



YLIOPPILASTUTKINTOLAUTAKUNTA
STUDENTEXAMENSNÄMNDEN

Selvitys adaptiivisuudesta sähköistetyn ylioppilastutkinnon kontekstissa

Joulukuu 2013

Outi Hakola
Tohtorikoulutettava,
HY / Opettajankoulutuslaitos

Hanna Mäenpää
Tohtorikoulutettava,
HY / Tietojenkäsittelytieteen laitos

Suvelahdenkatu 10 B
PL 50 00581 Helsinki

Söderviksgatan 10 B
PB 50 00581 Helsingfors

Puhelin / Telefon
0295 338 200

Faksi / Fax
(09) 762 274

etunimi.sukunimi@ylioppilastutkinto.fi
förnamn.efternamn@ylioppilastutkinto.fi

www.ylioppilastutkinto.fi
www.studentexamen.fi

Johdanto

Marraskuussa 2013 saimme tehtäväksemme tuottaa selvityksen ylioppilastutkintolautakunnan jäsenille ylioppilaskokeiden adaptiivisuuteen liittyen. Selvityksessä toivottiin käsiteltävän tulevan sähköistetyn ylioppilastutkinnon keskeistä teemaa, arviointiin liittyvää adaptiivisuutta sekä pedagogisesta että teknisestä näkökulmasta. Ylioppilastutkintolautakunnan toiveena oli siis avata adaptiivisuuden monimuotoiseksi koettua käsitettä sekä yleisesti osana (lukio-)opetuksen, opiskelun ja arvioinnin kokonaisuutta että tarkemmin juuri tietokonepohjaiseen testaamiseen liittyen. Tässä selvityksessä tulemmekin alkuun antamaan kuvauksen adaptiivisuuden eli mukautuvuuden erilaisista määritelmistä lähinnä koulukontekstissa, jonka jälkeen siirrymme yksityiskohtaisempaan ja teknisempään kuvailuun ja pohdintaan.

Johdannon jälkeen nostamme esiin ja esittelemme lyhyesti muutamia teemaan liittyviä tutkimuksia, selvityksiä ja kokemuksia adaptiivisesta testaamisesta. Viimeisenä raportin sisällöllisenä osana paneudumme sähköistetyn, tietokonepohjaisen testauksen teknisen toteutuksen perusteisiin, ja toteutuksen haasteisiin ja etuihin. Selvityksen tarkoituksena on antaa tarkennettu kuvaus adaptiivisesta testaamisesta osana päättökoe -tyyppistä koejärjestelmää. Toivomme, että tämä lyhyessä ajassa tuotettu dokumentti auttaisi ylioppilastutkintolautakunnan jäseniä muodostamaan yhteisen ymmärryksen adaptiivisuudesta ja laajentamaan käsitystään adaptiivisen testaamisen eri muodoista hyötyineen ja haasteineen. Varsinaisten järjestelmien määrittelyn sekä ainekohtaisten sovellusten laatimisen ja tarkemman kehittelyn jätämme lautakunnalle. Selvityksen lopussa kokoamme muutamia prosessin aikana syntyneitä ideoita joista testien ja testijärjestelmän laatijat voisivat lähteä omassa työssään liikkeelle, sekä kokoamme selvityksen pääajatuksen.

Koska ylioppilaskokeen sähköistämiseen liittyy laaja tietojärjestelmätyö, on tässä selvityksessä otettu mukaan myös erityinen adaptiivisen järjestelmän toimintaperiaatteeseen, kehitysohjon ja ylläpitoon liittyvä näkökulma. Koska alan asiantuntemusta on Suomessa rajatusti, on työssä kartoitettu lähinnä ulkomaisia tutkimus- ja sovellusprojekteja, joita tarkastelemalla voidaan saada tarkempi kuva erilaisista toteutusvaihtoehdoista.

1 Käsitteen määrittely

Käsitteellä *adaptiivisuus* voidaan tarkoittaa hyvin erilaisia asioita eri konteksteissa. Joka tapauksessa kyse on aina jonkin asian, ominaisuuden tai elementin *mukautuvuudesta* toisiin nähden. Mukautuvuuden kohteena voi olla esimerkiksi jokin tekninen laite tai sen osa, ihmisen käytös tai teoreettinen näkemys. Pedagogisessa, kasvatuksellisessa ja kehityspsykologisessa kontekstissa adaptiivisuuden katsotaan useimmiten liittyvän joko oppijan tai persoonan toimintaan, tai oppijoiden kehityksen/oppimistulosten arviointiin. Erityispedagogiikan ja kehityspsykologian näkökulmasta adaptiivisuus liittyy esimerkiksi henkilön kehitysasteeseen. On olemassa niin kutsuttua adaptiivista käyttäytymistä, jota on käsitelty paljon esimerkiksi down-syndroomaisten nuorten arkipäivän toimintaa kartoittavissa tutkimuksissa.

Tämän selvityksen tavoitteena on kuitenkin keskittyä pedagogiseen arviointikontekstiin, johon liittyy myös opetus- ja arviointiteknologian käyttösovelluksia. Tässä kontekstissa adaptiivisuus käsitetään oppimateriaalien ja testien mukautuvuutena oppijaan nähden. Eritasoista oppimateriaalia voi olla tietokonepohjaisessa järjestelmässä esimerkiksi tuhansia tehtäviä. Oppija voi valita tästä nk. tehtävä- ja tietopankista itselleen sopivan haastavia tekstejä ja tehtäviä, ja edetä opiskelussaan omaan tahtiinsa tasolta toiselle. Kone voi myös syöttää oppijalle tämän sen hetkisen tason mukaisia tehtäviä, ja päästää hänet eteenpäin seuraavalle tasolle vasta kun edelliselle tasolle ohjelmoidut taidot on järjestelmän arvion mukaan saavutettu. Tämän tyyliässä pedagogisessa ratkaisussa opettaja toimii enää opiskelun ja oppimisen ohjaajana tai mahdollistajana, aktiivisen toimijuuden ja oppimisen edistäjän roolin sijaan.

