

Cloze-testi luetun ymmärtämisen mittarina

3. ja 4.-luokkalaisilla

Miki Mäenpää
Rami Radwan
Pro gradu -tutkielma
Psykologian laitos
Jyväskylän yliopisto
Lokakuu 2019

JYVÄSKYLÄN YLIOPISTO

Psykologian laitos

RADWAN, RAMI & MÄENPÄÄ, MIKI: Cloze-testi luetun ymmärtämisen mittarina 3.-4.-
luokkalaisilla

Pro gradu -tutkielma, 33 s.

Ohjaaja: Jarkko Hautala

Psykologia

Lokakuu 2019

Tutkimuksessa tarkasteltiin, miten kehittämämme cloze-testi mittaa luetun ymmärtämistä 3.-4. -luokkalaisilla. Tutkimus on osa Niilo Mäki Instituutin ReadDrama-hanketta, jossa tutkitaan lukuteatterin vaikuttavuutta lukuharjoittelun muotona. Kaikki 114 osallistujaa tulivat samasta Jyväskylän alakoulusta, josta kerättiin aineisto kevätlukukauden 2019 aikana. Vertailtavaksi testiksi otettiin Suomessa yleisessä käytössä oleva Ala-asteen lukutesti eli ALLU. Testejä vertailemalla arvioitiin clozen validiteettia luetun ymmärtämisen testinä. Lisäksi osallistujilta mitattiin kognitiivisia taustatekijöitä luetun ymmärtämisen teorioita integroivaan malliin (eng. Integrated Model of Reading Comprehensio, IMREC) pohjautuen. Tarkoituksena oli selvittää, mitkä kognitiiviset prosessit ovat aktiivisina luetun ymmärtämisen aikana ja tarkastella niiden yhteyttä cloze-testiin. Mitattaviksi kognitiivisiksi taustatekijöiksi valittiin lukusujuvuus, sanaston laajuus, näönvarainen päättelykyky, kielellinen päättelykyky, lyhykestoinen muisti ja työmuisti. Testien välisiä yhteyksiä sekä kognitiivisten taustatekijöiden yhteyttä luetun ymmärtämisen testeihin tarkasteltiin korrelaatiomatriisin ja regressiomallien avulla.

Testien välinen vahva korrelaatio ja kummankin testin pääosin samankaltaiset korrelaatiot kognitiivisiin taustatekijöihin antavat viitteitä siitä, että testit mittaavat samaa taitoa. Tulokset viittaavat siihen, että cloze-testillä voi mitata validisti luetun ymmärtämistä 3.-4. -luokkalaisilla, kun sen tuloksia verrattiin Ala-asteen lukutestin tuloksiin. Kognitiivisista taustatekijöistä päättelykyky korostui cloze-testissä enemmän kuin ALLUssa. Päättelykyky vaikuttaa olevan keskeinen komponentti luetun ymmärtämisessä.

AVAINSANAT: luetun ymmärtäminen, cloze-testi, kognitiiviset taustatekijät, ala-asteen lukutesti, lukemisen arviointi, päättelykyky, alakoulu, oppilaat, sidoksisuus, aukot, johdonmukaisuus

Sisällys

JOHDANTO	1
Lukemisen arviointi.....	2
Luetun ymmärtämisen teorit ja mallit	3
Kognitiiviset teorit.....	4
Luetun ymmärtämisen kognitiiviset taustatekijät	8
Luetun ymmärtämisen mittaaminen.....	11
Cloze-testi.....	14
Tutkimuskysymykset.....	17
AINEISTO & MENETELMÄT	18
Menetelmät ja muuttujat.....	18
Cloze-testi.....	19
Esimerkkitekstejä cloze-testistä	19
Lukusujuvuus.....	20
Päättelykyky ja muisti	21
Sanasto.....	22
Ala-asteen lukutesti	23
Aineiston analysointi	23
TULOKSET.....	24
POHDINTA.....	27
LÄHTEET	34

JOHDANTO

Lukemisen taito on edellytys lähes kaikissa oppiaineissa pärjäämiselle. Lukutaidosta erotellaan yleensä mekaaninen osa, jossa yksittäisistä kirjaimista muodostetaan suurempia osasia kuten tavuja ja lopulta sanoja sekä ymmärtämisen osa, jossa yksittäisistä sanoista muodostetaan suurempia merkityksellisiä kokonaisuuksia. Suomen kielen lähes täydellinen kirjain-äännevastaavuus (Aro 2004; Seymour, Aro & Erskine 2003) mahdollistaa muita kieliä helpommin mekaanisen lukemisen oppimisen, mutta mekaaninen lukutaito ei vielä takaa sitä, että lukeminen johtaisi tekstin ymmärtämiseen. Lukemisen perimmäinen tavoite on saada tekstistä tietoa, joka tapahtuu tekstin ymmärtämisen välityksellä. Ymmärtäessään tekstiä lukija pystyy kokoamaan mielessään tekstistä kokonaisuuden, jossa tekstin välittämä tieto nivoutuu johdonmukaiseksi kokonaisuudeksi välittäen tekstin perimmäisen viestin tai opin.

Heikko lukutaito vaikuttaa kielteisesti lapsen alkavaan lukuinnostukseen ja opiskelumotivaatioon (Lyytinen ym., 2005; Ahonen, & Korkman, 2012) sekä myöhempään akateemiseen menestykseen. Lisäksi se nostaa syrjäytymisen riskiä nuorilla: 15-vuoden iässä nuoren aikuisen työttömyyttä ennustavat erityisesti koulun päättötodistuksen puuttuminen, huono lukutaito ja osallistumattomuus koulunkäyntiin (Caspi, Wright, Moffitt & Silva, 1998). Peruskoulun jälkeisestä koulutuksesta jättäytyminen johtaa useimmiten syrjäytymiseen työelämästä (Johansson & Vuori 1999), mikä näkyy etenkin miehillä.

Tämän vuoksi järjestelmällinen lukutaidon arviointi ja lukemisen ongelmiin puuttuminen on tärkeää jo varhaisessa vaiheessa. Suomessa tarpeeseen on vastattu kehittämällä useita lukemisen arviointimenetelmiä, jotka ovat kouluissa yleisesti käytössä. Niillä ei kuitenkaan välttämättä saavuteta tarkkaa kuvaa koululaisten lukutaidon tasosta. Lukemisen arviointiin liittyvä vuosikymmenien ajan jatkunut väittely ideaalin mittarin ominaisuuksista on vieläkin käynnissä. Esimerkiksi kognitiivisten taitojen roolit luetun ymmärtämisen selittäjinä korostuvat erilailta riippuen siitä, mitä lukemisen testiä käytetään. Lisäksi mittaamiseen kuluvat resurssit, niin ajalliset kuin taloudelliset, asettavat lisäpaineita uudistuksille. Teknologisen kehityksen myötä onkin oleellista pohtia testaamisen sähköistämistä ratkaisuna resurssien luomiin rajoitteisiin. Tämä herättää kysymyksiä siitä, miten tulevaisuudessa pystytään kehittämään arviointimenetelmiä niin, että niillä kyetään vastaamaan ajan vaatimuksiin.

Lukemisen arviointi

Suomalaisissa kouluissa lukutaitoa arvioidaan yleisesti Ala-asteen lukutestillä eli ALLUlla (Turunen, Alisaari, Poskiparta & Lindeman, 2018). Testi mittaa oppilaiden kielellistä tietoisuutta, teknistä lukemista ja luetun ymmärtämistä. Testi on normitettu vuonna 1995 sadoilla eri peruskouluilla käyttäen eri luokka-asteita (luokat 1-6) (Lindeman, 1998). Luokka-asteittainen normitus mahdollistaa lukutaidon tason vertailtavuuden läpi luokka-asteiden. Testi tehdään yleensä vuosittain, jolloin saadaan käsitys oppilaiden lukutaidon kehittymisestä ja mahdollisten interventioiden tarpeista. ALLU on ainut alakoululaisilla normitettu lukutaidon testi Suomessa.

Toimivan testin kannalta on tärkeää, että sen avulla voidaan tehdä tarkkoja johtopäätöksiä ja toimenpiteitä, jotta voidaan suunnata tehokkaasti resursseja lukemisen tukemiseen. Suomessa luetun ymmärtämistä testataan yleisesti joko tekstiin pohjautuvilla yksittäisillä tehtävillä, kuten avoimilla kysymyksillä, monivalintatehtävillä, oikein- ja väärin kysymyksillä ja tekstistä muistettujen asioiden määrällä tai niiden yhdistelmillä. (Kuisma & Myllylä 2013). Tutkimukset osoittavat, että nämä tehtävätyypit eivät välttämättä kata samoja kognitiivisia prosesseja ja ne saattavat korostaa eri voimakkuudella tiettyjen prosessien roolia luetun ymmärtämisessä (Cutting, 2006; Spear-Swerling, 2004; Nation & Snowling, 1997; Bowey, 1986).

Kuhunkin tehtävätyyppiin liittyy omanlaisia ongelmia, jotka heikentävät mahdollisuutta tehdä tuloksista tarkkoja johtopäätöksiä. Esimerkiksi lukemisen jälkeen täytettävissä monivalintatehtävissä oikea vastaus voi olla pääteltävissä, vaikka tekstiä ei olisi luettu ollenkaan (Keenan & Betjemann, 2006). Toisaalta kun teksti luetaan, sen haastavuus korostaa eri luetun ymmärtämisen osatekijöitä riippuen lukijan taitotasosta. (Amendum, Conrad & Hiebert, 2018). Tehtävissä, joissa täytetään tekstiin jätettyjä aukkoja, saattaa dekodeaus rooli korostua siinä määrin, ettei testillä tavoiteta korkeampia luetun ymmärtämiseen tarvittavia prosesseja (Alderson, 1979). Lisäksi erilaisten tehtävätyyppien käyttö aiheuttaa yleistettävyyss- ja vertailtavuusongelmia, koska eri testeistä saadut pistemäärät eivät säilytä oppilaiden välistä suhteellista järjestystä luetun ymmärtämisen taidossa (van den Broek & Espin, 2012).

ALLU on normitettu testi, joten siitä saatujen tulosten vertailtavuus on perusteltua, mutta yli 20 vuotta käytössä ollut testi ei ole kuitenkaan ongelmaton. Turusen, Alisaaren, Poskiparran & Lindemanin (2018) tutkimuksessa, jossa tarkasteltiin opettajien kokemuksia ALLUn käytöstä, 1009 opettajasta 88.9% arvioi testin kaipaavan uudistusta. Testiä päivitettiin, mutta uudistamisessa keskityttiin aluksi lähinnä vanhentuneiden sanojen ja ilmausten nykyaikaistamiseen, joten mahdolliset rakenteelliset puutteet jäivät huomiotta. Testissä käytetään mm. lukemisen jälkeen täytettäviä monivalintatehtäviä, jotka eivät välttämättä vaadi testattavalta tehtäviin liittyvän tekstin

lukemista (Keenan & Betjemann, 2006). ALLU normitettiin vuonna 1995, joten se ei välttämättä vastaa nykyajan standardeja (Turunen ym., 2018). ALLUa validoitaessa ei vielä ollut olemassa muita kriteerimuuttujaksi soveltuvia lukutaidon testejä, joten validoinnin kriteerimuuttujana oli käytettävä lukutaidon kouluarvioita (Lindeman, 1998), jotka eivät välttämättä kuvaa puhtaasti lukemista. Myöhemmäksi uudistamistavoitteeksi on kuitenkin otettu testin digitalisointi ja uuden normiaineiston kerääminen (Turunen ym., 2018). ALLUun sisältyy monenlaisia tekstitaidon tehtäviä, joista suositellaan tehtäväksi vähintään neljä tekstiä (Lindeman, 1998). Tästä syystä ALLUn tekeminen vie useita oppitunteja, kun se tehdään manuaalin mukaisella tavalla, mikä voi näkyä oppilaiden motivaatiossa ja jaksamisessa, jolloin nämä tekijät saattavat korostua tuloksissa epätarkoituksenmukaisesti. Etuna saattaa olla kuitenkin sen kyky kattaa suuri osa lukemisen taidoista sen sisältäessä erilaisia tekstitaitoja mittaavia osioita.

Testien taustalla on erilaisia teoreettisia lähtökohtia, jotka ovat omiaan vaikuttamaan siihen, miten luetun ymmärtämistä lähestytään ja miten se lopulta käsitetään. Lähestymistapa vaikuttaa edelleen mittarien valintaan ja siihen mitä tehtävätyyppettä testeissä käytetään. Se, mille tasolle oppilaiden luetun ymmärtämisen taito sijoittuu suhteessa muihin oppilaisiin, on siis riippuvainen testien taustalla vaikuttavista teorioista. Tämän vuoksi tarvitaan testiä, jossa yhdistellään eri osatekijöitä kuvaamaan luetun ymmärtämistä yhtenä ilmiönä ja kontrolloidaan mittaustapaan liittyviä ongelmia.

Luetun ymmärtämisen teorit ja mallit

Luetun ymmärtämisestä on muodostettu useita teorioita esimerkiksi behavioraalista, kognitiivisesta, konstruktivistisesta ja transaktionaalisesta näkökulmasta. Koska lukeminen on monimutkainen ja laaja prosessi, ei ole olemassa yhtä teoriaa, jolla voitaisiin selittää kaikkia lukemiseen liittyviä tekijöitä. Yhdistelemällä eri teorioita päästään kuitenkin kokonaisvaltaisemmin käsiksi siihen, mitä lukeminen on. Vaikka teorioissa keskitytään tarkoituksenmukaisesti yksittäisiin tekijöihin, kuten motivaation vaikutukseen lukemisessa, niin kokonaisvaltaisen kuvan saamiseksi on lukutaidosta muodostettava useampia teorioita yhdistäviä malleja. Monet lukemisen teorioiden pohjalta muodostetut mallit, kuten Simple view of reading (Hoover & Goughin, 1990), Inferential mediation model (Cromley & Azevedo, 2007), Verbal efficiency theory (Perfetti, 1985) ja Reading Systems Framework (Perfetti & Stafura, 2014) rakentuvatkin useista tekijöistä antaen suuremman selitysasteen lukutaidossa esiintyvälle varianssille.

Simple view of reading (SVR) on tutkimuksissa yleisesti käytetty malli, joka muodostuu kahdesta erillisestä, mutta yhdessä toimivasta komponentista: lukusujuvuudesta ja kielellisestä

ymmärtämisestä, joita molempia tarvitaan taitavaan lukemiseen. Tässä mallissa lukusujuvuudella tarkoitetaan kykyä muodostaa näkyvästä kirjoitusmerkistä tai kirjoitusmerkkien sarjasta jokin kielellinen representaatio eli toisin sanoen kykyä tunnistaa sanoja. Kielellinen ymmärtäminen taas on kykyä vastaanottaa semanttista informaatiota sanojen tasolla ja johtaa tästä lause- ja diskurssitulkitintoja.

Lukusujuvuuden (D) ja kielellisen ymmärtämisen (L) suhde luetun ymmärtämiseen (R) voidaan esittää kaavalla $R = D \times L$, joka saa arvon välillä 0-1. Lukusujuvuudella ja kielellisellä ymmärtämisellä on siis keskenään kertautuva vaikutus luetun ymmärtämiseen. Jos esimerkiksi lukusujuvuus on arvoltaan nolla, myös luetun ymmärtäminen saa arvon nolla, vaikka kielellinen ymmärtäminen olisi arvoltaan yksi, eli täydellinen. (Hoover & Goughin, 1990). SVR-mallin vahvuutena on sen suuri selitysaste lukutaidossa (Tilstra & McMaster ym., 2009), mutta heikkoutena kyvyttömyys erotella kaikkia niitä yksittäisiä tekijöitä, jotka lukutaitoon vaikuttavat. Malli ottaa siis huomioon useita tekijöitä, mutta esittää ne kahden yläkäsitteen sisällä.

