

Ylioppilastutkinnon kielikokeiden
kehittämisseminaari
Helsingissä 15.-16.11.1979

Raimo Konttinen

Monitasokokeen testiteoreettisista ja -teknisistä ongelmista: luentorunko

1. Johdanto

YTL:n esittämät ongelmat ovat niin vaikeita ja kompleksisia, että monitasokokeiden käytön teknisiä ongelmia on tässä vaiheessa ennenaikaista käsitellä kovin tarkasti. Esitys keskittyy monitasokokeiden käytön edellytyksien pohdintaan.

2. Yhteisten tehtävien käyttö eritasoisissa kokeissa ja arvosanojen verrantaminen.

Tarkoituksena on pohtia, millaisia kysymyksiä joudutaan käsittelemään ennenkuin eri kokeiden verrantamisen yksityiskohtia voidaan ryhtyä suunnittelemaan.

2.1. Erilaisia mahdollisuuksia

(a) yksi vai useampia kielikoesuorituksia (eli todistusmerkintöjä eli "kielitutkintoja")

(1) Yksi ja sama kielikoesuoritusmerkintä eritasoisista (saman kielen) kokeista.

(2) Useampia kielikoesuorituksia, jotka tarvittaessa voidaan siirtää samalle asteikolle (esim. lyhyen L = pitkän kokeen A)

(3) Useita erillisiä kielikoesuorituksia (samaan tapaan kuin pitkä matematiikka, lyhyt matematiikka, reaalikoe jne.)

(b) Yhdet vai useampia tehtäväsarjoja?

(1) Yksi ja sama tehtäväsarja kaikille

(2) Useampia, mutta yhteisiä osia ja tehtäviä sisältäviä sarjoja

(3) Useita erillisiä tehtäväsarjoja (kuten lyhyt ja pitkä matem. koe)

2.2. Eräitä ongelmia liittyen kielikoesuoritusten lukumäärään

(1) Missä määrin seuraavien seikkojen on tarpeen liittyä toisiinsa: kielikoesuoritus (pitkä/lyhyt koe), pakollinen/ylimääräinen, oppimäärä (lyhyt, pitkä tms.) Voiko osallistua heti tai myöhemmin sekä pitkään että lyhyeen kokeeseen? Voiko osallistua oppimäärää vastaamattomaan kokeeseen?

(2) Käykö A-L-asteikko liian ahtaaksi tapauksessa al. Vai olisiko päinvastoin toivottavaa, että esim. pitkän kokeen arvosanat

keskittyisivät välille C - L?

- (3) Vaihtoehtoon a2 liittyviä kysymyksiä: Etuna kai olisi se, että lyhyen oppimäärän arvosanat eivät rajoittuisi asteikon alapäähän. Myös vertailukelpoisuudesta olisi hyötyä, mutta olisi päätettävä, millaisia opinto-oikeuksia eri arvosanat antaisivat. Samoin korkeakoulut tarvitsivat ehkä arvosanojen tulkintaohjeita. Arvosanojen vastaavuuksiin totuttaisiin ehkä hitaasti.
- (4) Mitä vaihtoehdosta a3 seuraisi korkeakoulujen valintoihin. Mitä päätöksiä edellyttäisi opinto-oikeuksista?
- (5) Verrantamiongelmiä ratkaisumahdollisuudet riippuvat myös em. kaltaisten kysymysten ratkaisusta (siis esim. siitä, miten arvosanoja tullaan käyttämään)

2.3. Erilliset vai yhteiset tehtävät? Eräitä mahdollisuuksia.

- (1) Mahdollisuuksia mm. seuraavasti:

	1	2a	2b	2c	3
pitkä	A+P	A+P	A+P	A	A
lyhyt	A+P	B+P	P	P	B

- (2) Huomioon otettavia seikkoja
 - ops:ien erot
 - suoritustasoerot
 - eri oppimääriä opiskelleiden osallistuminen eri kokeisiin