Testi eli ylioppilaskoe voi myös mukautua kokelaan oman tieto- ja taitotason mukaisesti, mikä tarkoittaa että testijärjestelmän koepankkiin on valmiiksi syötetty eritasoisia kysymyspatteristoja ja kysymykset joko vaikeutuvat tai helpottuvat kokeen edetessä sen mukaan, miten kokelas osaa niihin vastata (Kupari 2009). Kokeen loputtua järjestelmä määrittää kokelaan taitotason ja antaa tälle tasoaan vastaavan arvion, joko kirjallisesti tai numeerisesti. Sähköiset testit, erityisesti adaptiiviset testimuodot, voidaan siis kalibroida tietyille taitotasoille opiskelijan mukaan, jolloin ne voivat lineaarisia testimuotoja nopeammin laskea kokelaan tarkan taitotason, eikä kokelaan tarvitse turhaan suorittaa hänelle liian helppoja tai liian vaikeita tehtäviä. Järjestelmä näin ollen laskee koko ajan kokelaan suoriutumista testissä, ja syöttää tälle ainoastaan sellaisia tehtäviä, jotka ovat lähellä edellistä tehtävää, josta kokelas on suoriutunut. Adaptiivinen testaustapa ei myöskään ole rajoittunut pelkkiin monivalintatehtäviin, mikä puolestaan on ominaista lineaarisille testeille. (Scheuermann & Björnsson, 2009). Tätä adaptiivisuuden määrittelyä käytämme myös tässä selvityksessä.

2 Adaptiivinen, tietokonepohjainen testaaminen tutkimuksen ja muutamien maailmalla suoritettujen kokeilujen valossa

Tietokoneella suoritettavan adaptiivisen arvioinnin kokeiluja on tehty useissa eri maissa jo usean vuoden ajan, ja monien lähteiden mukaan pisimmällä kokeiluissaan ovat Islanti, Tanska ja Hollanti. Näissä maissa nk. e-arviointia on toteutettu loppukokeissa sekä perusopetuksen että toisen asteen koulutuksen loppuvaiheessa. Yleisiä huomioita kokeiluista ovat olleet kokeiden keston lyheneminen, yksilöllisten ominaisuuksien huomioon ottamisen mahdollisuus (adaptiivisuus) sekä palautteen ja tulosten saamisen nopeus. Suomessa samanlaista koekäytäntöjen monipuolistamista ovat testanneet vuonna 2007 Länsi-Suomen läänin aikuislukiot opetushallituksen tuella. Kokeiluista on tuotettu myös kaksi raporttia Internet- ja verkkokokeisiin liittyen. Kokeiluissa tietokonetta hyödynnettiin sekä manuaalisen arvioinnin apuvälineenä että osittaisessa automaattiarvioinnissa. (Opetushallituksen muistio 2011 :2).

Kokeiden suorittamista vastasi tässä kokeilussa minkä tahansa sähköisen lomakkeen erilaisten kenttien täyttäminen. Jokainen kokelas sai oikeuden ohjelman käyttämiseen opettajalta, jonka jälkeen kokelaat pystyivät anonymisti lähettämään lomakekenttiä vastauksineen opettajalle. Kellään muulla ei ollut oikeutta vastausten näkemiseen kuin kokelaalla ja opettajalla. Oli myös opettajan päätettävissä, sulkeutuuko kenttä lopullisesti lähettämisen jälkeen, vai voiko kokelas vielä palata vastaukseensa myöhemmin ja lähettää sen uudelleen arvioitavaksi uusilla tiedoilla täydennettynä. Tässä kokeilussa keskiössä oli siis tietokonepohjaisen, lineaaristyyppisen arvioinnin testaaminen, mutta ilman adaptiivista järjestelmää. Vihervaaran (2009, 17) mukaan tämän kokeilun perusteella verkkoon soveltuvia kokeita olivat kuitenkin esimerkiksi Skypetentti, ryhmäkoee, suullinen dialogikoe, standardikokeet eli tenttiakvaario tai mekaaniset kokeet, kongressitentti, tietokoneavusteinen koe, esseet, raportit, oppimissalkut ja verkkopäiväkirjat (kts. tarkemmin: Opetushallituksen muistio 2011 :2).

Tenttiakvaariossa esimerkiksi käytetään sekä välitöntä, järjestelmän antamaa palautetta että opettajan myöhemmin antamaa henkilökohtaista palautetta. Tenttilanteessa ohjelma arpoo opettajan ennalta asettaman määrän koetehtäviä ja kokelas vastaa niihin ohjelman vaatimalla tavalla. Kokeen suoritus aika voi olla ennalta ohjelmoitu tai vapaa. Kokeen lopussa ohjelma antaa välittömän palautteen monivalintatehtävien onnistumisesta. Esseekysymykset menevät opettajalle arvioitaviksi, jonka jälkeen hän antaa lopullisen arvionsa kokelaalle heidän keskenään sopimallaan tavalla. Yksi tämänkaltaisten testien hyvistä puolista on, että niihin on useimmiten ohjelmoitu myös mahdollisuus kokelaalle saada testin tulokset esimerkiksi henkilökohtaiseen sähköpostiinsa tai arvosanajärjestelmään (esim. WebOodi). Tunnistautumiseen taas voidaan käyttää esimerkiksi pankkitunnuksia tai kansalaisvarmenteen sisältävää henkilösirukorttia. Koekysymykset eivät kuitenkaan, ainakaan toistaiseksi testatuissa malleissa, muokkautu kokelaan mukaan vaan kone arpoo koealueen sisältä kokelaasta riippumatta ennalta ohjelmoidun määrän kysymyksiä.

Iso-Britanniassa tehdyssä adaptiivisen testauksen pilottiprojektissa saatiin useita Suomeenkin yleistettävissä olevia tuloksia. Adaptiivinen testaus tarjosi tässä tapauksessa uutta tietoa oppilaan menestymisestä niin testin suorittamisen aikana kuin postuumisti, tilastollisena yhteenvedona. Testin suorittamisen aikana kerätyn metatiedon avulla voitiin mukauttaa kullekin käyttäjälle näytettävää testisisältöä. Käyttäjät voitiin jakaa edistymisensä perusteella eri tasoluokkiin, ja näyttää heille kuhunkin tasoon määriteltyjä tehtäviä. Tämä vähensi esimerkiksi vierekkäin istuvien oppilaiden lunttausmahdollisuuksia, sillä useimmiten tietokoneen ruudulla näkyvillä olevat testit ja vastaukset olivat erilaisia. Testin sähköistämisen ansiosta kokeen suunnittelijat voivat määrätä myös tehtävien tekemisen järjestyksen tarkasti pedagogisten metamallien avulla. Adaptiivisuuden erityisenä etuna tutkijat esittivät myös koeasetelman, jossa kysymys b) perustuu vahvasti kohdan a) osaamiseen. Mikäli oppilas vastaa kohtaan a) väärin, voidaan hänelle tarjota b)-vaihtoehdoksi kysymys jonka riippuvuus edellisestä kohdasta on vähäisempi ja täten mahdollistaa oppilaan parempi menestys testissä. Sähköisen testauksen avulla voitiin myös välttää tilanteet, joissa oppilas harppoi tehtävien yli tai korjaili hätäisesti aikaisempia vastauksiaan koetilaisuuden lopuksi (Cowan, 2007).

