

Eveliina Yli-Salomäki

# ÄÄNEENAJATTELUMENETELMÄN HAASTEET KÄYTETTÄVYYS- TESTAUKSESSA

Informaatioteknologian ja viestinnän tiedekunta  
Kandidaattitutkielma  
Toukokuu 2024

# TIIVISTELMÄ

Eveliina Yli-Salomäki: Ääneenajattelumenetelmän haasteet käytettävyydestä  
Kandidaattitutkielma  
Tampereen yliopisto  
Tietojenkäsittelytieteiden tutkinto-ohjelma  
Toukokuu 2024

---

Ääneenajattelu on yksi keskeisimmistä käytettävyyden tutkimisen menetelmistä. Siinä käyttäjä ajattelee ääneen tehdessään tehtäviä osana käytettävyydestä. Menetelmän avulla saadaan tietoa käyttäjän ajatusprosesseista ja voidaan havaita käytettävyysongelmia testattavassa järjestelmässä. Menetelmän käyttöön ja käytön ohjeistukseen liittyy kuitenkin vaihtelevuutta ja epäjohdonmukaisuutta, jotka voivat vaikuttaa sen käyttöön ja tuloksiin. Tämän tutkielman tarkoituksena onkin selvittää, mitä haasteita ääneenajattelumenetelmän käyttöön liittyy käytettävyydestä. Tulokset voivat auttaa menetelmää käyttäviä alan toimijoita ja opiskelijoita tiedostamaan, mitä menetelmän käytössä kannattaa huomioida, jotta siitä saa eniten hyötyä.

Tutkielma on kirjallisuuskatsaus, jonka aineistona on 20 vertaisarvioitua lehti- ja konferenssijulkaisua. Lähteiden perusteella löytyneet haasteet on jaettu neljään kategoriaan: käytännön haasteet, teorian haasteet, menetelmän eri varianttien haasteet ja saavutettavuuteen liittyvät haasteet. Variantteja ovat esimerkiksi klassinen, interaktiivinen ja retrospektiivinen variantti.

Tuloksena saatiin, että menetelmään liittyy monenlaisia haasteita. Keskeinen haaste menetelmän käytössä on se, miten osallistujat saadaan ajattelemaan ääneen ja varmistetaan, etteivät he sano asioita vain miellyttääkseen moderaattoria. Teoriaan liittyen keskeisenä haasteena puolestaan on yhtenäisen teorian puute, mikä vaikeuttaa esimerkiksi tutkimusten vertailua.

Menetelmän eri variantteihin liittyy myös omia haasteitaan. Klassisen variantin haasteena on, ettei se välttämättä sovellu käytettävyydestä, kun taas interaktiivisen variantin kohdalla haasteina ovat menetelmän suurempi vaikutus osallistujien työskentelyyn sekä moderaattorin toiminnan korostuminen. Retrospektiivisen variantin haasteiksi nousevat aikaavievuus ja testissä ilmenneiden asioiden unohtuminen. Ääneenajattelu ei myöskään sovi kaikkiin tilanteisiin, sillä se ei välttämättä toimi esimerkiksi lasten tai muistisairaiden kanssa. Eri kulttuureilla voi myös olla ominaispiirteitä, jotka vaikuttavat menetelmän käyttöön.

Vaikka menetelmän käyttöön liittyy useita haasteita, monet niistä ovat silti melko pieniä, eivätkä estä ääneenajattelun käyttöä. Tarvittaisiinkin lisätutkimusta perehtymään tarkemmin haasteisiin, esimerkiksi niiden syihin ja vakavuuteen erilaisissa käyttötilanteissa.

Avainsanat: ääneenajattelu, haasteet, käytettävyydestä, käytettävyys, arviointimenetelmät, saavutettavuus

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

# SISÄLLYSLUETTELO

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Johdanto</b> .....   | <b>1</b>  |
| <b>2</b> | <b>Käytettävyys, käytettävyydestaus ja ääneenajattelu</b> ..... | <b>2</b>  |
| 2.1      | Käytettävyys ja käytettävyydestaus                              | 2         |
| 2.2      | Ääneenajattelu  | 3         |
| 2.2.1    | Teoriatausta  | 3         |
| 2.2.2    | Variantit   | 4         |
| 2.2.3    | Käyttökonteksti   | 5         |
| <b>3</b> | <b>Ääneenajattelumenetelmän haasteet</b> .....                  | <b>6</b>  |
| 3.1      | Käytännön haasteet  | 6         |
| 3.2      | Teoriaan liittyvät haasteet                                     | 8         |
| 3.3      | Erilaisten varianttien haasteet                                 | 9         |
| 3.4      | Haasteet liittyen saavutettavuuteen                             | 11        |
| 3.4.1    | Lapset ja ikääntyneet   | 12        |
| 3.4.2    | Kuulo- ja näkövammaiset   | 13        |
| 3.4.3    | Kulttuuriset tekijät  | 14        |
| <b>4</b> | <b>Kehitysehdotukset ja varianttien haasteiden koonti</b> ..... | <b>14</b> |
| 4.1      | Kehitysehdotuksina puheviestintämalli ja paritestausta          | 15        |
| 4.2      | Koonti eri varianttien haasteista                               | 16        |
| <b>5</b> | <b>Pohdinta</b> .....   | <b>17</b> |
| <b>6</b> | <b>Yhteenveto</b> .....   | <b>18</b> |
|          | <b>Lähdeluettelo</b> .....                                      | <b>19</b> |

## 1 Johdanto

Ääneenajattelu (engl. think aloud) on keskeinen käytettävyytestauksessa käytetty menetelmä, jossa käyttäjä ajattelee ääneen tehtäviä tehdessään, ja sen avulla voidaan esimerkiksi löytää käytettävyysoongelmia testattavassa järjestelmässä. Nielsenin (1993) mukaan ääneenajattelu on mahdollisesti arvokkain käytettävyyden tutkimisen menetelmä. Menetelmän käytössä on kuitenkin epäjohdonmukaisuutta ja vaihtelevuutta liittyen esimerkiksi osallistujan ohjeistukseen, testiä vetävän moderaattorin keskeytyksiin ja teoriataustaan. Nämä voivat olla haasteita, jotka vaikuttavat menetelmän käyttöön ja tuloksiin.

Tutkimuskysymyksenä on selvittää, mitä haasteita ääneenajattelumenetelmän käyttöön liittyy käytettävyytestauksessa. Koska ääneenajattelu on erittäin keskeinen käytettävyytestauksessa käytetty menetelmä, sen mahdollisiin ongelmakohtiin perehtyminen on tärkeää ja mielenkiintoista. Tulokset voivat auttaa menetelmää käyttäviä alan toimijoita ja opiskelijoita tekemään päätelmiä siitä, mitä ääneenajattelumenetelmän käyttöön liittyen kannattaa ottaa huomioon, jotta menetelmästä saa suurimman hyödyn irti.

Tämä tutkielma on kirjallisuuskatsaus. Lähteet ovat artikkeleita ja konferenssijulkaisuja, joita on haettu tietokannoista, minkä lisäksi lähteenä on myös yksi kirja. Tärkeimmät tietokannat, joista tietoa on haettu, ovat ACM Digital Library, IEEE Explore ja ScienceDirect. Keskeisiä hakusanoja tiedonhaussa olivat puolestaan ”think aloud”, ”usability testing”, ”usability”, ”challenges”, ”problems”, ”method” ja ”theory”. Jotta löydettiin tietoa menetelmän käytön haasteista jonkin tietyn ryhmän kohdalla, lisättiin hakuun sellaisia sanoja kuten ”children”, ”elderly” ja ”cultural”.

Aineiston haussa tavoitteena oli erityisesti löytää tutkimuksia, jotka käsittelevät ääneenajattelumenetelmän haasteita. Tämän lisäksi pyrkimyksenä oli löytää tutkimuksia, jotka tarkastelevat ääneenajattelumenetelmää yleisemmällä tasolla tai teoreettisesta näkökulmasta. Tämä auttaa ymmärtämään menetelmän lähtökohdat ja sen taustalla olevaa teoriaa, ja mahdollisesti sitä, miten nämä kietoutuvat haasteisiin.

Haualla löytyneiden tutkimusten tiivistelmästä sai yleensä viitteitä siitä, ovatko ne relevantteja. Lisäksi lähteitä löytyi myös tutkimalla jonkun hyödylliseksi osoittautuneen tutkimuksen lähdeluetteloa ja etsimällä sieltä aiheeseen sopivia lähteitä. Lähteiden valinnassa on pyritty painottamaan uudempia tutkimuksia. Äänenajattelumenetelmää käytetään muillakin aloilla, joten tutkimukset on rajattu koskemaan sellaisia, jotka tarkastelevat menetelmää käytettävyytestauksen kannalta.

Ääneenajattelumenetelmää käytettävyytestauksessa on tutkittu jonkin verran, tosin osa keskeisestä tutkimuksesta on jo yli 20 vuotta sitten julkaistua. Tähän tutkielmaan liittyvä keskeinen aiempi tutkimus voidaan jaotella kolmeen kategoriaan. Ensinnäkin on tutkimuksia, joissa menetelmää on käsitelty teoriaan peilaten tai tarkasteltu sen käyttöä alan

ammattilaisten keskuudessa. Toinen keskeinen kategoria on tutkimukset, joissa vertailaan joko menetelmän erilaisia muunnelmia eli variantteja keskenään tai ääneenajattelumenetelmää verrataan johonkin toiseen käytettävyydestä käytettävään menetelmään. Lisäksi löytyy tutkimuksia, joissa menetelmää tarkastellaan jonkun käyttäjäryhmän näkökulmasta. Tällaisia ovat esimerkiksi lapset tai ihmiset, joilla on muistisairaus tai näkö- tai kuulovamma, tai eri kulttuureista tulevat käyttäjät. Yhteensä tutkielman lähteiksi valittiin 20 vertaisarvioitua lehti- ja konferenssijulkaisua.

