

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Kärkkäinen, Tommi; Juutinen, Sanna; Saarela, Mirka; Nissinen, Kari

Title: Lokidatan käyttö oppilaiden profiloimisessa - sovellus matematiikan PISA-aineistoon

Year: 2018

Version: Published version

Copyright: © Kirjoittajat & Suomen kasvatustieteellinen seura, 2018.

Rights: CC BY 4.0

Rights url: <https://creativecommons.org/licenses/by/4.0/>

Please cite the original version:

Kärkkäinen, T., Juutinen, S., Saarela, M., & Nissinen, K. (2018). Lokidatan käyttö oppilaiden profiloimisessa - sovellus matematiikan PISA-aineistoon. In J. Rautopuro, & K. Juuti (Eds.), PISA pintaa syvemmältä : PISA 2015 Suomen pääraportti (pp. 259-289). Suomen kasvatustieteellinen seura. Kasvatusalan tutkimuksia, 77. <http://urn.fi/URN:ISBN:978-952-5401-82-0>

11. Lokidatan käyttö oppilaiden profiloimisessa – sovellus matematiikan PISA-aineistoon

Johdanto

PISA (Programme for International Students Assessment) on joka kolmas vuosi toteutettava kansainvälinen koulutusjärjestelmävertailu, joka tuottaa julkisesti saatavilla olevaa laajaa aineistoa koulutuksen tilasta ja tuloksista sekä koulun ulkopuolella tapahtuvasta oppimisesta. PISA 2015 -tutkimuksessa oppilaiden osaamista oli ensimmäisen kerran mahdollista arvioida tietokoneavusteisen kokeen perusteella. Suurin osa osallistujamaista, Suomi mukaan lukien, toteuttikin kokeen tietokonepohjaisesti. Tehtävät esitettiin verkkoympäristössä ja opiskelija suoritti kokeen selainpohjaisesti. Oppilaiden vastausten automaattinen koodaus tapahtui www-tekniikoita (HTML, CSS, Javascript, XML, XLIFF) hyödyntämällä (OECD 2015). Oppilaan koetilanteessa suorittamat toiminnot tallennettiin lokitiedostoon käyttäen seuraavien Javascript-komponenttien ominaisuuksia: lomake (*Form*), raahaa ja pudota (*Drag and Drop*), valinta (*Selection*) ja mukautettu (*Ad hoc*, ei-

standardin interaktion tallennuksen tuki). Näin saatu lokiaineisto sisälsi seuraavat yksittäisen tehtävän ratkaisuprosessia kuvaavat muuttujat (OECD 2015): tehtävän ratkaisuun käytetty aika, ratkaisemiseen liittyvien toimintojen (oppilaan ja tietokoneen väliset interaktiot, esim. hiirellä klikkaukset) lukumäärä sekä opiskelijan vastaus ja sille annettu pisteytys.

Lokianalyysi on yksi tapa ymmärtää ja analysoida oppimista tietokonepohjaisessa oppimisympäristössä (*computer-based learning environment, CBLE*), jota käytetään erityisesti yhteistoiminnallisen oppimisen piirissä (Gress, Fior, Hadwin & Winne 2010; Winne, Hadwin & Gress 2010). Oppimisen tavoitteet, oppimistrategiat ja oppimisympäristössä toimimisen profiilit liitetään tällöin usein oppilaan itseohjautuvuuteen (Schraw 2010). Kun näitä tekijöitä pyritään analysoimaan täsmällisesti, seurataan oppilaan ja tietokoneen välistä interaktiota aika-toimintoaikasarjana (*tracing*) tuottamalla havaittavia representaatioita muun muassa kognitiiviseen, metakognitiiviseen ja motivaatioon liittyvistä tapahtumista (Winne 2010). Tässä artikkelissa ei kuitenkaan käsitellä ongelmanratkaisulokeja näin suurella tarkkuudella, vaan analysoidaan PISA 2015:n matematiikan kokeesta tallennetun lokiaineiston yhteenvedoa.

Azevedon (2015) mukaan lokien avulla voidaan analysoida erityisesti oppilaiden kognitiivista sitoutumista koetehtävien ratkaisuun. Tämä on tekijä, jossa voi olla eroja perinteisen kynä-paperikokeen ja tietokonepohjaisen testin välillä. Perinteisessä koetilanteessa koevastaukset tuotetaan kullekin tehtävälle lineaarisesti, kun taas tietokonepohjaisessa ympäristössä oppilas voi toimia lähtökohtaisesti huomattavasti dynaamisemmin ja joustavammin. Interaktiivisessa tietokoneympäristössä korostuu esimerkiksi kokeilemisen helppous, joka voi edistää tutkivaa ongelmanratkaisua tai nopeaa pintapuolista kokeilemistä. Tästä syystä myös PISA-kontekstissa, jossa ollaan siirtymässä entistä laajemmin tietokonepohjaisiin arviointeihin, olisi tärkeää ymmärtää oppilaan ja tietokoneen välisiä interaktioprosesseja ja niiden suhdetta kokeessa osoitettuun osaamiseen. Tämän tutkimuksen tavoitteena onkin selvittää, millaisia opiskelija- ja tehtäväprofiileja voidaan tunnis-

taa PISA 2015:n tietokonepohjaisen kokeen lokiaineiston yhteen-
vetoa analysoimalla. Erityisesti tarkastellaan sitä, mitä tunnistetut
profiilit voivat kertoa tehtävien ratkaisustrategioista ja toisaalta op-
pilaiden kognitiivisesta sitoutumisesta tehtävien ratkaisemiseen.

Tutkimusmenetelmänä artikkelissa sovelletaan kartoitettavaa op-
pimisanalytiikkaa (*descriptive learning analytics*), erityisesti mah-
dollisimman pieniin ennakko-oletuksiin nojaavaa robustia klus-
terointia (Rousseuw & Leroy 1987; Hettmansperger & McKean
1998; Kärkkäinen & Heikkola 2004; Äyrämö 2006; Hämäläinen,
Jauhiainen & Kärkkäinen 2017). Artikkelin rakenne on seuraava.
Johdannon jälkeen tarkastellaan erityisesti laajojen havaintoai-
neistojen oppimisanalytiikkaa. Seuraavassa luvussa kuvataan käy-
tetyt tutkimusmenetelmät (tutkimuksen aineisto ja analyysimene-
telmät). Lopuksi esitellään tutkimuksen tulokset ja tiiviit johtopää-
tökset.

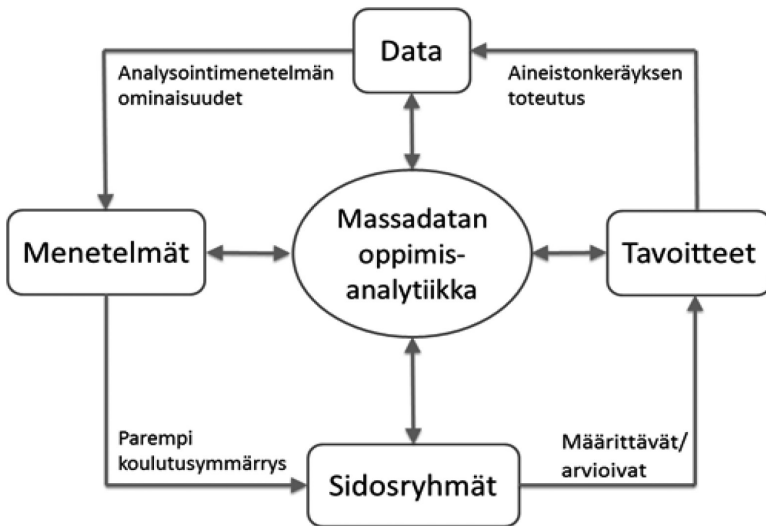
Laajojen havaintoaineistojen oppimisanalytiikka

Oppimisanalytiikka on jatkuvasti kasvava ja kehittyvä tieteenala
(Knight, Wise, Ochoa & Hershkovitz 2017), jossa sovelletaan mo-
nipuolisia kvantitatiivisia ja laskennallisia menetelmiä koulutusta
kuvaavien aineistojen tavoitteelliseen analyysiin. Siemens (2013)
ja Ferguson (2012) määrittelivät oppimisanalytiikan tarkoitukseksi
”mitata, kerätä, analysoida ja raportoida aineistoa oppijoista sekä
heidän ympäristöstään, tarkoituksena ymmärtää ja optimoida op-
pimista sekä tiloja missä oppimista tapahtuu”.

Yksi esimerkki koulutuksellisen massadatan lisääntymisestä
ovat ns. massiiviset avoimet verkkokurssit (*massive open online
course*, MOOC), joissa käyttäjien toimintaa verkkoympäristössä
voidaan tarkkailla, analysoida ja hyödyntää opetuksen suunnitte-
lussa (ks. esim. Kay, Reimann, Diepold & Kummerfel 2013; Reich,
Tingley, Leder-Luis, Roberts & Stewart 2014). Opetuksen suun-
nittelun ja oppimisprosessien adaptaation näkökulmasta tyypil-
lisiä haasteita ovat esimerkiksi taustaosaamiseltaan ja kulttuuri-
taustoiltaan heterogeeniset opiskelijaryhmät (esim. Liu ym. 2016).

Analysointimenetelmien näkökulmasta aineistojen koot ja erikoispiirteet, kuten esimerkiksi epästabiilit ja muuttuvat jakaumat sekä poikkeavat ja puuttuvat havainnot, edellyttävät syvällistä osaamista menetelmien valinnassa ja kehittämisessä (ks. Bergner, Colvin & Pritchard 2015; Bergner, Kerr & Pritchard 2015; Saarela 2017).