Luetun ymmärtämisen testaamisen kannalta on olennaista, kuinka hyvin testin pohjalla olevan mallin esittämiä komponentteja voidaan mitata. SVR-mallin tarkoituksena on jakaa luetun ymmärtäminen kahteen eri komponenttiin, mutta se ei tarjoa käytännön keinoja kummankaan komponentin muodostavien osatekijöiden mittaamiselle. Näiden pienempien osatekijöiden mittaamiseksi tarvitaan testin pohjalle yksityiskohtaisempia malleja, joiden avulla osatekijöihin liittyvät prosessit saadaan näkyviin ja mitattaviksi. Tarkkuuden lisääntyessä on kuitenkin otettava huomioon testaamisen resurssi vaatimusten kasvaminen.

Kognitiiviset teoriat

Pearsonin ja Hammin (2005) mukaan psykologian kognitiiviset teoriat yleistyivät 1970-luvulla ja näistä johdetut luetun ymmärtämisen kognitiiviset teoriat alkoivat hiljalleen vallata lukemisen tutkimuskenttää. Aiempi psykometrinen lähestymistapa lukutaidon mittaamisessa korosti behavioraalisiin ja sosiokonstruktivistisiin teorioihin pohjautuvia muuttujia, kuten sanaston laajuutta, lukunopeutta ja muistettavien asioiden määrää, jotka olivat helposti mitattavissa. Kognitiivisten teorioiden yleistymisen myötä keskeisimpänä lukutaidon indikaattorina alettiin nähdä luetun ymmärtämisen taito. Ongelmana oli kuitenkin se, ettei voitu tarkalleen sanoa mitä ymmärtäminen tarkoittaa eikä sitä voitu mitata luotettavasti. Kognitiivisten teorioiden edelleen kehittyessä alettiin saada parempaa kuvaa siitä, mitä ymmärtäminen tarkoittaa ja miten sitä voidaan epäsuorasti mitata.

Luetun ymmärtämisen monimutkaisen luonteen vuoksi yksikään teoria ei tavoita tyhjentävästi kaikkia siihen liittyviä tekijöitä. Siksi kattavan kuvan saamiseksi on syytä tarkastella luetun ymmärtämistä usean teorian kautta. Vaikka teoreettiset mallit toisaalta korostavat eri osa-alueita luetun ymmärtämisestä, niin niitä kaikkia myös yhdistää se, että pohjimmiltaan luetun ymmärtäminen muodostuu tekstistä lukijan muistiin rakennetuista johdonmukaisista skeemoista. (Kendeou, van den Broek, Helder & Karlsson 2014). Seuraavaksi esitellään lyhyesti joitain keskeisimpiä malleja niiden yhteneväisyyksien ja eroavaisuuksien havainnollistamiseksi, lopulta syventyen useita malleja integroivaan luetun ymmärtämisen malliin.

Artikkelissaan Perfetti & Stafura (2014) esittelevät Reading Systems Framework (RSF) mallin, joka on näkemys luetun ymmärtämisen vaatimuksista. Malli mukailee Kintschin (1988) kehittämää *construction integration* -mallia, jonka mukaan luetun ymmärtäminen koostuu ylhäältä alaspäin suuntautuvista ja alhaalta ylöspäin suuntautuvista prosesseista eli tiedostetuista ja tiedostamattomista prosesseista. RSF-malli määrittelee luetun ymmärtämisen vaatimuksiksi lingvististä, ortografista ja yleistä tietoa, jolloin lukija käyttää näitä tietoja joko rajoittavalla tavalla (luetun ymmärtäminen perustuu omaksuttuun tietoon) tai interaktiivisella tavalla, jossa luetun ymmärtäminen perustuu päättelyyn, jossa käytetään aiemman tiedon lisäksi tekstissä esiintyviä merkityksiä. Luetun ymmärtämisen voidaan siis tämän teorian nojalla nähdä vaativan sekä tietoa maailmasta että kognitiivisia prosesseja, joita vaaditaan merkityksen luomiseen luetusta tekstistä.

Minimalistinen malli on McKoonin ja Ratcliffin (1992) kehittämä teoria luetun ymmärtämisestä, jonka mukaan lukija muodostaa lukemisen hetkellä niitä johtopäätöksiä, joita tarvitaan, jotta teksti vaikuttaisi lokaalisti johdonmukaiselta. Lokaalisti tehtävillä johtopäätöksillä tarkoitetaan tekstin sisällä tapahtuvaa tiedonrakentamista, joka ei ota kantaa mahdollisiin ristiriitoihin tekstin ulkopuolisen tiedon kanssa. Luetun ymmärtäminen tarkoittaa siis johdonmukaisuuden luomista tekstistä silloinkin, kun teksti ei mene yhteen aiempien tietojen kanssa.

Current-state selection strategy -malli on Fletcherin ja Bloomin (1988) kehittämä malli, joka on kehitetty Kintschin ja van Dijkn (1978) tekstin ymmärtämismallista. Siinä ajatellaan, että jokaisen luetun virkkeen jälkeen lukijan keskittyminen siirtyy pieneen määrään lausekkeita, joita työstetään lyhytkestoisessa muistissa. Mallissa luetun ymmärtäminen nähdään rakentuvan prosessissa, jossa meneillään oleva lauseke tunnustetaan sen edellisiin lausekkeisiin viittaavien kausaalisten ominaisuuksien mukaan, mutta ei sen mukaan, mitä kausaalisia seurauksia se aiheuttaa myöhemmille lausekkeille. Tekstin pienet rakenneosat yhdistyvät näin vähitellen suuremmaksi rakennelmaksi. Lauseke, jolla ei vielä ole kausaalista linkkiä, on syytä pitää mielessä, kunnes se saa

yhteyden. Näin muistin rooli korostuu varsinkin teksteissä, joissa mielessä pidettäviä lausekkeita on paljon tai niitä on pidettävä siellä pitkään.

Konstruktivistinen malli (Graesser & Kreuz, 1993) kuvaa luetun ymmärtämistä kahdella sitä ohjaavalla oletuksella, joista ensimmäinen on selitysoletus. Siinä lukija pyrkii selittämään miksi erilaisia toimintoja, tapahtumia ja tiloja mainitaan tekstissä. Toinen oletus on johdonmukaisuusoletus, jossa lukija pyrkii rakentamaan merkityksellisen representaation, joka on johdonmukainen niin lokaalisti että globaalisti. Globaalisti johdonmukainen tarkoittaa tässä tekstin liittämistä johdonmukaisesti sen ulkopuolisen kontekstiin. Mallissa lukijan ymmärrys tekstistä pysyy matalalla tasolla, jos teksti on epäjohdonmukaista, lukijalta puuttuu tarvittava aiempi aihetietämys tai jos lukijalla on matalat lukutavoitteet.

Kun tuodaan yhteen eri mallien esittämiä komponentteja luetun ymmärtämiselle, voidaan havaita kielelliseen työmuistiin, aiempaan aihetietämykseen, sanaston laajuuteen, päättelykykyyn, johdonmukaisuuden luomiseen ja motivaatioon liittyviä teemoja. Teemoja painotetaan hieman eri tavoin ja kuvataan erilaisilla käsitteillä.

Luetun ymmärtämistä integroiva malli (eng. Integrated Model of Reading Comprehension, IMREC) on van den Broekin & Espinin (2012) luoma malli, jossa on yhteenkoottuna pääkomponentteja monista luetun ymmärtämisen teorioista ja malleista. Malli korostaa luetun ymmärtämisen moniulotteisuutta ja tarjoaa puitteet luetun ymmärtämisen kehittymisen ja siinä ilmenevien yksilöllisten erojen tarkastelulle. Mallissa luetun ymmärtämisestä erotellaan lopputulema sekä sen muodostumiseen johtanut tapahtumasarja, toisin sanoen tuotos ja prosessi, missä edellinen on jälkimmäisen seuraus.

IMREC-mallissa tuotoksella tarkoitetaan tekstistä luotua mielikuvaa, jonka laatu vaikuttaa siihen, miten hyvin lukija suoriutuu tekstiin liittyvissä tehtävissä, kuten uudelleenkertomisessa tai kysymyksiin vastaamisessa. Näin ollen tuotoksen laatu on käytännössä luetun ymmärtämisen tasoa ilmentävä indikaattori. Taitavan lukijan luomassa mielikuvassa esimerkiksi painottuvat enemmän tekstin keskeiset elementit eli sellaiset elementit joilla on useita yhteyksiä tekstin muihin elementteihin, kun taas vähemmän taitavilla lukijoilla tällaista valikointia tapahtuu vähemmän (van den Broek, Helder, & van Leijen-Horst, 2012). Havainto on yhdenmukainen myös muissa aiemmin esitetyissä malleissa havaitun johdonmukaisen skeemojen keskeisen roolin kanssa. Vaikka tuotos ja sen johdonmukaisuus itsessään osoittavat luetun ymmärtämisen tason, on diagnosoinnin ja sopivien tukitoimien suunnittelun kannalta tärkeää tarkastella myös tuotoksen muodostumisen taustalla olevia prosesseja (Perfetti & Adlof, 2012).

Jokainen uusi luettu lause aktivoi aiempaa aihetietämystä sekä siihen liittyviä konsepteja, joita lukija muodostaa prosessoidessaan tekstin aiempia lauseita (O'Brien & Albrecht, 1992). Van den

Broekin & Espinin (2012) mukaan prosessoidessaan tekstiä lukija rakentaa sen sisäisistä elementeistä yhteneväisen kokonaiskuvan, jossa sanat, lauseet ja kappaleet ovat tarkoituksenmukaisesti suhteutettu toisiinsa teksti-informaation sisältämien viittausten ja kausaalisuhteiden sekä kielioppisääntöjen perusteella. Kukin tekstin elementti saa identiteetin, joka syntyy, todentuu ja muuttuu suhteessa muihin identiteetteihin prosessoinnin aikana. Jotta johdonmukainen mielikuva tekstistä olisi mahdollinen, täytyy lukijan todeta tekstin identiteetit nojaamalla esimerkiksi aiempaan aihetietoon, kielen syntaksiin tai kontekstiin, jossa teksti esiintyy. Esimerkiksi lukiessaan arvostelevalta tekstistä sanaparin *antiikin aikainen* on lukijan kyettävä suhteuttamaan ilmaus kontekstiin ja aiempaan aihetietoonsa historiasta sekä ymmärtämään mihin tekstin elementteihin se viittaa: tekstin osa viittaa käsiteltävään aiheeseen liioitellusti liittämällä kritisoitavan kohteen yleisesti käytettyyn ilmaukseen menneestä aikakaudesta, mikä kertoo arvostellun kohteen olevan ajastaan jäljessä. Tähän tapaan tuotettu tulkinta tekstistä, joka perustuu tekstistä saatavalle tiedolle ja aiempaan tietoon aiheesta tunnetaan yleisemmin tilannemallin rakentamisena (Kintsch, 1988).

Van den Broekin & Espinin (2012) mukaan lukijan pyrkiessä luomaan tekstistä johdonmukaista mielikuvaa, hän tasapainottelee tämän pyrkimyksen vaatimien automaattisten sekä strategisten prosessien resurssivaatimusten ja tavoiteltavan johdonmukaisuuden tason välillä. Kunkin prosessin kognitiivinen resurssivaatimus on yhteydessä kyseisen prosessin automatisaation tasoon. Mitä enemmän jokin prosessi automatisoituu, sitä enemmän lukijalle jää resursseja vaativampien strategisten prosessien käyttämiselle. Täysin automatisoituneet prosessit käynnistyvät ilman lukijan tietoista kontrollia. Esimerkiksi sanojen dekadaus ja kirjainten tunnistus ovat taitavalla lukijalla pitkälti automatisoituneita prosesseja, jolloin näiden prosessien aiheuttama vähäinen kognitiivinen kuormitus antaa tilaa strategisille prosesseille, kuten keskeisten elementtien tunnistamiselle.

Muita strategisia prosesseja ovat esimerkiksi tekstin aiemman osion uudelleenlukeminen oleellisen tiedon hakemiseksi sekä oman etukäteistiedon tutkiminen edessä olevan tekstin merkitysten selittäjäksi. Niihin kuuluvat myös toiminnot, jotka liittyvät tekstin rakenteellisiin poikkeavuuksiin, kuten kursiivin tulkitsemiseen (Lemarié, Lorch, Eyrolle & Virbel, 2008). Toiminnot ovat opeteltavissa ja ne muodostavat lukijan yksilöllisen strategisten prosessien kokoelman, joka kuvaa keskeisesti luetun ymmärtämisen kehittyneisyyttä (LaBerge & Samuels, 1974). Tietoisesti muodostetut vaativimmatkin strategiset prosessit voivat muuttua harjoittelun myötä lähes automaattisiksi, jolloin vapautuneet kognitiiviset resurssit mahdollistavat yhä edistyneempien strategioiden soveltamisen (van den Broek & Espin, 2012).

Van den Broekin ja Espinin (2012) mukaan jokaisella lukijalla on tyypiltään ja voimakkuudeltaan vaihtelevat johdonmukaisuuden standardit (*eng. standards of coherence*). Ne määrittävät sitä semanttisen johdonmukaisuuden tasoa, johon luettaessa pyritään (van den Broek, Bohn-Gettler, Kendeou, Carlson & White, 2011; van den Broek, Risdén & Husebye-Hartman, 1995). Standardit ovat dynaamisesti yhteydessä lukutilanteeseen eli ne ohjaavat lukijaa tekstin edetessä, mutta voivat myös muuttua tekstistä saadun uuden informaation myötä. Johdonmukaisuuden standardien tyypeiksi on ehdotettu muun muassa referentiaalista, kausaalista, spatiaalista, temporaalista ja loogista standardia (Graesse, Singer & Trabasso, 1994). Lukijalla on yksilöllinen näkemys siitä, millä tasolla teksti-informaatiosta luodun mielikuvan tulee olla suhteessa näihin standardityyppeihin.

Strategisten prosessien käynnistyminen riippuu siitä informaatiosta, joka muodostuu automaattisissa prosesseissa ja siitä, millaiset johdonmukaisuuden standardit lukija omaksuu (van den Broek & Espin, 2012). Jos automaattisilla prosesseilla saavutettu mielikuva tekstistä vastaa lukijan johdonmukaisuuden standardeja, strategisia prosesseja ei oteta käyttöön (van den Broek & Espin, 2012). Esimerkiksi luettaessa uutisia hallitusneuvottelusta, voidaan automaattisilla prosesseilla saavuttaa yleinen käsitys neuvottelun osapuolista ja keskeisestä hallitusohjelmasta. Jos tämä saavutettu mielikuva on lukijan standardien mukainen, lukijan ei tarvitse käyttää strategisia prosesseja. Toisaalta mielenkiintoinen lause tai sana voi herättää lukijan halun saada yksityiskohtaisemman kuvan osapuolten välisistä suhteista, jolloin standardit saattavat muuttua ja vaatia strategisten prosessien käyttöönottoa missä tahansa lukemisen vaiheessa.

Yhteenvetona luetun ymmärtäminen nähdään IMREC-mallissa automaattisten ja strategisten prosessien kautta syntyvänä, johdonmukaisuudeltaan lukijan standardeja vastaavana tuotoksena, jossa luetun ymmärtämisen tasoa voidaan tarkastella sekä tuotoksen laadun että käytettyjen prosessien kautta.