Sähköisen, adaptiivisen ylioppilaskokeen toteutustapa mahdollistaa siis mm. tarkan tilastotiedon keräämisen. Tietoa voidaan myös rikastaa eri tehtävien suorittamiseen tai keskeneräiseen riippuvuuteen kohdistuvalla uudella tiedolla. Sähköisen koejärjestelmän tuottaman uuden tiedon avulla voidaan tarjota palautetta oppilaiden suoriutumisesta. Koetulosten läpinäkyvyyden avulla kokelas, hänen vanhempansa, opettajat, koulut ja opetukseen liittyvät organisaatiot voivat ohjata toimintaansa todellisiin oppimistuloksiin perustuvan tiedon avulla (Cowan, 2007). Uuden tiedon avulla voidaan myös tehdä sähköiseen testaamiseen liittyvää akateemista tutkimusta uudelta näkökulmalta.

Hollannissa puolestaan Cito -niminen testausfirma tekee oppilaitoksille kansallisia high stakes-kokeita, joiden avulla arvioidaan opiskelijoiden soveltuvuutta mm. peruskoulun jälkeiseen koulutukseen. Koesuoritusten perusteella kyetään laatijoiden mukaan arvioimaan myös koulutusjärjestelmän toimivuutta. Cito -testit koostuvat hollannin kieleen, matematiikkaan, ymmärrystaitoihin, opiskelutaitoihin ja "maailmatietouteen" (maantieto, biologia, historia) liittyvistä monivalintatehtävistä, eivätkä mittaa niinkään sisältötietoutta vaan enemmän kokelaan yleistä soveltuvuutta tietylle koulutuslalle. Peruskouluista noin 85% kerrotaan käyttävän kyseisiä päättökokeita. Kansallisten testien lisäksi Citolla on myös testejä mm. peruskoulunaikaisen edistymisen mittaamiseen, ammatilliseen koulutukseen, lukiokoulutukseen, aikuiskoulutukseen ja erityisopetukseen liittyen. Toinen kattava, tietokonepohjainen Cito -testi jota hollantilaiset koulut käyttävät, on lukiokoulutuksen päättövaiheen koe niille opiskelijoille, jotka haluavat jatkaa opintojaan yliopistossa. Testissä on sekä avoimia kysymyksiä että monivalintakysymyksiä. Kokeen painoarvo loppuarvostelussa on 50%. Toiset 50% arvosanasta muodostaa koulun itse hallinnoima sisäinen päättökoe. Tämän vuoksi lukioiden onkin toimitettava kopiot päättökokeistaan joka vuosi koulutuksen tarkistuslautakunnalle (Education

Inspectorate), joka valvoo koulujen kansallista yhteismitallisuutta. (Center on International Education Benchmarking, 2013).

Myös Euroopan komission raportti *“The Transition to Computer-Based Assessment - New Approaches to Skills Assessment and Implications for Large-Scale Testing”* vuodelta 2009 selvittää tietokonepohjaisen adaptiivisen testauksen hyötyjä ja haittoja kattavasti monesta eri näkökulmasta. Alussa keskitytään muutamiin testikokeiluihin mm. Tanskassa ja Islannissa, ja myöhemmin raportissa ruoditaan myös yleisellä tasolla kyseisen testaustavan periaatteita, soveltamisaloja ja resurssikysymyksiä. Raportti sisältää vahvan eurooppalaisen näkökulman tietokonepohjaiseen testaukseen, keskustelua ja tutkimuspohjaista tietoa sen periaatteista, kynä- ja-paperi -testauksen vertailua tietokonepohjaiseen testaukseen ja kyseisen testaustavan metodologiaa. Seuraavaksi esittelemme raportista muutamia keskeisiä ajatuksia, joita on hyvä ottaa huomioon myös suomalaisessa testauskontekstissa.

Teoksen toimittaneet Scheuermann ja Björnsson toteavat raportin johdannossa (s.7-11) tietokonepohjaisen testauksen olevan niin kompleksinen kokonaisuus ettei siihen laajojen tutkimusten ja kokeilujenkaan jälkeen ole yhtä selkeää ratkaisua tai lähestymistapaa. Testausjärjestelmän vaatimukset riippuvat esimerkiksi siitä, onko se tarkoitettu opiskelun aikaiseen taitojen arviointiin, onko testillä merkitystä koulutuksesta toiseen siirryttäessä ja onko se koulu- tai maakohtainen vai kansainvälinen taitojen mittari. Tietokonepohjaisilla testaustavoilla on kuitenkin todettu olevan niin paljon positiivisia vaikutuksia arvioinnin nopeutumiseen ja etenkin objektiivisuudesta eroon pääsemiseen, että niiden katsotaan helposti voittavan yksittäiset hankaluudet järjestelmien kehittämisessä.

Scheuermann ja Björnsson esittelevät myös Euroopan komission CRELL- keskuksen (Centre for Research on Lifelong Learning) aloittamaa tutkimusprojektia, jonka tarkoituksena on selvittää ja analysoida tietokonepohjaista testausta ja järjestelmien tehokkuutta laajan skaalan testauksissa. Myös Reykjavikissa sijaitseva The Educational Testing Institute tekee parhaillaan aiheeseen liittyvää tutkimusta, tavoitteenaan selvittää nimenomaan siirtymää kynä- ja-paperi -kokeista CAT-pohjaiseen testaukseen (Computer Adaptive Testing). Kyseinen toimija on myös järjestänyt muutamia kansainvälisiä seminaareja, joissa tietokonepohjaisen testauksen nykysuuntia on käsitelty niin eurooppalaisessa kuin laajemmassakin kontekstissa. Lisätietoja näistä löytyy osoitteesta <http://crell.jrc.eu.europa.eu>.