Oppimisanalytiikan päätavoitteena on siis tuottaa ja analysoida merkityksellisiä ja hyödyllisiä malleja oppimisympäristöstä saadun aineiston pohjalta. Saarela ja Kärkkäinen (2017) esittivät erityisesti laajoista koulutusvertailuista – kuten PISA – saatavan massadatan yleisen oppimisanalytiikan soveltamisen viitekehysten, jota on havainnollistettu kuviossa 1. Tämän artikkelin osalta yleinen konteksti eli kuvion 1 oikea alanurkka liittyy tavoitteisiin verrata ja ymmärtää erilaisia koulutusjärjestelmiä, jotta niiden tutkitusti toimivia piirteitä voitaisiin hyödyntää globaalisti. Astetta spesifimmin tarkastellaan tietokonepohjaisen kokeen ominaispiirteitä, jotta PISA-tutkimusta koordinoivat ja hyödyntävät tahot sekä



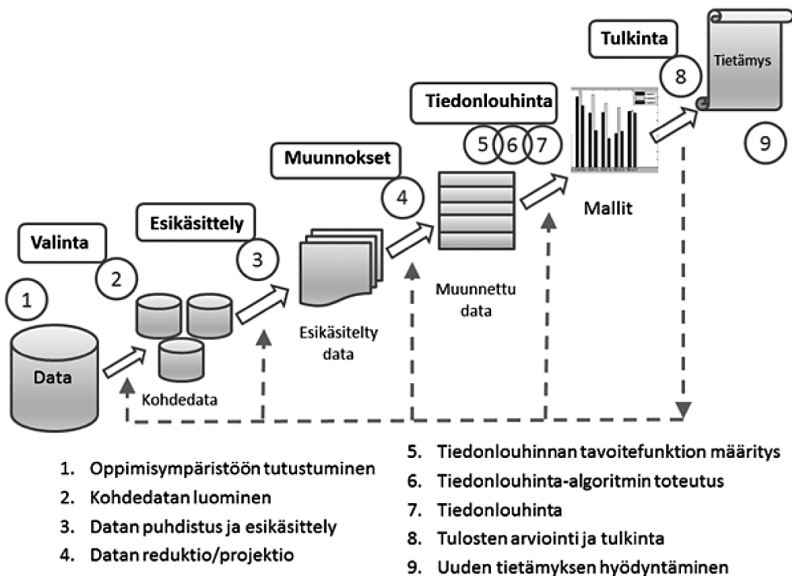
Kuvio 1. Massadatan oppimisanalytiikka. (ks. Saarela ja Kärkkäinen 2017)

koulutusteknologian käyttäjät ja kehittäjät voisivat paremmin ymmärtää saatuja tuloksia. Kuvion vasemmassa ylänurkassa korostuu analysointimenetelmien valinta suhteessa saatavilla olevan aineiston ominaisuuksiin sekä analytiikan yleisiin tavoitteisiin (ks. luku Tutkimusmenetelmät).

Chatti, Dyckhoff, Schroeder ja Thüs (2012) jaottelevat oppimisanalytiikassa käytettävät tekniikat seuraaviin kategorioihin: a) tilastolliset menetelmät, b) tiedon visualisoinnin menetelmät, c) tiedonlouhinnan menetelmät sekä d) sosiaalisten verkostojen analysointimenetelmät. Vaikka tiedonlouhinnan menetelmät muodostavat vain yhden kategorian tässä taksonomiassa, ne on yleisesti tunnustettu nopeimmin kasvaviksi ja monipuolisimmiksi oppimisanalytiikan tekniikoiksi, koska niiden avulla voidaan prosessoida suuria määriä strukturoimatonta aineistoa (Ferguson 2012; Rogers 2015). Olemassa olevien aineistojen lisääntyminen ja havaintomäärien kasvaminen on saanut oppimisanalytikot käyttämään perinteisten tilastollisten analyysimenetelmien lisäksi laajoille ja heterogeenisille aineistoille soveltuvia tiedonlouhintamenetelmiä (Ferguson 2012; Saarela & Kärkkäinen 2015b; Hershkovitz, Knight, Dawson, Jovanovic & Gasevic 2016; Joksimovic ym. 2016; Saarela & Kärkkäinen 2017).

Oppimisanalytiikassa pyritään huomioimaan monipuolisesti kaikkia niitä tekijöitä, jotka vaikuttavat oppimiseen tiettyssä oppimisympäristössä (Peña-Ayala 2017). Oppimisen luonteen paremmaksi ymmärtämiseksi aineistoa kerätään, muunnetaan ja esitetään tavalla, jolla erilaiset oppimisympäristöihin liittyvät sidoryhmät pääsisivät helposti käsiksi ja ymmärtäisivät oppimisanalytiikan tutkimukset ja tulokset (kuvio 1). Tämä voidaan saavuttaa esimerkiksi sillä, että analysoinnissa hyödynnetään eri mittaristoja ja tuloksia visualisoidaan erityyppisesti eri osapuolille, esimerkiksi opiskelijoille, opettajille, rehtoreille, vanhemmille tai koulutoimen johtajille (Ferguson & Shum 2012; Verbert, Duval, Klerkx, Govaerts & Santos 2013). Kuten Saarela ja Kärkkäinen (2017) osoittavat Suomen PISA-aineistoa hyödyntäen, tällaiset mittaristot voivat myös kuvastaa ja vetää yhteen merkityksellistä oppimistietoa suurelle yleisölle.

Oppimisanalyttisen tietämyksen tuottamista voidaan tukea hyödyntämällä yleistä tietämyksen hallinnan prosessia (Fayyad, Piatesky-Shapiro & Smyth 1996; Peña-Ayala 2017). Koulutukselliseen kontekstiin muokattuna tämä prosessi on kuvattu kuviossa 2. Kuvio sisältää eri vaiheet, joiden kautta oppimiseen liittyvää aineistoa prosessoidaan. Katkoviivoilla kuvataan prosessin iteratiivista luonnetta eli eri osioissa käytettävien menetelmien ja niiden tuotosten (joiden siirtymistä eteenpäin kuvataan paksuilla ylöspäin osoittavilla nuolilla) kehittymistä lopullisesti raportoitavaan muotoon. Kuvion 2 teknisesti keskeinen osa, tiedonlouhinta, viittaa lähestymistapaan (Hand, Mannila & Smyth 2001), jossa mielenkiintoiset ja käyttökelpoiset mallit nousevat aineistosta esiin ilman vahvojen alkuperäisten muodostamista. Tämä on suurin erottava tekijä tiedonlouhinnan ja perinteisen tilastollisen data-analyysin välillä. Tutkimusmenetelmä on täten lähtökohteisesti luonteeltaan kartoittava (Tukey 1977), mutta tämä ei sulje



Kuvio 2. Koulutuksellisen tietämyksen tuottamisen prosessi.

pois hypoteesitestausta esimerkiksi löydettyjen mallien piirteiden ja erojen analysoimiseksi. Kuten Saarela (2017, kuvio 5) kuvaa, ennustavat ja luokittelevat mallit, klusterointi ja relaatiotyypiset assosiaatiösäännöt ovat keskeisesti käytettyjä tiedonlouhintamenehtelmiä oppimisanalytiikassa.

Opiskelijoiden profilointi verkko-oppimisympäristöissä

Klusterointi eli yhtenäisten opiskelijajoukkojen profiilien määrittäminen on keskeinen tekniikka sekä oppimisanalytiikassa että muissa oppimisen aineistoja prosessoivissa tieteenaloissa (Saarela 2017). Yhteenveto suosituimmista oppimisanalyttisistä klusterointimenetelmistä ja niiden sovelluksista on esitetty Saarelan ja Kärkkäisen (2017) artikkelissa. Yleisimmin klusterointia sovelletaan haluttaessa mallintaa ja kuvata oppilaan toimintaa ja suoriutumista tietokonepohjaisessa oppimisympäristössä.

Klusterianalyysiä opetusympäristössä on sovellettu pääasiassa etsittäessä oppilaista erilaisia ryhmiä samankaltaisten tulosten mukaan, mutta myös tunnistamaan epäeettistä käyttäytymistä (ks. Romero & Ventura 2010). Toinen tapa soveltaa klusterointia opetuksessa on opiskelijoiden käyttäytymisen klusterointi verkko-oppimisympäristöissä. Tätä tapaa sovelletaan tässä tutkimuksessa. Verkko-oppimisjärjestelmissä käyttäjän kaikki klikkaukset voidaan tallentaa ja analysoida. Tämän seurauksena käyttäytymistä tällaisissa järjestelmissä on tutkittu laajasti. Esimerkiksi Desmarais ja Lemieux (2013) muodostivat klusteroimalla matematiikan ongelmanratkaisuympäristössä toimimisesta 1) tutkivan selailun, 2) lyhyiden harjoitussessioiden sekä 3) intensiivisen harjoittelun opiskeluprofiilit. Segedy, Kinnebrew ja Biswas (2015) puolestaan tiivistivät verkko-oppimisympäristön opiskelijoista kolme pääprofiilia: 1) usein toistuvan tutkimisen, 2) strategisen kokeilun, 3) sitoutuneen ja tuloksellisen oppimisen, sekä kaksi epäorientoitunutta ryhmää: 4) hämmentynyt arvailu ja 5) tehtävään sitoutumattomat. Verkko-oppimisympäristöä koskevissa tutkimuksissa

löydetyt opiskelijaprofiilit voidaan useimmiten jakaa karkeasti kahteen ryhmään: oppijoihin, jotka ajattelevat enemmän toimintaansa, ja oppijoihin, jotka soveltavat vähemmän oppimiseen suuntautunutta kokeilutapaa (Karavirta, Korhonen & Malmi 2006; Novotny 2004). Jälkimmäistä ryhmää pidetään usein opiskelijoina, jotka vain ”pelaavat järjestelmässä” (Baker, Corbett, Koedinger & Wagner 2004), mikä osoittaa, että nämä opiskelijat eivät luultavasti ole sitoutuneet oppimistehtäviin.

Aiemmat tutkimukset ovat osoittaneet, että klusterianalyysi voi tunnistaa tällaiset käyttäytymisryhmät verkko-oppimisympäristöissä. Esimerkiksi Hosseini, Brusilovsky, Yudelson ja Hellas (2017) klusteroivat opiskelijat ongelmanratkaisun perusteella käyttäen hierarkkista ja spektristä klusterointia. He löysivät kaksi opiskelijaklusteria: ”ajattelijat” (ts. oppilaat, jotka olivat rakentavampia) ja ”liikkujat” (ts. opiskelijat, jotka yrittivät kokeilla erilaisia strategioita lisäämättä tehtävien oikeellisuutta). Ei ole yllättävää, että ajattelijat suoriutuivat paremmin. Karavirta ym. (2006) löysivät viisi opiskelijaklusteria loppukokeiden pisteiden ja harjoitusten ratkaisemiseen käytetyn kokonaisajan perusteella. Klusterit olivat: ”läpipääsijät”, ”tavalliset”, ”iteraattorit”, ”kunnianhimoiset” ja ”lahjakkaat”. Läpipääsijät tekivät vähimmäistyön, joka tarvittiin kokeen suoritukseen. Tavalliset tähtäsivät korkeampaan arvosanaan kuin läpipääsijät mutta eivät tehneet odotettua enempää töitä. Iteraattorit käyttivät kokeilustrategiaa ja usein kaikkia heille tarjottuja yritysmahdollisuuksia, kun taas lahjakkaiden opiskelijoiden klusteri käytti vähiten yrityksiä ja saavutti silti korkeimmat tulokset. Kunnianhimoisten opiskelijoiden klusteri sai myös hyviä tuloksia, mutta he tarvitsivat enemmän yrityksiä kuin lahjakkaat opiskelijat.