Kirjallisuuskatsauksen tuloksena saatiin, että ääneenajattelumenetelmään liittyy neljänlaisia haasteita. Ryhmittelyn tarkoituksena on helpottaa haasteiden hahmottamista. Ensimmäisessä kategoriassa ovat menetelmän käytännön haasteet, kuten esimerkiksi se, miten osallistujat saadaan ajattelemaan ääneen ja välttämään siltä, että he sanovat asioita vain miellyttääkseen moderaattoria. Toisena kategoriana ovat teoriaan liittyvät haasteet. Näitä ovat esimerkiksi yhtenäisen teorian puute ja keskeisen, kognitiiviseen psykologiaan perustuvan teorian sopimattomuus käytettävyydestä.

Kolmantena kategoriana ovat ääneenajattelumenetelmän erilaisiin variantteihin liittyvät haasteet. Klassisen variantin haasteena on se, ettei se välttämättä sovellu käytettäväksi käytettävyydestä. Interaktiivinen variantti puolestaan voi vaikuttaa osallistujien työskentelyyn sekä korostaa moderaattorin roolia ja toimintaa. Retrospektiivisen variantin suhteen haasteina ovat aikaavievyys ja testissä ilmenneiden asioiden unohtuminen. Viimeisenä kategoriana ovat menetelmän saavutettavuuteen liittyvät haasteet. Menetelmän käytössä esimerkiksi lasten, muistisairaiden, näkö- ja kuulovammaisten tai eri kulttuurista tulevien osallistujien kanssa on tärkeää ottaa huomioon joitakin asioita, jotta menetelmä on saavutettava.

Tämä tutkielma koostuu kuudesta luvusta. Luku 1 on johdantoluku. Luvussa 2 määritellään keskeiset käsitteet, jotka auttavat pohjustamaan haasteiden käsittelyä. Luku 3 on tulosluku, jossa kuvataan tutkielman keskeiset tulokset ryhmiteltynä edellä kuvatun luokittelun mukaisesti. Luvussa 4 esitellään kehitysehdotuksia ja kootaan yhteen varianttien haasteita. Luvussa 5 pohditaan tuloksia ja luku 6 puolestaan kattaa yhteenvedon.

## **2 Käytettävyys, käytettävyydestä ja ääneenajattelu**

Tässä luvussa määritellään keskeiset tutkielmaan liittyvät käsitteet: käytettävyys, käytettävyydestä ja ääneenajattelu. Ääneenajattelua voidaan käyttää osana käytettävyydestä, ja käytettävyydestä puolestaan on yksi keskeinen tapa tutkia käytettävyyttä.

### **2.1 Käytettävyys ja käytettävyydestä**

Keskeinen käytettävyyden määritelmä on kansainvälisen standardisointijärjestön ISO:n määritelmä, jonka mukaan käytettävyys on se laajuus, missä määrin käyttäjät voivat saavuttaa tavoitteensa tuloksellisesti, tehokkaasti ja tyytyväisinä käyttäessään järjestelmää,

tuotetta tai palvelua (ISO 9241-11, 2018). On kuitenkin olemassa muitakin tapoja määrittellä käytettävyys. Nielsen (1993) esimerkiksi jakaa käytettävyyden viiteen osa-alueeseen: opittavuuteen, tehokkuuteen, muistettavuuteen, virheiden vähyyteen ja tyytyväisyyteen. Käytettävyyttä voidaan tutkia monin eri tavoin, esimerkiksi käytettävyystestauksella, haastattelulla, kyselylomakkeilla ja fokusryhmillä (Nielsen, 1993).

Yksi tapa saada tietoa käytettävyydestä on tehdä käytettävyystestausta. Käytettävyystestauksessa käyttäjä tekee tehtäviä testattavalla järjestelmällä. Sitä kautta saadaan suoraa tietoa käyttäjien ongelmista järjestelmän suhteen. Käytettävyystestaus voidaan toteuttaa laboratoriossa tai kentällä. (Nielsen, 1993.) Sitä voidaan tehdä myös etänä (Fan ym., 2020).

Nielsenin (1993) mukaan käytettävyystestaus voidaan jakaa formatiiviseen ja summatiiviseen arviointiin. Formattiivista arviointia käytetään osana iteratiivista kehitysvaihetta tarkoituksena selvittää, mikä toimii ja mikä ei, ja miten järjestelmää voidaan parantaa. Ääneenajattelu on tyypillinen menetelmä formatiivisessa arvioinnissa. Summatiivinen arviointi puolestaan keskittyy arvioimaan kokonaisvaltaisesti järjestelmän laatua, esimerkiksi silloin, kun halutaan verrata kahta järjestelmää toisiinsa. Summatiivisessa arvioinnissa tehdään tyypillisesti mittauksia, esimerkiksi mitataan tehtäviin käytettyä aikaa. Ääneenajattelua on kuitenkin toisinaan käytetty myös summatiivisessa käytettävyystestauksessa (McDonald ym., 2012).

## **2.2 Ääneenajattelu**

Nielsenin (1993) mukaan ääneenajattelu on mahdollisesti arvokkain käytettävyyden tutkimisen menetelmä. Ääneenajattelu on käytettävyystestauksessa käytetty menetelmä, jonka tarkoituksena on saada tietoa käytettävyystestauksen osallistujien ajatusprosesseista liittyen tehtävän suorittamiseen. Kyseisellä menetelmällä on tärkeä merkitys käytettävyyteen liittyvässä tiedonkeruussa. (Boren & Ramey, 2000.) Menetelmän avulla saadaan tietoa esimerkiksi järjestelmän käytettävyysongelmista (O'Brien & Wilson, 2023). Fanin ja kollegoiden (2020) kyselytutkimuksessa 86 % vastaajista sanoi käyttävänsä ääneenajattelua vetäessään käytettävyystestausta. Seuraavaksi tarkastellaan ääneenajattelu menetelmän teoriataustaa, variantteja ja käyttökontekstia.

### *2.2.1 Teoriatausta*

Ääneenajattelumenetelmän tausta on kognitiivisessa psykologiassa, ja käytettävyystestauksessakin menetelmä linkitetään useimmiten Ericssonin ja Simonin kognitiivisen psykologian protokolla-analyysiin 1980-luvun alkupuolelta. Käytettävyystutkimuksen kontekstissa kognitio itsessään ei kuitenkaan ole pääasiallinen mielenkiinnonkohde, vaan tarkoituksena on tukea järjestelmien kehittämistä löytämällä niistä puutteita. Lisäksi käytettävyystestauksessa verbalisointiin yhdistyy myös käyttäjän vuorovaikutus testattavan järjestelmän kanssa. (Boren & Ramey, 2000.)

Jos käytettävyytutkimuksen yhteydessä viitataan johonkin teoriaan, on kyseessä siis Borenin ja Rameyn (2000) mukaan yleensä Ericssonin ja Simonin malli. Malli on edelleenkin laajasti viitattu (O'Brien & Wilson, 2023). Kyseistä alkuperäislähdettä ei ollut saatavilla tutkielmaa tehdessä, joten sitä käsitellään muiden tutkimuksien kautta.

Boren ja Ramey (2000) avaavat tutkimuksessaan Ericssonin ja Simonin mallia, jossa keskeisenä käsitteenä on lyhytkestoinen muisti. Ericssonin ja Simonin mallissa on kolme verbalisoinnin tasoa. Ensimmäisessä tasossa ovat sellaiset ajatukset, jotka eivät vaadi muokkaamista ennen kuin ne sanallistaa, kuten numerosarjojen ääneen sanominen matemaattista tehtävää ratkaistaessa. Toisessa tasossa ovat sellaiset ajatukset, joita pitää muokata ennen sanoiksi pukemista, kuten mielikuvat ja abstraktit ajatukset. Kolmannessa tasossa puolestaan ovat sellaiset asiat, jotka vaativat enemmän kognitiivista prosessointia kuin tarvittaisiin tehtävien tekoon tai sanallistamiseen. Kolmannen tason verbalisointia voi syntyä esimerkiksi, jos henkilö suodattaa ajatuksiaan mainitsemalla vain tiettyyn asiaan liittyviä seikkoja tai jos henkilö hakee asioita pitkäkestoisesta muistista. Myös ulkoinen vaikutus, kuten moderaattorin kommentti, muuttaa verbalisoinnin kolmannen tason dataksi.

Kahden ensimmäisen tason mukaan tuotettu ääneenajattelun data on Ericssonin ja Simonin mallissa luotettavaa, mutta kolmannen tason ei. Ericssonin ja Simonin mallissa ääneenajattelun dataksi eivät myöskään sovellu esimerkiksi tunteet ja arvioinnit. (Boren & Ramey, 2000.)

Boren ja Ramey (2000) korostavat neljä keskeistä asiaa, joita Ericssonin ja Simonin mallista pitää huomioida, jos sitä haluaa soveltaa käytettävyyttestauksessa. Ensinnäkin moderaattorin pitäisi kerätä ainoastaan tasojen 1 ja 2 mukaista dataa. Toisekseen moderaattorin pitäisi antaa osallistujalle yksityiskohtaiset ohjeet ennen ääneenajattelun aloittamista, ja lisäksi osallistujien pitäisi saada harjoitella ennen varsinaisen käytettävyyttestin aloittamista. Kolmanneksi moderaattorin pitäisi muistuttaa osallistujaa ajattelemaan ääneen, jos osallistuja on hiljaa jonkun ennalta määrätyn ajan. Muistutuksen pitäisi olla mahdollisimman lyhyt ja sellainen, joka ei ohjaa käyttäjää. Ericsson ja Simon suosittelevat muistutukseksi esimerkiksi ”Jatka puhumista” (engl. ”Keep talking”). Neljänneksi moderaattorin ei pitäisi muutoin keskeyttää mitenkään, esimerkiksi esittämällä kysymyksiä tai kommentteja, sillä ne pilaavat ääneenajattelusta syntyvän aineiston.

