

Refleksioner over udarbejdelse af test i Blackboard

Annelise Grinsted

Lektor, ph.d.

Institut for Fagsprog, Kommunikation og Informationsvidenskab

Syddansk Universitet

annelise@sitkom.sdu.dk



Annelise Grinsted er ansat ved Syddansk Universitet (tidligere Handelshøjskole Syd) siden 1987. Hun har varetaget hverv som prorektor, dekan og institutleder.

1. Forhistorie

Vi er 2 medarbejdere / lektorer på Syddansk Universitet - Simon Clement og mig selv (AG) - der har påtaget os at udvikle 3 former for test i læringssystemet Blackboard, afprøve disse og derefter evaluere på resultaterne af denne proces. Vi har begge undervisning i Flensborg med hhv. danske og tyske studerende i forbindelse med det grænseoverskridende samarbejde mellem Syddansk Universitet og Flensburg Universitat.

Projektet blev udviklet pa folgende baggrund:

Det tyske uddannelsessystem er i sammenligning med det danske noget mere skolepraget, end vi kender det fra vores eget universitet - 22-25 konfrontationstimer per uge per semester med dertil svarende mindre forberedelsestid i det enkelte fag. Dette lagger ikke umiddelbart op til en fordybelse i stoffet.

De studerende har samtidig givet udtryk for, at det er vigtigt for dem at blive evalueret lobende med konsekvenser for karakteren, idet de argumenterer, at der ellers ikke er motivation for dem til at forberede sig - endsige fordybe sig.

Vi har indtil nu udprovet vores undervisningsforlob med frie seminaropgaver. I forhold til denne proveform er der stor uenighed blandt de studerende – kontrol af paratviden som individ versus den mere refleksive form, ofte i grupper. Som danske undervisere er vi dog ikke meget for helt at opgive at udprove i form af seminar eller anden fri skriftlig fremstilling, hvor en vis refleksion over stoffet kommer til udtryk. Derfor vil vi gerne eksperimentere med at kombinere paratviden >< refleksion.

Det er dermed vigtigt at papege, at karakteren baserer sig pa flere former for praestation, idet testene ikke kan sta alene som et udtryk for den studerendes tilegnelse af fagets indhold. I AG's forlob baserer den studerendes karakter sig pa 3 former for praestation:

- 1) *Præsentationer*: En gruppe af studerende (forskellig fra gang til gang) laver til hver undervisningstime en præsentation af en for faget relevant tekst. Dette stiller krav til den studerende om at strukturere indholdet og uddybe vanskelige begreber, så emnet bliver endnu mere tilgængeligt for de øvrige studerende. (20 % af karakteren) .
- 2) *Selvtest*: Den enkelte studerende (og underviseren) kan kontrollere, om han / hun har forstået vigtige elementer af og sammenhænge i fagets tekster. (20 % af karakteren).
- 3) *Opgave*: Den studerende eller en gruppe af studerende udvælger et for faget relevant emne og skriver en opgave på baggrund af en godkendt problemformulering. Dette stiller krav om refleksion og kritisk holdning til emnet og dets forskellige tilgange og udlægninger, og afspejler de studerendes evne til at bearbejde stoffet. (60 % af karakteren).

2. Vores overvejelser

Vi inddrager således selvtest i Blackboard som en del af forløbet for at imødekomme de studerendes ønske om "kontrol". Blackboard er det læringssystem, der anvendes i undervisningen på Syddansk Universitet.

Syddansk Universitet er et internationalt orienteret universitet med en stærk regional forankring. Anvendelsen af it understøtter og styrker denne profil. Globaliseringen og den teknologiske udvikling påvirker såvel forskning som undervisning på SDU, og universitetet må derfor berede sig på at kunne opfylde de krav samt udnytte de muligheder, der følger med denne udvikling. It indtager som værktøj og medium en særlig rolle i og efter omstillingen fra industrisamfund til videnssamfund. Det skal sikres, at studerende erhverver de nødvendige kompetencer, således at Universitetet også internationalt til stadighed vil være på højde med udviklingen. Implementeringen af e-læring repræsenterer således en vigtig konkurrenceparameter, der skal være med til at positionere SDU's uddannelser.

