

Kvalitet og retfærdighed

ALTE's Code of Practice



Nick Saville

University of Cambridge.
Local Examinations Syndicate, UK.

ALTE - *Association of Language Testers in Europe* (Foreningen af Sprogtestere i Europa) - blev dannet i 1990 da repræsentanter for otte institutioner deltog i et stiftende møde i Barcelona. Alle stiftende Medlemmer var udbydere af eksaminer og arbejdede med evaluering af deres egne sprog som fremmedsprog. På dette første møde blev det klart at alle havde en fælles interesse i høje standarder for sprogevaluering og for retfærdig behandling af de kandidater der tog deres eksaminer. Denne fælles interesse har været en samlende faktor, og har skabt et fokus for samarbejde i det forløbne årti. Så tidligt som i 1991 blev det besluttet at foreningens hovedmål skulle være:

„... at skabe professionelle standarder for alle trin i sprogtest-procesen“,
med det formål at give sprogtilegnere

adgang til eksaminer af høj kvalitet, som kunne hjælpe dem med at nå deres mål i livet, og som kunne have en positiv effekt på deres læringsoplevelser.

I de senere år er ALTE's arbejde blevet kendt i Europa, og medlemstallet er steget støt. Der er nu 27 institutionsmedlemmer og -tilknyttede, som repræsenterer 24 europæiske sprog.

Medlemmerne omfatter nu: Undervisningsministerier og andre ministerielle afdelinger, eksamenskommissioner, kulturinstanser, universitetsafdelinger og sprogundervisningsinstitutioner. I nogle tilfælde er der blevet dannet sprogtest-konsortier inden for rammerne af ALTE, f.eks. i Danmark, hvor Undervisningsministeriet samarbejder med Københavns Universitet og sprogcentre.

Inden for ALTE besluttede de stiftende Medlemmer at det var vigtigt for både eksamensudviklere og eksamensbrugere at følge en fastlagt *code of practice* (kodeks for praksis) som kunne sikre at evalueringsprocedurerne var af en høj kvalitet, og at alle interessenter blev retfærdigt behandlet. En sådan *code of practice* tager udgangspunkt i sunde principper for god evalueringspraksis, som gør det muligt at opnå høje standarder for kvalitet og retfærdighed. Diskussionen om hvad god praksis inden for ALTE er, viser at man er optaget af at der skal kunne redegøres for alle områder af den evaluering som ALTE's Medlemmer udfører. I denne henseende lægger foreningen vægt på betydningen af validering og den rolle som forskning og udvikling spiller for evalueringsprocesser.

Med publiceringen af sin første *Code of Practice* i 1994 opstillede ALTE de standarder som Medlemmerne havde sat som mål for deres sprogtests. Den trak på *The Code of Fair Testing Practices in Education* (Kodeks for Retfærdig Testpraksis inden for Uddannelser) udviklet af *Washington D.C. Joint Committee on Testing Practices* (1988), og hensigten var at der skulle være en bred formulering af hvad brugerne af ALTE-eksaminer kunne forvente, og hvilke roller og hvilket ansvar interessenterne havde i forsøgene på at opnå retfærdighed.

Brugergrupper

Code of Practice identificerer tre hovedgrupper af interessenter i test-processen:

- Opgavestillerne, (dvs. eksamens-

komitéer og andre institutioner som er Medlem af ALTE)

- De primære brugere - de kandidater der tager eksamen efter eget valg, gennem henvisning, eller af nødvendighed
- De sekundære brugere, som er kandidaternes sponsorer, eller andre som har en interesse i at kandidaterne tager eksamen (f.eks. forældre, lærere, arbejdsgivere, ansvarlige for optagelse osv.).

Desuden udstikker *Code of Practice* disse interessentgruppers ansvar inden for fire brede områder:

- Udvikling af eksaminer
- Tolkning af eksamensresultater
- Bestræbelse på retfærdighed
- Informering af eksaminander

Det er et vigtigt træk at *Code of Practice* understreger interessenternes fælles ansvar og fokuserer på eksamensbrugernes såvel som eksamensudviklernes forpligtelser i bestræbelserne på at opnå retfærdighed.

I Nordamerika har der inden for psykologiske og uddannelsesmæssige evalueringer været en lang tradition for at opstille standarder, således som det afspejles i 1999-udgaven af *Standards for Educational and Psychological Testing* (Standarder for uddannelsesmæssig og psykologisk testning (AERA / APA / NCME)). Tidligere udgaver af disse standarder har eksisteret siden 1950'erne, og har haft indflydelse på udvikling og brug af sprogpøver, især i USA. Der bliver jævnligt henvist til dem i litteraturen om sprogpøver (se f.eks. Bachman 1990 og Bachman & Palmer 1996).

Især inden for sprogevalueringssområdet har der været en stigende interesse for at fastlægge standarder, herunder for udviklingen af *codes of practice*, og en interesse for etiske spørgsmål. ILTA - *International Language Testing Association* - gennemførte en undersøgelse af internationale prøvestandarder (1995) og udgav i 2000 sin *Code of Ethics* (2000). Dette dokument præsenterer en række principper med kommentarer som „trækker på moralfilosofi og tjener som en ledetråd for god professionel adfærd.“

Retfærdighed?