### 2.2.2 *Variantit*

Ääneenajattelusta on erilaisia variantteja, jotka eroavat toisistaan toimintatapojen suhteen. Ensinnäkin menetelmä voidaan jakaa samanaikaiseen (engl. concurrent) ja retrospektiiviseen (engl. retrospective) ääneenajatteluun. Samanaikaisessa ääneenajattelussa käyttäjät ajattelevat ääneen tehdessään tehtäviä, kun taas retrospektiivisessä ääneenajattelussa käyttäjät ajattelevat ääneen jälkikäteen, tavallisesti katsoessaan videota testisessiossa (Fan ym., 2020). McDonald ja kollegat (2012) selvittivät kyselytutkimuksella, että

samanaikainen ääneenajattelu on alan ammattilaisten keskuudessa suosituin variantti, sillä sitä käytti 89 % vastanneista. Seuraavaksi suosituin oli retrospektiivinen ääneenajattelu, jota käytti 5 % vastaajista. Myös Fanin ja kollegoiden (2020) kyselytutkimuksessa samanaikainen ääneenajattelu osoittautui suositummaksi, sillä 61 % vastaajista käytti sitä lähes jokaisessa vetämässään käytettävyytestissä ja 91 % vastaajista yli puolessa käytettävyytesteistä. Retrospektiivisen ääneenajattelun kohdalla puolestaan vain 21 % ammattilaisista käytti sitä lähes jokaisessa vetämässään käytettävyytestissä, kun taas 61 % käytti sitä silloin tällöin tai vain harvoin.

Myös samanaikainen ääneenajattelu voidaan jakaa erilaisiin variantteihin. Yllä kuvattiin klassinen Ericssonin ja Simonin malli. Kuitenkin sen käytöstä poiketaan usein, jolloin käytetään esimerkiksi interaktiivista (rentoutunutta) ääneenajattelua, jossa moderaattori keskeyttää osallistujaa esimerkiksi kysyäksään kysymyksiä (O'Brien & Wilson, 2023). Interaktiivinen ääneenajattelu tuottaa Ericssonin ja Simonin mallin kolmannen tason aineistoa (Hertzum ym., 2009).

Olmsted-Hawala ja kollegat (2010) puolestaan nimeävät kolme eri variaation lähdettä menetelmän käytössä. Ensinnäkin ammattilaisten keskuudessa on variaatiota siinä, kuinka paljon osallistujia ohjeistetaan etukäteen ja annetaanko heidän harjoitella. Toisekseen variaatiota löytyy moderaattorien keskeytysten suhteen. Tämän lisäksi vaihtelua on myös siinä, annetaanko osallistujan ollessa hiljaa kehoitus jonkun tietyn ajan kuluttua vai satunnaisesti.

McDonaldin ja kollegoiden (2012) kyselytutkimuksen perusteella moni ammattilainen ei kuitenkaan aina noudata samaa menettelytapaa, vaan mukauttaa sitä esimerkiksi osallistujien ja tilanteen mukaan. Variaatio siinä, miten ääneenajattelumenetelmää käytetään, voi johtua myös osittain siitä, että käytettävyytestaukseen liittyvässä kirjallisuudessa on monenlaista neuvoa siitä, miten menetelmää sovelletaan (McDonald ym., 2012). Variantteihin palataan myöhemmin, kun luvussa 3.2 käydään läpi klassisen teorian haasteita ja luvussa 3.3 laajemmin varianttien haasteita.

### *2.2.3 Käyttökonteksti*

Käytettävyytestauksen yhteydessä ääneenajattelua voidaan käyttää kontrolloidussa laboratoriossa, kentällä käyttäjän ympäristössä (engl. field study) sekä etänä toteutettavassa käytettävyytestauksessa. Fanin ja kollegoiden (2020) kyselytutkimuksessa 75 % vastaajista käytti ääneenajattelua laboratoriotestauksessa, 72 % käytti sitä etätestauksessa ja 48 % käytti sitä kenttätestauksessa. Moderaattori on usein paikalla käytettävyytestissä. On kuitenkin mahdollista, että moderaattori ei ole paikalla, vaan käyttäjä on yksin. (Riihiho, 2015.)

Luvussa 2.1 kuvattiin formatiivisen ja summatiivisen arvioinnin erilaiset tavoitteet ja ajankohta tuotekehitysprosessissa. McDonaldin ja kollegoiden (2012) kyselytutkimuksessa 86 % vastanneista alan ammattilaisista sanoi käyttävänsä ääneenajattelua formatiivisessa käytettävyydestestauksessa ja 67 % summatiivisessa käytettävyydestestauksessa.

Ääneenajattelua voidaan käyttää sekä tieteelliseen tutkimukseen liittyvässä käytettävyydestestauksessa että käytännön käytettävyydestestauksessa esimerkiksi yrityksissä (McDonald ym., 2012). Tällä voi myös olla vaikutusta menetelmän käyttöön. McDonald ja kollegat (2012) havaitsivat kyselytutkimuksensa perusteella, että akateemisen tutkimuksen parissa työskentelevät vastaajat käyttivät ääneenajattelumenetelmää enemmän summatiivisessa käytettävyydestestauksessa kuin käytännön työssä olevat.

### **3 Ääneenajattelumenetelmän haasteet**

Tässä luvussa kuvataan kirjallisuudesta esiin nousseita haasteita liittyen ääneenajattelumenetelmän käyttöön käytettävyydestestauksessa. Ensin käsitellään ääneenajattelun haasteita liittyen käytettävyydestien toteutukseen ja analysointiin yleisellä tasolla. Seuraavaksi perehdytään menetelmän teoriaan liittyviin haasteisiin. Tämän jälkeen käsitellään ääneenajattelun eri varianttien haasteita. Lopuksi nostetaan esiin saavutettavuuteen liittyviä haasteita menetelmän käytössä.

#### **3.1 Käytännön haasteet**

Tutkimuksessaan Fan ja muut (2020) selvittivät, miten ja missä määrin ammattilaiset käyttävät ääneenajattelumenetelmää, ja mitä haasteita he sen käytössä kokevat. He teettivät kyselyn lähes kahdellasadalla UX-asiantuntijalla, jotka työskentelivät eri puolilla maailmaa ja erikokoisissa yrityksissä. Vastaajista 86 % käytti ääneenajattelua tehdessään käytettävyydestestausta. Tutkimuksessa havaittiin haasteita liittyen sekä menetelmän käyttöön käytettävyydestestauksessa että käytettävyydestien analysointiin.

Ensimmäinen testaukseen toteuttamiseen liittyvä haaste koski Fanin ja kollegoiden (2020) mukaan sitä, miten osallistujat saadaan ajattelemaan ääneen. Tähän voi vaikuttaa moni asia, kuten persoonallisuus ja tehtävän tyyppi. Myös esimerkiksi se, että osallistujina on lapsia, voi vaikeuttaa menetelmän käyttöä. Toisena haasteena on heidän mukaansa se, miten saadaan luotua sellainen ympäristö, että osallistujat uskaltavat ajatella rehellisesti ääneen, eivätkä sano asioita miellyttääkseen testaajia. Kolmantena haasteena on se, kuinka moderaattorin pitäisi toimia keskeyttämisen suhteen. Moderaattorin pitää varoa keskeyttämästä liian pian, mutta toisaalta se, että antaa käyttäjän jatkaa liian pitkään, voi turhauttaa käyttäjää ja vaikuttaa lopputestauksen ajan.

Samoja asioita nousi esiin McDonaldin ja kollegoiden (2012) kyselytutkimuksessa, jossa oli vastaajina 207 alan ammattilaista. Yhden vastaajan mukaan ääneenajattelun käyttö lisää ahdistusta osallistujissa. Toinen ei pitänyt ääneenajattelun tunkeilevasta

(engl. inquisitive) luonteesta, vaan suosi pareittain tehtävää käytettävyydestä (engl. constructive interaction), joka esitellään lyhyesti luvussa 4.1. Se tuntui vastaajasta luontevammalta eikä hänen mukaansa vääristä dataa niin paljon, sillä osallistujat eivät ole yhtä tietoisia tarkkailusta verrattuna yksin tehtävään ääneenajatteluun. Niin ikään Nielsen (1993) toteaa, että ääneenajattelu voi tuntua useimmista ihmisistä epäluonnolliselta, mikä voi vaikeuttaa testin vetämistä sekä vaikuttaa tuloksiin.

Osallistujien pyrkimys miellyttää moderaattoria nousi esiin myös McDonaldin ja kollegoiden (2012) kyselytutkimuksessa. Yksi heidän kyselytutkimukseensa vastanneista kommentoi, ettei kerro osallistujille, mistä asioista on kiinnostunut, sillä muuten osallistujat vastaavat yrittäen miellyttää moderaattoria. McDonaldin ja kollegoiden mukaan tietynlainen ohjeistus voikin vaikuttaa ääneenajattelusta syntyvän aineiston luotettavuuteen: ohjeidenannon suhteen pitää siis olla tarkkana. Moderaattorit voivat siis ohjeistaa käyttäjiä eri tavalla. Hertzumin (2016) mukaan ei ole usein selvää, miten moderaattorit ohjeistavat käyttäjiä ääneenajattelussa käytännön testauksessa.

Fanin ja kollegoiden (2020) mukaan ääneenajattelua käyttävien testien analysoimiseen liittyväksi haasteeksi puolestaan koettiin testistä nauhoitetun videon analysoinnin vaivalloisuus ja aikaavievyys, minkä takia vastaajat päätyivät usein vain luottamaan muistiinsa tai muistiinpanoihinsa. Tämä voi tutkijoiden mukaan kuitenkin johtaa puutteelliseen analyysiin. Voi myös olla, että analyysin helpottamiseksi moderaattorit keskeyttävät osallistujia enemmän, jotta saisivat täsmennyksiä heidän sanomisiinsa (McDonald ym., 2012). Osallistujien keskeyttämisestä voi kuitenkin seurata omia haasteitaan, joita käsitellään tarkemmin luvussa 3.3.

Ääneenajattelun yhtenä haasteena on myös se, että se ei välttämättä sovi kaikkiin tilanteisiin. Hertzumin (2016) mukaan ääneenajattelu sopii paremmin laboratorio-olosuhteisiin kuin kenttätestaukseen, sillä esimerkiksi julkisella paikalla tehtävä ääneenajattelu voi tuntua kohtuuttoman luonnottomalta ja raskaalta.

