

**Alakouluikäisten itsearvioinnin  
pysyvyys ja tarkkuus**

Leena Lahtinen

Erityispedagogiikan pro gradu -tutkielma  
Kevätlukukausi 2015  
Erityispedagogiikan laitos  
Jyväskylän yliopisto

# TIIVISTELMÄ

**Lahtinen, Leena. 2015. Alakouluikäisten itsearvioinnin pysyvyys ja tarkkuus. Erityispedagogiikan pro gradu -tutkielma. Jyväskylän yliopisto. Erityispedagogiikan laitos.**

Pro gradu -tutkielmassa tarkastellaan alakouluikäisten 1) matematiikan, lukemisen, tehtävissä ponnistelemisen, keskittymisen sekä oppimiseen liittyvien tunteiden itsearvioinnin pysyvyyttä testi-uusintatesti -reliabiliteetin avulla ja 2) matematiikan ja lukemisen taitojen itsearvioinnin tarkkuutta.

Tutkimus on osa Jyväskylän yliopiston ja Niilo Mäki Instituutin Minäpystyvyys ja oppimisvaikeusinterventiot -hanketta, johon osallistui 1344 oppilasta Keski-Suomesta ja Mikkelistä vuosina 2013–2015. Tutkimuksessa oppilasryhmät vastasivat kyselylomakkeisiin koulutettujen testaaajien johdolla kahtena eri ajankohtana.

Tutkimustulokset osoittivat testi-uusintatesti -reliabiliteetin olevan korkea matematiikan, lukemisen, tehtävissä ponnistelemisen, keskittymisen ja oppimiseen liittyvien tunteiden itsearvioinneissa. Matematiikan ja lukemisen taitojen sekä oppimiseen liittyvien tunteiden osalla tyttöjen itsearviot olivat pysyvämpiä kuin pojilla. Lisäksi vanhemmat oppilaat kykenivät arvioimaan minäpystyvyyttään ja taitojaan nuorempia oppilaita pysyvämmiin kahdella eri mittauskerralla, vaikkakin tuloksissa oli poikkeavuutta siinä, missä iässä tilastolliset erot syntyivät. Toisaalta matematiikan ja lukemisen itsearvioinnin tarkkuudessa ei ollut tilastollisesti merkitsevää eroa tyttöjen ja poikien eikä eri luokkatasojen välillä.

Tutkimustulosten perusteella itsearvioitavalla asialla on vaikutusta itsearvioinnin pysyvyyteen. Lisäksi itsearvioinnin tarkkuuden tulokset osoittavat sen, ettei 2.-5.-luokkalaisten sukupuolella eikä luokkatasolla ole merkitystä itsearvioinnin ja suoriutumisen väliseen yhteyteen. Tämän vuoksi itsearviointia on perusteltua käyttää jo alakoulussa.

Asiasanat: Itsearviointi, itsearvioinnin pysyvyys, itsearvioinnin tarkkuus, suoriutuminen, testi-uusintatesti -reliabiliteetti, validiteetti, minäpystyvyys

# SISÄLTÖ

Johdanto.....	4
Itsearviointin tarkkuuteen vaikuttavia tekijöitä.....	5
Itsearviointin reliabiliteetti ja validiteetti.....	11
Tutkimusongelmat.....	13
Metodi.....	14
Tutkimushanke.....	14
Osallistujat.....	14
Mittarit.....	15
Aineiston analysointi.....	18
Tulokset.....	19
Itsearviointin pysyvyys.....	19
Itsearviointin tarkkuus.....	23
Pohdinta.....	24
Tulosten tarkastelua.....	24
Itsearviointi erityisopetuksessa.....	28
Tutkimuksen luotettavuus, rajoitukset ja yleistettävyys.....	30
Jatkotutkimushaasteita.....	33
Lähteet.....	34
Liite.....	38

## JOHDANTO

Nykyajan konstruktivistinen oppimiskäsitys painottaa aktiivista, osallistavaa, itseohjautuvaa ja vuorovaikutteista oppimista. Tällaisella uudella oppimiskulttuurilla on myös vaikutuksensa oppilasarviointiin ja sen uudistamiseen. Nykyaikainen arviointi osallistaakin oppilaan oman oppimisensa prosessiin ja sen arviointiin. Tämän mahdollistamiseksi muita arviointimenetelmiä täydentämään on kehitetty itsearviointi, joka on kasvattanut suosiotaan osana koulujen oppilasarviointia.

Itsearvioinnin käsitteen kontekstisidonnaisuuden vuoksi täytyy selvittää, mitä itsearvioinnilla kussakin yhteydessä tarkoitetaan, mitä sillä arvioidaan ja mihin tarkoitukseen saatuja tuloksia sovelletaan (Huusko, 2008, 129). Koulukontekstissa itsearviointia voidaan kuvata arvioinnin kivijalaksi, joka muodostaa oppilasarvioinnin perustan (Halinen, 1995, 99), kun taas oppilasarvioinnissa oppilas arvioi oppimistaan tai työtään sitä kehittäääkseen (Alaoutinen, 2012; Huusko, 2008, 129).

Perusopetuslain (22 §) mukaan oppilasarvioinnin tulee arvioida monipuolisesti oppilaan oppimista, työskentelyä ja käyttäytymistä. Kun arviointiin sisällytetään monipuolisesti eri arviointimenetelmiä, siitä muodostuu kattava ja oppilaan kasvua tukeva prosessi. Itsearvioinnit täydentävätkin oppilaskeskeisyydellään perinteistä kynä-paperi- ja käyttäytymisen arviointia, jossa opettajalla on tavanomaisesti ollut suuri rooli (Sung, Chang, K., Chang, T., & Yu, 2010). Itsearviointikyvyn harjoittamisen tulisikin olla osa jokaisen koulun arviointijärjestelmää. Kun itsearviointi kulkee toimintatapana muun opiskelun ohessa, sitä ei koeta raskaaksi eikä ajanpuutteestakaan koidu ongelmia (Salmio & Vainio, 1995, 10).

Itsearviointia käytetäänkin paljon, mutta silti sen toimivuutta ja luotettavuutta on tutkittu vain vähän (Tari, 2010; Sung ym., 2010). Tutkimuksia itsearvioinnin tarkkuudestakin on vain lähinnä vanhemmilta oppilailta yläkoulusta, lukiosta ja korkeakoulusta (Alaoutinen, 2012; Beyer, 2002; Dupeyrat, Escribe, Huet, & Regner, 2011; Pallier, 2003), mutta tätä nuorempien itsearvioinnin luotettavuudesta on vain vähän tutkimustietoa. Koska edes kaikki lukio- ja korkeakouluikäiset eivät osaa arvioida taitojaan tarkasti (Beyer, 2002; Dupeyrat ym., 2011), on myös alettu kyseenalaistaa itsearvioinnin hyödyt ja luotettavuus (Lew, Alwis, & Schmidt, 2010; Mabe & West 1982). Lasten ja aikuisten välisten piirteiden merkittävien erojen vuoksi on tärkeää tutkia sitä, onko itsearviointia mielekästä toteuttaa lapsilla.

### **Itsearvioinnin tarkkuuteen vaikuttavia tekijöitä**

Aiemmissä tutkimuksissa itsearvioinnin tarkkuutta on tarkasteltu joko vertaamalla itsearviointia oppilaan omaan suoriutumiseen, vertaisarviointiin tai opettajan tekemään arviointiin. Tämän vuoksi erot aiempien tutkimustulosten välillä voivat johtua myös erilaisista tutkimusasetelmista itsearvioinnin tarkkuuden tutkimuksissa. Tässä tutkimuksessa itsearvioinnin tarkkuudesta puhuttaessa tarkoitetaan itsearvioinnin ja oman suoriutumisen välistä yhteyttä. Blatchfordin (1997) mukaan itsearvioinnin tarkkuuden tutkimisessa paras tapa onkin verrata itsearviointia omaan suoriutumiseen.

Itsearvioinnin tarkkuus lisääntyy iän myötä (Alaoutinen, 2012; Blatchford, 1997; Kaderavek, Gillam, Ukrainetz, Justice, & Eisenberg, 2004). Erimielisyyttä on kuitenkin siitä, missä iässä tämä tapahtuu. Täsmällistä ikää itsearviointitaitojen tarkoiksi kehittymiseen onkin haastavaa määritellä, koska itsearvioinnin tarkkuutta on tutkittu enimmäkseen vain lukio- ja korkeakouluikäisillä opiskelijoilla sekä alle kouluikäisillä lapsilla. Nuorten aikuisten itsearviointi on luotettavaa (Blatchford, 1997), mutta kouluikäisiä nuoremmilla on taipumusta yliarvioida taitojaan (Blatchford, 1997; Kaderavek ym., 2004).

Koulutulokas osaa arvioida fyysiset ominaisuutensa jo kohtuullisen tarkasti, mutta psyykkisten ominaisuuksien arviointi on huomattavasti

vaikeampaa (Aho, 2004). Koulun alkaessa lapsen käsitys itsestään on hyvin positiivinen, mikä ilmenee lapsen taipumuksena yliarvioida taitojaan (Blatchford, 1997; Kaderavek ym., 2004). 7-vuotiaat eivät siis ole vielä tarkkoja saavutustensa arvioijia.

Toisella ja kolmannella luokalla lasten itsearvioinnit ovat keskimäärin luotettavia, koska tällöin itsearviointitaidot ovat kehittyneet niin, että ne vastaavat opettajien ja luokkakavereiden tekemiä arviointeja (Aho, 2004; Blatchford, 1997). Myös Piaget'n vaiheteoria tukee tätä itsearvioinnin tarkkuuden lisääntymistä; käsitys kun oman minän pysyvyydestä syntyy hänen mukaansa 7–10 vuoden iässä (Ihme, 2009, 97). Kymmenvuotias erottaa jo todellisen minäkuvansa ihanneminäkuvastaan (Aho, 2004), mikä voi vaikuttaa itsearvioinnin tarkkuuteen.

Epätarkkuus itsearvioinnissa saattaa johtua ensinnäkin itsearviointeihin tottumattomuudesta, jolloin oppilaalta puuttuu kyky ja itseluottamus arvioida omia taitojaan (Ng & Earl, 2008; Olin & Sullivan, 2004). Toiseksi itsearviointien epätarkkuus voi selittyä sillä, että 7-vuotiaan lapsen ajattelu on vielä konkreettista, kun taas itsearviointiin liittyvät metakognitiiviset taidot vaativat ajattelua formaalien operaatioiden tasolla (Kaderavek ym., 2004). Kolmanneksi lasten suurelliset arviot taidoistaan saattaa johtua siitä, että lapset tulkitsevat todellisuutta aikuisia enemmän halujensa ja toiveidensa mukaan (Butler, 1990). Butlerin mukaan lasten käsitys omista taidoistaan onkin vielä eriytymätön, jolloin kouluikäistä nuoremmat tuntevat olevansa hyviä, jos he ovat ponnistelleet kovasti tehtävässä ja päässeet tavoitteeseen.

Lisäksi koulutulokkaat eivät ole vielä tottuneet vertaamaan taitojaan vertaisten taitoihin, opettajan arviointeihin tai saamiinsa arvosanoihin (Kaderavek ym., 2004). Vanhemmilla lapsilla onkin nuorempia lapsia enemmän kokemusta omien taitojen arvioinnista ja opettajalta saadusta palautteesta sekä tietoa aiemmasta suoriutumisestaan, mitkä tukevat oman arviointikyvyn kehittymistä (Blatchford, 1997; Mabe & West, 1982). Kokemus itsearvioinnista syntyykin usein iän myötä, kun takana on mahdollisesti säännöllisesti vuosittain, kuukausittain tai jopa viikoittain kertynyttä harjoitusta oman oppimisen arvioinnista.

Itsearviointitaidon oppiminen ei olekaan itsestänselvyys, vaan se on muiden taitojen tavoin harjoittelua vaativaa. Jos omien taitojen arvioimista kuitenkin harjoitellaan säännöllisesti, itsearviointien tarkkuus lisääntyy (Alaoutinen, 2012; Olina & Sullivan, 2004). Itsearviointitaitojen opettaminen onkin välttämätöntä, jotta menetelmän avulla saatu tieto olisi luotettavaa.