2.4. Ongelmia

- (1) Jos yritetään suoraan tuottaa yhteisiä osioita, syntyvät osiot eivät ehkä ole samassa määrin eri ops:a edustavia tai yhtä keskeisiä niissä. Vaatii ops:iä analysointia (ks. myöh.)
- (2) Mitä pienempiä eri kokeisiin osallistuvien erot, sitä paremmin yhteisiä osioita voidaan käyttää. Periaatteessa yhteisten tehtävien pitäisi olla, paitsi sisällöllisesti, myös vaikeustasoltaan edes jotenkin sopivia eri kokeisiin osallistuville. Ovatko esim. finskan tasoerot liian suuria?
- (3) Jo pelkästään kovin vaikeat (vaikkakaan eivät aivan mahdottomat) tehtävät tuottavat helposti ongelmia (syntyy kiistaa, kuuluvatko oppimäärään eivät ehkä erottele, monivalintaosioissa arvaaminen tällöin todennäköisempää jne.)
- (4) Jos käytetään yhteisellä osalla verrannettuja kokeita, joiden arvosanat ovat vertailukelpoisia, jokin (esim. pitkä) koe ankkuroi myös toisen kokeen asteikon ja vuosien kuluessa voi syntyä ongelmia, jos (todennäköisimmin lyhyen kurssin) taso muuttuu voimakkaasti.

3. Monitasokokeiden käyttömahdollisuuksia

3.1. Verrantaminen

- (1) Tarkoittaa vertailukelpoisten arvosanojen hankkimista, vaikka koeteh-
tävät ovat osittain kokonaan erilaisia.
- (2) Verrantaminen yhteisten tehtävien avulla. Tällöin todistukseen voitai-
siin kirjata myös esim. lyhyen kokeen tulosta vastaava pitkän kokeen
arvosana.
 - Yhteisen osan oltava mielellään koko koetta edustava (sisältäen eri
osakoikeita).
 - Yhteisen osan oltava joidenkin asiantuntijoiden mukaan vähintään 1/5
koko kokeesta tai moniosioisessa kokeessa vähintään 15-20 osion mit-
tainen.
 - Verrantaminen ei ole mielekäästä, jos yhteinen osa ei mittaa samaa kuin
koko koe kaikissa kokeissa (ts. eri kokeiden on oltava periaatteessa
vaihtokelpoisia, mutta vain vaikeustasoltaan ja yksityiskohdiltaan
erilaisia. Empiirinen selvitys.
 - Verrantamiskeinoja on runsaastikin, mutta kutakin sovellusta varten
niitä on kyllä syytä ensin kokeilla. KTL:ssä on käytetty standardoin-
tia ja latentin piirteen malleja (erityisesti Raschin mallia).
Verrantamiskeinoista olisi luultavasti valittava jokin hyvin selkeä
ja helposti ymmärrettävä.
- (3) Olisi syytä harkita myös arvosana-asteikoiden summittaisempaa verran-
tamista niin, että eri kokeiden arvosanojen vastaavuus voitaisiin kyl-
lä ilmoittaa, mutta kukin eritasoisista kokeista arvosteltaisiin
erikseen omalla (mahd. (I)A - L-)asteikollaan. Kukin oppilas saisi
vain yhden ("oman") kokeen arvosanan. Tällöin yhteiselle osalle ei
tarvitsisi asettaa yhtä suuria edustavuusvaatimuksia.
- (4) Olisi ehkä syytä harkita myös eri kerroilla pidettävien kokeiden
verrantamista, jolloin voitaisiin tarkkailla tason muutoksia.
- (5) Summittaisemmassa verrantamisessa eli eri kokeiden ja eri kerroilla
pidettyjen kokeiden tason ja vaativuuden seurannassa voitaisiin käyt-
tää myös oppilasotokselle ennen kirjoituksia esitettyjä yo-kokeen
ulkopuolista koetta.

3.2. Mahdollisuuksia yksilöllisemmän kokeen suuntaan

- (1) Äärimahdollisuutena täysin tailored testing, jolloin oppilas saa
uusii tehtäviä sen mukaan miten hän on menestynyt aiemmissa. Näin saa-
daan reliaabeleita mittaustuloksia huomattavasti tavanomaista ly-
hyemmällä kokeilla. Edellyttää yksilökokeita (esim. tietokoneen

ohjaamana) ja tulevat ehkä kysymykseen vain puhumiskokeissa.