Luetun ymmärtämisen kognitiiviset taustatekijät

Kognitiivisissa malleissa esitettyihin prosesseihin liittyy taustatekijöitä, joilla on jokaisella yksilöllinen vaikutus luetun ymmärtämiseen. Esimerkkinä kognitiivisista taustatekijöistä on muisti, jota tarvitaan teksteissä esiintyvien keskeisten elementtien mielessä pitämiseen. Lisäksi tarvitaan päättelykykyä eri tekstielementtien välisten suhteiden ymmärtämiseen. Kendeoun, van den Broekin, Helderin ja Karlssonin (2014) mukaan ymmärtääkseen lukemaansa, lukijan tulee prosessoida visuaalinen ärsyke (painetut sanat), tunnistaa ja aktivoida hänen fonologiset, ortografiset ja semanttiset representaatiot sekä liittää nämä representaatiot yhteen luodakseen yhtenäisen kuvan

lukemastaan. Jotta lukija ymmärtäisi tekstin kokonaisuudessaan, on hänen prosessoitava ja liitettävä yksittäiset ideat yhteen luodakseen koherentin mentaalisen representaation koko tekstistä (Kendeou ym., 2014). Luetun ymmärtämisen tiedetäänkin olevan vahvasti yhteydessä muun muassa lukijan työmuistiin, päättelykykyyn, lukusujuvuuteen ja aiempaan tietoon (Kendeou ym., 2014). Vaikka jokaisella näistä tekijöistä on oma vaikutuksensa luetun ymmärtämiseen, niitä ei voi tarkastella toisistaan erillisinä vaikuttajina. Esimerkiksi sanaston laajuuden on huomattu helpottavan päättelyä varten tarvittavien representaatioiden valintaa ja työmuistin puolestaan helpottavan valittujen representaatioiden integrointia lukijan tilannemalliin (Calvo, 2005).

Hannonin ja Danemanin (2001) tutkimuksessa työmuistilla voitiin selittää 21 % luetun ymmärtämisen vaihtelusta yliopisto-opiskelijoilla. Cainin, Oakhillin & Bryantin (2004) tutkimuksessa työmuistin laajuus kertoi yksilöllisestä vaihtelusta luetun ymmärtämisessä 8–11-vuotiailla sen jälkeen, kun yksittäisten sanojen lukutaito sekä verbaalinen ja kielellinen taito oli otettu huomioon. Nation, Adams, Bowyer-Crane ja Snowling (1999) esittivät, että työmuistin ja luetun ymmärtämisen yhteyden taustalla ovat verbaaliset ja semanttiset taidot. Heidän mukaan heikoilla lukijoilla on tiettyjä semanttisia heikkouksia, jotka rajoittavat kykyä varastoida verbaalista informaatiota lyhytkestoisessa muistissa. Cainin ym. (2004) esittämän luetun ymmärtämistä ennustavan mallin mukaan työmuisti nähdään resurssina, joka mahdollistaa monen sellaisen prosessin käyttämisen, jota tekstistä luotavan representaation rakentamiseen tarvitaan. Prosessit jaetaan alemman tason prosesseihin, kuten sanantunnistukseen sekä ylemmän tason prosesseihin, kuten sanoihin liittyvien päätelmien tekemiseen. Tällöin esimerkiksi sanantunnistuksen ollessa hidasta, jää vähemmän työmuistin resursseja korkeamman tason prosessoinnille, mikä voi näkyä heikompana luetun ymmärtämisenä. Vaikka työmuistin toiminta on oleellinen taustatekijä monissa prosesseissa, prosesseilla ja työmuistilla on myös oma vaikutuksensa, kun niiden yhteisvaikutus kontrolloidaan (Seigneuric, Ehrlich, Oakhill & Yuill, 2000).

Päätelykyky on oleellinen osa luetun ymmärtämistä, jos lukija tarvitsee teksti-informaation lisäksi tekstin ulkopuolista tietoa luodakseen tekstistä paikallisesti ja globaalisti johdonmukaisen (Graesser ym., 1994; Long & Chong, 2001). Päätelmien avulla lukijat luovat merkityksellisiä yhteyksiä tekstin elementtien ja oleellisen taustatiedon välillä (Oakhill, Cain, & Bryant, 2003; van den Broek, 1990). Päätelyn avulla siis täytetään ne tekstin aukot, jotka lukijan on täytettävä vastatakseen johdonmukaisuuden standardeihinsa (van den Broekin & Espinin, 2012). Oakhill (1984) tutki miten päätelykyky toimii, jos haettu informaatio mainitaan luettavassa tekstissä vain implisiittisesti. Tutkimuksessa lapset jaettiin kahteen ryhmään. Ensimmäisessä ryhmässä oli lapsia, joiden luetun ymmärtäminen oli sujuvaa ja toisessa ryhmässä lapsia, joilla oli todettu spesifi luetun ymmärtämisen vaikeus. Tutkimuksessa heidän tuli lukea tekstejä ja vastata niihin liittyviin

kysymyksiin. Joihinkin kysymyksiin oli tekstissä suora vastaus, kun taas toisissa vastaus vaati päättelykykyä, koska vastaajan oli käytettävä yleistietoa tekstin tiedon kanssa puuttuvan informaation saamiseksi. Esimerkiksi yhdessä tekstissä päähenkilö pyöräilee kouluun, mutta pyörää ei mainita missään vaiheessa suoraan, vaan se implikoidaan avainlauseessa “Hän polki niin lujaa kuin pystyi.”. Siitä, kuinka lapsi vastasi kysymykseen “Miten päähenkilö meni kouluun?”, kyettiin arvioimaan hänen päättelykykyään. Sujuvat lukijat vastasivat useammin oikein kysymyksiin, jotka vaativat päättelyä, mutta lapset, joilla oli S-RCD vastasivat oikein lähinnä kysymyksiin, joihin oli suoraan tekstistä saatava vastaus. Kummankin ryhmän tulokset paranivat, kun heille annettiin mahdollisuus lukea teksti uudestaan, mutta suhteelliset erot päättelykyvyssä säilyivät silti ryhmien välillä. Oakhill (1984) päätteli tästä, että päättelykyky on olennainen erottava tekijä taitavien ja heikkojen luetun ymmärtäjien välillä, ja ettei muistamisella ole roolia päättelykyvyssä.

Luetun ymmärtämisen ja lukusujuvuuden yhteyttä voidaan pyrkiä selittämään niiden jakamalla semanttisilla ja lauseopillisilla prosesseilla, joita käytetään kielen prosessointiin lause- tai virketasolla. (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003). Eri tutkimuksissa lukusujuvuus on jaettu erilaisiin komponentteihin. Jaottelusta riippuen lukusujuvuuden on havaittu olevan yhteydessä luetun ymmärtämiseen hieman eri tavoin. Klaudan & Guthrien (2008) mukaan yhteyden selittäjäksi on muodostunut kaksi vallitsevaa teoreettista suuntausta, joista ensimmäinen keskittyy lukusujuvuuteen sanatasolla ja toinen lauseopillisella tasolla. Automatiikan (LaBerge & Samuels, 1974) ja verbaalisen tehokkuuden teorian (Perfetti, 1985) mukaan etenkin lukusujuvuuden nopeuskomponentin harjaantuminen helpottaa luetun ymmärtämistä.

Klaudan & Guthrien (2008) tutkimuksessa tarkasteltiin lukusujuvuuden kolmen komponentin yhteyttä luetun ymmärtämiseen viidesluokkalaisilla. Komponentteja olivat yksittäiset sanat, lauseopilliset yksiköt ja kokonaiset kappaleet. Luetun ymmärtämistä selittävässä hierarkkisessa regressiomallissa aiempi tieto selitti vaihtelusta 50 %, päättelykyky 8 %, sanojen lukunopeus 10 %, lauseopillinen prosessointi 5 % ja kappaletason prosessointi 2 %. Lopullisessa mallissa jokainen muuttuja ennusti merkitsevästi luetun ymmärtämistä ja yhdessä muuttujat selittivät 75 % luetun ymmärtämisen taidon vaihtelusta. Meta-analysissaan García & Cain (2014) selvittivät lukusujuvuuden ja luetun ymmärtämisen yhteyttä eri ikäisillä englanninkielisillä lukijoilla. Heidän mukaansa ne olivat vahvasti yhteydessä toisiinsa kaikilla ikätasoilla. He huomasivat myös, että tämä yhteys välittyi ainakin iällä, kuullun ymmärtämisen tasolla, lukusujuvuuden arvioinnin testin luonteella, ymmärrettävän tekstin lajilla, avun saannista lukusujuvuuden kanssa ja sillä, luettiinko teksti ääneen vai ei. Interventioissa, joissa keskitytään lukusujuvuuden kehittämiseen, saavutetaan yleensä merkittävää taitojen parantumista niin lukusujuvuudessa kuin luetun ymmärtämisessä

aloittelevien lukijoiden ja lukemisen kanssa kamppailevien lukiolaisten kanssa (Chard, Vaughn, & Tyler, 2002; Kuhn & Stahl, 2003)

Lukijan aiemman tiedon vaikutus luetun ymmärtämiseen näkyy muun muassa sanaston laajuudessa sekä lukijan tiedossa eri tekstilajeista. Rickettsin, Nationin ja Bishopin (2007) pitkittäistutkimuksessa tarkasteltiin sanaston laajuuden merkitystä luetun ymmärtämisessä heikoilla ja taitavilla lukijoilla. Korrelaatio luetun ymmärtämisen ja sanaston laajuuden välillä oli .63, missä sanaston laajuus selitti merkittävän osan luetun ymmärtämisen vaihtelusta, kun kronologinen ikä, yksittäisten sanojen lukeminen ja nonverbaalinen päättelykyky oli huomioitu. Cainin ym. (2004) tutkimuksessa lukijan tieto tekstilajeista oli positiivisesti yhteydessä luetun ymmärtämisen tasoon senkin jälkeen, kun toiset kielellisen taidon mittarit oli otettu huomioon.

Tutkimuksissa eri komponenttien vaikutuksen luetun ymmärtämiseen on havaittu muuttuvan iän myötä. Esimerkiksi yksittäisten sanojen lukeminen on tärkein selittäjä luetun ymmärtämiselle aloittelevilla lukijoilla (Juel, Griffith, & Gough, 1986), mutta myöhemmissä vaiheissa muista komponenteista tulee tärkeämpiä selittäjiä. Onkin huomioitava, että eri komponentit luetun ymmärtämiselle korostuvat eri ikävaiheissa.

Luetun ymmärtämisen mittaaminen

Jotta luetun ymmärtämistä voidaan mitata luotettavasti, mittari täytyy rakentaa pohjautuen johonkin luetun ymmärtämisen teoriaan. Kuten aiemmin todettiin, ei ole teoriaa, jota käytettäisiin yksimielisesti selittämään luetun ymmärtämistä. On vain teorioita, jotka kattavat eri alueita luetun ymmärtämisestä ja painottavat siihen vaikuttavia muuttujia eri painotuksin. Katettujen alueiden käsitteistä saadaan mitattavia kohteita, joista voidaan koostaa luetun ymmärtämisen tasosta kertova mittari. Useiden teorioiden myötä on siis luonnollisesti syntynyt useita eri mittaustapoja. Siksi mittareita ja mittaustuloksia tarkasteltaessa on huomioitava niiden pohjalla oleva teoria, jotta ymmärretään mitä tarkalleen ottaen mitataan. Kahden viime vuosikymmenen teorit luetun ymmärtämisestä ovat huomattavan yhtenäisiä pääkomponenteissaan (Sabatini, Albro & O'Reilly, 2012). Ongelmana on kuitenkin se, etteivät eri teorit päädy samanlaisiin luetun ymmärtämisen arviointeihin. Esimerkiksi Rimrodtin, Lightmanin, Robertsin, Dencklan, ja Cuttingin (2005) tutkimuksessa vertailtiin kolmen luetun ymmärtämisen testin tuloksia ja havaittiin, että 43.5 %:lla tutkittavista oli ainakin yhden testin mukaan luetun ymmärtämisen vaikeuksia, mutta vain 9.4 %:lla oli yksimielisesti kaikkien testien mukaan vaikeuksia. Myös Keenan, Betjemann & Olson (2008) vertailivat tutkimuksessaan Yhdysvaltojen suosituimpia luetun ymmärtämisen testejä keskenään. He päättelivät, että kohtalaiset interkorrelaatiot testien välillä kertoivat testien mittaavan eri luetun

ymmärtämisen taitoja. Lisäksi Francisin, Snowin, Augustin, Carlsonin, Millerin ja Iglesiaksen (2006) mukaan yksidimensionaaliset pyrkimykset luetun ymmärtämisen arviointiin johtavat väistämättä epätarkkuuteen luetun ymmärtämistä selittäessä.

Oma vaikutuksensa mittarin luotettavuuteen ja vertailtavuuteen on sillä, minkälainen lähtökohta mittarin validoimiseksi on otettu. Konstruktiovaliditeetti liittyy abstraktien käsitteiden ja piirteiden mittaamiseen, missä piirteet tulee operationalisoida ennen mittausta. Näiden käsitteiden tai piirteiden ominaisuudet tulee avata validiteetin osoittamiseksi (Gray, 2013). Koska luetun ymmärtäminen on abstrakti käsite, johon liittyy moniulotteisia ja latenteja muuttujia, konstruktiovaliditeetin käyttö on parhaiten sopiva lähtökohta mittarin validoimiseksi.

Konstruktiovaliditeetin konsepti on ajan saatossa kuitenkin muuttunut. Alkuperäisen konseptin mukaisesti validin mittarin kehittäminen vaatii yleisesti tuetun, tarkasti tutkitun teorian, josta muodostetaan teorian mukaiset operationalisoidut, mitattavat muuttujat (van den Broek & Espin, 2012). Luetun ymmärtämisen mittaamisen tapauksessa tämä tarkoittaisi, että kaikista toisiaan täydentävistä teorioista olisi johdettava kokoelma mitattavia muuttujia, jotka yhdessä selittäisivät luetun ymmärtämistä kauttaaltaan. Tällä tavalla validoitava mittari olisi van den Broekin & Espinin (2012) mukaan kuitenkin epäkäytännöllinen eikä mukailisi nykyisiä käsityksiä konstruktiovaliditeetista, joiden mukaan validiteetti tulisi osoittaa muodostamalla hypoteeseja mittarista saatujen tulosten merkityksestä käytännön tasolla. Silloin osoitetaan sekä mittarin että sen taustalla olevan teorian validiteetti. Esimerkiksi jokin mittari voidaan alkuperäisen käsityksen mukaan nähdä validina, jos sen tulokset osoittavat teoriasta johdettujen hypoteesien mukaiset yhteydet mitattavaan ominaisuuteen, kun taas nykyisen käsityksen mukaan validi mittari vaatii, että mitatuista tuloksista tehdyt johtopäätökset ja niiden mukaiset käytännön toimenpiteet vaikuttavat teoriasta johdetun hypoteesin mukaisesti tutkittavaan ilmiöön. Konstruktiovaliditeetin muutos tuottaa haasteita eri aikakausilla saatujen tutkimustulosten vertailtavuuteen.