Lapsen iän lisäksi sukupuolen rooli on huomioitava itsearvioinnin tarkkuutta tutkittaessa, koska sukupuolen merkitys tuloksiin vaihtelee tutkimusten perusteella sen mukaan, missä asioissa itsearviointia on tarkasteltu (Kaderavek ym., 2004; Pallier, 2003). Aiempien tutkimustulosten yleistämisessä täytyy olla varovainen, koska suurin osa tutkimuksista on toteutettu joko aikuisilla tai nuorilla aikuisilla. Lisäksi sukupuolten välisiä eroja on tutkittu hyvin erilaisissa asioissa, eikä kattavaa tutkimusta ole tehty yhden oppiaineen tai taidon osalta.

Naisten ja miesten välillä itsearvioinnin tarkkuudessa havaittiin eroja matematiikan, liikkumisen ja poliittisen osaamisen itsearvioinneissa, joissa naiset aliarvioivat kykynsä, kun taas miehet olivat itsearvioissaan joko tarkkoja tai taitonsa yliarvioivia (Beyer, 2002; Dupeyrat ym., 2011). Lisäksi Pallierin (2003) tutkimuksessa miehet olivat hieman naisia tarkempia itsearvioimaan yleistietämystään. Testissä painottuivat kuitenkin tieteen ja teknologian kysymykset, joista miehet suoriutuvat yleensä naisia paremmin. Päinvastoin 5–12-vuotiaat tytöt olivat saman ikäisiä poikia tarkempia tarinan tuottamiseen liittyvässä itsearvioinnissa (Kaderavek ym., 2004). Toisaalta äidinkielen, historian ja maantiedon taitojen itsearvioinnin tarkkuudessa ei havaittu sukupuolten välisiä eroja (Beyer, 2002).

Yhtenä selittäväenä tekijänä mahdollisiin sukupuolieroihin saattavat olla sukupuolistereotypiat: Monet naiset pohtivat itseään ja tuntemuksiaan miehiä enemmän, ja tämän vuoksi saattavat tuntea itsensä paremmin (Keltikangas-Järvinen, 1998, 72). Lisäksi naisten usko omaan kykyihinkin näyttää olevan vahvempi humanistisissa aineissa, kun taas miesten minäkäsitys on keskimäärin vahva matemaattisissa aineissa (Dupeyrat ym., 2011; Pallier, 2003). Naisten pessimistiset itsearvioinnit matematiikassa saattavat johtua siitä, että heillä on keskimäärin poikia heikompi käsitys osaamisestaan mate-

matiikassa (Dupeyrat ym., 2011). Itsearviointin tarkkuuden sukupuolieroja voi selittää myös poikien tyttöjä hitaampi psykologinen kehitys (Verhofstadt-Denève, Schittekatte, & Van Leeuwen, 2003).

Itsearviointin tarkkuuteen vaikuttavien seikkojen tarkastelussa iän ja sukupuolen lisäksi on oleellista huomioida myös se, milloin ja missä tilanteissa lapset ovat taipuvaisia yliarvioimaan taitonsa, ja milloin he päinvastoin kykenevät tarkkaan kykyjensä arviointiin. Lasten ajatellaan pystyvän arvioimaan jokapäiväiseen elämään liittyviä konkreettisia asioita realistisesti (Butler, 1990). Lisäksi tehtävän tuttuus vaikuttaa arvioinnin tarkkuuteen: Jos arvioitavasta tehtävästä on vain vähän kokemusta, itsearviointin tarkkuus kärsii (Lew ym., 2010).

Lukemisen ja matematiikan osalta suomalaislapsille kertyy kokemusta koulun alkamisesta lähtien. Blatchfordin tutkimuksessa (1997) oppilaiden matematiikan ja lukemisen itsearviointit poikkesivat toisistaan: Lukemisessa itsearviointit näyttivät muuttuvan tarkoiksi 7 ja 16 ikävuoden välillä, kun taas matematiikassa tämä tapahtui jo 7–11-vuotiaana. Tämän tutkimuksen mukaan nuoremmat oppilaat arvioivat olevansa parempia matematiikassa kuin lukemisessa, kun taas vanhemmat oppilaat kokivat päinvastoin, vaikka todellisessa suoriutumisessa ei ollut oppiaineiden välillä eroja.

Iän, itsearviointikokemuksen, sukupuolen sekä arvioitavan asian tuttuuden lisäksi erilaiset oppilaan yksilölliset ominaisuudet vaikuttavat siihen, kuinka tarkasti oppilas kykenee arvioimaan taitojaan. Vaikka itsearviointin tarkkuuteen vaikuttavia henkilökohtaisia piirteitä on tutkittu suhteellisen vähän, eroavaisuuksia on löydetty muun muassa metakognitiivisista taidoista, koulusuoriutumisesta, älyllisestä kyvykkyydestä, itseluottamuksesta, minäpystyvyydestä sekä itsearviointiin sitoutumisesta ja siinä ponnistelemisestä.

Niin metakognitiivisten taitojen, itsesäätelytaitojen kuin myös kyvyn arvioida omia taitoja on todettu olevan tehokkaan oppijan piirteitä (Lew ym., 2010; Olin & Sullivan, 2004). Kun metakognitiiviset taidot kehittyvät, vahvistuvat samalla myös itsesäätelytaidot, joihin oleellisesti sisältyvät itsearviointitaidot (Alaoutinen, 2012). Itsearviointin tarkkuutta lisäävät meta-



kognitiiviset taidot kehittyvät iän ja harjoituksen myötä, ja niissä on keskeistä tietoisuus ja kontrolli omasta oppimisprosessista, kriittinen ajattelu, oman toiminnan suhteuttaminen asetettuihin tavoitteisiin, havaitun tiedon tulkitseminen sekä johtopäätösten teko (Alaoutinen, 2012; Lew ym., 2010; Lyytinen, 1995, 38; Olina & Sullivan, 2004; Sung ym., 2010). Taidon harjoittelu näyttää lisäävän tietoisuutta omasta osaamisesta, mikä edelleen edesauttaa omien kykyjen tarkkaa itsearviointia (Ng & Earl, 2008).

Koulumenestyksen yhteydestä itsearvioinnin tarkkuuteen on ristiriitaista tutkimustietoa. Yleisesti ottaen on ajateltu, että oppilailta, joilla on oppimisvaikeuksia (Stone & May, 2002), tai jotka muusta syystä menestyvät koulussa heikommin, on taipumusta yliarvioida osaamistaan (Kaderavek ym., 2004; Sung, 2010). Toisaalta koulussa hyvin menestyvien ja opettajien arvioiden (Sung, 2010), sekä taitavien yliopisto-opiskelijoiden itsearviointien ja vertaisten arviointien todettiin olevan keskenään samansuuntaisia (Alaoutinen, 2012). Toisaalla taas kyvykkäiden oppilaiden todettiin aliarvioivan taitonsa, kun oli kyse haastavassa tehtävässä suoriutumisen itsearvioinnista, kun taas helppojen tehtävien itsearvioinneissa he olivat koulussa heikommin menestyviä tarkempia taitojensa itsearvioijia (Burson, Larrick, & Klayman, 2006).

Mahdollinen yhteys koulumenestyksen ja itsearvioinnin tarkkuuden välillä voi selittyä sillä, että menestyjillä on heikommin koulussa pärjääviä realistisempi käsitys omista taidoistaan (Sung ym., 2010). Päinvastoin koulussa heikosti menestyvien itsearviointien epätarkkuus saattaa johtua näiden puutteellisista taidoista arvioida työnsä laatua, asettaa tavoitteita ja arvioida siihen käytetyn ponnistelun määrää, sekä tarkkailla omaa edistymistään (Sung ym., 2010).

Itsearvioinnin tarkkuuteen heikentävästi vaikuttava ominaisuus on myös liika itseluottamus, jolla oli yhteyttä suoriutumisen yliarviointiin (Ng & Earl, 2008). Toisaalta myös itseluottamuksen puute voi näkyä itsearvioinneissa suorituksen yliarviointina. Oppija saattaa antaa itsestään todellisuutta myönteisemmän arvion joko pönkittääkseen itseluottamustaan tai luodakseen opettajalle itsestään kyvykkään vaikutelman (Olina & Sullivan, 2004;

Sung ym., 2010). Monet uskovat omia kykyjä liioittelevien itsearviointien vaikuttavan opettajien käsitykseen oppilaan taidoista, ja näin saavan itsearvioinneista etua myös opettajan tekemässä arvioinnissa (Lew ym., 2010).

Itseluottamuksen puute itsearvioinneissa näkyy myös näihin panostamattomuutena. Panostuksen puute voi johtua siitä, että opiskelijat kokevat itsensä arvioimisen hankalana ja epämiellyttävänä sekä puolueettomana pysymisen haastavana (Olina & Sullivan, 2004). Panostamattomuuden lisäksi oppilaissa on yksilöllisiä eroja siinä, kuinka sitoutuneita he ovat itsearviointiin (Lew ym., 2010; Mabe & West, 1982). Itsearvioinnin kehittäviä vaikutuksia oppimiseen epäilevät ovat epätarkempia arvioissaan kuin ne, jotka luottavat itsearvioinnin parantavan oppimista (Lew ym., 2010).

Minäpystyvyys tarkoittaa yksilön uskomuksia omista taidoistaan ja kyvystään oppia (Pajares & Schunk, 2001). Minäpystyvyysuskomukset vaikuttavat siihen, miten ihminen tuntee, ajattelee, käyttäytyy ja motivoi itsensä erilaisiin tehtäviin (Bandura, 1994). Minäpystyvyysuskomuksilla on myös suuri merkitys siihen, millaisia tehtäviä yksilö valitsee tehtäväkseen, kuinka paljon tämä ponnistelee näiden tehtävien eteen ja millaiseen lopputulokseen hän lopulta pystyy (Bandura, 1994; Schunk & Meece, 2006). Ponnistellemekin ahkerasti niissä tehtävissä, joissa uskomme suoriutuvamme hyvin, kun taas emme motivoitu tekemään niitä tehtäviä, joissa ajattelemme epäonnistuvamme (Schunk & Meece, 2006).

Minäpystyvyydessä on neljä lähdettä. Onnistumisen kokemukset vaikuttavat olevan merkittävin tekijä vahvan minäpystyvyyden rakentumisessa (Bandura, 1994; Schunk & Meece, 2006; Pajares & Schunk, 2001). Toiseksi kokemukset vertaisten onnistumisista rakentavat uskoa siihen, että itsekin kykenee suoriutumaan tehtävästä (Bandura, 1994). Kolmanneksi muilta saatu kannustus ja palaute vahvistavat minäpystyvyyttä (Pajares & Schunk, 2001). Lisäksi oppimiseen liittyvät positiiviset ja negatiiviset tunnetilat ja fyysiset reaktiot vaikuttavat minäpystyvyyden rakentumiseen (Schunk & Meece, 2006).

Omien kykyjen myönteinen arvioiminen kasvattaa uskoa itseän ja motivaatiota oppimiseen, koska oppilaat ajattelevat oppivansa ja olevansa

kyvykkäitä kehittymään myös jatkossa (Schunk, 2003). Adeyemi (2012) tutki matematiikan itsearviointien kasvattavan minäpystyvyyttä matematiikassa. Omien taitojen huonoksi itsearvioiminenkaan ei välttämättä heikennä minäpystyvyyttä, jos oppilas näkee epäonnistumisensa johtuvan panostuksen puutteesta ja kokee pystyvänsä menestymään (Olina & Sullivan, 2004; Schunk, 2003).

Siis myös minäpystyvyydellä voi olla vaikutusta itsearvioinnin tarkkuuteen. Tutkimuksen mukaan korkea minäpystyvyys saattaa aiheuttaa epätarkkuutta itsearviointiin, koska yksilö ajattelee olevansa parempi mitä todellisuudessa onkaan (Vancouver & Kendall, 2006). Toisaalta toisessa tutkimuksessa minäpystyvyydellä ja tarkalla itsearvioinnilla ei todettu olevan yhteyttä (Ng & Earl, 2008).

