



**TURUN
YLIOPISTO**

Ensimmäisen luokan oppilaiden määrällisen päättelyn testaamisen kehittäminen ja pilotointi

Kasvatustiede, opettajankoulutuslaitos
pro gradu -tutkielma

Laatijat:
Joonas Pulli
Tommi Jumppainen

16.11.2023
Turku

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu
Turnitin OriginalityCheck -järjestelmällä.

Pro gradu -tutkielma

Oppiaine: Kasvatustiede, opettajankoulutuslaitos

Tekijät: Joonas Pulli & Tommi Jumppainen

Otsikko: Ensimmäisen luokan oppilaiden määrällisen päättelyn testaamisen kehittäminen ja pilotointi

Ohjaaja: Professori Minna Hannula-Sormunen

Sivumäärä: 45 sivua, 4 liites.

Päivämäärä: 16.11.2023

Matematiikan oppimisen ja matemaattisten taitojen kehityksen kannalta yksi tärkeimmistä tavoitteista on saada oppilaat oppimaan ongelmanratkaisu- ja päättelytaitoja. Matemaattiset päättelytaidot ovat lapselle uuden oppimisen sekä varhaisten matemaattisten taitojen kehityksen perusta. Määrälliseksi päättelyksi kutsutaan taitoa jäsenellä määrällistä tietoa ja tulla sen avulla uusiin päätelmiin. Matemaattisia päättelytaitoja ja määrällistä päättelyä voidaan aiempien tutkimusten perusteella testata luotettavasti kuviovaakatehtävillä.

Tämän pro gradu -tutkielman tarkoituksena oli kehittää määrällisen päättelyn taidon testaamista ensimmäisen luokan oppilailla. Tutkimus on osa Turun yliopiston oppimisanalytiikan tutkimusinstituutin saamaa toimeksiantoa arvioida kaksivuotisen esiopetuksen vaikutuksia lasten osaamiseen ja oppimisvalmiuksiin. Tässä tutkielmassa esitetty kehitystyö on toteutettu yhdessä Turun yliopiston oppimisanalytiikan tutkimusinstituutin kanssa.

Määrällisen päättelyn taidon testaamista varten kehitettiin 26 kuviovaakatehtävää. Tehtävien avulla luotiin kolme samantasoista testiversiota ja testiversioiden oli tarkoitus osoittaa, että oppilaat ymmärtävät tehtävätyypin ja samalla erotella oppilaiden osaamista. Testiversioiden pilotointiin osallistui 405 oppilasta yhdeksästä eri koulusta ja 25 eri luokasta. Tutkittavat suorittivat testin virtuaalisella ViLLE-alustalla käyttäen tabletteja ja kuulokkeita. Tutkittavien vastaukset analysoitiin SPSS 27 -tilasto-ohjelmalla. Jokaisen testiversioiden tuloksista esitetään frekvenssitaulukot. Testiversioista muodostettiin summamuuttujat keskinäistä vertailua varten.

Tulosten perusteella määrällisen päättelyn taitoa voidaan testata ja arvioida ensimmäisen luokan oppilailla. Suuri osa tutkittavista oppilaista osasi ensimmäiset tehtävät. Lisäksi testiversiot erottelivat oppilaita osaamisen mukaan tarkoituksenmukaisesti ja luotettavasti. Testiversioiden sisäiset reliabiliteetit olivat hyvät. Kaksi kolmesta testiversiosta tuotti odotetun kaltaiset tulokset.

Tässä tutkielmassa on esitetty ja raportoitu määrällisen päättelyn testaamisen kehitystyö. Määrällistä päättelyä voidaan mitata luotettavasti jo nuoriltakin oppilailta. Yleisesti erot oppilaiden määrällisen päättelyn taidoissa ovat suuria jo ensimmäisen luokan lopulla. Turun yliopiston oppimisanalytiikan tutkimusinstituutti jatkaa tässä tutkimuksessa kuvattua pilotoinnin jälkeen määrällisen päättelyn taidon testaamisen ja arvioinnin kehitystyötä.

Avainsanat: määrällinen päättely, matemaattiset taidot, kuviovaakatehtävät, kaksivuotisen esiopetuksen kokeilu

Sisällysluettelo

1	Johdanto	4
2	Matemaattiset ongelmanratkaisutaidot	5
2.1	Määrällinen päättely	6
2.2	Määrällisen päättelyn mittaaminen	8
2.3	Kuviovaakatehtävät.....	9
2.4	Testaamisen kehitystyö ja luotettavuus	11
3	Matemaattisten taitojen kehitys ennen koulua ja koulun alkaessa	13
3.1	Matemaattisten taitojen tukeminen ennen koulua	14
3.2	Matematiikka esiopetuksen opetussuunnitelmissa.....	15
4	Tutkimusongelmat	18
5	Tutkimusmenetelmät	19
5.1	Tutkittavat.....	19
5.2	Kuviovaakatehtävien kehittäminen	19
5.3	Kolmen testiversioiden luominen	20
5.4	Aineistonkeruu	23
5.5	Aineiston käsittely.....	24
5.6	Tutkimuksen eettisyys.....	25
6	Tulokset	26
6.1	Osaamisen arviointi korjaamattomasta aineistosta.....	26
6.2	Osaamisen arviointi korjatusta aineistosta	27
6.3	Testiversioiden vertailu	29
6.4	Ankkuritehtävien tulosten ristiintaulukointeja	31
7	Pohdinta	34
7.1	Määrällisen päättelyn taidot	35
7.2	Testiversioiden toimivuus	36
7.3	Tutkimuksen luotettavuus	37
7.4	Tutkimuksen eettisyys.....	40
	Lähteet	42
	Liitteet	46

1 Johdanto

Tämän tutkimuksen tarkoituksena on kehittää ja pilotoida määrällisen päättelyn tehtäviä 1.-luokkalaisilla oppilailla ja arvioida oppilaiden määrällisen päättelyn taidon osaamista. Tässä tutkimuksessa esiteltävä testin kehittämistyö ja pilotointi toteutettiin yhdessä Turun yliopiston oppimisanalytiikan tutkimusinstituutin (myöhemmin nimellä ViLLE-tiimi) kanssa.

Kehittämistyön taustalla on Turun yliopiston oppimisanalytiikan tutkimusinstituutin saama toimeksianto tutkia ja arvioida kaksivuotisen esiopetuksen kokeilun vaikutuksia. Eduskunnan päätöksen mukaisesti laissa kaksivuotisen esiopetuksen kokeilusta (1046/2020) on säädetty, että Suomessa kokeilun avulla tutkitaan kaksivuotisen esiopetuksen vaikutuksia tasa-arvoon ja koulutukseen.

Tämän työn tarkoituksena on kehittää ja esitellä määrällisen päättelyn taidon testiä. Tutkimuskysymyksenä on, voidaanko kehitettävän testin avulla arvioida luotettavasti peruskoulun ensimmäisen luokan aloittaneita oppilaita ja heidän määrällisen päättelyn taitojaan. Tässä tutkimuksessa esitelty pilotointi suoritetaan kevätlukukaudella, kun taas virallinen testaus tullaan tekemään 1. luokan oppilaille syys-lokakuussa. Tämän tutkielman taustalla on tarve tietää, millaiset matemaattisen päättelyn taidot juuri koulunsa aloittavilla oppilailla on. Lisäksi halutaan lisätietoa, kuinka määrällistä päättelyä voitaisiin luotettavasti tulevaisuudessa mitata ja arvioida.

Dwyer, Gallagher, Levin ja Morley (2003) avaavat artikkelissa määrällisen päättelyn termiä. Matemaattisia ongelmanratkaisutaitoja voidaan arvioida määrällisen päättelyn taitoa testaamalla. Määrällisen päättelyn tehtävät soveltuvat pienille oppilaille, sillä määrällistä päättelyä voidaan testata ilman lukutaidon vaatimusta. Määrällisen päättelyn taidon testaamiseen käytettäviä tehtäviä voi vaikeuttaa helposti päättelyprosessien määrää ja kompleksisuutta lisäämällä (Dwyer ym. 2003). Matemaattiset päättelytaidot ovat lapselle tärkeä osa kaiken uuden oppimista sekä perusta matemaattisten taitojen varhaiselle kehitykselle (Alexander, White & Daugherty 1997). Lapsen hyvä matemaattinen osaaminen esikoulussa ennustaa positiivisesti hänen matemaattisten taitojen kehitystään jatkossa (Aunola, Leskinen, Lerkkanen & Nurmi 2004). Matemaattisten taitojen ja matemaattisen ajattelun kehityksestä sekä arvioinnista on tärkeää saada lisää tutkimusta ja tietoa, sillä aiemmat tutkimukset painottavat pienten lasten matemaattisten taitojen kehityksen merkitystä. Matemaattisia taitoja lapset oppivat ja lasten tulisi oppia monipuolisesti jo ennen koulutaipaleensa alkamista (Aunola ym. 2004).

2 Matemaattiset ongelmanratkaisutaidot

Matemaattiset ongelmanratkaisutaidot ovat olennainen osa matematiikkaa. Matemaattisilla ongelmanratkaisutaidoilla tarkoitetaan kykyä ratkaista ongelmia, joihin vaaditaan matemaattisten tietojen, taitojen ja strategioiden monipuolista hyödyntämistä (Leppäaho 2007; Leppäaho 2018). Tieto ongelmanratkaisutaitojen osaamisesta on olennaista, kun ollaan kiinnostuneita taidosta käyttäen matemaattisia taitoja (Silver 1987, 33).

Matemaattinen ongelmanratkaisu löytyy opetussuunnitelmista ympäri maailmaa ja ongelmanratkaisutaitojen opettamista pidetään tärkeänä (Liljedahl & Santos-Trigo 2019). Kilpatrick, Swafford & Findell (2001) toteavat, että lasten ja nuorten on opittava ajattelemaan matemaattisesti oppiakseen uusia asioita. Ihmiset, jotka eivät kykene matemaattiseen päättelyyn, joutuvat usein tahtomattaan ulos yhteiskunnan järjestelmistä. Matemaattisten taitojen puutteet heikentävät kykyä ja mahdollisuuksia toimia jokapäiväisten tehtävien parissa (Kilpatrick ym. 2001, 1). Monet matemaattiseen ongelmanratkaisuun liittyvät aiheet yhdistyvät muiden matematiikan opetuksen aiheiden kanssa (Tjoe 2019, 4).

Kilpatrickin ym. (2001) mukaan ennen koulun alkamista lapset ovat vertailleet omien ja muiden lelujen tai vaikkapa karkkien määrää. Lapset ovat laskeneet askelmien määrää portaita kävellessään tai olleet ruokapöydässä, jossa jaetaan kakkua kaikille syöjille. Päiväkodissa olevista lapsista monet pystyvät suorittamaan pieniä laskutoimituksia, jotka vaativat yhteen- ja vähennyslaskun periaatteita. Monet lapset osaavat käyttää kerto- ja jakolaskun periaatteita ennen koulun alkua (Kilpatrick ym. 2001, 6). Taidot käsitellä lukumääriä, käsitteitä ja tietoja kehittyvät lapsuudessa (Hannula-Sormunen, Mattila, Räsänen & Ruusuvirta 2018, 160). Jo esikouluikäiset osoittavat tutkitusti omaavansa päättelyn taitoja. Pienten lasten päättelytaidot ovat uuden oppimisen ja erityisesti varhaisten matemaattisten taitojen kehityksen perusta (Alexander ym. 1997).

Perusopetuksen opetussuunnitelman perusteet (2014) korostavat vuosiluokkien 1–2 matematiikan opetuksessa matemaattista ajattelua, joka on luovaa, täsmällistä ja loogista. Opetuksen tulee kehittää kykyä ratkaista ongelmia sekä käsitellä tietoa. Opetuksen tulee myös yleisesti auttaa oppilaita ymmärtämään matematiikan merkitys elämässä. Yksi arvioinnin perusteista on matematiikan käyttäminen ongelmanratkaisussa (Opetushallitus 2014, 128, 130).

2.1 Määrällinen päättely

Dwyerin ym. (2003) mukaan 1920-luvulta lähtien muutamia filosofeja ja psykologeja on kiinnostanut määrien päättelemisen taidot osana ihmisen kognitiivista kykyä. Nämä tutkijat toivoivat ongelmanratkaisutaitojen ja määrällisen päättelyn saavan enemmän jalansijaa matematiikasta ja fysiikasta kaavioiden ja yhtälöiden ulkoa opetteluun sijaan. Päättelyn taitojen kehitys nähtiin tärkeänä oppimisen ja opettamisen kannalta (Dwyer ym. 2003).

Määrällinen päättely sisältää termin määrällinen. Thompsonin (1993) mukaan, jos jokin asia on määrällinen, se saa mitattaessa jonkin lukumääräisen arvon. Esimerkiksi mitattaessa pituutta saadaan jokin lukuarvo. Kahta määrää, kuten esimerkiksi kahta pituutta voidaan kuitenkin vertailla keskenään tietämättä niiden tarkkoja arvoja. Vaikka mitattavissa olevien kohteiden tarkkoja pituuksia ei tiedettäisi, voidaan päätellä toisen olevan toista pidempi (Thompson 1993).

Dwyer ym. (2003) määrittelevät määrällisen päättelyn yleiseksi taidoksi jäsenellä määrällistä informaatiota. Määrällinen päättely on yleinen taito ymmärtää ja käyttää erilaisten määrien suhteita ja päästä näiden avulla uusiin päätelmiin. Nunes, Bryant, Evans & Barros (2015) ovat luoneet samankaltaisen määritelmän määrälliselle päättelylle. Määrällinen päättely on taito käsitellä määrällistä informaatiota ja tulla vertailujen avulla uusiin johtopäätöksiin.

Määrällinen päättely on loogista ja yhtenäistä ajattelua, joka sisältää tietoa määristä, kuten matemaattisia yhteyksiä eri asioiden välillä (Nunes ym. 2015). Määrällistä päättelyä on esimerkiksi sen ymmärtäminen, että jos yksi kakku jaetaan kahdelle ihmiselle tai kaksi kakkua jaetaan neljälle ihmiselle, saavat kaikki saman verran (Barros Baertl 2020, 2).

Dwyerin ym. (2003) määrittelyn mukaan määrällinen päättely jakautuu kuuteen erilaiseen osaamisalueeseen (kuva 1). Ensimmäiseksi on osattava lukea ja ymmärrettävä annettuja informaatioita, kuten taulukoita tai geometrisia kuvioita. Toiseksi määrälliseen päättelyyn linkittyy kyky tulkita informaatiota, joka sisältää määriä ja tehdä siitä oikeita päätelmiä. Kolmantena osa-alueena on kyky ongelmanratkaisuun. Neljänneksi määrälliseen päättelyyn liittyy kyky arvioida päättelyn avulla saatuja vastauksia ja pohtia vastauksen järkevyyttä ja oikeellisuutta. Viidenneksi tulisi olla kykyä kommunikoida määrällistä tietoa esimerkiksi sanallisesti, numeroin tai lukumäärin. Vielä viimeiseksi mainitaan kyky ymmärtää matemaattisten tai tilastollisten metodien rajallisuuksien ymmärtäminen (Dwyer ym. 2003).



Kuva 1. Määrällisen päättelyn osa-alueet Dwyerin ym. (2003) mukaan.

Määrällisen päättelyn taito kehittyy luonnostaan matematiikan opetuksessa sekä jokapäiväisessä elämässä, jossa olemme tekemisissä mitä moninaisimpien lukumäärien kanssa. Määrällisen päättelyn käyttäminen arjessa tai arvioitavassa tehtävässä vaatii sen käyttäjältä matemaattisten perustietojen ja -taitojen osaamista. Määrällinen päättely taitona voidaan kuitenkin selkeästi erottaa aritmeettisistä taidoista (Dwyer ym. 2003). Barros Baertl (2020) toteaa, että määrällinen päättely ja aritmeettiset taidot ovat erilliset taidot, joita voi käyttää samanaikaisesti, ja jotka luonnollisesti linkittyvät toisiinsa.

Kun lapsi näkee ja päättelee määriä, hän voi käyttää tiedostamattaan yhteen- tai vähennyslaskua, tai jopa kertolaskun käytänteitä, jos kyse on moninkertaisista määristä. Lapsi voi myös päätellä oikean vastauksen käyttämättä numeroita tai aritmeettisiä prosesseja. Tällöin kyse on luonnollisesta määrien ja lukumäärien suhteiden päättelystä (Barros Baertl 2020, 56). Aikaisemmissa tutkimuksissa on esitetty määrällisen päättelyn taitojen alkavan jo varhaisesta lapsuudesta ja jatkuvan aikuisuuteen asti. Päättelytaitoja voidaan mitata ja ennustaa jo alle 7-vuotiailta lapsilta (Dwyer ym. 2003; Alexander ym. 1997).

