



**UNIVERSITY
OF OULU**

KATSEENSEURANTA KÄYTETTÄVYYDEN TUTKIMISEN TUKENA

Oulun yliopisto
Tietojenkäsittelytiede
LuK-tutkielma
Johannes Turpeinen
2023

Tiivistelmä

Avainsanat

Katseenseuranta, käytettävyystudkimus, käytettävyysestaus, käytettävyysongelmat

Ohjaaja

Ft, yliopistonlehtori Mikko Rajanen

Motivaationa tälle tutkielmalle oli se, että katseenseuranta teknologiana kiinnostaa minua paljon sekä se, että käyttäjälähtöisen suunnittelun ja helppokäyttöisten järjestelmien kysyntä on kasvavaa. Halusin selvittää sen, voiko katseenseurannan avulla parantaa järjestelmien käytettävyyttä niin, että siihen käytetyt resurssit olisivat saadun lisäarvon perusteella katseenseurannan tuoman lisätyön arvoista. Lisäksi minua kiinnosti se, kuinka luotettavasti katseenseurannan avulla saatua dataa voidaan hyödyntää.

Tutkimusmenetelmänä toimi kirjallisuuskatsaus, jossa perehdyttiin tutkimuksiin katseenseurannan luotettavuudesta ja tehokkuudesta, käytettävyyden tutkimisesta, katseenseurannan hyödyntämisestä käytettävyyden parantamisessa, sekä katseenseurannan käytöstä muiden mittausten kanssa. Eri tutkimuksien avaaminen ja vertailu mahdollisti sitten aiheesta luotettavan pohdinnan ja johtopäätöksiä tekemisen.

Tutkimuksen tärkeimmät havainnot olivat se, että katseenseurannan hyödyntäminen voi oikeinkäytettynä tuoda uusia näkökulmia käytettävyyden parantamiseen. Oikeanlainen hyödyntäminen on kuitenkin avainasemassa, sillä mikäli saatua dataa ei hyödynnetä oikein, tai resurssit eivät ole riittävät, katseenseurannan käyttäminen ei tuo huomattavaa lisäarvoa. On myös tärkeä muistaa se, ettei katseenseurannan käyttö korvaa muita käytettävyyden tutkimismenetelmiä, vaan se on lisätyökalu. Tutkimus toi yhteen monia erilaisia näkökulmia katseenseurannan käytöstä.

Katseenseurannan käyttö käytettävyyden parantamisessa ei ole suurelle osalle tuttu käsite, joten tämä tutkimus voi tuoda uutta tietoa sen hyödyntämisestä. Tutkimus auttaa myös arvioimaan sitä, olisiko katseenseurannan käytöstä mahdollisesti hyötyä lukijan oman projektin käyttöliittymän käytettävyyden parantamisessa. Tutkimus kertoo myös asioista, joita on tärkeä ottaa huomioon, mikäli katseenseurantaa halutaan hyödyntää mahdollisimman laadukkaasti.

Keskeiset käsitteet

EEG, Aivosähkökäyrätutkimus

Katseen polku, kuvaa käyttöliittymään sen järjestyksen, jossa katse etenee eri osista toisiin

Katseenseuranta, Menetelmä silmän liikkeen ja katseen sijainnin tallentamiseksi kuvaamalla pupillin liikkeitä (Carter & Luke, 2020)

Sakkadit, silmän äkillisiä nytkähdysliikkeitä

Lämpökartta, Piirtää näkymään sen, kuinka monesti katse kiinnittyy tiettyyn kohtaan (Joseph & Muruges, 2020)

Sisällysluettelo

Tiivistelmä.....	2
Keskeiset käsitteet.....	3
1. Johdanto.....	5
2. Tutkimusmenetelmät.....	6
3. Aiempi tutkimus.....	7
3.1 Katseenseurannan luotettavuus ja tehokkuus.....	7
3.2 Käytettävyyden tutkiminen.....	8
3.3 Katseenseurannan hyödyntäminen käytettävyyden parantamisessa.....	10
3.4 Katseenseurannan käyttö muiden mittausten kanssa.....	14
4. Pohdinta.....	16
5. Johtopäätökset.....	18
Lähteet.....	20