Mittarin validiteetin, taustateorian sekä lukijan ominaisuuksien lisäksi Fletcherin (2006) mukaan luettavan tekstin ominaisuudet vaikuttavat siihen, millaisia päätelmiä luetun ymmärtämisestä voidaan tehdä. Jos luettava teksti on ikätasoon nähden kielellisesti liian vaikea tai helppo, ei voida tehdä luotettavia päätelmiä luetun ymmärtämisestä. Tekstin ominaisuudet vaikuttavat myös siihen, minkä laatuiseen ongelmaan ymmärtämisvaikeuden voidaan ajatella liittyvän. Esimerkiksi vaikeaa sanastoa sisältävässä tekstissä luetun ymmärtämiseen saattaa vaikuttaa enemmän lukijan sanaston laajuus kuin lukusujuvuus. Diagnostisessa mielessä on siis syytä ottaa huomioon luettavan tekstin rakenne ja vaikeustaso, jotta saadaan tarkempi kuva ymmärtämisen vaikeuden taustalla olevista syistä. Deane ym. (2006) osoittivat, että koulussa kolmannella ja kuudennella luokalla luettavien tekstien eroavaisuuksia voitiin ilmentää tekstin

ominaisuuksia kuvaavalla yhdeksän faktorin mallilla, joka selitti 63 % tekstien vaihtelusta. Eroavaisuuksia kuvaavia faktoreita olivat esimerkiksi tekstissä esiintyvän narratiivisen diskurssin määrä sekä kausaalisten sanojen määrä. Faktorimalli tarjoaa viitekehyksen sille, mitä tekijöitä lukutehtäviä muodostettaessa voisi ottaa huomioon, jotta tekstin ominaisuuksien vaikutus saataisiin kontrolloitua.

Amendumin, Conradin & Hiebertin (2018) meta-analyysi tekstin haastavuuden yhteydestä lukusujuvuuteen ja luetun ymmärtämiseen peruskouluikäisillä osoitti, että tekstin vaikeuden lisääntyessä oppilaiden lukemisen tarkkuus ja lukunopeus vähenivät etenkin vähemmän taitavilla lukijoilla. Myös luetun ymmärtäminen oli tarkasteltavasta tutkimuksesta riippuen yhteydessä tekstin haasteellisuuteen joko negatiivisessa suhteessa tai ei-merkittävästi. Onkin siis oleellista ottaa huomioon tutkittavien taitotaso suhteessa tekstin haastavuuteen, jotta tutkimuksessa saatavat tulokset heijastelisivat todenmukaisesti eri tekijöiden vaikutusta lukemisen taitoihin. Jokin tekstin ominaisuus saattaa ylikorostaa yksittäisten muuttujien vaikutusta tarkasteltaessa lukemisen taitoja, jolloin tutkittavasta ilmiöstä ei muodostu yhtä tarkkaa kuvaa kuin lukijan taitotasoon suhteutetuilla testeillä.

IMREC -mallin tekijöiden mukaan luetun ymmärtämisen testissä täytyy huomioida lukemisen moniulotteisuus ja mittaamisen tarkoitus. Mittaamisen tarkoituksena voi olla seulonta, kehityksen seuranta sekä vahvuuksien ja heikkouksien tunnistaminen. He ehdottavat teoreettisessa viitekehysessään luetun ymmärtämisen testiä muodostaessa huomioitavaksi taulukko 1:ssä esitetyt faktorit, jotka on todettu aiemmissa tutkimuksissa merkittäviksi tekijöiksi.

TAULUKKO 1. Luetunymmärtämistestin rakentamisessa huomioitavat faktorit (IMREC)

Faktorit

- a) Kielellisen työmuistin kapasiteetti
 - b) Aiempi aihetietämys
 - c) Päättelykyky
 - d) Tietämys johdonmukaisuuden standardeista ja niiden soveltaminen
 - e) Huomion kohdistaminen keskeiseen tietoon
 - f) Kielelliset ja lukemiseen liittyvät perustaidot (esim. dekkoodaus)
 - g) Sanaston laajuus
 - h) Lukemisen aikainen ja jälkeinen sensitiivisyys rakenteelliseen keskiöön
 - i) Tieto tekstilajeista
 - j) Tekstikohtainen koherenssin ja johtopäätelmien tuottamisen taito
 - k) Motivaatio lukemiseen
-

Cloze-testi

Cloze-testi on luetun ymmärtämisen testityyppi, jossa on cloze-proseduurin mukaisesti käsiteltyjä tekstejä. Cloze-proseduuri on tekniikka, jossa luettavaan tekstiin tehdään aukkoja poistamalla systemaattisesti sanoja ennalta määrättyjä sääntöjä noudattaen. Aukkoja sisältävän tekstin ymmärtäminen asettaa tiettyjä vaatimuksia tietyille luetun ymmärtämisen komponenteille riippuen siitä, miten aukko on valittu (Gellert & Elbro, 2013).

Alun perin aukkoja muodostettiin vakioidun sanavälin säännöllä, jossa tekstistä poistettiin joka n:nes sana, yleensä joka viides. Tällä tavoin luodun testin huomattiin korreloivan melko hyvin muiden luetun ymmärtämisen testien kanssa (Bormuth, 1967), mutta korrelaation katsottiin syntyvän lähinnä dekodauksen vaikutuksesta. Dekoodaus on merkittävä tekijä luetun ymmärtämisessä, joten kohtuullinen korrelaatio laajemmille luetun ymmärtämisen testeille oli ymmärrettävää. Testillä ei kuitenkaan päästy käsiksi luetun ymmärtämiseen sen laajemmassa merkityksessä (Alderson, 1979).

Aukkojen muodostamiseen käytetty sääntö vaikuttaa siis siihen, mitä osia luetun ymmärtämisestä kyetään mittaamaan. Bachmanin (1985) mukaan valtaosassa clozen validiteettitutkimuksista cloze-testit on muodostettu poistamalla sanoja vakioituilla väleillä, eikä tutkimuksissa ole tehty erottelua eri säännöillä muodostetuista testeistä, mikä voi olla syy tutkimustulosten epä johdonmukaisuudelle. Bachman teki yhden tekstin pohjalta kaksi cloze-testiä, joista ensimmäisessä oli vakioidusti ja toisessa harkintaa käyttäen luodut aukot. Hän lajitteli aukkojen täydentämiseen vaaditut tiedot neljään eri tyyppiin: Lauseen, virkkeen ja tekstin sisäiseen sekä tekstin ulkopuoliseen tietoon. Suurin osa vakioitujen aukkojen cloze-testissä olivat joko lauseen sisäistä tai tekstin ulkopuolista tietoa vaativia aukkoja, mutta vain muutamat virkkeen ja tekstin sisäistä tietoa vaativia aukkoja. Toisaalta harkittujen aukkojen cloze-testissä testin tekijä voi päättää sisällyttää enemmän myös virkkeen ja tekstin sisäistä tietoa vaativia aukkoja.

Gellert ja Elbro (2013) esittävät, että aukot voidaan laatia enemmän ymmärtämistä vaativiksi esimerkiksi huomioimalla pronominaaliset viittaukset, leksikaaliset yhtenäisyydet ja konjunktiot, jotka ovat yhtenäisyyttä luovia sidoksia. Nämä sidokset tuovat tekstiin johdonmukaisuutta läpi lauseiden. Cloze-testi, jossa sidoksisuus huomioitiin, selitti 16 % enemmän testin varianssista kuin cloze-testi, jossa sidoksisuutta ei huomioitu, vaikka dekodauksen ja sanaston vaikutus kontrolloitiin (Gellert & Elbro, 2013).

Halliday and Hasan kuvailevat (1976) kirjassaan tekstin yhtenäisyyden vaatimusta seuraavasti: Tekstissä olevan aukon täyttö on ehtona tekstin yhtenäisyydelle, jos yhden tekstielementin tulkinta vaatii täytettävän tekstielementin olemassaoloa samassa tekstissä. Tämän

yhtenäisyyden ehtona on suhde, jossa ilmenee sekä viitattu osa ja osa, joka toimii viittaajana. Esimerkiksi virkkeessä, jossa ensimmäinen sana on monivalintainen aukko “_____ oli karitsa, josta *hän* piti kovasti” (Koululla/*Maijalla*/Tulppuloilla/Pihassa), poistettuun sanaan viitataan seuraavassa lauseessa persoonapronominilla, joka ei muuten viittaisi mihinkään. Tällöin voidaan päätellä persoonapronominin edellyttävän aukon kohdalle sanaa, johon se viittaa kieliopillisesti oikein. Jos tällaista sidoksisuutta sisältävä sana on poistettu tekstistä, lukijan täytyy osata hakea ratkaisuvihjeitä aukon sisältävän lauseen ulkopuolelta. On siis perusteltua, että harkitusti laaditulla cloze-testillä voidaan päästä käsiksi myös luetun ymmärtämisen syvempiin prosesseihin.

Levenstonin, Nirin and Blum-Kulkan (1984) mukaan tekstin makrorakenteen uudelleenrakentamiskykyä mittaavalla testillä päästään käsiksi lukemisen ylempiin taitoihin. Van Dijk & Kintschin (1983) määrittelemänä makrorakenne edustaa johdonmukaiseksi kokonaisuudeksi organisoituneen tekstin ydintä. Se ohjaa koodausta, palauttamista ja olennaisten tekstin kohtien uudelleentuottamista. Makrorakenteiden muodostaminen on näin edellytys kokonaisvaltaisen ymmärryksen ja merkityksellisen oppimisen onnistumiselle. Levenstonin ym. (1984) mukaan makrorakenteen uudelleenrakentamiskykyä vaativan cloze-testin sisältämistä aukkoista suurin osa vaatii lauseen ulkopuolisista vihjeistä tehtäviä johtopäätöksiä. Tämän tapaista cloze-testiä on kutsuttu yleisesti diskurssi-clozeksi.

Toisaalta, jos tarkastelee Widdowsonin (1979) määrittelemää diskurssi- ja teksti-clozea, Levenstonin diskurssi-clozen määritelmä näyttää täyttävän vain teksti-clozen kriteerit, jotka huomioivat ainoastaan tekstin sisäisen johdonmukaisuuden. Widdowsonin (1979) kriteeristöä mukailleen ero teksti-clozen ja diskurssi-clozen välillä näkyy siinä, että ensimmäinen mittaa kielitaitoa ja jälkimmäinen vaatii sen lisäksi sitä, että lukija kykenee ymmärtämään kommunikaatiota yhtenä kokonaisuutena, jossa käy ilmi jaetut oletukset tekstin kirjoittajan kanssa, kirjoittajan intentioiden ymmärtäminen ja kaikki ne tekstin ymmärtämisen prosessit, jotka liittyvät tekstin tulkitsemiseen. Diskurssi-cloze, joka reflektoi lukijan taitoja seurata informaatiota tekstin läpi ja taitoja käyttää informaatiota niin kontekstuaalisista että kielellisen ympäristön vihjeistä, käyttää siis yksikköinä teemaa ja reemaa, jotka tarjoavat kriteerit tällaisille aukkovalinnoille (Deyes, 1984). Tästä näkökulmasta katsottuna sidossanojen harkittu poistaminen ei takaa sitä, että tehtävä vaatii luetun ymmärtämistä, vaan ennemminkin kielioppisääntöjen osaamista.

Cloze-testi on siis kehittynyt vahvasti dekodeesta mittaavasta testistä myös ylempiä lukemisen taitoja mittaavaksi testiksi, mutta vielääkään ei ole konsensusta siitä, mitä nämä ylempät taidot ovat ja kuinka paljon kielellistä päättelyä lukijalta vaaditaan onnistuakseen tällaisissa cloze-testeissä.

Clozen kehittyneemmissäkin muodoissa on vajavaisuuksia, joihin pyritään edelleen vastaamaan. Harkittuun aukkojen valintaan liittyy esimerkiksi riski tutkijan aiheuttamalle vinoumalle valituissa sanoissa. Kleijnin, Pander Maatin ja Sandersin (2019) nykyaikaisessa Hybrid Text Comprehension clozessa eli HyTeC-clozessa aukkovaihtoehdot valitaan kahden kriteerin, arvattavuuden ja arvaamattomuuden perusteella. Ensinnäkään aukkojen vastaukset eivät saa olla liian helposti pääteltävissä paikallisesta informaatiosta kielen käytön yleisiä konventioita käyttäen. Toiseksi aukkojen vastaukset eivät saa olla sanoja, jotka voidaan päätellä vain ekstratekstuaalista tietoa käyttäen eli ilman minkäänlaisia kontekstuaalisia vihjeitä. Kun näitä kriteerejä noudattaen on karsittu aukoiksi kelpaamattomat sanat, niin jäljelle jääneestä sanojen joukosta poistetaan mekaanisesti sen verran sanoja, että lopullisessa versiossa joka kymmenes sana on muutettu aukoksi. Kun käytetään harkinnan lisäksi mekaanista poistotapaa, vähennetään harkintaan liittyvän vinouman vaikutusta.

Yleisin tapa pisteyttää cloze-tehtäviä on antaa piste vain täsmälleen oikeasta vastauksesta, mutta esimerkiksi testattaessa englantia vieraana kielenä, käytetään myös väljempää pisteytystä, jossa synonyymista tai semanttisesti johdonmukaisesta sanasta aukon täyttämiseksi annetaan piste (Alderson, 1979). Pisteyttäessä täsmälleen oikeasta vastauksesta vähennetään tarkastusurakkaa, mutta haittana saattaa olla se, että täsmällisen vastauksen vaatiminen saattaa antaa vääristyneen kuvan joidenkin testattavien luetun ymmärtämisen taidoista. Semanttisella pisteytyksellä on ollut korkeampi näennäisvaliditeetti (Kleijn ym., 2019). Silti useat tutkijat käyttävät eksaktia pisteytystä helppouden ja sen vuoksi, että sillä on suuri korrelaatio semanttiseen pisteytystapaan ($r = .9$; Alderson, 1979).

Myös vastaustavoissa eri cloze-testien välillä löytyy eroavaisuuksia. Pike tutki (1979) englantia vieraana kielenä testattavia (*TOEFL examinees*) perulaisia, chileläisiä ja japanilaisia, joita testattiin monivalinta-cloze-testillä ja vapaatäyttöisellä cloze-testillä. Näiden väliset korrelaatiot vaihtelivat kansalaisuudesta riippuen välillä .70 - .89 ja kalibroitu korrelaatio välillä .86 - .99. Myös Hinofotis & Snow (1980) havaitsivat, että englantia vieraana kielenä -tasokokeella oli monivalinta-cloze-testin kanssa .63 ja vapaatäyttöisen cloze-testin kanssa .71 korrelaatio, mistä tutkijat päättelivät, että nämä kaksi clozea tuottavat samankaltaista informaatiota. Edelleen Brown (1980) englantia vieraana kielenä -tasokokeen korreloivan monivalinta-cloze-testin kanssa .89 ja .88 (täsmällisistä vastauksista) tai .90 (hyväksytyistä vastauksista) vapaatäyttöisen cloze-testin kanssa. Valitulla vastaustavalla ei siis näyttäisi olevan merkittävää vaikutusta testitulokseen.

Clozea on kritisoitu paljon, mutta suurin osa kritiikistä on kohdistunut perinteisellä cloze-proseduurilla laadittuihin testeihin. HyTeC on pyrkinyt vastaamaan kritiikkiin muuttamalla kritiikin kohteena olleita proseduurin vaiheita (taulukko 2). Toisaalta clozessa on myös selkeitä etuja

suhteessa muihin testausmenetelmiin (Kleijn ym., 2019). Cloze-testi on helppo rakentaa, ja sitä voidaan käyttää useissa eri tekstityypeissä. Tekstin vaikeustasoa voidaan tarkastella systemaattisesti, koska kaikille teksteille tehdään sama käsittely, jolloin aukkojen täyttämisen vaikeus määräytyy tekstin vaikeuden mukaan (Klare, 1976; Miller & Coleman, 1967). Lisäksi proseduurilla vältetään esimerkiksi tekstin jälkeisiin avoimiin kysymyksiin liittyvältä subjektiivisuusongelmalta, jossa tulkintaa tapahtuu niin testattavan kuin tutkijankin puolesta. Clozella saadaan myös helposti osioita suhteessa tekstin pituuteen, ja ne jakautuvat tasaisesti tekstin eri osiin, mikä saattaa olla huomattavasti vaikeammin toteutettavissa tekstin jälkeisillä kysymyksillä.