Aikaisemmissa tutkimuksissa on tutkittu matemaattisen osaamisen ja kehittymisen taustatekijöitä. Tutkimusten mukaan määrällisen päättelyn taito ennustaa hyvin tutkittavan

matemaattista kehitystä ja osaamista tulevina vuosina. Määrällisen päättelyn testeissä menestyminen on todettu olevan jopa aritmeettisia taitoja parempi ennustaja, kun halutaan ennustaa tutkittavan matemaattista kehittymistä (Nunes, Bryant, Barros & Sylva 2012). Nunes ym. (2007) ovat lisäksi aiemmin todenneet, että testit, joilla mitataan koulunsa aloittavan lapsen loogista ajattelua, ennustavat hyvin lapsen matemaattisten taitojen oppimisen kehittymistä seuraavan vuoden ajan. Tämä loogisen päättelyn ja oppimisen kehityksen yhteys oli myös erillinen yleisen kognitiivisen kyvykkyyden ja työmuistin tuloksista (Nunes ym. 2007).

2.2 Määrällisen päättelyn mittaaminen

Määrällisen päättelyn tutkimuksen teoriatausta antaa mahdollisuuden testata ja arvioida määrällistä päättelyä itsenäisenä taitona. Tutkimuksen tulosten validiteetin ja reliabiliteetin näkökulmasta on tärkeää, että ymmärretään määrällinen päättely ja sen ero aritmetiikasta. Dwyerin ym. (2003) mukaan määrällisen päättelyn taidon testaamista ja mittaamista suunniteltaessa tulee kiinnittää huomiota siihen, miten päättelyä mitataan. Samalla on hyvä miettiä, mikä on sellaista osaamista ja tietoa, jonka voidaan katsoa olevan yleistä tietoa kaikkien testin tekijöiden kontekstissa. Nunesin ym. (2015) tutkimuksessa mitattiin 5–6-vuotiaiden lasten määrällisen päättelyn taitoja. Tutkimuksessa osoitettiin, että määrällisen päättelyn taitoa voi luotettavasti mitata tämän ikäisillä lapsilla (Nunes ym. 2015).

Jotta voidaan arvioida lasten määrällisen päättelyn taitoa, tehtävissä on oltava erilaisia kysymyksiä, jotka liittyvät määriin ja lukujen suhteisiin. Määrällistä päättelyä testaavat kysymykset voivat sisältää tai olla sisältämättä lukuja (Barros Baertl 2020, 60). Dwyerin ym. (2003) mukaan päättelytehtäviä voidaan vaikeuttaa lisäämällä itsenäisten päättelyprosessien määrää, joita tehtävässä tarvitaan oikean vastauksen saamiseksi. Aikaisempien kokemusten mukaan määrällisen päättelyn tehtäviä voidaan vaikeuttaa hyvinkin haastaviksi käyttämällä vain yksinkertaisia aritmeettisia taitoja. Ratkaisevaa päättelytehtävän vaikeuden kannalta on tehtävässä oikeaan vastaukseen vaaditun päättelyn vaativuus ja kompleksisuus (Dwyer ym. 2003). Käytetyin testityyppi määrällisen päättelyn testaamiseen on viiden vaihtoehdon monivalintatehtävä. Vastausvaihtoehtojen avulla monivalintatehtävä antaa erilaisia päättelyn mahdollisuuksia, oikeaa vastausta voi päätellä tehtävänannon lisäksi myös vastausvaihtoehtojen avulla (Dwyer ym. 2003).

Dwyerin ym. (2003) mukaan määrällisen päättelyn tehtävissä on otettava huomioon sen yhteys matematiikkaan, vaikka määrällinen päättely on erillinen taito verrattuna aritmeettisiin

taitoihin. Määrällisen päättelyn mittaaminen on tämän takia rajoittuneempaa kuin määrällinen päättely ylipäätään jokapäiväisessä elämässämme. Määrällisen päättelyn taidon testaamisessa ja arvioinnissa on tärkeää erottaa ne tiedot ja taidot, joita testataan sekä ne tiedot, jotka kaikilla tutkimukseen osallistujilla oletetaan olevan testihetkellä. Tehtävät tulee suunnata sellaiselle joukolle, josta mahdollisuuksien mukaan kaikki osaavat ja tuntevat tehtävässä sisältönä käytetyt matemaattiset sisällöt. Tutkijoiden ja testaajien on tärkeää miettiä tarkkaan, mitä matemaattisia tietoja ja taitoja sekä luku- ja kirjoitustaitoja kaikilla testattavilla oletetaan olevan. Määrällisen päättelyn taitoa on mahdollista tutkia ilman lukutaidon vaatimusta (Dwyer ym. 2003).

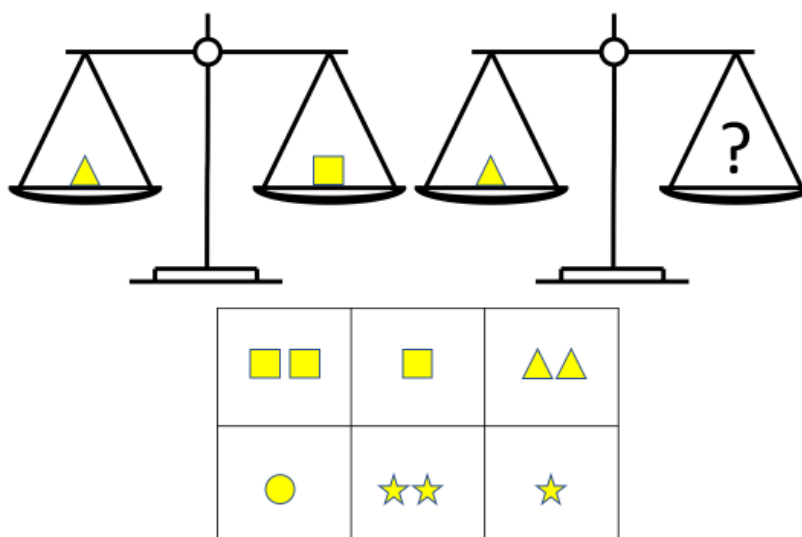
Dwyer ym. (2003) toteavat, että tarvittavan laskutaidon lisäksi tehtävissä on otettava huomioon, että niissä käytetyt kuvat ja kuviot ovat visuaalisesti mahdollisimman selkeät sekä ymmärrettävät kaikille vastaajille. Lisäksi lukutaito ja lukunopeus on otettava huomioon tehtävänannoissa ja vastauksissa. Ongelmanratkaisutehtävissä yksi muuttuva tekijä on tehtävään käytetty aika. On tärkeää pohtia, onko vastausaika tutkittava muuttuja ja onko vastausaikaa testissä ylipäätään rajattu (Dwyer ym. 2003).

2.3 Kuviovaakatehtävät

Tasapainovaa'an analogiaan perustuvia kuviovaakatehtäviä (eng. figure weights) käytetään uusimmassa lasten älykkyyttä mittaavassa testissä Weschler intelligence scale for children V (WISC-V). Kuviovaakatehtävät muodostavat määrällisen päättelyn osatestin (Weiss, Saklofske, Holdnack & Prifitera 2016). Kuviovaakatehtävässä lapsi näkee yhden tai kaksi vaakaa, jotka ovat tasapainossa erilaisten painojen avulla. Lisäksi esitetään vaaka, josta toiselta puolelta puuttuu asia tai asioita. Lapsen pitää annettujen tietojen sekä esillä olevien vaihtoehtojen avulla päätellä oikea vastausvaihtoehto, jotta vaaka olisi tasapainossa (Weiss ym. 2016).

Kuviovaakatehtävät vaativat tutkittavalta analogista päättelyä. Analoginen päättely tarkoittaa yhdestä asiayhteydestä opitun säännön siirtämistä ja käyttämistä toisessa kontekstissa. Analogisen päättelyn kehitys alkaa jo varhaislapsuudesta (Halinen ym. 2016, 54; Gentner & Maravilla 2018). Gentnerin ja Maravillan mukaan analoginen päättely on yksi olennainen osa ihmisen kognitiivista ajattelua. Taito käyttää aiemmin opittuja tietoja ja sääntöjä uusissa tilanteissa on merkittävä oppimisen kannalta. Analogista päättelyä voidaan testata esittämällä perustieto, jonka avulla tulisi päästä seuraavassa tilanteessa oikeaan vastaukseen (Gentner & Maravilla 2018).

Kuviovaakatehtävät mittaavat luotettavasti määrällistä päättelyä (Weiss ym. 2016). Kuvassa 2 esitetään yksi tässä tutkimuksessa käytetystä kuviovaakatehtävästä.



Kuva 2. Esimerkki kuviovaakatehtävästä.

Kuviovaakatehtäviin ohjeistaessa on hyvä kiinnittää lasten huomio visuaalisesti näyttäen, millaisesta tehtävästä on kyse ja missä vastausvaihtoehdot ovat, sekä miten niistä valitaan lapsen mielestä oikea vastaus (Wahlstrom, Weiss & Saklofske 2016, 40). Kuviovaakatehtävät vaativat luonnollisesti kykyä myös työmuistilta, mutta verrattuna aritmeettiseen testiin, visuaalisten kappaleiden käytön nähdään keventävän työmuistin tarvetta, sillä lapsi voi tehtävän aikana helposti katsoa määriä uudestaan (Weiss ym. 2016, 11).

Tasapainovaaka on yksinkertaisuudessaan monipuolinen didaktinen väline matemaattisten suhteiden havainnollistamiseen (Otten, Van den Heuvel-Panhuizen & Veldhuis 2019; Lehtonen 2022). Matemaattisesti yhtäsuuruusmerkki korvaa sanaparin ”sama kuin”.

Yhtäsuuruusmerkin molemmin puolin on asioita jollain tapaa saman verran (Knuth, Stephens, McNeil & Alibali 2006). Tasapainovaa’an käytöstä varsinkin pienten oppilaiden kanssa on positiivisia tuloksia, kun halutaan korvata yhtäsuuruusmerkki lapsille konkreettisemmalla vastineella (Otten ym. 2019).

Tasapainovaa’an analogiaa on käytetty aiemmissa tutkimuksissa osana digitaalisia tehtäviä. Aikaisemmassa tutkimuksessa virtuaalisen tasapainovaa’an avulla oppilaat ovat päätelleet

abstraktien kuvioiden, ympyröiden, kolmioiden ja neliöiden, välisiä suhteita annettujen tietojen perusteella (Otten ym. 2019). Lehtonen (2022) kirjoittaa väitöskirjassaan, että tasapainovaakatehtävät vaativat oletuksen, että oppilas osaa peruslaskutoimitukset, yhteen- ja vähennyslaskun sekä kertomisen sekä jakamisen logiikan.

2.4 Testaamisen kehitystyö ja luotettavuus

Tutkimuksen luotettavuus perustuu suoraan verrannollisesti käytetyn testin luotettavuuteen. Luotettavuutta kuvataan kahdella termillä, validiteetilla ja reliabiliteetilla (Metsämuuronen 2009, 74). Tärkeä osa taidon testaamisen kehittämistyötä on selvittää, toimiko testi halutulla tavalla ja mitattiinko luotettavasti haluttua taitoa (Downing 2006). Kaikissa testeissä on mittausvirheitä, mutta reliabiliteetin tarkastelulla voidaan luotettavammin arvioida testin tuloksia (Linn 2006; Metsämuuronen 2009). Reliabiliteetti mittaa, kuinka samanlaisia tai poikkeavia tuloksia saataisiin eri mittauskerroilla. Toistettavuuden mittana käytetään reliaabeliuskerrointa. Mittauksen sisäisen yhtenevyyden reliabiliteetti mitataan Cronbachin alphan avulla (Metsämuuronen 2009, 75).

Validiteetin tarkastelun avulla tutkitaan, ovatko mittauksessa ja tutkimuksessa käytetyt käsitteet teorian mukaiset. Lisäksi voidaan tarkastella, ovatko tutkimuksessa ja testissä käytetyt käsitteet oikein operationalisoitu ja kattavatko käsitteet riittävän laajasti tutkittavan ilmiön (Metsämuuronen 2009, 74). Testaamisen tulosten ja koko tutkimuksen validiteettia voidaan parantaa ottamalla huomioon tutkittava ilmiö, käsitteet ja niiden operationalisointi kaikissa testin kehityksen eri vaiheissa (Downing 2006). Validiteetin kannalta testikehittäjien on hyvä olla tietoisia mahdollisista luotettavuutta rajoittavista tekijöistä ja raportoida niistä (Linn 2006).

Metsämuuronen (2002) määrittelee kirjassaan validiteettia. Validiteetilla tarkoitetaan mittauksen sisäistä luotettavuutta. Tutkimuksen luotettavuutta voidaan käsitellä sisäisen ja ulkoisen validiteetin avulla. Ulkoisen validiteetin tarkastelu on tutkimuksen yleistettävyyden tarkastelua. Sisäinen validiteetti voidaan jakaa sisällön validiteettiin, rakennevaliditeettiin sekä kriteerivaliditeettiin. Kun kyseessä on taidon testaamisen ja mittaamisen kehittäminen, luotettavuuden tarkastelun tulee olla olennainen osa tutkimusta. Kun halutaan tarkastella sisällön validiteettia, on tärkeää pohtia, ovatko tutkimuksessa käytetyt käsitteet teorian mukaisia ja oikein operationalisoitu. Käsitteiden tulee myös kattaa tarpeeksi laajasti tutkittava ja mitattava ilmiö. Kriteerivalidius on osa sisäistä validiutta ja sen tarkastelua. Kriteerivaliditeetissa testillä saatuja arvoja verrataan tunnettuun arvoon, joka toimii

validiteetin kriteerinä (Metsämuuronen 2002, 32–37). Yksinkertaisesti validiteetin keskeinen sisältö luotettavuuden tutkimisessa on se, mitataanko todella asiaa, jota on tarkoitus mitata (Metsämuuronen 2009, 74).

Downing (2006) on laatinut taitoja mittaavan testin kehittämiseen kattavan ohjeistuksen, joiden avulla testin kehittämisen validiteettia voi parantaa. Linn (2006) kirjoittaa, että testin kehittämistyössä olennaista on aloittaa yleisestä suunnitelmasta. Aivan aluksi kaikkein olennaisinta on ymmärtää, mikä on testaamisen päämäärä eli tarkoitus. Tarkoituksen ymmärtäminen auttaa testin kehittäjiä valitsemaan esimerkiksi parhaimmat tehtävätyypit ja tehtävänannot halutuille testattaville (Linn 2006).

Kehittämistyön edetessä syvennytään teoriaan testattavan taidon taustalla sekä pohditaan metodeja, joilla haluttuja taitoja halutaan ja mahdollisesti voidaan testata (Downing 2006). Linn (2006) tarkentaa, että taustateorian avulla voidaan pohtia ja päättää, mitä taitoa tarkalleen halutaan testata, ja mitkä taidot eivät saa vaikuttaa testin tuloksiin. Taustateorian hallitsemisen jälkeen voidaan ryhtyä suunnittelemaan yksittäisiä tehtäviä ja esitestaamaan valitun tehtävätyypin toimivuutta. Tehtävien suunnittelussa tulee huolehtia monipuolisesti ja tarkasti, että tehtäviä on riittävästi ja että tehtävät ovat riittävän monipuolisia (Linn 2006). Linn (2006) kirjoittaa, että kun halutaan testata esimerkiksi ongelmanratkaisutaitoa, on testiä kehitettäessä tärkeä varmistua siitä, ettei tehtävien aiheisisältö muodostu erottelevaksi tekijäksi, vaan erot tutkittavien välillä syntyvät tutkittavien taidon osaamisessa. Yksi yleisimmistä tehtävätyypeistä ongelmanratkaisutehtävien arviointiin on monivalintatehtävä. Monivalintatehtävä on yleisesti käytetty, nopea tehdä sekä nopeasti koneella tarkastettava tehtävätyyppi (Linn 2006).

Koko kehittämistyö ja sen tulokset tulee raportoida tarkasti. Testin kehittämistyöstä on sen päätyttyä oltava systemaattinen ja yksityiskohtainen kuvaus kehittämistyön vaiheista ja sen tuloksista (Downing 2006). Koko tutkimuksen ajan on otettava huomioon, että mitattaessa ihmisen osaamista tarkat mittaukset eivät ole mahdollisia. Oppimisella ja osaamisella ei ole fyysistä muotoa, jota mitata, vaan osaaminen on hypoteettinen käsite. Oleellista on todeta, että mittarilla saatu tulos osaamisesta on arvio, jonkinasteinen kuva todellisuudesta ja mitatusta ominaisuudesta ja sen määrästä (Metsämuuronen & Räsänen 2018).

3 Matemaattisten taitojen kehitys ennen koulua ja koulun alkaessa

Tässä tutkimuksessa testin kehittämisen ja pilotoinnin kannalta on tärkeää selvittää lasten määrällisen päättelyn taidon osaaminen ennen koulun aloitusta. Seuraavissa kappaleissa taustoitetaan, mitä matemaattisia taitoja voimme testin kehittämisen kannalta olettaa lapsien osaavan. Lisäksi kappaleissa pohjustetaan, miksi lasten matemaattisten päättelytaitojen testaaminen on yhteiskunnallisesti tärkeää.