For at SDU nationalt såvel som internationalt til stadighed er på højde med udviklingen, er det nødvendigt, at uddannelserne giver kandidaterne kompetencer, som stemmer overens med de krav, der stilles på arbejdsmarkedet. Derfor må universitets undervisnings- og læringsformer understøtte erhvervelsen af sådanne kompetencer.

Dette stiller et krav til Universitetet om, at nye kompetencer hos underviserne udvikles. SDU satser på, at e-læring inddrages i både heltids- og deltids-uddannelser. Der er investeret mange ressourcer i udviklingen af en nødvendig og fremtidssikret infrastruktur. Denne skal vedligeholdes og i nødvendigt omfang udbygges yderligere, men vigtigst er det, at SDU i de kommende år arbejder intensivt med at sikre de pædagogiske kvalifikationer hos underviserne, som anvendelse af e-læring kræver.

Uddrag af Syddansk Universitets E-læringsstrategi 2005

"Udnyttelse af potentialerne i e-læring til understøttelse af undervisning, læring og forskning på Syddansk Universitet"

Udarbejdelse af test i Blackboard betyder, at vi som undervisere ikke skal bruge tiden på at kontrollere den enkelte studerendes løbende indsats. Således kan de studerende løbende kontrollere sig selv og underviseren kan – uden en stor arbejdskrævende indsats i form af gennemretning af opgaver - få feedback til at genoptage emner periodevis, som kan vise sig at have været vanskelige for de studerende. Imidlertid skal man bestemt ikke undervurdere den arbejdsindsats, der

ligger i at udarbejde testene, som dog forhåbentligt kan bruges mere end en gang i en aktualiseret form.

Vi er i gang med - i hvert vores forløb – at udarbejde 3 forskellige former for test til selvkontrol hver tredje uge i løbet af et semester. I AG's forløb vil det være obligatorisk for de studerende at teste sig selv hver tredje uge i en af de tilbudte prøveformer og en forudsætning for adgang til den endelige eksamen, hvorimod det i SC's forløb er et tilbud til de studerende:

- 1) Multiple Choice
- 2) True / False
- 3) Fill in the blank

Der tages udgangspunkt i obligatoriske tekster, undervisningens indhold og sprogligt niveau. På dette tidlige tidspunkt i forløbet, vil jeg – fra mit fag i ”Den Økonomiske Udvikling i Latinamerika” - give et konkret eksempel på de overvejelser / vanskeligheder der er ved at formulere spørgsmålene. Formulering af spørgsmål vil naturligvis komme til at indgå i den endelige evaluering i det omfang, der kan fremskaffes brugbare kommentarer.

Samtidig vil jeg gerne give et bud på, hvad disse test muligvis kan teste. Til dette formål vil jeg tage udgangspunkt i **Blooms Taksonomi**, skønt der er andre mange andre måder at anskue viden på (eksempelvis Lars Qvortrups videnskematik: kvalifikationer, kompetencer, kreativitet og kultur).

- **Viden:** gengivelse af indlært stof.
Nøgleord: viden om fakta, viden om måder og midler til at håndtere fakta, viden om generelle principper og teorier.
- **Forståelse:** anvende det indlærte, men ikke nødvendigvis således, at det indlærte stof forstås til bunds.
Nøgleord: gengive med egne ord, fortolke, ekstrapolation.
- **Anvendelse:** anvendelse af generelle ideer, teorier, principper, procedurer og metoder i konkrete nye situationer.
Nøgleord: forandre, demonstrere.
- **Analyse:** nedbrydning og afdækning af relationen mellem enkeltdele.
Nøgleord: analyse af elementer og principper, skille ad.
- **Syntese:** sammensætning af enkeltdele således, at de udgør en ny helhed.
Nøgleord: kombinere, abstrahere, skabe.
- **Vurdering:** Bedømmelse af et givet materiale. Bedømme ud fra forskellige kriterier.
Nøgleord: kritisere, rose, støtte, konkludere, opsummere.

Til testene har jeg valgt en model, hvor de studerende i hver test får 30 spørgsmål svarende til ca. 10 spørgsmål til hver tekst (én tekst fra hver uge testen dækker). Testen skal bestå med 50 % korrekte svar, svarende til et 6-tal.

