class sizes become larger. One way of addressing this issue is to re-examine the nature of feedback, and who provides it (e.g. teacher, peer, self), in relation to its effectiveness in supporting learning processes.

In the next section a conceptual model of formative assessment and feedback is presented that centres on the processes inherent in learner self-regulation. A key feature of the model that differentiates it from everyday understandings of feedback is that students are assumed to occupy a central and active role in all feedback processes. They are always actively involved in monitoring and regulating their own performance both in relation to desired goals and in terms of the strategies used to reach these goals. The student also actively constructs his or her own understanding of feedback messages derived from external sources (Ivanic, Clark and Rimmershaw, 2000; Black and Wiliam, 1998). This is consistent with the literature on student-centred and social constructivist conceptions of learning (Lea, Stephenson & Troy, 2003; Palinscar, 1998).

The conceptual model of self-regulation outlined in this paper draws on earlier work by Butler and Winne (1995). Their paper stands out as one of the few available to provide a theoretical synthesis of thinking about feedback and self-regulation. Following a presentation of the conceptual model, seven principles of good feedback practice are proposed; these are aligned to the model and backed up by a review of the research literature on assessment and feedback.

Relating the recent feedback research to the conceptual model adds significant value to this area of study. First, the model provides a coherent educational rationale to draw together some quite diverse research findings on formative assessment and feedback. Secondly, the model and seven principles offer complementary tools that teachers might use to think about the design of, and to evaluate, their own feedback procedures. In that context, after describing each principle we identify some related feedback strategies that teachers might easily implement.

A Conceptual Model of processes of self-regulation and internal feedback.

Figure 1 presents a conceptual model of self-regulation and feedback that synthesises current thinking in these areas. The top part of Figure 1 is based on a model originally published by Butler and Winne (1995). Processes internal to the learner are depicted inside the shaded area. This shows how the learner monitors and regulates learning and performance. It also shows the crucial role of internally generated feedback in these processes. Pintrich and Zusho (2002) provide the following working definition of self-regulation:

Self-regulated learning is an active constructive process whereby learners set goals for their learning and monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features of the environment. (p64)

This definition fits the purpose of this paper in that it recognises that self-regulation applies not just to cognition but also to motivational beliefs and overt behaviour. It also recognises that there are limits to learner self-regulation; for example, the teacher usually devises the learning task and determines the assessment requirements (see below).

In the model, an academic task set by the teacher (A), in class or set as an assignment, is shown as the trigger to initiate self-regulatory processes in the student (shown at centre of diagram). Engagement with the task requires that the student draw on prior knowledge and motivational beliefs (B) and construct a personal interpretation of the meaning of the task and its requirements. Based on this internal conception, the student formulates his or her own task goals (C). While there would normally be an overlap between the student's goals and those of the teacher, the degree of overlap may not be high (e.g. if the student wishes only to pass the assignment). The student's goals might also be fuzzy rather than clear (e.g. a vague intention or task orientation). Nonetheless, these goals would help shape the strategies and tactics (D) that are used by students to generate outcomes, both internal (E) and externally observable (F). Internal outcomes refer to changes in cognitive or affective/motivational states that occur during task engagement (e.g. increased understanding, changes in self-perceptions of ability). Externally observable outcomes refer to tangible products (e.g. essays) and behaviours (e.g. student presentations).

Monitoring these interactions with the task and the outcomes that are being cumulatively produced generates *internal feedback* at a variety of levels (i.e. cognitive, motivational and behavioural). This feedback is derived from a comparison of current progress against desired goals. It is these comparisons that help the student determine whether current modes of engagement should continue as is or if some type of change is necessary. For example, this self-generated feedback might lead to a re-interpretation of the task or to an adjustment of internal goals or of tactics and strategies. The student might even revise his or her domain knowledge or motivational beliefs which, in turn, might influence subsequent self-regulation.

In the model, external feedback to the student (G) might be provided by the teacher, by a peer or by other means (e.g. a placement supervisor, a computer). This additional information might augment, concur or conflict with the student's interpretation of the task and the path of learning. However, to produce an effect on internal processes or external outcomes the

student must actively engage with these external inputs. In effect, the teachers' feedback responses would have to be interpreted, constructed and internalised by the student if it were to have a significant influence on subsequent learning (Ivanic, Clark & Rimmershaw, 2000).

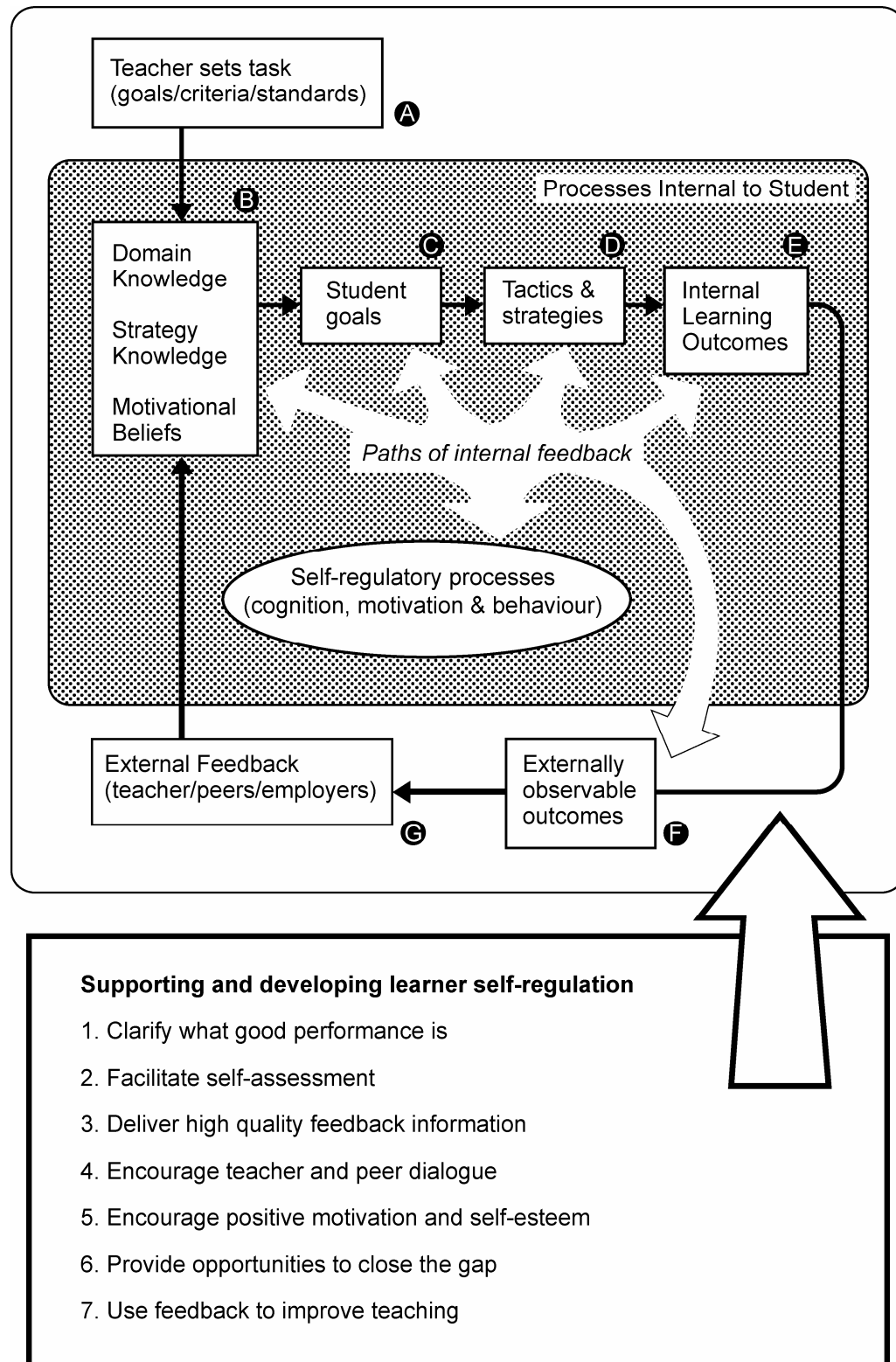


Figure 1. A model of self-regulated learning and the feedback principles that support and develop self-regulation in students.