McDonaldin ja kollegoiden (2012) kyselytutkimuksessa nousee esiin syitä, miksi alan ammattilaiset eivät käytä ääneenajattelua toteuttaessaan käytettävyydestä. Esimerkiksi yksi vastaaja sanoi, ettei käytä ääneenajattelua summatiivisessa käytettävyydestä. Kuten luvussa 2.1 mainittiin, summatiivisessa käytettävyydestä mitataan usein erilaisia asioita. Tähän viittaa myös toinen vastaaja, joka sanoi, ettei käyttäisi ääneenajattelua, kun tehtävien suoritusaikaa mitataan. (McDonald ym., 2012.) Ääneenajattelumenetelmä siis vääristää tuloksia, jos halutaan esimerkiksi mitata tehtävään käytettyä aikaa.

Ääneenajattelun vaikutus tehtävien tekoon havaittiin myös Riihiahon (2015) tutkimuksessa, jossa testattiin ääneenajattelun ja moderaattorin läsnäolon vaikutusta. Osallistujat jaettiin neljään ryhmään, jotka erosivat sen suhteen, ajattelivatko he ääneen vai työs-

kentelivätkö he hiljaa sekä sen suhteen, työskentelivätkö he yksin vai moderaattorin seurassa. Tuloksena saatiin, että ääneenajattelu hidasti tehtävien suorittamista verrattuna hiljaiseen työskentelyyn. Myös Hertzum ja kollegat (2009) havaitsivat, että ääneenajattellessa tehtävien suoritus aika on suurempi hiljaiseen työskentelyyn verrattuna.

### **3.2 Teoriaan liittyvät haasteet**

Ääneenajattelumenetelmään liittyy haasteita myös teorian suhteen. Tällaisia ovat muun muassa teorian puute ja epäyhtenäisyys toimintatapojen suhteen, mikä näkyy esimerkiksi vaihtelevana ohjeistuksena.

Teoriaan liittyvänä haasteena ääneenajattelumenetelmää käsittelevissä ja hyödyntävissä tutkimuksissa voidaan pitää sitä, että menetelmää ei kytketä mihinkään teoriaan. Boren ja Ramey (2000) havaitsivat, että usein ei viitata mihinkään lähteeseen ääneenajattelumenetelmän käytön perustelemiseksi. Jos kuitenkin viitataan teoriaan, on se heidän mukaansa lähes aina Ericssonin ja Simonin klassinen malli, joka esiteltiin luvussa 2.3.1. O'Brienin ja Wilsonin (2023) mukaan malli on tänä päivänä edelleen laajasti viitattu.

Borenin ja Rameyn (2000) mukaan ei kuitenkaan ole kunnolla tutkittu sitä, sopivatko kognitiivisen psykologian verbalisaatiomallit käytettävyydestänsä. Myös O'Brien ja Wilson (2023) huomauttavat, että Ericssonin ja Simonin malli pohjautuu kognitiivisen psykologian suhteellisen kapealle alalle.

Sekä käytettävyydestänsä liittyvässä kirjallisuudessa että ammattilaisten toiminnassa esiintyy poikkeamia klassiseen teoriaan nähden. Havaitut poikkeamat teoriasta kuitenkin vaihtelevat ammattilaisten keskuudessa eivätkä usein kytkeydy mihinkään teoriaan. Esimerkkeinä poikkeamista teoriasta on, että moni ammattilainen ei ohjeista osallistujaa asiaankuuluvalla tavalla ennen testauksen aloittamista tai sen aikana. Lisäksi osallistujia keskeytetään epä johdonmukaisesti. (Boren ja Ramey, 2000.)

Samoin Fan ja kollegat (2020) havaitsivat kyselytutkimuksessaan, että ääneenajattelumenetelmää käytetään monella eri tavalla. Esimerkiksi toisin kuin klassisessa teoriassa, harva käytettävyydestänsä pyytää osallistujaa harjoittelemaan ääneenajattelua ennen varsinaista käytettävyydestänsä. Testaajat myös ohjeistavat käyttäjiä sanomaan ääneen muitakin asioita kuin alkuperäisen teorian mukaan pitäisi, kuten tunteitaan, palautetta ja suunnittelusuosituksia.

Voikin olla, että kaikissa tilanteissa klassinen teoria ei sovellu käytettävyydestänsä. Boren ja Ramey (2000) huomauttavatkin, että käytettävyydestänsä voi ilmetä tilanteita, joita kognitiiviseen psykologiaan pohjautuva klassinen teoria ei käsittele. Tällainen on esimerkiksi järjestelmän kaatumisen, jolloin moderaattori joutuu puuttumaan tilanteeseen, minkä jälkeen teorian mukaisen monologin jatkaminen voi tuottaa osallistujalle vaikeuksia. Lisäksi testaajat voivat olla kiinnostuneita saamaan klassisen teorian mukaan kiellettyä aineistoa, kuten osallistujan mielipiteitä, ideoita ja tunteita. Voi myös

olla, että klassisen teorian mukainen menettely on liian monimutkainen testaa-  
jien tarpeisiin, joten he yksinkertaistavat sitä. (Boren & Ramey, 2000.)

On kuitenkin hyvä huomioda, että haasteena ei niinkään suoraan voida pitää sitä, että klassista teoriaa ei noudateta, sillä kuten todettua, se ei välttämättä sovellu aina käytettävyydestä. Pikemminkin haasteena on siis teorian puute ja vaihtelu menettelytavoissa. Menetelmän käyttöön liittyvä suuri variaatio voi vaikeuttaa tulosten vertailua sekä tieteellisten tutkimusten suhteen että yritysmaailmassa (Boren & Ramey, 2000). Vaihteluus menetelmän käytössä voi johtaa erilaisiin ongelmiin myös reliabiliteetin suhteen (Olmsted-Hawala ym., 2010). Jos Ericssonin ja Simonin klassinen teoria ei sovellu käytettävyydestä, pitäisi miettiä minkälainen teoria soveltuisi. Boren ja Ramey (2000) korostavatkin teorian tärkeyttä, sillä jos ääneenajattelua ei käytetä systemaattisesti ja täsmällisellä tavalla, sitä ei heidän mukaansa voi kutsua metodiksi.

Erityisesti teorian tärkeys korostuu tieteellisessä tutkimuksessa esimerkiksi tutkimusten vertailtavuuden ja toistettavuuden suhteen. Kuitenkin, kuten Boren ja Ramey (2000) mainitsevat, vertailtavuus on tärkeää myös käytännön käytettävyydestä. Lisäksi heidän mukaansa teoriaan pohjautuvat ja systemaattisesti käytettävät menetelmät mahdollistaisivat vakiintuneiden toimintatapojen opettamisen alan uusille tulokkaille.

Teoriaan liittyviä ongelmia on kuitenkin myös klassiselle mallille vaihtoehtoisissa toimintatavoissa. Moni käyttää interaktiivista ääneenajattelua, jolla voidaan saada kattavampaa kolmannen tason aineistoa (O'Brien & Wilson, 2023). Alan kirjallisuudesta löytyy ohjeita tämän toteuttamiseen, mutta interaktiiviselle ääneenajattelulle ei ole vakiintunut toimintatapaa, ja lisäksi kirjallisuudessa annetaan ristiriitaisia ohjeita (McDonald ym., 2016). Myös Boren ja Ramey (2000) korostavat, että jos käytettävyydestä halutaan poiketa Ericssonin ja Simonin klassisesta teoriasta, esimerkiksi saadakseen mahdollisesti kiinnostavampaa aineistoa, pitäisi tällaisenkin datan keräämisen perustua johonkin teoriaan.

### **3.3 Erilaisten varianttien haasteet**

Ääneenajattelumenetelmää käytetään siis monella eri tavalla. Tässä luvussa perehdytään klassisen, interaktiivisen ja retrospektiivisen variantin haasteisiin.

Kaksi keskeistä samanaikaisen ääneenajattelun muotoa ovat siis Simonin ja Ericssonin klassisen teorian mukainen ääneenajattelu sekä interaktiivinen ääneenajattelu, jota voidaan kutsua myös rentoutuneeksi ääneenajatteluksi. Interaktiivisessa ääneenajattelussa moderaattori keskeyttää osallistujaa esimerkiksi kysyäkseen kysymyksiä (O'Brien & Wilson, 2023). Luvussa 3.2 on jo tuotu esiin, kuinka klassisen mallin keskeisenä haasteena on, ettei se välttämättä sovi käytettävyydestä. Voikin olla, että moderaattorit pyrkivät tarkoituksella keskeytyksien avulla luomaan kolmannen tason aineistoa, koska kokevat sen hyödyllisemmäksi (Boren & Ramey, 2000). Klassisen teorian mukainen ää-

neenajattelu voi johtaa aineistoon, joka on epämääräistä ja vaikeaa tulkita, joten interaktiivista ääneenajattelua saatetaan käyttää, koska sen koetaan tuottavan erilaista ja hyödyllisempää aineistoa (O'Brien & Wilson, 2023).

Interaktiivisessa ääneenajattelussa on kuitenkin omat haasteensa. Moderaattorin keskeytykset interaktiivisessa ääneenajattelussa voivat sisältävää enemmän riskejä kuin hyötyjä, sillä ne voivat esimerkiksi vaikuttaa aineiston reliabiliteettiin ja validiteettiin (McDonald ym., 2012). Myös O'Brien ja Wilson (2023) toteavat interaktiivisen ääneenajattelun aiheuttavan haasteita validiteetin kanssa.

Interaktiivinen ääneenajattelu vaikuttaa tehtävissä suoriutumiseen (O'Brien & Wilson, 2023). Tämä tuli ilmi Alhadretin ja Mayhew'n (2017) tutkimuksessa, jossa testattiin kolmea erilaista ääneenajattelun varianttia. Yksi näistä oli klassinen Ericssonin ja Simonin teoriaan pohjautuva versio, jossa moderaattori ainoastaan muistuttaa, jos osallistuja on hiljaa 15 sekuntia. Toinen oli Borenin ja Rameyn (2000) ehdottama puheviestintäteoria, jossa moderaattori äänтелеe myötäilevästi ("Mm hmm"). Kolmas variantti oli interaktiivinen ääneenajattelu, jossa moderaattori aktiivisesti keskeytti. Jokaisessa heillä oli 20 osallistujaa. Tuloksina saatiin esimerkiksi, että interaktiivisessa ääneenajattelussa tehtävien suorittaminen oli hitaampaa, ja hiiren painalluksien sekä katsottujen sivujen määrä oli suurempi. Osallistujat myös kokivat tilanteen häiritsevämmäksi kuin kahdessa muussa versiossa olleet osallistujat. Kuitenkin on huomioitava, että löydettyjen käytettävyysongelmiin määrässä ei ollut tilastollisesti merkitsevää eroa eri varianttien välillä.