### **Itsearvioinnin reliabiliteetti ja validiteetti**

Koska tutkimuksen luotettavuus on suoraan verrannollinen mittarin luotettavuuteen (Metsämuuronen, 2009, 74), tässä tutkimuksessa tarkastellaan myös itsearvioinnin reliabiliteettia ja validiteettia. Itsearvioinnin mittareiden reliabiliteettia on tutkittu mittarin kehittämiseen tähtäävissä tutkimuksissa. Tutkimus on kuitenkin toteutettu yläkoulu-, lukio- ja korkeakouluikäisillä opiskelijoilla, eikä lasten itsearvioinnin mittarien reliabiliteetista ei ole nykyaikaista tutkimustietoa matematiikan, lukemisen, tehtävissä ponnisteleminen, keskittymisen tai tunteiden osa-alueilla. Tämän vuoksi seuraavaksi tarkastelen reliabiliteettia yleisemmin.

Perinteisesti reliabiliteetti määritellään mittauksen tai mittarin pysyvyytenä tai toistettavuutena, mutta reliabiliteetti sisältää myös mittauksen johdonmukaisuuden ja yhtenevyyden (Bollen, 1989, 220). Täydellisen reliabiliteetti mittaus on tarkka, eikä se sisällä virhettä (Guilford, 1965, 438). Toisaalta mittaustulos ei ole koskaan virheetön, vaan siihen jää aina mittarista tai mittaolosuhteista johtuvaa satunnaisvirhettä, jota voidaan tarkastella sen avulla, kuinka samanlaisina tulokset pysyvät mittauskerrasta toiseen (Nummenmaa, 2009, 346–349). Mitä suurempi on vaihtelu vastausten välillä,

sitä matalampi on reliabiliteetti, ja mitä yhtenevämmät ovat vastaukset eri mittauksissa, sitä korkeampi on reliabiliteetti (Bollen, 1989, 207).

Reliabiliteetin arvioimiseen ei ole yhtä parasta keinoa, vaan sen valintaan vaikuttavat muun muassa tutkijan tavoitteet, aineiston luonne sekä itse mittari (Guilford, 1965, 446). Reliabiliteettia voidaan arvioida testi-uusintatesti -menetelmän avulla, jossa säästetään aikaa, kun menettelyyn tarvitaan vain yksi mittari ja kaksi mittauskertaa (Nummenmaa, 2009, 355). Menetelmä on käytännöllinen myös yksinkertaisuutensa ja helppoutensa vuoksi, mutta toisaalta sillä on myös rajoituksensa (Bollen, 1989, 212).

Testi-uusintatesti -menetelmän teoreettiset ja käytännölliset rajoitukset liittyvät ensinnäkin siihen, että tutkittava saattaa muistaa ensimmäisen tutkimuskertansa vastaukset myös toisella tutkimuskerralla, ja toisella mittauskerralla testin toimintaperiaate on jo ensimmäisestä kerrasta tuttu (Bollen, 1989, 211; Nummenmaa, 2009, 349). Toisena menetelmän rajoituksena on se, että mittauksen kohteena olevan ominaisuuskin voi muuttua, jos mittauskertojen välinen aika on liian pitkä (Nummenmaa, 2009, 355).

Edellä mainittujen rajoitusten lisäksi myös tutkittavasta itsestään tai testausolosuhteista johtuvat tekijät, kuten väsymys, valaistus tai ohjeistuksen kuuluvuus, saattavat vaihdella mittauskerrasta toiseen, mikä saattaa vaikuttaa mittauksen reliabiliteettiin (Guilford, 1965, 447). Panadero ja Romero (2014) vertasivat tutkimuksessaan ohjeistettua itsearviointia ei-ohjeistettuun, ja havaitsivat ohjeistuksella ja itsearvioinnin tarkkuudella olevan yhteyttä. Mittariasioista ei tulisi kuitenkaan tehdä estettä itsearvioinnin toteuttamiselle, koska huolellisesti suunniteltu ja toteutettu tutkimus sekä joustava kysymyksenasettelu on usein riittävä mittauksen lähtökohta (Lyytinen, 1995, 42).

Reliabiliteettitarkastelun lisäksi tarvitaan myös tutkimuksen validiuden tarkastelua, jonka tarkoituksena on arvioida sitä, kuinka hyvin tarkastelun kohteena olevan ominaisuuden mittaaminen onnistuu (Bollen, 1989, 184; Metsämuuronen, 2009, 125). Validiteettitarkastelussa onkin tärkeä arvioida muun muassa tutkimuksen yleistettävyyttä, käsitteiden vastaavuutta teorian kanssa sekä mitattavan asian operationalisointia (Metsämuuronen,

2009, 125–128). Validiteetin uhat on oleellista huomioida jo etukäteen, koska jälkikäteen tutkimusta koskevien valintojen muuttaminen on haastavaa (Metsämuuronen, 2009, 75).

Suurin osa itsearvioinnin validiteettia tarkastelevista tutkimuksista ovat havainneet matalasta kohtalaiseen validiteetteja (Mabe & West, 1982; Sung, 2010). Syyt joidenkin tutkimusten matalaan validiteettiin ovat pitkälti samoja kuin syyt epätarkkaan itsearviointiin: Ensinnäkin paine olla vertaisten kanssa samantasoinen vaikuttaa itsearvioinnin validiteettiin ja reliabiliteettiin (Sung, 2010). Toiseksi tutkimuksen validiteettia heikentää se, että ihminen tiedostamattaan havaitsee itsessään usein vain niitä asioita, joita on hyväksyttävää havaita (Keltikangas-Järvinen, 1998, 71). Kolmanneksi kaikki ihmiset eivät osaa arvioida itseään objektiivisesti, vaan nämä saattavat arvioida itsensä sijasta sitä, mitä muut hänestä ajattelevat tai toiveita siitä, millainen hän haluaisi olla (Blatchford, 1997; Olina & Sullivan, 2004). Lisäksi ihmisellä on luontainen taipumus yliarvioida taitojaan (Ng & Earl, 2008) ja ajatella olevansa hieman parempi kuin toinen ihminen (Keltikangas-Järvinen, 1998, 72).

## **Tutkimusongelmat**

Tämän tutkimuksen ensimmäisessä vaiheessa tarkoituksena on tarkastella itsearvioinnin pysyvyyttä testi-uusintatesti -reliabiliteetin avulla viidessä eri ulottuvuudessa: matematiikassa, lukemisessa, tehtävissä ponnistelemissa, keskittymisessä ja oppimiseen liittyvissä tunteissa. Tutkimuksen toisessa vaiheessa tutkitaan sukupuolen ja luokkatason vaikutusta matematiikan ja lukemisen taitojen itsearvioinnin ja suoriutumisen väliseen yhteyteen. Näitä selvitetään seuraavien kysymysten avulla:

- 1) Onko eri ulottuvuuksien itsearvioinnin pysyvyydessä eroa?
- 2) Onko sukupuolella tai luokkatasolla yhteyttä eri ulottuvuuksien itsearvioinnin pysyvyyteen?

-----

- 3) Onko sukupuolella tai luokkatasolla yhteyttä matematiikan ja lukemisen taitojen itsearvioinnin tarkkuuteen?

## **METODI**

### **Tutkimushanke**

Tämä tutkimus on osa Minäpystyvyys ja oppimisinterventiot -hanketta, joka on Jyväskylän yliopiston, Niilo Mäki Instituutin ja Jyväskylän kaupungin yhteistyössä toteuttama ja Suomen Akatemian rahoittama hanke. Vuosina 2013–2015 toteutettavan tutkimusprojektin tarkoituksena on selvittää alakouluikäisten oppilaiden minäpystyvyyden kokemuksen vaikutusta oppimiseen lukemisessa ja matematiikassa. Hanke sisältää alku- ja loppumittauksia, joiden välissä osalle tutkimukseen osallistuneista toteutetaan lukemisen sujuvuutta ja yhteenlaskustrategioita kehittäviä interventioita. Hankkeessa puureudutaan myös tarkkaavuuden ja toiminnanohjauksen pulmien kuntouttamiseen. Luokan- ja erityisopettajat toteuttavat nämä interventiot oppilaiden omissa kouluissaan.

### **Osallistujat**

Oman tutkimukseni ensimmäinen vaihe kohdistui hankkeen kahteen kyselylomakkeella suoritettuun alkumittaukseen, joista ensimmäinen toteutettiin marraskuussa 2013 ja toinen tammikuussa 2014. Tutkimuksen toisen vaiheen mittareina olivat marraskuun 2013 mittaus matematiikan ja lukemisen osalta, sekä oppilaiden matematiikan ja lukemisen suoriutumisen testit. Tutkimukseen osallistui Keski-Suomesta ja Itä-Suomesta vapaaehtoisiksi ilmoittautuneiden opettajien luokilta yhteensä 1344 2.-5.-luokkalaista lasta. Tyttöjä oli 649 ja poikia 695. Otokoko kuitenkin vaihteli muuttujakohtaisesti satunnaisista puuttuvista tiedoista johtuen. Tyttöjen ja poikien jakautuminen eri luokka-asteille on esitetty taulukossa 1.

TAULUKKO 1. Osallistujat sukupuolittain ja luokka-asteittain.

Luokka-aste	Tytöt	Pojat	Yhteensä
2. lk N	103	104	207
%	7,7	7,7	15,4
3. lk N	239	238	477
%	17,8	17,7	35,5
4. lk N	174	198	372
%	12,9	14,7	27,7
5. lk N	133	155	288
%	9,9	11,5	21,4
Yht. N	649	695	1344
%	48,3	51,7	100,0

Prosentit ovat osuuksia koko osallistujajoukosta.

## Mittarit

Minäpystyvyys ja oppimisvaikeus-interventiot -hankkeen kyselylomake on laadittu Niilo Mäki Instituutin ja Jyväskylän erityispedagogiikan laitoksen tutkijoiden yhteistyönä. Lomake perustuu Banduran (2006) minäpystyvyysteoriaan, ja siinä on hyödynnetty myös aiemmin julkaistuja minäpystyvyys -kyselyitä. Mittarit on esitettävä ja muokattu uudelleen näiden perusteella.

Yhdeksässä ryhmätestinä toteutetussa alku- ja loppumittauksen kyselylomakkeessa lapsen tehtävänä on esimerkiksi valita itseään kuvaavin vaihtoehto siitä, kuinka varma hän on, että hän pystyy tai ei pysty tekemään erilaisia kouluun ja oppimiseen liittyviä asioita. Toiseksi lapsi arvioi sitä, kuinka pitkälle hän itse, vertainen, yläkoululainen tai aikuinen ehtii lukea tai laskea sivun mittaista tekstiä tai yhteenlaskuja puolen minuutin aikana. Kolmanneksi osallistujan tehtävänä on rastittaa itseään kuvaavin vaihtoehto väittämistä, joissa käsitellään omaa osaamista ja omia vahvuuksia. Neljänneksi lapsi erottelee sanoja yhteen kirjoitetuista kirjainjonoista eli sanakettjuista. Lisäksi lapsen lukemisen sujuvuutta testataan niin, että lapsi merkitsee mahdollisimman monta esitettyä väittämää joko oikeaksi tai vääräksi.

Tutkimukseni ensimmäisen vaiheen mittarit liittyivät kyselylomakkeen minäpystyvyys -tehtäviin ja minäpystyvyyden lähteet -osioihin. Minäpysty-

vytteen liittyvässä mittarissa osallistujan tehtävänä oli arvioida omaa pystyvyyttään kouluun ja oppimiseen liittyvissä kysymyksissä 7-portaisella asteikolla (1 = Täysin varma, että en pysty - 7 = Täysin varma, että pystyn). Mittarin 48 kysymystä käsittelivät matematiikkaan, lukemiseen, tehtävissä ponnisteleminen, keskittymiseen ja oppimistilanteiden tunteisiin liittyviä minäpystyvyyssuskomuksia.