- (2) Hyvin mahdollinen olisi kaksivaiheinen koeohjelma, jossa ensimmäisen kokeen (tai muun mittauksen) tuloksen perusteella annettaisiin toinen (edellistä paremmin oppilaan suoritustasoa vastaava) koe. Alkukokeena voi olla myös oppilaan oma arvio taidoistaan, arvosana, yhteinen (yo-kokeen ulkopuolinen) koe tms.
- (3) Jos se, minkä oppimäärän oppilas on suorittanut ei sido hänen koevalintaansa, kaksivaiheinen koe voisi olla käyttökelpoinen. Kuitenkin kysymykseen tulisi ehkä vain tällainen koejärjestely, ei aito monitasokoe.
- (4) Yleensä monitasokokeiden (esim. kaksivaiheisen testauksen) käyttö edellyttää, että koetehtäväpankki on olemassa, se pysyy suurinpiirtein samana ja kaikista tehtävistä on olemassa tietyt tiedot. On mahdollista muutoinkin: KTL:n menettely. On epäiltävissä, että teknisesti etevimpiä mentettelyitä ei oikein vielä sulatettaisi ja joka tapauksessa niitä pitäisi ensin kokeilla.
- (4) Monitasokokeet teknisessä merkityksessä soveltuisivat ehkä paremmin lukion aikana pidettäviksi kokeiksi.

4. Eräitä muita YTL:n kysymysluetteloön liittyviä näkökohtia

IIA1. Kokeen osuutta voidaan ajatella muutettavaksi kahdella tavalla: Sitä helpotetaan ja/tai sen hajontaa pienennetään. (Nämä voivat vaikuttaa myös - mutta ei välttämättä - toisiinsa). Kuullunymmärtämiskoetta voisi yrittää tehdä helpommaksi tai nimenomaan selkeämmäksi niin, että sen tulokseen eivät missään käytännössä esiin tulevassa tilanteessa vaikuta irrelevantit tekijät ja sellaiset koulu- ja aluekohtaiset erot, joita ei kuitenkaan voida poistaa. Tämä edellyttää puheaines- ja tehtävätaksonomian laatimista eli riittävän yksityiskohtaista analyysiä ja kuvausta siitä, millaiset äänitteet (sisällöltään, tasoltaan jne.) ja millaiset kysymykset tai tehtävät sekä millaiset osiotyypit voivat tulla kysymykseen.

IIA2. Olisiko mahdollista ryhtyä kehittämään osittain opettajan arviointeihin perustuvia kokeita (combined internal - external exams). Erityisesti puhumistaidon mittaamiseen. YTL tutkisi tällöin ainoastaan yleistasoa (ohjeita, opettajan arvosana plus/miinus yksi-menettely, otokseen perustuva koulun tason arviointi tms.).

Voisiko kommunikatiivisuuden ymmärtää laajemminkin? Myös tekstin ymmärtämisen kokeessa voitaisiin pitää silmällä sitä, mitä informaatiota se on välittänyt oppilaalle (ei: mitä jokin sana tai rakenne merkitsee). Myös rakennekoe voi

olla kovin erilainen riippuen siitä, millaiseksi ajatellaan kommunikaatiotilanne. Olisiko tarpeen tehdä siis myös sellainen tehtävätyyppien kuvaus, jossa kuvataan, millaisia kommunikaatiotilanteita tehtävät voivat ja saavat sisältää.

IIB. edellyttää ops:n analyysiä.

III. Osiotyyppeiden pohdinta on tärkeätä. Sitä pitäisi jatkaa tehtävätyyppien, tehtäväaihioiden (item forms) ideointiin ja käsittelyyn (ks. edellä myös IIA2).

Koetulosten yleistettävyydestä tarkastelut olisivat hyvin aiheellisia: Olisivatko tulokset samoja eri osiotyypeillä, eri tehtävätyypeillä, olettaen eri kommunikaatiotilanteita, erilaista puhe- tai tekstiainesta käyttäen, käyttäen toisia kirjoitelman aiheita. Näitä kaikkia olisi syytä tutkia. Välittömästi voitaisiin lisätä tekstien määrää, sillä tekstin luonteen on todettu vaikuttavan tulokseen. Luultavasti sama pätee ja ehkä vielä selveremminkin kirjoitelman aiheisiin. Sama näkyy jopa rakennekokeessa (kokemukset KTL:n ruotsin kielen kokeista).