Myös Hertzumin ja kollegoiden (2009) tutkimuksessa havaittiin interaktiivisen ääneenajattelun suurempi vaikutus osallistujien käytökseen. He vertasivat sekä klassista että interaktiivista ääneenajattelua hiljaiseen työskentelyyn, kun osallistujat käyttivät neljää nettisivua. Interaktiivinen ääneenajattelu vaikutti osallistujien käytökseen monella tavalla ja johti suurempaan henkiseen kuormittuneisuuteen (engl. mental workload) kuin klassinen ääneenajattelu tai hiljainen työskentely. Interaktiivisessa ääneenajattelussa käyttäjät tekivät merkitsevästi enemmän painalluksia ja vierityksiä hiirellä kuin hiljaisessa työskentelyssä. Myös klassisessa ääneenajattelussa hiiren painalluksien ja kirjoittamisen määrä oli hieman suurempi kuin hiljaisessa työskentelyssä, mutta ei niin merkitsevästi. Tähän liittyen Hertzum ja muut tulkitsivat, että interaktiivinen ääneenajattelu muutti osallistujien tapaa suorittaa tehtäviä, toisin kuin klassinen ääneenajattelu, joka pääosin vain hidasti suoritusta. Yhtenä selityksenä tähän he tarjosivat sen, että interaktiivinen ääneenajattelu johti siihen, että erityisesti tason 3 verbalisaatiot häiritsivät osallistujien ajatusprosesseja, jonka takia heidän täytyi toistaa järjestelmässä tekemiään toimia. Heidän mukaansa kolmannen tason verbalisaatio vaikuttaa selvästi järjestelmän käyttöön, toisin kuin ensimmäisen ja toisen tason verbalisaatio, joilla voi olla rajallinen vaikutus.

O'Brien ja Wilson (2023) haastattelivat ammattilaisia, jotka käyttävät interaktiivista ääneenajattelua toteuttaessaan käytännön käytettävyydestä. Tutkijat halusivat selvittää, mitä haasteita ja hyötyjä ammattilaiset interaktiivisessa ääneenajattelussa näkevät. Haastateltavat tunnistivat interaktiivisen ääneenajattelun haasteeksi esimerkiksi sen, että osallistuja alkaa arvostelemaan ja ylianalysoimaan järjestelmää sen sijaan, että käyttäisi sitä. Ammattilaisten mielestä on myös mahdollista, että moderaattorin liialliset tai huonot kysymykset tai epäsoviva äänensävy voivat saada käyttäjät kokemaan, että heitä kuultellaan tai testataan. Tutkijoiden mielestä onkin mahdollista, että interaktiivinen ääneenajattelu korostaa moderaattorin ja osallistujan välistä vallan epätasapainoa verrattuna klassiseen ääneenajatteluun. Interaktiivinen ääneenajattelu on kuitenkin suosittu menetelmä käytännön työssä. Voi olla, että interaktiivista ääneenajattelua ei ole tutkittu tarpeeksi erityisesti käytännön kontekstissa, menetelmää oikeasti käyttävien ammattilaisten parissa. (O'Brien & Wilson, 2023.)

Yksi vaihtoehto samanaikaiselle menetelmälle on myös retrospektiivinen ääneenajattelu, jossa osallistajat ajattelevat äänen jälkikäteen tehtävät tehtyään. Hertzum (2016) toteaa, että retrospektiivinen ääneenajattelu samalla, kun osallistuja katsoo videota testauksesta, voi olla toimiva ratkaisu. Se mahdollistaa järjestelmän käyttämisen ja kommentoinnin erottamisen, jolloin ääneenajattelu ei vaikuta osallistujan käytökseen tehtäviä tehdessä. Lisäksi se voi helpottaa kenttätestauksissa, kun osallistujan ei tarvitse tehdä ääneenajattelua kentällä. Aiemmin tässä tutkielmassa nousi myös esiin, kuinka tehtävien suorittamiseen kuluvan ajan mittaaminen on samanaikaisen menetelmän haaste. Retrospektiivisen ääneenajattelun kanssa tätä haastetta ei ole.

Tässäkin variantissa on omat haasteensa. Alhadreti ja Mayhew (2018) vertaisivat samanaikaista ääneenajattelua, retrospektiivistä ääneenajattelua ja hybridimenetelmää, joka oli kahden edellä mainitun yhdistelmä. Osallistujia oli 20 jokaisessa variantissa. Retrospektiivistä menetelmää käyttäneet löysivät vähiten käytettävyysongelmia. Tutkijat pohivat selitykseksi sitä, että osallistujat saattoivat unohtaa mainita ongelmista, vaikka olisivatkin havainneet ne tehtäviä tehdessään. Lisäksi tutkimuksessa havaittiin, että retrospektiivinen ja hybridimenetelmä veivät noin kaksikertaisesti enemmän aikaa samanaikaiseen menetelmään verrattuna. Myös McDonaldin ja kollegoiden (2012) kyselytutkimuksessa vastaajat mainitsevat retrospektiivisen ääneenajattelun aikaavievyyden sen haittapuolena.

### **3.4 Haasteet liittyen saavutettavuuteen**

Yhtenä ääneenajattelumenetelmän haasteena voidaan nähdä se, että se ei välttämättä sovellu kaikkiin tilanteisiin. Tässä luvussa käsitellään menetelmän haasteita saavutettavuuden näkökulmasta, eli kuinka toimiva menetelmä on tietyillä käyttäjäryhmillä sekä eri kulttuureissa. Menetelmän saavutettavuutta tarkastellaan lasten, muistisairaiden, näkö- ja kuulovammaisten sekä eri kulttuurien näkökulmasta.

### 3.4.1 Lapset ja ikääntyneet

Gibson ja kollegat (2016) havaitsivat, että ääneenajattelu ei toiminut dementiaa sairastavien ihmisten kohdalla. Tutkimuksessa dementiaa sairastavat ihmiset käyttivät sovellusta, ja samalla arvioitiin erilaisten käytettävyyden arvioinnin menetelmien toimivuutta käytettävyydestestauksessa. Tutkijat havaitsivat, että osallistujilla oli vaikeuksia verbalisoida ja selostaa tekemisiään. Tähän eivät auttaneet myöskään moderaattorin kehotukset ja muistutukset. Ääneenajattelu dementiaa sairastavien kohdalla ei tuottanut tutkimuksessa luotettavia tuloksia. Muistisairauksista valtaosa on ikäihmisillä (Muistisairaudet: Käypä hoito -suositus, 2023), ja ikääntyvien määrä Suomessakin kasvaa, joten tämä voi olla hyvä ottaa huomioon erityisesti, kun käytettävyydestestauksessa on osallistujina ikäihmisiä.

Toisaalta Wood ja kollegat (2021) havaitsivat erään ääneenajattelun sovelluksen toimivan hyvin muistisairaiden kanssa. He testasivat etänä tehtävää summatiivista käytettävyydestausta muistisairaiden kanssa, joiden dementia oli lievä tai kohtalainen. Osallistujilla ei ollut testauksessa avustajia. Osallistujilla oli käytettävyydestestauksessa mahdollisuus käyttää sellaista ääneenajattelun sovellusta, jossa he ohjasivat moderaattoria ääneenajattelemalla. Moderaattori käytti järjestelmää, jonka oli jakanut osallistujalle ruudunjaon avulla. Neljä kymmenestä osallistujasta käytti tätä menetelmää. Kokemukset olivat positiivisia. Ääneenajattelumenetelmän soveltaminen tällä tavalla lisäsi myös yhdessä työskentelyn tunnetta. Toki tämänkaltainen ääneenajattelu poikkeaa tavallisesta ääneenajattelusta, mutta se kuvastaa, että ääneenajattelua voidaan mahdollisesti muokata saavutettavammaksi, jolloin sitä voidaan helpommin käyttää muistisairaiden kanssa.

On ymmärrettävää, että ääneenajattelua ei voi käyttää pienten lasten kanssa. Esimerkiksi Hanna ja kollegat (1997) toteavat, että nuorilla lapsilla on aikuisia enemmän vaikeuksia verbalisoida ajatuksiaan. Kuitenkin heidän mukaansa 11–14-vuotiaiden lasten ikäryhmässä ääneenajattelua voi olla mahdollista käyttää, erityisesti ikähaarukan vanhimpien kanssa. Jotkut lapset voivat kuitenkin tuntea vaivautuneisuutta siitä, että heitä katsellaan ja kuunnellaan. Donker ja Reitsma (2004) havaitsivat, että 70:stä 6–7-vuotiaasta lapsesta vain 28 ajatteli ääneen, vaikka lapsia oli ohjeistettu siihen. Tosin tähän vaikutti varmasti se, että moderaattori ei testin aika kehottanut lapsia ääneenajattelemaan, jolloin ääneenajattelu oli täysin vapaaehtoista. Lapset olivat myös melko nuoria.

Kuitenkin tutkimusten perusteella ääneenajattelua voi käyttää lasten kanssa ja siitä on hyötyä. Kuten muistisairaiden, myös lasten kohdalla menetelmän mukauttaminen voi auttaa. Lapsien kohdalla voidaan mukauttaa ääneenajattelua niin, että lapsia pyydetään sanomaan ääneen tekemisiään ajatusten sijaan (ääneenjuttelu, engl. talk aloud) (Donker & Reitsma, 2004).

Donker ja Markopoulos (2001) havaitsivat, että lasten suuri motivaatio kompensoi ääneenajattelun vaikeuksia, sillä lapsilla oli tapana yrittää parhaansa, vaikka he olisivat

kokeneet sen haastavaksi. He testasivat ääneenjuttelua 8–14-vuotiailla lapsilla ja vertasivat sitä strukturoituun haastatteluun ja kirjalliseen kyselyyn. Ääneenajattelun avulla löydettiin eniten käytettävyysoongelmia. Myös Donkerin ja Reitsman (2004) tutkimuksessa ääneenjuttelun avulla löydettiin tärkeitä käytettävyysoongelmia, ja lisäksi lapset ilmaisivat myös mielipiteitään.