Toinen tutkimuksen ensimmäisen vaiheen mittareista koostui 33 väittämistä, jotka liittyivät minäpystyvyyden lähteisiin, kuten onnistumisen kokemuksiin, kokemuksiin toisten onnistumisista, muilta saatuun kannustukseen ja oppimiseen liittyviin tunnetiloihin. Tässä osiossa osallistujan tuli arvioida väittämän paikkansapitävyyttä omalla kohdallaan edelleen 7-portaisella asteikolla (1 = Ei totta - 7 = Totta).

Suurin osa näiden mittareiden kysymyksistä ja väittämistä liittyi matematiikkaan, lukemiseen, ponnisteleminen, keskittymiseen sekä oppimiseen ja oppimistilanteisiin liittyviin tunteisiin. Matematiikan 21 kysymyksessä ja lukemisen 24 kysymyksessä oppilas ensinnäkin itsearvioi minäpystyvyyttään matematiikan ja lukemisen taidoissa. Toiseksi lapsi arvioi matematiikkaan ja lukemiseen liittyviä onnistumisen kokemuksiaan, attribuutioitaan sekä kokemuksia omasta pystyvyydestään toisten onnistumisten kautta. Lisäksi lapselle näytettiin 8 erilaista vaikeutuvaa laskua, ja tämän tehtävänä oli arvioida, kuinka varma hän on, että hän pystyy laskemaan sellaisia laskuja nopeasti. Lukemisen osalla lapsen tuli arvioida 10 eripituisen tekstin avulla sitä, kuinka varma hän on, että hän pystyy lukemaan näkemänsä tekstin puolessa minuutissa.

Tehtävissä ponnisteleminen kysymykset ja osa keskittymiseen liittyvistä kysymyksistä kuuluivat minäpystyvyyden mittarin itsesäätelävän oppimisen kysymyksiin, joissa lapsi itsearvioi kykyään ponnistella tehtävissä ja keskittyä koulunkäyntiin. Lisäksi lapsi itsearvioi kyselylomakkeessa oppimiseensa liittyviä tunnetiloja. Mittarin viiden osa-alueen kysymykset ja väittämät ulottuvuuksittain on kuvattu yksityiskohtaisemmin liitteessä 1.

Lasten testaajina toimivat hankkeen kouluttamat testaajat, jotka saivat yhdenmukaisen ohjeistuksen sekä kirjalliset testausohjeet. Tehtävät tehtiin



koulussa testajaan johdolla yhteen tahtiin niin, että testaja luki ohjeistuksen ja kysymykset, ja oppilaat vastasivat näihin. Kun kaikki olivat valmiita, siirryttiin seuraavaan kysymykseen tai tehtävään. Kyselylomakkeen tehtäviin vastaaminen kesti 45–60 minuuttia.

Tutkimuksen toisessa vaiheessa lapsen itsearviointia verrattiin tämän suoriutumiseen matematiikassa ja lukemisessa. Seuraavaksi kuvaan niin matematiikan osaamisen mittareina olleita yhteen- ja vähennyslaskutehtäviä sekä aritmetiikan testejä kuin myös lukemisen sujuvuuden mittareita lukusujuvuutta, sanaketju-tehtävää ja Kilpikonnan tekstiä.

Yhteen- ja vähennyslaskutaitoa arvioitiin kumpaakin 60 laskua sisältävän kaksisivuisen tehtävän avulla. Yhteenlaskutehtävässä lapsen tuli laskea mahdollisimman nopeasti ja tarkasti kymmentä pienemmistä luvuista koottuja laskuja (esim.  $9 + 7$ ). Vähennyslaskujen osalta vähennettävä oli pienempi kuin 20 (esim.  $16 - 9$ ). Kummankin tehtävän tekemiseen oli aikaa kaksi minuuttia.

Aritmeettisiä taitoja mittaavassa testissä lapsen tuli laskea edelleen nopeasti ja tarkasti mahdollisimman monta laskua 30 tehtävästä. Laskutehtävät sisälsivät niin yhteen-, vähennys-, kerto- ja jakolaskuja kuin myös näiden yhdistelmiä. Toisissa laskuissa lapsen tehtävänä oli laskea tehtävään vastaus, toisissa taas päätellä laskusta puuttuva luku (esim.  $\_\_\_ - 3 = 10$ ). Tehtävien ratkaisemiseen oli aikaa kolme minuuttia. Kaikissa matematiikan tehtävissä oikein ratkaistusta laskusta sai yhden pisteen.

Lukusujuvuustehtävässä oppilas luki äänettömästi mahdollisimman monta lausetta kolmen minuutin aikana. Lapsen tehtävänä oli merkitä mahdollisimman monta väittämää 70:stä joko paikkaansa pitäväksi tai vääräksi (esim. "Kaloilla on höyhenet." tai "Mehua voi juoda.>"). Varsinaista testiä edelsi kolme harjoitusosiota. Lauseet pyrittiin muodostamaan niin, että niiden totuudellisuuden määrittely olisi helppoa. Lukemissujuvuuden mittana käytettiin kolmessa minuutissa oikein arvioitujen lauseiden lukumäärää.

Sanaketju-tehtävässä arvioitiin lukusujuvuutta niin, että lapsen tehtävänä oli erotella pystyviivoin sanoja 78:sta yhteen kirjoitetusta kirjainjonosta eli sanaketjusta. Ennen varsinaista testiä tehtävän tekemistä harjoiteltiin

kuuden kirjainjonon avulla. Yhdessä ketjussa sanoja oli kahdesta neljään. Jokaisesta oikein erotellusta sanasta sai yhden pisteen. Tehtävän tekemiseen oli aikaa kolme ja puoli minuuttia.

Myös kilpikonna-teksti toimi lukemisen sujuvuuden mittarina. Siinä lapsi luki 90 sekunnin ajan Kilpikonna-aiheista tekstiä mahdollisimman nopeasti ja tarkasti. Lukemisen mittana oli luettujen sanojen määrä sekä lukemisen virheettömyys.

## **Aineiston analysointi**

Aineiston analysoinnissa käytettiin tilastollisia analyysimenetelmiä. Analyysit suoritettiin SPSS 22 -ohjelmiston avulla. Suurin osa minäpystyvyyden ja sen lähteiden mittareiden kysymyksistä ja väittämistä liittyi viiteen osa-alueeseen: matematiikkaan, lukemiseen, ponnisteleamiseen, keskittymiseen ja tunteisiin. Tutkimukseen valittiinkin näihin osa-alueisiin liittyvät 77 kysymystä ja väittämää. Muuttujat käännettiin samansuuntaisiksi, jonka jälkeen ne tiivistettiin osa-alueita vastaaviksi keskiarvosummamuuttujiksi. Summamuuuttujien reliabiliteetti oli 0.86–0.90.

Matematiikan, lukemisen ja keskittymisen ulottuvuudet muodostuivat sekä minäpystyvyyteen että minäpystyvyyden lähteisiin liittyvistä muuttujista. Mittarin tarkempaa tutkimista ja sen kehittämistä varten minäpystyvyyden kysymyksistä ja minäpystyvyyden lähteiden väittämistä muodostettiin omat summamuuttujansa, ja osioiden pysyvyyttä testattiin testi-uusintatesti -reliabiliteetin avulla. Toistomittauskorrelaatioiden avulla tutkittiin myös summamuuttujien yksittäisten kysymysten pysyvyyttä.

Tutkimuksen ensimmäisessä vaiheessa itsearviointin pysyvyyttä tarkasteltiin eri ulottuvuuksien osalta Pearsonin korrelaatiokertoimien avulla testi-uusintatesti -reliabiliteettina. Tyttöjen ja poikien sekä eri luokkatasojen toistomittausten korrelaatioiden eroja tutkittiin laskemalla todennäköisyys korrelaatioiden erojen tilastolliseen merkitsevyyteen. Tämän laskemiseksi mittauksen 1 ja mittauksen 2 väliset korrelaatiot muutettiin molemmissa ryhmissä vastaamaan Fisherin jakaumaa, koska Pearsonin korrelaatiokertoimen jakauma ei ollut normaali. Soperin (2014) laskuria apuna käyttäen

saatiin kaksisuuntainen todennäköisyys z-arvon tilastolliselle merkitsevyydelle, jota verrattiin valittuun merkitsevyytasoon (0.05).

Tutkimuksen toisessa vaiheessa tutkittiin sukupuolen ja luokkatason yhteyttä matematiikan ja lukemisen itsearviointin tarkkuuteen. Matematiikassa suoriutumisen mittarit (yhteen- ja vähennyslaskutestit sekä aritmetiikan taitoja mittaava testi) yhdistettiin keskiarvosummamuuttujaksi ( $\alpha = 0.92$ ). Myös lukemisen suoriutumista mittaavista lukemisen sujuvuuden testeistä rakennettiin keskiarvosummamuuttuja ( $\alpha = 0.91$ ). Keskiarvosummamuuttujien rakentamisen jälkeen itsearviointin tarkkuutta tarkasteltiin korrelaatioiden avulla, ja näiden tilastollinen merkitsevyys tutkittiin Pearsonin riippumattoman korrelaatiotestin avulla (samoin kuin tutkimuksen ensimmäisessä vaiheessa).

## TULOKSET

### Itsearviointin pysyvyys

Itsearviointin pysyvyyden tutkimiseksi muuttujat tiivistettiin viideksi eri keskiarvosummamuuttujaksi. Nämä viisi ulottuvuutta koskivat matematiikkaa, lukemista, tehtävissä ponnistelemista, keskittymistä ja oppimiseen liittyviä tunteita.

Matematiikan ja lukemisen osalla itsearviot olivat ulottuvuuksista pysyvimpiä ( $r = 0.76$  ja  $0.75$ , tässä järjestyksessä). Keskittymisen itsearvioinneissa oli näitä hieman vähemmän pysyvyyttä ( $r = 0.70$ ). Ulottuvuuksista vähiten mittauskertojen välistä pysyvyyttä oli tehtävissä ponnistelemisen ja oppimistilanteissa esiintyvien tunteiden itsearvioinneissa ( $0.65$  ja  $0.65$ ).

TAULUKKO 2. Itsearviointin pysyvyys ulottuvuuksittain ja osioittain sekä osioiden korrelaatioiden erojen tilastolliset merkitsevyydet.

<b>Ulottuvuus</b>	<b>Testi-uusintatesti -reliabiliteetti</b>	<b>Korrelaatioiden erojen merkitsevyys p</b>
<b>Matematiikka</b>	<b>0.76</b>	
Minäpystyvyys	0.71	
Minäpystyvyyden lähteet	0.74	.180
<b>Lukeminen</b>	<b>0.75</b>	
Minäpystyvyys	0.71	
Minäpystyvyyden lähteet	0.68	.174
<b>Ponnisteleminen</b>	<b>0.65</b>	
Minäpystyvyys		
<b>Keskittyminen</b>	<b>0.70</b>	
Minäpystyvyys	0.58	
Minäpystyvyyden lähteet	0.66	.001
<b>Oppimiseen liittyvät tunteet</b>	<b>0.65</b>	
Minäpystyvyyden lähteet		

Matematiikan, lukemisen ja keskittymisen ulottuvuudet sisälsivät sekä minäpystyvyyteen että sen lähteisiin liittyviä muuttujia. Matematiikan ja lukemisen minäpystyvyyden ja sen lähteiden pysyvyydessä ei ollut tilastollisesti merkitsevää eroa. Toisaalta keskittymisen ulottuvuuden minäpystyvyyden ja sen lähteiden pysyvyydessä oli tilastollisesti merkitsevä ero. Summa-  
muuttujien ja niiden osioiden testi-uusintatesti -reliabiliteetit sekä korrelaatioiden erojen tilastolliset merkitsevyydet on esitetty taulukossa 2.

Sukupuolen yhteyttä itsearviointin pysyvyyteen tarkasteltaessa huomattiin, että matematiikan ja lukemisen sekä oppimiseen liittyvien tunteiden osalla tyttöjen itsearviot olivat pysyvämpiä kuin pojilla (taulukko 3). Toisaalta tilastollisesti merkitseviä eroja ei ilmennyt sukupuolten välillä keskittymisen eikä tehtävissä ponnisteleminen itsearvioinneissa (taulukko 3).

TAULUKKO 3. Tyttöjen ja poikien itsearvioinnin pysyvyys ulottuvuuksittain sekä korrelaatioiden erojen tilastolliset merkitsevyydet.