Tekst eksempel på spansk (med efterfølgende dansk oversættelse for at kunne forstå pointerne):

Spansk:

Hay que reconocer de entrada que el impulso de la modernización económica de América Latina tuvo un origen externo. El crecimiento de las economías europea y estadounidense provocó una rápida demanda de materias primas de las que América Latina se encontraba bien surtida. Los países que fueron capaces de estabilizar con rapidez su vida política fueron los mejor situados para responder a tal demanda. En Brasil, México, Argentina y Chile, el Estado consolidado fue capaz de crear las condiciones propicias al desarrollo económico. Otros países como Venezuela o Guatemala, que atravesaron largos períodos de estabilidad, consecuencia de la dominación de los caudillos, conocieron progresos económicos espectaculares. Inspiraban confianza, de manera que tanto el comercio como las inversiones extranjeras se dirigían allí con entusiasmo. En todos ellos, la estabilidad política y el crecimiento económico se reforzaron mutuamente. El incremento en los ingresos de exportación permitió a los gobiernos aplazar las reformas fiscales, mientras el clima de prosperidad proporcionaba un indiscutible afianzamiento a las clases dirigentes. Un orden oligárquico estable se estaba poniendo en marcha.

Dansk:

Man skal være klar over, at starten på Latinamerikas økonomiske modernisering havde sin oprindelse uden for landets grænser. Væksten i de europæiske og nordamerikanske økonomier skabte en stor efterspørgsel efter råstoffer, som Latinamerika havde et bredt udvalg af. De lande som var i stand til hurtigt at stabilisere deres politiske situation var bedst rustet til at imødekomme en sådan efterspørgsel. I Brasilien, Mexico, Argentina og Chile var en konsolideret stat i stand til at skabe betingelser der fremmede den økonomiske udvikling. I andre lande som Venezuela eller Guatemala, der oplevede lange perioder med stabilitet som følge af statsoverhovedernes (oversætter kommentar: termen "caudillo" refererer til stærke mænd, som fx Spaniens Franco) stærke position, oplevede bemærkelsesværdige økonomiske fremskridt. De indgød tillid i en sådan grad, at såvel handel som udenlandske investeringer med stor entusiasme gik den vej. I alle disse lande forstærkede den politiske stabilitet og den økonomiske vækst hinanden gensidigt. Stigningen i eksportindtægterne gjorde det muligt for regeringerne at udsætte en reform af skattesystemet, mens velstanden skaffede de herskende klasser en indiskutabel støtte. Et oligarkisk system var ved at blive indført.

Jeg har valgt i meget høj grad at genbruge emnerne i samtlige 3 test i de formuleringer, der er nødvendige for den givne testform. Dette har jeg gjort for at kunne imødegå kritik af, at indholdet i en test er sværere end i den anden – de studerende kan frit vælge mellem, hvilken testtype de vil tage som den første, og i det tilfælde, at de ikke består første gang, så den anden og tredje test. Jeg har givet dem den option at tage alle tre test og fastsætte karakteren efter det bedste resultat. Jeg er samtidig klar over, at selve testformen kan give meget forskellige resultater. Jeg har ikke lagt skjul på over for de studerende, at jeg anser "nemhedsgraden" af de 3 test former som følger: 1) Multiple Choice, 2) True / False og 3) Fill in the blank.

Når jeg vurderer "nemhedsgrad" forudsætter jeg, at den studerende har læst og forstået det gennemgåede stof – "nemhedsgrad" handler således ikke om, hvorvidt den studerende kan gætte sig frem til det korrekte svar. Jeg baserer min vurdering på, at i Multiple Choice har den studerende

rent faktisk det korrekte svar foran sig, hvorimod i True / False skal den studerende ind og vurdere et givent udsagn, og i Fill in the Blank skal den studerende selv formulere sit svar. Det skal så i sidste ende vurderes, om antagelse om ”nemhedsgrad” er korrekt.

Hvordan kan der stilles spørgsmål til indholdet i det ovenstående tekststykke, som rummer en grundlæggende viden for forståelse af den fremtidige udvikling i de latinamerikanske økonomier?