Matemaattiset taidot ja käsitteet kehittyvät jo ennen koulun alkamista (Aunola ym. 2004; Kilpatrick ym. 2001, 5). Varhaislapsuudessa lapset alkavat kehittää käsitystä esimerkiksi esineiden koosta: onko jokin iso tai pieni ja onko jokin osa jostain tai kokonainen (Aunola ym. 2004). Matemaattisten taitojen ja käsitteiden oppiminen jatkuu varhaiskasvatuksessa ja esikoulussa (Kilpatrick ym. 2001, 5). Matemaattisia taitoja ja käsitteiden hallintaa lapsi voi hyödyntää uusissa tilanteissa ja rakentaa uuden oppimista vanhan, jo omaksutun tiedon päälle. Hyvät matemaattiset käsitteelliset taidot omaava lapsi pystyy uudenlaisen tehtävän äärellä valitsemaan monesta erilaisesta vaihtoehdosta oikean tavan ratkaista tehtävää (Kilpatrick ym. 2001, 5, 127).

Jo pienillä lapsilla on taitoja matemaattisten ongelmien ratkaisemiseen, esittämiseen ja vastausten selittämiseen. Yleisesti pienet lapset ovat luonnostaan positiivisia matematiikkaa kohtaan ja alle kouluikäisillä on paljon matemaattisen taitavuuden piirteitä (Kilpatrick ym. 2001, 6). Aunola ym. (2004) toteavat, että lapsen laskentakyky esikoulun alussa ennustaa sekä matemaattista suoriutumista että matemaattista kehitystahtia. Mitä parempia lapsen matemaattiset kyvyt olivat, sitä parempi oli heidän matemaattinen toimintakykynsä ja kehitystahtinsa (Aunola ym. 2004). Halberda, Mazzocco & Feigenson (2008) ovat todenneet, että matemaattisen päättelyn osaamisen erot ovat suuria 14-vuotiailla. Nämä erot juontuvat päiväkodissa tehtyihin testeihin asti. Korrelaatio pysyy merkittävänä, vaikka muut kognitiiviset ja taidolliset muuttujat ovat vakioitu (Halberda ym. 2008).

Hannula-Sormusen ym. (2018) mukaan ihmiselle on luontaista löytää ympäristöstä lukumääräisyyksiä. Ympäristön täytyy tukea huomion kiinnittymistä lukumääriin. Pienenkin lapsen kehitysympäristöt, esimerkiksi päiväkotit tai esikoulu ja niiden henkilökunta sekä vertaiset, vaikuttavat matemaattisten taitojen kehittymiseen. Matemaattinen ajattelu on havaintojen, tapahtumien ja esineiden välisten suhteiden ja säännönmukaisuuksien löytämistä ja pohtimista (Hannula-Sormunen ym. 2018).

Aunolan ym. (2004) tutkimuksen tulosten mukaan lapsen matemaattinen suorituskkyky esikoulun ja kahden ensimmäisen vuoden aikana peruskoulussa kehittyi kumulatiivisesti. Yksilölliset erot matemaattisessa suorituskvyvyssä pysyivät tasaisina koko tarkastelujakson ajan eli esikoulusta peruskoulun toisen luokan loppuun asti. Lisäksi nämä erot kasvoivat suuremmiksi myöhemminä vuosina. Lapset, joilla oli mittausten alussa hyvät matemaattiset taidot, kehittyivät enemmän verrattuna lapsiin, joilla oli heikot matemaattiset taidot (Aunola ym. 2004).

3.1 Matemaattisten taitojen tukeminen ennen koulua

Aunio, Hannulan ja Räsänen (2004) mukaan lapsen matemaattisten taitojen kehityksessä lapsen oma toiminta on tärkeässä roolissa. Riittävä harjoittelu, matemaattisesti hyvät oppimisympäristöt sekä vuorovaikutus ja sosiaalinen tuki ovat tärkeitä. Lapsen omaa ajattelua ja toimintaa ei voi korvata ulkopuolisella toiminnalla. Ei ole kuitenkaan itsestään selvää, että kaikki lapset automaattisesti kiinnittäisivät huomiota matematiikkaan, ympäristön matemaattisiin tekijöihin ja lukumääriin. Suuri osa lapsista näkee maailman matemaattisena ympäristönä ja tutkii ympäristöään matematiikan näkökulmasta, esimerkiksi pohtien ulkona olevien autojen lukumäärää tai matkan pituutta. Tällöin matematiikan ja matemaattisen ajattelun harjoittelun määrä on luonnollisesti valtava jo pienenä lapsena. Toiset lapset voivat keskittyä aivan muihin teemoihin ympäristössään, kuten väreihin ja tunnelmaan. Nämä erot lasten omassa huomion kiinnittämisessä ympäristöönsä aiheuttavat valtavan eron matemaattisen ajattelun harjoittelussa (Aunio ym. 2004).

Lukumäärien hahmottamisen nähdään olevan luonnollinen ja synnynnäinen taito. Lisäksi ympäristö tukee lukumäärien hahmottamista ja matemaattisiin aiheisiin tutustumista (Aunio ym. 2004). Tutkijoille on vielä epäselvää, mitkä taidot ovat synnynnäisiä ja mitkä enemmänkin harjoiteltavia taitoja. Matemaattisten taitojen on kuitenkin nähty muodostavan oma taitoalueensa erona esimerkiksi kognitiiviseen kyvykkyyteen. On todettu, että matemaattisen ajattelun tukeminen on tärkeässä asemassa, eivätkä yleiset kognitiivisia taitoja tukevat toimet riitä matemaattisten taitojen kehittämiseen varhaislapsuudessa. Lapsen matemaattisten taitojen hyvään kehitykseen päästään harjoittelemalla matematiikkaa (Aunio ym. 2004).

Spontaani huomionkiinnittäminen lukumääriin on kehitettävä taito (Aunio ym. 2004). Se on myös positiivisesti yhteydessä matemaattisiin taitoihin. Lisäksi on tärkeää huomioida, että lapset, jotka eivät spontaanisti kiinnitä huomiota lukumääriin, saisivat ohjattua harjoittelua

huomionkiinnittämiseen matemaattisiin tekijöihin ympäristössä (Hannula & Lehtinen 2005). Hannula-Sormusen, Lehtisen ja Räsänen (2015) mukaan lapsen taito kiinnittää huomiota lukumääriin ja siitä syntyvä matemaattinen osaaminen on nähtävissä vielä kuuden vuoden jälkeen lapsen matemaattisessa osaamisessa. On myös todettu, että matemaattiset taidot ovat erillään verbaalisesta ja kognitiivisesta älykkyydestä. Koska spontaania huomionkiinnittämistä lukumääriin voidaan harjoitella ja sen yhteys myöhempään matemaattisiin taitoihin on selvä, spontaanin huomionkiinnittämisen taitojen kehittäminen ja tukeminen varhaiskasvatuksessa ja esikoulussa olisi tärkeää lapsen matemaattisten taitojen kehityksen kannalta (Hannula-Sormunen ym. 2015).

Aunio ym. mukaan (2004) olisi tärkeää, että Suomessa viimeistään esikoulussa kiinnitetään huomiota lapsen varhaisiin matemaattisiin taitoihin. Lapsen varhaiset matemaattiset taidot ja niiden kehittyminen ennaltaehkäisevät tutkitusti matemaattisia oppimisvaikeuksia jatkossa. Olisi tärkeää, että varhaiskasvatuksen ja esikoulun oppimisympäristöt tukisivat aidosti matematiikan oppimista ja opettamista (Aunio ym. 2004).

3.2 Matematiikka esiopetuksen opetussuunnitelmissa

Suomessa on ollut vuodesta 2021 käynnissä kaksivuotinen esiopetuksen kokeilu. Eduskunnan päätöksen mukaisesti laissa kaksivuotisen esiopetuksen kokeilusta (1046/2020) on säädetty, että Suomessa kokeilun avulla tutkitaan kaksivuotisen esiopetuksen vaikutuksia tasa-arvoon ja koulutukseen. Tarkoituksena on esimerkiksi selvittää kaksivuotisen esikoulun käymisen vaikutuksia lapsen kehitys- ja oppimisedellytyksiin (Laki kaksivuotisen esiopetuksen kokeilusta 1046/2020 § 1).

Laissa säädetään, että kaksivuotista esiopetusta on annettava yhteensä 1400 tuntia, vähintään 700 tuntia kumpanakin vuotena (Laki kaksivuotisen esiopetuksen kokeilusta 1046/2020 § 7). Lain mukaan seurannan arvioija ja toteuttaja voi teettää arviointeja kohdejoukkoon kuuluville lapsille. Näillä arvioinneilla voidaan esimerkiksi seurata lasten kehitys- ja oppimisedellytysten kehitystä (Laki kaksivuotisen esiopetuksen kokeilusta 1046/2020 § 10).

Lain pohjalta Opetushallitus on julkaissut Kaksivuotisen esiopetuksen opetussuunnitelman perusteet 2021 (myöhemmin Opetushallitus 2021), jossa on avattu kaksivuotisen esiopetuksen pedagogisia tavoitteita ja sisältöjä tarkemmin. Kaksivuotisen esiopetuksen opetussuunnitelma perustuu pitkälti Esiopetuksen opetussuunnitelmaan vuodelta 2014, mutta uusissa esiopetuksen opetussuunnitelman perusteissa tunnistetaan, että kaksivuotisessa esiopetuksessa

on enemmän aikaa keskittyä pitkäjänteisesti saavuttamaan opetussuunnitelman tavoitteet jokaisen lapsen kohdalla yksilölliseen tahtiin. Kokeiluperusteissa korostuvat asioiden havainnointi, tutkiminen ja leikki. Lisäksi on otettu huomioon esiopetuksen nivelmäisyys varhaiskasvatuksen ja perusopetuksen välissä (Opetushallitus 2021, 7).

Kokeilun opetussuunnitelmassa määrätään tavoitteet, jotka esiopetuksen tulisi täyttää. Kokeiluperusteissa määrätään oppimisen alueita, kuten erilaisia matemaattisia osa-alueita ja niiden sisältöjä. Oppimisenalueet on jaettu viideksi kokonaisuudeksi, joista yksi on ”Tutkin ja toimin ympäristössäni”. Tähän oppimisenalueeseen sisältyy matemaattisten taitojen tukeminen. Keskeistä on lasten omien havaintojen ja toiminnan sekä esiopetuksessa syntyvien tilanteiden hyödyntäminen ohjauksessa (Opetushallitus 2021, 48). Opetushallitus (2021) toteaa Kaksivuotisen esiopetuksen kokeilun opetussuunnitelmassaan, että ”matemaattinen toiminta kehittää lapsen taitoja monipuolisesti”. Opetussuunnitelmassa mainitaan matemaattisten havaintojen arvostamisen sekä rikastamisen tukevan lasten kehittymistä ajattelijoina. Esiopetuksen tehtävänä on lisäksi tukea lasten matemaattisen ajattelun kehittymistä ja kiinnostusta matematiikkaa kohtaan (Opetushallitus 2021, 48).

Esiopetuksen tavoitteena on luoda mahdollisuuksia luvun, muutoksen, geometristen käsitteiden sekä avaruudellisen hahmottamisen ja mittaamistaitojen kehittymiselle. Tavoitteena on tarjota oppimisen iloa ja tukea kaikkien lasten matemaattisia valmiuksia (Opetushallitus 2021, 48). Nimenomaan tutkiva toiminta ja oppiminen ovat keskiössä matemaattisten taitojen oppimisessa esikoulussa. ”Lapset opettelevat vertailemaan, luokittelemaan, sekä järjestelemään havaintojen tai mittausten pohjalta saatuja tietoja”. Lisäksi lapsia rohkaistaan tekemään omia päätelmiä ja keksimään ratkaisuja ongelmiin (Opetushallitus 2021, 49).

Matematiikan ja numeerisen päättelyn oppimisen osalta kaksivuotisen esiopetuksen kokeilun opetussuunnitelman perusteissa (2021) on lisäyksiä verrattuna esiopetuksen opetussuunnitelmaan. Kaksivuotisen esiopetuksen opetussuunnitelmassa (2021) on aiempaa tarkemmin avattu matemaattisten taitojen käsitteitä, joihin esiopetuksessa tulisi tutustua ja joita lasten tulisi oppia. Kirjallisesti aiempaa tarkemmin avattuja käsitteitä ovat matemaattiset suhteet, lukukäsite, geometriset käsitteet, mittaaminen ja aika sekä aritmeettiset taidot. Opetussuunnitelmassa tunnistetaan, että nämä osa-alueet eivät kehity yksitellen, vaan usein samanaikaisesti yhdessä. Matemaattisten suhteiden oppimisen perusajatuksena on havainnointi, niiden sanallistaminen ja käsitteellistäminen. Opetuksessa keskitytään lasten

omiin spontaaneihin havaintoihin ja annetaan heille mahdollisuuksia kehittyä ongelmanratkaisutaidoissa sekä matemaattisloogisessa ajattelussa (Opetushallitus 2021, 50–51).

4 Tutkimusongelmat

Määrällisen päättelyn taidon testaamisesta ja arvioinnista alakouluikäisillä oppilailla on Suomessa hyvin vähän tutkimusta. Turun yliopiston oppimisanalytiikan tutkimusinstituutin saama toimeksianto arvioida kaksivuotisen esiopetuksen kokeilun vaikutusta luo tarpeen tutkia ja kehittää lasten määrällisen päättelyn taidon testausta. Varsinainen testaus suoritetaan tulevaisuudessa suurelle joukolle oppilaita, joten testin kehittämistyössä on otettava huomioon taustateoria, tehtävät, testiversiot, ohjeistukset ja laitevaatimukset sekä realistiset ajalliset ja taloudelliset resurssit.

Tässä tutkimuksessa kehitettiin ja pilotoitiin määrällisen päättelyn tehtäviä, joista koostettiin pilotointiin kolme erilaista testiversiota. Tutkittava oppilas vastasi satunnaisesti yhteen kolmesta testiversiosta eli jokaisella testiversiolla oli eri vastaajajoukko. Kaikissa kolmessa testiversiossa testattiin määrällistä päättelyä kymmenen monivalintatehtävän avulla.

Vastaajajoukkojen vertailua varten kaikissa testiversioissa oli kaksi samaa tehtävää, jokaisen testiversion tehtävät 3 ja 6. Näitä tehtäviä kutsuttiin ankkuritehtäviksi. Ankkuritehtävien osaamisen vertailun perusteella pystytään vertailemaan vastaajajoukkoja toisiinsa.

Tutkimuksessa kehitettyjen määrällisen päättelyn tehtävien, luotujen testiversioiden ja niiden pilotoinnin avulla pystyttiin kysymään seuraavat tutkimuskysymykset:

1. Minkälaisia määrällisen päättelyn taitoja 1.-luokkalaisilla oppilailla oli, kun taitoa testattiin kuviovaakatehtävillä?
2. Miten pilotointivaiheen kolme määrällisen päättelyn testiversiota toimivat, kun tavoiteltiin normaalijakautunutta vastausaineistoa?
3. Miten tehtävätyyppi yleisesti soveltui 1.-luokkalaisille?

5 Tutkimusmenetelmät

Tämän tutkielman tarkoituksena oli kehittää, pilotoida sekä arvioida määrällisen päättelyn tehtäviä ja testiversioita. Määrällisen päättelyn taitoa mitattiin kuviovaakatehtävillä. Tämä pro gradu -opinnäytetyö oli osa määrällisen päättelyn testin pilotointia kaksivuotisen esiopetuksen arviointia varten. Kehitimme ja pilotoimme määrällisen päättelyn testiä yhdessä Turun oppimisanalytiikan tutkimusinstituutin eli ViLLE-tiimin kanssa. Kuviovaakatehtävistä ja testiversioista esitellään tässä tutkielmassa vain yksittäisiä tehtäviä. Turun oppimisanalytiikan tutkimusinstituutti testaa tulevaisuudessa oppilaita tässä tutkimuksessa esitellyillä tai niiden kaltaisilla testiversioilla. Ei ole tarkoituksenmukaista, että testit tai niiden kehitysversiot ovat julkisesti nähtävillä etukäteen.

5.1 Tutkittavat

Tässä tutkimuksessa tutkittiin keväällä 2023 ensimmäistä luokkaa käyviä oppilaita. Tutkittavat olivat erään Varsinais-Suomen kunnan alueen alakouluista. Tutkittavat valikoituivat tutkimukseen ViLLE-tiimin ja osallistuneen kunnan yhteistyöllä. Tutkimukseen osallistuneita oppilaita oli yhteensä 405 ja osallistuneita kouluja yhdeksän. Kouluja valittiin eri puolilta kuntaa ja tutkimukseen osallistui sekä suomenkielisiä että ruotsinkielisiä kouluja. Tutkittavia oppilaita oli 25 eri luokalta. Osaamiserot matematiikassa maan eri osien sekä maaseudun ja kaupunkien välillä ovat TIMMS- ja PISA-tutkimusten mukaan Suomessa vähäiset. Samoin suomenkielisten ja ruotsinkielisten matematiikan osaamiserot ovat kaventuneet (Kupari & Hiltunen 2018).