Some supporting research

There is considerable research evidence to show that effective feedback leads to learning gains. Black and Wiliam (1998) drew together over 250 studies of feedback carried out since 1988 spanning all educational sectors. These studies focused on real teaching situations and the selection included teacher-made assessments and self and peer assessments. A meta-analysis of these studies revealed that feedback produced significant benefits in learning and achievement across all content areas, knowledge and skill types and levels of education. While the bulk of Black and Wiliam's data came from the school sector, their review and that of others (e.g. Hattie, 1987; Crooks, 1988) provide convincing evidence of the value of feedback in promoting learning. In addition, there is a large body of complementary research studies demonstrating the effects of self and peer feedback on learning (e.g. Boud, 1995; Boud, Cohen & Sampson, 1999). Nonetheless, while the work of Black and others has had an important influence on teaching practices in schools (Black, Harrison, Lee, Marshall and Wiliam, 2003) it has so far had much less influence on higher education.

One of the most influential papers underpinning the Black and Wiliam review, and the writings of other researchers (e.g. Yorke, 2003) is that of Sadler (1989). Sadler identified three conditions necessary for students to benefit from feedback in academic tasks. He argued that the student must know:

- i. what good performance is (i.e. must possess a concept of the goal or standard being aimed for);
- ii. how current performance relates to good performance (for this, students must be able to compare current and good performance);
- iii. how to act to close the gap between current and good performance.

From this analysis Sadler (1989) made an important observation: for students to be able to *compare* actual performance with a standard (as suggested by ii), and take action to close the gap (iii) then they '*must already possess some of the same evaluative skills as their teacher*' (Sadler, 1989). For some writers, this observation has led to the conclusion that, as well as improving the quality of feedback messages, teachers should focus much more effort on strengthening the skills of self-assessment in their students (Yorke, 2003; Boud, 2000). Sadler's argument, that students are already generating their own feedback, also helps account for the common finding that students still make significant progress in their learning in HE even when the external feedback they receive is quite impoverished (especially true in many large enrolment classes).

Although Sadler's writings are consistent with the argument in this paper, his focus on 'control theory and closing gaps' has been interpreted by some as too limited a basis to account for the range of effects produced by feedback (Gibbs, 2004). This paper addresses this concern by re-positioning formative assessment and feedback within a wider framework that encompasses self-regulation of motivation and behaviour as well as of cognition. For example, feedback is involved when students actively control their study time or their interactions with others (behaviour) and when they monitor and control motivational beliefs to adapt to the demands of the course (e.g. choosing a personal goal orientation).

Despite the appeal of self-regulation as a construct, it is important to recognise some basic assumptions underlying its use. While it is assumed that students can self-regulate internal states and behaviour as well as some aspects of the environment, this does not mean that the student always has full control. Learning tasks set by teachers, marking regimes and other course requirements are not under students' control even though students still have latitude to self-regulate within such constraints. Also, students often learn in implicit or unintentional ways without explicit regulation (e.g. aspects of some skills such as reading are automated).

There is a large body of empirical evidence, mainly published in the US, showing that learners who are more self-regulated are more effective learners: they are more persistent, resourceful, confident and higher achievers (Pintrich, 1995; Zimmerman and Schunk, 2001). Also, the more learning becomes self-regulated, the more students assume control over their learning and the less dependent they are on external teacher support when they engage in regulatory activities (Zimmerman and Schunk, 2004). Importantly, this research also shows that any student, even those 'at risk', can learn to become more self-regulating (Pintrich and Zusho, 2002). The development of self-regulation in students can be facilitated by structuring learning environments in ways that make learning processes explicit, through meta-cognitive training, self-monitoring and by providing opportunities to practise self-regulation (Schunk and Zimmerman, 1994; Pintrich, 1995). The unique contribution of this paper is to identify how formative assessment and feedback processes might help foster self-regulation. [It is beyond the scope of this paper to summarise the literature on self-regulation but a useful first text might be that by Zimmerman and Schunk (2001)].

Seven principles of good feedback practice: Facilitating self-regulation

From the self-regulation model and the research literature on formative assessment it is possible to identify some principles of good feedback practice. These are shown at the bottom of Figure 1. Good feedback practice is broadly defined here as anything that might strengthen the students' capacity to self-regulate their own performance. A synthesis of the research literature led to the following seven principles:

Good feedback practice:

1. helps clarify what good performance is (goals, criteria, expected standards);
2. facilitates the development of self-assessment (reflection) in learning;
3. delivers high quality information to students about their learning;
4. encourages teacher and peer dialogue around learning;
5. encourages positive motivational beliefs and self-esteem;
6. provides opportunities to close the gap between current and desired performance;
7. provides information to teachers that can be used to help shape the teaching.

The following sections provide the rationale for each principle in terms of the self-regulation and the associated research literature. Specific strategies that teachers can use to facilitate self-regulation are proposed after the presentation of each principle.

1. Helps clarify what good performance is.

Students can only achieve learning goals if they understand those goals, assume some ownership of them, and can assess progress (Sadler, 1989; Black & Wiliam, 1998). In academic settings, understanding goals means that there must be a reasonable degree of overlap between the task goals set by students and the goals originally set by the teacher. This is logically essential given that it is the students' goals that serve as the criteria for self-regulation (Figure 1). However, there is considerable research evidence showing significant mismatches between tutors' and students' conceptions of goals and of assessment criteria and standards.

Hounsell (1997) has shown that tutors and students often have quite different conceptions about the goals and criteria for essays in undergraduate courses in history and psychology and that poor essay performance is correlated with the degree of mismatch. In a similar vein, Norton (1990) has shown that when students were asked to rank specific assessment criteria for an essay task they produced quite different rankings from those of their teachers, emphasising content above critical thinking and argument. Weak and incorrect conceptions of goals not only influence what students do but also the value of external feedback information. If students do not share (at least in part) their teacher's conceptions of

assessment goals (and criteria and standards) then the feedback information they receive is unlikely to 'connect' (Hounsell, 1997). In this case, it will be difficult for students to evaluate discrepancies between required and actual performance. It is also important to note here that feedback not only has a role in helping guide students towards academic goals but, over time, it also has a role in helping clarify what these goals are (Sadler, 1989).

One way of clarifying task requirements (goals/criteria/standards) is to provide students with written documents containing statements that describe assessment criteria and/or the standards that define different levels of achievement. However, many studies have shown that it is difficult to make assessment criteria and standards explicit through written documentation or through verbal descriptions in class (Rust, Price & O'Donovan, 2003). Most criteria for academic tasks are complex, multidimensional (Sadler, 1989) and difficult to articulate; they are often 'tacit' and unarticulated in the mind of the teacher. As Yorke (2003) notes:

Statements of expected standards, curriculum objectives or learning outcomes are generally insufficient to convey the richness of meaning that is wrapped up in them (Yorke, 2003, p480)

Hence there is a need for strategies that complement written materials and simple verbal explanations. An approach that has proved particularly powerful in clarifying goals and standards has been to provide students with 'exemplars' of performance (Orsmond, Merry and Reiling, 2002). Exemplars are effective because they make explicit what is required and they define a valid standard against which students can compare their work.

Other strategies that have proved effective in clarifying criteria, standards and goals include: (i) providing better definitions of requirements using carefully constructed criteria sheets and performance level definitions; (ii) increasing discussion and reflection about criteria and standards in class (e.g. before an assignment); (iii) involving students in assessment exercises where they mark or comment on other students' work in relation to defined criteria and standards; (iv) workshops where students in collaboration with the teacher devise or negotiate their own assessment criteria for a piece of work. These strategies exemplify increasing levels of self-regulation

2. Facilitates the development of self-assessment (reflection) in learning

As suggested earlier, one effective way to develop self-regulation in students is to provide them with opportunities to practise regulating aspects of their own learning and to reflect on that practice. Students are (to some extent) already engaged in monitoring gaps between internally set task goals and the outcomes that they are generating (both internal and external). This monitoring is a by-product of such purposeful engagement in a task (Figure 1). However, in order to build on this, and to develop systematically the learner's capacity for self-regulation, teachers need to create more structured opportunities for self-monitoring and the judging of progression to goals. Self-assessment tasks are an effective way of achieving this, as are activities that encourage reflection on learning progress.