Kizilcec, Piech ja Schneider (2013) käyttivät osittavaa klusterointia tutkiakseen opiskelijoiden sitoutumista MOOC-kurssilla. He löysivät neljä opiskelijaklusteria: ”suorittavat”, ”tarkistavat”, ”irtipäästävät” ja ”kokeilevat”. Kun suorittavat opiskelijat olivat tarpeeksi sitoutuneita yrittämään kaikkia tehtäviä, tarkistavat opiskelijat enimmäkseen vain katselivat videoluennot. Irtipäästävät opiskelijat sitoutuivat mukaan kurssin alussa mutta lopettivat sit-

ten tehtävien tekemisen. Kokeilevat opiskelijat sitoutuivat vain osiin kurssin alueista, esimerkiksi katsomalla vain valittuja videoita. Oppilaiden käyttäytymisprofiilien avulla voidaankin usein arvioida oppilaan sitoutumista kurssin tai tehtävän suorittamiseen ja siten koko oppimisprosessiin.

Tutkimusmenetelmät

Tämän tutkimuksen tavoitteena on pyrkiä tunnistamaan erilaisia opiskelija- ja tehtäväprofiileja PISA 2015:n tietokonepohjaisen kokeen lokiaineistoa analysoimalla. Seuraavassa kuvataan tarkemmin tätä tarkoitusta varten muokattu aineisto ja sen prosessoinnissa käytetyt oppimisanalysimenetelmät.

Aineisto

Aineistona toimi suomalaisten koululaisten vuoden 2015 matematiikan PISA-kokeesta tallennettu ja muokattu lokidata. Kaksi tietolähdettä – tehtäväkohtaiset (*item*) tiedot ja tehtävien vaikeutta kuvaavat osioanalyysin parametrit (OECD 2016) – yhdistämällä muodostettiin ensin yhdistetty aineisto, joka sisälsi kaikki yksittäisen tehtävän ratkaisutapahtumaa kuvaavat muuttujat. Havaintoyksikkönä tässä aineistossa on oppilas-tehtäväpari, toisin sanoen muuttujat kuvaavat oppilaan toimintaa tietyn matematiikan tehtävän ratkaisemisessa. Aineisto käsitti 2 426 oppilasta (ts. PISA-kokeeseen osallistuneista 5 882 oppilaasta vain n. 41 % teki matematiikan tehtäviä) ja vastauksia 50 eri tehtävään. Kukin oppilas sai ratkaistavaksi useita tehtäviä, ja sama tehtävä annettiin ratkaistavaksi usealle oppilaalle etukäteen suunnitellun tehtävärotaation mukaisesti. Kaikki oppilaat eivät kuitenkaan tehneet täsmälleen samoja tehtäväkokonaisuuksia eivätkä välttämättä yrittäneet tai ehtineet yrittää ratkaista kaikkia heille osoitettuja tehtäviä. Siten oppilaskohtaista aineistoa löytyi 1–17 ratkaistusta tehtävästä (ks. kuvio 8). Analysoitavassa aineistossa oli kaikkiaan 34 812 oppilas-tehtävähavaintoa. Lokitietojen ohella aineistoon yhdistettiin

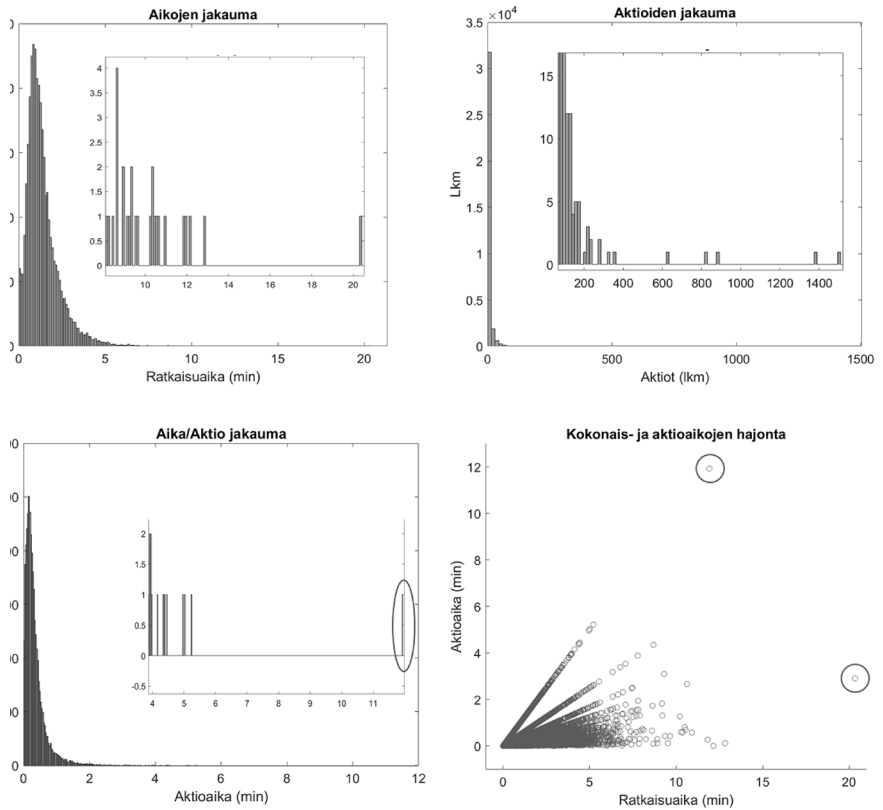
kullekin tehtävälle kansainvälisestä PISA-aineistosta osioanalyysillä estimoitu vaikeustaso kuvaava parametri.

Tarkasteltavat muuttujat olivat seuraavat:

- 1) Ratkaisuun kulunut kokonaisaika – *Aika* (alkuperäinen yksikkö millisekunteja, muunnettiin minuuteiksi).
- 2) Aktioiden eli opiskelijan tehtävän ratkaisun aikana suorittamien toimintojen (klikkaukset, tuplaklikkaukset, näppäimistön painallukset, raahaa-ja-pudota tapahtumat) lukumäärä – *Aktiot*.
- 3) *Aktioaika* eli aika jaettuna aktioiden määrällä: muuttuja kuvaa yksittäiseen aktioon keskimäärin kulunutta aikaa.
- 4) Pisteytetty vastaus – *Tulos*. Suurin osa vastauksista oli pisteytetty muotoon ”oikein” (1) tai ”väärin” (0), mutta joissakin tehtävissä oli myös vaihtoehto ”osittain oikein”. Tälle vaihtoehdolle annettiin lukuarvo 0,5.
- 5) Tehtävän vaikeustaso. Tämä on estimoitu kansainvälisestä PISA-aineistosta kaksiparametrinen Raschin mallin avulla. Estimoinnin suoritti yhdysvaltalainen ETS-tutkimuslaitos (Educational Testing Service), joka koordinoi PISA 2015 -aineiston keruuta ja toteutti datatiedostojen rakentamisen. Matematiikan tehtävien vaikeustasoestimaatit vaihtelivat välillä –1,805–1,705.

Muuttujien frekvenssijakaumia on havainnollistettu kuviossa 3. Lisäksi kuvio sisältää hajontakuvan ratkaisu- ja aktioaikojen yhteisistä arvoista. Havaintojoukosta poistettiin kolme selkeästi poikkeavaa havaintoa: yksi, jossa aktioiden määrä oli nolla; yksi, jossa ratkaisuaika oli yli 20 minuuttia sekä yksi, jossa aktioaika oli yli 10 minuuttia (kahta jälkimmäistä on havainnollistettu kuvion 3 alaosassa). Analysoitujen havaintoyksiköiden lukumäärä oli siis 34 809.

Aineiston läpikäynnin yhteydessä tallennettiin erikseen myös tehtävän vaikeustaso sekä suomalaisten oppilaiden oikeiden vastusten osuus. Yleisesti PISA-tutkimuksissa oppilaan osaaminen (*proficiency*) estimoidaan tilastollisella osioanalyysimallilla, jossa on keskeistä estimoida ja ottaa huomioon oppilaalle osoitettujen tehtävien vaikeustasot. Oppilaan osaaminen on latentti suure, jota PISA-

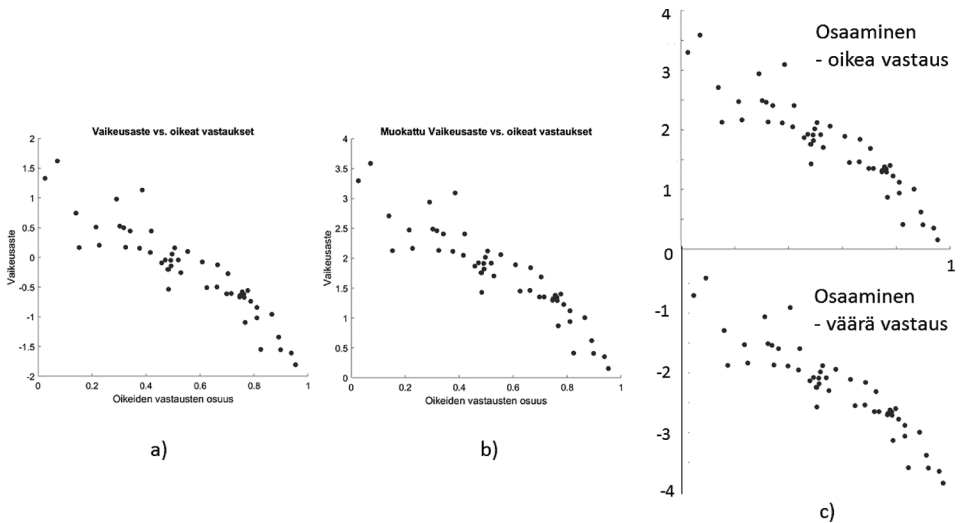


Kuvio 3. Ratkaisuaikojen, aktioiden lukumäärien sekä aktioiden yksikköaikojen jakaumat sisältäen suurien arvojen zoomaukset. Alimpana oikealla on aikojen ja yksikköaktioaikojen hajontakuvi, jossa on havainnollistettu poistettuja havaintoja.

tutkimuksissa arvioidaan niin sanotuilla todennäköisillä tuloksilla (*plausible values*). Nämä ovat realisaatioita oppilaan ”arvioitua osaamistasoa” kuvaavasta posteriorijakaumasta, joka määrätään bayesiläisen tilastotieteen keinoin yhdistämällä havaintoaineistosta saatu oppilaan taustamuuttujia, koetuloksia ja tehtyjen tehtävien vaikeustasoa koskeva informaatio oppilaiden (latentin) osaamisen kansallista vaihtelua kuvaavaan priorijakaumaan (ks. esim.

Rutkowski, Gonzalez, Joncas & von Davier 2010; Saarela 2017). Tässä tutkimuksessa kuitenkin rajoitetaan pelkästään tietokoneympäristössä suoritettuja tehtäviä kuvaavan aineiston analysointiin. Siksi oppilaiden tehtävissä osoittamalle osaamiselle johdettiin yksinkertaisempi muuttuja, joka huomioi sekä yksittäisen oppilaan vastauksen oikeellisuuden että tehtävän vaikeustason.