TAULUKKO 2. Perinteinen cloze-proseduuri (Oller & Jonz, 1994) verrattuna HyTeC-clozen proseduurin (Kleijn, Pander Maat & Sanders, 2019)

Ominaispiirteet	Perinteinen cloze	HyTeC-cloze
1. Aukonvalintastrategia	Mekaaninen	Harkittu + mekaaninen
2. Valintasuhde	20 %	10 %
3. Valintaetäisyys	Vakioitu, joka 5. sana	Vaihteleva
4. Aukkojen määrä (per 300 sanaa)	60	30
5. Ei-poistettavat sanat	Yleisnimet, numerot	Paikallisesti ennakoitavat ja arvaamista vaativat sanat
6. Vastaustapa	Avoin	Avoin
7. Pisteytys	Eksakti	Semanttinen

Tutkimuskysymykset

Tämän tutkimuksen päätarkoituksena oli tarkastella digitalisoidun cloze-testin validiteettia luetun ymmärtämisen mittarina vertaamalla sitä Suomessa yleisesti käytössä olevaan ja normitettuun alasteen lukutestiin eli ALLUun. Cloze-testin validiteetin tarkastelu on tärkeää myös siksi, koska Suomessa ei toistaiseksi ole nopeaa tapaa luetun ymmärtämisen mittaamiselle. Lisäksi tutkimuksessa tarkasteltiin kummankin testin yhteensopivuutta IMREC-mallin (van den Broek & Espin, 2012) viitekehityksessä esitettyihin mittarin rakentamisen oleellisiin faktoreihin (taulukko 3), jotka perustuvat luetun ymmärtämisen mittaamisen kirjallisuudessa yleisesti hyväksytyihin tekijöihin. Testien yhteensopivuus faktoreihin kertoo testien validiteetista luetun ymmärtämisen mittarina.

Tutkimuksen toisena tavoitteena oli tarkastella, painottuuko päättelykyvyn rooli cloze-testissä. Koska cloze-testi pyrittiin muodostamaan kirjallisuuteen pohjautuen erityisesti päättelykykyä vaativaksi valitsemalla aukoiksi sidoksisuutta sisältäviä sanoja sekä aukkoja, joiden ratkaiseminen vaatii vihjeiden hakemista lauseen ulkopuolelta, oletettiin päättelykyvyllä olevan merkittävämpi yhteys cloze-testin tulokseen kuin muilla tutkimuksen muuttujilla. Nykyaikaisen cloze-proseduurin mukaisesti luotujen testien on havaittu vaativan luetun ymmärtämisen korkeampia prosesseja (Ellert & Elbro, 2013; Kleijn ym., 2019), joten oletimme päättelykyvylle jäävän selitysvoimaa sen jälkeen, kun luetun ymmärtämisen alempien komponenttien vaikutus on kontrolloitu. Nojaten aiempiin tutkimustuloksiin ALLUn tulosten oletettiin olevan yhteydessä kaikkiin kognitiivisiin taustamuuttujiin. Aiempaan kirjallisuuteen nojaten ALLUn tulosten oletettiin painottuvan päättelykykyä lukuun ottamatta samalla tavoin kuin cloze-testissä.

Tiivistettynä tutkimuskysymykset olivat:

1. Mittaako cloze-testi luetun ymmärtämistä?
2. Painottaako cloze-testi päättelykykyä, kun testien ja muuttujien keskinäiset yhteydet on kontrolloitu?

AINEISTO & MENETELMÄT

Tutkimukseen osallistui kaiken kaikkiaan 114 keskisuomalaisen koulun 3.-4. luokan oppilasta (62 poikaa, 51 tyttöä). Kolmannen luokan oppilaita oli 41 ja neljännen luokan oppilaita 73. Heistä 110 osallistui yksilö- ja ryhmämittauksiin. Osallistujia tutkittiin kevään 2019 aikana kolmena eri mittauskertana sekä yksittäin että ryhmittäin.

Menetelmät ja muuttujat

Cloze-testi ja ALLU tehtiin ryhmämittauksin tietokoneella Niilo Mäki Instituutin kehittämällä verkkovälitteisellä Arvio –teknologialla. Myös lukusujuvuus mitattiin tietokoneavusteisesti ryhmissä. Paperilla tehtävillä ryhmämittauksilla kerättiin tietoa näönvaraisesta päättelykyvystä ja sanaston laajuudesta. Kielellisestä päättelykyvystä, työmuistista ja ilmaisevan lukemisen taidoista kerättiin tietoa yksilömittauksissa paperilla. Lisäksi ennakkotietoja selvitettiin paperille täytettävällä alkukyselyllä, joka tehtiin lukemiseen vaikuttavista taustatekijöistä, kuten temperamentista, lukumotivaatiosta ja -tottumuksista. Ilmaisevan lukemisen tehtävä ja suurin osa lukemiseen vaikuttavista taustatekijöistä jätettiin kuitenkin analyysivaiheessa pois. Taustatekijöistä huomioitiin oppilaan äidinkieli, sukupuoli ja luokka.

Cloze-testi

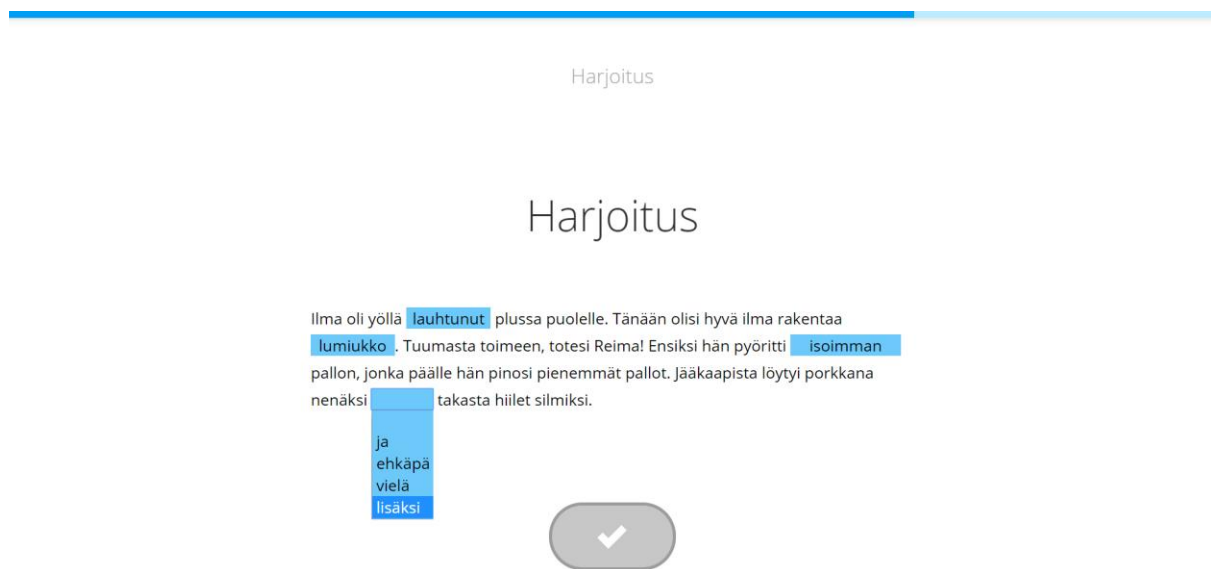
Cloze-testin teksteissä oli lauseita, joista puuttui jokin tekstin kontekstin tai kielioppisääntöjen avulla pääteltävissä oleva sana, joka valittiin neljästä vaihtoehdosta. Tehtävä koostui 16 tekstistä, joissa kussakin oli neljä täydennettävää aukkoa. Tekstit olivat lapsille tuttuihin konteksteihin liittyviä, pituudeltaan 33-59 sanan mittaisia lyhyitä tarinoita ja koko tehtävän tekemiseen oli käytettävissä 20 minuuttia. Tutkijat muodostivat teksteihin aukkoja tarkoituksenmukaisesti niin, että kussakin tekstissä oli aukko, jonka täyttämiseen vaadittava vihje oli sijoitettu lähelle eli yhden virkkeen sisälle aukosta tai kauas eli muualle kuin aukon sisältävään virkkeeseen. Teksteihin oli sijoitettu myös hämääviä vihjeitä, joiden tarkoitus oli ohjata lukijaa muuhun kuin aukon oikeaan ratkaisuun. Vastausvaihtoehdot valittiin siten, että yksi oli vihjeiden perusteella selkeästi oikea vastaus, kaksi olivat tekstiin sopivia, mutta ei vihjeiden kanssa yhteneviä vastauksia ja yksi oli kieliopillisista syistä selkeästi väärä vastaus. Oikeasta vastauksesta sai yhden pisteen, sopivasta ja väärästä vastauksesta nolla pistettä. Mahdolliseksi kokonaispistemääräksi muodostui 0-64 pistettä. Aukkojen ja vihjeiden ominaisuuksia käsitellään vielä tarkemmin analyysivaiheessa.

Esimerkkitekstejä cloze-testistä

Simon naapureilla on lemmikkieläimenä *käärme / kyy / lisko / peto*. Sitä kutsutaan *madoksi / matelijaksi / suomuksi / loheksi*, vaikka se onkin paljon suurempi kuin kastemato. Simo ei ole koskaan aikaisemmin nähnyt näin erikoista lemmikkiä. Otus on melko suuri, eikä se *taida / jaksaa / halua / osaa* liikkua kovin usein. Nimittäin Simo on nähnyt vain harvoin, kun se *kiemurtelee / syö / sihisee / lentää* isossa lasisessa laatikossaan.

Taina heräsi aamulla innostuneena. *Huomenna / Tänään / Ylihuomenna / Seuraavana* olisi luokkaretkipäivä! Hän oli odottanut retkeä *luontomuseoon / eläinpuistoon / metsään / sirkukseen* ja halusi erityisesti nähdä ilveksen ja huuhekajan. Hän pakkasi laukkuunsa kameran, jotta saisi otettua kuvia edellä/nyt/tässä/myöhemmin mainituista eläimistä. Tainalla oli seuraavana päivänä hauskaa valokuvatessaan eläimiä, *ja / kunnes / vaikka / miksi* hän sai otettua monta hyvää valokuvaa näytteille asetetuista täytetyistä eläimistä!

KUVA 1. Cloze-testin näkymä tietokoneella



Lukusujuvuus

Lukusujuvuutta mitattiin luokan kokoisissa ryhmissä tietokoneilla. Sen mittaamiseen käytettiin tietokoneversioitua lukusujuvuuden testiä (Luksu), jonka on havaittu toimivan kohtuullisena lukusujuvuuden mittarina (Suokas, 2008). Oppilaat suorittivat testit yksin tietokoneilla siten, että ruutuun ilmestyvään väitteeseen (esim. "Rikkinäinen pesukone korjataan sairaalassa.") vastattiin klikkaamalla hiirellä joko vihreää ("pitää paikkansa") tai punaista ("ei pidä paikkaansa") painiketta. Ennen kuin oppilaat pääsivät varsinaiseen tekstiin, oli heidän suoritettava viisi harjoituskierrosta siten, että oppilaan vastatessa väärin tuli ruudulle ilmoitus "Väärä vastaus!". Oppilas sai tämän jälkeen korjata väärän vastauksensa ja edetä, kunnes oli saanut kaikkiin harjoitustehtäviin oikeat vastaukset. Harjoituskierrosten jälkeen alkoi varsinainen testi, jossa oppilaan oli samaan tapaan määrä päätellä, oliko väitelause totta vai ei, käyden läpi niin monta väitettä kuin kahden minuutin (120s.) aikana ehti tai käyden läpi kaikki 70 väitettä. Toisin kuin harjoituksen aikana, testi ei antanut oppilaalle palautetta siitä vastasiko hän oikein vai väärin. Mitä enemmän oppilas antoi oikeita vastauksia väittämiin, sitä paremmaksi arvioitiin hänen lukusujuvuutensa.

KUVA 2. Lukusujuvuus

Harjoitus

Maitoa voi juoda.



Päätelykyky ja muisti

WISC-IV (Wechsler Intelligence Scale for Children) on kouluikäisten lasten älykkyyden mittaamiseen tarkoitettu arviointimenetelmä. Se koostuu 15 osatestistä, jotka muodostavat neljä älykkyyden osa-alueita: kielellinen päättely, näönvarainen päättely, työmuisti ja prosessointinopeus. Yksittäiset osatestit eivät anna tarkkaa kuvaa tutkittavan kognitiivisista kyvyistä, mutta yksittäisten osatestien tuloksilla voidaan tehdä suuntaa antavia hypoteeseja tutkittavan kykyprofiilista (Lindeman, 1998). Lukuun ottamatta prosessointinopeutta, kultakin älykkyyden osa-alueelta otettiin käyttöön yksi osatesti.

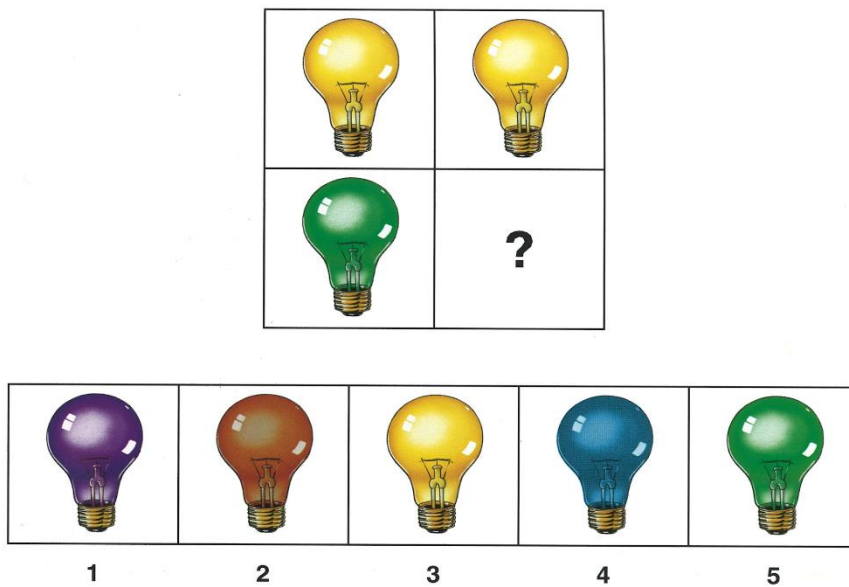
Yksilömittauksissa kerättiin tietoa kielellisestä päättelykyvystä sekä työmuistin toiminnasta. Kielellistä päättelyä testattiin WISC-IV:n samankaltaisuudet -osatestillä, jossa oppilaan tuli kertoa mitä yhteistä kullakin esitetyllä sanaparilla on. Sanapareja yhdistävien käsitteiden muodostaminen edellyttää oppilaalta kykyä ymmärtää kuulemaansa, kykyä erottaa epäoleelliset ja oleelliset ominaisuudet toisistaan sekä kykyä kielelliseen ilmaisuun. Vastaukset pisteytettiin manuaalin mukaisesti välillä 0-39, jossa suurempi pistemäärä kertoo paremmasta kielellisestä päättelykyvystä. Manuaaliin merkitty mittarin reliabiliteetti on .87.

WISC-IV:n työmuisti-indeksistä mitattiin työmuistia ja lyhytkestoista muistia numerosarjatehtävillä, jossa tutkija esitti oppilaalle piteneviä numerosarjoja, jotka oppilaan tuli toistaa ensimmäisessä vaiheessa samassa järjestyksessä ja toisessa vaiheessa päinvastaisessa

järjestyksessä. Samassa järjestyksessä toistaminen mittaa lyhytkestoisen muistin kapasiteettia, ja päinvastaisessa järjestyksessä toistaminen työmuistin kapasiteettia. Vastaukset kummastakin vaiheesta pisteytettiin manuaalin mukaisesti välillä 0-16, jossa suurempi pistemäärä kuvaa suurempaa muistikapasiteettia. Manuaaliin merkitty mittarin reliabiliteetti on .731.