I 2000 blev en publikation med oplæg, redigeret af Antony Kunnan, udgivet i serien *Cambridge Studies in Language Testing Series* baseret på den *19th Language Testing Research Colloquium* (LTRC Orlando 1997). Denne publikation med titlen *Fairness and validation in language assessment* fokuserer på retfærdighed inden for to områder: *Section One - Concept and Context* og *Section Two - Standards, Criteria and Bias*. Kunnan selv har udviklet en 'retfærdighedsmodel' som søger at integrere de traditionelle begreber validitet og reliabilitet med nyere begreber som fravær af skævvægtning (*absence of bias*), retfærdig adgang (*fair access*) og retfærdighed i administreringen (*fair administration*).

Mens ALTE's *Code of Practice* opridser sine generelle principper i 18 erklæringer¹, giver den kun ringe praktisk vejledning til dem der skal bruge dem, om hvordan disse principper skal implementeres, eller hvordan der kan opstilles standarder og garanteres for dem. I et forsøg på at tage fat på dette problem

udviklede ALTE Medlemmerne et supplerende dokument med titlen *Principles of Good Practice for ALTE examinations* (Principper for god praksis ved ALTE eksaminationer). Meningen med dette dokument var at det skulle give en mere detaljeret beskrivelse af hvordan principperne kunne omsættes til sprogprøvepraksisser som ALTE Medlemmerne skulle indføre for at nå deres mål om høje professionelle standarder.

Inspirationen til hvordan man skulle opnå god praksis kom fra forskellige kilder, bl.a. Bachmans arbejde med *usefulness*, Messickas arbejde med validitet samt tidligere udgaver af *APA Standards* (f.eks. 1985). Selv om dokumentet ikke blev udgivet i sin helhed, blev dele af det senere inkluderet i *User's Guide for Examiners* (1997), udgivet af ALTE på Europarådets vegne (dette dokument har fået navneændring til *Language Examining and Test Development* og er under udgivelse af Europarådet som tillæg til *Common European Framework*, der skulle markere Det Europæiske Sprogår 2001).

I slutningen af 1990'erne havde en række andre projekter høj prioritet for ALTE, herunder arbejdet med ALTEs *Can Do Project*, der har udgjort det empiriske grundlag for tilknytningen af ALTE niveauerne til *Common European Framework*. Men i 1999/2000 rettede ALTE igen sin opmærksomhed mod *Code of Practice* og genetablerede en arbejdsgruppe der skulle bringe projektet videre.

I 2001 mødtes Arbejdsgruppen tre gange. Indtil nu er der kommet to vægtige

resultater ud af det, som blev diskuteret af den samlede Medlemskreds på mødet i Budapest i november:

- En revideret udgave af *Principles of Good Practice for ALTE Examinations* (2001) er blevet udviklet og har cirkuleret i Medlemskredsen,
- *Code of Practice* er blevet omformet og udvidet som en tjekliste der bliver anvendt af Medlemskredsen som en del af et *Quality Management System* (QMS).

Den reviderede version af *Principles of Good Practice for ALTE Examinations* (2001) er baseret på tidligere udgaver, men er blevet ajourført og bearbejdet på flere punkter. Den beskæftiger sig mere detaljeret med de centrale spørgsmål om validitet og reliabilitet og ser på disse spørgsmål i relation til den påvirkning (*impact*) eksaminationer har på den enkelte og samfundet. Denne version trækker ligesom de tidligere udkast på de reviderede *AERA/APA/NCME Standards* (1999) - især i afsnittene om validitet og reliabilitet - såvel som på Bachman og Palmers arbejde, 1996.

Kvalitetsvurdering

Den nye dimension der er blevet indført i arbejdet på dette trin, er koncepterne *Total Quality Management* (TQM) og *Quality Management Systems* (QMS).

QM-systemer er i stort omfang blevet taget i brug både inden for produktions- og serviceområdet med henblik på at forbedre kvaliteten så 'kundens krav' kan imødekommes. Den samme tilgang er på det seneste blevet taget i brug i mange uddannelsesmæssige sammen-

hænge over hele Europa, herunder inden for nationale uddannelsessystemer og af internationale NGO'er. *Customer focus* er et fælles tema både i forretnings- og uddannelsesmæssige sammenhænge, og det samme gælder ALTE som eksamensudbyder. ALTE Medlemskarens største 'kunder' er de kandidater der tager eksaminer (de primære brugere) og som kræver evalueringssystemer af høj kvalitet for at få en retfærdig behandling. Ved at anvende QM-tilgangen håber man at målsætningen om 'fortsat forbedring' vil bidrage til at sikre at standarderne fortsat højnes, og at 'kundernes' krav til eksaminerne fortsat imødekommes.

Anvendelsen af en QM-tilgang har ført til en bearbejdning af *Code of Practice* med henblik på at lade den afspejle de praktiske aspekter af evalueringsarbejdet inden for ALTE's Medlemskreds. Målet er at fastsætte brugbare procedurer og programmer for forbedring som i sidste ende vil kunne anvendes til at garantere minimumskvalitetsstandarder baseret på principperne i *Code of Practice*.