5.2 Kuviovaakatehtävien kehittäminen

Tutkittavien lasten määrällisen päättelyn taitoja testattiin yhdessä ViLLE-tiimin kanssa kuviovaakatehtävillä. WISC-V:n käyttämiin kuviovaakatehtäviin tutustuttiin yhdessä ViLLE-tiimin asiantuntijan kanssa ja saimme itsellemme talteen WISC-V:n kuviovaakatehtävät sekä skandinaavisille lapsille luodut normitaulukot. Ohjeistusta normitaulukoiden lukemiseen saimme ViLLE-tiimin asiantuntijalta. Normitaulukon avulla pystytään lukemaan tietyn kuviovaakatehtävän keskimääräinen osaamisikä tai tietyn ikävaiheen osaamistaso.

Kuviovaakatehtäviin sekä aiheen teoriataustaan tutustumisen perusteella kehitettiin 26 uutta kuviovaakatehtävää. Jokaiselle tekemällemme kuviovaakatehtävälle laskettiin keskimääräinen osaamisikä normitaulukon avulla. Tämän tutkimuksen tehtävistä pyrittiin luomaan osittain samankaltaisia alkuperäisten tehtävien kanssa. Tehtävien kuvat ja kuviot luotiin PowerPoint -

sovelluksella. Tämän tutkimuksen kuviovaakatehtävissä vastausvaihtoehtoja on kuusi, joista yksi vastausvaihtoehto on oikea vastaus. Tehtävissä oikeasta vastauksesta sai yhden pisteen ja väärästä vastauksesta nolla pistettä.

5.3 Kolmen testiversiön luominen

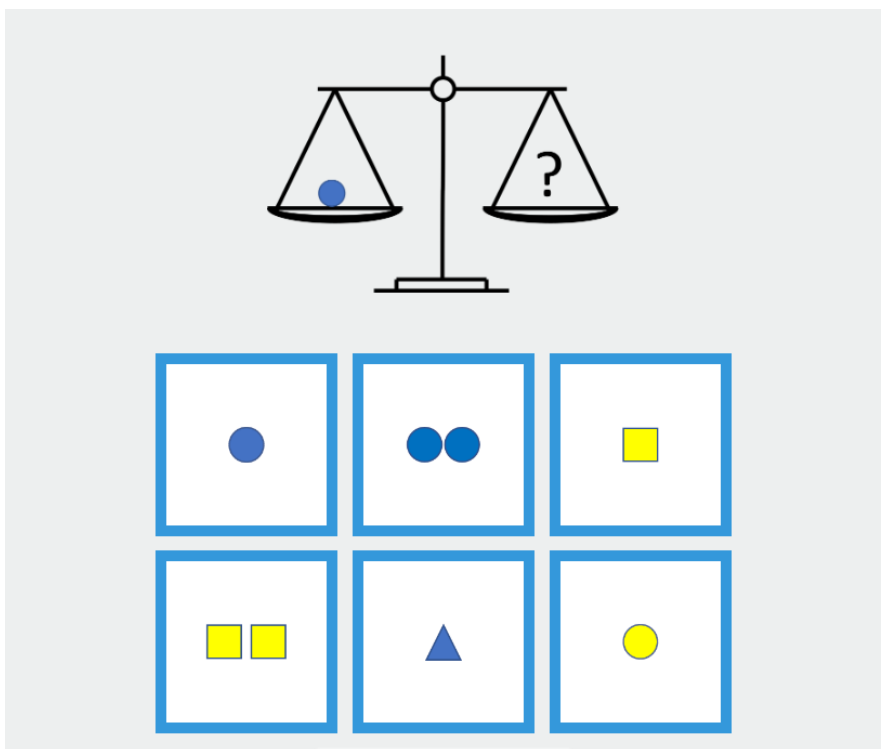
Tässä pilotoititutkimuksessa määrällistä päättelyä mitattiin 26:lla eri kuviovaakatehtävällä. Tehtävistä luotiin kolme eri testiversiota, joista jokaisessa oli kymmenen tehtävää. Kolmen eri testiversiön avulla pystyttiin vertailemaan erilaisia tehtäviä sisältäviä testiversioita keskenään. Tarkoituksena oli testata kuviovaakatehtävien ja testiversioiden toimivuus 1.-luokkalaisilla oppilaille sekä saada eroja näkyviin oppilaiden osaamisen välillä. Testiversioista pyrittiin saamaan mahdollisimman samantasoisia, kuitenkin mahdollistaen pienet erot versioiden välillä. Lisäksi testiversiot suunniteltiin niin, että tuloksista saataisiin mahdollisimman normaali jakauma. Testin yksittäisissä kuviovaakatehtävissä ei ollut käytössä aikarajaa. Mikäli tutkittava vastasi neljä kertaa peräkkäin väärin, testi kyseisen tutkittavan osalta päättyi.

Yhdessä kymmenen tehtävän testiversiossa oli kolme tehtävää, joiden keskimääräinen osaamisikä on noin 6 vuotta sekä neljä tehtävää, joiden keskimääräinen osaamisikä oli noin 7–8 ikävuotta. Lisäksi testiversioissa oli kolme tehtävää, joiden keskimääräinen osaamisikä oli yli 8 ikävuotta, jopa yli 12 ikävuoteen asti. Kaikkien tehtävien keskimääräinen osaamisikä löytyy taulukosta 1. Jokaisessa testiversiossa tehtävät 3 ja 6 olivat samat. Näitä tehtäviä kutsuttiin ankkuritehtäviksi, joiden avulla pystyttiin vertailemaan vastaajajoukkoja. Testin keskeisiksi tehtäviksi haluttiin useampi tehtävä, joiden keskimääräinen osaamisikä on noin 7–8 ikävuotta. Testiversioiden viimeisten tehtävien tarkoitus oli olla niin vaikeita, että vain harva tutkittava pystyy vastaamaan niihin oikein.

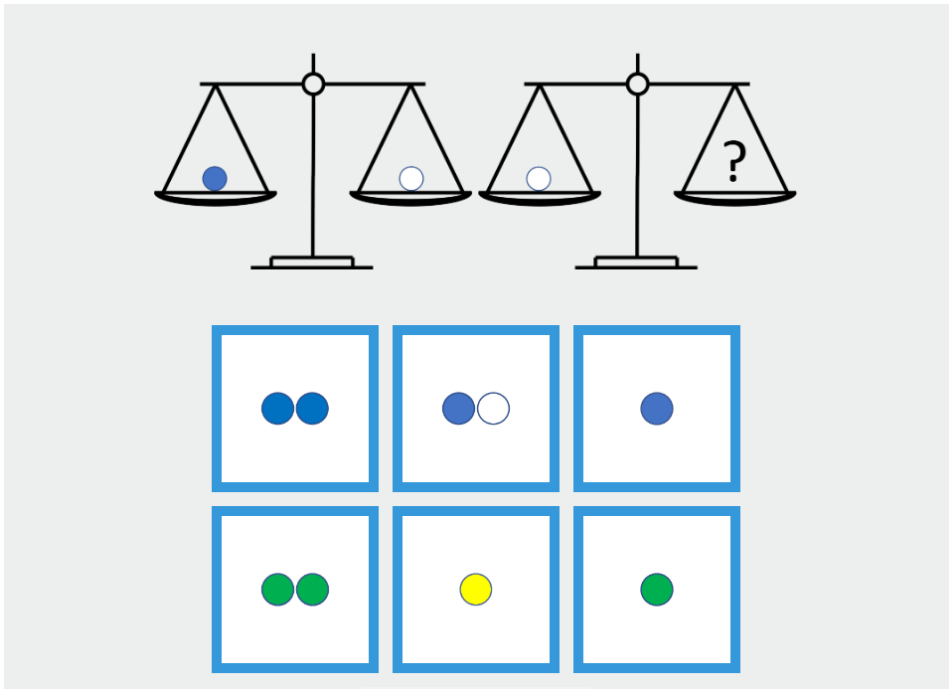
Taulukko 1. Testiversioiden tehtäväkohtaiset keskimääräiset osaamisiät.

Testiversio 1	Keskim. osaamisiikä	Testiversio 2	Keskim. osaamisiikä	Testiversio 3	Keskim. osaamisiikä
1	6v0kk	1	6v0kk	1	6v0kk
2	6v2kk	2	6v2kk	2	6v2kk
3	6v2kk	3	6v2kk	3	6v2kk
4	6v10kk	4	6v6kk	4	6v6kk
5	7v2kk	5	7v2kk	5	6v10kk
6	7v6kk	6	7v6kk	6	7v6kk
7	8v2kk	7	8v10kk	7	8v2kk
8	9v2kk	8	9v2kk	8	8v10kk
9	9v6kk	9	11v0kk	9	9v6kk
10	12v10kk	10	12v10kk	10	12v10kk

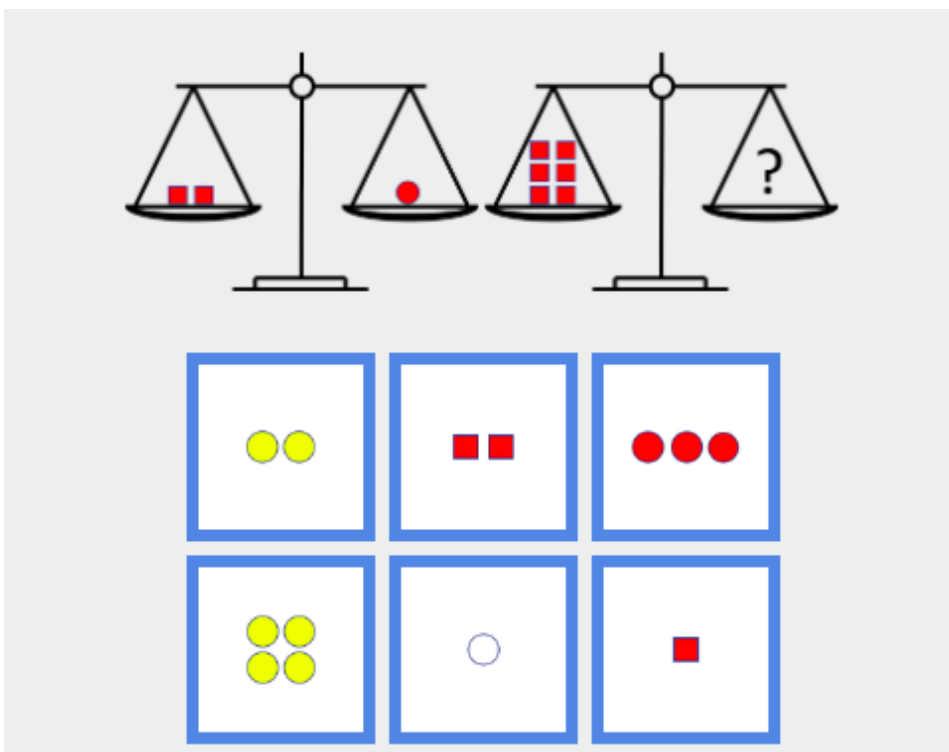
Testiversioiden tehtävät luotiin siis asteittain vaikeutuviksi. Ensimmäisten tehtävien, (kuva 4) ja (kuva 5), haluttiin olevan erittäin helppoja, jotka mahdollisimman moni tutkittava osaisi tehdä. Toinen ankkuritehtävä edusti keskitason osaamista ikäluokassa (kuva 6). Jokaisen testiversioiden viimeinen tehtävä oli tarkoituksella tutkittavalle ikäluokalle jo hyvinkin haastava (kuva 7).



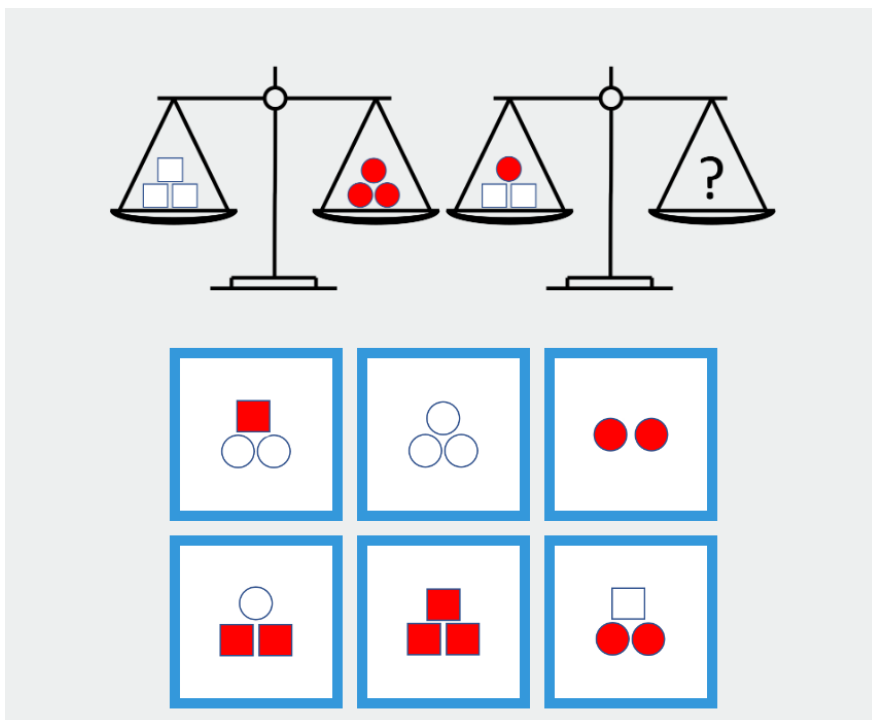
Kuva 4. Testiversio 1, tehtävä 1.



Kuva 5. Testiversio 2, tehtävä 2.



Kuva 6. Testiversiot 1,2 & 3, tehtävä 6, ankkuritehtävä.



Kuva 7. Testiversio 3, tehtävä 10.

Tutkimusta varten tekemämme kolme testiversiota siirrettiin yhteistyössä ViLLE-tiimin kanssa ViLLE:n virtuaaliselle alustalle. Testaamista varten luotiin kirjallinen ohjeistus sekä ohjevideo gif-muodossa. Ohjeiden ja ohjevideon avulla haluttiin antaa yksinkertainen esimerkki kuviovaakatehtävästä sekä selkeyttää tutkittaville vaa'an logiikka ja toimintaperiaate. Lisäksi ohjeen tarkoitus oli näyttää, miten tutkittava valitsee mielestään oikean vaihtoehdon ja vastaa kysymykseen. Ohjeistus pyrittiin pitämään mahdollisimman lyhyenä ja selkokieleisenä tutkittavien iän vuoksi (Liite 4).

5.4 Aineistonkeruu

Aineistonkeruu suoritettiin osallistuvissa luokissa yksi luokka kerrallaan. ViLLE-tiimi loi alustaksi nettisivun, jolla testi suoritettiin. Testausalustana toiminut nettisivu oli sekä suomeksi että ruotsiksi. Sivulla oli oppilaille ensiksi lyhyt ohje, jonka avulla testin suorittaminen onnistui. Lisäksi oli lisätietoja luokan opettajille sekä rehtoreille. Opettaja pystyi suorittamaan testin halutessaan itsenäisesti luokkansa kanssa. Opettajan ohjeissa annettiin ohjeet ennakkovalmisteluihin sekä testin suorittamiseen. Laite- ja ohjelmistovaatimuksissa kerrottiin, millä selaimilla testin tekeminen onnistuu. Niissä myös ohjeistettiin kuulokkeiden käyttämiseen. Tutkimusta varten oli mahdollista saada kuulokkeet lainaan ViLLE-tiimiltä. Koulun oli mahdollista pyytää tukihenkilö tutkimusryhmästä testin

tekemisen tueksi. Suuri osa kouluista pyysi tukihenkilön paikalle. Oppilaiden ei tarvinnut kirjautua henkilökohtaisilla tunnuksilla, vaan he pääsivät testiin lukemalla QR-koodin tai kopioimalla nettilinkin selaimen.

Testitilanteeseen sovittiin tutkijaryhmän kanssa yhteiset ohjeistukset. Tabletit ja kuulokkeet laitettiin paikoilleen ennen oppilaiden saapumista luokkaan. Kun oppilaat asettuivat paikoilleen, tutkijat kertoivat olevansa Turun yliopistolta ja tekevänsä tutkimusta. Oppilaille painotettiin, ettei tutkimuksen vastauksilla ole merkitystä arvosanoihin. Lisäksi kerrottiin, ettei kukaan näe heidän vastauksiaan yksilöllisesti. Tutkittaville oppilaille kerrottiin, että tutkimuksesta voi kieltäytyä ja osallistuminen on vapaaehtoista. Ennen testin aloittamista kerrottiin, että ennen varsinaisia tehtäviä on ohjeet. Tutkittavia kehoitettiin keskittymään ja tutustumaan ohjeisiin huolella. Tämän jälkeen oppilaita ohjeistettiin kirjautumaan testiin QR-koodin avulla. Testin aikana tutkijaryhmän jäsenet ja opettajat saivat toimia vain teknisenä tukena.