Näönvaraista päättelykykyä mitattiin luokkatilanteeseen sovitulla WISC-IV:n matriisipäättelytehtävällä, jossa täydennettiin 35 kuvasarjasta puuttuva neljäs osa jollain annetuista vastausvaihtoehdoista. Tehtävän tekemiseen oli käytettävissä 10 minuuttia. Vastaukset pisteytettiin manuaalin mukaisesti välillä 0-35, jossa suurempi pistemäärä kuvaa parempaa näönvaraista päättelykykyä. Mittarin manuaaliin merkitty reliabiliteetti on .925.

KUVA 3. Matriisipäättelytehtävä



8

Sanasto

Sanaston laajuutta mitattiin ryhmätesteillä luokittain. Oppilaiden tuli valita 17 sanan listasta jokaiselle sanalle 4 vaihtoehdosta merkitystä parhaiten vastaava sana. Jokaisesta oikeasta vastauksesta sai yhden pisteen ja väärästä vastauksesta ei menettänyt pistettä. Oppilaat pisteytettiin välillä 0-17.

KUVA 4. Sanaston laajuutta mittaava tehtävä

SANASTOTEHTÄVÄ

*Mikä sanoista A-D vastaa merkitykseltään parhaiten numeroitua isoilla kirjaimilla kirjoitettua sanaa?
Ympyröi vastauksesi.*

1. OIKEASTAAN A) itse asiassa B) varmaankaan C) ehkä D) reilusti
2. AINES A) osa B) tarvike C) neste D) pala
3. MATELIJA A) etana B) käärme C) rapu D) mato
4. TOISINAAN A) usein B) harvoin C) silloin tällöin D) koskaan
5. SKEITTILAUTA A) lumilauta B) astinlauta C) pyykkilauta D) rullalauta

Ala-asteen lukutesti

Oppilaat suorittivat ALLUn käsikirjan ohjeiden mukaisesti kouluhenkilökunnan johdolla ja oppilaiden suorituksista saatiin ALLUn kokonaispistemäärät. Kokonaispistemäärät muutettiin analyyseissa suhteellisiksi oikeellisuusprosentteiksi.

Aineiston analysointi

Aineiston tilastolliseen analysointiin käytettiin Statistical Package for the Social Sciences- ohjelman versiota 24. Muodostimme ensin regressiomallin, jossa cloze-testin tulosta selitettiin ALLUn tuloksella. Mallin perusteella poistimme seuraavista analyysivaiheista sellaiset koehenkilöt, joiden standardoitu residuaali mallissa oli itseisarvoltaan suurempi kuin 2 sekä koehenkilöt, joiden tulos cloze-testissä jäi arvaustason alapuolelle (oikeat vastaukset < 25 %). Poistettuja koehenkilöitä oli yhteensä 12. Poistetuista koehenkilöistä seitsemällä oli normaalijakauman mukaiset pisteet ALLUssa ja kognitiivisissa mittareissa, mutta erittäin matalat pisteet cloze-testissä, minkä katsoimme johtuvan vaikeudesta motivoitua kokeen suorittamiseen. Kolme poistetuista koehenkilöistä ei puhunut äidinkielenään suomea, mikä näytti vaikuttavan suoriutumiseen kaikissa testeissä. Lisäksi poistettiin kaksi koehenkilöä, joista toinen suoriutui selkeästi muita paremmin ja toisella oli puuttuvaa tietoa. Yhteensä 98 koehenkilöä valittiin mukaan seuraaviin analyysivaiheisiin.

Seuraavaksi tarkastelimme luetun ymmärtämisen testien ja kognitiivisten taustatekijöiden välisiä yhteyksiä muodostamalla korrelaatiomatriisin, jossa muuttujina olivat koehenkilöiden vastausten oikeellisuusprosentit cloze-testissä ja ALLUssa sekä lukusujuvuuden, sanaston laajuuden, näönvaraisen päättelykyvyn, kielellisen päättelykyvyn, lyhytkestoisen muistin ja pitkäkestoisen muistin tehtävien pistemäärät.

Seuraavassa vaiheessa muodostimme korrelaatiomatriisiin pohjautuen testien oikeellisuusprosenttia selittävät hierarkkiset regressiomallit. Malleissa luetun ymmärtämisen testi toimi riippuvana muuttujana. Riippumattomiksi muuttujiksi lisättiin ensimmäisellä askeleella sukupuoli ja luokka. Toisella askeleella lisättiin lyhytkestoinen muisti ja työmuisti. Kolmannella askeleella malliin lisättiin lukusujuvuus, neljännellä sanaston laajuus, viidennellä kielellinen päättelykyky ja viimeisellä askeleella näönvarainen päättelykyky.

Seuraavaksi tarkastelimme kahdella regressiomallilla, minkä verran luetun ymmärtämisen testit selittivät toisiaan itsenäisesti. Malleista tallennettiin residuaalit, joiden yhteyttä kognitiivisiin mittareihin tarkasteltiin kahdella regressiomallilla, jossa residuaalit asetettiin riippuvaksi muuttujaksi ja niiden selittäjiksi lisättiin suoraan kaikki kognitiiviset taustatekijät.

TULOKSET

Cloze-testi oli voimakkaasti yhteydessä ALLUun ($r = .606$, $p < .01$). Korrelaatio lukusujuvuuden ($r = .401$, $p < .01$), sanaston laajuuden ($r = .404$, $p < .01$) ja näönvaraisen päättelykyvyn ($r = .528$, $p < .01$) kanssa oli kohtalainen. Korrelaatio kielelliseen päättelykykyyn ($r = .383$, $p < .01$) ja lyhytkestoiseen muistiin ($r = .204$, $p < .05$) oli heikko. Yhteys työmuistiin ei ollut tilastollisesti merkitsevä.

ALLU korreloi kohtalaisesti lukusujuvuuden ($r = .436$, $p < .01$) cloze-spesifisen sanaston ($r = .489$, $p < .01$) ja näönvaraisen päättelykyvyn ($r = .452$, $p < .01$) kanssa. Korrelaatio cloze-spesifisen sanaston kanssa oli kohtalainen. Korrelaatio lyhytkestoisen muistin kanssa oli heikko ($r = .273$, $p < .05$) ja korrelaatio työmuistin kanssa ei ollut tilastollisesti merkitsevä.

Yhteenvetona korrelaatiot kognitiivisiin taustatekijöihin olivat kummallakin testillä pääosin samankaltaisia, mutta clozella oli hieman suurempi korrelaatio näönvaraiseen päättelykykyyn ja ALLUlla sanaston laajuuteen sekä lyhytkestoiseen muistiin.

TAULUKKO 3. Korrelaatiomatriisi tutkittavista muuttujista

Muuttujat	1	2	3	4	5	6	7	8
1 Cloze-testi	-							
2 ALLU	.606**	-						
3 Lukusujuvuus	.401**	.436**	-					
4 Sanaston laajuus	.404**	.489**	.169	-				
5 Näönvarainen päättelykyky	.528**	.452**	.182	.183	-			
6 Kielellinen päättelykyky	.383**	.381**	.233*	.383**	.153	-		
7 Lyhytkestoinen muisti	.204*	.273*	.166	.109	.102	.034	-	
8 Työmuisti	.039	.000	.099	.081	.071	.037	.415**	-

** p < .01, * p < .05

Hierarkkinen regressiomalli (taulukko 4) selitti kokonaisuudessaan 51.8 % cloze-testin vaihtelusta. Sukupuolen, luokan, lukusujuvuuden, sanaston laajuuden, kielellisen päättelykyvyn ja näönvaraisen päätelykyvyn selitysosuus oli tilastollisesti merkitsevä. Kun katsottiin standardoituja regressiokeroimia mallin viimeiseltä askeleelta, näönvarainen päättelykyky ennusti parhaiten testin tulosta ($\beta = .363$, $p < .001$). Sen sijaan lyhytkestoinen muisti ja työmuisti eivät olleet yhteydessä cloze-testin tuloksiin.

ALLUa selittävä regressiomalli selitti kokonaisuudessaan 56.2 % testin vaihtelusta. Lyhytkestoisen muistin, työmuistin, lukusujuvuuden, sanaston laajuuden, kielellisen päätelykyvyn ja näönvaraisen päätelykyvyn selitysosuus oli tilastollisesti merkitsevä. Kun katsottiin standardoituja regressiokeroimia mallin viimeiseltä askeleelta, sanaston laajuus ennusti parhaiten testin tulosta ($\beta = .333$, $p < .001$). Sen sijaan sukupuoli ja luokka eivät olleet yhteydessä ALLUn tuloksiin.

TAULUKKO 4. Vertailu luetunymmärtämistestin oikeellisuusprosenttia selittävistä hierarkkisista regressiomalleista

Muuttujat	Cloze			ALLU		
	β	ΔR^2	R^2	β	ΔR^2	R^2
Askel 1:		.087*	.087*		.040	.040
1. Sukupuoli	-.152			-.096		
2. Luokka	.101			-.166*		
Askel 2:		.045	.132*		.084*	.124*
3. Lyhytkestoinen muisti	.128			.215**		
4. Työmuisti	-.069			-.159*		
Askel 3:		.147***	.279***		.180***	.304***
5. Lukusujuvuus	.251**			.299***		
Askel 4:		.089**	.367***		.170***	.474***
6. Sanaston laajuus	.190*			.333***		
Askel 5:		.041*	.409***		.028*	.503***
7. Kielellinen päättelykyky	.196*			.164*		
Askel 6:		.109***	.518***		.060**	.562***
8. Näönvarainen päättelykyky	.363***			.268**		

*** $p < .001$, ** $p < .01$, * $p < .05$. Huom. β = standardoitu regressiokerroin mallin viimeiseltä askeleelta

Regressiomalleissa, joissa luetun ymmärtämisen testit olivat ristikkäin riippuvana ja riippumattomana muuttujina, selitysaste oli kummallekin 36.8 %. Clozea selittävän mallin residuaalia (Res 1) ennustavassa regressiomallissa kognitiiviset mittarit selittivät 13.6 % vaihtelusta. Näönvarainen päättelykyky oli ainoa tilastollisesti merkitsevä muuttuja (taulukko 5). Näönvarainen päättelykyky ennusti parhaiten sitä osaa clozesta, jota ALLU ei selittänyt ($\beta = .278$).

ALLUa selittävän mallin residuaalia (Res 2) ennustavassa regressiomallissa kognitiiviset mittarit selittivät kokonaisuudessaan 16.2 % vaihtelusta. Ainoa merkitsevä muuttuja oli sanaston laajuus, joka selitti parhaiten sitä osaa ALLUsta, jota cloze ei selittänyt ($\beta = .228$).

TAULUKKO 5. Clozea ja ALLUa selittävien regressiomallien residuaalien yhteys kognitiivisiin mittareihin

	Res 1 (Cloze)	Res 2 (ALLU)
Muuttujat	β	β
1. Näönvarainen päättelykyky	.278**	.078
2. Sanaston laajuus	.037	.228*
3. Lukusujuvuus	.101	.148
4. Lyhytkestoinen muisti	-.009	.186
5. Työmuisti	.006	-.134
6. Kielellinen päättelykyky	.118	.051

*** $p < .001$, ** $p < .01$, * $p < .05$.

Res 1 = residuaali mallista, jossa clozen tulosta selitettiin ALLUn tuloksella.

Res 2 = residuaali mallista, jossa ALLUn tulosta selitettiin clozen tuloksella.

POHDINTA

Tutkimuksemme päätarkoituksena oli tarkastella digitalisoidun cloze-testin validiteettia luetun ymmärtämisen mittarina 3.-4.-luokan oppilailla vertaamalla sitä Suomessa yleisesti käytettyyn ja normitettuun ALLUun. Lisäksi tarkastelimme molempien luetun ymmärtämisen testien tulosten yhteyttä IMREC-mallin viitekehyksessä esitettyihin faktoreihin, keskittyen erityisesti eroavaisuuksiin päättelykyvyssä.

Tuloksemme osoittavat, että cloze-testin tulos ennustaa melko hyvin ALLUn tulosta. Testien välinen vahva korrelaatio (.606) ja kummankin testin pääosin samankaltaiset korrelaatiot kognitiivisiin taustatekijöihin antavat viitteitä siitä, että testit mittaavat samaa taitoa. Havaitsimme yhteyksien voimakkuuden poikkeavan hieman sanaston laajuuden, näönvaraisen päättelykyvyn ja lyhytkestoisen muistin kohdalla, mikä oli odotettavaa, sillä eri tehtävätyyppien on havaittu korostavan eri voimakkuudella tiettyjen prosessien roolia luetun ymmärtämisessä (Bowey, 1986; Cutting, 2006; Nation & Snowling, 1997; Spear-Swerling, 2004).

Sanaston laajuuden vaikutuksen korostuminen ALLUssa selittyy mahdollisesti sillä, että clozessa sanaston vaikutus pyrittiin kontrolloimaan lyhyillä, arkisilla ja tuttuja sanoja sisältävillä tarinoilla, kun taas ALLUun sisältyy laajemmin aiempaa tietoa vaativia tekstejä (Lindeman, 1998). Lisäksi sanaston laajuuden mittari kehitettiin tätä tutkimusta varten suuntaa antavaksi mittariksi, joten sen yhteyksistä luetun ymmärtämisen testeihin ei voi tehdä vahvoja johtopäätöksiä, ennen kuin mittarin validiteetti on määritelty.

Cloze-testin yhteys päättelykykyyn oli suurempi kuin ALLUssa. Tulos on tavoitteemme mukainen, koska pyrimme rakentamaan cloze-testin HyTeC-proseduuria mukailleen erityisesti päättelykykyä vaativaksi testiksi. Päättelykyvyn on havaittu olevan taitaville lukijoille ominainen piirre (Oakhill, 1984) sekä tukevan tekstin makrorakenteen muodostamista, joka on kokonaisvaltaisen ymmärryksen ja merkityksellisen oppimisen onnistumisen edellytys (Van Dijk & Kintsch, 1983).

Eroavaisuus lyhytkestoisen muistin kohdalla selittyy osin tehtävätyyppien erojen vuoksi. ALLUssa luetaan tarinamuotoisten tekstien lisäksi myös tietotekstejä, joiden on havaittu asettavan suurempia vaatimuksia muistille (Graesser & McNamara, 2011). Yllättävää oli, ettei työmuisti ollut yhteydessä kumpaankaan luetun ymmärtämisen testiin, vaikka aiemmissa tutkimuksissa työmuistin on havaittu olevan yksi tärkeimmistä luetun ymmärtämisen komponenteista (Hannon & Daneman, 2001; Seigneuric, Ehrlich, Oakhill & Yuill, 2000). Tulosta saattaa selittää testiolosuhteet ja mittaustapaan liittyvät ongelmat. Lapsilla oli usein vaikeuksia ymmärtää ohjeita työmuistia mittaavassa tehtävässä. Lisäksi heidän motivaatio saattoi olla mittaushetkellä matalampi kuin testin alussa, sillä työmuisti oli viimeinen testattava taustamuuttuja. Vaihtelu työmuistitehtävän pistemäärässä oli pientä, sillä noin 77 % tutkittavista sai tehtävästä 5-7 pistettä, vaikka teoreettinen vaihteluväli oli 0-16 pistettä. Koska tehtävän erottelevuus otoksessa oli heikko, yhteyttä luetun ymmärtämiseen on vaikeaa arvioida. Cloze-testin osalta tulosta voidaan selittää myös Cainin ym. (2004) luetun ymmärtämistä ennustavalla mallilla, jossa työmuisti nähdään resurssina luetun ymmärtämisen korkeammille prosesseille. On mahdollista, että tehtävän asettama resurssivaatimus työmuistille oli pieni, sillä alempien kognitiivisten taitojen vaikutusta pyrittiin kontrolloimaan. Kaikkien lasten ominainen työmuistikapasiteetti saattoi näin riittää tehtävissä suoriutumiseen siten, että he saivat täytettyä johdonmukaisuuden standardinsa ylittämättä työmuistinsa kapasiteettia. Lisäksi kirjallisuudessa usein sekoitetaan työmuistin ja lyhytkestoisen muistin käsitteet keskenään (Cowan, 2008), joten tarkan erottelun tekeminen ei ole välttämättä perusteltua.