5. Ratkaisun etsiminen kriteeri- eli sisältöviitteisten kokeiden suunnalta

5.1. Kaksi (ei toisensa poissulkevaa) lähestymistapaa: (1) empiirinen, normeihin perustuva verrantaminen ja (2) sisältöihin (mitattaviin sisältöihin, tai-toihin, tietoihin) perustuva verrantaminen. Suhteellinen ja absoluuttinen arvostelu hieman uudessa muodossa ja käytössä. ~~Tavassa voidaan erottaa seuraavat vaiheet:~~

- (1) Edellinen perustuu paljolti saatuihin tuloksiin: Arvosanat annetaan kokemuksen ohella saatujen jakaumien perusteella ja verrantaminen perustuu yhteisen ankkuriosan tuottamiin tuloksiin.
- (2) Jälkimmäisessä nojaututaan suuremmassa määrin ennalta päätettyihin periaatteisiin: Kokeet laaditaan ops-analyysiin perustuen ja arvosanat annetaan paljolti päätettyjen yleisten periaatteiden varassa. Esim. uudenmuotoinen kielikoe.

5.2. Eräs mahdollisuus: Kaksi ~~ops:iin ankkuroitua~~ sisällöllisesti ja tavoitteiltaan ops:iin ankkuroitua sisältölähtöistä (domain-referenced) koetta, suppea ja ^{laajempi} ~~ylämpi~~, alempi tmv.)

(1) Laadintavaiheet (rinnalla nykyistä muistuttava menettely):

Säädökset	Säädökset
Ops:t	ops:t
Mittausalueen erittely	- - - -
- ops:ien konkretisointi	
- kommunikaatiotilant.	

erittely

- tehtävätyyppien tuottaminen
- osiotyyppien erittely
- mahdol. osioaihioiden (item forms) tuottaminen

Kokeen rakenne (ops:n mukaan)

- mittavat osa-alueet
- niiden painotus
- osakokeiden rakenne

Kokeen rakenne kokemuksen pohjalta

Arvosteluperiaatteet, arvosanojen antaminen

Arvosteluperiaatteet ym.

Kokeiden laadinta em. pohjalta

Asiantuntijat laativat kokeet

Arvosteluperiaatteiden soveltaminen (ja mahd. tulosten perusteella tapahtuva tarkentaminen)

Arvosteluperiaatteet

- (2) Verrantaminen: Kokeet olisivat vertailukelpoisia loogisin perustein, sisällöllisesti ja tavoitteiltaan. Silti edelleenkin olisi tarpeen seurata empiirisin selvityksin tason muutoksia, mutta empiiriset tulokset eivät sanelisi arvosana-asteikkoa tai verrantamista.
- (3) Ratkaisuun sisältyisivät ja olisi pakkokin sisällyttää useimmat YTL:n kysymysten ratkaisuihin (koetyypit, painotus jne.). Ratkaisu-
perusteena eivät nyt olisi pelkästään asiantuntevat mielipiteet, vaan lukion opetus ja ops:t. Erityisesti ongelma IIB saattaisi edellyttää tämän tapaista menettelyä.
- (4) Ratkaisu edellyttää, että mahdollisimman pian laaditaan sen edellyttämä työohjelma ja organisoidaan sen toteutus. Ratkaisu ei ole halpa, mutta sekä yo-tutkinnon että lukion opetuksen kehittämistä ajatellen lähes välttämätön ja myös taloudellinen.
- (5) Esitetyn ratkaisun edellyttämistä toimenpiteistä saa käsityksen tutustumalla KTL:ssä tehtyyn ja KH:n rahoittamaan peruskoulun opetus-
suunnitelmien konkretisoinnin valmisteluun ja koulusaavutuskokeiden laadintaan. Konkretisointityön tulokset, ns. perustavoitemuistiot ovat tulosta koko siitä 70-luvun tutkimuksesta, jota KH on rahoittanut KTL:ssä. Ilman sitä ja pelkästään tutkijoiden ja kouluhallituksen virkamiesten työryhmätyöskentelyn varassa niitä ei voisi ajatellakaan

7.

syntyneiksi. Näihin pohjautuva koulusaavutuskokeiden laadinta edellyttää myös suurta asiantuntemusta, hyvin koulutettuja osioiden laatijoita ja jatkuvaa kokemusten kartuttamista empiirisin selvityksin. KTL:n johdon kanta on se, että mikäli tällaiselle työlle ei luoda riittäviä työskentelyedellytyksiä, on edesvastuutonta antaa tutkijoiden aikaa pois heidän varsinaisista tehtävistään (toimeksiantotutkimuksesta ja laitoksen tutkimuksesta).