Jakob Wandall Tanskan opetushallituksesta esittelee komission raportissa Tanskan uusia kansallisia kokeita, jotka suoritetaan CAT:in avulla (s.45-50). Prosessi lähti käyntiin vuonna 2006, ja lopullisten testausmuotojen tuli olla raportin mukaan käytössä vuoteen 2010 mennessä. Lisää tietoja hankkeen etenemisestä löytyy esimerkiksi osoitteesta evaluating.uvm.dk (tanskaksi). Wandall toteaa Tanskan hankkeesta raportoidessaan saman kun useat muutkin CAT:ia tutkineet; testeistä on luotava useita versioita ja jokaista versiota tulee pilotoida useita kertoja monilla eri kokelasjoukoilla. Tanskassa ensimmäisen testiversion tarkastelussa esiin

nousi mm. ongelmia tehtävien laadun ja määrän suhteen. Tulokset eivät olleet totuudenmukaisia, joten testipankkia oli laajennettava ja tehtäviä eri osioineen pilotoitava uudelleen. Raportin julkaisuvaiheessa testejä oli kehitetty 2,5 vuoden ajan, ja edelleen hankkeessa oli useita ongelmia.

Lähtökohdaksi tanskalaiset olivat valinneet saman määritelmän adaptiivisen testaamisen luonteesta kuin tässäkin selvityksessä: adaptiiviset testit mukautuvat kokelaan taitotason mukaan testin edetessä. Ensimmäinen kokelaalle tuleva tehtävä on vaikeustasoltaan keskiluokkaa (suhteutettuna testin kohderyhmään, esim. peruskoulun päättövaiheen äidinkielen taidot). Mikäli kokelas vastaa annettuun tehtävään oikein, seuraava tehtävä on automaattisesti vaikeampi. Mikäli hän vastaa tähän väärin, seuraava tehtävä on puolestaan tasoltaan hieman helpompi. Tällä tavoin testi muokkautuu kokelaan mukaan, mikä tosin aiheuttaa myös sen, että eritasoiset kokelaat saavat eri määrän tehtäviä ratkottavakseen. Toisaalta kukaan ei joudu kuitenkaan tekemään itselleen "tarpeettomia" tehtäviä vaan ohjelma syöttää ainoastaan sellaisia tehtäviä, joihin laskee kokelaan olevan kykeneväinen vastaamaan. Adaptiivinen testaustapa antaakin siis eniten relevanttia ja todenmukaista tietoa kokelaan oikeasta taitotasosta eikä ainoastaan siitä, ovatko ennalta suunnitellut tehtäväosiot olleet hänelle liian helppoja tai vaikeita.

Tanskassa CAT:ia on Wandallin mukaan testattu seuraavissa aineissa seuraavin osioin, kansallista testausta kehiteltäessä: 1) *fysiikka/kemia: profiili 1: energia ja energian eri muodot, profiili 2: ilmiöt ja aineet, profiili 3: sovellukset ja näkökulmat*, 2) *matematiikka: profiili 1: numerot ja algebra, profiili 2: geometria, profiili 3: käytännön sovellukset*, ja 3) *lukeminen (tanskan kieli): profiili 1: yleinen kielen ymmärtäminen, profiili 2: dekadaus/tulkitseminen, profiili 3: luetunymmärtäminen*. Jokaisessa oppiaineessa on siis ollut kolme opetussuunnitelmien mukaista osiota, jotka jokainen kokelas on suorittanut omalla taitotasollaan.

Kyseessä oleva testijärjestelmä on ollut koulukohtaisesti yhteydessä testin laatijan hallinnoimaan sivustoon (evaluing.uvm.dk), ja sekä opettajalla että oppilailla on ollut pääsy kyseiselle sivustolle vapaasti omalta koneeltaan. Jokaiselle kokelaalle ja opettajalle on jaettu oma, henkilökohtainen käyttäjätunnus ja salasana, jolla kirjautua järjestelmään koetilanteessa. Itse koetta on päässyt suorittamaan vasta kun opettaja on omalta koneeltaan avannut koulukohtaisen yhteyden, ja ilmoittanut yhteydestä kokelaille. Kokeen kesto on ollut 45 minuuttia, jonka aikana kokelaat ovat saaneet vastata niin moneen kysymykseen kuin ovat ehtineet. Keskimäärin vastauksia on ehditty rekisteröidä järjestelmään noin 50-80 kysymyksen verran em. ajassa. Mikäli kokelas on tarvinnut enemmän aikaa suoritukseensa, opettajalla on ollut valtuudet antaa hänelle siihen mahdollisuus omalta koneeltaan käsin. Testijärjestelmän keskiössä on keskustietokone, joka rekisteröi jokaisen kokelaan antaman vastauksen, arvioi sen ja syöttää kokelaan tason mukaisesti seuraavat tehtävät. Tehtävien vaikeustason määrittelyssä järjestelmä käyttää apunaan Raschin skaalaa (tai "theta-skaalaa"). Jokaista testiosiota kohden järjestelmään on syötetty 500 eritasoista tehtävää, jotka on esitestattu 500-700 kokelaalla ennen varsinaista testausta.

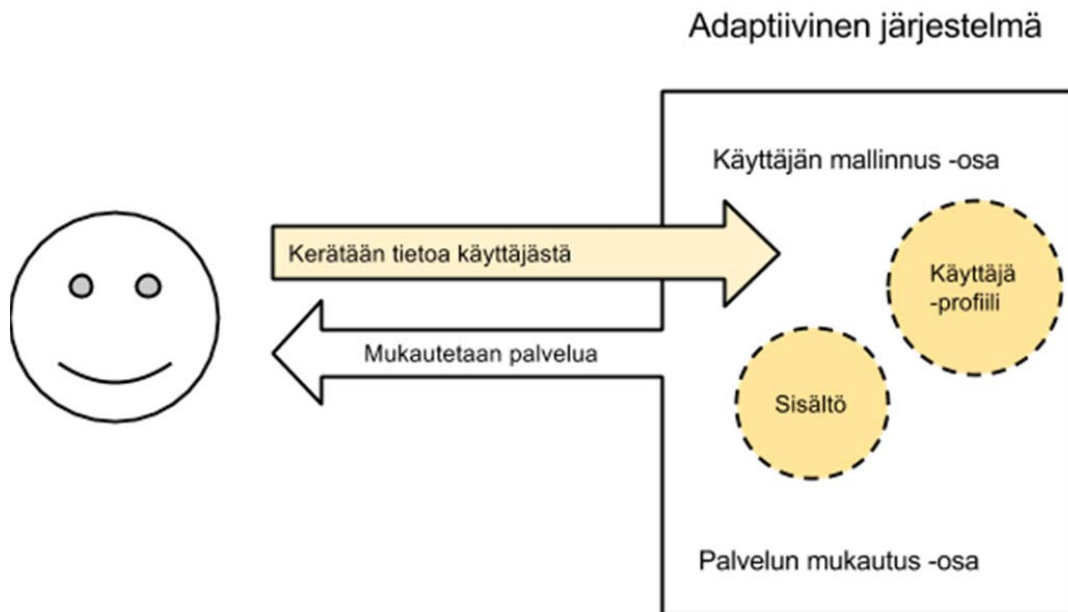
Wandall toteaa kuitenkin, ettei kyseinen testijärjestelmä ainakaan toistaiseksi taivu testaamaan kokelaiden kykyä suullisen eikä kirjoitetun kielitaidon osalta. Näitä taitoja testataan Tanskan peruskouluissa siis edelleen ainoastaan opettajajohtoisesti.