1. Johdanto

Katseenseuranta on menetelmä, jonka avulla voidaan ymmärtää, miten ihmiset käyttävät visuaalista tietoa ja mihin he keskittävät katseensa jotain katsoessaan. Tämä saavutetaan nauhoittamalla silmän pupillien liikettä. Tutkielmassa perehdytään siihen, onko katseenseuranta luotettava ja tehokas tapa käytettävyyden tutkimiseen ja parantamiseen. Lisäksi perehdytään käytettävyyssarvioinnin menetelmiin käymällä läpi erilaisia menetelmiä ja ohjeistuksia, joita käyttämällä katseenseurantaa voidaan hyödyntää mahdollisimman laadukkaasti.

Motivaationa tälle aiheelle on käyttäjälähtöisen suunnittelun ja helppokäyttöisten järjestelmien kysynnän kasvu, sillä katseenseuranta tuo uusia näkökulmia tuotteiden sekä käyttäjärajapintojen käytettävyyden tutkimiseen ja optimoimiseen. Katseenseurannan hyödyntämisen tutkimiseen liittyy useita haasteita. Usein katseenseurantaa käytetään vain suoraan datan tuottamiseen ja tämän datan tutkimiseen, eikä siinä oteta huomioon katseenseurannan yhdistämistä ihmisen käyttäytymisen psykologisiin puoliin. Näin kävi muun muassa Chandran, Sharman, Malhotran, Jhan ja Mittalin (2015) tutkimuksessa, kun pelkällä katseenseurannalla saatu data oli haastavaa hyödyntää. He kuitenkin tunnistivat ongelman ja osasivat ottaa sen huomioon tuloksia käsitellessään, joka unohtuessaan voi vääristää tutkimustuloksia.

Katseenseuranta on myös teknologiana verrattain uutta sekä nopeasti kehittyvää, joten haasteena voi olla uudet teknologiat, jotka jäävät hyödyntämättä tutkimuksissa. Nopeasta kehityksestä kertoo esimerkiksi se, että Schnipken ja Toddin (2000) tutkimus katseenseurannan hyödyntämisestä käytettävyyden tutkimuksessa päättyi lopputulokseen, jossa ilman laajaa seulontaa tutkimukseen osallistuvien kesken, katseenseurannan hyödyntäminen luotettavasti onnistui vain 37,5 prosentin kohdalla osallistujista. Heidän lopputuloksensa oli siis se, etteivät he kokeneet katseenseurannan olevan heidän tutkimuksensa hetkellä tarpeeksi luotettavaa, jotta sitä voitaisiin käyttää käytettävyyden tutkimiseen ja tukemiseen. Kirjallisuuskatsauksen aikana kävi myös ilmi, että usein testiryhmät ovat verrattain homogeenisiä, eivätkä ole välttämättä järjestelmien todellisia loppukäyttäjiä. Tästä on esimerkkinä Wangin, Antonenkon, Celepkolon, Jimenezin, Fieldman E:n ja Fieldman A:n (2018) tutkimus, jossa testikäyttäjänä toimi 35 naispuolista yliopisto-opiskelijaa, jotka olivat kaikki 18–21-vuotiaita. On siis tärkeää muistaa tarpeellinen skeptisyys tuloksien toistettavuudesta ja perehdyttävä useiden eri tutkimuksien tuloksiin aiempaa tutkimusta läpikäydessä. Tämä homogeenisyys johtui useimmissa tapauksissa todennäköisimmin resurssien puutteesta.