Over the last decade there has been an increasing interest in self-assessment in higher education (Boud, 1995). Research shows that, when suitably organised, self-assessment can lead to significant enhancements in learning and achievement. For example, McDonald and Boud (2003) have shown that training in self-assessment can improve students' performance in final examinations. Also, Taras (2001; 2002; 2003) has carried out a number of studies on student self-assessment in higher education which have shown positive benefits. In one study, students were trained in self-assessment under two conditions: self-assessment prior to peer and tutor feedback and self-assessment with integrated tutor feedback. The latter condition involved students self-assessing after they had received tutor feedback. The results showed that while both conditions benefited learning, self-assessment with integrated tutor feedback

helped students identify and correct more errors (those that they or peers had not been aware of) than self-assessment prior to peer or tutor feedback. Interestingly, this study not only shows the benefits of integrating external and internal feedback but it also shows ways of helping students internalise and use tutor feedback.

In developing self-assessment skills it is important to engage students in both identifying standards/criteria that will apply to their work (discussed in principle 1 above) and in making judgements about how their work relates to these standards (Boud, 1986). While structured opportunities for training in self-assessment are important there are other ways of supporting the development of these skills. One approach is to provide students with opportunities to evaluate and provide feedback on each other's work. Such peer processes help develop the skills needed to make objective judgements against standards, skills which are transferred when students turn to producing and regulating their own work (Boud, Cohen and Sampson, 1999; Gibbs, 1999). Another approach is to create frequent opportunities for reflection by students during their study. Cowan (1999) identifies ways that this can be done both in the context of simple classroom activities and during longer-term projects.

Other examples of structured reflection and self-assessment are varied and might include students: (i) requesting the kinds of feedback they would like when they hand in work; (ii) identifying the strengths and weaknesses in their own work in relation to criteria or standards before handing it in for teacher feedback; (iii) reflecting on their achievements and selecting work in order to compile a portfolio; (iv) reflecting before a task on achievement milestones and reflecting back on progress and forward to the next stage of action (Cowan, 1999).

3. Delivers high quality information to students about their learning.

While research shows that teachers have a central role in developing their students' own capacity for self-regulation, they are also a crucial source of external feedback. Feedback from teachers is a source against which students can evaluate progress and check out their own internal constructions of goals, criteria and standards. Moreover, teachers are much more effective in identifying errors or misconceptions in students' work than peers or the students themselves. In effect, feedback from teachers can help substantiate student self-regulation.

In the research literature there is little consensus about what constitutes good quality external feedback. Quality is defined quite broadly and tends to be discussed in relation to student needs and teacher-defined goals. For example, most researchers and textbook writers (e.g. Freeman and Lewis, 1998) are concerned that feedback to students might be delayed, not relevant or informative, that it might focus on low level learning goals or might be overwhelming in quantity or deficient in tone (i.e. too critical). For these researchers, the way forward is to ensure that feedback is provided in a timely manner (close to the act of learning production), that it focuses not just on strengths and weaknesses but also on offering corrective advice, that it directs students to higher order learning goals and that it involves some praise alongside constructive criticism. While each of these issues is important, there is a need for a more focused definition of quality in relation to external feedback, a definition that links more closely to the idea of self-regulation. Hence it is proposed here that:

Good quality external feedback is information that helps students trouble-shoot their own performance and self-correct: that is, it helps students take action to reduce the discrepancy between their intentions and the resulting effects.

In this context, it is argued that where feedback is given it is important that it is related to (and that students understand its relation to) goals, standards or criteria. Moreover, from this definition it is clear that external feedback should also help convey to students an appropriate

conception of the goal. This is not always the case. For example, it has become common practice in recent years to devise feedback sheets with assessment criteria as a way of informing students about task requirements and of providing consistent feedback in relation to goals (where there are a number of assessors). However, Sadler (1983) has argued that the use of criteria sheets often has unwanted effects in relation to essay assessments: for example, if there are a large number of criteria (12-20) this may convey to the student a conception of an the essay as a list of things to be done (ticked off) rather than as a holistic process (e.g. involving the production of a coherent argument supported by evidence). So as well as relating feedback to criteria and goals, teachers should also be alert to the fact that instruments they use to deliver feedback might adversely influence students' conceptions of the expected goals.

In the literature on essay assessment, some researchers have tried to formulate guidelines regarding the quantity and tone of feedback comments that, when analysed, show a close correspondence with the principle underlying the above definition of feedback quality. For example, Lunsford (1997) examined the written feedback comments given by writing experts on students' essays. From his analysis he made two proposals. Firstly, that three well thought out feedback comments per essay was the optimum if the expectation was that students would act on these comments. Secondly, and more importantly, these comments should indicate to the student how the reader (the teacher) experienced the essay as it was read (i.e. playing back to the students how the essay worked) rather than offer judgemental comments. Such comments would help the student grasp the difference between his or her intentions (goals) and the effects of the writing. Lunsford also advises that the comments should always be written in a non-authoritative tone and where possible they should offer corrective advice (both about the writing process as well as about content) instead of just information about strengths and weaknesses. In relation to self-regulation, Lunsford's reader-response strategy supports the shift from feedback provided by the teacher to students' evaluating their own writing.

The literature on external feedback is undeveloped in terms of how teachers should frame feedback comments, what kind of discourse should be used, how many comments are appropriate and in what context they should be made. Much more research is required in this area. One fruitful area of investigation is that, currently being conducted by Gibbs and Simpson (in press), on the relationship between feedback and the time students spend on task. They have shown that if students receive feedback often and regularly it enables better monitoring and self-regulation of progress by students. Other research is investigating the strengths of alternative modes of feedback communication (e.g. audio feedback, computer feedback) and of alternative ways of producing feedback information (e.g. poster productions where students get feedback by comparing their work with that of other students) (Hounsell, 2004; Hounsell & McCune, 2003).

Further strategies that increase the quality of teacher feedback based on the definition given above and on traditional research include: (i) making sure that feedback is provided in relation to pre-defined criteria but paying particular attention to the number of criteria; (ii) providing timely feedback – this means before it is too late for students to change their work (i.e. before submission) rather than just, as the research literature often suggests, soon after submission; (iii) providing corrective advice, not just information on strengths/weaknesses; (iv) limiting the amount of feedback so that it is actually used; (v) prioritising areas for improvement; (vi) providing online tests so that feedback can be accessed anytime, any place and as many times as students wish.

4. Encourages teacher and peer dialogue around learning.

In the self-regulation model, for external feedback to be effective it must be understood and internalised by the student before it can be used to make productive improvements. Yet in the research literature (Chanock, 2000; Hyland, 2000) there is a great deal of evidence that students do not understand the feedback given by tutors (e.g. 'this essay is not sufficiently analytical') and are therefore not able to take action to reduce the discrepancy between their intentions (goals) and the effects they would like to produce (i.e. the student may not know what to do to make the essay 'more analytical'). External feedback as a transmission process involving 'telling' ignores the active role the student must play in constructing meaning from feedback messages and of using this to regulate performance.

One way of increasing the effectiveness of external feedback and the likelihood that the information provided is understood by students is to conceptualise feedback more as *dialogue* rather than as information transmission. Feedback as dialogue means that the student not only receives initial feedback information but also has the opportunity to engage the teacher in discussion about that feedback. Some researchers maintain that teacher-student dialogue is essential if feedback is to be effective in higher education (Laurillard, 2002). Freeman and Lewis (1998) argue that the teacher 'should try to stimulate a response and a continuing dialogue – whether this be on the topics that formed the basis of the assignment or aspects of students' performance or the feedback itself' (p51). Discussions with the teacher help students to develop their understanding of expectations and standards, to check out and correct misunderstandings and to get an immediate response to difficulties.