Anvendelsen af *Quality Management* er baseret på følgende nøglebegreber, som er taget fra litteraturen om QM:

- organisationen
- selvevaluering og gensidig evaluering blandt elever / studerende på samme niveau (*Peer Monitoring*)

I denne tilgang er det vigtigt at identificere de roller og det ansvar som nøgleinteressenterne i ALTE-institutionerne har, og at anvende systemet på en fleksibel måde i overensstemmelse med hver organisations særtræk (dvs. de

forskellige ALTE-Medlemmer og deres interessent-grupper). I projektets nuværende fase har Arbejdsgruppen fokuseret på de forskellige organisatoriske faktorer inden for ALTE-Medlemskredsen og på den række af forskelligheder der er. I bestræbelserne på at fastsætte standarder er det ikke målet at få alle ALTE-Medlemmer til at tilpasse sig den samme evalueringsmodel for alle de 24 sprog der er repræsenteret, og det er vigtigt at anerkende de forskellige sproglige, uddannelsesmæssige og kulturelle kontekster inden for hvilke eksaminerne bliver udviklet og anvendt. Det er nødvendigt med en passende balance mellem behovet for at garantere professionelle standarder over for brugerne og behovet for at tage højde for ALTE-organisationernes forskelligartede organisatoriske træk, og de sammenhænge hvori deres eksaminer anvendes.

Medlemmernes selvevaluering

Som det næste skridt anbefaler Arbejdsgruppen at alle Medlemmer identificerer deres nuværende stærke sider og de områder inden for deres egen organisation der trænger til øjeblikkelige forbedringer. På denne baggrund vil det så være muligt at opnå de ønskede resultater af kort- og langsigtede udviklinger. Målet bør være:

- at fastsætte et minimum for acceptable standarder
- at etablere *best practice* modeller
- at have fortløbende udvikling som mål (at bevæge sig mod *best practice*)

Der er sandsynligvis ikke noget ALTE-Medlem som vil opnå *best practice* inden for alle områder, og det vil altid

være muligt at lave forbedringer. Målet for alle Medlemmer bør være at fortsætte med at udveksle ekspertise og gradvist hæve standarden over tid, dvs. at stræbe efter *best practice* modeller gennem en løbende udviklingsproces.

I et QM-system af denne slags bliver standarder ikke pålagt 'udefra', men iværksat gennem systemets måder af fungere på, og de procedurer der anvendes til at holde øje med standarder, er fra begyndelsen af baseret på bevidsthedsskabelse og selv-evaluering. Overvågning fra eksterne ligestillede kommer ind på et senere trin for at bekræfte at man lever op til minimumsstandarderne. I sin nuværende form er *Code of Practice* blevet bearbejdet så den kan fungere som et bevidsthedsskabende redskab i denne fase i projektet. Det bearbejdede format afspejler nu de fire aspekter i test-udviklingscyklussen (*test development cycle*), som alle ALTE-Medlemmer er bekendt med:

- Eksamensudvikling
- Administration af eksaminerne (herunder retning, karaktergivning og information om resultaterne)
- Analyse og efter-eksaminationsopsamling

Som opfølgning på mødet i Budapest i november 2001 blev det besluttet at alle Medlemmer skulle færdiggøre de procedurer der skulle indgå i de reviderede *Code of Practice* tjeklister, og sende dem til Arbejdsgruppen. I 2002 vil Arbejdsgruppen lave yderligere analyser og rapportere tilbage til Medlemskredsen i forbindelse med deres møder i april og november.

Note

1 ALTE's *Code of Practice* kan ses på:
www.alte.org/about/code/index.cfm

Litteratur

Bachman L.F.: *Fundamental Considerations in Language Testing*. Oxford: OUP, 1990.
Bachman, L. F. and Palmer, A.: *Language Testing in Practice*. Oxford: OUP, 1996.
Code of Ethics. ILTA, 2000.
The **Code of Fair Testing Practices in Education**. Washington D.C. Joint Committee on Testing Practices, 1988.
Code of Practice. ALTE, 1994.
Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge: Council of Europe / CUP, 2001.

Handbook of Language Examinations and Examination Systems. ALTE, 1998.

Kunnan, A.: *Fairness and Validation in Language Assessment*. Cambridge: UCLES / CUP, 2000.

Linn, R.L. ed.: *Educational Measurement*. 3rd edition. New York: ACE / Macmillan, 1989.

Messick, S.A.: Test validity and the ethics of assessment. i: *American Psychologist*, 1980, 35.

Messick, S.A.: *Validity*, i: Linn, op.cit.

Milanovic, M. and Saville, N.: *Principles of Good Practice for ALTE Examinations*. 1991, 1993.

Standards for Educational and Psychological Testing. Washington: AERA / APA / NCME, 1985/99.

Task Force on Testing Standards. ILTA, 1995.

Users Guide for Examiners. ALTE / Council of Europe, 1997.

Oversat af Michael Svendsen Pedersen