Ulottuvuus	Testi-uusintatesti -reliabiliteetti	Korrelaatioiden erojen merkitsevyys p
<b>Matematiikka</b>		
tytöt	0.80	
pojat	0.70	.000
<b>Lukeminen</b>		
tytöt	0.79	
pojat	0.71	.005
<b>Ponnisteleminen</b>		
tytöt	0.66	
pojat	0.65	.747
<b>Keskittyminen</b>		
tytöt	0.72	
pojat	0.68	.245
<b>Oppimiseen liittyvät tunteet</b>		
tytöt	0.73	
pojat	0.56	.000

Eri-ikäisten oppilaiden itsearvioinnin pysyvyyttä tutkittaessa havaittiin, että kaikissa ulottuvuuksissa vanhemmilla oppilailla oli kyky itsearvioida nuorempia oppilaita yhtenevämmiin kahdella eri mittauskerralla (taulukko 4). Ulottuvuuksissa eroja oli siinä, minkä luokkatasojen välillä erot testi-uusintatesti -reliabiliteetissa olivat tilastollisesti merkitseviä.

Matematiikan osalla neljännen ja viidennen luokan oppilaiden itsearviot olivat kolmannella luokalla olevia pysyvämpiä. Lukemisen itsearviot olivat neljäsluokkalaisilla pysyvämmät kuin toisella ja kolmannella luokalla opiskelevilla. Tehtävissä ponnistelemisen itsearvioinneissa viidesluokkalaiset kykenivät arvioimaan taitojaan kakkos- ja neljäsluokkalaisia pysyvämmiin, kun taas keskittymisen itsearvioinnissa viidesluokkalaiset arvioivat taitojaan kolmosluokkalaisia pysyvämmiin. Oppimiseen liittyviä tunteita viidesluokkalaiset kykenivät arvioimaan niin neljännellä, kolmannella, kuin myös toisella luokalla opiskelevia pysyvämmiin. Mittausten väliset korrelaatiot ja niiden tilastolliset merkitsevyydet on esitetty taulukossa 4.

TAULUKKO 4. Eri-ikäisten itsearvioinnin pysyvyys ulottuvuuksittain sekä korrelaatioiden erojen tilastolliset merkitsevyydet.

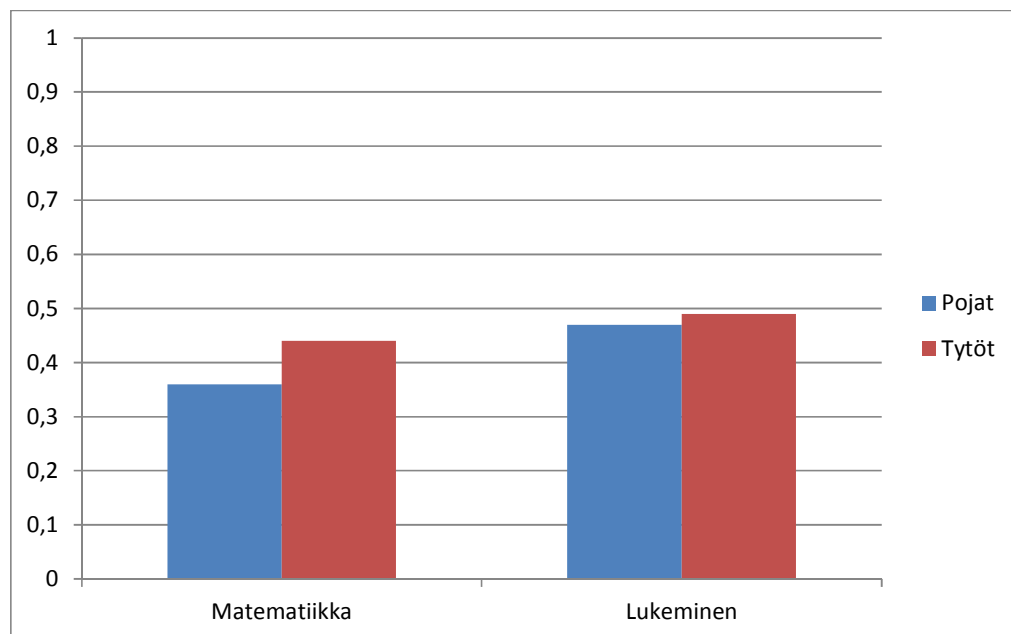
Ulottuvuus	Testi-uusintatesti -reliabiliteetti	Korrelaatioiden erojen merkitsevyys p
Matematiikka		.220 <sup>23</sup>
2. lk	0.77	.509 <sup>24</sup>
3. lk	0.71	.724 <sup>25</sup>
4. lk	0.80	<b>.008</b> <sup>34</sup>
5. lk	0.79	<b>.041</b> <sup>35</sup>
		.715 <sup>45</sup>
Lukeminen		.589 <sup>23</sup>
2. lk	0.70	<b>.025</b> <sup>24</sup>
3. lk	0.73	.406 <sup>25</sup>
4. lk	0.81	<b>.012</b> <sup>34</sup>
5. lk	0.74	.649 <sup>35</sup>
		.077 <sup>45</sup>
Ponnisteleminen		.081 <sup>23</sup>
2. lk	0.56	.354 <sup>24</sup>
3. lk	0.67	<b>.008</b> <sup>25</sup>
4. lk	0.63	.284 <sup>34</sup>
5. lk	0.73	.173 <sup>35</sup>
		<b>.025</b> <sup>45</sup>
Keskittyminen		.632 <sup>23</sup>
2. lk	0.69	.647 <sup>24</sup>
3. lk	0.66	.142 <sup>25</sup>
4. lk	0.72	.193 <sup>34</sup>
5. lk	0.77	<b>.009</b> <sup>35</sup>
		.185 <sup>45</sup>
Oppimiseen liittyvät tunteet		<b>.005</b> <sup>23</sup>
2. lk	0.50	<b>.004</b> <sup>24</sup>
3. lk	0.68	<b>.039</b> <sup>25</sup>
4. lk	0.69	.826 <sup>34</sup>
5. lk	0.65	.449 <sup>35</sup>
		.361 <sup>45</sup>

2 = 2. lk, 3 = 3. lk, 4 = 4. lk, 5 = 5. lk

Mittausten välisiä korrelaatioita tarkasteltaessa havaittiin, että testi-uusintatesti -reliabiliteetti oli korkeampi matematiikassa ja lukemisessa kuin muissa ulottuvuuksissa. Lisäksi aineistossa oli mittari ainoastaan matematiikassa ja lukemisessa suoriutumisen. Näiden seikkojen vuoksi tutkimuksen toisessa vaiheessa tarkasteltiin itsearviointin ja suoriutumisen välistä yhteyttä ainoastaan matematiikan ja lukemisen osalta. Vaikka testi-uusintatesti -reliabiliteetissa oli eroja myös sukupuoli- ja luokka-astevertailussa, jatkotutkimuksessa päätettiin tutkia koko osallistujajoukkoa, koska matematiikan ja lukemisen reliabiliteettikertoimet olivat korkeat ( $r = 0.70-0.81$ )

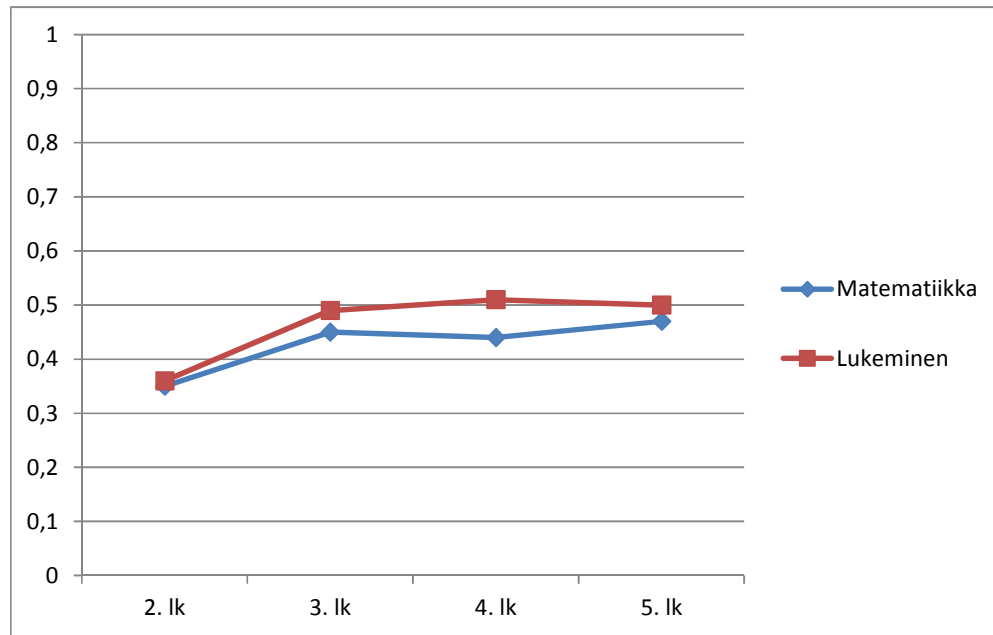
### Itsearviointin tarkkuus

Tyttöjen ja poikien välillä ei ollut tilastollisesti merkitsevää eroa matematiikan ( $p = .104$ ) eikä lukemisen ( $p = .671$ ) itsearviointin tarkkuudessa. Itsearviointin ja suoriutumisen väliset korrelaatiot sukupuolittain on esitetty kuviossa 1.



KUVIO 1. Itsearviointin ja suoriutumisen väliset korrelaatiot sukupuolittain.

Myöskään eri-ikäisten oppilaiden matematiikan ja lukemisen itsearvioinnin tarkkuudessa ei ollut tilastollisesti merkitsevää eroa, vaikkakin 5.-luokkalaisten itsearvioinnin ja suoriutumisen välillä oli esimerkiksi 2.-luokkalaisten itsearviointia ja suoriutumista enemmän yhteyttä. Itsearvioinnin ja suoriutumisen väliset korrelaatiot luokka-asteittain on esitetty kuviossa 2.



KUVIO 2. Itsearvioinnin ja suoriutumisen väliset korrelaatiot luokka-asteittain.

## POHDINTA

Tämän tutkimuksen tarkoitus oli kahtalainen: Ensinnäkin tutkimuksessa tutkittiin itsearvioinnin pysyvyyttä testi-uusintatesti -reliabiliteetin avulla. Toiseksi tutkimuksessa tarkasteltiin sukupuolen ja luokkatason yhteyttä itsearvioinnin tarkkuuteen.

### Tulosten tarkastelua

Matematiikan, lukemisen, tehtävissä ponnistelemisen, keskittymisen ja oppimiseen liittyvien tunteiden itsearvioinnin pysyvyys oli tässä tutkimuksessa korkea. Kuitenkin matematiikan ja lukemisen itsearvioinneissa oli



ulottuvuuksista eniten pysyvyyttä. Itsearviointin tarkkuuden tutkimuksen mukaan tehtävän tuttuus ja konkreettisuus näyttävät vaikuttavan siihen, kuinka tarkasti lapsi osaa arvioida taitojaan (Butler, 1990; Lew ym., 2010). Suomalaislapsilla onkin kokemusta matematiikan ja lukemisen harjoittelemisesta ja itsearvioinneista jo alkuopetusvuosilta lähtien. Itsearviointikokemuksen lisäksi lapsi saa koulutaipaleellaan palautetta matematiikan ja lukemisen taidoistaan opettajalta sekä vertaisilta, mikä kasvattaa tietoisuutta omasta osaamisesta. Tietoisuus omista vahvuuksista ja kehittämiskohteista tukee edelleen kykyä arvioida omaa osaamistaan tarkasti.