Tutkittava oppilas sai itselleen tehtäväksi satunnaisesti yhden kolmesta testiversiosta. Kymmenen kysymyksen testiin tutkittavilta kului aikaa keskimäärin 3–10 minuuttia. Oikeasta vastauksesta tehtävään tutkittava sai yhden pisteen. Väärästä vastauksesta sai 0 pistettä. Mikäli tutkittava vastasi neljä kertaa peräkkäin väärin, testi päättyi tutkittavan osalta. Tulokset tallentuivat reaaliaikaisesti ViLLEn järjestelmään. Tämän tutkimuksen aineistonkeruun lisäksi tutkittavat oppilaat tekivät samalla kaksi muuta pilottitestiä. Määrällisen päättelyn testiosio suoritettiin näistä ensimmäisenä.

5.5 Aineiston käsittely

Turun oppimisanalytiikan tutkimusinstituutti lähetti aineistonkeruun jälkeen Excel-tiedoston määrällisen päättelyn testiosion tuloksista. Aineistossa ei ollut henkilötietoja, joten vastaajia ei pystynyt yksilöimään. Tiedosto muokattiin tähän tutkimukseen sopivaksi ja poistimme ylimääräisen aineiston, kuten vastausajan. Muokkaamisen jälkeen aineisto siirrettiin SPSS 27 -tilasto-ohjelmaan, jossa aineiston tarkempi analysointi suoritettiin. Jokaista kolmea testiversiota tarkasteltiin erikseen ja niistä tulostettiin frekvenssitaulukot sekä korjaamattomasta että korjatusta aineistosta. Korjatun aineiston pohjalta luotiin jokaisesta testiversiosta summamuuttujat, joiden avulla testiversioita pystyttiin vertailemaan keskenään. Vertailuissa tutkittiin testiversioiden tulosten normaalijakautuneisuutta. Lisäksi jokaisesta kolmesta testiversiosta luotiin histogrammi. Ankkuritehtävien avulla suoritettiin eri

testiversioihin vastanneiden vastaajajoukkojen välillä osaamisen vertailua ristiintaulukoinnin avulla.

5.6 Tutkimuksen eettisyys

Tässä tutkimuksessa tutkittavat olivat 1. luokan oppilaita. Tutkielmassa noudatettiin tutkimuseettisen neuvottelukunnan ohjeita 2019 (myöhemmin TENK 2019). ”Eettisiä periaatteita noudatetaan aina ihmisiä tutkittaessa riippumatta siitä, minkä ikäinen tutkittava on. Lapsen tulee saada vaikuttaa itseään koskeviin asioihin kehitystään vastaavasti” (TENK 2019, 9). Tutkittaville annettiin tietoa tutkimuksesta sellaisella tavalla, jonka he pystyvät ikätasoonsa nähden ymmärtämään. Tämä otettiin huomioon pitämällä ohjeet lyhyinä ja mahdollisimman helposti ymmärrettävinä.

Tutkimusta varten ei kerätty henkilötietoja, eikä tutkittavia pystynyt yksilöimään vastausten perusteella. Tutkijaryhmä tietää ainoastaan, mistä kouluista ja miltä luokilta aineisto kerättiin. Tutkittaville kerrottiin tutkimusohjeita annettaessa, että kukaan ei tule tietämään heidän henkilötietojaan tai pysty yksilöimään vastauksia.

”Suurille vastaajamäärille kohdennettavissa kyselytutkimuksissa riittää huoltajan informointi tutkimuksesta niin, että he voivat halutessaan kieltää lastansa osallistumasta tutkimukseen. Suuri vastaajamäärä on vähintään 400 henkilölle kohdennettu tutkimus” (TENK 2019, 10). Tässä tutkimuksessa vastaajamäärä oli yli 400. Huoltajia informoitiin tutkimuksesta koulun viestintäjärjestelmässä Wilma-viestillä suomeksi (Liite 1), ruotsiksi (Liite 2) ja englanniksi (Liite 3). Halutessaan huoltajilla oli mahdollisuus kieltää lapseltaan osallistuminen tutkimukseen.

Vaikka vanhempia oli informoitu tutkimuksesta ja tällä tavoin saatu suostumus lapsen osallistumisesta tutkimukseen, oli tutkittavilla lapsilla mahdollisuus kieltäytyä tutkimuksen tekemisestä vedoten itsemääräämisoikeuden ja vapaaehtoisuuden periaatteeseen. Tutkittaville kerrottiin aineistonkeruun aluksi, että on mahdollisuus olla osallistumatta tutkimukseen. Lisäksi painotettiin, ettei tutkimukseen osallistuminen tai tutkimuksesta kieltäytyminen vaikuta oppilaan arviointiin.

6 Tulokset

Tämän tutkimuksen tavoitteena oli luoda, kehittää ja pilotoida määrällisen päättelyn taitoa mittaavia testiversioita. Tarkoituksena oli, että mahdollisimman moni tutkittava osaisi aloittaa testin oikeilla vastauksilla. Samalla testin pilotoinnin tavoitteena oli, että testi erottelisi tutkittavia tasaisesti oikeiden vastausten osalta. Tässä tulososiossa esitetään tuloksia määrällisen päättelyn taidon testiversioista. Lisäksi vertaillaan testiversioita keskenään ja alustavasti etsitään parasta testiversiota tulevaisuuden jatkokehitystä varten. Tuloksia arvioitaessa on hyvä muistaa, että tämä pilotointi suoritettiin ensimmäisen luokan oppilaille keväällä, kun varsinainen osaamisen arviointi tullaan suorittamaan ensimmäisen luokan oppilaille heti syksyllä.

6.1 Osaamisen arviointi korjaamattomasta aineistosta

Tämän tutkimuksen tarkoituksena oli esitellä ja arvioida noin 7-vuotiaiden lasten määrällisen päättelyn taitoa tulevaa taitojen arviointia varten. Testiversio 1 teki 137 vastaajaa, testiversio 2 teki 133 vastaajaa ja testiversio 3 teki 135 vastaajaa. Aineistossa kaikki puuttuvat vastaajat johtuvat testin asetuksista, jossa neljällä peräkkäisellä väärällä vastauksella testi tutkittavan osalta päättyi. Tutkimuskysymysten kannalta oleellista oli selvittää, kuinka testityyppi sopii 1.-luokkalaisten tekemäksi. Tulososiossa esitetyissä taulukoissa 2–7 keskiarvo osoittaa tehtävän onnistumisprosentin. Lisäksi summa osoittaa, kuinka moni tutkittava vastasi tehtävään oikein. Ensimmäinen tehtävä onnistui tutkittavilta eri testiversioissa 89 % (Testiversio 1), 87 % (Testiversio 2) ja 90 % (Testiversio 3) todennäköisyydellä. Toisen tehtävän onnistui tekemään 91 % (Testiversio 1), 82 % (testiversio 2) ja 82 % (Testiversio 3) osuus tutkittavista.

Puuttuvien vastaajien määrästä voidaan nähdä, kuinka moni vastaaja on testin aikana pudonnut testiosiosta neljän peräkkäisen väärän vastauksen takia. Testiversiossa 1 toiseksi viimeiseen tehtävään on vastannut 137:sta vastaajasta enää 69 ja viimeiseen 64 vastaajaa. Testiversiossa 2 133:sta vastaajasta toiseksi viimeiseen kysymykseen on päässyt 98 vastaajaa ja viimeiseen kysymykseen 91 vastaajaa. Testiversiossa 3 kahden viimeisen kysymyksen vastaajamäärät ovat 81 ja 77 vastaajaa 135:sta aloittaneesta tutkittavasta. Tarkat tunnusluvut testiversioiden 1, 2 ja 3 korjaamattomasta aineistosta ovat nähtävissä taulukoissa 2–4.

Taulukko 2. Testiversion 1 tunnusluvut korjaamattomasta aineistosta.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
N	137	137	137	137	133	128	96	74	69	64
Puuttuu	0	0	0	0	4	9	41	63	68	73
Keskiarvo	,89	,91	,53	,23	,34	,44	,45	,54	,30	,25
Summa	122	124	72	31	45	56	43	40	21	16

Taulukko 3. Testiversion 2 tunnusluvut korjaamattomasta aineistosta.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
N	133	133	133	133	128	126	122	113	98	91
Puuttuu	0	0	0	0	5	7	11	20	35	42
Keskiarvo	,87	,82	,80	,77	,31	,52	,57	,34	,21	,32
Summa	116	109	106	103	40	65	69	38	21	29

Taulukko 4. Testiversion 3 tunnusluvut korjaamattomasta aineistosta.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
N	135	135	135	135	135	129	119	101	81	77
Puuttuu	0	0	0	0	0	6	16	34	54	58
Keskiarvo	,90	,82	,73	,66	,41	,45	,38	,47	,17	,31
Summa	121	111	98	89	56	58	45	47	14	24

6.2 Osaamisen arviointi korjatusta aineistosta

Puuttuvat vastaukset edellisen kappaleen aineistossa johtuvat siitä, että tutkittava oli vastannut neljä kertaa peräkkäin väärin. Testiversioiden puuttuvat arvot korvattiin 0:lla eli väärällä vastauksella. Puuttuvien tietojen korvaus eli imputointi on tutkijan sivistynyt arvaus mittaustuloksista (Nummenmaa 2009, 159). Koska puuttuvat vastaukset puuttuvat vastaajilta, jotka ovat vastanneet testissä neljä kertaa peräkkäin väärin, oletetaan myös puuttuvat vastaukset vääriksi vastauksiksi.

Aineiston korjauksen jälkeen tarkastettiin testiversioiden reliabiliteetit Cronbachin alphan avulla. Mittauksen reliabiliteetin avulla arvioidaan, kuinka paljon mittausvirhettä tulos sisältää. Reliabiliteetin tarkastelun avulla voidaan arvioida sekä kehittää mittalaitteena toimivia kysymyksiä. Reliabiliteetti kertoo, kuinka hyvin testin kysymykset mittaavat samaa asiaa. (Nummenmaa 2009, 351, 357). Testiversion 1 alphakerroin oli 0,831. Testiversion 2 alphakerroin oli 0,768 ja testiversion 3 alphakerroin oli 0,771. Kaikissa testiversioissa

reliabiliteetti on hyvä, kun hyvän Cronbachin alphan rajana pidetään 0,7:aa (Nummenmaa 2009, 378).

Ensimmäisten tehtävien osaaminen on kerrottu edellisessä luvussa. Keskimmäisten tehtävien osaaminen eri testiversioissa oli vaihtelevaa. Testiversiossa 1 tehtävän 4 osasi tutkittavista vain 23 %, viidennen tehtävän osasi 33 % ja kuudennen 41 % (taulukko 5). Testiversiossa 2 neljännen tehtävän osasi 77 %, viidennen 30 % ja kuudennen tehtävän 49 % tutkittavista (taulukko 6). Testiversiossa 3 tehtävien 4, 5 ja 6 osaamisprosentit olivat 66 %, 41 % ja 43 % (taulukko 7).

Viimeisten tehtävien oli tarkoitus olla haastavia, mutta tutkittavalle ikäryhmälle mahdollisia. Testiversiossa 1 tehtävän 9 ja tehtävän 10 osaamisprosentit kaikista testin tekijöistä olivat 15 % ja 12 % (taulukko 5). Testiversiossa 2 kahden viimeisen tehtävän osaamisprosentit olivat 16 % ja 22 % (taulukko 6). Testiversiossa 3 tehtävät 9 ja 10 osasi 10 % ja 18 % kaikista testin tekijöistä (taulukko 7).

Taulukko 5. Testiversion 1 tunnusluvut korjatusta aineistosta.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
N	137	137	137	137	137	137	137	137	137	137
Keskiarvo	,89	,91	,53	,23	,33	,41	,31	,29	,15	,12
Summa	122	124	72	31	45	56	43	40	21	16

Taulukko 6. Testiversion 2 tunnusluvut korjatusta aineistosta.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
N	133	133	133	133	133	133	133	133	133	133
Keskiarvo	,87	,82	,80	,77	,30	,49	,52	,29	,16	,22
Summa	116	109	106	103	40	65	69	38	21	29

Taulukko 7. Testiversion 3 tunnusluvut korjatusta aineistosta.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
N	135	135	135	135	135	135	135	135	135	135
Keskiarvo	,90	,82	,73	,66	,41	,43	,33	,35	,10	,18
Summa	121	111	98	89	56	58	45	47	14	24

6.3 Testiversioiden vertailu

Toisena tutkimusongelmana oli hyvän määrällisen päättelyn taitoa mittaavan testiversion kehittäminen. Tämän tutkimuksen oleellisena osana oli kolmen samankaltaisen testiversion kehittäminen ja pilotointi. Tässä kappaleessa esitetään testiversioiden keskinäinen vertailu. Testiversioiden pilotoinnin oleelliset tulokset ja tunnusluvut löytyvät taulukosta 8.

Taulukko 8. Tunnuslukuja testiversioista 1,2 & 3.

	Testiversio 1	Testiversio 2	Testiversio 3
N	137	133	135
Tehtävien lkm.	10	10	10
Keskiarvo	4,16	5,23	4,91
Mediaani	3	5	4
Keskihajonta	2,62	2,43	2,46
Vinous	,67	-,21	,27
Huipukkuus	-,64	-,64	-,91
Minimi	0	0	1
Maksimi	10	10	10

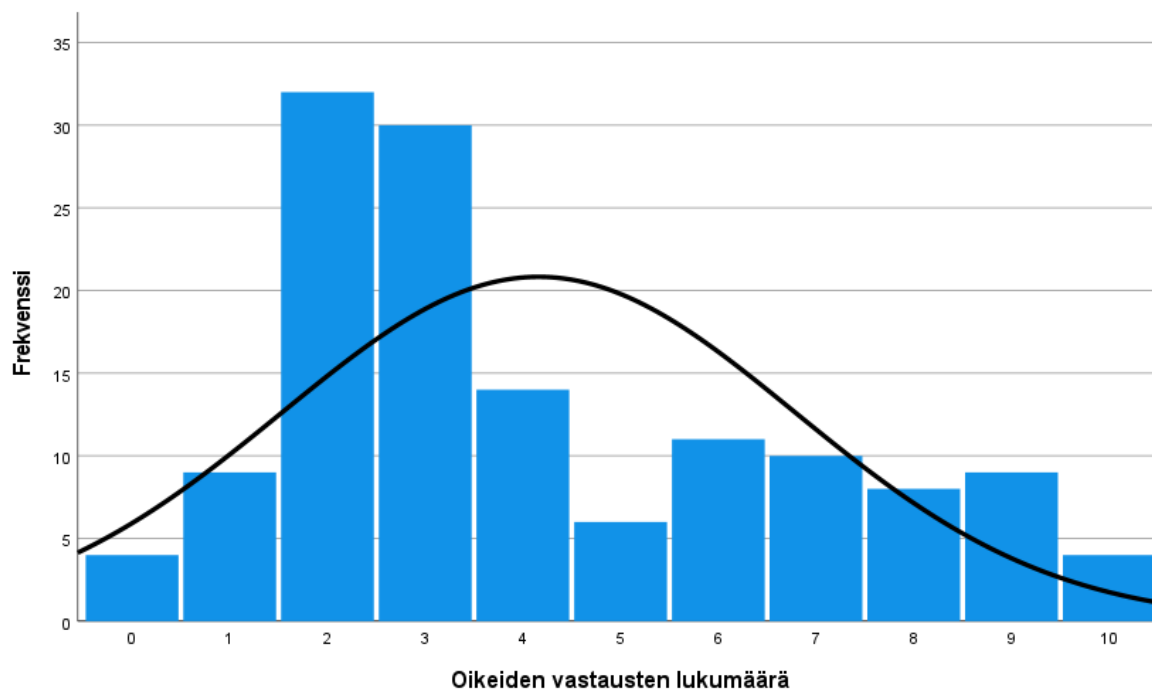
Tunnuslukutaulukkoon muuttujien arvot on pyöristetty kahden desimaalin tarkkuudella. Testiversioissa oli kymmenen kysymystä, joista oikealla vastauksella sai yhden pisteen. Minimipistemäärä testissä oli 0 pistettä ja maksimipistemäärä 10 pistettä. Testiversioiden keskiarvot poikkeavat hieman toisistaan, pienimmän keskiarvon ollessa 4,16 (testiversio 1) ja suurimman ollessa 5,23 (testiversio 2). Pienen tehtävämäärän takia keskihajonnat ovat suuret, selvästi yli 2 kaikissa testiversioissa.

Vinousarvon avulla tarkastellaan, kuinka suuri osa havainnoista on keskiarvoa pienempiä tai suurempia. Mikäli vinousarvo $g_1 > 0$, niin jakauma on oikealle vino. Oikealle vino jakauma tarkoittaa, että suurin osa havainnoista on keskiarvoa pienempiä. Vasemmalle vinossa jakaumassa $g_1 < 0$, suurin osa havainnoista on keskiarvoa suurempia (Nummenmaa 2009, 71). Testiversio 1 ja testiversio 3 ovat oikealle vinoja. Testiversio 2 on vasemmalle vino.