Nathan A. Thompson ja David J. Weiss toteavat omassa artikkelissaan "*Computerized and Adaptive Testing in Educational Assessment*" (s.127-134) että CAT:in edut ovat selkeästi sen haasteita suurempia. Heidän esittelemänsä mallin mukaan kokelas voi päättää testin jopa jo 10 osatehtävän jälkeen, mikäli testille laadittu tasovaatimus täyttyy. Tämä ei kuitenkaan päde laajemman skaalan päättökokeissa. Joka tapauksessa CAT-testaus nopeuttaa testitilanteita huomattavasti, ja järkevöittää tehtäviä kokelaan näkökulmasta. Tehtäviä myös tarvitaan vain noin puolet lineaarisiin testeihin verrattuna, sillä ne ovat valmiiksi skaalattuja ja testattuja. Testaustavalla on myös kokelaiden motivaatioon kytkeytyviä positiivisia vaikutuksia; kun jokainen saa ainoastaan tasolleen soveltuvia kysymyksiä, turhautuminen liian helppojen tai liian vaikeiden tehtävien kohdalla poistuu. Tämä vaikuttaa myös kokelaan kokemaan stressiin.

3 Tietokonepohjaisen testauksen teknisen toteutuksen perusteet

Tässä osiossa perehdymme tarkemmin adaptiivisen testin rakenteeseen, ja annamme ehdotelmia sellaisen mahdollisesta sovelluksesta suomalaisessa kontekstissa.

Testin sisältö: ohjetekstit, taustatiedot, kysymykset ja vastaukset tallennetaan tietojärjestelmään. Adaptiivisuutta toteuttavan järjestelmän toiminta perustuu siis käyttäjän profiilitiedon keräämiseen ja järjestelmän toiminnan mukauttamiseen profiilitiedon ja sisällöstä rakennettujen metamallien avulla (Brusilovsky, Millán, 2007). Seuraavassa kuvassa hahmotetaan adaptiivisen järjestelmän osia ja sen toiminnan vuorovaikutteista luonnetta.



Jotta voitaisiin toteuttaa sisällön henkilökohtaista mukauttamista, on järjestelmän käyttäjät voitava yksilöidä. Tätä varten voidaan rakentaa koejärjestelmän ja jonkin olemassaolevan henkilötunnistusjärjestelmän liitos. Tunnistautumistapaa suunnitellessa on otettava huomioon sen kuormittavuuden, luotettavuuden ja järjestelmien liitoksen tarjoaman tiedon määrä. Tarkoituksenmukainen tunnistautumismenetelmä on oppilaalle ja koejärjestelijälle yksinkertainen, mutta tarjoaa samalla sähköisen kokoon tuottavalle järjestelmälle riittävät tiedot opiskelijan perusprofiilin rakentamiseen.

Opiskelijan perustietojen saatavuuden ja täydentävän tiedon keruun strategiset päätökset muodostavat ensimmäisen hahmotelman koejärjestelmän arkkitehtuurista. Strategiaan liittyviä kysymyksiä ovat: Mitä tietoa on helposti saatavilla? Mikä määrä opiskelijan perustietoa on riittävä koejärjestelmän toiminnan tueksi? Mitä tietoa tarvitaan, jotta voidaan tarjota oppilaalle hänen tasoaan vastaava koe? Mistä tiedot olisi saatavilla helpoiten, ilman manuaalista syöttämistä? Voidaanko tunnistautuminen toteuttaa luotettavasti myös ilman verkkoyhteyttä?

Mukauttaakseen toimintaansa, on adaptiivisesti toimivan järjestelmän kerättävä kustakin käyttäjästä perustietoja *täydentävää* profiilitietoa. Profiilitiedon keruu voidaan toteuttaa eksplisiittisesti eli perustuen käyttäjän tietoiseen vuorovaikutukseen. Tietoa voidaan kerätä myös ilman käyttäjän varsinaista osallistumista, implisiittisesti (Gauch, Speretta 2007). Seuraava taulukko antaa esimerkkejä eri tyyppisestä käyttäjäprofiilitiedosta.

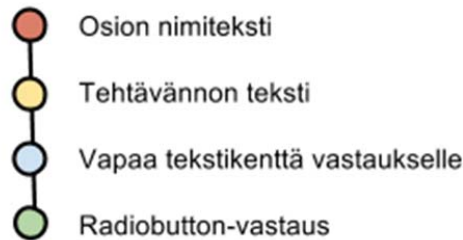
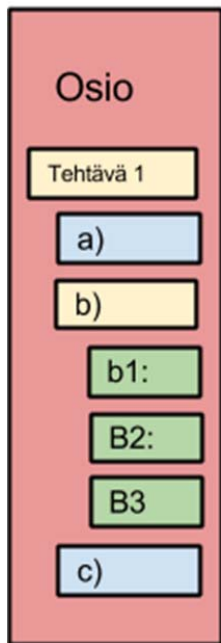
	Tyyppejä	Esimerkkejä
Eksplisiittinen profiilitieto	Käyttäjän itsensä aktiivisesti tarjoama tieto	Kiinnostuksen kohteet, koevastaukset
Implisiittinen profiilitieto (Ei aktiivisia toiminpiteitä)	Ulkopuolisen tahon antama taustatieto	Adaptiiviseen järjestelmään ennalta ladatut tiedot, kuten koulukohtaisen opiskelijarekisterin tiedot: opiskelijanumero, osoite, puhelinnumero jne. Eri järjestelmien liitosten avulla saadut tiedot (esim. Vetuma)
	Käyttäjän tietämättä talletettu tieto	Eteneminen näytettävässä sisällössä, viipyminen sisällössä, navigaatiopolku
	Muunnettu tieto	Syntymäaika → ikä → käyttäjä on alaikäinen Käyttäjä on opiskellut aihetta 7 vuotta → käyttäjä tietää paljon asiasta X → ekspertti käyttäjä on kiinnostunut asiasta y → kannattaa näyttää kysymys, joka liittyy teemaan Y Käyttäjä on edennyt askelelle X → näytetään askeleet Y,Z