Toisaalta tehtävissä ponnistelemisen ja oppimiseen liittyvien tunteiden itsearvioinnit eivät olleet yhtä pysyviä kuin matematiikan ja lukemisen itsearvioinnit. Joidenkin lasten erilaiset vastaukset eri mittauskerroilla saattavat olla merkki siitä, että näillä ei ole joko kokemusta abstraktien asioiden arvioimisesta, tai edes kehityksellistä valmiutta hahmottaa ja arvioida ei-konkreettisia asioita. Mittarin ponnistelemista, keskittymistä ja tunteita käsittelevät kysymykset ja väittämät oli pyritty muotoilemaan ymmärrettäviksi ja lapsen kehitystasoon sopiviksi. Lasten vastausten heilahteleminen mittarin ansioista huolimatta voi johtua myös siitä, että lapsen käsitys matematiikasta ja lukemisesta on harjoittelamisen ja saadun palautteen vuoksi pysyvämpi kuin käsitys omista tunteistaan tai minäpystyvyydestään, jotka saattavat vaihdella eri päivinä ja erilaisissa tehtävissä.

Matematiikan ja lukemisen sekä oppimiseen liittyvien tunteiden osalla tyttöjen itsearvioiden todettiin olevan pysyvämpiä kuin pojilla. Tyttöjen vastausten pysyvyyteen saattaa vaikuttaa ensinnäkin näiden poikia varhaisempi fyysinen ja psyykkinen kehitys. Toiseksi Verhofstadtin ym. (2003) mukaan tyttöjen oletetaan olevan poikia huolellisempia tekemisissään ja hillitympiä käytöksessään, joten tyttöjen kyky keskittyä voi olla myös syynä itsearviointin pysyvyyteen. Lisäksi stereotyyppisesti ajateltuna tytöiltä sallitaan poikia enemmän emotionaalista käyttäytymistä, joten kenties tytöt myös kykenevät tunnistamaan, tiedostamaan ja analysoimaan negatiivisia tunteitaan poikia paremmin, mikä voi vaikuttaa tunteiden itsearviointin pysyvyyteen.

Vanhempien oppilaiden itsearviot olivat nuorempien itsearvioita pysyvämpiä kaikissa ulottuvuuksissa. Tähän voi vaikuttaa ensinnäkin se, että matematiikkaa ja lukemista harjoitellaan koulussa päivittäin, jolloin on luonnollista, että viidesluokkalaisella on kakkosluokkalaista vakaammat taidot näiden asioiden itsearvioinnissa. Toiseksi oppilasta pyydetään kiinnittämään huomiota käyttäytymiseensä, kehoitetaan ponnistelemaan ja jaksamaan tehtävien tekemisessä, sekä hänelle opetetaan kouluvuosien aikana tunteiden tunnistamista ja niiden käsittelemistä. Lisäksi alakoulun aikana kehittyvät psykologiset valmiudet ja metakognitiiviset taidot saattavat olla syynä vanhempien oppilaiden nuorempia pysyvämpiin itsearviointeihin.

Matematiikan osalla neljännen ja viidennen luokan oppilaiden itsearviot olivat kolmannella luokalla olevia pysyvämpiä. Tämä voi olla merkki siitä, että matemaattisten uskomusten ja taitojen itsearvioiminen on vielä kolmannella luokalla olevalle haastavaa. Toisaalta itsearvioinnin pysyvyys neljännen ja viidennen luokan välillä kertoo siitä, että matematiikasta on kertynyt vuosien varrella harjoitusta ja palautetta, jolloin myös käsitys omasta osaamisesta on totuudenmukainen. Itsearvioinnin pysyvyys neljännen ja viidennen luokan välillä voi myös kertoa siitä, että mittari on toiminut tämän ikäisillä oppilailla luotettavasti.

Lukemisenkin itsearviot olivat vanhemmilla oppilailla nuorempia pysyvämmät, kun neljäsluokkalaiset kykenivät arvioimaan osaamistaan toisella ja kolmannella luokalla opiskelevia pysyvämmiin. Viidennellä luokalla olijat eivät kuitenkaan enää olleet neljäsluokkalaisia pysyvämpiä itsearvioijia, mikä voi kertoa siitä, että lukemisen osalla lapsi näyttää kykenevän arvioimaan taitojaan pysyvästi matematiikkaa aiemmin. Tähän voi olla selityksenä lukemisen tuttuus; monella lapsella on kertynyt kokemusta lukemisesta matematiikkaa enemmän jo ennen koulun alkamista.

Tehtävissä ponnistelemisen itsearvioinneissa viidesluokkalaiset kykenivät arvioimaan taitojaan kakkos- ja nelosluokkalaisia pysyvämmiin, kun taas keskittymisen itsearvioinnissa viidesluokkalaiset arvioivat taitojaan kolmosluokkalaisia pysyvämmiin. Oppimiseen liittyviä tunteita viidesluokkalaiset kykenivät arvioimaan niin neljännellä, kolmannella, kuin myös

toisella luokalla opiskelevia pysyvämmin. Viidennelle luokalle saakka pysyvämmiksi muuttuneet itsearviointit voivat selittyä sillä, että taidot tunnistaa omia kykyjä abstrakteissa asioissa alkuluokilla eivät ole vielä kehittyneet. Alaluokkien matalat toistomittausreliabiliteetit voivat johtua yksilöllisten itsearviointitaitojen lisäksi myös sillä, että mittarin kysymykset ovat teemoiltaan tämän ikäisille vielä liian haastavia.

Oppimiseen liittyvien tunteiden itsearviot eivät olleet kakkosluokkalaisten osalla pysyviä. Myös aiempien tutkimusten mukaan 7-vuotias yliarvioi itseään ja taitojaan (Blatchford, 1997; Kaderavek ym., 2004), mikä voi johtua siitä, että taidot oman osaamisen tunnistamiseen eivät ole vielä kehittyneet. Kolmannella luokalla opiskelevien itsearviointit olivat kuitenkin jo toisella luokalla olevia pysyvämmin. Itsearviointitaidot kehittyvätkin harjoittelun myötä nopeasti, ja tutkimuksen mukaan kolmannella luokalla itsearviointikyky on jo tarkka (Blatchford, 1997).

Itsearvioinnin ja suoriutumisen välistä yhteyttä tutkittaessa selvisi, ettei sukupuolella ollut yhteyttä siihen, kuinka tarkasti lapset osaavat itsearvioida matematiikkaan ja lukemiseen liittyviä minäpystyvyyssuskomuksiaan. Tutkimusten mukaan sukupuolen merkitys itsearvioinnin tarkkuuteen riippuu siitä, missä asioissa itsearviointia on tarkasteltu (Kaderavek ym. 2004; Pallier, 2003). Matematiikan osalta sukupuolen merkitsemättömyys itsearvioinnin tarkkuuteen on ristiriidassa aiemman tutkimuksen kanssa, jonka mukaan naiset ovat taipuvaisia aliarvioimaan matematiikan taitojaan (Beyer, 2002). Toisaalta lukemisen osalla tämän tutkimuksen tulos noudatti aiemman tutkimuksen havaintoja siitä, että äidinkielen itsearvioinnissa sukupuolella ei ole merkitystä itsearvioinnin tarkkuuteen (Beyer, 2002). Kuitenkin tutkimustulosten tulkinnassa täytyy säilyttää varovaisuus, koska aikaisempi tutkimus on toteutettu pääosin aikuisille opiskelijoille. Niin kuin aikaisemmista tutkimuksista tiedämme, aikuistuuessa myös itsearviointitaidot paranevat psyykkisen kehittymisen ja itsearvioinnin harjoittamisen myötä, mikä vaikuttaa myös itsearvioinnin tarkkuuteen (Alaoutinen, 2012; Olina & Sullivan, 2004).

Tutkimustulosten mukaan sukupuolen tavoin myöskään luokkatasolla ei ollut merkitystä itsearviointien tarkkuuteen. Vaikka ulkomaisen tutkimuksen mukaan luotettava itsearviointi on jopa aikuiselle haastavaa (Beyer, 2002; Dupeyrat, 2011), Suomessa jo alkuopetusikäinen saattaa kyetä arvioimaan osaamistaan luotettavasti, koska kokemusta itsearviointien tekemisestä on todennäköisesti muualla maailmassa asuvia enemmän. Vaikka aiemmissa tutkimuksissa on todettu pienten lasten itsearviointien olevan epätarkkoja, onkin epätodennäköistä, että kaikki lapset olisivat huonoja itsearvioijia. Butlerin (1990) mukaan lasten leikeissä sattuisi enemmän vahinkoja ja onnettomuuksia, jos he aina yliarvioisivat taitonsa.

Eri kulttuureissa kasvaminenkin voi vaikuttaa itsearviointien tarkkuuteen, koska kaikkialla maailmassa ei ole totuttu tekemään itsearviointeja säännöllisesti. Monet tutkimustulokset vielä lukio- ja korkeakouluikäistenkin itsearviointien epätarkkuudesta voivat johtua siitä, että tutkimukset on toteutettu eri puolilla maailmaa. Suomessa itsearviointeja aletaan toteuttaa jo esikouluikäisillä, minkä vuoksi harjoittelun tuloksena suomalaislapset myös osannee itsearvioida luotettavasti muita lapsia nuorempina.

### **Itsearviointi erityisopetuksessa**

Tämän tutkimuksen mukaan siis jo alakouluikäiset oppilaat osaavat itsearvioida matematiikkaan ja lukemiseen liittyviä minäpystyvyyssuhteita tarkasti. Yleis- ja erityisopetuksen kannalta tieto lasten kyvystä tarkkaan itsearviointiin on myönteinen. Kuitenkin itsearviointien pysyvyyttä tutkittaessa havaittiin myös se, että sukupuolella ja luokkatasolla oli osassa ulottuvuuksista merkitystä testi-uusintatesti -reliabiliteetteihin. Koska itsearviointi vaatii metakognitiivisia taitoja (Alaoutinen, 2012), erityisopetuksessa tulisi kiinnittää huomiota erityisesti niiden lasten itsearviointitaitojen tukemiseen, joilla on pulmia oman osaamisen tiedostamisessa, tunnistamisessa ja arvioinnissa.

Erityisopetusta saavalla oppimisvaikeuksisella lapsella on monesti myös haasteita metakognitiivisissa taidoissa, eikä siis käsitys omista kyvyistä aina ole totuudenmukainen. Aikuisina usein oletamme, että jo alakoulu-

ikäinen lapsi pystyy itsenäisesti säätelemään omaa oppimistaan. Harvemmin ajattelemme kuitenkaan itsearviointin haastavuutta. Onhan mahdollista, että jotkut ihmiset eivät osaa koskaan arvioida työtään tai osaamistaan luotettavasti.

Itsearviointin haastavuuden vuoksi opettajan täytyy antaa palautetta tehdyistä itsearvioinneista ja myös neuvoa lasta näiden tekemisessä. Oppilasta voi esimerkiksi kehottaa vertaamaan suoritustaan tämän edelliseen suoritukseen. Lisäksi opettaja voi auttaa oppilasta hahmottamaan osaamistaan keskustelemalla siitä, mikä tehtävässä onnistui ja mitä olisi voinut tehdä toisin. Säännöllisen itsearviointin harjoittamisen myötä lapsi huomaa taitonsa kehittyvän ja hän saattaa motivoitua opiskelemaan aiempaa enemmän, jolloin myös oppimistulokset paranevat. Lisäksi itsearviointitaitojen kehittyessä oppilas oppii ottamaan vastuuta oppimisestaan, jolloin opettajan rooli opetuksessa pienenee ja voi muuttua oppilasta ohjaavaksi.

Palautteen saamisella voi olla itsearviointitaitojen kehittymisen lisäksi merkitystä myös minäpystyvyyteen (Panadero, Alonso-Tapia, & Reche, 2013). Joillakin lapsilla ei ole uskoa omaan osaamiseensa, minkä vuoksi koulussa kognitiivisen taitoharjoittelun ohessa on oleellista harjaannuttaa myös uskomuksia ja asenteita, jolloin myönteiset minäpystyvyyssuskomukset tuottavat onnistumisen kokemuksia. Onnistumisten myötä oppilas sitoutuu todennäköisesti myös koulutehtäviinsä, jolloin oppimisestakin tulee tehokkaampaa.