Unfortunately, with large class sizes it can be difficult for the teacher to engage in dialogue with students. Nonetheless, there are ways that teachers might increase feedback dialogue even in these situations. One approach is to structure small group break-out discussions of feedback in class after students have received written comments on their individual assignments. Another approach is to use classroom technologies. These technologies help collate student responses to in-class questions (often multiple-choice questions) using handset devices. The results are feed back to the class visually as a histogram. This collated feedback has been used as a trigger for peer discussion (e.g. 'convince your neighbour that you have the right answer') and teacher-managed discussion in large classes (e.g. Nicol and Boyle, 2003; Boyle and Nicol, 2003).

These studies identify another source of external feedback to students – their peers. Peer dialogue enhances in students a sense of self-control over learning in a variety of ways. Firstly, students who have just learned something are often better able than teachers to explain it to their classmates in a language and in a way that is accessible. Secondly, peer discussion exposes students to alternative perspectives on problems and to alternative tactics and strategies. Alternative perspectives enable students to revise or reject their initial hypothesis and construct new knowledge and meaning through negotiation. Thirdly, by commenting on the work of peers, students develop detachment of judgement (about work in relation to standards) which is transferred to the assessment of their own work (e.g. 'I didn't do that either'). Fourthly, peer discussion can be motivational in that it encourages students to persist (see, Boyle and Nicol, 2003). Finally, it is sometimes easier for students to accept critiques of their work from peers rather than tutors.

Dialogical feedback strategies that support self-regulation include: (i) providing feedback using one-minute papers in class (see, Angelo and Cross, 1993); (ii) reviewing feedback in tutorials where students are asked to read the feedback comments they have been given earlier on an assignment and discuss these with peers (they might also be asked to suggest strategies to improve performance next time); (iii) asking students to find one or two examples of feedback comments that they found useful and to explain how they helped (iv) having students give each other descriptive feedback on their work in relation to published criteria before submission; (iv) group projects especially where students discuss criteria and standards before the project begins.

5. *Encourages positive motivational beliefs and self-esteem*

Motivation and self-esteem play a very important role in learning and assessment as is shown in Figure 1. Studies by Dweck (1999) show that depending on their beliefs about learning students possess qualitatively different motivational frameworks. These frameworks affect both students' responses to external feedback and their commitment to the self-regulation of learning.

Research in school settings has shown that frequent high stakes assessment (where marks or grades are given) has a 'negative impact on motivation for learning that militates against preparation for lifelong learning' (Harlen & Crick, 2003). Dweck (1999) argues that such assessments encourage students to focus on performance goals (passing the test, looking good) rather than learning goals (mastering the subject). In one study, Butler (1988) demonstrated that feedback comments alone increased students' subsequent interest in learning when compared with two other controlled situations, one where only marks were given and the other where students were given feedback and marks. Butler argued that students paid less attention to the comments when given marks and consequently did not try to use the comments to make improvements. This phenomenon is also commonly reported by academics in higher education.

Butler (1987) has also argued that grading student performance has less effect than feedback comments because it leads students to compare themselves against others (ego-involvement) rather than to focus on the difficulties in the task and on making efforts to improve (task-involvement). Feedback given as grades has also been shown to have especially negative effects on the self-esteem of low ability students (Craven, Marsh & Debus, 1991).

Dweck (1999) has interpreted these findings in terms of a developmental model that differentiates students into those who believe that ability is fixed and that there is a limit to what they can achieve (the 'entity view') and those that believe that their ability is malleable and depends on the effort that is input into a task (the 'incremental view'). These views affect how students respond to learning difficulties. Those with an entity view (fixed) interpret failure as a reflection of their low ability and are likely to give up whereas those with an incremental view (malleable) interpret this as a challenge or an obstacle to be overcome and increase their effort. Grant and Dweck (2003) have confirmed the validity of this model within higher education as have Yorke and Knight (2003) who found that about one-third of a sample of 2269 undergraduates students in first and final years, and across a range of disciplines, held beliefs in fixed intelligence.

Although this is an under-explored area of research in HE, there is evidence that teachers can have a positive or negative effect on motivation and self-esteem. They can influence both the goals that students set (learning or performance goals) as well as their commitment to those goals. Praising effort and strategic behaviours, and focusing students through feedback on learning goals, leads to higher achievement than praising ability or intelligence. The latter can result in a learned-helplessness orientation (Dweck, 1999). As Black and Wiliam (1998) note, feedback that draws attention away from the task and towards self-esteem can have a negative effect on attitudes and performance. In other words, it is important that students understand that feedback is an evaluation, not of the person but of the performance in context. This holds true whether the feedback derives from an external source or is generated through self-assessment.

These studies on motivation and self-esteem are important - they help explain why students often fail to self-regulate. In terms of teaching practice they suggest that motivation and self-esteem are more likely to be enhanced when a course has many low-stakes assessment tasks, with feedback geared to providing information about progress and achievement, rather than

high stakes summative assessment tasks where information is only about success or failure or about how students compare with their peers (e.g. grades). Other strategies that help encourage high levels of motivation and self-esteem include: (i) providing marks on written work only after students have responded to feedback comments (Gibbs, 1999); (ii) allocating time for students to re-write selected pieces of work – this would help change students' expectations about purpose and learning goals; (iii) automated testing with feedback; (iv) drafts and resubmissions.

6. Provides opportunities to close the gap between current and desired performance.

So far, feedback has been discussed from a cognitive or informational perspective and from a motivational perspective. However, in terms of self-regulation we must also consider how feedback influences behaviour and the academic work that is produced. According to Yorke (2003), two questions might be asked regarding external feedback. First, is the feedback of the best quality and second, does it lead to changes in student behaviour? Many writers have focused on the first question but the second is equally important. External feedback provides an opportunity to close a gap between current performance and the performance expected by the teacher. As Boud notes:

The only way to tell if learning results from feedback is for students to make some kind of response to complete the feedback loop (Sadler, 1989). This is one of the most often forgotten aspects of formative assessment. Unless students are able to use the feedback to produce improved work, through for example, re-doing the same assignment, neither they nor those giving the feedback will know that it has been effective. (Boud, 2000, p158)

In the self-regulation model (Figure 1), Boud's arguments about closing the performance gap might be viewed in two ways. First, closing the gap is about supporting students while engaged in the act of production of a piece of work (e.g. essays, presentations). Second, it is about providing opportunities to repeat the same 'task-performance-external feedback cycle' by, for example, allowing resubmission. External feedback should support both processes: it should help students to recognise the next steps in learning and how to take them, both during production and in relation to the next assignment.

Supporting the act of production requires the generation of concurrent or intrinsic feedback that students can interact with while engaged in an assessment task. This feedback would normally be built into the task (e.g. a group task with peer interaction, or a computer simulation) or the task might be broken down into components each associated with its own feedback. Many forms of electronic feedback (e.g. online simulations) can be automatically generated to support task engagement (Bull & McKenna, 2004). Providing feedback at sub-task level is not significantly different from other forms of feedback described in this paper.

In higher education, most students have little opportunity to use directly the feedback they receive to close the performance gap especially in the case of planned assignments. Invariably they move on to the next assessment task soon after feedback is received. While not all work can be re-submitted, many writers argue that re-submissions should play a more prominent role in learning (Boud, 2000). Also, greater emphasis might need to be given to providing feedback on work-in-progress (e.g. on structures for essays, plans for reports, sketches) and to encouraging students to plan the strategies they might use to improve subsequent work (Hounsell, 2004).

The following are some specific strategies to help students use external feedback to regulate and close the performance gap: (i) provide feedback on work in progress and increase opportunities for resubmission; (ii) introduce two stage assignments where feedback on stage one helps improve stage two (Gibbs, 2004); (iii) teachers might model the strategies they

would use to close a performance gap in class (e.g. model how to structure an essay when given a new question); (iv) specifically provide some 'action points' alongside the normal feedback provision; (v) involve students in groups in identifying their own action points in class after they have read the feedback on their assignments. The latter strategy would integrate feedback into the teaching and learning process and involve the students more actively in the generation and planned use of feedback.

