

Referat fra møtet om bokstavkarakterer.

Intensjonen med møtet, som ble holdt på initiativ fra matematisk institutt ved Universitetet i Oslo, var å komme frem til en nasjonal konsensus om bruken av de nye bokstavkarakterene. Dessuten har Det nasjonale realfagsmøtet bedt om at det utvikles fagspesifikke evalueringskriterier for mastergraden. Møtet hadde derfor en naturlig todeling – først ble bokstavkarakterbruken på bachelornivå diskutert og deretter bruken på masternivå.

Tilstede:

Møtet ble ledet av Tom Lindstrøm.

Bachelornivå

ECTS-skalaen er i utgangspunktet ment å være en europeisk konverteringsskala som skal gjøre det enklere å sammenligne karakterer fra forskjellige institusjoner i forskjellige land, men i Norge har vi bestemte å bygge vårt nye bokstavkaraktersystem over ETCS-lesten.

Prinsippet ved ECTS-skalaen er helt og holdent relativt – karakteren bestemmes ved kandidatens rangering blant de andre i kullet - A for dem blant de 10% beste, B for dem blant de 25% beste etc. Da det kun kandidater som har bestått eksamen har en karakter – og karakterens eksistens er selvsagt en nødvendig betingelse for at karaktereren skal kunne konverteres! – angir ikke ECTS-skalaen noen strykgrense.

Vårt nasjonale system - som for bachelorgradens vedkommende er anbefalt av UHR - opererer med en verbal definisjon av karakterene i tillegg til en relativ prosentvis fordelingsdefinisjon. Den relative skalaen er noe mere fuzzy enn ECTS-skalaen – en A for eksempel gis til kandidater blant de 8-12% beste. Karakterene har fått navn: Fremragende, Meget godt, Godt, Nokså godt, Tilstrekkelig og Ikke bestått. Heller ikke i den norske beskrivelsen presiseres strykgrensen annet enn med tautologien *en prestasjon som ikke tilfredsstillir minimumskravene*.

Ved NTNU er allerede flere interpretasjoner av det nye karaktersystemet i bruk. De har flere konverteringstabeller som oversetter gamle tallkarakterer til bokstavkarakterer i bruk, og de har en egen verbaldefinisjon av bokstavkarakterene som avviker noe fra den som UHR anbefaler. Dessuten har de en scoreskala: Det gis poeng på hver oppgave, normert slik at total score er 100 poeng, og bokstavkarakterene tilordnes

intervaller på scoreskalaen; en score mellom 100 og 85, for eksempel belønnes med en A. Magnus Landstad og Harald Hanche-Olsen redegjorde. De viste også statistikk fra NTNU, i de siste årene har E og F vært dominerende karakterer (på MAT1 i høst var det 32% E og 31% F).

100 - 85 - 75 - 65 - 55 - 35

100 - 85 - 70 - 60 - 50 - 40

Det er flere problemer med en relativ skala à la ETCS som ble fremhevet. For det første en manglende strykdefinisjon, og for det andre spørsmålet om hvilken populasjon man skal legge til grunn i den relative plasseringen. Dessuten er det en iboende spenning i et slikt karaktersystem, det skal både fungere relativt og absolutt.

I de store grunnkursene – iallfall hvis man ser dem nasjonalt og over noen tid – vil kunne fungere som egne referansepopulasjoner, mens kurser lengre oppe i systemet blir for små selv om man ser dem over tid. Det var også mange som mente at studentenes prestasjoner på en naturlig måte blir bedre jo lenger opp i systemet de kommer, og at dette bør reflekteres i karakterene. Synspunktene i debatten spente mellom å burde praktisere relativsystemet strengt - altså en relativ plassering innen hvert kurs - og å burde betrakte hele studentmassen (innen matematikk) på lavere grad som referansepopulasjon. Etter hvert ble det konsensus om det siste synspunktet.

Det ble enighet om å bruke en scoreskala for alle kurs på bachelornivå, og det var en del diskusjon om hvor grensene mellom de forskjellige karakterene skulle gå. TL viste frem statistikk fra begynnerkursene i Oslo gjennom de siste 10 år. Den viser en ganske jevn fordeling med en hvis stigende frekvens mot dårligere karakterer, men fordelingen ganske fjern en normalfordeling, og viser heller ingen "kamelfordeling" slik mange kanskje forventet. Tilsvarende statistikker ble vist fra NTNU og Bergen, og begge bekreftet fordelingen. Dette tilsier at intervallene på scoreskalaen ikke bør være uniforme, men heller smale i endene og brede på midten. En intervallinndeling som passer svært godt med statistikken er:

100 - 92 - 78 - 62 - 49 -40

Det var generell enighet om at matematikk miljøene i Norge må prøve å etablere en mest mulig felles praksis for hvordan de nye karakterene skal brukes.

Etter noe debatt ble det enighet om at vi må forlange av kandidatene at de har 40% eller mere av oppgaven korrekt for å bestå. I praksis vil man alltid skjele til strykprosenten og det vil fortsatt bestå et utstrakt amnesti for dem nær strykgrensen, men møtet mente det prinsipielt er viktig å holde på 40 av 100 for å stå – særlig i disse når-pengene-i-kassen-klinger-tider.