Itsearviointien harjoittaminen ja niistä keskusteleminen auttavat lasta ottamaan vastuuta omasta oppimisestaan ja tulemaan tietoisiksi itsestään, jolloin myös epärealistiset käsitykset itsestä ja suoriutumisesta vähenevät, ja itsearviointi lähenee todellisen suoriutumisen tasoa. Oppilaiden itsearvioinneista ja lasten kanssa käydyistä keskusteluista myös opettaja saa arvokasta tietoa siitä, kuinka hän voi kehittää omaa opetustaan, ja näin edistää myös oppilaiden oppimista. Lisäksi opettaja saa keskustelujen kautta tietoa oppilaan ajatusmaailmasta ja siitä, onko oppilas ymmärtänyt arvioitavan asian samoin kuin opettaja. Toisinaan opettaja saattaa arvioida oppilaan osaamis-

ta, ja oppilas taas yrittämisen määrää, jolloin itsearviointi ei ole yhtenevä opettajan arvioinnin kanssa.

### **Tutkimuksen luotettavuus, rajoitukset ja yleistettävyys**

Erot eri ulottuvuuksien itsearvioinnin pysyvyydessä voivat johtua esimerkiksi lapsen kyvyttömyydestä itsearvioida tälle vieraita asioita. Erilaiset tulokset saattavat olla myös merkki siitä, että itse mittari ei ole toiminut luotettavasti. Vaikka 7-portainen asteikko antaa mahdollisuuden tuntemusten monipuoliseen kuvaamiseen, asteikko voi olla tutkimuksen nuoremmille osallistujille vanhempia osallistujia vaikeampi hahmottaa. Sävyerot ovat pieniä esimerkiksi vaihtoehdon 6 (Melko varma, että pystyt tekemään kysytyn asian) ja 5 välillä (Ehkä pystyt tekemään kysytyn asian). Nopeatempoisessa testitilanteessa tilanteeseen sopivimman vaihtoehdon valinta voi olla haastavaa. Tämänkin haasteen ennaltaehkäisemiseksi vastaustekniikkaa 7-portaisella asteikolla harjoiteltiin ennen kyselylomakkeen täyttämistä kolmen eri harjoituskysymyksen avulla.

Korrelaatioiden vaihtelussa täytyy huomioida myös testi-uusintatesti -menetelmän rajoitukset, joista yhtenä voi olla mittarin reliabiliteetti. Tässä tutkimuksessa mittarin reliabiliteettia tutkittiin toistomittauksen avulla viidessä eri ulottuvuudessa, ja näiden ulottuvuuksien toistomittauskorrelaatiot osoittautuivat korkeiksi. Kuitenkin tehtävissä ponnistelemisen ja oppimiseen liittyvien tunteiden itsearvioinnit eivät olleet aivan yhtä pysyviä kuin esimerkiksi matematiikan ja lukemisen itsearvioinnit. Näiden ulottuvuuksien tulosten tulkinnassa tulee olla varovainen, koska mittari ei ole välttämättä toiminut riittävän luotettavasti tehtävissä ponnistelemisen ja tunteiden itsearvioinneissa.

Mittarin tarkempaa tutkimista ja sen kehittämistä varten tutkimuksessa tarkasteltiin myös summamuuttujien osioiden ja yksittäisten kysymysten pysyvyyttä. Matematiikan ja lukemisen summamuuttujien osioiden pysyvyydessä ei ollut tilastollisesti merkitsevää eroa. Tämän vuoksi näissä ulottuvuuksissa summamuuttujat oli perusteltua muodostaa sekä minäpystyvyyden että sen lähteiden muuttujista. Keskittymisen ulottuvuuden osioiden

välillä oli tilastollisesti merkitsevä ero. Tämä kahdesta hieman erityyilisestä osiosta muodostettu summamuuttuja saattoi vaikuttaa keskittymisen itsearvioinnin pysyvyyteen, koska lapsen usko siihen, kuinka hyvin hän pystyy keskittymään voi poiketa siitä, kuinka hyväksi hän arvioi keskittymisen taitonsa.

Suurin osa matematiikan, lukemisen ja keskittymisen ulottuvuuksien yksittäisten muuttujien testi-uusintatesti -reliabiliteeteista ylsi melko korkeiksi tai korkeiksi. Kuitenkin yhden matematiikan ja lukemisen minäpystyvyyden osion ensimmäisten kysymysten testi-uusintatesti -reliabiliteetit olivat matalia (ks. liite 1: Kuinka varma olet, että pystyt lukemaan näin pitkän tekstin / laskemaan tällaisia laskuja nopeasti puolessa minuutissa?). Orientoituminen uudenlaisen tehtävän ensimmäiseen kysymykseen on saattanut olla lapsille haastavaa. Mittarin kehittämisessä tulisikin huomioida osioiden ensimmäisten kysymysten matalat korrelaatiot esimerkiksi tekeillä myös tämän osion alkuun yksi harjoittelukysymys. Tämän kysymyksen avulla lapset saisivat harjoitella uudenlaiseen tehtävään vastaamista.

Toiseksi testi-uusintatesti -reliabiliteettiin voi vaikuttaa mittausten välinen aika. Jotta mitattava ominaisuus ei ehtisi muuttua mittauskertojen välillä, luotettavan tuloksen saamiseksi aikavälin mittausten välillä tulisi olla noin kaksi viikkoa, muttei ehdottomasti yli puolta vuotta (Metsämuuronen, 2009, 141). Itsearviointitaidot on hitaasti muuttuva ominaisuus, joten mittauskertojen välisen kaksikuukautisen ei pitäisi olla liian pitkä aikaväli luotettavien tulosten saamiseksi.

Kolmanneksi toistomittausten korrelaatioon saattaa vaikuttaa kyselylomakkeen kysymysten ja edellisen kerran vastausten muistaminen. Tämän epäkohdan pienentämiseksi tutkimukseen valittiin hankkeen kaksi ensimmäistä mittauskertaa. Ensimmäisellä mittauskerralla testin toimintaperiaate on lapselle uusi. Toisella kerralla kahden kuukauden kuluttua testi on osallistujalle tuttu, mutta sen yksityiskohdat eivät todennäköisesti ole enää kirkkaina lapsen mielessä.

Lisäksi yksilölliset ja testausolosuhteisiin liittyvät tekijät voivat vaikuttaa testi-uusintatesti -reliabiliteettiin. Isossa oppilasryhmässä toteutettava

testaus saattoi vaikuttaa heikentävästi joidenkin lasten kykyyn vastata tehtäviin huolellisesti. Huolimatonta vastaamista saattoi aiheuttaa myös tehtävämönisteet, joiden täytyi olla melko pitkiä, jotta tutkimuksen monet testit saatiin suoritettua. Kaikki lapset eivät jaksaneetkaan keskittyä jopa tunnin kestävään testaukseen, vaan jotkut päätyivät arvaamaan vastauksia tai valikoimaan joko vaihtoehdon 1, 4 tai 7. Kuitenkaan näiden epäluotettavien vastausten ei pitäisi vaikuttaa suuren otannan tutkimustuloksiin. Silti mittarin kehittämisessä olisi hyvä huomioida lasten erilainen kyky keskittyä ja ponnistella. Testitilanteessa turhautuville ja väsyville olisi tarpeellista järjestää mahdollisuus suorittaa tehtävät esimerkiksi kahdessa eri osassa.

Tutkimuksen reliabiliteetin arvioimisen lisäksi tarvitaan myös validiteettitarkastelua. Ensinnäkin mittarit on laadittu teoriaan perustuen ammattilaistutkijoiden yhteistyönä, ja ne on joko esitestattuja tai jo julkaistuissa tutkimuksissa käytettyjä. Validiteetin uhkia pyrittiin ennakoimaan myös kouluttamalla testajat työhönsä sekä antamalla heille yhdenmukaiset kirjalliset ohjeet, joiden mukaan testaukset toteutettiin.

Lisäksi mittarit koostuvat useiden muuttujien tai testien summista, joiden reliabiliteettien testattiin olevan erittäin korkeat. Summamuuttujat kuvaavat ulottuvuuksia monipuolisesti, mutta niissä on myös rajoituksensa: Summamuuttujat muodostettiin niin minäpystyvyyden kuin myös taidon tason arvioinnin muuttujista. Kuitenkaan kaikilla lapsilla luottamus omaan osaamiseensa ei välttämättä ole sama kuin se, kuinka hyväksi oma taito arvioidaan.

Tutkimustuloksia voidaan yleistää koskemaan Keski-Suomen 2.-5.-luokkalaisia oppilaita tämän alueen oppilaiden suuren osallistujamäärän vuoksi. Toisaalta Suomen pienten koulutuksellisten erojen vuoksi ei ole ongelmallista varovaisesti yleistää tutkimustulosten koskevan koko Suomen tämän ikäisiä lapsia.

Tutkimustulosten yleistämiseen liittyy kuitenkin myös rajoituksia aikaisemman tutkimustiedon vähäisyyden vuoksi. Itsearviointin testi-uusintatesti -reliabiliteetin tuloksia on mahdoton peilata vähäiseen aiempaan tutkimustietoon, joten yhteyksiä on luotu yleiseen tutkimukseen ja itsear-



vioinnin luotettavuutta käsittelevään tutkimustietoon. Lisäksi aiemmissa tutkimuksissa itsearviointin tarkkuutta on tutkittu vertaamalla itsearviointeja joko omaan suoriutumiseen, vertaisarviointiin tai opettajan arviointiin. Tämän tutkimuksen tuloksia on tarkasteltu suhteessa näihin kaikkiin tutkimusasetelmiin, joten mittaustuloksissa saattaa olla asetelmista johtuvia eroavaisuuksia.

Kun puhutaan tarkasta itsearvioinnista, tarkoitetaan sillä yleensä sitä, että kahden vertailtavan vastauksen yhtenevyyden lisäksi tuloksessa ei ole yli- ja aliarviointia. Tässä tutkimuksessa yli- ja aliarvioimista ei kuitenkaan ollut mahdollisuutta arvioida, jolloin näennäisesti tarkka itsearviointi on saattanut sisältää testi-uusintatestissä yli- tai aliarviointia. Tällaisessa tilanteessa tutkimuksen tuloksista ei ole mielekästä tuottaa pitkälle meneviä johtopäätöksiä. Lisäksi itsearvioinnissa lapsen tehtävänä oli pääosin arvioida minäpystyvyyttään matematiikassa ja lukemisessa, mikä saattaa poiketa siitä, miten tämä suoriutuu näissä tehtävissä. Jonkun lapsen kohdalla voi olla niin, että minäpystyvyys matematiikassa on heikko, mutta lapsi silti suoriutuu hyvin matematiikan tehtävästä. Tällöin lapsen itsearviointitaito näyttää tutkimustuloksissa heikolta, vaikka todellisuudessa tämän itsearvio minäpystyvyydestään voi olla hyvinkin totuudenmukainen.

Edellä mainituista rajoituksista huolimatta tämän tutkimuksen tulokset antavat suuntaviivoja sille, minkä asioiden itsearvioiminen on ollut lapselle helppoa ja mikä on toisaalta ollut haastavaa. Tutkimus valottaa myös sukupuolen ja luokkatason yhteyttä itsearviointin pysyvyyteen sekä itsearviointin tarkkuuteen.

## **Jatkotutkimushaasteita**

Itsearviointin luotettavuus ajankohtaisena ja mielenkiintoisena ilmiönä tarvitsee jatkotutkimusta. Etenkin lasten itsearviointin luotettavuuden tutkimusta eri oppiaineissa tulisi kartoittaa lisää. Lisäksi olisi tärkeää tutkia sitä, millä keinoilla lapsia voitaisiin jo pienenä tukea niin, että matematiikkaa ja lukemista haastavampina itsearvioitavina todetut tehtävissä ponnisteleminen, keskittyminen ja tunteiden itsearvioiminen olisivat luotettavia. Inter-

ventiotutkimuksen avulla olisi hyödyllistä kartoittaa myös sitä, vaikuttaako poikien keskittymisen harjoittaminen siihen, että näiden itsearviointi olisi tyttöjen itsearvioinnin kanssa yhtä pysyvää matematiikan, lukemisen ja oppimiseen liittyvien tunteiden itsearvioinnissa.