7. *Provides information to teachers that can be used to help shape the teaching.*

Good feedback practice is not only about providing accessible and usable information that helps students improve their learning, but it is also about providing good information to teachers. As Yorke (2003) notes:

The act of assessing has an effect on the assessor as well as the student. Assessors learn about the extent to which they [students] have developed expertise and can tailor their teaching accordingly (York, 2003, p482)

In order to produce feedback that is relevant and informative and meets students' needs, teachers themselves need good data about how students are progressing. They also need to be involved in reviewing and reflecting on this data and in taking action to help support the development of self-regulation in their students.

In the self-regulation model (Figure 1) information about students only becomes available when the learning outcomes are translated into public performances and products. Teachers help generate this public information about students through a variety of methods – by setting assessment tasks, by questioning of students in class and by observing behaviour (e.g. presentations). Such information helps teachers uncover student difficulties with subject matter (e.g. conceptual misunderstandings) and with study methods.

Frequent assessment tasks, especially diagnostic tests, can help teachers generate cumulative information about students' levels of understanding and skill so that they can adapt their teaching accordingly. This is one of the key ideas behind the work in the US of Angelo and Cross (1993). They have shown how teachers can gain regular feedback information about student learning within large classes by using variants of the one-minute paper – questions that are posed to students before a teaching session begins and responded to at the end of the session (e.g. What was the most important argument in this lecture? What question remains uppermost in your mind now at the end of this teaching session?). These strategies can be adapted to any classroom situation or discipline. Moreover, they help develop in students important meta-cognitive skills such as the ability to think holistically and to identify gaps in understanding (Steadman, 1998).

As well as giving feedback to the teacher, one-minute papers can also be used to provide feedback to the student (e.g. when teachers replay some of the student responses to the one-minute paper in class at the next teaching session). Indeed, this approach allows teachers and students to share, on a regular basis, their conceptions about both the goals and processes of learning (Stefani & Nicol, 1997) thus supporting academic self-regulation.

Other strategies available to teachers to help generate and collate quality information about student learning include (i) having students request the feedback they would like when they make an assignment submission (e.g. on a proforma with published criteria); (ii) having students identify where they are having difficulties when they hand in assessed work; (iii) asking students in groups to identify 'a question worth asking', based on prior study, that they would like to explore for a short time at the beginning of the next tutorial.

Conclusion and future work

This paper has argued that conceptions of assessment have lagged behind conceptions of learning in HE. While students have been given more responsibility for learning in recent years there has been far greater reluctance to give them increased responsibility for assessment processes (even low stakes formative processes). Yet, if students are to be prepared for learning throughout life they must be provided with opportunities to develop the capacity to regulate their own learning as they progress through higher education. This paper has identified ways in which formative assessment and feedback might be organised so as to support this development. It has provided some key principles of good feedback practice that address a wide spectrum - the cognitive, behavioural and motivational aspects of self-regulation. How might teachers use the ideas in this paper? One practical proposal is that teachers examine current assessment practices in relation to the self-regulation model and to the seven principles. An audit of this kind might help identify where assessment practices might be strengthened. However, the seven principles presented here do not exhaust all that teachers might do to enhance self-regulated learning in HE classrooms. They merely provide a starting point. The research challenge is to refine these principles, identify gaps and to gather further evidence about the potential of formative assessment and feedback to support self-regulation.

Acknowledgements

Thanks to David Boud (University of Technology, Sydney) and to Graham Gibbs (Oxford University, UK) for feedback on a draft of this article. Thanks to the Learning and Teaching Support Network (now the Higher Education Academy, UK) for funding the Student Enhanced Learning through Effective Feedback (SENLEF) project which led us to review the assessment literature. Also thanks to our SENLEF project colleagues Charles Juwah, Bob Matthew, David Ross and Brenda Smith, for their input.

References

- Angelo, T. & Cross, P. (1993) *Classroom Assessment Techniques* (New York, Jossey-Bass).
- Barr, R. B. and Tagg, J. (1995) A New Paradigm for Undergraduate Education, *Change*, 27(6), 13-25.
- Black, P., Harrison, C., Lee, C., Marshal, B. & Wiliam, D. (2003) *Assessment for Learning: putting it into practice* (Maidenhead, Open University Press).
- Black, P. & Wiliam, D. (1998) Assessment and classroom learning, *Assessment in Education*, 5(1), 7-74.
- Boud, D. (1986) *Implementing Student Self-Assessment* (Sydney, Higher Education Research and Development Society of Australia).
- Boud, D. (1995) *Enhancing learning through self-assessment* (London, Kogan Page).
- Boud, D. (2000) Sustainable assessment: rethinking assessment for the learning society, *Studies in Continuing Education*, 22(2), 151-167.
- Boud, D., Cohen, R. and Sampson, J. (1999) Peer learning and assessment, *Assessment and Evaluation in Higher Education*, 24(4), 413-426.
- Boyle, J.T. and Nicol, D. J. (2003) Using classroom communication systems to support interaction and discussion in large class settings, *Association for Learning Technology Journal*, 11(3), 43-57.
- Bull, J. & McKenna, C. (2004) *Blueprint for computer-assisted assessment* (London, RoutledgeFalmer).
- Butler, D.L. & Winne, P.H. (1995) Feedback and self-regulated learning: a theoretical synthesis, *Review of Educational Research*, 65(3), 245-281.
- Butler, R. (1987) Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest and performance, *Journal of Educational Psychology*, 78(4), 210-216.
- Butler, R. (1988) Enhancing and undermining intrinsic motivation: the effects of task-involving and ego-involving evaluation on interest and involvement, *British Journal of Educational Psychology*, 58, 1-14.
- Chanock, K. (2000) Comments on essays: do students understand what tutors write?, *Teaching in Higher Education*, 5(1), 95-105.
- Cowan, J. (1999) *Being an innovative university teacher* (Buckingham, SRHE & Open University Press).
- Craven, R.G., Marsh, H. W., & Debus, R.L. (1991) Effects of internally focused feedback on the enhancement of academic self-concept, *Journal of Educational Psychology*, 83(1), 17-27.
- Crooks, T.J. (1988) The impact of classroom evaluation practices on students, *Review of Educational Research*, 58, 438-481.

- DeCorte, E. (1996) New perspectives on learning and teaching in higher education, in: A. Burgen (Ed) *Goals and Purposes of Higher Education in the 21st Century* (London, Jessica Kingsley Publishers).
- Dweck, C. (1999) *Self-theories: Their Role in Motivation, Personality and Development* (Philadelphia: Psychology Press).
- Freeman, R. & Lewis, R. (1998) *Planning and Implementing Assessment* (London, Kogan Page).
- Garcia, T. (1995) The role of motivational strategies in self-regulated learning, in: P. R. Pintrich (Ed) *Understanding self-regulated learning* (San Francisco, Jossey-Bass).
- Gibbs, G. (1999) Using assessment strategically to change the way students learn, in: S. Brown & A. Glasner (Eds) *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches* (Buckingham, SRHE/Open University Press).
- Gibbs, G. (2004). Personal communication.
- Gibbs, G and Simpson, C. (in press) Does your assessment support your students' learning?, *Learning and Teaching in Higher Education, Volume 1*.
- Grant, H. & Dweck, C.S. (2003) Clarifying achievement goals and their impact, *Journal of Personality and Social Psychology*, 85, 541-553.
- Harlen, W. & Crick, R.D. (2003) Testing and motivation for learning, *Assessment in Education*, 10(2), 169-207.
- Hattie, J. A. (1987) Identifying the salient facets of a model of student learning: a synthesis and meta-analysis, *International Journal of Educational Research*, 11, 187-212.
- Higgins, R., Hartley, P. & Skelton, A. (2001) Getting the message across: the problem of communicating assessment feedback, *Teaching in Higher Education*, 6(2), 269-274.
- Hounsell, D. (1997) Contrasting conceptions of essay-writing, in: F. Marton, D. Hounsell and N. Entwistle (Eds) *The Experience of Learning* (Edinburgh, Scottish Academic Press).
- Hounsell, D. (2004) Reinventing feedback for the contemporary Scottish university, paper presented at Quality Enhancement Conference on Assessment, University of Glasgow, 4 June.
- Hounsell, D. & McCune, V. (2003) Students' experiences of learning to present, in: C. Rust (Ed) *Improving student learning theory and practice – ten years on* (Oxford, Oxford Centre for Staff and Learning Development), 108-119.
- Hyland, P. (2000) Learning from feedback on assessment, in: A. Booth and P. Hyland (Eds) *The practice of university history teaching* (Manchester, Manchester University Press).
- Ivanic, R., Clark, R. and Rimmershaw, R. (2000) What am I supposed to make of this? The messages conveyed to students by tutors' written comments, in: M.R. Lea and B. Stierer, (Eds) *Student Writing in Higher Education: New Contexts* (Buckingham, SHRE/Open University Press).
- Laurillard, D. (2002) *Rethinking University Teaching: a conversational framework for the effective use of learning technologies*, 2nd edition (London, RoutledgeFalmer).