Det ble presisert at det selvsagt ikke er meningen at en skala som ovenfor skal brukes med prøyssisk absolutisme, men at karaktersetting alltid må gjøres med skjønn og lempe, og at grensen derfor i praksis ved hver enkelt sensur vil justeres.

Det var også enighet at man, om ikke løpende, men iallfall om noen år skulle komme tilbake til skalaen, analysere hvorledes den har fungert og eventuelt revurdere den - at man så og si har en gjensidig rapporteringsplikt.

Mastergrad.

Etter lunch diskuterte møtet karaktergivingen på masternivået. Tom Lindstrøm innledet og blant annet viste ham frem statistikk over karakterene som er gitt på hovedfagsoppgavene i ren matematikk ved UiO i løpet av de siste 10 år. Den viste en overraskende stor andel av svært gode karakterer, se følgende tabell.

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1-2.5	2.6-3.	3.1-3.	3.6-4.0
											0	5	
2.3	12.9	18.0	13.5	7.3	10.7	6.2	7.3	7.9	2.8	7.3	2.9	0	0.6
2.3	15.2	33.2	46.7	54.0	64.7	70.9	78.2	86.1	88.9	96.4	99.4	99.4	100

Tilsvarende statistikker fra NTNU og UiB bekrefter disse dataene. Vi ser at ca 15% av hovedfagskandidatene fikk 1.2 eller 1.1 og at hele 33% fikk 1.3 eller bedre og at ca 47% fikk en bedre karakter enn 1.4.

Hvis vi vil bruke bokstavkarakterene må en måte som er noe i nærheten av det intensjonene er, betyr dette at grensen mellom en A og en B bør gå på 1.2, eller sagt annerledes at vi reserverer karakter A for oppgaver som i gammel setting ville fått 1.1 eller 1.2, mens karakter B benyttes for oppgaver som på gamlemetoden ville blitt verdsatt til en karakter mellom 1.3 og 1.4. Halvparten av kandidatene vil da få karakteren A eller B. Karakteren C kan brukes mellom 1.5 og 2.0, mens D for eksempel går fra 2.0 til 2.5, og E derifra og ut.

Diskusjonen på møtet dreide seg stort sett om hvor grensen mellom A og B burde gå, det var svært lite snakk om grensen mellom B og C, men slik referenten oppfattet stemningen, kan en grense på 1.4 muligens være kontroversiell. De fremmøtte fra Oslo og Bergen var stort sett på linje og støttet den strenge praksisen som reserverer A for gamlekarakterene 1.1 og 1.2. De andre var noe reservert, men under visse forutsetninger, som vi kommer tilbake til nedenfor, kunne de også støtte en slik praksis.

Argumentene for en streng praksis var flere. Vi reviderer systemet og anledningen bør brukes til å utvide den effektive karakterskalaen. Med for store A-intervaller og B-intervaller, vil i praksis - iallfall blant den mest ambisøse delen av studenter - A og B bli de eneste karakterene (70% med en B-grense på 1.7 og nesten 80% med 1.8) noe som lett vil føre til at det ikke blir en seier å få A, men det blir et nederlag å få B. En slik situasjon vil vi unngå (en lignende situasjon på bergverkslinjen ved NTH ble beskrevet – A-gutta var de som kom inn på petroleumslinjen, mens B-gjengen måtte nøye seg med gruver – og det førte til ganske dårlig klima blant studentene). Det er altså ønskelig å benytte alle karakterene A, B, C og D, og for å få plass på den effektive delen av skalaen, blir intervallene noenlunde som skissert ovenfor.

Den topptunge karakterfordelingen på oppgaven, tyder vel også på et visst inflatorisk press i karaktersystemet, og man bør benytte anledningen til å revaluere når man innfører et nytt system.

Et argument mot en streng praksis er at det kan føre til at våre kandidater får et handicap på arbeidsmarkedet og at potensielle matergradsstudenter velger et annet fag for å sikre seg en bedre karakter. Forbeholdet til enkelte, særlig ML; som vi nevnte, dreide seg nettopp om dette, og støtten var betinget av at våre nabofag har en tilnærmet lik praksis.

Det var stort sett konsensus om Toms verbale beskrivelser, men muligens formulert uten for sterke referanser til stipendiatstillinger. I den endelige versjonen kan det heller ikke være referanser til tallkarakterer, men i følgebrevet (katekismen!) er det naturlig. Selv om tallkarakterene etterhvert skal forsvinne helt kan vi ikke se bort ifra at de for tiden – iallfall innen matematikkmiljøet – gir den beste beskrivelsen av en prestasjon.

NMR vil ta kontakt med fagrådene i nabofag og i tillegg statistikere, anvendte matematikere, HiA og NLH under utarbeidelse av forslag til fagspesifikke kriterier. Studiedekan TL tar kontakt med de andre realfagsfakultetene.