3. Multiple Choice

Spørgsmål:

I slutningen af det 19. århundrede og i begyndelsen af det 20. århundrede forstærkedes den økonomiske vækst i de latinamerikanske lande af:

1. eksportindtægter
2. den politiske stabilitet (X)
3. statsoverhovedernes stærke position
4. tillid til fremtiden

Overvejelser:

1) Valgmuligheder: Jeg har valgt 4 svarforslag af følgende grunde:

- a. 2 svarforlag bliver til True / False
- b. 3 svarforslag bliver til 1 rigtigt, 1 mere eller mindre forkert og 1 til overvejelse
- c. 4 svarforslag bliver til 1 rigtigt, 1 mere eller mindre forkert og 2 til overvejelse

hvilket sidste stiller lidt større krav til den studerendes viden.

2) Sværhedsgrad: Svarmulighederne skal formuleres således, at der er ét entydigt rigtigt svar, men de øvrige muligheder (kaldet ”distractors”) må ikke være af en sådan karakter, at de omgående kan fravælges pga. deres manglende sammenhæng med spørgsmålet.

1. eksportindtægter (er ikke helt hen i skoven, i den forstand at det er en faktor, der hører til den økonomiske vækst, der nævnes i selve spørgsmålet)
2. den politiske stabilitet (det korrekte svar)
3. statsoverhovedernes stærke position (er ikke helt hen i skoven, i den forstand at det er en medvirkende / underliggende faktor til den politiske stabilitet)
4. tillid til fremtiden (er ikke helt hen i skoven, men nok det mest forkerte svar, idet det er en implicit faktor)

3) Tid: Da svarmulighederne er relativt korte, får den studerende - i gennemsnit - 1 minut til hvert af spørgsmålene. Da nogle af spørgsmålene er lettere end andre og derfor kan besvares hurtigt (såfremt den studerende kan sit stof, naturligvis), vil der være lidt mere tid til de lidt vanskeligere spørgsmål.

- 4) Rækkefølge: Der ligger en overvejelse i rytmen af korrekte svar – i hvert fjerde spørgsmål er det korrekte svar nr. 4, 10 rigtige svar i rap er nr. 2, etc.

Styrker og svagheder i Multiple Choice Test

Styrker:

- Tester et bredt udvalg af emner på relativ kort tid.
- Giver en pålidelig og objektiv karaktergivning i forhold til de forudbestemte svar.
- Kan rettes hurtigt af en computer.
- Rettelser via en computer muliggør en analyse af de stillede spørgsmål med henblik på vanskelige emner for de studerende.
- Resultaterne kan bruges til en opfølgning efter hver test.
- Der kan opbygges en stor pulje af spørgsmål, som på sigt kan reducere forberedelsestiden for underviseren.
- Testresultatet påvirkes ikke af den studerendes evne til at skrive.

Svagheder:

- Det tager lang tid at udarbejde plausible (ikke helt forkerte) svarforslag.
- Der er restriktioner på den studerendes svar, idet han / hun skal vælge blandt underviserens svarforslag.
- Tager ikke højde for den enkelte studerendes måde at indlære på.
- Gætværk kan være et resultat (men med plausible svarforslag ("distractors") vil resultatet være intelligent gætværk).
- Det samme spørgsmål kan gentages i forskellige formuleringer, hvilket kræver en skarp kontrol af formuleringen i forhold til det rigtige svar.
- Spørgsmål bør testes forud og revideres grundigt for at sikre spørgsmålenes entydighed.

4. True / False

Spørgsmål:

Den politiske ustabilitet i slutningen af det 19. århundrede og i begyndelsen af det 20. århundrede og den økonomiske vækst forstærkede hinanden gensidigt. Stigningen i eksportindtægterne gjorde det muligt for regeringerne at udsætte en reform af skattesystemet, mens velstanden skaffede de herskende klasser en indiskutabel støtte. Et oligarkisk system var ved at blive indført.