Yksinään ALLUn ja cloze-testin tulosten oikeellisuusprosentteilla voitiin selittää noin kolmasosa toistensa vaihtelusta, mutta yhdessä kognitiivisten taustamuuttujien kanssa selitysaste kasvoi clozessa 55 %:iin ja ALLUssa 59 %:iin. Tulosten samankaltaisuus vaikuttaa lupaavalta ottaen huomioon, että tutkimuksessa käytettiin cloze-testin pilottiversiota, jonka suorittamiseen vaadittiin oppilailta vain 20 minuuttia. Kun lisäksi huomioidaan luetun ymmärtämisen moniulotteisuus, ovat molemmat selitysasteet kontekstiin nähden korkeita. Testien yhdenmukainen sopivuus IMREC-mallin faktoreihin viittaa siihen, että molemmat testit mittaavat suureksi osaksi samoja taitoja. Faktoreiden selitysvoimat kuitenkin painottuvat testien välillä toisistaan eriävästi, joten ne antavat hieman erilaisen kuvan luetun ymmärtämisestä. Mallien selittämättä jäänyttä osaa

voisi mahdollisesti pienentää esimerkiksi tarkastelemalla tutkimuksesta ulosjätettyjen IMREC-mallin faktoreiden vaikutusta.

IMREC-mallissa esitetään 11 faktoria (taulukko 1), jotka luetun ymmärtämisen mittaamisessa tulisi ottaa huomioon mahdollisimman tarkan kuvan saamiseksi. Pyrimme kontrolloimaan lukemiseen liittyvien perustaitojen sekä kielellisen työmuistin kapasiteetin vaikutuksen käyttämällä testissä vain lyhyitä tekstejä. Työmuistia kontrolloitiin myös niin, ettei tehtäviin vastaamiseen ollut aikarajaa ja niihin vastattiin lukemisen yhteydessä. Numerosarjatehtävät ovat kokonaisuudessaan WISC-IV:n työmuisti-indeksin osatestejä, joten lyhytkestoisen muistin voi ajatella kuvaavan yhtä työmuisti-indeksin osa-aluetta. Tulos antaa viitteitä siitä, että työmuistin vaikutus cloze-testin selittäjänä saatiin kontrolloitua. Sitä ei voi kuitenkaan todeta varmuudella, sillä tulos saattaa kertoa myös muistiin liittyvien mittareiden valintaan liittyvistä ongelmista. Työmuistin yhteyden tarkempaan tarkasteluun olisi tarpeellista ottaa mukaan myös muita työmuisti-indeksin osatehtäviä.

Motivaation vaikutusta pyrittiin kontrolloimaan lyhyillä, lapsille suunnatuilla arvoitus-tyyppisillä teksteillä, joihin lasten oli helppo kiinnittyä. Tehtävä oli haastava, mikä näkyi siinä, ettei kukaan saanut siitä täysiä pisteitä. Siitä kuitenkin pystyi selviytymään, eikä siinä ollut suurta epäonnistumisen uhkaa, koska se ei esimerkiksi vaikuttanut kouluarvosanaan. Lasten motivaatiota saattoi heikentää tietokoneella tehtävä tutkimustilanne, sillä koulussa tehtävät testit tehdään perinteisesti paperille eli tutkimustilanne erosi tavallisesta koetilanteesta, jolloin siihen ei välttämättä suhtauduttu samalla vakavuudella. Selvitimme lisäksi alkukyselyllä tutkittavien motivaatiota lukemiseen liittyen, mutta tutkimuskysymysten rajaamiseksi jätimme motivaatiodatan analyysivaiheessa pois. Motivaatiolla on tärkeä rooli lukemisessa, mutta sen tutkimuskentällä on ongelmia. Niitä ovat esimerkiksi motivaation rakenteiden viitekehysten määrittely ja rakenteisiin sopivien mittareiden kehittäminen (Conradi, Jang & McKenna, 2014). Monet hyväksyvätkin, että motivaatio on moniulotteinen, useasta tekijästä koostuva käsite (Baker & Wigfield 1999; Guthrie & Wigfield, 2000; Schiefele, Schaffner, Möller & Wigfield, 2012; Wigfield & Guthrie, 1997). Lisäksi motivaatioon liittyvien faktoreiden mittaaminen voi olla haasteellista, sillä tutkijoiden kesken ei ole yksimielisyyttä siitä, miten niihin liittyviä komponentteja tulisi operationalisoida (Morgan & Fuchs, 2007; Petscher 2010; Schiefele ym., 2012; Watkins & Coffey, 2004).

Aiemman aihetietämyksen, tekstilajitietämyksen sekä sanaston laajuuden vaikutusta pyrimme kontrolloimaan käyttämällä teksteissä arkipäiväisiä tarinoita, jotka liittyivät alakoululaisille tuttuihin aiheisiin, kuten harrastuksiin ja koulunkäyntiin. Sanaston laajuutta mittaavaan tehtävään valitut sanat olivat cloze-teksteistä poimittuja, hankalaksi arvioituja sanoja. Tutkittavat saivat keskimäärin 11.9 pistettä 2.3 pisteen hajonnalla, mikä viittaa siihen, että cloze-testissä käytetty sanasto oli tutkittaville tuttua. Ei kuitenkaan ole varmaa onnistuimmeko luomaan tekstejä, jotka

kokonaisuudessaan vastasivat lapsille tuttuja asioita. Fletcherin (2006) mukaan tekstin haasteellisuuden on oltava sopiva suhteessa lukijan taitotasoon, jotta voidaan tehdä tarkkoja päätelmiä luetun ymmärtämisestä. Saadaksemme tietoa luetun ymmärtämisestä myös vähemmän taitavilta lukijoilta teimme teksteistä tarkoituksella helppoja, sillä Amendumin ym. (2018) meta-analyysi peruskouluikäisillä osoitti, että tekstin vaikeus vaikuttaa negatiivisesti lukemisen tarkkuuteen ja lukunopeuteen etenkin vähemmän taitavilla lukijoilla. Toisaalta tekstin helppoudella saattaa olla kuitenkin myös käänteinen vaikutus esimerkiksi taitavien lukijoiden motivaatioon. Jotta voitaisiin kertoa tarkemmin tekstin sopivuudesta lukijoiden taitotasoon, olisi tulevaisuudessa syytä määrittää tarkasti tekstin haasteellisuuden taso käyttämällä esimerkiksi Deanen ym. (2006) esittämää tekstin haasteellisuutta kuvaavaa faktorimallia. Vaikka tutkimuksessamme ei tarkasteltu tekstin haasteellisuutta kriittisesti, on aiemmissa tutkimuksissa havaittu, että tarinamuotoista tekstiä on merkittävästi helpompi lukea, ymmärtää ja palauttaa mieleen kuin tietotekstiä silloinkin, kun tekstin tuttuus ja sanasto on kontrolloitu (Graesser & McNamara, 2011; Haberlandt & Graesser, 1985).

Tutkimuksessamme erityisenä kiinnostuksen kohteena oli johdonmukaisuuden standardien soveltamisen sekä päättelykyvyn vaikutukset, sillä kirjallisuuden perusteella ne ovat osoittautuneet tärkeiksi cloze-testillä tavoitettaviksi luetun ymmärtämisen komponenteiksi. Johdonmukaisuuden standardeja tarkasteltiin poistettujen sanojen sidoksisuuden kautta. Lukija pyrkii luomaan tekstistä yhteneväistä kokonaiskuvaa, jossa sanat, lauseet ja kappaleet ovat tarkoituksenmukaisesti suhteutettu toisiinsa teksti-informaation sisältämien viittausten ja kausaalisuhteiden sekä kielioppisääntöjen perusteella (van den Broek & Espin, 2012). Jos tekstistä on harkiten poistettu sidoksisuutta sisältäviä sanoja, jotka lukijan tulee päätellä, se asettaa erityisiä vaatimuksia yhtenäisen kokonaiskuvan luomiselle (Gellert ja Elbro, 2013). Sidoksisuuden voidaan siis nähdä olevan ehtona sille, että tekstin eri elementit saavuttavat johdonmukaisuuden. Johdonmukaisuuden standardien vaikutus tutkimuksen cloze-testissä on vain spekulatiivista, sillä tehtävissä oli myös muun tyyppisiä aukkoja, mutta tarkastelimme tutkimuksessamme vain kokonaispistemääriä. Tarkemman kuvan saamiseksi johdonmukaisuuden standardien roolista olisi jokaista aukkoa tutkittava cloze-testissä yksitellen, jotta ne voitaisiin luokitella sidoksisuuden perusteella johdonmukaisuuden standardeja vaativiin ja lokaalisti pääteltäviin aukkoihin.

Johdonmukaisuuden standardit määrittävät semanttisen johdonmukaisuuden tasoa, johon lukija pyrkii (van den Broek ja Espin, 2012). Johdonmukaisuuden luomisessa päättelykyvyllä on keskeinen rooli erityisesti silloin, kun lukija tarvitsee tekstin ulkopuolista informaatiota (Graesser ym., 1994; Long & Chong, 2001). Tutkimuksessamme johdonmukaisuuden standardien soveltamisen oletettiin näkyvän päättelykyvyn vaatimuksissa erityisesti aukoissa, joissa oikea

vastaus on pääteltävä käyttämällä sellaista aiempaa aihetietoa, jonka 3.-4. luokkalaisten oletetaan jo omaksuneen. Jokaisella lukijalla on kuitenkin tyypiltään ja voimakkuudeltaan vaihtelevat johdonmukaisuuden standardit (van den Broek & Espin, 2012), mikä saattaa näkyä cloze-testissämme siten, että osalle lukijoista aukkoon voi tuntua sopivan useampi vastausvaihtoehto, kun taas osalle lukijoista vain yksi vaihtoehto näyttäytyy oikeana, kun tavoitteena on kokonaisvaltaisesti johdonmukainen käsitys tekstistä.

Tarkastellaksemme johdonmukaisuuden standardien eroavaisuuksia koehenkilöittäin, pyrimme aluksi pisteyttämään aukkojen vastausvaihtoehdot siten, että selvästi väärästä vastauksesta sai nolla pistettä, kahdesta lähes oikeasta vaihtoehdosta sai yhden pisteen ja oikeasta vastauksesta sai kaksi pistettä. Havaitsimme kuitenkin, että tällaisella pisteytystavalla nollan pisteen vastauksia ei esiintynyt tarpeeksi, jotta niiden avulla voisi tehdä päätelmiä johdonmukaisuuden standardeista. Oletamme tämän johtuvan siitä, että väärät vastausvaihtoehdot olivat liian selvästi väärin, jolloin pisteytystavan erottelukyky oli heikko. Lisäksi on hankalaa määrittää objektiivisesti vaihtoehtojen oikeellisuus, joten rajanveto yhden ja kahden pisteen vastausten välillä voi olla epäselvää. Päädyimme lopulta pisteyttämään vaihtoehdot siten, että ainoastaan täysin oikeasta vastauksesta sai pisteen. Havaitsimme, ettei näiden pisteytystapojen välillä ollut merkittävää eroa, kun arvioimme lasten luetun ymmärtämistä cloze-testillä.

Mittasimme koehenkilöiltä sekä kielellistä että näönvaraista päättelykykyä saadaksemme toisaalta yleistä kuvaa päättelykyvyn vaikutuksista ja toisaalta saadaksemme erottelua millaista päättelykykyä luetun ymmärtämisen testeissä tarvitaan. Tuloksissamme näönvarainen päättelykyky korostui kielellistä päättelykykyä enemmän. Kirjallisuuden perusteella oletimme kielellisen päättelykyvyn korostuvan enemmän, sillä näönvaraisen päättelykyvyn yhteydestä luetun ymmärtämiseen emme löytäneet aiempaa tietoa johon nojata. Saimme tutkimuksemme avulla tietoa siitä, mikä rooli yleiseen päättelykykyyn sisältyvällä näönvaraisella päättelykyvyllä voi olla luetun ymmärtämisessä. Tulostemme perusteella sillä näyttäisi olevan merkittävä vaikutus, joka ei ole riippuvainen muista mittaamistamme luetun ymmärtämisen kognitiivisista komponenteista. Toisaalta joidenkin tutkijoiden mukaan WISC IV:n matriisipäättelytehtävässä pärjääminen saattaa välittyä osittain kielellisen päättelykyvyn kautta (Dugbartey, Sanchez, Rosenbaum, Mahurin, Davis & Townes, 1999; Sattler & Dumont, 2008), joten tuloksemme ei välttämättä kerro puhtaasti näönvaraisen päättelykyvyn painotuksesta, vaikka pyrimmekin kontrolloimaan kielellisen päättelykyvyn vaikutusta WISC IV:n samankaltaisuudet -tehtävällä.

Houskeeperin (2011) mukaan samankaltaisuudet -tehtävä on yhteydessä Cattell-Horn-Carroll-älykkyysteoriassa (Flanagan & Dixon, 2013) määriteltyyn kiteytyneen älykkyiden faktoriin, kun taas kuviomatriisitehtävä on yhteydessä joustavan älykkyiden faktoriin. Joustava älykkyys koostuu

muun muassa induktiivisesta ja deduktiivisesta päättelytaidosta sekä ongelmanratkaisutaidoista, kun taas kiteytynyt älykkyys on kulttuurin, opetuksen ja muiden oppimiskokemusten kautta sisäistettyä kielellistä tietämystä (Flanagan & Dixon, 2013). Matriisipäättelytehtävä voidaan siis nähdä päättelykyvyn testinä, joka mittaa yleistä päättelykykyä puhtaammin kuin kielellistä päättelykykyä mittaava sanaparitehtävä. Lisäksi sanaparitehtävässä oikean vastauksen pisteyttämiseen liittyy epämääräisyyttä. Huomasimme pisteyttäessämme sanaparitehtävää, että ohjeiden mukainen pisteytys ei ollut yksioikoista, vaan monien vastausten kohdalla oikea vastaus saattoi näyttäytyä vääränä ja väärä vastaus oikeana riippuen siitä, millaisella logiikalla vastausta lähestyi. Belkin, LoBellon, Rayn & Zacharin tutkimuksessa (2002) havaittiin, että tutkinto-opiskelijat tekivät merkittävän määrän virheitä testatessaan lapsia ja pisteyttäessään heidän tuloksiaan WISC-III:n sanaparitehtävällä. Kuviomatriisitehtävän pisteytyksessä puolestaan ei ole tulkinnan varaa, sillä oikea vastaus valitaan annetuista vaihtoehdoista.

Tutkimuksemme osoittaa, että cloze on helppokäyttöinen ja testaajan tarpeisiin mukautettava testi, joka täyttää perustellusti luetun ymmärtämisen testin vaatimuksia. Testi ei vaadi paljon resursseja myöskään opetushenkilökunnalta tai testattavilta oppilailta, sillä se voidaan tehdä verkkoympäristössä koulun omalla laitteistolla yhden oppitunnin aikana.