#### *3.4.2 Kuulo- ja näkövammaiset*

Ääneenajattelumenetelmän käyttöä on tutkittu myös kuulovammaisten ihmisten kanssa. Chenin ja kollegoiden (2023) tutkimuksessa testattiin ääneenajattelun käyttöä etänä tehtävässä käytettävyysongelmien testauksessa. Osallistujina oli 12 yliopisto-opiskelijaa, jotka olivat kuuroja tai huonokuuloisia. Tutkimuksessa havaittiin, että osallistujien viittomat ja navigointi saattoivat olla eri tahdissa. Lisäksi osallistujat käyttivät usein visuaalisesti kuvailevia viittomia eivätkä käyttöliittymässä esiintyviä eksplisiittisiä termejä. Nämä voivat johtaa haasteisiin yksiselitteisen aineiston keräämisessä. Chenin ja muiden mukaan onkin suositeltavaa, että tutkijat ymmärtävät kuurojen ja huonokuuloisten käyttämiä eri kommunikointitapoja ja ottavat huomioon niiden ajalliset ja tilalliset riippuvuudet.

Niin ikään Dingman ja kollegat (2021) havaitsivat jaksollisuutta kuurojen ja huonokuuloisten ääneenajattelussa. Heidän tutkimuksessaan osallistujat ääneenajattelivat käyttäessään podcast-alustan prototyyppejä. Osallistujat, jotka käyttivät viittomakieltä, viittosivat ensin sen, mitä halusivat tuoda esiin, ja tämän jälkeen aloittivat tehtävien tekemisen. Tämä toimintatapa antoi tietoa siitä, mitä osallistuja aikoi tehdä, mutta se ei paljastanut lennossa ilmeneviä ajatuksia, jolloin ääneenajattelumenetelmän jotkut hyödyt jäivät saamatta.

Strainin ja kollegoiden (2007) mukaan näkövammaisten käyttäjien kohdalla yhdeksi haasteeksi nousee ruudunlukijan audio, jolloin ääneenajattelun käyttö voi olla rajoitunutta. Yhtäaikainen ruudunlukijan kuunteleminen sekä ääneenajattelu voi olla haastavaa. Lisäksi Strain ja kollegat havaitsivat, että käytettävyydestin tarkkailijoille oli usein vaikea pysyä mukana ruudunlukijan mahdollisesti puhetta paljon nopeammassa tahdissa.

Strain ja kollegat ehdottavat kolmea erilaista ääneenajattelun muunnelmaa, joita he testasivat 10 näkövammaisella käyttäjällä. Ensimmäisenä menetelmänä he testasivat ruudunlukijan pysäyttämistä ääneenajattelun ajaksi. He havaitsivat, että tämä toimi erityisesti osallistujilla, joille ääneenajattelu ja ruudunlukijan käyttö oli vaivatonta, ja kun tehtävät eivät olleet kognitiivisesti vaativia. Menetelmän haittapuolena oli tehtävien teon luonnollisen rytmin keskeytyminen. Toisena tapana oli retrospektiivinen ääneenajattelu, jossa osallistujat ajattelivat ääneen jokaisen tehtävän jälkeen. Tätä tapaa käytettiin erityisesti, jos osallistuja sai tehtävän tehtyä erittäin nopeasti tai toisaalta, jos tehtävän tekeminen vaati paljon ponnistelua sivulla, jonka saavutettavuus oli huono ja osallistujien piti keskittyä ruudunlukijan audioon. Kolmantena tapana oli muokattu retrospektiivinen ää-

neenajattelu, jossa tehtävän teon jälkeen moderaattori pyysi osallistujaa käymään käyttöliittymä läpi ja selittämään mitä heidän mielestään tapahtui. Moderaattori saattoi halutesaan tiedustella lisätietoa. (Strain ym., 2007.)

### *3.4.3 Kulttuuriset tekijät*

Kulttuuriset tekijät voivat vaikuttaa ääneenajatteluun, mikä voi globaalissa maailmassa olla tärkeää huomioida. Ensinnäkin moderaattorin ja osallistujien ominaisuudet pitää ottaa huomioon. Clemmensenin (2011) tutkimuksessa intialaiset moderaattorit nostivat esiin ikäerot: maaseudulta tulevaa nuorta osallistujaa voi pelottaa vanhempi moderaattori. Tämän voisi olettaa vaikeuttavan ajatusten sanomista ääneen rehellisesti.

Kulttuuri voi näkyä moderaattorien ja osallistujien erilaisissa toimintatavoissa. Clemmensen (2011) havainnoi kahden muun tutkijan kanssa käytettävyydestä Intiassa, Kiinassa ja Tanskassa. He havaitsivat, että Tanskassa osallistujat ajattelivat aktiivisesti ääneen, ja tanskalaiset moderaattorit eivät montaa kertaa muistuttaneet ääneenajattelusta. Osallistujilla oli tapana antaa mielipiteitään ja kehitysehdotuksia testattavasta tuotteesta, johon moderaattorit vastasivat tiedustellen lisää tietoa. Tutkijoiden mukaan ääneenajattelumenetelmä Tanskassa oli melko rentoutunut ja vaikutti enemmän haastattelulta. Myös Intiassa käyttäjät ajattelivat aktiivisesti ääneen. Siellä moderaattorit muistuttivat käyttäjää. Tutkijat kuitenkin havaitsivat, että oli hankala erottaa, milloin moderaattorit kysyivät kysymyksiä ja milloin antoivat muistutuksia jatkaa ääneenajattelua. Kiinassa puolestaan ääneenajattelu ajautui usein retrospektiiviseksi tai selittämiseksi. Osallistujat keskittyivät usein tekemään tehtävää eivätkä ajatelleet ääneen, ja monesti moderaattorit eivät muistuttaneet heitä.

Niin ikään Shi (2008) havaitsi, että kiinalaiset käyttäjät eivät ajatelleet aktiivisesti ääneen. Hän havainnoi kollegoidensa kanssa käytettävyydestä viidessä kiinalaisessa yrityksessä. Hänen mukaansa moderaattorit muistuttivat osallistujia ääneenajattelusta ja kysyivät kysymyksiä, mutta ääneenajattelu oli enemmän selittämistä tai retrospektiivistä ääneenajattelua. Hänen mukaansa kiinalaisten käyttäjien vaikeus ajatella ääneen tehtäviä tehdessä voi johtua holistisesta ajattelutavasta, jota voi olla vaikeampaa verbalisoida kuin länsimaissa yleisempää analyttistä ja lineaarista ajattelutapaa.

## **4 Kehitysehdotukset ja varianttien haasteiden koonti**

Tässä luvussa esitellään ensin puheviestintämalli, jota on esitetty käytettävyydestä ajatellen kehitettynä parannuksena klassiseen malliin. Lisäksi luvussa mainitaan lyhyesti myös yksintehtävälle ääneenajattelulle vaihtoehtoinen paritestausta. Toisessa alaluvussa on koonti eri varianttien haasteista sekä muutamasta hyödyistä, joita tutkielmassa on tullut ilmi.

#### 4.1 Kehitysehdotuksina puheviestintämalli ja paritestausta

Tässä alaluvussa käydään läpi vielä lyhyesti esimerkkejä toimintatavoista, joita on ehdotettu vaihtoehtoina ääneenajattelun haasteisiin. Jo aiemmin tässä tutkielmassa on käyty läpi erilaisia variantteja, jotka voivat haasteistaan huolimatta toimia ratkaisuna joissakin tilanteissa.

Keskeisenä haasteena korostui teorian puute ja epä johdonmukainen käyttö. On luonnollisesti järkevää olla viittaamatta teoriaan, jota ei noudata. Ratkaisuna klassisen teorian haasteisiin käytettävyydestä Boren ja Ramey (2000) ehdottavat puheviestintäteoriaan pohjautuvaa vaihtoehtoa. Se ottaa huomioon puheen sosiaalisen luonteen, missä myös kuuntelija vaikuttaa puhujaan, ja kuuntelijan hiljaisuudellakin on merkitys. Teorian mukaan moderaattori on aktiivinen kuuntelija, jolloin monologin sijaan kyse on dialogista. Klassisessa teoriassa moderaattori keskeyttää ainoastaan, kun osallistuja on hiljaa. Puheviestintämallin mukaan on parempi, jos moderaattori huomioi kuuntelevansa melko tasaisesti tulevilla, mutta huomaamattomilla ilmaisuilla, kuten ”mm hm” tai ”uh huh”, jolloin huomiot ovat luonnollinen osa vuorovaikutusta. Näin voidaan saada samanlaista mahdollisimman jatkuvaa ja keskeytymätöntä ääneenajatteluaineistoa kuin myös klassisen teorian pyrkimyksenä on. Yhtenä mahdollisena haasteena moderaattorille voi kuitenkin olla huomiointien ajoittaminen sopiviin kohtiin niin, että ne sopivat osallistujan puheen sekaan (Boren & Ramey, 2000.)

Koska puheviestintämalli on tehty käytettävyydestä ajatellen, Boren ja Ramey (2000) ottavat siinä myös huomioon käytettävyydestä ominaispiirteitä, kuten prototyyppivaiheessa olevan järjestelmän epävakauden ja yllättävät tilanteet. Lisäksi he tiedostavat, että moni käytettävyydestä haluaa klassisessa mallissa kiellettyä kolmannen tason aineistoa, joten he ohjeistavat myös siihen liittyen.

Boren ja Ramey (2000) kuitenkin huomauttavat, että kyseistä mallia on vielä tutkittava. Tätä tutkielmaa varten luetuissa tutkimuksissa se ei ollut kovin laajasti näkyvillä, paitsi luvussa 3.3 mainitussa Alhadretin ja Mayhew'n (2017) tutkimuksessa, jossa puheviestintämalli oli yksi vertailtavista varianteista. Alhadreti ja Mayhew myös mainitsevat, että mikään tutkimus ei ole tarkastellut mallia yksityiskohtaisesti, joten sen vaikutuksesta ei ole todisteita.