True

False (X)

Overvejelser:

- 1) Entydighed: Når man formulerer en udtalelse, som kan være sand eller falsk, virker det så uendeligt logisk, hvorfor man har en tendens til at overveje, hvordan man kan "sløre" det tilsyneladende iøjnespringende logiske svar. Dette kan forvirre den studerende.
- 2) Antydning: Sætningsstrukturen i den oprindelige tekst af *den politiske stabilitet og den økonomiske vækst* er ændret i denne udtalelse til *den politiske ustabilitet*. Samtidig for-

søger jeg ved den udvidede udtalelse at give en antydning af det korrekte svar gennem følgende oplysning: ... *mens velstanden skaffede de herskende klasser en indiskutabel støtte*... Har den studerende læst og forstået teksten, vil han / hun hurtigt kunne regne sig frem til det korrekte svar.

- 3) Tid: Da der kun er to mulige svar, får den studerende - i gennemsnit - 1 minut til hver af spørgsmålene. Da nogle af spørgsmålene er lettere end andre og derfor kan besvares hurtigt (såfremt den studerende kan sit stof, naturligvis), vil der være lidt mere tid til de lidt vanskeligere spørgsmål.
- 4) Rækkefølge: Der ligger en overvejelse i rytmen af spørgsmål med True og False – hver anden, 2 rigtige 2 forkert, en kombination, 10 rigtige i rap, etc.

Styrker og svagheder i True / False Test

Styrker:

- Tester et bredt udvalg af emner på meget kort tid.
- Giver en pålidelig og objektiv karaktergivning i forhold til de forudbestemte svar.
- Kan rettes hurtigt af en computer.
- Rettelser via en computer muliggør en analyse af de stillede spørgsmål med henblik på vanskelige emner for de studerende.
- Resultaterne kan bruges til en opfølgning efter hver test.
- Der kan opbygges en stor pulje af spørgsmål, som på sigt kan reducere forberedelsestiden for underviseren.
- Testresultatet påvirkes ikke af den studerendes evne til at skrive.

Svagheder:

- Det tager lang tid at udarbejde meningsfyldte spørgsmål.
- Underviseren kan forfalde til at stille meget trivielle spørgsmål.
- Tager ikke højde for den enkelte studerendes måde at indlære på.
- Lægger op til en del gætværk.
- Udtalelserne bør testes forud og revideres grundigt for at sikre spørgsmålenes entydighed.

5. Fill in the Blank

Spørgsmål:

I slutningen af det 19. århundrede og i begyndelsen af det 20. århundrede var de lande, som var i stand til hurtigt at stabilisere deres politiske situation, bedst rustet til at imødekomme en efterspørgsel efter råvarer. De indgød tillid i en sådan grad, at såvel handel som udenlandske investeringer med stor entusiasme gik den vej. Den politiske _____ og den økonomiske vækst forstærkede hinanden gensidigt.

stabilitet

Overvejelser:

- 1) Vidensområde: I denne form for test er det 2 vidensområder, der bliver testet på spansk: den faglige og den sproglige viden.

Den politiske _____ og den økonomiske vækst forstærkede hinanden gensidigt.
(La estabilidad política y el crecimiento económico se reforzaron mutuamente).

Som det fremgår af den spanske tekst er ”estabilidad” et hunkønsord: La _____ política. Når der så skal indsættes et ord, skal der derfor tages højde for kønnet. Dette bliver endnu tydeligere, hvor et adjektiv skal indsættes.

- 2) Entydighed: Det er vigtigt, at der kun er et rigtigt svar. Hvis der er mulighed for mere end et svar, må spørgsmålet forkastes, med mindre den grammatiske indgang kan afgøre hvilket ord, der er det korrekte.
- 3) Antydning: I dette eksempel er der ved den udvidede udtalelse en antydning af det korrekte ord gennem følgende oplysning: ... *var de lande som var i stand til hurtigt at stabilisere deres politiske situation...* På den måde bliver det ikke det rene gætværk, idet en indsætningsøvelse kan åbne op for mange mulige fortolkninger.
- 4) Tid: Da der skal laves en analyse af teksten, som i gennemsnit er længere end de spørgsmål, der stilles i de andre testformer, for at kunne indsætte det rigtige ord, får den studerende - i gennemsnit – lidt længere tid end i de to andre testformer, nemlig 1½ minut til hvert af spørgsmålene. Da nogle af spørgsmålene er lettere end andre og derfor kan besvares hurtigt (såfremt den studerende kan sit stof, naturligvis), vil der være lidt mere tid til de lidt vanskeligere spørgsmål.