Oppilaan yksittäisessä tehtävässä osoittamaa osaamista kuvaavan muuttujan rakentamista havainnollistetaan kuviossa 4. Kuvion kaikkien kolmen graafin x-akselina toimii kullekin tehtävällä annettujen oikeiden vastausten osuus (0 = ei yhtään oikeaa vastausta, 1 = kaikki vastaukset oikein). Aluksi tehtävän vaikeusaste (y-akseli kuvioissa 4 a ja 4 b muunnettiin positiiviseksi siten, että arvo nolla vastaa maksimaalisen helppoa tehtävää, jonka kaikki oppilaat osaisivat ratkaista. Tätä varten kuvion 4 a pisteiden perusteella

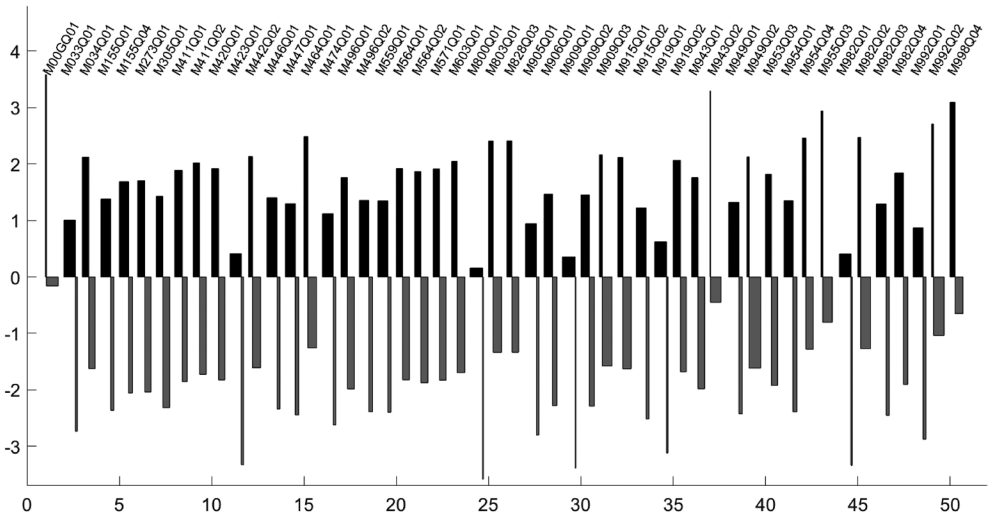


Kuvio 4. Tehtävässä osoitetun osaamisen estimointi suhteessa oikeiden vastausten osuuteen (x-akseli). a) Alkuperäinen tehtävän vaikeusaste (y-akseli) suhteessa oikeiden vastausten osuuteen, b) vaikeustason (y-akseli) muuntaminen positiiviseksi, c) oikeiden ja väärin vastausten ilmentämän osaamisen määrä (y-akseli) suhteessa oikeiden vastausten osuuteen.

määriteltiin lineaarinen regressiosuora, joka saa arvon $-1,9627$ pisteessä $x = 1$ (eli kun oikeiden osuus on 1). Tämän jälkeen alkupe-
räisistä vaikeustasoista vähennettiin kyseinen arvo, jolloin päädyttiin kuvion 4 tilanteeseen b. Jos oppilas oli osannut ratkaista tehtävän oikein, valittiin kyseinen vaikeusaste arvo esittämään osoitettua osaamista (ks. kuvio 4 c yläosa). Vastaavasti jos oppilas ratkaisi tehtävän väärin, osoitetun osaamattomuuden arvoksi valittiin negatiivinen, ”käänteinen” (suhteessa kuvion 4 y-akseliin) vaikeustaso (ks. kuvio 4 c alaosa). Eli helppoon, paljon osattuun tehtävään väärin vastaaminen osoitti suurta negatiivista osaamista ja puolestaan vaikeaan tehtävään väärin vastaaminen osoitti vain lievää negatiivista osaamista. Matemaattisesti väärin vastatun muuttujan arvoksi siis valittiin Tätä vaikeustasoparametrien avulla johdettua oppilaan osaamista mittaavaa muuttujaa kutsutaan jatkossa nimellä *Osaaminen*.

Tehtäviin väärin vastanneiden oppilaiden ”osaamattomuus” (negatiivinen osaaminen) saatiin siis lineaarisella muunnoksella tehtävän vaikeustasosta ja siihen oikein vastanneiden osuudesta. Koska tehtävän vaikeustaso ilmentää kansainvälisellä tasolla osoitettua suoriutumista, voidaan suomalaisten oppilaiden tuloksia verrata kansainväliseen aineistoon tarkastelemalla oikeiden vastausten ja osaamista mittaavan muuttujan välisiä suhteita. Tässä suhteessa jo kuvio 4 a on mielenkiintoinen, sillä sen tulisi ilmentää täsmällistä negatiivista korrelaatiota (laskevaa suoraa), jos suomalaiset opiskelijat olisivat matematiikan tehtäviä ratkoessaan osoittaneet täsmälleen samanlaista osaamista kuin koko kansainvälinen populaatio. Käyrän lievä kaareutuminen alaspäin liikuttaessa oikeiden vastausten osuudesta 1 kohti osuutta 0 antaisi ymmärtää, että suomalaisten koululaisten suhteellinen osaaminen on ollut kansainvälistä keskitasoa parempaa helpommille tehtäville mutta huonontunut keskimääräistä nopeammin tehtävien vaikeutuessa.

Tarkennetaan tulkintaa suomalaisten oppilaiden osoittamasta osaamisesta suhteessa tehtävän yleiseen vaikeustasoon eli koko kansainväliseen oppilaspopulaatioon. Kuviossa 5 havainnollistetaan täsmälleen samaa aineistoa kuin kuviossa 4 a siten, että pylvään paksuus esittää oikeiden (musta) ja väärin (harmaa) vastausten



Kuvio 5. Matematiikan PISA-tehtävissä osoitettu osaaminen.

suhteellisia osuuksia ja pylvään pituus tehtävässä osoitettua osaamista. Tässäkin kuviossa koko PISA-tutkimuksen mukainen osaaminen ilmeni tasaisesti ohenevina pylväinä niiden pituuden kasvessa sekä positiivisen että negatiivisen y-akselin suuntaan. Näin ei kuitenkaan ole. Esimerkiksi tehtävissä M305Q01 (tehtävä n:o 7 vaaka-akselilla) ja M949Q02 (tehtävä n:o 39) on tavanomaista suurempi määrä vääriä vastauksia ja puolestaan tehtävässä M998Q04 (tehtävä n:o 50) selvästi suurempi määrä oikeita vastauksia kuin mitä kansainvälisestä aineistosta estimoidut osaamistasot ennakoisivat.

Menetelmät

Kuten edellinen luku esittää, tutkimuksen aineisto muodostuu kymmenistä tuhansista havainnoista, jotka ovat luonteeltaan diskreettejä (aika, aktiot, kuinka monta oikeaa vastausta, vaikeustaso, osaaminen). Tutkimuksen tavoitteena on pyrkiä löytämään kiinnostavia profiileja tästä aineistosta. Menetelmällisesti kyseessä on siis ns. ohjaamattoman oppimisen (*unsupervised learning*).

ning) tehtävä (Hand ym. 2001). Klusterointi onkin käytetyin oppimisanalytiikan menetelmä, jolla etsitään aineiston sisältämiä piileviä ryppäitä ja määritetään näiden profiileja (ed. luku; Saarela & Kärkkäinen 2017). Sen lisäksi, että tässä tutkimuksessa klusteroitavat muuttujat ovat diskreettejä eivätkä siten täytä oletusta normaalijakautuneesta satunnaismuuttujasta, ovat niiden jakaumat Pareto-tyyppisiä ja hyvin vinoja kuvion 3 esittämällä tavalla. Tästä syystä käytettävän klusterointimenetelmän tulee voida käsitellä erityyppisiä jakaumia ja perustua mahdollisimman pieniin lähtöoletuksiin aineistosta ja sen sisältämien muuttujien jakaumista. Kuten Saarela ja Kärkkäinen (2015a) esittävät, tämäntyyppisten kvantisoitujen muuttujien virhejakauman lähtöoletus onkin tasainen jakauma.

Robusteja, ei-parametrisia menetelmiä käytetään tilastotieteen piirissä silloin, kun data-analyysi perustuu ei-normaalijakautuneisiin lähtöoletuksiin (ks. Rousseeuw & Leroy 1987; Hettmansperger & McKean 1998; Kärkkäinen & Heikkola 2004). Tällöin kutakin muodostettua datan osajoukkoa, klusteria, edustaa jokin muu tunnusluku kuin tyypillisesti normaalijakaumaoletukseen perustuva keskiarvo. Yksinkertaisimpia robusteja ei-parametrisia tunnuslukuja ovat mediaani ja spatiaalinen mediaani, joista mediaani perustuu reunajakaumiin eli on luonteeltaan yksikulotteinen kun taas spatiaalinen mediaani on aidosti monimuuttujainen (Rousseeuw & Leroy 1987; Hettmansperger & McKean 1998; Kärkkäinen & Heikkola 2004).

Tässä tutkimuksessa käytettiin klusterointialgoritmina K-SpatialiMediaani++-menetelmää (Hämäläinen ym. 2017), joka perustuu kutakin klusteria edustavan spatiaalisen mediaanin laskemiseen peräkkäisten ylirelaksointien avulla (Kärkkäinen & Äyrämö 2004, 2005). Laskennat toteutettiin Matlab-ohjelmistolla. Muuttujille *Aika* ja *Aktioaika* suoritettiin ennen klusterointia muunnos ja muuttujalle *Aktiot* muunnos. Näiden muunnosten avulla muuttujien arvojakauma tasoittuu ja tulee helpommin eroteltavaksi (Baeza-Yates & Ribeiro-Neto 1999). Ohjaamattoman oppimisen skenaarion mukaisesti meidän täytyy pystyä estimoimaan klustereiden sisältöjen lisäksi myös niiden lukumäärä. Tähän käytetään ns. klusteri-

validointi-indeksejä, joista suosituimpia on laajasti ja perinpohjaisesti vertailtu useissa artikkeleissa (Jauhainen & Kärkkäinen 2017; Hämäläinen ym. 2017). Myöskään näissä ns. sisäisissä klusterivalidointi-indekseissä ei tehdä mitään ennako-oletuksia klustereiden lukumäärästä. Käytetyt indeksit olivat Pakhiran, Bandyopadhyayn ja Maulikin PBM-indeksi, Silhouetti-indeksi, WB-indeksi, k*-klusterointivirhe sekä Wemmert-Gançarskin indeksi. Lisäksi käytettiin erityyppisiä tiedon ja mallien visualisointitekniikoita tulkintojen ja johtopäätösten tekemisen helpottamiseksi.