On myös joitakin yksintehtävälle ääneenajattelulle vaihtoehtoisia toimintatapoja. Yksi näistä on paritestausta, jossa käyttäjät testaavat järjestelmää pareittain tavoitteenaan yhdessä ratkaista tehtäviä. Se mainittiin lyhyesti luvussa 3.1, jossa eräs McDonaldin ja kollegoiden (2012) kyselyyn vastannut ammattilainen sanoi suosivansa paritestausta, sillä se tuntui hänestä luontevammalta, ja osallistujat olivat vähemmän tietoisia tarkkailusta. Niin ikään Nielsen (1993) toteaa paritestausta suurimman vahvuuden olevan tilanteen luonnollisuus verrattuna yksin tehtävään ääneenajatteluun. Lisäksi menetelmä voisi toi-

mia erityisesti lapsilla. Toki tälläkin menetelmällä on omat haasteensa. Esimerkiksi pareilla voi olla erilaisia tapoja käyttää järjestelmää tai he eivät pysty työskentelemään yhdessä. Lisäksi käyttäjiä tarvitaan kaksinkertainen määrä yksintehtävään testaukseen verrattuna. (Nielsen, 1993.)

#### 4.2 Koonti eri varianttien haasteista

Taulukkoon 1 on vielä koottu eri varianttien keskeisiä haasteita. Varianteista on mukana luvussa 3 käsitelty klassinen, interaktiivinen ja retrospektiivinen variantti. Lisäksi vertailuun on sisällytetty myös luvussa 4.1 esitelty puheviestintämalli.

Taulukko 1. Koonti eri varianttien haasteista ja hyödyistä

|                          | Klassinen  | Puheviestintämalli  | Interaktiivinen  | Retrospektiivinen  |
|--------------------------|--|---|--|--|
| Haasteet moderaattorille | Aineisto ei välttämättä tarpeeksi hyödyllistä<br>Kognitiivisen psykologian teorian sopivuus<br>Ei huomioi käytettävyydestä ominaispiirteitä<br>Tehtäviin kuluva ajan mittaaminen | Huomioimisen (esim. ”mm hm”) ajoittaminen sopiviin kohtiin<br>Ei tutkittu paljoo<br>Tehtäviin kuluva ajan mittaaminen | Klassisen teorian mukaan aineisto ei enää kunnollista<br>Osallistujat voi alkaa arvostella ja ylläanalysoida järjestelmää<br>Moderaattorin taitojen merkitys kasvaa<br>Validiteetin vaarantuminen<br>Tehtäviin kuluva ajan mittaaminen | Aineiston puutteellisuus, jos käyttäjä unohtaa testissä mieleen tulleita asioita<br>Aikaavievyys |
| Haasteet käyttäjälle     | Epäluonnollisuus<br>Monologi vaikeaa   |   | Voi tuntua kuulustelulta<br>Kuormittava, muuttaa tapaa tehdä tehtäviä  | Tehtäviä tehdessä mieleen tulleiden asioiden unohtaminen<br>Aikaavievyys                         |
| Hyödyt moderaattorille   | Tuottaa 1. ja 2. tason aineistoa, joka voi joskus olla tavoitteena   | Kehitetty käytettävyydestä varten   | Tuottaa mahdollisesti hyödyllisempää aineistoa   | Tehtäviin kuluva ajan mittaaminen onnistuu   |
| Hyödyt käyttäjälle       | Ei vaikuta tehtävien suoritukseen yhtä paljon kuin interaktiivinen variantti   | Ottaa huomioon puheen sosiaalisen luonteen, jossa kuulijain hiljaisuudella on merkitys                                | Voi tuntua luonnollisemmalta, sillä moderaattorin keskeytysten myötä osallistujan ei tarvitse olla yksin äänessä   | Ei vaikuta osallistujan toimintaan tehtäviä tehdessä   |

Taulukosta nähdään, että kaikkiin variantteihin liittyy omanlaisiaan haasteita. Usein myös jonkun variantin haaste on toisen variantin hyöty. Tämä kuvastaa sitä, että ei ole täydellistä varianttia, joka sopisi kaikkiin tilanteisiin, vaan kaikissa on joitain etuja, mutta myös haastavia puolia.

Taulukkoon on myös listattu muutama hyöty, joita tutkielmaa tehdessä on noussett esiin. Hyödyt eivät kuitenkaan ole kattavia, koska tässä tutkielmassa on keskitytty haasteisiin.

## 5 Pohdinta

Luvussa 3 esiteltyjen tulosten perusteella voidaan havaita, että ääneenajattelumenetelmän käyttöön käytettävyydestä liittyvät erilaisia haasteita. Nämä haasteet koskettavat esimerkiksi menetelmän käytäntöjä, teoriaa, varianteja ja saavutettavuutta. On kuitenkin hyvä huomioida joitakin asioita tutkielman aineistonkeruuseen, tutkimusmenetelmään ja tulosten tulkintaan liittyen.

Tutkielman tekemiseen vaikutti lähteiden saatavuus. Tietoa hakiessa kävi ilmi, että ei vaikuta löytyvän kovin montaa tutkimusta, jotka käsittelevät eksplisiittisesti menetelmään liittyviä haasteita niin, että niiden otsikossa tai tiivistelmässä olisi suoraan mainittu haaste-sana tai sen synonyymi. Tästä huolimatta lähde saattoi antaa tietoa menetelmään liittyvästä haasteesta. Tämän takia välillä saattoi olla hankalaa selvittää, mitkä tutkimukset ovat relevantteja, ja haasteiden löytäminen vaati melko paljon omaa tulkintaa ja päättelyä. Voi olla, että relevantteja tutkimuksia on jäänyt huomaamatta.

Lisäksi on hyvä huomioida, että edellisessä luvussa esitetyt haastekategoriat eivät ole mitään yleisiä tutkimuksista löytyneitä kategorioita. Ne on tehty tätä tutkielmaa varten luettujen lähteiden perusteella, ja ryhmittelyn päätarkoitus on helpottaa haasteiden kokonaisvaltaista hahmottamista. Tämän tutkielman keskeinen anti onkin toimia ryhmittelevänä tekijänä erilaisista aikaisempien tutkimuksien löytämisestä haasteista. Kuitenkin haasteiden ryhmittely aiheutti ajoittain hankaluuksia, sillä sama haaste liittyi välillä useampaan kategoriaan tai eri kategorioiden haasteet liittyivät toisiinsa. Jatkossa voitaisiinkin perehtyä lisää siihen, miten haasteita kannattaisi ryhmitellä, ja onko jokin keskeinen kategoria jäänyt käsittelemättä.

Lähteisiin liittyen on huomioitava se, että moni keskeinen lähde on jo melko vanha. O'Brien ja Wilson (2023) huomauttavat, kuinka käytännöt ja asenteet ovat voineet muuttua esimerkiksi McDonaldirin ja kollegoiden tutkimuksesta vuodelta 2012, sillä sitä seuranneen vuosikymmenen aikana alalle on tullut lukuisia uusia käytettävyydestäajia.

On myös hyvä pohtia saatujen tulosten merkittävyyttä. Vaikka menetelmään liittyy haasteita, monet löydettyistä haasteista eivät vaikuta kovin vakavilta tai estä menetelmän käyttöä. Tutkimuksessa havaittujen haasteiden kääntöpuolella on usein hyötyjä, jotka saattavat olla haasteita paljon suurempia. Esimerkiksi interaktiiviseen ääneenajatteluun

liittyy haasteita, mutta siihen liittyy myös monia hyötyjä, ja variantin suosion perusteella voisi olettaa, että hyödyt koetaan haasteita suuremmiksi. Lisäksi myös vaihtoehtoisissa toimintatavoissa on omia haasteita, joten pitää vain pyrkiä valitsemaan sellainen variantti tai menetelmä, johon liittyy kyseisessä käyttötilanteessa vähiten haasteita tai jonka hyödyt ovat suurimmat haasteisiin nähden.

Tuloksia tulkittaessa on myös hyvä ottaa huomioon, että ääneenajattelumenetelmän käyttöön vaikuttaa moni asia, mikä puolestaan voi vaikuttaa haasteisiin. Kuten aiemmin tutkielmassa on tullut ilmi, ääneenajattelumenetelmää voidaan käyttää hyvin erilaisissa tilanteissa. Menetelmää käytetään sekä tieteellisessä tutkimuksessa että käytännön testauksessa yrityksissä. Lisäksi sitä käytetään sekä formatiivisessa että summatiivisessa käytettävyydestauksessa. Myös testattavat järjestelmät tai laitteet voivat vaihdella suuresti. Lisäksi esimerkiksi käytettävyydestin osallistujan persoonallisuus, kulttuuritausta ja muut ominaisuudet voivat vaikuttaa ääneenajatteluun. Tällainen suuri vaihtelevuus olosuhteissa voi tehdä selvien johtopäätösten vetämisestä vaikeaa menetelmän haasteiden suhteen. Esimerkiksi tutkielmassa käsitelty vaihtelevuus menetelmän käytössä ammattilaisten keskuudessa voi johtua juuri siitä, että he käyttävät menetelmää hyvin erilaisissa tilanteissa, eikä näin ollen käytäntöjen vaihtelevuus välttämättä ole itsessään suuri haaste.

## 6 Yhteenveto

Tässä tutkielmassa etsittiin kirjallisuuskatsauksella vastausta siihen, mitä haasteita ääneenajattelumenetelmän käytössä on käytettävyydestaukseen liittyy. Ääneenajattelu on keskeinen käytettävyyden tutkimisen menetelmä, jossa käyttäjä ajattelee ääneen käytettävyydestauksessa. Sen avulla saadaan tietoa käyttäjän ajatusprosesseista tehtäviin liittyen, ja näin voidaan löytää käytettävyysoongelmia.

Lähdekirjallisuudesta löytyneiden haasteiden perusteella haasteet ryhmiteltiin neljään kategoriaan. Ensimmäisenä kategoriana ovat käytännön haasteet. Tällainen haaste on esimerkiksi se, miten käyttäjät saadaan ajattelemaan ääneen, sillä menetelmä voi tuntua ihmisistä epäluonnolliselta. Osallistujat voivat myös yrittää miellyttää moderaattoria. Lisäksi ääneenajattelu vaikuttaa tehtävien suoritusajkaan eikä sitä välttämättä kannata käyttää samanaikaisesti tehtävien teon kanssa, jos halutaan mitata tehtäviin käytettyä aikaa. Ääneenajattelu voi olla hankalaa muissakin tilanteissa, kuten kenttätestauksessa.