Dokumentin rakennetta koskevia malleja

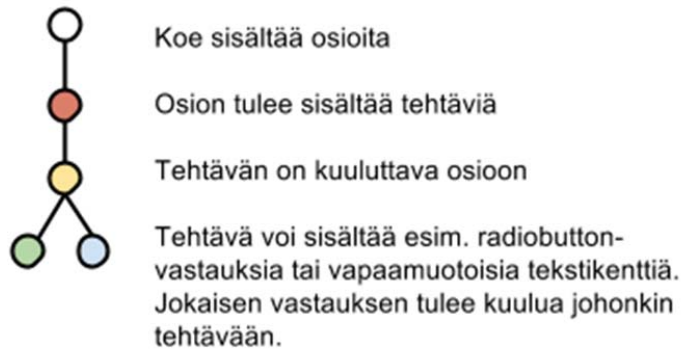
Jotta sisällön esittämistä voidaan mukauttaa vastaamaan käyttäjien tarpeita, tarvitaan käyttäjäprofiilitiedon lisäksi meta-tietoa esitettävästä sisällöstä: sen rakenteesta, eri rakenteiden välisestä hierarkiasta, sisällössä navigointiin liittyvistä tavoitteista, teemoista ja sisällössä esiintyvistä käsitteistä (Micarelli, Sciarrone, Marinilli 2007). Seuraavassa kuvassa annetaan

esimerkki web-sivun muodossa toteutetun koedokumentin rakenteeseen liittyvästä Document Object Model (DOM) -mallista ja sen osien välille määritellystä hierarkiamallista.

Koedokumentin malli (DOM)



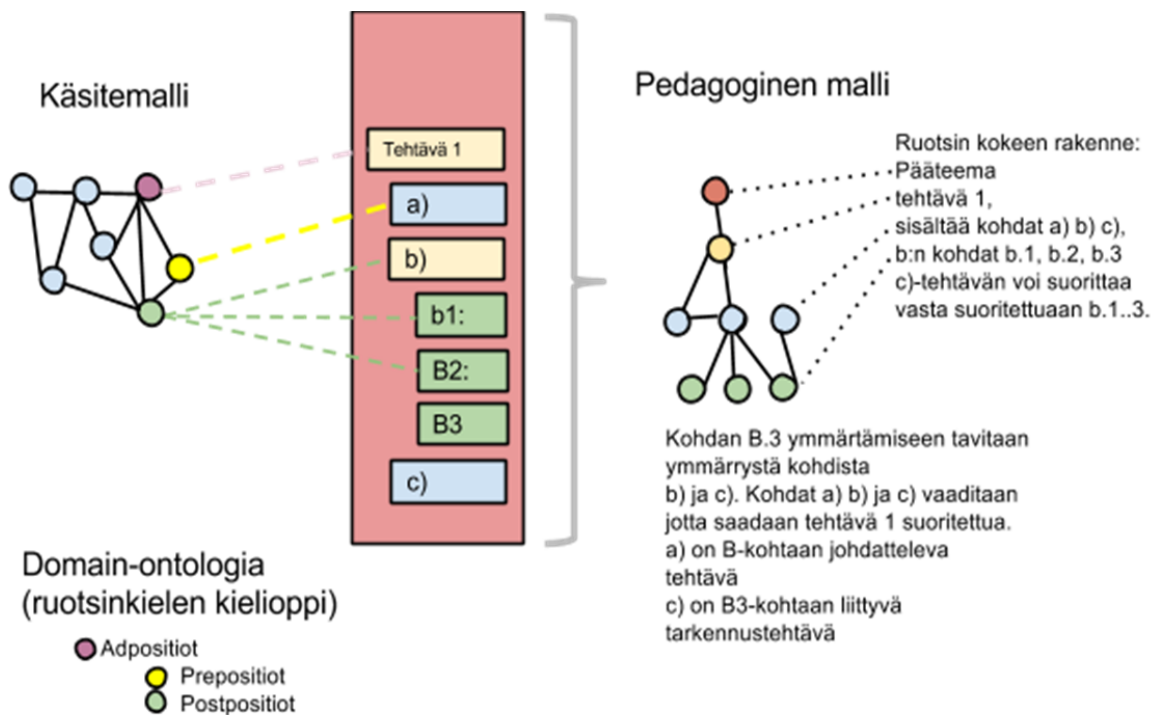
Hierarkiamalli



Dokumentin sisältöä koskevia malleja

Dokumentin sisältöä käsittelevistä malleista tässä yhteydessä esitetään käsitemalli ja pedagoginen malli. Käsitemallin avulla voidaan jokaiseen tehtävään liittää tietoa esimerkiksi siitä, mitä kieliopin osa-alueita kukin koekysymys käsittelee (Conlan, Wade et al. 2002). Käsitemalleja voidaan rakentaa kieliopin lisäksi minkä tahansa taksonomian, ontologian tai vapaasti määriteltävissä olevan teeman ympärille.

Kokeen pedagogisen mallin avulla määritetään se polku, jonka opiskelija koesisällössä kulkee. Pedagogisia malleja voidaan rakentaa useita, vastaamaan eri tasoisten opiskelijaryhmien tarpeita. Käytettävä malli valitaan kunkin koejärjestelmään kirjautuneen käyttäjän profiilitietojen perusteella. Hyvin onnistuneen käsitteellisen ja pedagogisen mallinnuksen avulla voidaan varioida kokeita eri vuosina valitsemalla kokeisiin saman tasoisia tehtäviä eri tema-alueilta (Conlan, Wade et al. 2002).