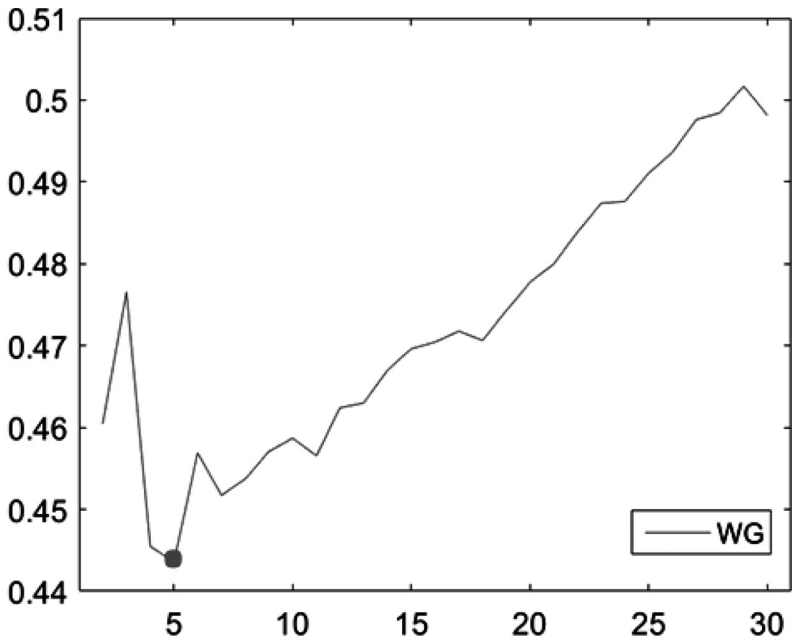
Tulokset

Tässä luvussa kuvataan kahden eri klusteroinnin toteutukset, tulokset ja tulosten tulkinnat.

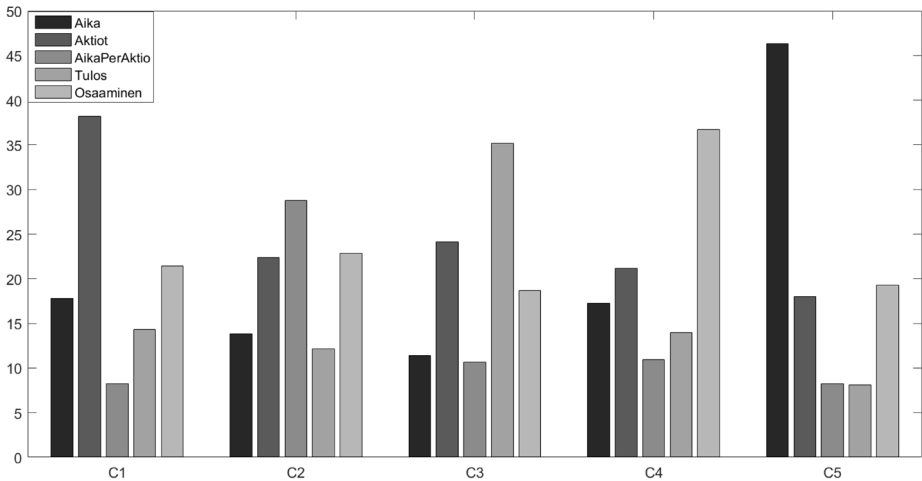
Koko aineiston klusterointi

Ensimmäisessä vaiheessa K-SpatiaaliMediaani++-klusterointimenetelmää sovellettiin 34 809 havainnon muokattuun aineistoon, jossa yksittäinen havainto esitti opiskelija-tehtäväparia. Tällöin esimerkiksi saman opiskelijan vastaukset 1–17 eri tehtävään (kuvio 9) oli irrotettu toisistaan. Wemmert-Gançarskin klusterivalidointi-indeksin ehdotuksesta (kuvio 6) aineisto jakautui viiteen eri klusteriin C1–C5 (PBM- ja Silhouette-indeksit ehdottivat kolmea klusteria, mutta valitsimme viisi, jotta pystyimme analysoimaan hie-man useampaa ratkaisuprofiilia).

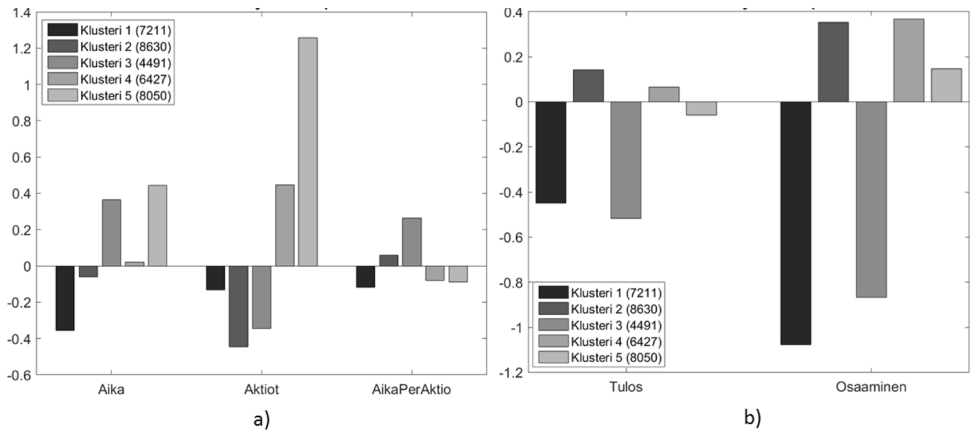
Klustereiden muuttujaprofiileja eli klusteriin kuuluvan aineiston spatiaalista mediaania on havainnollistettu kuvioissa 7 ja 8. Klustereiden koot ja prosenttiosuudet koko havaintoaineistossa olivat 7 211 (21 %), 8 630 (25 %), 4 491 (13 %), 6 427 (18 %) ja 8 050 (23 %).



Kuvio 6. Wemert-Gancarskyn klusterivalidointi-indeksin arvot havaintoaineistolle.



Kuvio 7. Viiden saadun klusterin muuttujaprofiilit.



Kuvio 8. Klusterien profiilit muuttujittain. Viiden klusterin poikkeamat koko aineiston yleisestä käyttäytymisestä (spatiaalinen mediaani) muuttujille *Aika*, *Aktio* ja *Aktioaika* (a) sekä muuttujille *Tulos* ja *Osaaminen* (b).

Viiden eri klusterin tulkinta erityisesti kuvion 8 pohjalta on seuraava:

- Klusteri 1: ”nopeasti ratkaistut, paljon erityisesti helpompia väärin”. Sisältää keskimääräistä nopeammin ratkaistuja tehtäviä, jotka menevät paljon väärin erityisesti helpommilla tehtävillä (*Osaaminen* merkittävästi pienempi kuin *Tulos*). On ajateltavissa, että tässä havaintojoukossa tehtäviin ja niiden ratkaisuun ei ole paneuduttu.
- Klusteri 2: ”vähemmällä aktioilla hyvää osaamista”. Klusterista löytyy hyvin erityisesti vaikeammista tehtävistä (*Osaaminen* suurempi kuin *Tulos*) suoriutumista. Koska erityisesti aktioiden määrä on koko joukon pienin, on tässä joukossa mietitty ja paneuduttu ratkaisemiseen.
- Klusteri 3: ”paljon harkintaa mutta silti virheitä”. Tässä klusterissa ratkaisemiseen on käytetty tyypillistä enemmän aikaa ja erityisesti yksittäistä aktiota on harkittu eniten. Tulosten perusteella tästä ei ole kuitenkaan seurannut erityisesti haastavampien tehtävien oikeita ratkaisuja.

- Klusteri 4: ”enemmän aktioita ja hyvää osaamista”. Mielenkiintoinen klusteri, jossa keskimääräistä suurempi interaktioiden eli tutkivan ongelmanratkaisun määrä on johtanut parhaaseen kokonaisosaamiseen.
- Klusteri 5: ”työläimmät interaktiotehtävät, joissa vaikeammissa osaamista”. Havainnot, joissa on käytetty eniten aikaa ja muita ryhmiä merkittävästi enemmän aktioita tehtävien ratkaisuyrityksissä. Tämä onkin johtanut keskimääräistä hieman parempaan osaamiseen eli onnistumisiin erityisesti hieman vaikeammissa tehtävissä.

Koko aineiston klusteroinnin tulkinta

PISA-tutkimuksissa tallennettu lokiaineisto, jossa yksi havaintopiste kuvaa yhden oppilaan yhden tehtävän ratkaisemisyritystä, on tulkinnalliselta kannalta hankala klusteroitava. Klusterien muodostumiseen vaikuttavat samanaikaisesti sekä oppilaat että heille ratkaistavaksi annettujen tehtävien luonne, ja tämä kaksiulotteisuus tekee edellä saaduista klustereista vaikeasti tulkittavia. Tehtävä, jossa lyhyen tekstin perusteella valitaan yksi neljästä vastausvaihtoehdosta, vaatii lähtökohtaisesti oppilaalta vähemmän aikaa ja aktioita kuin tehtävä, jossa annetaan taustatietoina enemmän tekstiä tai kuvia ja oppilas joutuu esimerkiksi kokeilemaan erilaisia tekijäkombinaatioita lopputulokseen pääsemiseksi. Lisäksi klusterien muodostumiseen vaikuttaa se, miten oppilaat ovat asennoituneet erilaisiin tehtäviin ja kokeen tekemiseen yleensä. Esimerkiksi klusterissa 1 on voinut ilmetä huolimattomuusvirheitä tai tehtävää ei ole täysin ymmärretty tai oppilas ei ole ollut täysin motivoitunut yrittämään parastaan. Klustereiden 2 ja 4 ominaispiirteet voivat taas johtua joko tehtävien lähtökohtaisesti vähäisemmästä tai suuremmasta interaktiovaatimuksesta tai oppilaan tavasta paneutua tehtävien ratkaisemiseen. Klusteri 3 vaikuttaa hyvin vahvasti sisältävän ongelmanratkaisuyrityksiä, jotka yksinkertaisesti ovat olleet oppilaalle liian vaativia. Myös klusterin 5 suuri aktioiden määrä voi johtua joko tehtävien luonteesta tai oppilaan tavasta pyrkiä ratkaisemaan niitä.