Tutkimuksen tärkeimpiä tuloksia oli se, että katseenseuranta voi tuoda lisäarvoa käytettävyyden tutkimiseen, kunhan riskit on minimoitu ja osataan tutkia oikeita asioita, sekä katseenseurantaa käytetään yhdessä muiden menetelmien kanssa. Suurimpia riskejä ovat liian pienet resurssit, datan vääränlainen kerääminen, sekä datan vääränlainen hyödyntäminen. Tutkielmassa käydään ensin läpi tutkimusmenetelmiä, jossa käydään läpi kirjallisuuskatsauksen hakulausekkeita, hakuprosessia, sekä haun rajauksia. Tämän jälkeen käydään läpi keskeisiä käsitteitä, jossa katseenseuranta ja siihen liittyvä termistö ja käsitteet avataan. Sen jälkeen tarkastellaan aiempaa tutkimusta aiheesta, jossa tutkimukset käydään läpi. Aiemman tutkimuksen tarkastelun jälkeen on pohdintaosa, jossa tarkastellaan eri tutkimuksien päälöydöksiä merkitystä ja suhteita toisiinsa, sekä tehdään sen perusteella löydöksistä tulkintoja. Lopuksi tutkielmassa on yhteenvedo työn tuloksista.

2. Tutkimusmenetelmät

Kirjallisuuskatsauksen teokset ovat etsitty Google Scholarista. Rajasin tulokset siten, että niiden julkaisupäivän tulisi olla aikaisintaan vuonna 2009. Lähteissä on kuitenkin viisi tätä ennen julkaistua teosta, joista vanhin on vuodelta 1991. Se on vanhin löytämäni tutkimus, jossa katseenseurannan hyödyntämistä perinteisessä käytettävyytestauksessa tutkittiin (Benel, Ottends & Horst, 1991). Tämän teoksen mukaan ottaminen oli aiheen avaamisen kannalta tärkeää, sillä se kertoo katseenseurannan mahdollistavan teknologian kehitymisestä.

Käyttämäni hakulausekkeita:

- “eye tracking”
- ”eye tracking” AND usability
- ”eye tracking” AND “user testing”
- usability
- HCI OR “human computer interaction”
- HCI AND “eye tracking”
- “human computer interaction” AND “eye tracking”

Katseenseurantaa käytetään myös esimerkiksi videopeleissä, lääketieteessä sekä psykologiassa ja näistä aiheista löytyi tietyillä hakulausekkeilla tutkimuksia. Oli siis erityisen tärkeä tarkistaa löydetyistä artikkeleista se, että ne liittyivät juuri ihmisen ja tietokoneen välisen vuorovaikutuksen tutkimiseen tai olivat muuten aiheeltaan sellaisia, että ne liittyivät tutkimuksen aiheeseen. Yksi artikkeli kertoo esimerkiksi siitä, miten fyysisten laitteiden käytettävyyttä voi parantaa katseenseurannan avulla. Koin artikkelin kuitenkin sellaiseksi, että sen mukaan ottaminen ja läpikäyminen toisi lisäarvoa tähän tutkimukseen, sillä katseenseurantaan käytetty teknologia oli vastaavaa kuin käyttöliittymien tutkimisessa käytetty teknologia. Tutkimuksessa on myös lähteenä tieteellisten nettisivujen kirjoituksia, mutta sieltä löytyvää tietoa ei käsitellä kirjallisuuskatsauksessa yhtenä vertailtavana tutkimuksena.

3. Aiempi tutkimus

Tässä luvussa käydään aiempaan tutkimukseen pohjaten läpi ensin sitä, miten luotettavana katseenseurannalla hankittua dataa voidaan pitää. Joidenkin tutkimuksien tulokset tästä ovat varsin ristiriitaisia, mutta syyt ristiriitaisuuksille ovat selitettävissä vaihtelevilla toteutuksilla ja resursseilla. Sen jälkeen tutustutaan perinteisempiin tapoihin tutkia käytettävyyttä, jonka jälkeen tutustutaan siihen, miten katseenseurantaa voidaan hyödyntää käytännössä käytettävyyden parantamiseen. Lopuksi käydään läpi sitä, miten eri mittauksia voidaan hyödyntää yhdessä katseenseurannan kanssa mahdollisimman suuren hyödyn saamiseksi.