Testin suoritusta koskeva tieto

Opiskelijan vastaukset koemateriaalissa tallennetaan kunkin opiskelijan henkilökohtaiseen profiilitietoon. Kun järjestelmällä on ajantasainen kuva siitä, missä kohtaa koedokumentin sisältöä kokeen suorittaja kulkee, voidaan opiskelijan navigoimista eri koesisältöjen välillä helpottaa mukauttamalla järjestelmän tarjoamia valikoita. Osia valikosta voidaan piilottaa, niihin pääsyä voidaan rajoittaa tai osia voidaan ottaa kokonaan pois käytöstä. Mukauttamisen avulla voidaan tarjota suoraa ohjausta sille, kuinka käyttäjän tulee edetä sisällössä (Brusilovsky, 2007). Myös opiskelijan todellista etenemistä koskeva navigaatiopolku (toteutunut pedagoginen malli) voidaan tallentaa käyttäjäprofiiliin myöhempää tarkastelua varten. Mallia voidaan rikastaa tallentamalla käyttäjän profiilitietoon suoritukseen liittyvää metatietoa (tehtävän suoritus-aika, vastauksen korjailu).

Johtopäätös

Adaptiivisen testauksen toteuttamiseksi on olemassa runsaasti erilaisia kaupallisia järjestelmiä, joita voidaan käyttää vertailukohteina Suomen kansallisen järjestelmän rakentamisprojektissa. International Association for Computer Adaptive Testing (IACAT) on organisaatio, jonka

tarkoituksena on edistää adaptiivisen testauksen käyttöä ja tutkimusta. Sen verkkosivuille (<http://www.iacat.org/>) on koottu tieto CAT-ohjelmistojen valmistajista. Organisaatio julkaisee myös "The Journal of Computerized Adaptive Testing"-lehteä ja järjestää vuosittaista, adaptiiviseen tietokoneavusteiseen testaamiseen liittyvää konferenssia.

Vaikka sähköisen ylioppilaskokeen toteutukseen valittaisiin valmisohjelmisto, on käyttäjäprofiilien, koemateriaalin ja siihen liittyvien metatietojen suunnittelu laaja kehitysprojekti. Profiilitiedon ja metamallien välisen vuorovaikutuksen kompleksisuuden takia myös järjestelmän testaus ennen käyttöönottoa on monimutkaista, koetilanteen simulointia ja kokonaisvaltaista näkemystä vaativaa työtä. Onnistuneen suunnittelu- ja kehitystyön avulla voidaan tarjota opiskelijalle sujuva ja hänen tarpeitaan tukeva koe-elämys. Sen avulla myös tarjoutuu yhteiskunnalle uutta tietoa opiskelijoiden suorituksista mikä mahdollistaa opetuksen reaktiivisen mukauttamisen ja uusia mahdollisuuksia testausprosessin jatkuvaan parantamiseen akateemisen tutkimuksen avulla.

Yhteenveto

Adaptiivinen testaaminen on tässä selvityksessä esille tulleen tiedon valossa monimutkainen, mutta oikein suoritettuna luotettava ja etenkin aikaa säästävä tapa arvioida esimerkiksi koulutuksen päättövaiheessa olevia oppijoita. Testausta voidaan käyttää kaikissa oppiaineissa, mutta kielitaidon suullisesta testauksesta ei ole juuri saatavilla olevaa tutkimustietoa. Opettaja voi keskittyä kattavammin myös kurssien aikana ilmeneviin pedagogisiin haasteisiin (esim. opetuksen eriyttäminen) ja käyttää enemmän aikaa substanssin hallitsemisen ohjaamiseen. Automaattiarviointi vähentää lisäksi arvioinnin inhimillisistä syistä johtuvaa vaihtelua ja siten tasa-arvoistaa kokelaiden asemaa, ja sen mahdollistama tiedonkeruu avaa uudenlaisen näkymän opetuksen suunnitteluun koetulosten läpinäkyvyyden avulla. Adaptiivinen testaus lisää myös mahdollisuuksia suunnitella koulutuksen aikaista adaptiivisuutta. Erityistä huomiota tulee siis edelleen kiinnittää opetuksen eriyttämiseen ja mukauttamiseen jo koulutuksen aikana, ei ainoastaan loppukokeessa. Opiskelijat hyötyvät tiedollisesti ja taidollisesti eniten laadukkaasta, nykyaikaisesta ja monimediaisesta opetuksesta ja opiskelusta, joka on tarkoituksenmukaista ja tulevaisuuden kannalta heille relevanttia. Mitä opetetaan, sitä testataan; tämän koulutuksen peruseriaatteen mukaisesti adaptiivisuuden on näytävä sekä opetuksessa että testauksessa.

Myös Euroopan komission julkaiseman raportin mukaan kokemukset useissa maissa suoritetuista tietokonepohjaisista adaptiivisista kokeilutesteistä ovat positiivisia, mutta kokeilujen pohjalta on noussut esiin myös monia perustavanlaatuisia ongelmia. Suurimpina haasteina ovat testipankkien laajuus/kapeus ja tehtäväosioiden riittävä pilotointi sekä käytetty verkko ja palvelin. Kansallisessa, samaan aikaan tapahtuvassa testauksessa palvelimen on sijaittava yhdessä paikassa ja verkon oltava kapasiteetiltaan riittävä suorittamaan samaan aikaan tuhansien

kokelaiden hallinnointia. Näiden ongelmien poistamiseksi on kiinnitettävä runsaasti huomiota mm. pilotointiin järjestelmää valitessa ja muokatessa. (Hingel, 2009).

Kolmen kansainvälisen testausalan toimijan mukaan (Pearson, ETS ja College Board, 2010) tietokonepohjainen, adaptiivinen testaaminen tulee kuitenkin lähiaikoina valtaamaan runsaasti enemmän alaa etenkin summatiivisessa arvioinnissa, sillä sen hyödyt ovat heidänkin mukaansa selkeitä: 1) laaja-alaisen kehittymisen seurannan mahdollistuminen läpi oppijan koulu-uran (tiedot ovat jo automaattisesti sähköisessä järjestelmässä, josta opetuksen järjestäjä saa ne tarvittaessa), 2) testausajat lyhentyvät kun ihmisten käsin tekemä työmäärä vähentyy, 3) objektiivisuus arvioinnissa kasvaa ja 4) testiohjelmistojen kehittymisen myötä myös avointen kysymysten arviointi tapahtuu järjestelmässä, ei ainoastaan opettajavetoisesti. Haasteena on kuitenkin saada järjestelmiin tehtäväosioita, joiden vaihtelu takaisi jokaisen kokelaan luotettavan ja totuudenmukaisen taitotason arvioinnin. Tehtäviä tulee olla odotettuun päättötasoon nähden tason alapuolelta, oppiaineen asetetulta päättötasolta, ja tason yläpuolelta. Adaptiivisuus ei siis edellytä monivalintoja, vaan riippuen valittavasta toteutustavasta mikä tahansa koedokumentin osa voi olla adaptiivisen testauksen kohteena. Myös avointen kysymysten ja esseiden arviointia varten on kehitetty testiohjelmiä. Keskeisintä päättökokeen muuntamisessa sähköiseen muotoon on kuitenkin eri ohjelmien ja niiden toimintaa ohjaavien käyttäjäprofiili-, ja sisältömallien huolellinen suunnittelu, testaus ja pilotointi mahdollisimman kattavasti sekä normaaliopetukseen osallistuneiden kokelaiden että eriytetyn opiskelusuunnitelman mukaan opiskelleiden kokelaiden osalta.