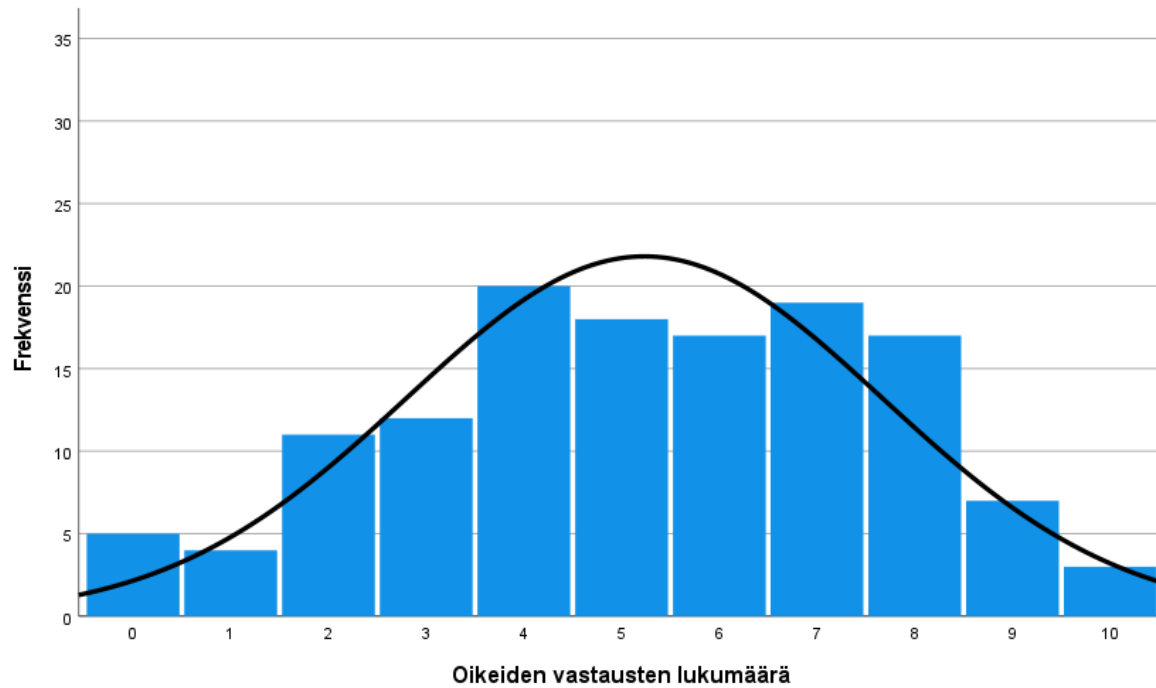
Huipukkuuskerroin g_2 kuvaa aineiston huipukkuutta. Standardoidun normaalijakauman huipukkuuskerroin olisi 0. Jos huipukkuuskerroin $g_2 > 0$, niin jakauma on terävähuippuinen, jos $g_2 < 0$, jakauma on litteähuippuinen eli huiputon (Nummenmaa 2009, 72). Kaikki kolme tutkittua testiversiota ovat litteähuippuisia eli huiputtomia, huipukkuuskertoimen ollessa negatiivinen jokaisen testiversion osalta.

Testiversioissa 1 ja 2 oli tutkittavia, jotka eivät saaneet yhtään vastausta oikein. Nämä tutkittavat ovat siis vastanneet neljään ensimmäiseen kysymykseen väärin ja testi heidän osaltaan päättyi. Testiversiossa 3 pienin saatu oikeiden vastausten määrä oli 1 oikein. Jokaisen testiversion tekijöissä oli tutkittavia, jotka saivat kaikki kymmenen tehtävää oikein.

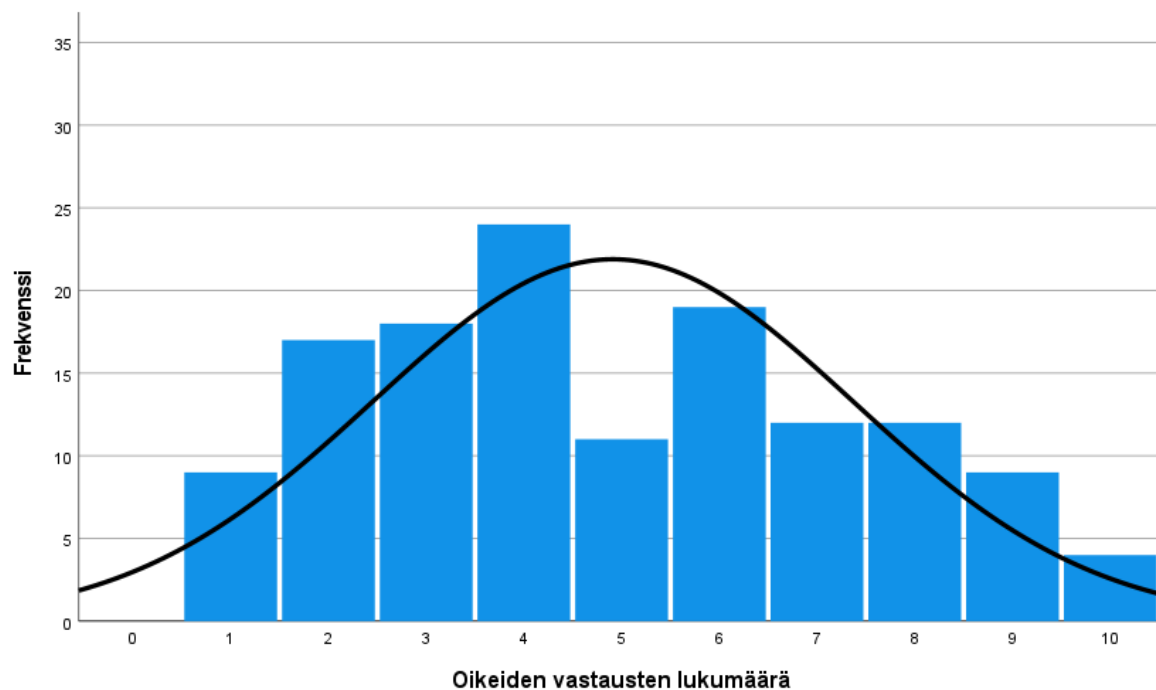
Histogrammi vastaa pylväsdiagrammia, kun halutaan kuvata jatkuvia muuttujia. Histogrammi kuvaa jatkuvan muuttujan frekvenssijakaumaa. Histogrammin yksittäiset pylväät kuvaavat, kuinka monta minkäkin suuruista havaintoa aineistossa on. Pylvään korkeus kertoo havaintojen lukumäärän suuruudesta. Histogrammin avulla voidaan helposti ja selkeästi havainnollistaa jakauman muotoa, esimerkiksi jakauman vinoutta. (Nummenmaa 2009, 81–82). Kuvissa 8, 9 ja 10 on esitetty kaikkien testiversioiden tulosten kuvaaja ja normaalijakauma.



Kuva 8. Testiversion 1 kuvaaja ja normaalijakauma.



Kuva 9. Testversion 2 kuvaaja ja normaalijakauma.



Kuva 10. Testversion 3 kuvaaja ja normaalijakauma.

6.4 Ankkuritehtävien tulosten ristiintaulukoiteja

Testiversio 1 sai huomattavan erilaiset tulokset verrattuna kahteen muuhun testiversioon. Testversion keskiarvo (4,16) oli testiversioista pienin ja vinousarvo (0,67) itseisarvoltaan suurin. Koska jokaisessa testiversiossa ankkuritehtävinä oli samat tehtävät 3 ja 6, pystyimme vertaamaan samojen tehtävien avulla vastaajajoukkoja toisiinsa.

Ristiintaulukointi on yksi käytetyimmistä ja yksinkertaisimmista aineiston analysoinnin menetelmistä. Sen avulla voidaan yksinkertaisesti ja tehokkaasti analysoida tutkimusaineiston muuttujien suhteita. Ristiintaulukointi yhdistettynä khiin neliö -testiin sopii monen tutkimusongelman käsittelyyn. Tyypillisesti ristiintaulukointianalyyssissä on kaksi muuttujaa, joiden välisistä yhteyksistä ollaan kiinnostuneita. (Tähtinen, Laakkonen & Broberg 2016, 166). Ristiintaulukoinnin avulla selvitetään, muodostavatko eri vastaajaryhmien tulokset ankkuritehtäviin tilastollisesti merkitsevän eron. Lähtöoletus khiin neliö -testille on, että otos on satunnaisotos (Tähtinen ym. 2016, 167). On huomattava, että khiin neliö -testi testaa ristiintaulukointia kokonaisuudessaan.

Taulukko 9. Ristiintaulukointi ankkuritehtävästä 1.

		Ristiintaulukointi				
			Testi 1	Testi 2	Testi 3	Yhteensä
Ankkuritehtävä 1 (tehtävä 3)	Väärin	Havaittu frekvenssi	65 _a	27 _b	37 _b	129
		% testiluokissa	47,4 %	20,3 %	27,4 %	31,9 %
	Oikein	Havaittu frekvenssi	72 _a	106 _b	98 _b	276
		% testiluokissa	52,6 %	79,7 %	72,6 %	68,1 %
	Yhteensä	Havaittu frekvenssi	137	133	135	405
		% testiluokissa	100,0 %	100,0 %	100,0 %	100,0 %

Each subscript letter denotes a subset of Vastaajat categories whose column proportions do not differ significantly from each other at the ,05 level.

$$\chi^2 = 24,75; df = 2; p < 0,001$$

Jos halutaan tarkastella taulukon yksityiskohtia, tällöin voidaan analysoida riveittäin toteutettavia z-testejä (Tähtinen ym. 2016, 169). Taulukosta 9 voidaan tulkita, että testiversion 1 vastaajien tulokset eroavat verrattuna testiversion 2 ja 3 vastaajien tuloksiin. Tarkemmin tarkasteltuna keskiarvovertailulla testiversion 1 vastaajat erosivat ankkuritehtävässä 1 testiversion 2 vastaajista tilastollisesti merkitsevästi ($z = -4,7$; $p < 0,001$) sekä testiversion 3 vastaajista tilastollisesti merkitsevästi ($z = -3,4$; $p < 0,001$). Testiversioiden 2 ja 3 vastaajien välillä ei ollut tilastollisesti merkitsevää eroa ankkuritehtävän 1 perusteella ($z = 1,4$; $p = 0,172$).

Taulukko 10. Ristiintaulukointi ankkuritehtävästä 2.

		Ristiintaulukointi				
			Testi 1	Testi 2	Testi 3	Yhteensä
Ankkuritehtävä 2 (tehtävä 6)	Väärin	Havaittu frekvenssi	81 _a	68 _a	77 _a	226
		% testiluokissa	59,1 %	51,1 %	57,0 %	55,8 %
	Oikein	Havaittu frekvenssi	56 _a	65 _a	58 _a	179
		% testiluokissa	40,9 %	48,9 %	43,0 %	44,2 %
	Yhteensä	Havaittu frekvenssi	137	133	135	405
		% testiluokissa	100,0 %	100,0 %	100,0 %	100,0 %

Each subscript letter denotes a subset of Vastaaajat categories whose column proportions do not differ significantly from each other at the ,05 level.

$$\chi^2 = 1,88; df = 2; p = 0,392$$

Taulukosta 10 nähdään, että eri testiversioiden vastaajien osaamisessa ankkuritehtävään 2 ei ole tilastollisesti merkitsevää eroa ($\chi^2 = 1,88; df = 2; p = 0,392$).

7 Pohdinta

Tämän tutkimuksen tarkoituksena oli kehittää ja pilotoida määrällisen päättelyn tehtäviä ensimmäisen luokan oppilaille. Tehtävänä oli myös pilotoida määrällisen päättelyn taitoa testaavia testiversioita jatkokehitystä varten. Testiversioiden tehtävien tulisi vaikeutua tasaisesti, jotta testin avulla voidaan erotella tutkittavien osaamista. Lisäksi tutkimuksen tarkoituksena oli selvittää, kuinka määrällistä päättelyä mittaavat kuviovaakatehtävät soveltuvat Suomessa 1.-luokkalaisten oppilaiden testaamiseen. Kehittämistyö tapahtui yhdessä Turun yliopiston oppimisanalytiikan tutkimusinstituutin kanssa ja tutkimus oli osa kaksivuotisen esiopetuksen kokeilua sekä siihen liittyvää osaamisen arviointia.

Yhteistyö Turun yliopiston oppimisanalytiikan tutkimusinstituutin kanssa alkoi alkuvuodesta 2023. Määrälliseen päättelyyn, kuviovaakatehtäviin sekä taustateoriaan perehtyminen olivat iso osa tutkimusta. Kevään 2023 aikana suunniteltiin ja toteutettiin kuviovaakatehtävät ja testiversiot sekä niiden siirtäminen ViLLE-alustalle. Kaikki kuviovaakatehtävät sekä kuviovaakatehtävissä käytetyt kuvat ja objektit on luotu tässä tutkielmassa kuvattua kehitystyötä ja tutkimusta varten. Lisäksi yhteistyössä ViLLE:n kanssa luotiin testaamista varten ohjeet sekä kirjallisena että puhuttuna audiotiedostona. Lisäksi teimme ohjevideon gif-muodossa. Roolimme yhteistyössä oli toimia tehtävien sekä ohjeiden laatimisessa pedagogisena asiantuntijana sekä käytännön toteuttajana.

Kuviovaakatehtävien kehittämisen jälkeen arvioimme keskimääräiset osaamisiät jokaiselle tehtävälle WISC-V:n normitaulukkoa apuna käyttäen. Kun jokaisella tehtävällä oli keskimääräinen osaamiskä, loimme itse kolme kymmenen kysymyksen testiversiota. Kuten aiemmin tässä tutkimuksessa on mainittu, testiversioiden tavoitteena oli, että ne erottelisivat vastaajia tasaisesti. Tämän tutkielman vahvuutena pidetään myös sitä, mitä kaikkea on tehty ennen aineistonkeruuta. Iso osa testin kehittämistyöstä on tapahtunut jo ennen aineiston keruuta ja tulosten analysointia. Yleisesti ottaen todetaan, että yhteistyö ViLLE-tiimin kanssa sujui hyvin. Tehtävien sekä testiversioiden kehitystyössä onnistuttiin.

Taidon testaamisen kehittämistyössä on oleellista selvittää, toimiko testi halutulla tavalla ja mitattiinko luotettavasti haluttua taitoa (Downing 2006). Seuraavaksi esitellään tutkittavien määrällisen päättelyn taitoja kuviovaakatehtävien osaamisen perusteella ja arvioidaan testiversioiden toimivuutta. Lisäksi pohdimme tämän tutkimuksen luotettavuutta ja eettisyyttä.

7.1 Määrällisen päättelyn taidot

Tämän tutkimuksen yhtenä tavoitteena oli selvittää ensimmäisellä luokalla olevien oppilaiden määrällisen päättelyn taitoa. Tarkoituksena oli selvittää, miten 1.-luokkalaiset oppilaat suoriutuvat määrällisen päättelyn taitoa mittaavista kuviovaakatehtävistä. Tulosten perusteella suuri osa tutkittavista oppilaista sai ensimmäisen ja toisen tehtävän oikein. Riippuen testiversiosta ensimmäisen tehtävän osasi vastaajista 89 %, 87 % ja 90 %. Toisenkin tehtävän osasi 91 %, 82 % ja 82 % vastaajista. Ennen kuviovaakatehtävien tekemistä olisi tärkeää, että lasten huomio olisi kiinnitetty tehtävän kannalta oleellisiin asioihin ja että lapset olisivat ymmärtäneet, millaisesta tehtävästä on kyse sekä miten tehtävään vastataan (Wahlstrom, Weiss & Saklofske 2016). Tulosten perusteella ensimmäiset kaksi tehtävää osattiin hyvin, joten oppilaat olivat ymmärtäneet tehtävän ja suurella osalla tutkittavista taidot riittivät vastaamaan oikein. Tulokset antavat vahvistusta sille, että pienten oppilaiden kanssa tasapainovaakaa voidaan käyttää tehtävissä (Otten ym. 2019).

Jokaisessa testiversiossa oli vastaajia, jotka osasivat kaikki kymmenen tehtävää. Tulokset osoittavatkin, kuinka suuret erot määrällisen päättelyn taidon osaamisessa ovat jo 1. luokan keväällä. Jokaisen testiversion yli 130 tutkittavan joukossa oli noin kymmenen vastaajaa, jotka saivat enimmillään yhden tehtävän oikein. Samalla jokaisessa yli 130 tutkittavan joukossa oli noin kymmenen vastaajaa, jotka saivat 9 tai 10 vastausta oikein. On kuitenkin hyvä muistaa, että osaamisen erot ovat voineet mahdollisesti kasvaa 1. luokan aikanakin. Aunolan ym. (2004) mukaan matemaattinen suorituskyyky kehittyy ensimmäisillä luokilla kumulatiivisesti, joten osaamisen erot voivat olla suuremmat 1. luokan kevätlukukaudella kuin syyslukukaudella.