3.1 Katseenseurannan luotettavuus ja tehokkuus

Schnipke & Todd (2000) tutkivat katseenseurannan avulla kerätyn datan luotettavuutta vuonna 2000 huonoin tuloksin. Dataa saatiin kerättyä luotettavasti vain 37.5 % kohdalla osallistujista. Heidän mukaansa katseenseuranta ei ole käytännöllinen työkalu käytettävyystudkimukseen. Kuitenkin paljon vanhemmassa tutkimuksessa Benel, Ottens & Todd (1991) arvioivat katseenseurannan käyttöä käytettävyyden tutkimisessa seitsemällä osallistujalla hyvin tuloksin ja heidän mukaansa katseenseurannan käyttö on dramaattisesti lisännyt kerättävää käytettävyydataa, jonka avulla käyttöliittymiä saadaan parannettua.

Vanhemmissa tutkimuksissa tulokset ovat olleet siis varsin ristiriitaisia. Schnipke & Todd (2000) selittävät aiempien tutkimuksien positiivisia tuloksia sillä, että niissä osallistujiksi valikoitiin vain niitä, joiden kohdalla katseenseurannan käyttö onnistui varmemmin, kun taas he käyttivät katseenseurantaa jokaisen osallistujan kohdalla. Tällainen osallistujien tarkka seulominen luo ongelmatilanteen varsinkin silloin, jos todellisia loppukäyttäjiä ei saada osallistumaan tarpeeksi käytettävyyso Ongelmien löytämiseen. Schnipken & Toddin (2000) tutkimuksen mukaan heidän käyttämänsä katseenseurannan avulla ei saada luotettavaa dataa, mikäli osallistuja käyttää silmälasia, hänen pupillinsa eivät heijasta tarpeeksi valoa, iiris on liian vaalea, pupillit ovat liian isot, kasvot heijastavat liikaa valoa, silmät sekä silmäluomet häiritsevät seurantaa tai silmät ovat liian kuivat.

Nämä edellä mainitut tutkimukset avaavat kuitenkin vain katseenseurannan historiaa. Katseenseurannan teknologia on kehittynyt siten, että katseenseuranta antaa paljon luotettavampaa dataa, eikä häiriötekijöitä ole niin paljon. Koester, Brøsted, Jakobsen, Malmros & Andreasen (2017) tutkivat katseenseurannan käyttöä käytettävyyden parantamisessa lääketieteellisten laitteiden kohdalla ja heidän tutkimuksessaan katseenseurannan avulla löydettiin käytettävyyso Ongelmia luotettavasti. He suosittelivatkin katseenseurantaa lisätyökaluksi käytettävyystudkimuksiin. Kuitenkin hekin huomauttivat siitä, että riittävä aika, resurssit sekä tekniset taidot on otettava huomioon, jotta lisäarvoa saadaan. Myös valo-olosuhteiden tulee olla optimaaliset.

Koesterin ym. (2017) tutkimuskysymykset aiheesta olivat seuraavat; mitkä ovat katseenseurannan käytön hyödyt ja haitat käytettävyystudkimuksessa, miten katseenseurantaa voidaan yhdistää muihin tapoihin tutkia käytettävyyttä, sekä auttaako katseenseurannan lisääminen tutkimiseen saavuttamaan tuloksia nopeammin ja tehokkaammin. Tutkimus tuotti seitsemän päälöydöstä; osallistajat vilkuilivat tutkittavaa kohdetta käytettävyysohjeiden aikana häiriten näin ohjeiden vastaanottamista, kukaan

osallistujista ei katsonut tärkeää nappia tietyssä näkymässä vieraillessaan, aiempaa kokemusta omaavat käyttäjät eivät olleet yhtä tehokkaita kuin kokemattomat, tärkeitä epäonnistumista merkkavia vihjeitä jäi huomaamatta, vastaavista laitteista aiempaa kokemusta omaavat käyttäjät eivät lukeneet ohjeita yhtä hyvin, yksi vastaaja ei katsonut tuotteen logoa vaan ainoastaan luki sen nimen, sekä viimeisimpänä se, että käyttäjät huomasivat erot tuotteen kuvassa tuotelaatikossa sekä tuotteessa itsessään. Näistä seitsemästä löydöksestä viisi oli sellaisia, joita valmistajan tekemä analyysi ei ollut huomionnut. Näistä viidestä löydöksestä kaksi oli sellaista löydöstä, joiden huomaaminen koettiin erityisen tärkeäksi. Tärkein löydöksistä oli se, että ne osallistujat, joilla oli kokemusta vastaavasta laitteesta, eivät lukeneet ohjeita ollenkaan tai jättivät viimeiset rivit lukematta.