Toisena kategoriana ovat teoriaan liittyvät haasteet. Tällaisena nousi esiin se, kuinka menetelmää ei usein yhdistetä mihinkään teoriaan. Yleisesti viitattu klassinen teoria on kognitiiviseen psykologiaan perustuva protokolla-analyysi, mutta siitäkin poiketaan ammattilaisten keskuudessa monella eri tavalla, eikä klassinen teoria välttämättä kaikilta osin sovellu käytettävyydestaukseen. Toisaalta vaihtoehtoinen interaktiivinen malli ei ole

myöskään vakiintunut. Teorian puute ja epäyhtenäinen käyttö voivat esimerkiksi vaikeuttaa tulosten vertailua.

Kolmantena kategoriana ovat menetelmän variantteihin liittyvät haasteet. Tässä perehdyttiin klassisen, interaktiivisen ja retrospektiivisen variantin haasteisiin. Klassisen variantin keskeisenä haasteena on, että se ei vaikuta tuottavan tarpeeksi hyödyllistä aineistoa. Täten moni käyttääkin interaktiivista varianttia. Sen haasteena on puolestaan mahdollinen muutos osallistujien tavassa tehdä tehtäviä. Se voi myös saada osallistujat arvostelemaan ja ylianalysoimaan järjestelmää. Interaktiivisen variantin haasteena on myös moderaattorin toiminnan suurempi vaikutus, ja lisäksi variantti voi myös vaikuttaa menetelmän validiteettiin. Retrospektiivisessä variantissa haasteina taas ovat testissä ilmenneiden asioiden unohtuminen ja menetelmän aikaavievyys.

Viimeisenä kategoriana ovat ääneenajattelun saavutettavuuteen liittyvät haasteet. Menetelmää ei aina voi suoraan soveltaa esimerkiksi lasten, muistisairaiden ikääntyneiden tai näkö- ja kuulovammaisten parissa, vaan vaaditaan mahdollisesti erilaisia toimintamalleja. Esimerkiksi kuulo- ja näkövammaisilla haasteena on menetelmän samanaikainen käyttö tehtävien teon kanssa viittomisesta ja ruudunlukijasta johtuen. Kuitenkin menetelmää voidaan muokata saavutettavammaksi. Samoin on tärkeää huomioida kulttuuriset ominaispiirteet menetelmän käytössä. Esimerkiksi tutkimuksissa on havaittu, että kiinalaiset käyttäjät eivät usein ajatelleet ääneen.

Tuloksista voidaan havaita, että ääneenajattelumenetelmään liittyy monenlaisia haasteita, mutta monet niistä ovat melko pieniä. Tutkielmassa esitettiin lyhyesti myös erilaisia vaihtoehtoisia toimintamalleja, kuten puheviestintämalli ja paritestausta, mutta niihinkin liittyy omia haasteitaan. Täten mikään yksi ääneenajattelumenetelmän variantti tai vaihtoehtoinen tapa ei vaikuta olevan aina toista parempi vaan käyttötilanne vaikuttaa siihen, mikä voisi olla paras menettelytapa. Kuitenkin yleisesti vaikuttaisi siltä, että vakiintunut teoria ja menettelytavat eri varianteille helpottaisivat esimerkiksi tulosten vertailua sekä yksittäisessä käytettävyysohjelmassa että tutkimusten kesken.

Tutkielman tulokset voivat auttaa ääneenajattelumenetelmää käyttäviä alan toimijoita ja opiskelijoita havaitsemaan ja välttämään mahdollisia ongelmakohtia sekä tiedostamaan, mitä menetelmän käytössä kannattaa huomioida, jotta siitä saa mahdollisimman suurta hyötyä. Jatkossa olisi mielenkiintoista perehtyä vielä tarkemmin erilaisiin haasteisiin sekä siihen, minkälaisissa konteksteissa ne esiintyvät, ja mitkä asiat niihin vaikuttavat. Lisäksi haasteiden todellista vakavuutta voisi tutkia lisää.

## Lähdeluettelo

Alhadreti, O. & Mayhew, P. (2017). To intervene or not to intervene: An investigation of three think-aloud protocols in usability testing. *Journal of Usability Studies*, 12(3), 111–132. <https://dl.acm.org/doi/10.5555/3190862.3190865>

- Alhadreti, O. & Mayhew, P. (2018). Rethinking thinking aloud: A comparison of three think-aloud protocols. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3173574.3173618>
- Boren, T. & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication*, 43(3), 261–278. <https://doi.org/10.1109/47.867942>
- Chen, S., Kirst, D., Wang, Q. & Huang, Y. (2023). Exploring think-aloud method with deaf and hard of hearing college students. *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, 1757–1772. <https://doi.org/10.1145/3563657.3595980>
- Clemmensen, T. (2011). Templates for cross-cultural and culturally specific usability testing: Results from field studies and ethnographic interviewing in three countries. *International Journal of Human-Computer Interaction*, 27(7), 634–669. <https://doi.org/10.1080/10447318.2011.555303>
- Dingman, B., Tigwell, G.W. & Shinohara, K. (2021). Interview and think aloud accessibility for deaf and hard of hearing participants in design research. *Proceedings of the 23<sup>rd</sup> International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21)*, 1–3. <https://doi.org/10.1145/3441852.3476526>
- Donker, A & Markopoulos, P. (2001). Assessing the effectiveness of usability evaluation methods for children. *Advances in Human-Computer Interaction I: Proceedings of the Panhellenic Conference with International Participation in Human-Computer Interaction, PC-HCI 2001*, 409–410.
- Donker, A. & Reitsma, P. (2004). Usability testing with young children. *Proceedings of the 2004 International Conference on Interaction Design and Children (IDC)*, 43–48. <https://doi.org/10.1145/1017833.1017839>
- Fan, M., Shi, S. & Truong, K.N. (2020). Practices and challenges of using think-aloud protocols in industry: An international survey. *Journal of Usability Studies*, 15(2), 85–102. <https://dl.acm.org/doi/10.5555/3532708.3532711>
- Gibson, A., McCauley, C., Mulvenna, M., Ryan, A., Laird, L., Curran, K., Bunting, B., Ferry, F. & Bond, R. (2016). Assessing usability testing for people living with dementia. *REHAB '16: Proceedings of the 4th Workshop on ICTs for Improving Patients Rehabilitation Research Techniques*, 25–31. <https://doi.org/10.1145/3051488.3051492>
- Hanna, L., Ridsden, K. & Alexander, K.J. (1997). Guidelines for usability testing with children. *Interactions*, 4(5), 9–14. <https://doi.org/10.1145/264044.264045>
- Hertzum, M. (2016). Usability testing: Too early? Too much talking? Too many problems? *Journal of Usability Studies*, 11(3), 83–88. <https://dl.acm.org/doi/10.5555/2993219.2993220>
- Hertzum, M., Hansen, K.D. & Andersen, H.H.K. (2009). Scrutinising usability evaluation: Does thinking aloud affect behaviour and mental workload? *Behaviour & Information Technology*, 28(2), 165–181. <https://doi.org/10.1080/01449290701773842>

- ISO 9241-11. (2018). SFS-EN ISO 9241-11:2018, Ihmisen ja järjestelmän vuorovaikutuksen ergonomia. Osa 11: Käytettävyys. Määritelmiä ja käsitteitä. International Organization for Standardization.
- McDonald, S., Edwards, H.M. & Zhao, T. (2012). Exploring think-alouds in usability testing: An international survey. *IEEE Transactions on Professional Communication*, 55(1). <https://doi.org/10.1109/TPC.2011.2182569>
- McDonald, S., Zhao, T. & Edwards, H.M. (2016). Look who's talking: Evaluating the utility of interventions during an interactive think-aloud. *Interacting with Computers*, 28(3), 387–403. <https://doi.org/10.1093/iwc/iwv014>
- Muistisairaudet: Käypä hoito -suositus. (11.12.2023). Suomalaisen Lääkäriseuran Duodecimin, Societas Gerontologica Fennican, Suomen Geriatri -yhdistyksen, Suomen Neurologisen Yhdistyksen, Suomen Psykogeriatrisen Yhdistyksen ja Suomen Yleislääketieteen yhdistyksen asettama työryhmä. Suomalainen Lääkäriseura Duodecim. [www.kaypahoito.fi/hoi50044](http://www.kaypahoito.fi/hoi50044)
- Nielsen, J. (1993). *Usability Engineering*. Academic Press.
- O'Brien, L. & Wilson, S. (2023). Talking about thinking aloud: Perspectives from interactive think-aloud practitioners. *Journal of User Experience*, 18(3), 113–132. <https://dl.acm.org/doi/abs/10.5555/3605526.3605528>
- Olmsted-Hawala, E.L., Murphy, E.D., Hawala, S. & Ashenfelter, K.T. (2010). Think-aloud protocols: A comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 10)*, 2381–2390. <https://doi.org/10.1145/1753326.1753685>
- Riihiahho, S. (2015). *Experiences with usability testing: Effects of thinking aloud and moderator presence*. [Väitöskirja, Aalto-yliopisto]. <https://urn.fi/URN:ISBN:978-952-60-6227-3>
- Shi, Q. (2008). A field study of the relationship and communication between Chinese evaluators and users in thinking aloud usability tests. *NordiCHI '08 Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges*, 344–352. <https://doi.org/10.1145/1463160.1463198>
- Strain, P., Shaikh, A.D. & Boardman, R. (2007). Thinking but not seeing: Think-aloud for non-sighted users. *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, 1851–1856. <https://doi.org/10.1145/1240866.1240910>
- Wood, R., Dixon, E., Elsayed-Ali, S., Shokeen, E., Lazar, A. & Lazar, J. (2021). Investigating best practices for remote summative usability testing with people with mild to moderate dementia. *ACM Transactions on Accessible Computing*, 14(3), 1–26. <https://doi.org/10.1145/3460942>