Lähteet

- Brusilovsky, P. 2007. Adaptive navigation support. Teoksessa: The adaptive web. P. Brusilovsky, A. Kobsa, ja W. Nejdl (Toim.). Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin, Heidelberg. S. 263-290.
- Brusilovsky, P. & Millán, E. 2007. User models for adaptive hypermedia and adaptive educational systems. Teoksessa: The adaptive web, P. Brusilovsky, A. Kobsa, ja W. Nejdl (Toim.). Lecture Notes in Computer Science, Vol. 4321. Springer-Verlag, Berlin, Heidelberg. S. 3-53.
- Center on International Education Benchmarking: Netherlands. 2013.
<http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/netherlands-overview/netherlands-instructional-systems/> Luettu 01.12.2013
- Conlan, O., Wade, V., Bruen, C., ja Gargan, M. 2002. Multi-model, Metadata Driven Approach to Adaptive Hypermedia Services for Personalized eLearning. Teoksessa: Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems(AH '02). P. De Bra, P. Brusilovsky, ja R. Conejo (Toim.). Springer-Verlag, London, UK, UK. S. 100-111.
- Cowan, M. 2007. Assessment in the twenty-first century: the role of computerised adaptive testing in national curriculum subjects. JOURNALISSA: Teacher Development: An international journal of teachers' professional development. S. 241-257. DOI: 10.1080/13664530100200136.
- Gauch, S., Speretta, M., Chandramouli, A., ja Micarelli, A. 2007. User profiles for personalized information access. Teoksessa: The adaptive web. P. Brusilovsky, A. Kobsa, ja W. Nejdl (Toim.). Lecture Notes In Computer Science, Vol. 4321. Springer-Verlag, Berlin, Heidelberg. S. 54-89.
- Kupari, P. 2009. Koulutuksen kansainvälisen arviointitutkimuksen kysymyksiä. Virkaanastujaisesitys. Jyväskylän yliopiston Koulutuksen tutkimuslaitos. Saatavilla: https://www.jyu.fi/ajankohtaista/arkisto/2009/10/tiedote-2009-10-15-10-01-09-110710/at_download/attachment. Luettu: 02.12.2013.
- Micarelli, A., Sciarrone, F. ja Marinilli, M. 2007. Web document modeling. Teoksessa: The adaptive web. P. Brusilovsky, A. Kobsa, ja W. Nejdl (Toim.). Lecture Notes In Computer Science, Vol. 4321. Springer-Verlag, Berlin, Heidelberg. S. 155-192.

Opetushallituksen muistio 2011 :2. Tieto ja viestintäteknikka opetuskäytössä -Välineet, vaikutus ja hyödyt. Tilannekatsaus, toukokuu 2011. Heino, T., Honkasalo, R., Kiesi, E., Koivisto, J., Koskinen, K., Nyyssölä, K., Packalen, P., Vähähyppä, K. Opetushallitus.
http://www.oph.fi/download/132877_Tieto-_ja_viestintateknikka_opetuskaytossa.pdf. Luettu 01.12.2013.

Pearson, ETS, College Board. 2010. Some Considerations Related to the Use of Adaptive Testing for the Common Core Assessments.
<http://research.collegeboard.org/sites/default/files/publications/2012/8/ccss-2010-2-adaptive-testing-common-core-assessments.pdf>. Luettu 02.12.2013.

Scheuermann, F. & Björsson, J. (Toim.). 2009. The Transition to Computer-Based Assessment. New Approaches to Skills Assessment and Implications for Large-Scale Testing. European Commission. 211 s. Saatavilla: <http://www.gesci.org/assets/files/reporttransition.pdf>. Luettu 04.12.2013.

Thompson, N.A., & Weiss, D. J. Computerized and Adaptive Testing in Educational Assessment. Teoksessa: Scheuermann, F. & Björsson, J. (Toim.). 2009. The Transition to Computer-Based Assessment. New Approaches to Skills Assessment and Implications for Large-Scale Testing. European Commission. S. 127-134. Saatavilla: <http://www.gesci.org/assets/files/reporttransition.pdf>. Luettu 04.12.2013.

Vihervaara, E. (Toim.). 2009. Koe oppimisympäristönä. Länsi-Suomen läänin aikuislukioiden raportti koekäytänteiden kehittämistä lukiassa II. Saatavilla: http://issuu.com/eviherva/docs/koe_oppimisymp_riist_n_versio_lopullinen_4.8. Luettu 04.12.2013.

Wandall, J. National Tests in Denmark – CAT as a Pedagogic Tool. Teoksessa: Scheuermann, F. & Björsson, J. (Toim.). 2009. The Transition to Computer-Based Assessment. New Approaches to Skills Assessment and Implications for Large-Scale Testing. European Commission. S. 45-51. Saatavilla: <http://www.gesci.org/assets/files/reporttransition.pdf>. Luettu 04.12.2013.

Lisää tietoa adaptiivisesta testaamisesta

Tietokonepohjaista testausta tutkiva taho: International Association for Computer Adaptive Testing (IACAT), www.iacat.org

Artikkeli (yleistajuinen): Adaptive Testing Evolves to Assess Common-Core Skills
<http://www.edweek.org/dd/articles/2012/10/17/01adaptive.h06.html> Luettu 28.11.2013.

Kirja: Computerized Adaptive Testing: Theory and Practice. Linden, Wim J., Glas, Cees A.W. (Toim.). 2000. 323 s. Springer.

Käytettyjä hakusanoja

Computer Adaptive Testing (CAT), tailored testing, e-assessment, adaptivity, adaptive testing, adaptive assessment, adaptive exams, educational measurement, adaptiivisuus, kokeet ja adaptiivisuus, mukautuvuus, testing technology