Lea, S.J., Stephenson, D. & Troy, J. (2003) Higher education students' attitudes to student-centred learning: beyond 'educational bulimia', *Studies in Higher Education*, 28(3), 321-334.

Lunsford, R. (1997) When less is more: principles for responding in the disciplines, in: M. Sorcinelli and P. Elbow (Eds) *Writing to learn: strategies for assigning and responding to writing across the disciplines* (San Francisco, Jossey-Bass).

McDonald, B. & Boud, D. (2003) The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations, *Assessment in Education*, 10(2), 209-220.

Nicol, D.J. (1997) *Research on Learning and Higher Education Teaching*, UCoSDA Briefing Paper 45 (Sheffield, Universities and Colleges Staff Development Agency).

Nicol, D.J. & Boyle, J.T. (2003) Peer Instruction versus Class-wide Discussion in large classes: a comparison of two interaction methods in the wired classroom, *Studies in Higher Education*, 28(4), 457-473.

Norton, L. S. (1990) Essay writing: what really counts?, *Higher Education*, 20(4), 411-42.

Orsmond, P., Merry, S. & Reiling, K. (2002) The use of formative feedback when using student derived marking criteria in peer and self-assessment, *Assessment & Evaluation in Higher Education*, 27(4), 309-323.

Palinscar, A.S. (1998) Social constructivist perspectives on teaching and learning, *Annual Review of Psychology*, 49, 345-375.

Pintrich, P. R. (1995) *Understanding self-regulated learning* (San Francisco, Jossey-Bass).

Pintrich, P. R. and Zusho, A. (2002) Student motivation and self-regulated learning in the college classroom, in: J. C. Smart and W.G. Tierney (Eds) *Higher Education: Handbook of Theory and Research*, Volume XVII (New York, Agathon Press).

Rust, C., Price, M. and O'Donovan, B. (2003) Improving students' learning by developing their understanding of assessment criteria and processes, *Assessment and Evaluation in Higher Education*, 28(2), 147-164.

Sadler, D. R. (1983) Evaluation and the improvement of academic learning, *Journal of Higher Education*, 54(1), 60-79.

Sadler, D.R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*, 18, 119-144.

Sadler, D.R. (1998) Formative assessment: revisiting the territory, *Assessment in Education*, 5(1), 77-84.

Schunk, D.H. & Zimmerman, B.J. (1994) *Self-regulation of learning and performance: issues and educational applications* (Lawrence Erlbaum Associates, New Jersey).

Steadman, M. (1998) Using classroom assessment to change both learning and teaching, *New Directions for Teaching and Learning*, 75, 23-35.

Stefani, L. & Nicol, D. (1997) From teacher to facilitator of collaborative enquiry, in: S. Armstrong, G. Thompson and S. Brown (Eds) *Facing up to Radical Changes in Universities and Colleges* (London, Kogan Page).

Taras, M. (2001) The use of tutor feedback and student self-assessment in summative assessment tasks; towards transparency for students and tutors, *Assessment and Evaluation in Higher Education*, 26(6), 605-614.

Taras, M. (2002) Using assessment for learning and learning from assessment, *Assessment and Evaluation in Higher Education*, 27(6), 501-510.

Taras, M. (2003) To feedback or not to feedback in student self-assessment, *Assessment and Evaluation in Higher Education*, 28(5), 549-565.

Yorke, M (2003) Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice, *Higher Education*, 45(4), 477-501.

Yorke, M. & Knight, P. (2004) Self-theories: some implications for teaching and learning in higher education, *Studies in Higher Education*, 29(1), 25-37.

Zimmerman, B.J. & Schunk, D.H. (2001) *Self-regulated learning and academic achievement: theoretical perspectives* (Lawrence Erlbaum Associates, New Jersey).

Zimmerman, B.J. & Schunk, D.H. (2004) Self-regulating intellectual processes and outcomes: a social cognitive perspective, in D. Y. Dai & R. J. Sternberg (Eds) *Motivation, emotion and cognition* (Lawrence Erlbaum Associates, New Jersey).

Patient OPUS ID:

Student:

Grupp:

Datum:

Ansvariga lärare:

Prov godkänt:

Prov underkänt:

Planerad behandling	
Tand nr	
Klass	
Yta	
Orsak	
Granntandsyta (intakt fyllning/tand)	
Antagonist i normal relation	
Granntand i normal relation	
Preparation	
Kariesfri	
Putsad granntandsfyllning	
Torrläggning m kofferdam	
Om granntandsskada har uppstått under behandlingen, underkänns provet i sin helhet	
Sammanfattad bedömning	GK/UK

Estetiskt resultat

	Utmärkt (perfekt, går inte göra bättre)	Acceptabelt (godkänt resultat)	Behöver åtgärdas (mindre defekter som går att åtgärda)	Oacceptabelt (Orsakar iatrogena skador)	POÄNG
Anatomi	4	3	2	1	
Färg/Translucens	4	3	2	1	
Ytfinish	4	3	2	1	

Max poäng 12. Godkänt 11.

Funktionella egenskaper

	Utmärkt (perfekt, går inte göra bättre)	Acceptabelt (godkänt resultat)	Behöver åtgärdas (mindre defekter som går att åtgärda)	Oacceptabelt (Orsakar iatrogena skador)	POÄNG
Kantanslutning	4	3	2	1	
Approx kontakt	4	3	2	1	
Överskott	4	3	2	1	
Occlusion Artikulation	4	3	2	1	
Patient- tillfredsställelse	4	3	2	1	

Max poäng 20 Godkänt 18.

Patientmötet		
Godkänt	Studenten visar under examinationen kunskap i hantering av instrument och material samt ett evidensbaserat arbetssätt	
	Genomför en relevant anamnes samt väljer adekvata undersökningsmoment. Utför dessa på ett säkert sätt med fullgod struktur, gott handlag inklusive ett gott omhändertagande av patienten.	
	Utför undersökning med god arbetsställning.	
	Visar flexibilitet och handlingsberedskap i anamnes- och undersökningssituationen.	
	Visar god kommunikationsförmåga med patienten och visar huvudsakligen ett aktivt lyssnande och anpassat språk	
	Visar ett gott förhållningssätt till patienten	
Underkänt	Studenten visar under examinationen brister i kunskap i hantering av instrument och material samt ett icke-evidensbaserat arbetssätt	
	Patientmötet genomförs ej enligt kriterierna för Godkänd.	
Sammanfattad bedömning		GK/UK

Efterföljande diskussion		
Godkänt	Utför en kortfattad och relevant redogörelse avseende sammanfattning av anamnes och undersökning.	
	Visar förmåga att i diskussion motivera och reflektera över genomförd anamnes, valda undersökningsmoment, mål, förslag till åtgärder samt utvärdering.	
	Kan motivera materialval och materialbehandling.	
	Kan redogöra för materialens påverkan på oral och allmänhälsa samt miljö	
	Visar förmåga att kritiskt reflektera över sin egen insats och fortsatt lärande.	
Underkänt	Efterföljande diskussion genomförs ej enligt kriterierna för Godkänd eller med alltför mycket hjälp och guidande frågor.	
Sammanfattad bedömning		GK/UK

Proven skall utföras under terminerna 7-8 med minst 1 godkänt prov / termin. I StuDentiGroup ska aktiviteten "Godkänd för fyllningsprov" vara ibockad då student anses nått tillräcklig klinisk färdighet avseende restorationer (minst 5 st). (Aktiviteten finns under Nivå 4/Kliniska prov och färdigheter)

Lämpliga tänder för prov planeras med karieslärare, förslagsvis i samband med undersökning. De två proven skall utföras på restorationer av klass II och III (ev.IV alternativt klass II med kuspörlust).