Tuloksemme tukevat IMREC-mallin käyttämistä teoriapohjana luetun ymmärtämisen mittaria rakentaessa. Sen perusteella valitut kognitiiviset komponentit selittivät merkittävästi suuren osan molempien luetun ymmärtämisen testien vaihtelusta, vaikka esimerkiksi motivaation vaikutusta ei huomioitu tarkasti. Tämä antaa suuntaa sille, mitä tekijöitä validin luetun ymmärtämisen testin rakentamisessa on otettava huomioon. IMREC-mallin laatijoiden mukaan luetun ymmärtämisen testissä täytyy huomioida lukemisen moniulotteisuus ja mittaamisen tarkoitus, joka voi olla seulonta, kehityksen seuranta tai heikkouksien ja vahvuuksien tunnistaminen. Tutkimuksemme perusteella cloze-testillä voidaan tehdä samanlaista seulontaa kuin ALLUlla, mutta tarvitaan jatkotutkimuksia siitä, voidaanko clozella seurata myös luetun ymmärtämisen kehitystä ja siinä esiintyvien heikkouksien ja vahvuuksien yksilöllistä tunnistamista kliinisissä tarkoituksissa.

Van den Broekin & Espinin (2012) mukaan mittarin validiteetti tulisi nykykäsityksen mukaan osoittaa muodostamalla hypoteeseja mittarista saatujen tulosten merkityksestä käytännön tasolla. Käytännön tason hypoteesien muodostaminen oli kuitenkin hankalaa tutkimuksemme ollessa vasta pilottivaiheessa. Tarkastelimme validiteettia perinteisemmän käsityksen mukaisesti, jossa testin validiteetin takaamiseksi käytetään yleisesti tuettua, tarkasti tutkittua teoriaa, josta muodostetaan teorian mukaiset operationalisoidut, mitattavat muuttujat. Validiteetin vahvistamiseksi jatkossa tulisi tarkastella, miten testin tuloksiin perustuvat käytännön tason toimenpiteet ovat yhteydessä toimenpiteitä seuraaviin mahdollisiin muutoksiin esimerkiksi tutkimalla vaikuttaako heikosti

clozessa pärjänneelle oppilaalle annettu tukiopetus positiivisesti clozessa pärjäämiseen. Esimerkiksi Yuill & Joscelyne (1988) havaitsivat päättelytaitojen kehittyvän lapsilla, joita opetettiin kohdistamaan huomionsa tekstin avainsanoihin.

Validiteetin määrittämistä hankaloittaa luetun ymmärtämisen tutkimuskentän vakiintumattomat käytänteet ja puuttuva konsensus ideaalin mittarin ominaisuuksista. Vaikka luetun ymmärtämisen kognitiivisista taustatekijöistä löytyy paljon tutkimuksia, on vaikea löytää sellaisia tutkimusasetelmia, joissa on käytetty yhteneväisiä mittareita ja kontrolloitu samoja taustamuuttujia. Mittarit ja muuttujat vaihtelevat muun muassa sen mukaan, miten kussakin tutkimuksessa luetun ymmärtäminen on käsitetty. Päätimme siten käyttää IMREC-mallia viitekehysenä, sillä sen yhdistäessä useita teorioita ja malleja, voidaan se nähdä mahdollisena lähtökohtana tutkimuskentän käytänteiden vakiinnuttamiselle. Kognitiivisina mittareina käytimme laajasti tutkittuja ja käytettyjä WISC IV:n osatestejä tai niiden sovitettuja versioita. Lisäksi tutkimustulosten yleistettävyyteen ja vertailuun vaikuttaa käytetty kieli. Tutkimuksissa käytetään usein kieltä, jossa ei ole säännöllistä ortografiaa toisin kuin suomen kielessä, jolloin tulosten vertailtavuus ei ole yksiselitteistä. Vaikka cloze-testiä on tutkittu maailmalla paljon, ei tämän tyyppistä cloze-testiä ole tutkittu Suomessa aiemmin. Tutkimuksemme luo siis perustaa suomenkieliselle cloze-testin kehittämislle.

Cloze-testiä on aiemmin kritisoitu siitä, että se mittaa pääasiassa lukemisen alempia komponentteja, kuten lukusujuvuutta tai kielioppitietämystä sekä siitä ettei aukkojen täyttäminen vaadi lauseen ulkopuolisen tiedon käyttämistä (Alderson, 1979; Kintsch & Yarbrough, 1982; Klein-Braley & Raatz, 1984; Pearson & Hamm, 2005; Shanahan ym., 1982). Clozen puolesta puhuvien tutkijoiden mukaan edellä mainittuun ongelmaan voidaan vastata harkituilla aukkovalinnoilla, joissa huomioidaan esimerkiksi sanojen arvattavuus ja sidoksisuus, jotka asettavat vaatimuksia korkeammille prosesseille (Kleijn, Pander Maat & Sanders, 2019). Myös tuloksemme tukevat tätä, sillä regressiomallissa kielellinen päättelykyky ja näönvarainen päättelykyky lisäsivät cloze-testin tulosten selitysastetta senkin jälkeen, kun alempien lukemisen taitojen vaikutus kontrolloitiin.

Muodostamamme cloze-testi, joka pohjautuu IMREC-mallin faktorirakenteeseen ja HyTeC-proseduuriin, näyttää suurelta osin vastaavan kritiikkiin, joka on kohdistunut perinteiseen cloze-testiin ja luetun ymmärtämisen mittaamiseen yleisesti. Tuloksemme osoittavat, että cloze-testi toimii luetun ymmärtämisen mittarina ja meillä on viitteitä siitä, että testillä voi seuloa luetun ymmärtämisen vaikeuksista kärsiviä yhtä hyvin kuin ALLUlla, joka on ainut normitettu luetun ymmärtämisen testi Suomessa. Cloze-testimme vastaa myös resurssien luomiin rajoitteisiin, mikä edistää lukutaidon järjestelmällistä arviointia ja varhaista puuttumista mahdollisiin ongelmiin. Testimme vaivaton muovailtavuus kielii testin kyvystä mukautua ajan mukana tuleviin uusiin muutosvaatimuksiin.

LÄHTEET

- Alderson, J. C. (1979). The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly*, 219-227.
- Amendum, S. J., Conradi, K., & Hiebert, E. (2018). Does text complexity matter in the elementary grades? A research synthesis of text difficulty and elementary students' reading fluency and comprehension. *Educational Psychology Review*, 30(1), 121-151.
- Aro, M. 2004. Learning to read: The effect of orthography. Jyväskylä: Jyväskylän yliopisto. Jyväskylä Studies in Education, Psychology and Social Research.
- Bachman, L. F. (1985). Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, 19, 535-556. DOI: 10.2307/3586277.
- Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly*, 34, 452-477.
- Belk, M. S., LoBello, S. G., Ray, G. E., & Zachar, P. (2002). WISC-III administration, clerical, and scoring errors made by student examiners. *Journal of Psychoeducational Assessment*, 20(3), 290-300.
- Bowey, J. A. (1986). Syntactic awareness in relation to reading skill and ongoing reading comprehension monitoring. *Journal of experimental child psychology*, 41(2), 282-299.
- Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, 64(3), 311-317.
- Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of educational psychology*, 96(1), 31.
- Calvo, M. G. (2005). Relative contribution of vocabulary knowledge and working memory span to elaborative inferences in reading. *Learning and individual differences*, 15(1), 53-65.
- Caspi, A., Wright, B. R. E., Moffitt, T. E., & Silva, P. A. (1998). Early failure in the labor market: Childhood and adolescent predictors of unemployment in the transition to adulthood. *American sociological review*, 63(3), 424.
- Conradi, K., Jang, B. G., & McKenna, M. C. (2014). Motivation terminology in reading research: A conceptual review. *Educational psychology review*, 26(1), 127-164.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in brain research*, 169, 323-338.

- Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2), 311.
- Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific studies of reading*, 10(3), 277-299.
- Deane, P., Sheehan, K. M., Sabatini, J., Futagi, Y., & Kostin, I. (2006). Differences in text structure and its implications for assessment of struggling readers. *Scientific Studies of Reading*, 10(3), 257-275.
- Deyes, T. (1984). Towards an authentic 'discourse cloze'. *Applied Linguistics*, 5(2), 128-137.
- Dugbartey, A.T., Sanchez, P.N., Rosenbaum, J.G., Mahurin, R.K., Davis, J.M., & Townes, B.D. (1999). WAIS-III Matrix Reasoning test performance in a mixed clinical sample. *The Clinical Neuropsychologist*, 13(4), 396-404.
- Flanagan, D. P., & Dixon, S. G. (2013). The Cattell-Horn-Carroll theory of cognitive abilities. *Encyclopedia of special education: A reference for the education of children, adolescents, and adults with disabilities and other exceptional individuals*.
- Fletcher, C. R., & Bloom, C. P. (1988). Causal reasoning in the comprehension of simple narrative texts. *Journal of Memory and language*, 27(3), 235-244.
- Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific studies of reading*, 10(3), 323-330.
- Francis, D. J., Snow, C. E., August, D., Carlson, C. D., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *Scientific Studies of Reading*, 10(3), 301-322.
- García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84(1), 74-111.
- Gellert, A. S., & Elbro, C. (2013). Cloze tests may be quick, but are they dirty? Development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment*, 31(1), 16-28.
- Graesser, A. C., & Kreuz, R. J. (1993). A theory of inference generation during text comprehension.
- Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in cognitive science*, 3(2), 371-398.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological review*, 101(3), 371.
- Gray, D. E. (2013). *Doing research in the real world*. Sage.

- Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 403–422). Mahwah: Erlbaum.
- Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General*, 114, 357–374.
- Halliday, M. A. K. (1980). Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of educational psychology*, 93(1), 103.
- Hinofotis, F. B., & Snow, B. G. (1980). An alternative cloze testing procedure: Multiple-choice format. *Research in language testing*, 129-133.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing*, 2(2), 127-160.
- Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research & Practice*, 18(4), 237-245.
- Johansson, A., & Vuori, J. (1999). *Työttömät nuoret syrjäytymisvaarassa. Työllistämisen- ja aktivointiprojektin aloittaneiden nuorten tausta ja terveystilanne*. Helsinki: Työterveyslaitos.
- Keenan, J. M., & Betjemann, R. S. (2006). Comprehending the Gray Oral Reading Test without reading it: Why comprehension tests should not include passage-independent items. *Scientific studies of reading*, 10(4), 363-380.
- Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 12(3), 281-300.
- Kendeou, P., van den Broek, P., Helder, A., & Karlsson, J. (2014). A cognitive view of reading comprehension: Implications for reading difficulties. *Learning disabilities research & practice*, 29(1), 10-16.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological review*, 95(2), 163.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review*, 85(5), 363.
- Kintsch, W., & Walter Kintsch, C. B. E. M. A. F. R. S. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.

- Kintsch, W., & Yarbrough, J. C. (1982). Role of rhetorical structure in text comprehension. *Journal of educational psychology*, 74(6), 828.
- Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational psychology*, 100(2), 310.
- Kleijn, S., Pander Maat, H., & Sanders, T. (2019). Cloze testing for comprehension assessment: The HyTeC-cloze. *Language Testing*, 0265532219840382.
- Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test1. *Language Testing*, 1(2), 134-146.
- Kuisma, H., & Myllylä, M. (2013). " Lukemaan kumminkin oppii lukemalla, että se on se tärkein, että lapsi lukee riittävästi": opettajien käsityksiä luetun ymmärtämisestä ja keinoja tukea luetun ymmärtämisen kehittymistä.
- LaBerge, D., & Samuels. S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293–323.
- Lemarié, J., Lorch Jr, R. F., Eyrolle, H., & Virbel, J. (2008). SARA: A text-based and reader-based theory of signaling. *Educational Psychologist*, 43(1), 27-48.
- Lindeman, J. (1998). ALLU – Ala-asteen lukutesti [ALLU – Reading Test for Primary School]. University of Turku, Finland: The Center for Learning Research.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological review*, 99(3), 440.
- Morgan, P. L., & Fuchs, D. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation?. *Exceptional children*, 73(2), 165-183.
- Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of experimental child psychology*, 73(2), 139-158.
- Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, 67, 359–370.
- Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British journal of educational psychology*, 54(1), 31-39.
- Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and cognitive processes*, 18(4), 443-468.
- Oller, J. W., & Jonz, J. (1994). *Cloze and coherence*. Associated University Presse.
- Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices—past, present, and future. *Children’s reading comprehension and assessment*, 13-69.

- Perfetti, C. (1985). *Reading ability*. New York: Oxford University Press.
- Perfetti, C. A., & Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. *Measuring up: Advances in how we assess reading ability*, 3-20.
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22-37.
- Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading*, 33(4), 335-355.
- Pike, L. W. (1979). An evaluation of alternative item formats for testing English as a foreign language.
- Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, 11(3), 235-257.
- Rimrodt, S., Lightman, A., Roberts, L., Denckla, M. B., & Cutting, L. E. (2005, February). Are all tests of reading comprehension the same. In Poster presented at the 33rd annual international neuropsychological society meeting.
- Sattler, J. M., & Dumont, R. (2008). WISC-IV Subtests. *JM Sattler, Assessment of Children: Cognitive Foundations*, 316-363.
- Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, 47, 427-463.
- Seigneuric, A., Ehrlich, M. F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and writing*, 13(1-2), 81-103.
- Seymour, P. H. K., Aro, M. & Erskine, J. M. 2003. Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174.
- Shanahan, T., Kamil, M. L., & Tobin, A. W. (1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly*, 229-255.
- Spear-Swerling, L. (2004). FOURTH GRADERS' PERFORMANCE ON A STATE-MANDATED ASSESSMENT INVOLVING TWO DIFFERENT MEASURES OF READING COMPREHENSION. *Reading Psychology*, 25(2), 121-148.
- Suokas, M. (2008). *Lukusujuvuustestin kehittäminen ja arviointi*. Jyväskylän yliopisto. Psykologian pro gradu -tutkielma.
- Turunen, T., Alisaari J., Poskiparta, E. & Lindeman, J. (2018). ALLUn käyttökokemuksia ja luetun ymmärtämisen arviointi vuosina 1995 ja 2017. *Oppimisen ja oppimisvaikeuksien erityislehti*.

- Van Dijk, T. A., Kintsch, W., & Van Dijk, T. A. (1983). Strategies of discourse comprehension.
- Van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. *Comprehension processes in reading*, 423-445.
- Van den Broek, P., & Espin, C. A. (2012). Connecting cognitive theory and assessment: Measuring individual differences in reading comprehension. *School Psychology Review*, 41(3), 315.
- Van den Broek, P., Bohn-Gettler, C. M., Kendeou, P., Carlson, S., & White, M. J. (2011). When a reader meets a text: The role of standards of coherence in reading comprehension.
- Van den Broek, P., Helder, A., & Van Leijenhorst, L. (2012). Sensitivity to structural centrality: Developmental and individual differences in reading comprehension skills. In *Reading-From Words to Multiple Texts* (pp. 146-160). Routledge.
- Van den Broek, P., Risdien, K., & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In Part of this research was reported at the Annual Meeting of the American Educational Research Assn, Chicago, 1991. Lawrence Erlbaum Associates, Inc.
- Watkins, M. W., & Coffey, D. Y. (2004). Reading Motivation: Multidimensional and Indeterminate. *Journal of Educational Psychology*, 96(1), 110.
- Widdowson, H. G. 1979. *Explorations in Applied Linguistics*. Oxford Oxford University Press.
- Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89, 420-432.
- Yuill, N., & Joscelyne, T. (1988). Effect of organizational cues and strategies on good and poor comprehenders' story understanding. *Journal of educational Psychology*, 80(2), 152.