Ensimmäisen luokan oppilaiden tulokset määrällisen päättelyn tehtävissä olivat pääpiirteittäin odotetun laiset. Kaksi kolmesta testiversiosta tuotti odotetun kaltaiset osaamistulokset. Jokaisessa testiversiossa neljännessä tehtävästä eteenpäin oppilaan tuli käyttää yhteenlaskun tai kertolaskun periaatteita. Tehtäviä 4, 5 ja 6 osasi kaikkiaan keskimäärin hieman alle puolet vastaajista. Tuloksia tukevat myös aikaisemmat tutkimukset. Aiemmin Kilpatrick ym. (2001) ovat todenneet, että lapset pystyvät hyödyntämään osaamiaan matemaattisia käsitteitä uusissa tilanteissa. Lisäksi hyvät matemaattiset taidot omaava lapsi pystyy uuden tehtävän parissa käyttämään tarkoituksenmukaisia keinoja ratkaistakseen tehtäviä (Kilpatrick ym. 2001, 5, 127). On kuitenkin tuloksia yleistäessä muistettava, että matemaattisten taitojen ja määrällisen päättelyn taidon testaaminen on aina rajoittuneempaa kuin määrällinen päättely lapsen

jokapäiväisessä elämässä (Dwyer ym. 2003). Yksittäinen testi on aina yksinkertaistettu versio laajan ja monipuolisen taidon testaamista varten. Tämän tutkimuksen tuloksia määrällisen päättelyn taidoista voidaan pitää hyvänä taustatietona tulevia testauksia varten.

7.2 Testiversioiden toimivuus

Tämän tutkimuksen toisen tutkimuskysymyksen avulla oli tarkoitus selvittää testiversioiden toimivuus ja vertailla kolmea eri testiversiota keskenään. Tavoitteena oli tutkia, toimivatko testiversiot tarkoituksessaan erotella oppilaita luotettavasti ja tasaisesti määrällisen päättelyn taidon osaamisen mukaan. Testiversioista parhaiten toimivat testiversio 2 ja testiversio 3. Molempien testiversioiden keskiarvo oli noin 5, vinousarvot ja huipukkuuskertoimet olivat hyvät sekä testiversion 3 vastaajista kaikki 135 vastaajaa saivat vähintään yhden tehtävän oikein. Testiversiot 2 ja 3 saivat tulokset, josta testaamisen kehitystyötä olisi tulevaisuudessa hyvä jatkaa.

Testiversio 1:n osattujen tehtävien keskiarvo oli pienin. Samalla vinousarvo oli itseisarvoltaan suurin (0,67), joten normaalijakautuneisuus oli testiversioista heikoin. Aineiston analysointivaiheessa suoritettujen ristiintaulukointien perusteella testiversion 1 vastaajat saivat ensimmäisen ankkuritehtävän (T3) perusteella merkitsevästi muiden testiversioiden tekijöitä heikommat tulokset. Tämän tuloksen perusteella voidaan olettaa testiversion 1 vastaajajoukon olleen tasoltaan kahta muuta ryhmää heikompi. Tämä vaikeuttaa testiversion 1 analysointia ja mielekkäämpää olisikin jatkaa tulevaisuudessa testiversioiden 2 ja 3 jatkokehitystä.

Osalla tutkittavista testi jäi kesken neljän peräkkäisen väärän vastauksen takia. Testistä neljällä peräkkäisellä väärällä vastauksella tippuneiden määrä kasvoi tasaisesti loppua kohden, joten tehtävien voidaan sanoa vaikeutuneen asteittain. Testityypiksi valittu monivalintatehtävä on yleisin määrällisen päättelyn taitoa mittaamaan käytetty testityyppi (Dwyer ym. 2003). Arvaaminen vaihtoehtojen avulla ei pääosin vaikuttanut auttavan, kun tutkittavan osaaminen loppui. Tehtävätyyppi ja vastausvaihtoehtojen määrä toimivat tältä osin toivotusti. Lisäksi on otettava huomioon, että monivalintatehtävissä testistä putoaminen voi aiheuttaa tuloksiin vääristymiä. Testistä pudonneet vastaajat eivät päässeet vastaamaan kaikkiin kysymyksiin ja näin ollen eivät pääse edes arvaamaan, vaan heidän vastaamatta jääneet vastauksensa on tämän tutkimuksen aineistossa korjattu vääräksi.

Yksittäisissä tehtävissä oli nähtävissä selvä osaamisen lasku verrattuna edelliseen ja seuraavaan tehtävään. Tällaisia tehtäviä olivat testiversion 1 tehtävät 3 ja 4, sekä testiversion 2 tehtävä 5. Tarkempaan laadulliseen tarkasteluun tulisi tulevaisuudessa ottaa, miksi näiden tehtävien kohdalla osaaminen oli heikkoa. Toisaalta yksittäiset pudotukset ovat luonnollisia, koska eroja tutkittavien välille oli tarkoitus saada ainoastaan kymmenen kysymyksen avulla.

7.3 Tutkimuksen luotettavuus

Tämän tutkimuksen luotettavuus perustuu tutkimuksessa käytetyn testin luotettavuuteen (Metsämuuronen 2009, 74). Lisäksi tähän tutkimukseen kuuluu itse testin kehittämistyö. Tutkimuksen validiteetin ja reliabiliteetin kannalta on tärkeä pohtia, mitä asioita tässä tutkimuksessa kehitetyillä testiversioilla testasimme. Kuviovaakatehtävillä voidaan mitata määrällistä päättelyä (Weiss ym. 2016). Tässä tutkimuksessa on luotu määrällistä päättelyä mittaavista tehtävistä erilaisia testiversioita, ja pilotoitu tehtävätyypin sekä testiversioiden käyttöä tavallisissa luokkahuoneissa valmiiksi tehdyillä ohjeistuksilla. Tämän tutkimuksen tekijöillä ei ole kuitenkaan tiedossa, kuinka tässä tutkimuksessa saatuja tuloksia aiotaan tulevaisuudessa käyttää kaksivuotisen esiopetuksen vaikutusten arviointiin.

Testaamisen tulosten ja koko tutkimuksen validiteettia parannetaan ottamalla huomioon tutkittava aihe, käsitteet ja niiden operationalisointi testin kehityksen eri vaiheissa (Downing 2006). Validiteetin varmistamisen kannalta testikehittäjien on oltava tietoisia mahdollisista testin luotettavuutta rajoittavista tekijöistä ja raportoida niistä (Linn 2006). Seuraavissa kappaleissa on esitelty tämän tutkimuksen validiteettia ja reliabiliteettia vahvistavia ja heikentäviä tekijöitä mahdollisimman tarkasti ja totuudenmukaisesti.

Määrällisessä päättelyssä on kyse monimutkaisesta ja monipuolisesta taidosta, johon linkittyvät monet matemaattiset ja kognitiiviset taidot. Tässä tutkimuksessa ei ole taustatietona vastaajien aritmeettisiä taitoja, joten niiden vaikutusta tuloksiin ei tiedetä. Taustateorian mukaan lapset ovat jo varhaiskasvatuksessa ja esikoulussa taitavia käyttämään aritmeettisiä taitoja päättelyn tukena (Kilpatrick ym. 2001). Määrällistä päättelyä arvioitaessa on tärkeää varmistua siitä, että kaikki tutkittavat hallitsevat tehtävissä käytetyt aritmeettiset sisällöt (Dwyer ym. 2003). Tässä tutkimuksessa luodut ja esiteltyt määrällisen päättelyn tehtävät pyrittiin luomaan siten, että aritmeettisten taitojen vaatimus olisi vähäistä ja tehtävissä vaadittaisiin nimenomaan määrällisen päättelyn taitoja. Aritmeettisten taitojen taustavaikutusta tämän tutkimuksen tuloksiin on vaikea arvioida. Toisaalta, kun kyse on kaksivuotisen esikoulun vaikutusten arvioinnista, aritmeettisten taitojen osaamisen näkyminen myös määrällisen päättelyn testin

tuloksissa voi olla arvokasta, kun halutaan arvioida ja kehittää yleisesti lasten matemaattista osaamista ja opetusta.

Jatkotutkimuksissa olisi kiinnostavaa tutkia, ymmärsivätkö kaikki tutkittavat tasapainovaa'an analogian. Tasapainovaakaa on aikaisemmin käytetty onnistuneesti pienten lasten opetuksessa (Otten ym. 2019). Tasapainovaaka voi olla nykylapsille vieras ja abstrakti asia, jonka toiminnan logiikan heikko ymmärrys vaikuttanee tutkittavan tuloksiin. Tässä tutkimuksessa luotiin kirjallinen ohjeistus sekä ohjevideo havainnollistamaan tasapainovaa'an toimintaa, mutta on mahdollista, ettei ohjevideo tai sanalliset ohjeet auttaneet kaikkia 6–7-vuotiaita tutkittavia ymmärtämään tarpeeksi nopeasti tasapainovaa'an toimintaa. Lisäksi oppilaiden kielelliset vaikeudet voivat vaikeuttaa esimerkiksi ohjeiden lukemista, kuuntelemista ja ymmärtämistä. Nämä seikat haastavat testauksen sisäistä validiteettia. Jälkikäteen voidaan pohtia, ollaanko tehtävien ja testiversioiden avulla mitattu ainoastaan haluttua taitoa ja sen osaamista.

Testitilanteessa tehtyjen havaintojen perusteella voidaan todeta ohjeistusten toimineen hyvin. Testitilanteen alkaessa ohjeistettiin ja muistutettiin tutkittavia seuraamaan ja lukemaan ohjeita tarkasti. Kuten Wahlstrom ym. (2016) kirjoittavat artikkelissaan kuviovaakatehtävien ohjeistuksista, myös tämän testin ohjeistuksissa kiinnitettiin lasten huomio visuaalisesti näyttäen, minkälaisesta tehtävästä on kyse sekä miten lapsi valitsee ja vastaa mielestään oikean vaihtoehdon. Harva tutkittava testitilanteen käynnistyttyä kysyi neuvoa kuviovaakaan tai tasapainovaakaan liittyen. Lasten kysymykset liittyivät enemmän tietoteknisiin pulmiin, esimerkiksi mistä painamalla pääsee eteenpäin testialustalla. Tutkimusryhmästä tulleen tukihenkilön avulla ohjeiden tarkkaa lukemista ja ohjevideon katsomista sekä kuuntelemista painotettiin varmastikin enemmän kuin tulevaisuuden testitilanteessa, jos opettaja teettää yksin testin luokkansa oppilaille. Yleisesti voidaan todeta, että tasapainovaaka ei ohjeistusten jälkeen aiheuttanut tutkittavissa kysymyksiä ja suuri osa tutkittavista osasi testin ensimmäiset tehtävät.

Kuviovaakatehtävien valinta määrällisen päättelyn taidon testaamiseen perustuu osaksi siihen, että tutkittavia voidaan testata tulevaisuudessakin samankaltaisilla kuviovaakatehtävillä. Päättelytehtäviä, kuten kuviovaakatehtäviä, voidaan vaikeuttaa käytännössä loputtomasti lisäämällä tehtävässä vaadittujen päättelyprosessien määrää sekä päättelyn kompleksisuutta (Dwyer ym. 2003). Sisäisen validiteetin arvioinnissa on otettava huomioon, että määrien suhteiden vertailu on tarpeeksi suuressa osassa, eivätkä tehtävät vaikeutuisi vain

päätelyprosessien määrän tai kompleksisuuden perusteella. Tavoite määrällisen päättelyn testaamisessa on mitata määrien välisten suhteiden vertailun hallintaa. Lisäksi sisällön validiteetin pohdinnan kannalta on hyvä huomata, että tässä tutkimuksessa kaikkialla ihmisen elämässä tapahtuva määrällinen päättely ja määrällisen päättelyn taito on operationalisoitu testaamisen mahdollistamiseksi kuviovaakatehtäviin. Tulevaisuudessa olisi mielenkiintoista tutkia suomalaisten lasten määrällisen päättelyn taitoa laadullisemmin esimerkiksi tehtävien ja haastatteluiden avulla, jolloin lasten osaamista sekä omaa ajattelua päätelyprosessien taustalla saataisiin tutkittua.

Eri testiversioita vertailtaessa ja koko tutkimusta sekä sen tuloksia arvioitaessa on otettava huomioon, että jokaisella testiversiolla oli omat vastaajajoukkonsa. Vastaajajoukot muodostuivat satunnaisotannalla, jokainen oppilas sai satunnaisesti tehtäväksi yhden kolmesta testiversiosta, eikä oppilaan luokka tai koulu vaikuttanut siihen, minkä testiversion tutkittava sai tehtäväkseen. Useamman vastaajajoukon avulla tähän tutkimukseen pystyttiin sisällyttämään kolmen erilaisen testiversion pilotointi. Määrällisen päättelyn testin lisäksi oppilaat tekivät aineistonkeruussa muitakin testejä, joten ei ollut tarkoituksenmukaista teettää tutkittavilla useita kymmeniä määrällisen päättelyn tehtäviä. Tiedossa oli testin kehittämistyön alusta asti, että eri vastaajajoukot tulevat tekemään eri testiversiot. Tämän vuoksi testiversioihin suunniteltiin ja sisällytettiin ankkuritehtäviä. Ankkuritehtävien perusteella pystyttiin vertailemaan eri vastaajajoukkoja. Testiversion 1 vastaajajoukko olikin ensimmäisen ankkuritehtävän perusteella tilastollisesti merkitsevästi kahta muuta vastaajajoukkoa heikompi. Tämä tieto vaikeuttaa testiversion 1 tarkastelua ja analysointia. Samalla se osoittaa laajemmin, että vaikka vastaajia jokaiselle testiversiolle on ollut yli 130 oppilasta, vastaajajoukot voivat erota osaamiseltaan merkittävästi toisistaan. Tämän tutkielman luotettavuuden kannalta lukijan on hyvä ottaa huomioon, että jokaisessa testiversiossa oli satunnaisesti valikoitu vastaajajoukko. Tämä valinta tehtiin, jotta pilotointiin saadaan mukaan useampi testiversio, mutta tutkittavien lasten ei tarvitse tehdä kaikkia 26 tehtävää.

Tulosten osalta luotettavuutta haastaa eri vastaajajoukkojen lisäksi tehtävien määrä, joka oli jokaisen testiversion osalta kymmenen tehtävää. Tulevaisuudessa oppilaita arvioidaan usean osatestin avulla, mutta testaukseen aikaa on rajallisesti. Testin on oltava lyhyt, mutta oppilaiden osaamista erotteleva. Lisäksi on luonnollista, ettei lapsia voi lopullisessa versiossa kuormittaa liiallisella määrällä tehtäviä arvioitaessa osaamista ja taitoja. Tehtävien määrä on

aina kompromissi taidon osaamisen arvioinnin luotettavuuden ja tutkimukseen käytettävissä olevien resurssien suhteen.

Kun halutaan tarkastella testiversioiden luotettavuutta ja mittausvirheitä, yksi tapa on tarkastella reliabiliteettia (Linn 2006; Metsämuuronen 2009). Testiversioiden korjattujen aineistojen Cronbachin alphas ($0,831$; $0,768$; $0,771$) tukevat testiversioiden luotettavuutta. Tämän tutkimuksen tulokset antavat samankaltaisia tuloksia kuin Nunesin ym. (2015) tutkimus, jossa osoitettiin, että määrällisen päättelyn taitoa voi luotettavasti mitata alle 7-vuotiailta lapsilta. Testiversioiden reliabiliteetit osoittavat, että tehtävät mittaavat luotettavasti samaa asiaa. Nämä tulokset ovat yhteneväiset aiempien tutkimusten kanssa, joissa kuviovaakatehtävien on todettu mittaavan luotettavasti määrällistä päättelyä (Weiss ym. 2016). On samalla huomioitava, että tässä tutkimuksessa esitetty pilottitestaus suoritettiin yli puoli vuotta myöhemmin kuin todellinen testaus tullaan suorittamaan. On oletettavaa, että tehtävien osaaminen olisi ollut syyslukukaudella heikompaa. Kuitenkin testiversioiden vertailua ja tehtävien välistä vertailua toisiinsa pystytään riittävän luotettavasti tekemään.

Testin kehittämistyöhön ja pilotointiin kuuluu koko kehittämistyön systemaattinen raportointi sen vaiheista ja tuloksista (Downing 2006). Tässä tutkielmassa on pyritty esittämään kehittämistyön vaiheet ja tulokset mahdollisimman selkeästi. Johdannon ja tutkimuskysymysten avulla osoitetaan tämän tutkielman yhteiskunnallinen tarve sekä tutkijoiden perehtyneisyys aiheeseen. Tutkimusmenetelmissä kuvaamme mahdollisimman tarkasti, kuinka testiversioiden kehittämistyötä, aineistonkeruuta sekä aineiston analysointia on suoritettu. Tulososiossa olemme esitelleet tärkeimmät tulokset niin tämän työn kuin yleisesti kasvatustieteen kannalta. Pohdinnassa olemme pohtineet työn eri vaiheita, jotta lukijakin voi arvioida työn onnistumista ja luotettavuutta.