Prov bokas i Kvalitetsprovspärmen. Proven skall bedömas av 2 lärare. Dessa har varsitt protokoll och gör individuell slutbedömning av restoration efter utfört prov. CAR lärare ansvarar för provets genomförande och avslutar med gemensam Reflektion.

Vid underkänt prov får nytt prov göras tidigast efter att ett visst antal fyllningar utförts. Kriterierna för godkänt prov skall finnas tillgängliga på kurswebben. Godkänt prov registreras i StuDentiGroup och underskrivet formulär lämnas till kurssekreterare . Proven sparas i pärm.

E-Calib grund för bedömning hittar du på:

" <http://zep01793.dent.med.uni-muenchen.de/moodle/login/index.php>"

För godkänt resultat på kursen gäller:

- Närvaro, minst 90 %, av schemalagd tid samt minst 70 % patientbehandling av kursens bokningsbara pass
- Att studenten uppvisar prov på klinisk mognad
- Godkänd fallpresentation samt granskning av annan students fallpresentation
- Godkända kliniska moment

Kriterier avseende invisnings moment:

- Studenten skall komma överens med lärare om lämplig patient och tidpunkt att utföra momentet.
- Studenten redogör för de ingående kriterierna innan invisning.
- Patienten (patientfallet) skall grundligt introduceras/beskrivas för lärare innan behandlingen/momentet utförs.
- Proven utförs med hög grad av självständighet termin 5

Parodontal undersökning, diagnostik och terapiplanering av patient med parodontala problem (ej patient med "enbart" gingivit)

Följande faktorer skall beaktas vid bedömning av undersökning, diagnostik och terapiplanering:

- den studerandes egen värdering av genomförd undersökning
- över/under diagnostik
- tolkning av undersökningsdata
- har studenten använt erforderliga diagnostiska hjälpmedel?

Av de ingående delarna;

1. PLI
2. MBI
3. BoP
4. Fickstatus
5. Furkationsinvolveringar
6. Mobilitet

Ev. kan komplettering göras på annan patient när det gäller punkt 5 och 6.

Den parodontala diagnostiken skall göras på såväl tand som individnivå.

- **Utföra terapiplanering inom tandhygienistens kompetensområde** och där så erfordras i samråd med tandläkare eller annan vårdgivare. Terapiplaneringen skall vara tydligt utformad i journalen. I enlighet med Tandvårdslagen (1985:125) skall vården om möjligt utformas och genomföras i samråd med patienten.

Utvärdera denna nya patient under termin 5.

Utvärdering/Epikris, Parodpatient (patient med parodontala problem, ej "enbart" gingivit).

- Efter avslutad behandling skall studenten själv kunna utforma en sammanfattande sjukdomshistoria, behandling/behandlingsresultat och prognos/prognoser ev. recall.
- Denna skall införas i lämplig mall; epikrismall/parodutvärdering/kariesriskbedömning efter godkännande av lärare

Kariesregistrering, diagnostik och terapiplanering

Syftet är att studenten självständigt skall kunna kariesregistrera en (*om möjligt*) sanerings patient vad gäller:

- primär- och sekundärkaries på kron- och rottytor (diagn. av D3 är önskvärt)
- bestämning av kariesskadans allvarlighetsgrad och behandling

Följande faktorer bedöms:

- att polering och ev. depuration har föregått kariesregistreringen
- att nödvändiga röntgenbilder finns tillgängliga
- att sond, spegel och bomullsrullar finns tillgängliga
- tolkning av undersökningsdata
- över/underdiagnostisering
- journalföring med uträknade kariologiska index (ex. DMFT)
- den studerandes egen värdering av undersökningen
- Utföra terapiplanering inom tandhygienistens kompetensområde och där så erfordras i samråd med tandläkare eller annan vårdgivare. Terapiplaneringen skall vara tydligt utformad i journalen. I enlighet med Tandvårdslagen (1985:125) skall vården om möjligt utformas och genomföras i samråd med patienten.

Utvärdera denna patient om möjligt under terminen. Att att du har en patient från termin 4 som du utvärderar under termin 5.

Utvärdering/Epikris, Kariespatient

- Efter avslutad behandling skall studenten själv kunna utforma en sammantagen sjukdomshistoria, behandling/behandlingsresultat och prognos/prognoser ev. recall.
- Denna skall införas i lämplig mall; epikrismall/kariesriskbedömning efter godkännande av lärare

Undersökning, diagnostik och terapiplanering av Sanering/revisionspatient

(Se kariologisk, parodontologisk undersökning, diagnostik och terapiplanering)

- fullständig undersökning av sanerings- eller revisionspatient, oavsett patientens munstatus

Infiltrationsanestesi (ett arbetsområde/invisning)

Syftet är att använda infiltrationsanestesi när denna metod är lämplig

- injektionen skall läggas vid korrekt region och på korrekt sätt
- studenten redogör för anatomiska strukturer
- studenten redogör för val av anestesimedel
- uppvisar korrekt handhavande av säkerhetssprutan
- studenten visar och redogör för hur man aspirerar
- studenten skall kunna bedöma om optimal effekt har uppnåtts samt
- visar på gott omhändertagande av patienten före, under tiden och efter injektionen

Ledningsanestesi (ett arbetsområde/invisning)

Syftet är att kunna använda ledningsanestesi när denna metod är lämplig

- injektionen skall läggas på korrekt sätt (ett arbetsområde/invisning)
- studenten redogör för anatomiska strukturer
- studenten redogör för val av anestesimedel
- korrekt handhavande av säkerhetssprutan
- studenten visar och redogör för hur man aspirerar
- studenten skall kunna bedöma om optimal effekt har uppnåtts
- visar på gott omhändertagande av patienten före, under tiden och efter injektionen

Slutinvisning av patient (JS & kliniklärare)

- Student redovisar den patient som planeras som examinationspatient/patientfallsredovisning för kursansvarig/examinator och kliniklärare.
- Studenten skall själv ha genomfört betydande delar av behandling.
- Stor vikt läggs vid att studenten ser helheten i behandlingen och har ett kritiskt förhållningssätt till utförd behandling och framtida prognos.
- Studenten skall uppvisa kunskaper om och förståelse för, a) munhålets betydelse för det allmänna välbefinnandet, och b) allmäntillståndets inverkan på munhälsan.
- Journal etc. skall vara korrekta och signerade av lärare.

- Vid inisningen är patienten ej närvarande. Ej heller någon patient vid detta pass. Passet är förbokad i schemat.

Patientfallsredovisning II (JS)

- På termin 5 sker en examination av ett patientfall. Se separata anvisningar. Såväl skriftlig sammanställning av patientfallet som en muntlig redovisning och granskning av annat arbete samt leda diskussion ingår i examinationen.

Rapport 2017 fra Programsensor, Det Medicinsk-odontologiske fakultet under programudvalg for medicin: Profesjonsstudiet i medicin

Programsensor:

Charlotte Ringsted

Professor and Director, Centre for Health Sciences Education

Prodekan, Faculty of Health, Aarhus University

Aarhus, Denmark

Denne rapport er foretaget på baggrund af følgende dokumenter: Nyhedsbreve (feb; apr/maj; aug; dec); Referater fra Styregruppemøder (19. jan; 26. april; 27. sept); referater fra møder i Programudvalget for medicin 18 jan (temamøde – emnebeskrivelser); 22. feb; 5. apr; 7. juni; 13. sep; 1. nov; 6. dec) samt Notat om evaluering diskuteret på PUM 13. sept 2017.