Styrker og svagheder i Fill in the Blank Test

Styrker:

- Tester et bredt udvalg af emner på relativ kort tid.
- Giver en pålidelig og objektiv karaktergivning i forhold til de forudbestemte svar.
- Kan rettes hurtigt af en computer.
- Rettelser via en computer muliggør en analyse af de stillede spørgsmål med henblik på vanskelige emner for de studerende.
- Resultaterne kan bruges til en opfølgning efter hver test.
- Der kan opbygges en stor bank af spørgsmål, som på sigt kan reducere forberedelsestiden for underviseren.

Svagheder:

- Det tager lang tid at udarbejde meningsfyldte udtalelser, der tester en relevant viden.
- Underviseren kan forfalde til at stille meget trivielle spørgsmål.
- Vanskeligt at udarbejde udtalelser, hvor der kun er ét eneste korrekt svar (spansk er et ordrigt sprog).
- Tager ikke højde for den enkelte studerendes måde at indlære på.

- Gætværk kan være et resultat (men i langt mindre omfang end True / False test).
- Udtalelserne bør testes forud og revideres grundigt for at sikre spørgsmålenes entydighed.

6. Overvejelser fælles for alle tre testformer:

- 1) Ramme: I et fag af denne karakter, hvor der arbejdes med en periode på mere end 100 år, vil det være nødvendigt at indsætte den periode, spørgsmålet handler om – det giver en ramme og samtidig en oplysning til den studerende.
- 2) Formulering:
 - Det er svært helt at undgå, at lærerens idiosynkrasier når ind i formuleringen af spørgsmål og svarforslag.
 - Det forudsættes, at formuleringen er så entydig, at spørgsmål eller udtalelser kun kan fortolkes således, som underviseren har tænkt / forestillet sig det.

3) Blooms Taksonomi:

Denne form for test vil kunne teste:

gengivelse af indlært stof (a),
forståelse af det indlærte (b).

Men vil vanskeligere kunne teste:

nedbrydning og afdækning af relationen mellem enkeltdele (d),
sammensætning af enkeltdele således, at de udgør en ny helhed (e).

Men vil ikke kunne teste:

anvendelse af generelle ideer, teorier, principper, procedurer og metoder i konkrete nye situationer (c),
bedømmelse af et givet materiale, bedømmelse ud fra forskellige kriterier (f).

Dette er de første overvejelser i forbindelse med udarbejdelse af testene. Vi har ikke i skrivende stund opnået konkrete resultater, da semestret i Flensborg løber fra slutningen af oktober 2005 til slutningen af februar 2006.

Referencer

Artwork, True-False Test Items:

<http://artwork.asu.edu/arts/teachers/assessment/forced2.htm>

Ascilite: Online Assessment

<http://www.ascilite.org.au/newsletter/3.02/>

Blooms Taksonomi:

http://it.bt.aau.dk/it/education/reports/blooms_taksonomi.pdf

Centre for Information and Computer Sciences, Computer Assisted Assessment Review

Exercises: <http://www.ics.ltsn.ac.uk/resources/assessment/caa/reviews/>

Designing and Managing MCQs:

<http://web.uct.ac.za/projects/cbe/mcqman/mcqappc.html>
E-læringspædagogik
http://217.61.111.32/Elring_for_undervisere/Study/Main/ses02/Content_aim.htm
Instructional Evaluation: Summative Strategies
<http://teach.lanec.edu/fpd/resources/LENS/module6.pdf>
Multiple Choice Tests:
http://www.flaguide.org/cat/multiplechoicetest/multiple_choice_test6.php
Murdoch University Teaching and Learning Center, Multiple Choice Tests:
<http://www.tlc.murdoch.edu.au/eddev/evaluation/mcq/mctests.html>
Syddansk Universitets E-læringsstrategi: Integration af it i undervisning, læring og forskning på
Syddansk Universitet:
<http://intern.sdu.dk/enheder/elearn/vedtagelser/strategi2004/elearnpdf/>
The Joint Information Systems Committee:
http://www.jisc.ac.uk/index.cfm?name=mle_briefings_1