Muokatun aineiston klusterointi

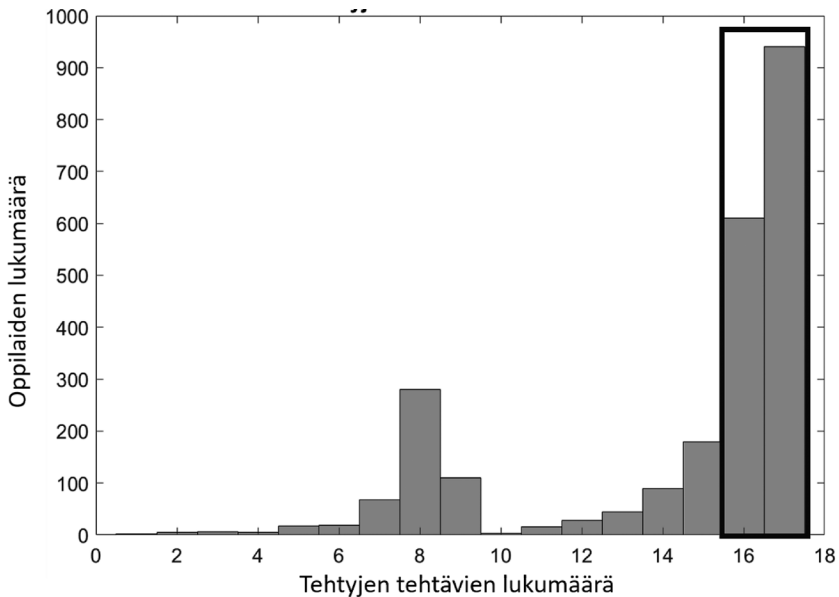
Edellä saatujen klustereiden luonnetta on helpompi ymmärtää, jos niiden muuttujaprofiileja katsotaan ”ulottuvuus kerrallaan”, toisin sanoen erikseen oppilaiden ja tehtävien näkökulmista. Seuraavaksi tarkastellaankin sitä, mitä edellä kuvatut muuttujaprofiilit voivat kertoa toisaalta matematiikan tehtävistä ja toisaalta niitä ratkaisseista oppilaista. Tätä varten muodostettiin kaksi uutta havaintomatriisia, joissa muuttujina toimi viisi eri ratkaisuprofiilia. Havainnot muodostuivat puolestaan joko 50 tehtävästä (tehtävämatriisi) tai 2426 oppilaasta (oppilasmatriisi). Matriisit muokattiin vielä yhtenäiseen muotoon laskemalla kunkin tehtävän tai oppilaan prosenttiosuudet kuhunkin ratkaisuprofiiliklusteriin kuulumiselle.

Aluksi tehtävämatriisista etsittiin sellaisia tehtäviä, joiden vastausyryksistä yli puolet (eli yli kaksinkertainen määrä verrattuna klustereiden perusosuuksiin, joiden maksimi oli 25 %) kuuluvat vain tiettyyn ratkaisuprofiiliin. Tällöin kyseinen ratkaisuprofiili edustaa enemmistöä tehtävän ratkaisuyryksistä. Tehtäväkohtaisen analyysin perusteella saimme tällä tavalla profiloitua vain 15 tehtävän (30 % kaikista 50 tehtävästä) luonteen suhteessa viiteen ratkaisuprofiiliin. Lukumääräisesti korostuivat toinen (kuusi tehtävää) ja viides (viisi tehtävää) ratkaisuprofiili, jotka ilmensivät pienintä ja suurinta oppilaan ja tietokoneen välisten interaktioiden määrää. Kokonaisuutena se, että suurin osa tehtävistä ei sijoittunut selvästi mihinkään klusteriin, kuvaa sitä, että saadut viisi klusteria profiloivat enemmän oppilaita kuin tehtäviä. Tämä on sikäli ymmärrettävää, että tehtävärotaatiosta huolimatta PISA-koeasetelmassa pyritään siihen, että kaikille oppilaille tulisi ratkaistavaksi mahdollisimman monipuolinen tehtäväkokoelma, jossa on sekä paljon että vähän aikaa ja aktioita vaativia tehtäviä. Se, että sama tehtävä voi sijoittua useaan klusteriin yhtä aikaa, kertookin siis siitä, että sitä ovat ratkaisseet toimintatavoiltaan erilaiset oppilaat: yksi oppilas on voinut antaa vastauksen nopeasti tai huolimattomasti, kun taas toinen oppilas on voinut miettiä kauan, selailla tehtävää edestakaisin tai muuttaa jo kerran antamaansa vastausta. Se, että klusterit profiloivat selvemmin oppilaita kuin tehtäviä, viittaa myös siihen, että tietyn toimintatavan ja motivaation omaava oppilas on toimi-

nut kokeessa suurin piirtein samalla tavalla (esim. käyttänyt paljon aikaa) tehtävästä riippumatta.

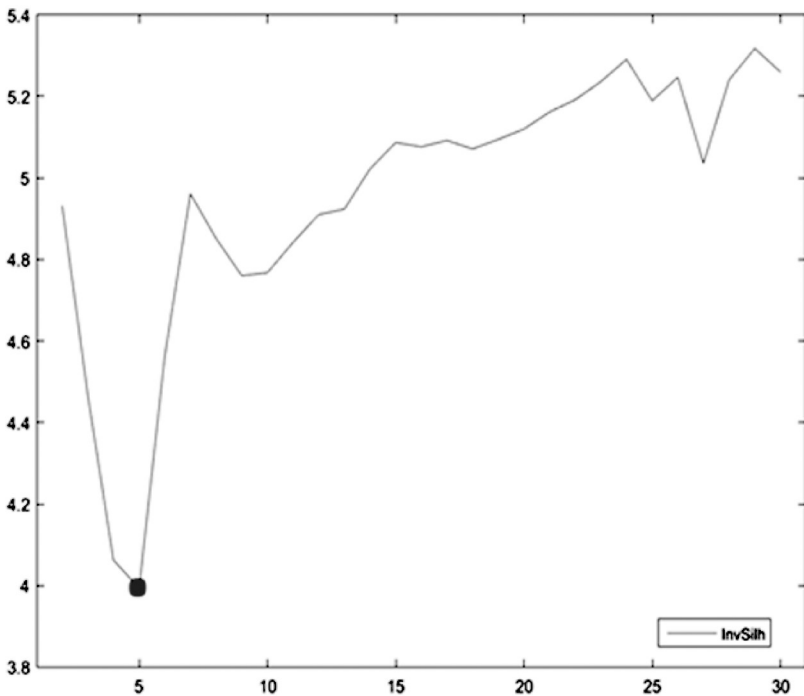
Oman pohdintansa tuloksiin tuovat tehtävien ominaisuudet sekä se, kuinka PISA-kokeen tapahtumalaskuri laskee kuhunkin tehtävään käytettyjen toimintojen (action) määrän. Jos esimerkiksi tehtävässä käytetään testin omaa laskinta ja laskimen käytössä tapahtuvat painallukset lasketaan tapahtumiksi lokitiedostoon, nosta se helposti tiettyjen tehtävien tapahtumien määrää oppilasta kohden. Myös tehtävien rotaatiosta johtuva tehtävien järjestys kokeessa voi muuttaa vastaajien motivaatiota ja tarkkaavaisuutta tehtävien tekemisessä ja näin ollen nostaa tai laskea ajankäyttöä tai toimintojen määrää tehtävien teossa.

Koska tehtävämatriisin klusterointi ei osoittautunut hedelmälliseksi, seuraavassa keskitytään oppilasmatriisin klusterointiin. Oppilasaineistossa oli melko paljon oppilaita, jotka olivat tehneet vain vähän matematiikan tehtäviä. Kuten kuvio 9 esittää, havainto-



Kuvio 9. Oppilaiden tekemien tehtävien jakauma (kaksi oppilasta oli tehnyt vain yhden tehtävän, 941 oppilasta maksimimäärän eli 17 tehtävää).

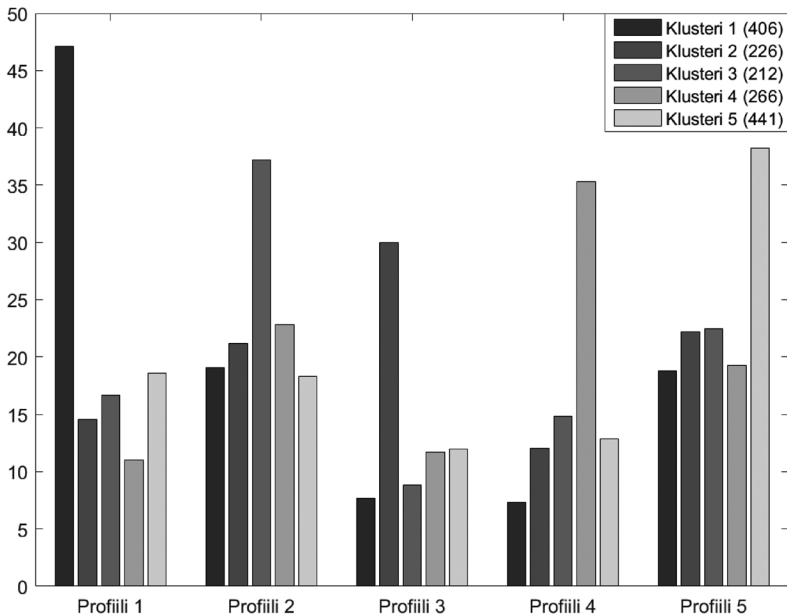
joukosta löytyi yhteensä 1551 oppilasta, jotka olivat tehneet vähintään 16 eli yhtä vaille maksimimäärän (17) yhdelle oppilaalle annettuja tehtäviä. Valitsimme tämän osajoukon oppilaista tarkempaan analyysiin, sillä nämä opiskelijat kuvaavat luotettavimmin koko kokeen tilannetta. Sovelsimme samaa K-SpatiaaliMediaani++-menetelmää aineistoon, jossa muuttujina toimivat opiskelijan tekemien tehtävien prosenttiosuudet viidessä eri ratkaisuprofiiliklusterissa. Silhouette-indeksi (kuvio 10) ehdotti viittä eri opiskelijaprofiilia ja PBM-indeksi tuki tätä valintaa, joten valitsimme klustereiden lukumääräksi viisi. Viiden syntyneen oppilasklusterin koot ja prosenttiosuudet aineistossa olivat 406 (26 %), 226 (15 %), 212 (14 %), 266 (17 %) sekä 441 (28 %).