I 2017 har programsensor været i Bergen til halvdagsseminar 9 feb og halvdagsmøde med programudvalg 10 feb.

Der har været afholdt 3 halvdagsseminarer i 2017:

23. maj. Tema: 12. semester i studieplan Medicin 2015

28. september. Tema: Skriftlige vurderingsformer

21. november, Tema: Frafald fra medicinstudiet; mentorordning; læringsmiljø

Denne rapport indeholder beskrivelse og kommentering af perspektiver vedr. Organisation; Arbejdet med studieplanen; Evaluering; og Generelle betragtninger.

Organisation

Den nuværende organisation vil fortsætte indtil sommeren 2018. Organisationen har fungeret fint indtil nu. Det er planen at projektgruppens arbejde skal overføres til PUM og der skal etableres en arbejdsgruppe herunder for at sikre kontinuitet i arbejdet.

Arbejdet med studieplanen

Arbejdet med studieplanen er forløbet godt. I 2017 blev nyt 5. og 9. semester implementeret. De sidste semesterplaner, for 11. og 12. semester, er blevet leveret til tiden. Udfordringen på disse semestre er integration af mange fag ind mod samme problemstilling og en klinisk sammenhæng. Specielt 12. semester skal ses som arbejdsforberedelse. Det er en udfordring at beskrive læringsmål for integrerede fag. Der er god opmærksomhed på en tendens til at fagene bliver for ambitiøse i fagbeskrivelsen – det er svært at sætte faget ind i helheden. Der har været diskussion om vanskeligheden ved semesteropdeling og forslag om, at 11. og 12. semester skal ses under et og udprøves med brug af mappevurdering. Dette arbejdes der videre med.

Der er lavet plan for de elektive perioder, som skal implementeres i jan 2018. Der efterspørges flere valgfagskurser – også gerne fra andre professionsstudier end medicin.

Der er i 2017 implementeret studenteraktiverende undervisning i form af TBL i flere semestre og der arbejdes med integrering af IT i undervisningen. Der er imidlertid også udfordring med mange forelæsninger og med, at disse ikke har ændret sig i den nye studieplan. Der er forslag om peer-supervision evt. med ældre studerende som middel til at udvikle forelæsningerne.

Der udarbejdes plan for tværprofessionel læring på studiet.

Mandat for akademisk søjle og hovedopgaven slås sammen – og det giver god mening.

Professionssøjlen har ansvar for undervisning i kommunikation, tidlig praksis og praktiske færdigheder. Tidlig praksis – som reelt betyder indlagte praksisperioder af kortere varighed på flere semestre – har være en stor succes og får positive evalueringer. Der har dog været klage over et ophold – og denne klage er der blevet taget godt hånd om. Der arbejdes fortsat nationalt med en liste og definition af færdigheder, som alle i landet skal opnå på medicinstudiet.

I 2017 har der været tiltagende udfordring med implementering af ny studieplan samtidig med overgangsordninger for gammel studieplan, hvilket har medført mange dobbelthold. Dette er krævende både for undervisere og administration.

Kommentar

Mange af de innovative tiltag, som den oprindelige plan indeholdt begynder nu at blive foldet ud og det ser ud til, at de bliver positivt modtaget. De elektive kurser er hver især spændende, men man bør overveje hvorledes de er tænkt koblet til kernecurriculum. Hvad er det overordnede rationale for netop disse tilbud? Er der en overordnet strategi, som de elektive kurser spiller ind i? I det hele taget bliver beskrivelse af rationalet for de enkelte fag og semestre vigtig som pejlemærke til både undervisere og studerende, så man kan bevare det store overblik.

OSCE

I foråret og i efteråret 2017 afholdtes 3. og 4. pilot for OSCE. Antal stationer er gradvist blevet udvidet. Fremover vil OSCE foregå i Medicinsk Færdighedscenter. Gruppen omkring OSCE er styrket med flere medlemmer og der er udnævnt leder af gruppen. Første egentlige OSCE eksamen afholdes for 3. studieår på ny ordning i start 2018.

Kommentar

Det har været en god plan for udvikling og implementering af OSCE-eksamen. Det foreslås, at man i 2018 laver pålidelighedsberegning og system for beslutning om beståelsesgrænse. Sådanne beregninger, med brug af G-theory og D-studies, kan bidrage til beslutning om det nødvendige antal stationer og antal bedømmere. Vejledning kan evt. hentes hos CAMES, Rigshospitalet i København, hvor man har mange års erfaring med OSCE.

Evaluering

Evalueringer af de enkelte kurser har tidligere ligget i institutterne og PUM her ikke være involveret. Men, med den nye studieplan ønsker PUM nu skriftlige tilbagemeldinger på tværs af semestre. Det er ønskeligt med en række fællesspørgsmål. Disse skal være supplement til de evalueringer, der foretages i STUND. Der bliver også mulighed for semesterstyrene, at stille supplerende spørgsmål til specifikke fagelementer. Arbejdet med nye rutiner for evaluering foretages i samarbejde med Enhed for Læring.

Hidtidig praksis viser, at der stilles mellem 30-70 spørgsmål, og at kvaliteten af spørgsmålene er varierende. Der ønskes derfor opbygget en spørgsmålsbank med kvalitetssikrede spørgsmål. Der søges en samordning med det, der foregår i STUND, ud fra betragtningen om, at der blive for mange spørgsmål til de studerende. Tæt dialog mellem STUND og semesterstyrerne pointeres derfor som væsentligt.

Der er også kommet evalueringsdata fra studiebarometret.no. PUM har noteret, at svarprocenten kun var 39, og diskuteret mulige tiltag i forhold til at øge denne.

PUM modtog i april 2016 et notat om 10 perspektiver ved den nye studieplan, som det ville være ønskeligt at evaluere. Det er spørgsmålet, hvor langt man er kommet med disse 10 perspektiver.

Kommentar

Der ses gode overvejelser vedr. evaluering og samordning med andre instanser, der arbejde med samme felt. Det skal også fremhæves, at PUM vil have skriftlige rapporter – det vil have betydning for opfølgning på evalueringerne og læringen fra dem.

30-70 spørgsmål er mange, og det anbefales at man overvejer, at noget af evalueringen kan foregå mundtligt ved dialog mellem underviser og studerende, f.eks. i den sidste uge af undervisningen. Særligt i en tid, hvor man ønsker at stille spørgsmål til den nye studieordning, er det vigtigt, at standardevalueringer ikke tager for meget plads i listen af spørgsmål til de studerende.

Det anbefales, at man overvejer, hvorledes evalueringerne evt. skal bruges til at demonstrere undervisernes meritter. I nogle lande ligger undervisningsevalueringer til grund for forfremmelse eller ansøgning om akademiske stillinger.

Notatet vedr. plan for evaluering fra apr 2016 bør genbesøges, og det skal overvejes, hvad der er praktisk muligt og hvilke dele af planen, som PUM finder mest værdifuld at prioritere. Arbejdsgruppen har været opmærksom på, det store arbejde, der er lagt op til og anbefalet, at arbejdsdelingen bliver mere klar. Med planen for koordinering af STUND undervisningsevaluering i tæt dialog med semesterstyrer og med rapport til PUM ser det ud til, at den del lykkes fint.

Generelle betragtninger

Det er imponerende som planen bliver overholdt, og der er meget god fremdrift i projektet. Implementering af nye kurser, undervisnings- og eksamensformer forløber godt. Der er god understøttelse af implementeringen ved temadage og kurser udbudt af Enhed for Læring. Der er behov for udvikling af standard kvalitetssikringsmekanismer – både for kurser og eksamener – og behov for en revideret evalueringsplan for hele projektet.

Efter planen er alle semestre implementeret i juni 2019 og det første hold, som har fulgt ny studieplan dimitterer i juni 2021. Man skal overveje om man vil publicere plan og procesforløbet som artikel – der har for nylig være indkaldt til særnummer af tidsskriftet Medical Teacher om curriculum revision. Dette kunne være en forløber for en artikel om evaluering af ny studieplan.

Programsensor

Dec 2017



Charlotte Ringsted