7.4 Tutkimuksen eettisyys

Tässä tutkimuksen toteutuksessa mukana olivat tämän tutkielman kirjoittajat yhteistyössä Turun yliopiston oppimisanalytiikan tutkimusinstituutin kanssa. Kuten aiemmin menetelmäosiossa on selostettu, tutkittaville ja heidän perheilleen annettiin tietoa tutkimuksesta. Wilma-järjestelmän kautta lähetetyn huoltajatiedotteen (Liite 1) avulla saatiin suoritettua informointi niin, että huoltajien oli mahdollista kieltää lapsen osallistuminen tutkimukseen. Kuten tutkimuseettinen neuvottelukunta (2019) ohjeistaa, informointi riittää, kun tutkittavia oli yli 400. Vanhempien informoinnin lisäksi ennen jokaista aineistonkeruuta luokassa kerrottiin lapsille tutkimuksesta ja annettiin mahdollisuus kieltäytyä osallistumasta

tutkimukseen. Informoinnin sijasta olisi tutkimuseettisesti ollut parempi pyytää tutkittavilta ja heidän huoltajiltaan suostumuksellinen tutkimuslupa.

Tätä tutkielmaa varten kirjoittajien ViLLE-tiimiltä saama aineisto ei sisällä tunnisteita oppilaista eikä edes luokista tai kouluista. Vastauksia ja tutkittavia ei pysty tässä tutkimuksessa kerätyn ja analysoidun aineiston perusteella yhdistämään toisiinsa.

Tutkielmassa ja sen vaiheissa alaikäisten tutkimukseen liittyvät eettiset asiat on pyritty mahdollisimman hyvin ottamaan huomioon ja kaikki tutkimuksen vaiheet on raportoitu mahdollisimman selkeästi sekä totuudenmukaisesti.

Lähteet

- Alexander, P., White, C., Daugherty, M. (1997). Analogical reasoning and early mathematics learning. Teoksessa Lyn, D. (toim.), *Mathematical Reasoning: Analogies, Metaphors and Images*. Mahwah (NJ): Lawrence Erlbaum Associates.
- Aunio, P., Hannula, M. & Räsänen, P. (2004). Matemaattisten taitojen varhaiskehitys. Teoksessa T. Ahonen, P. Kupari, P. Malinen & P. Räsänen. (toim.), *Matematiikka - näkökulmia opettamiseen ja oppimiseen*. Jyväskylä: Niilo Mäki Instituutti, 198–221.
- Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology*, 96(4), 699–713. <https://doi.org/10.1037/0022-0663.96.4.699>
- Barros Baertl, R. (2020). *Factors Contributing to Children’s Understanding of Fractions: Quantitative Reasoning and Arithmetic*. University of Oxford.
- Downing, S. (2006). Twelve steps for effective test development. Teoksessa Downing, S. & Haladyna, T. (toim.), *Handbook of Test Development*. Routledge.
- Dwyer, C., Gallagher, A., Levin, J., & Morley, M. (2003). What is quantitative reasoning? Defining the construct for assessment purposes. *ETS Research Report Series*. (2), i-48.
- Gentner, D. & Maravilla, F. (2018). Analogical reasoning. Teoksessa L. J. Ball & V. A. Thompson (toim.), *International Handbook of Thinking & Reasoning*. NY, NY: Psychology Press.
- Halberda, J., Mazocco, M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*. 455(7213), 665–668. <https://doi.org/10.1038/nature07246>
- Halinen, I., Hotulainen, R., Kauppinen, E., Nilivaara, P., Raami, A., & Vainikainen, M.-P. (2016). *Ajattelun taidot ja oppiminen*. Jyväskylä: PS-kustannus.
- Hannula, M. & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, 15(3), 237–256. <https://doi.org/10.1016/j.learninstruc.2005.04.005>
- Hannula-Sormunen, M., Lehtinen, E. & Räsänen, P. (2015). Preschool children’s spontaneous focusing on numerosity, subitizing, and counting skills as predictors of their mathematical performance seven years later at school. *Mathematical Thinking and Learning*, 17(2–3), 155–177. <https://doi.org/10.1080/10986065.2015.1016814>
- Hannula-Sormunen, M., Mattinen, A., Räsänen, P. & Ruusuvirta, T. (2018). Varhaisten matemaattisten taitojen perusta: synnynnäiset valmiudet, tietoinen toiminta ja

- vuorovaikutus. Teoksessa Joutsenlahti, J., Räsänen, P., & Silfverberg, H. (toim.), *Matematiikan opetus ja oppiminen*. Jyväskylä: Niilo Mäki Instituutti.
- Kaksivuotisen esiopetuksen kokeilun opetussuunnitelman perusteet 2021 (2021). Helsinki: Opetushallitus. (Viitattu 12.1.2023).
https://www.oph.fi/sites/default/files/documents/Kaksivuotisen_esiopetuksen_kokeilun_opetussuunnitelman_perusteet_2021.pdf
- Kilpatrick, J., Swafford, J. & Findell, B. (2001). *Adding it Up: Helping Children Learn Mathematics*. Washington, D.C: National Academies Press.
<https://doi.org/10.17226/9822>
- Knuth, E., Stephens, A., McNeil, N., & Alibali, M. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, 37(4), 297–312.
- Kupari, P. & Hiltunen, J. (2018). Matemaattiset taidot kansainvälisen arviointitutkimuksen valossa. Teoksessa Joutsenlahti, J., Räsänen, P., & Silfverberg, H. (toim.), *Matematiikan opetus ja oppiminen*. Jyväskylä: Niilo Mäki Instituutti.
- Laki kaksivuotisen esiopetuksen kokeilusta 2020/1046. Annettu Helsingissä 17.12.2020.
<https://www.finlex.fi/fi/laki/ajantasa/2020/20201046>
- Lehtonen, D. (2022). 'Now I Get It!': Developing a Real-world Design Solution for Understanding Equation-solving Concepts. <https://urn.fi/URN:ISBN:978-952-03-2250-2>
- Leppäaho, H. (2007). Matemaattisen ongelmanratkaisutaidon opettaminen peruskoulussa. Ongelmanratkaisukurssin kehittäminen ja arviointi. *Jyväskylä Studies in Education, Psychology and Social Research* 298. Jyväskylä: Jyväskylän yliopisto.
- Leppäaho, H. (2018). Ongelmanratkaisun opettamisesta. Teoksessa Joutsenlahti, J., Räsänen, P., & Silfverberg, H. (toim.), *Matematiikan opetus ja oppiminen*. Jyväskylä: Niilo Mäki Instituutti.
- Liljedahl, P., & Santos-Trigo, M. (2019). *Mathematical Problem Solving: Current Themes, Trends, and Research*. Cham: Springer International Publishing AG.
- Linn, R. (2006). The standards for educational and psychological testing: Guidance in test development. Teoksessa Downing, S. & Haladyna, T. (toim.), *Handbook of Test Development*. Routledge.
- Metsämuuronen, J. (2002). *Mittarin rakentaminen ja testiteorian perusteet* (2. uud. p.). Helsinki: International Methelp.

- Metsämuuronen, J. (2009). Tutkimuksen tekemisen perusteet ihmistieteissä: tutkijalaitos (4. laitos.). Helsinki: International Methelp.
- Metsämuuronen, J. & Räsänen, P. (2018). Opettaja mittaajana. Teoksessa Joutsenlahti, J., Räsänen, P., & Silfverberg, H. (toim.). Matematiikan opetus ja oppiminen. Jyväskylä: Niilo Mäki Instituutti.
- Nummenmaa, L. (2009). Käyttäytymistieteiden tilastolliset menetelmät (1. p., uud. laitos.). Helsinki: Tammi.
- Nunes, T., Bryant, P., Barros, R., and Sylva, K. (2012). The relative importance of two different mathematical abilities to mathematical achievement. *British Journal of Educational Psychology*, 82(1), 136–156.
- Nunes, T., Bryant, P., Evans, D., & Barros, R. (2015). Assessing quantitative reasoning in young children. *Mathematical Thinking and Learning*, 17(2–3), 178–196.
- Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A. & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology*, 25(1), 147–166
- Otten, M., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (2019). The balance model for teaching linear equations: A systematic literature review. *International Journal of STEM Education*, 6(30). <https://doi.org/10.1186/s40594-019-0183-2>
- Silver, E. (1987). Foundations of cognitive theory and research for mathematics problem-solving instruction. Teoksessa Schoenfeld, A. (toim.), *Cognitive Science and Mathematics Education*. Hillsdale, N.J: Lawrence Erlbaum.
- Thompson, P. (1993). Quantitative reasoning, complexity, and additive structures. *Educational Studies in Mathematics*, 25(3), 165–208.
- Tjoe, H. (2019). “Looking Back” to solve differently: familiarity, fluency, and flexibility. Teoksessa Liljedahl, P., & Santos-Trigo, M. (toim.), *Mathematical Problem Solving: Current Themes, Trends, and Research*. Cham: Springer International Publishing AG.
- Tutkimuseettinen neuvottelukunta. (2019). Ihmiseen kohdistuvan tutkimuksen eettiset periaatteet ja ihmistieteiden eettinen ennakoarviointi Suomessa. Tutkimuseettisen neuvottelukunnan ohje 2019. Tutkimuseettisen neuvottelukunnan julkaisuja 3/2019. Toinen, uudistettu painos. Helsinki: Tutkimuseettinen neuvottelukunta. https://tenk.fi/sites/default/files/2021-01/Ihmistieteiden_eettisen_ennakoarvioinnin_ohje_2020.pdf (Viitattu 29.5. 2023).
- Tähtinen, J., Laakkonen, E., & Broberg, M. (2020). Tilastollisen aineiston käsittelyn ja tulkinnan perusteita. Turun yliopisto.

- Wahlstrom, D., Weiss, L. & Saklofske, D. (2016) Practical issues in WISC-V administration and scoring. Teoksessa Weiss, L., Saklofske, D., Holdnack, J., & Prifitera, A. (toim.), WISC-V Assessment and Interpretation: Scientist-Practitioner Perspectives. San Diego: Elsevier Science & Technology.
- Weiss, L., Saklofske, D., Holdnack, J., & Prifitera, A. (2016). WISC-V: Advances in the assessment of intelligence. Teoksessa Weiss, L., Saklofske, D., Holdnack, J., & Prifitera, A. (toim.), WISC-V Assessment and Interpretation: Scientist-Practitioner Perspectives. San Diego: Elsevier Science & Technology.

Liitteet

Liite 1. Huoltajatiedote FIN



Arvoisa huoltaja,

Turun yliopiston oppimisanalytiikan tutkimusinstituutti ja Turun kaupunki kehittävät yhteistyössä luku- ja kirjoitustaitojen ja matematiikan oppimisen arvioinnin tehtäviä tulevien tutkimusten käyttöön.

Osana tätä yhteistyötä lapsenne luokka osallistuu yhden oppitunnin aikana uusien tehtävien kokeiluun. Tehtäviä on kolme:

- 1) Vaakatehtävä, missä lapsen pitää päätellä, mitä pitää laittaa vaakaan, jotta se olisi tasapainossa.
- 2) Oikeinkirjoitustehtävä, jossa lapsi kuulee kuulokkeista sanan, joka kirjoitetaan tietokoneen/tabletin näppäimillä ruudulle.
- 3) Sanantunnistus, missä lapsi kuulee sanan ja hänen pitää nappia painamalla kertoa, oliko sana oikea sana vai keksitty sana.

Tehtävät tehdään luokassa oppitunnin aikana tietokoneella/tabletilla. Luokassa on opettajan apuna tutkimusavustaja auttamassa lapsia tehtävien tekemisessä.

Lapsista **ei** kerätä mitään tunnistetietoja. Siis lapsen nimeä, ikää tai sukupuolta ei kysytä. Tutkimusaineistossa ei ole mitään tietoa lapsestanne. Vain kaikkien lasten suoritukset näissä tehtävissä.

Mikäli ette halua, että lapsenne osallistuu näiden tehtävien tekemiseen koululuokassa ja nimettömien suoritustietojen käyttämiseen tutkimuksessa, ilmoitatteko siitä opettajalle.

Tuloksia tullaan käyttämään tutkimustehtävien kehittämiseksi sekä opinnäytetöiden aineistona. Aineisto, joka ei sisällä mitään tietoa oppilaasta, luokasta tai koulusta tullaan säilyttämään tutkimustarkoituksiin.

Lisätietoja tutkimuksesta antavat

Mikko-Jussi Laakso
apulaisprofessori / johtaja
Oppimisanalytiikan tutkimusinstituutti
Turun yliopisto

milaak@utu.fi

Pekka Räsänen
työelämäprofessori/varajohtaja
Oppimisanalytiikan tutkimusinstituutti
Turun yliopisto

pekka.j.rasanen@utu.fi

Liite 2. Huoltajatiedote SV



Bästa vårdnadshavare,

Forskningsinstitutet för lärandeanalys vid Åbo universitet och Åbo stad samarbetar för att utveckla uppgifter för bedömning av läs- och skrivkunnighet samt matematikinläring för framtida studier.

Som en del av detta samarbete kommer ditt barns klass att delta i ett experiment med nya uppgifter under en lektion. Det finns tre uppgifter:

- 1) En våguppgift uppgift där barnet ska avgöra vad som ska läggas på vågen för att balansera den.
- 2) En stavningsuppgift där barnet hör ett ord via hörlurar, som sedan ska skrivas på skärmen med hjälp av tangenterna på datorn/pekplattan.
- 3) Ordigenkänning, där barnet hör ett ord och ska sedan trycka på en knapp för att avgöra om ordet är ett riktigt ord eller ett påhittat ord.

Uppgifterna görs i klassrummet under lektionstid på en dator/pekplatta. Läraren har en forskningsassistent i klassrummet som hjälper barnen med uppgifterna.

Ingen identifierande information samlas in om barnen. Det vill säga, barnets namn, ålder eller kön frågas inte. Det finns ingen information om ditt barn i undersökningsmaterialet. Endast barnens prestationer från dessa uppgifter.

Om du inte vill att ditt barn ska delta i dessa uppgifter i klassrummet och att anonyma resultatuppgifter ska användas i undersökningen, vänligen informera läraren.

Resultaten kommer att användas för utveckling av uppgifter och publiceras som examensarbete. Uppgifter utan information om barnen, klassen eller skolan sparas för forskningsändamål.

För mer information om studien, vänligen kontakta

Mikko-Jussi Laakso
biträdande professor / chef
Forskningsinstitutet för lärandeanalys
Åbo universitet

milaak@utu.fi

Pekka Räsänen
arbetslivsprofessor / vice chef
Forskningsinstitutet för lärandeanalys
Åbo universitet

pekka.j.rasanen@utu.fi

Liite 3. Huoltajätiedote ENG



Dear Caregiver,

The Research Institute for Learning Analytics at the University of Turku and the City of Turku are collaborating to develop tasks for assessing literacy and mathematics learning for the forthcoming research projects.

As part of this collaboration, your child's class will participate in a one-lesson trial of the new learning tasks. There will be three tasks:

- 1) A balance task, where the child must decide what to put on the scale to balance it.
- 2) A spelling task, where the child hears a word and needs to write what she/he heard.
- 3) Word recognition, where the child hears a word and has to press a button to indicate if the word is a real word or a made-up word.

The tasks are done in class during the regular classroom lessons using the computer/tablet and headphones. The teacher has a research assistant in the classroom to help the children with the tasks.

No identifying information is collected from the children. In other words, the child's name, age or gender will not be asked. There is no information about your child in the survey data. We collect only the performance of all children on these tasks.

If you do not want your child to participate in these tasks in the classroom and to have anonymous performance data used for research, please let the teacher know.

The results will be used for task development and published as thesis works. The data without any information related to the children, class or school will be saved for research purposes.

For more information about the survey, please contact

Mikko-Jussi Laakso
Associate professor / director
Research Centre for Learning analytics
University of Turku

milaak@utu.fi

Pekka Räsänen
Professor of practice/vice director
Research Centre for Learning analytics
University of Turku

pekka.j.rasanen@utu.fi

Liite 4. Ohjevideon puhuttu ja kirjoitettu teksti

Puhutut ohjeet ohjevideossa:

”Tämä on vaaka.” (Heti ohjevideon alkuun)

”Vaaka on tasapainossa, kun molemmilla puolilla on yhtä paljon.” (Kun videossa vaaka tasapainottuu).

”Mitä vaa’an toiselle puolelle tulee?” (Kun on kolmio toisella puolella ja kysymysmerkki toisella puolella.)

”Valitse oikea vaihtoehto.”

”Oikea vastaus on sininen kolmio.” (Kun vastaus tulee esiin.)

”Vaaka on tasapainossa.” *”Molemmilla puolilla on saman verran.”* (Lopuksi)

Ohjeteksti, joka näkyy ohjevideon jälkeen yhteenvetona:

Tämä on vaaka.

Vaaka on tasapainossa, kun molemmilla puolilla on yhtä paljon.

Mitä vaa’an toiselle puolelle tulee?

Valitse oikea vaihtoehto ja siirry seuraavaan kysymykseen.