## LÄHTEET

- Adeyemi, A. A. (2012). Effect of peer and self-assessment on male and female students' self-efficacy and self-autonomy in the learning of mathematics. *Gender & Behaviour, 10*(1), 4492–4508.
- Aho, S. (2004). Minä. Teoksessa S. Aho & K. Laine (toim.) *Minä ja muut. Kasvaminen sosiaaliseen vuorovaikutukseen*. Keuruu: Otava, 16–67.
- Alaoutinen, S. (2012). Evaluating the effect of learning style and student background on self-assessment accuracy. *Computer Science Education, 22*(2), 175–198.
- Bandura, A. (1994). Self-efficacy. *Encyclopedia of human behavior, 4*, 71–81.
- Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents, 5*, 307–337.
- Beyer, S. (2002). The effects of gender, dysphoria, and performance feedback on the accuracy of self-evaluations. *Sex Roles, 47*(9/10), 453–464.
- Blatchford, P. (1997). Students' self-assessment of academic attainment: Accuracy and stability from 7 to 16 years and influence of domain and social comparison group. *Educational Psychology, 17*(3), 345–359.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology, 90*(1), 60–77.
- Butler, R. (1990). The effects of mastery and competitive conditions on self-assessment at different ages. *Child Development, 61*(1), 201–210.

- Dupeyrat, C., Escribe, C., Huet, N., & Regner, I. (2011). Positive biases in self-assessment of mathematics competence, achievement goals, and mathematics performance. *International Journal of Educational Research*, 50(4), 241–250.
- Guilford, J. P. (1965). *Fundamental statistics in psychology and education*. New York: McGraw-Hill.
- Halinen, I. (1995). Itsearviointia Eläintarhan ala-asteella. Teoksessa B. Kilpinen, K. Salmio, L. Vainio, & A. Vanne (toim.) *Itsearviointin teoriaa ja käytäntöä*. Helsinki: Yliopistopaino, 99–110.
- Huusko, M. (2008). Itsearviointi kehittävän arvioinnin menetelmänä. Teoksessa E. Korkeakoski & H. Silvennoinen (toim.) *Avaimia koulutuksen arvioinnin kehittämiseen*. Jyväskylä: Gummerus, 127–138.
- Ihme, I. (2009). *Arviointi työvälineenä: Lasten ja nuorten kasvun tukeminen*. Jyväskylä: PS-kustannus.
- Kaderavek, J. N., Gillam, R. B., Ukrainetz, T. A., Justice, L. M., & Eisenberg, S. N. (2004). School-age children's self-assessment of oral narrative production. *Communication Disorders Quarterly*, 26(1), 37–48, 56, 58.
- Keltikangas-Järvinen, L. (2001). *Hyvä itsetunto*. Juva: WSOY.
- Kerlinger, F. N. (1986). *Foundations of behavioral research*. 3. painos. Texas, FW: Harcourt Publishers.
- Lew, M. D., Alwis, A. W., & Schmidt, H. G. (2010). Accuracy of students' self-assessment and their beliefs about its utility. *Assessment & Evaluation in Higher Education*, 35(2), 135–156.
- Lyytinen, H. K. (1995). Johdatus oppilaitoskohtaiseen itsearviointiin. Teoksessa B. Kilpinen, K. Salmio, L. Vainio, & A. Vanne (toim.) *Itsearviointin teoriaa ja käytäntöä*. Helsinki: Yliopistopaino, 37–56.
- Mabe, P. A. & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280–296.
- Metsämuuronen, J. (2009). *Tutkimuksen tekemisen perusteet ihmistieteissä*. Jyväskylä: Gummerus.

- Ng, J. R. & Earl, J. K. (2008). Accuracy in self-assessment: The role of ability, feedback, self-efficacy and goal orientation. *Australian Journal of Career Development, 17*(3), 39-50.
- Nummenmaa, L. (2009). Käyttäytymistieteiden tilastolliset menetelmät. Keuruu: Tammi.
- Olina, Z. & Sullivan, H. J. (2004). Student self-evaluation, teacher evaluation, and learner performance. *Educational Technology Research and Development, 52*(3), 5-22.
- Pajares, F. & Schunk, D. (2001). The development of academic self-efficacy. Teoksessa A. Wigfield & J. Eccles (toim.) *Development of Achievement Motivation*. San Diego: Academic Press, 1-27.
- Pallier, G. (2003). Gender differences in the self-assessment of accuracy on cognitive tasks. *Sex Roles, 48*(5-6), 265-276.
- Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation, 39*(3), 125-132.
- Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice, 21*(2), 133-148.
- Perusopetuslaki. (2003). Oppilaan arviointi. 22 §.  
<https://www.finlex.fi/fi/laki/ajantasa/1998/19980628>. Luettu 22.1.2015.
- Salmio, K. & Vainio, L. (1995). Johdanto. Teoksessa B. Kilpinen, K. Salmio, L. Vainio, & A. Vanne (toim.) *Itsearviointin teoriaa ja käytäntöä*. Helsinki: Yliopistopaino, 7-11.
- Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 19*(2), 159-172.
- Schunk, D. H. & Meece, J. L. (2006). Self-efficacy development in adolescence. *Self-Efficacy Beliefs of Adolescents, 5*, 71-96.

- Soper, D.S. (2015). Significance of the Difference between Two Correlations Calculator [Software]. Saatavissa myös:  
<http://danielsoper.com/statcalc3/calc.aspx?id=104>. Luettu 22.1.2015.
- Stone, A. C., & May, A. L. (2002). The accuracy of academic self-evaluations in adolescents with learning disabilities. *Journal of Learning Disabilities*, 35(4), 370–383.
- Sung, Y., Chang, K., Chang, T., & Yu, W. (2010). How many heads are better than one? The reliability and validity of teenagers' self- and peer assessments. *Journal of Adolescence*, 33(1), 135–145.
- Tari, J. J. (2010). Self-assessment processes: The importance of follow-up for success. *Quality Assurance in Education*, 18(1), 19–33.
- Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology*, 91(5), 1146–1153.

## LIITE

### Minäpystyvyyden kysymyksistä ja minäpystyvyyden lähteiden väittämistä muodostetut viisi keskiarvosummamuuttujaa sekä yksittäisten muuttujien testi-uusintatesti -reliabiliteetit.

#### Matematiikka

##### Minäpystyvyys:

Kuinka varma olet, että pystyt

...oppimaan laskemaan yhteenlaskuja nopeammin? (0.34)

...oppimaan laskemaan niin, että teen vähemmän virheitä? (0.44)

...laskemaan yhteenlaskuja nopeasti mielessä? (0.51)

...laskemaan kaupassa, kuinka paljon sinulla on rahaa? (0.55)

...laskemaan kaupassa, riittävätkö rahasi ostoksiin tai kuinka paljon saat rahaa takaisin? (0.55)

...pelaamaan lautapelejä, joissa pitää laskea päässä nopeasti? (0.51)

...laskemaan tällaisia laskuja nopeasti? (0.33)

...laskemaan tällaisia laskuja nopeasti? (0.45)

...laskemaan tällaisia laskuja nopeasti? (0.56)

...laskemaan tällaisia laskuja nopeasti? (0.56)

...laskemaan tällaisia laskuja nopeasti? (0.56)

...laskemaan tällaisia laskuja nopeasti? (0.55)

...laskemaan tällaisia laskuja nopeasti? (0.56)

...laskemaan tällaisia laskuja nopeasti? (0.59)

##### Minäpystyvyyden lähteet:

Olen aina ollut hyvä matematiikassa. (0.62)

Osaan hyvin matematiikkaa. (0.68)

Osaan laskea hyvin vaikeitakin matematiikan tehtäviä. (0.63)

Ajattelen usein, että jonain päivänä olen taitava laskija. (0.51)

Voin oppia matematiikkaa, jos teen kovasti töitä. (0.37)

En voi tulla paremmaksi matematiikassa, vaikka harjoittelisin kovasti. (0.59)

Minusta tulee hyvä matematiikassa. (0.48)

## Lukeminen

### **Minäpystyvyys:**

Kuinka varma olet, että pystyt

...oppimaan lukemaan nopeammin? (0.11)

...oppimaan lukemaan niin, että teen vähemmän virheitä? (0.42)

...oppimaan lukemaan niin, että ymmärrät kaiken, mitä luet? (0.47)

...lukemaan kaikki TV-ohjelmien tekstitykset helposti? (0.61)

...lukemaan internetissä pitkiä tekstejä? (0.55)

...lukemaan helposti ohuen kirjan? (0.32)

...lukemaan helposti paksun kirjan (esim. Harry Potter-kirjan)? (0.57)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.19)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.40)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.52)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.54)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.54)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.57)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.54)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.56)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.52)

...lukemaan näin pitkän tekstin puolessa minuutissa? (0.45)

### **Minäpystyvyyden lähteet:**

Olen aina ollut hyvä lukija. (0.56)

Osaan lukea hyvin. (0.57)

Osaan lukea hyvin vaikeitakin tekstejä. (0.57)

Ajattelen usein, että jonain päivänä olen taitava lukija. (0.53)

Voin oppia lukemaan paremmin, jos teen kovasti töitä. (0.42)

En voi tulla paremmaksi lukijaksi, vaikka harjoittelisin kovasti. (0.53)

Minusta tulee hyvä lukija. (0.38)

## Ponnisteleminen

### **Minäpystyvyys:**

Kuinka varma olet, että pystyt

...tekemään kaikki kotitehtävät, vaikka ne olisivat sinulle vaikeita?

...ponnistelemaan myös vaikeissa tehtävissä?

...tekemään parhaasi, vaikka huomaisit kokeen olevan sinulle vaikea?

...tekemään hyvin kaikki tehtävät koulussa, vaikka ympärillä olisi muuta mielenkiintoista tekemistä?

...tekemään kaikki tehtävät koulussa silloinkin, kun ne tuskastuttavat sinua?

...tekemään kaikki kotitehtävät, vaikka luettavat tekstit olisivat sinulle vaikeita?

...tekemään kaikki matematiikan kotitehtävät, vaikka ne olisivat sinulle vaikeita?

## Keskittyminen

### **Minäpystyvyys:**

Kuinka varma olet, että pystyt

...keskittymään hyvin tehtäviin luokassa? (0.46)

...seuraamaan tarkasti opettajan antamia ohjeita? (0.44)

...keskittymään lukemiseen silloinkin, kun se tuntuu tylsältä? (0.42)

...keskittymään laskemiseen silloinkin, kun se tuntuu tylsältä? (0.47)

...keskittymään hyvin lukemiseen luokassa? (0.42)

...keskittymään hyvin matematiikan tehtäviin luokassa? (0.43)

### **Minäpystyvyyden lähteet:**

Olen aina jaksanut keskittyä läksyihin. (0.53)

Kuuntelen aina opettajaa, kun hän opettaa. (0.55)

Olen aina jaksanut keskittyä oppitunneilla opetukseen. (0.54)

Olen aina jaksanut keskittyä luokassa opettajan antamiin tehtäviin. (0.51)



## Oppimiseen ja oppimistilanteisiin liittyvät tunteet

### **Minäpystyvyyden lähteet:**

Ahdistun, kun joudun vastaamaan matematiikan tunnilla.

Ahdistun, kun tiedän, että joudun lukemaan ääneen tunnilla.

Ahdistun, kun alan tehdä matematiikan tehtäviä.

Ahdistun, kun aloitan lukemisen.

Tunnen kehossani jännitystä, kun minun pitää tehdä matematiikan tehtäviä.

Tunnen kehossani jännitystä, kun minun pitää lukea.

Kokeet saavat minut ahdistumaan.

Opettajan kysymyksiin vastaaminen jännittää.

Pelkään tehdä virheitä luokassa.

Vaikeudet oppimisessa ahdistavat minua.

Pysyn aina rauhallisena, vaikka tehtävien tekemisessä tulisi ongelmia.

Saan vastattua kokeissa, vaikka minua jännittäisikin, kuinka pärjään.

Hermostun, kun teen läksyjä.

Hermostun koulutunnilla vastoinkäymisistä niin, että en pysty tekemään tehtäviäni.

Hermostun koulutunnilla vastoinkäymisistä niin, että en pysty keskittymään.