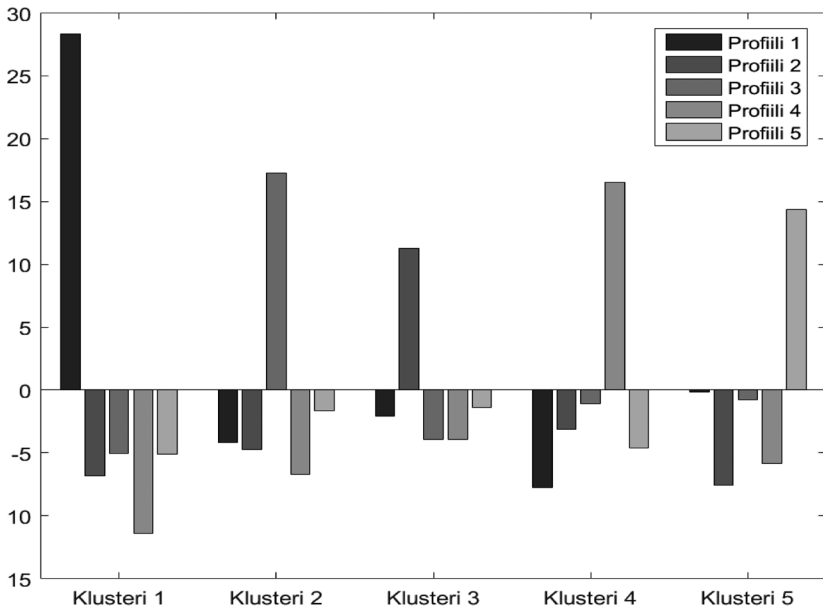
Kuvio 10. Käänteisen Silhouette-klusterivalidointi-indeksin arvot opiskelijoiden klustereihin kuulumisdatalle.

Opiskelijaklusterien tulkinta

Viiden opiskelijaklusterin prototyyppinä, jotka siis edustavat prosenttiosuuksia eri tehtäväratkaisuprofiileihin kuulumiselle, on havainnollistettu kuvioissa 11 ja 12. Kuvioista 11, joka esittää eri ratkaisuprofiiliryhmiin kuulumisosuudet klustereittain, huomataan, että kaikissa muodostuneissa oppilasryhmissä on havaintoja kaikista viidestä eri ratkaisuprofiilista. Siten kaikki oppilaat ovat koheen kuluessa käyttäneet useampia erilaisia työskentelytapoja (vähän/paljon aikaa, vähän/paljon aktioita, vähän/paljon aikaa aktiota kohti). Toisaalta kuvioista 12, jossa kuvataan eroja klustereittain koko aineiston ratkaisuprofiiliosuuksiin, ilmenee selkeästi, että kaikissa viidessä eri oppilasryhmässä on eri ratkaisuprofiili dominoivana. Näin ollen, vaikka oppilas sai ratkaistavakseen erityyppisiä tehtäviä, jotka vaativat erilaista työskentelyä (esim. paljon/vähän aktioita), hänellä oli lokitetietojen perusteella yksi työskentelytapa, joka oli muita tapoja vallitsevampi.



Kuvio 11. Opiskelijaklustereiden prototyytit muuttujittain.



Kuvio 12. Opiskelijaklustereiden prototyytit klustereittain poikkeamina tyypillisestä käyttäytymisestä (spatiaalinen mediaani).

406 opiskelijan klusterissa 1, johon kuului 26 prosenttia oppilaista, korostuu ratkaisuprofiili 1 eli toimintaa luonnehtii tulkinta ”nopeasti ratkaistut – paljon helppoja tehtäviä väärin”, joka viittaa alentuneeseen motivaatioon kokeessa. Toisessa, 226 oppilaan klusterissa (15 % oppilaista) ilmenee ”paljon harkintaa, mutta silti virheitä” – heille tehtävät ovat olleet vaikeita. Klusterin 3 (14 % oppilaista) oppilaat osoittivat ”vähemmällä interaktiolla hyvää osaamista” eli kykenivät rauhallisella ja harkitsevalla otteella ja tarvittaessa aikaa käyttämällä ratkaisemaan tehtäviä. Klusteriin 4 (17 % oppilaista) kuuluneet osoittivat ”enemmän interaktiota ja hyvää osaamista” eli saivat todennäköisesti tutkivalla ja kokeilevalla ongelmanratkaisuoitteella ja mahdollisesti jo kerran antamia vastauksia korjaamalla paljon oikeita vastauksia. Viimeinen ja suurin (28 %) klusteri on hankalimmin karakterisoitavissa. Sitä luonnehtii tulkinta ”työläät interaktiotehtävät, joista vaikeimmissa osaa-

mista”. Klusteri sisältää oppilaita, jotka uutteralla työllä onnistuivat ratkaisemaan vaikeampia tehtäviä.

Yhteenvedona voidaan todeta, että lähes maksimimäärän tehtäviä tehneiden oppilaiden joukossa on selvästi nähtävissä viiden erilaisen ratkaisuprofiilin mukaista käyttäytymistä, mutta yleiset profiilit eivät sellaisenaan kokonaisuudessaan kuvaa oppilaiden ratkaisuorientaatiota, sillä luonnollisesti myös oppilaille osoitettujen tehtävien luonne vaikuttaa sekä oppilaiden ongelmanratkaisukäyttäytymiseen että osoitettuihin oppimistuloksiin. Selvimmin tulkittavissa oli ensimmäinen klusteri, johon kuului 26 prosenttia matematiikan kokeen tehneistä oppilaista. Siihen kuuluneet oppilaat olivat lokitietojen perusteella tehneet merkittävän osan tehtävistä verraten nopeasti ja heikoin tuloksin. Voidaan arvioida, että noin neljäsosa kokeeseen osallistuneista suomalaisoppilaista on ollut melko heikosti motivoitunut ja haluton näkemään vaivaa matematiikan PISA-tehtävien ratkaisemisessa.

Johtopäätöksiä

Artikkelissa etsittiin matematiikan tehtävien ratkaisuprofiileja ja -strategioita tietokonepohjaisessa PISA 2015 -kokeessa suomalaisten koululaisten aineiston perusteella. Kokonaisuudessaan tutkimus ilmensi kuvion 1 mukaista tutkimusprosessia, jossa keskeistä oli valita tutkimusmenetelmät käytettävissä olleen ja muokatun aineiston ominaisuuksien perusteella. Tuloksia varten tarvittiin alkuperäisen aineiston valintaa (suomalaisten oppilaiden matematiikan tietokonepohjaiset tulokset, tehtävien vaikeusestimaatit), esikäsittelyä (poikkeavien havaintojen poisto, logaritminen skaalaus) ja muunnoksia (uusi osaamista mittaava muuttuja, tehtävä- ja opiskelijamatriisit) koulutuksellisen tietämyksen tuottamisen prosessin (kuvio 2) mukaisesti. Tutkimus ei perustunut hypoteesitestaukseen vaan tietämyksen tuottamiseen (Fayyad ym. 1996) eksploraatiivisen data-analyysin (Tukey 1977) ja robustien tilastollisten menetelmien ja tunnuslukujen avulla (Kruskal & Wallis 1952; Rousseeuw

& Leroy 1987; Hettmansperger & McKean 1998; Kärkkäinen & Heikkola 2004).

Matematiikan kokeesta verkkoympäristössä tallennetuista loki-tiedoista löytyi viisi erilaista ratkaisuprofiilia, joiden tulokset pystyttiin – erityisesti artikkelissa kehitetyn osaamisen mittarin avulla – liittämään erityyppisiin ratkaisustrategioihin ja konteksteihin (selvimmin siihen, kuinka motivoituneesti ratkaisuun oli pyritty). Vaikka havaintoyksikköinä toimivat oppilas-tehtäväparit, pystyttiin toistetulla klusteroinnilla liittämään kokeeseen osallistuneet oppilaat hyvin vahvasti eri ratkaisuprofiileihin. Tulokset muistuttivat tältä osin Azevedon (2015) esittämän lokianalyysin tuloksia oppilaiden kognitiivisesta sitoutumisesta ongelmanratkaisuun sekä Hosseinin ym. (2017) ja Karavirran ym. (2006) tuloksia oppilaiden erilaisista sitoutumisen tasoista opiskeluun. Tätä ilmensivät erityisesti tutkivaa ongelmanratkaisua ja nopeaa pintapuolista läpikäyntiä edustaneet klusteriprofiilit. Saadut profiilit olivat luonteeltaan samantyyppisiä kuin aikaisemmissa tutkimuksissa esitetty (Desmarais & Lemieux 2013; Segedy ym. 2015). Ne osoittivat, kuinka sitoutuneesti oppilas toimi ongelmanratkaisutilanteessa; tämä on tekijä, jonka avulla voidaan muun muassa arvioida oppilaiden yleistä suoriutumista opinnoista. Löydettyjen profiilien pohjalta olisikin mahdollista ennakoita ja seuloa oppilaiden toimintaa verkko-oppimisympäristöissä puuttuen heidän toimintatapaansa ja auttaen erityisesti heikkoa kognitiivista sitoutumista osoittavia oppilaita.

Tutkimuksen perustavoite – tuottaa uutta tietoa tietokoneympäristössä toteutettavasta ongelmanratkaisusta ja suomalaisista koululaisista tällaisina ongelmanratkaisijoina – saavutettiin, vaikka yksittäisen oppilaan yhden tehtävän ratkaisemisesta oli lähtökohtaisesti käytettävissä hyvin vähän erottelevaa tietoa. Myös aiemman tutkimuksen kanssa samassa linjassa olleet klusteriprofiilit osoittivat PISA-lokiaineiston tutkimuksen olevan aiheellista ja hyödyllistä jatkokäyttöä varten. Klusterointia kehittämällä on mahdollista identifioita oppilaiden ongelmakohtia ja vaikeita alueita ja tarvittaessa tarkastella esimerkiksi koulujen opetussuunnitelmia saadun informaation valossa. Voidaankin todeta, et-

tä tässä käytetyt hyvin tuoreet, oppimisanalyttisen (myös PISA-aineiston) tutkimuksen yhteydessä kehittyneet ja vielä kehittyvät analysointimenetelmät osoittautuivat lupaaviksi. Samaa voidaan sanoa koko oppimisanalyttisestä lähestymistavasta.

Tutkimusta voitaisiin jatkaa ja syventää monella tavalla. Kuten Saarela ja Kärkkäinen (2015b) osoittavat (ks. myös Saarela 2017), analyysi voitaisiin kohdistaa otoksen sijasta koko suomalaisen opiskelijapopulaatioon käyttäen otanta-asetelman mukaisia painoja. Kaikki artikkelissa käytetyt menetelmät voidaan yleistää tähän tarkoitukseen. Samoin tulosten laajempi analysointi tehtäväkohtaisesti, eri kategorioissa ja erityisesti suhteessa vastanneiden opiskelijoiden monipuoliseen tausta-aineistoon ja ominaisuuksiin olisi luonteva tie jatkotutkimukselle. Tosin se, että PISA-tehtävät on pidettävä salaisina, vaikeuttaa tehtäväkohtaisten analyysien julkaisemista. Vertailevaa kansainvälistä tutkimusta veisi luontevasti eteenpäin useampien maiden oppilaiden sisällyttäminen analysointeihin.

Lähteet

- Azevedo, R. 2015. Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, 50(1), 84–94.
- Baker, R., Corbett, A., Koedinger, K. & Wagner, A. 2004. Off-task behavior in the cognitive tutor classroom: when students “game the system”. Teoksessa *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. New York: ACM, 383–390.
- Baeza-Yates, R. & Ribeiro-Neto, B. 1999. *Modern information retrieval* (Vol. 463). New York: ACM press.
- Bergner, Y., Colvin, K. & Pritchard, D. E. 2015. Estimation of ability from homework items when there are missing and/or multiple attempts. Teoksessa *Proceedings of the 5th International Conference on Learning Analytics and Knowledge*. New York: ACM, 118–125.
- Bergner, Y., Kerr, D. & Pritchard, D. E. 2015. Methodological challenges in the analysis of MOOC data for exploring the relationship between discussion forum views and learning outcomes. Teoksessa *Proceedings of the 8th International Conference on Educational Data Mining*. Educational Data Mining Society, 234–241.
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U. & Thüs, H. 2012. A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4 (5–6), 318–331.

- Desmarais, M. C. & Lemieux, F. 2013. Clustering and visualizing study state sequences. *Teoksessa Proceedings of the 6th International Conference on Educational Data Mining*, Memphis: International Educational Data Mining Society, 224–227.
- Fayyad, U., Piatesky-Shapiro, S. & Smyth, P. 1996. Extracting useful knowledge from volumes of data. *Communications of the ACM*, 39 (11), 27–34.
- Ferguson, R. 2012. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4 (5–6), 304–317.
- Ferguson, R. & Shum, S. B. 2012. Social learning analytics: five approaches. *Teoksessa Proceedings of the Second International Conference on Learning Analytics and Knowledge*. New York: ACM, 23–33.
- Gress, C. L., Fior, M., Hadwin, A. F. & Winne, P. H. 2010. Measurement and assessment in computer-supported collaborative learning. *Computers in Human Behavior*, 26(5), 806–814.
- Hand, D., Mannila, H. & Smyth, P. 2001. *Principles of data mining*. Boston: MIT press.
- Hettmansperger, T. P. & McKean, J. W. 1998. *Robust nonparametric statistical methods*. London: Edward Arnold.
- Hershkovitz, A., Knight, S., Dawson, S., Jovanovic, J. & Gasevic, D. 2016. About "learning" and "analytics". *Journal of Learning Analytics*, 3 (2), 1–5.
- Hosseini, R., Brusilovsky, P., Yudelson, M. & Hellas, A. 2017. Stereotype Modeling for Problem-Solving Performance Predictions in MOOCs and Traditional Courses. *Teoksessa: Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. New York: ACM, 76–84.
- Hämäläinen, J., Jauhainen, S. & Kärkkäinen, T. 2017. Comparison of internal clustering validation indices for prototype-based clustering. *Algorithms*, 10, 1–14.
- Jauhainen, S. & Kärkkäinen, T. 2017. A simple cluster validation index with maximal coverage. *Teoksessa Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2017)*, 293–298.
- Joksimovic, S., Manataki, A., Gasevic, D., Dawson, S., Kovanovic, V. & Kereki, I. F. 2016. Translating network position into performance: importance of centrality in different network configurations. *Teoksessa Proceedings of the 6th International Conference on Learning Analytics and Knowledge*. New York: ACM, 314–323.
- Karavirta, V., Korhonen, A., Malmi, L. 2006. On the use of resubmissions in automatic assessment systems. *Computer science education* 16 (3), 229–240.
- Kay, J., Reimann, P., Diebold, E. & Kummerfeld, B. 2013. MOOCs: So Many Learners, So Much Potential... *IEEE Intelligent Systems*, 28 (3), 70–77.
- Kizilcec, R. F., Piech, C. & Schneider, E. 2013. Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. *Teoksessa Proceedings of the third international conference on learning analytics and knowledge*. New York: ACM, 170–179.

- Knight, S., Wise, A. F., Ochoa, X. & Hershkovitz, A. 2017. Learning analytics: Looking to the future. *Journal of Learning Analytics*, 4 (2), 1–5.
- Kruskal, W. & Wallis, W. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47 (260), 583–621.
- Kärkkäinen, T. & Heikkola, E. 2004. Robust formulations for training multi-layer perceptrons. *Neural Computation*, 16 (4), 837–862.
- Kärkkäinen, T. & Äyrämö, S. 2004. Robust clustering methods for incomplete and erroneous data. *WIT Transactions on Information and Communication Technologies*, 33, 101–112.
- Kärkkäinen, T. & Äyrämö, S. 2005. On computation of spatial median for robust data mining. *Teoksessa Proceedings of Sixth Conference on Evolutionary and Deterministic Methods for Design, Optimization and Control with Applications to Industrial and Societal Problems. EUROGEN, Munich*, 1–14.
- Liu, Z., Brown, R., Lynch, C., Barnes, T., Baker, R. S., Bergner, Y. & McNamara, D. S. 2016. MOOC learner behaviors by country and culture; an exploratory analysis. *EDM*, 16, 127–134.
- Novotny, E. 2004. I dont think i click: A protocol analysis study of use of a library online catalog in the internet age. *College & Research Libraries* 65 (6), 525–537.
- OECD 2015. PISA 2015 technical report. Luku 18. Paris: OECD Publishing.
- OECD 2016. Classification and scaling information of PISA 2015 Main Survey Items, in PISA 2015 Results (Volume I). Paris: OECD Publishing. <http://dx.doi.org/10.1787/9789264266490-table127-en>
- Peña-Ayala, A. 2017. Learning analytics: fundamentals, applications and trends: a view of the current state of the art to enhance e-learning. Cham: Springer International Publishing.
- Reich, J., Tingley, D. H., Leder-Luis, J., Roberts, M. E. & Stewart, B. 2014. Computer assisted reading and discovery for student generated text in massive open online courses. *Journal of Learning Analytics*, 2 (1), 156–184.
- Rogers, T. 2015. Critical realism and learning analytics research: epistemological implications of an ontological foundation. *Teoksessa Proceedings of the 5th International Conference on Learning Analytics and Knowledge*. New York: ACM, 223–230.
- Rousseeuw, P. J. & Leroy A. M. 1987. Robust regression and outlier detection. New York: John Wiley & Sons.
- Romero, C. & Ventura, S. 2010. Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 40 (6), 601–618.
- Rutkowski, L., Gonzalez, E., Joncas, M. & von Davier, M. 2010. International large-scale assessment data issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142–151.
- Saarela, M. 2017. Automatic knowledge discovery from sparse and large-scale educational data: case Finland. *Väitöskirja, Jyväskylän yliopisto, Informaatioteknologian tiedekunta*.

- Saarela M. & Kärkkäinen, T. 2015a. Analysing student performance using sparse data of core bachelor courses. *Journal of Educational Data Mining*, 7(1), 3–32.
- Saarela M. & Kärkkäinen, T. 2015b. Do country stereotypes exist in PISA? A clustering approach for large, sparse, and weighted data. *Proc. of the 8th International Conference on Educational Data Mining*, 156–163.
- Saarela, M. & Kärkkäinen, T. 2017. Knowledge discovery from the programme for international student assessment. *Teoksessa Learning Analytics: Fundamentals, Applications, and Trends (229–267)*. Springer International Publishing.
- Schraw, G. 2010. Measuring self-regulation in computer-based learning environments. *Educational Psychologist*, 45(4), 258–266.
- Segedy, J. R., Kinnebrew, J. S. & Biswas, G. 2015. Using coherence analysis to characterize self-regulated learning behaviours in open-ended learning environments. *Journal of Learning Analytics*, 2(1), 13–48.
- Siemens, G. 2013. Learning analytics: the emergence of a discipline. *American Behavioral Scientist*, 57, 1380–1400.
- Tukey, J. W. 1977. *Exploratory data analysis*. Pearson.
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S. & Santos, J. L. 2013. Learning analytics dashboard applications. *American Behavioral Scientist*, 57 (10), 1500–1509.
- Winne, P. H. 2010. Improving measurements of self-regulated learning. *Educational Psychologist*, 45(4), 267–276.
- Winne, P. H., Hadwin, A. F. & Gress, C. 2010. The learning kit project: Software tools for supporting and researching regulation of collaborative learning. *Computers in Human Behavior*, 26(5), 787–793.
- Äyrämö, S. 2006. *Knowledge mining using robust clustering*. Väitöskirja, Jyväskylän yliopisto, Informaatioteknologian tiedekunta.

Liite A.

Taulukko A.1 Yhteenveto niistä tehtävistä, joissa yli puolet yrityksistä kuului tiettyyn ratkaisuprofiiliin.

Klusteri 1	Klusteri 2	Klusteri 3	Klusteri 4	Klusteri 5
Tehtävät ovat nopeasti ratkaistut – ja paljon erityisesti helpompia tehtäviä ratkaistu väärin	Tehtävissä on tehty vähemmän interaktiota ja osoitettu hyvää osaamista	Tehtävissä on käytetty paljon harkintaa mutta ratkaisuissa on silti virheitä	Tehtävissä on tehty enemmän interaktiota ja osoitettu hyvää osaamista	Tehtävät ovat olleet työllimpiä interaktiotehtäviä, joissa vaikeammassa on osoitettu osaamista
Tehtävät				
M033Q01	M305Q01	M442Q02	M800Q01	M155Q04
M423Q01	M474Q01	M954Q04	M915Q01	M273Q01
	M905Q01		M420Q01	M496Q02
	M949Q01		M915Q02	
	M982Q03			
	M982Q01			
Tehtävien ominaisuuksia				
Nämä ovat nopeita tehtäviä, joissa on lyhyt saateteksti ja periaatteessa vain yksi painallus tarpeen.	Näissä tehtävissä on ollut jonkin verran enemmän saatetekstiä, lisäksi mukana on ollut kuvia tai liitetaulukko, jota piti tutkia.	Klusteri 2:ssa olevia tehtäviä vaikeampia tehtäviä, joissa on paljon taustamateriaalia, jota tutkittava oli tarkasti.	Tehtävissä ei oletuksena luulisi tarvittavan paljon aktiota; tutkitaan lähde materiaalia ja lasketaan lukuja yhteen.	Tehtävissä saatetekstin määrä vaihtelee; liitteenä yleensä kuvia tai taulukoita, joista vastaus päätellään. Yhdessä pitää soveltaa las-kukaavaa.
Tehtävien luonne klusteroinnin perusteella				
Vahvistaa tulkintaa, että erityisesti näissä tehtävissä on tehty huolimattomuusvirheitä eikä olla täysin ymmärretty tehtävää	Tulkinta, jonka mukaan tehtävän luonne on vähemmän interaktiivinen eli aktiota vaativa	Suomalaisille koululaisille vaikeasti lähestyttäviä tehtäviä	Näiden tehtävien luonne edellyttää enemmän interaktiota	Tehtävät, jotka työlläytensä takia eivät vahvasti erottele osaamista vaikka hieman vaikeammista onkin selviytytty hieman keskimääräistä paremmin
Nopeaan, mutta virheeliiseen sisältävään päättelyyn suomalaisia koululaisia houkuttelevat tehtävät				