Katseenseurannan aiheuttamat haittapuolet löydettiin haastattelemalla osallistujia. 13 osallistujasta kaksi oli sitä mieltä, että lasit aiheuttivat hieman epä mukavuutta ja yksi osallistuja sanoi, että lasit istuivat huonosti. Viisi osallistujista oli myös sitä mieltä, että lasit olivat huomattavissa häiritsevästi testauksen aikana ja yksi osallistuja sanoi, että lasien pitäminen saattoi vaikuttaa siihen, miten hän oli vuorovaikutuksessa tutkittavan laitteen kanssa, sillä hän keskittyi paljon siihen, mihin hän katsoi.

On tärkeä huomata se, että kyseisessä tutkimuksessa katseenseuranta suoritettiin siis päähän laitettavilla lasilla, eikä kuvaamalla osallistujien silmiä etäämpää, toisin kuin joissain muissa tutkimuksissa. Fyysisen laitteen tutkiminen katseenseurannalla ei kuitenkaan ole käytännössä toteutettavissa etäältä kuvattuna, koska testikäyttäjien tulee katsoa laitetta eri kulmista ja muutenkin liikutella itse laitetta. Tätä ongelmaa ei ole, mikäli käyttöliittymä on esimerkiksi tietokoneen näytöllä.

3.2 Käytettävyyden tutkiminen

Käyttöliittymien käytettävyyden arvioiminen on tapa varmistaa se, että käyttöliittymät ovat mukautettu sopiviksi käyttäjille sekä toiminnoille, joita järjestelmän tulee tehdä, eikä käytöstä saa johtua mitään negatiivisia tuloksia. Käytettävyyсарvioinnin tavoitteena on siis tutkia järjestelmän tehokkuutta sinä mielessä, kuinka hyvin järjestelmä toimii ja täyttää ne tehtävät, joita varten se on suunniteltu. Käyttö ei saa myöskään viedä liikaa aikaa tai vaivaa (Bastien, 2010). Tässä luvussa perehdytään käytettävyyden arvioimistapoihin, joihin katseenseurannan käyttö ei perinteisesti kuulu.

Käytettävyysongelmat voivat olla subjektiivisia tarkoittaen sitä, että eri käytettävyydestä tekevät tiimit voivat tuoda esille eri käytettävyysongelmaa, eikä jokin tiimi huomioi jotain tiettyä piirrettä käytettävyysongelmana. Vaihtoehtoisesti jokin käytettävyysongelma voi jäädä vain huomaamatta. Tästä hyvä esimerkki on Kessnerin, Woodin, Dillionin & Westin (2001) tekemät huomiot siitä, että kun kuusi eri ammatikseen käytettävyyttä testaavaa tiimiä tutkivat saman valintaikkunan käytettävyyttä, löydettiin yhteensä 36 eri käytettävyysongelmaa, mutta yksikään ongelma ei ollut sellainen, jonka jokainen tiimi olisi löytänyt. Viisi tiimeistä löysi kaksi ongelmista, neljä tiimeistä löysi neljä ongelmista, kolme tiimiä löysi seitsemän ja yksi tiimi löysi 18 ongelmaa. Tiimien välisten löydöksiä määrät vaihtelivat siis paljon.

Tämän ongelman ratkomiseksi käytettävyystiimit voivat noudattaa erilaisia käytettävyydestä testauksen malleja, jotka on luotu helpottamaan käytettävyysongelmien löytämistä. Yksi tunnetuin malli on Jakob Nielsenin 10 heuristisen arvioinnin ohjetta. Nämä ovat vapaasti suomennettuna järjestelmän tilan näkyvyys, järjestelmän ja oikean

maailman vastaavuudet, käyttäjän vapaus ja kontrolli, johdonmukaisuus ja standardit, virheiden ehkäisy, tunnistaminen muistamisen sijasta, käytön joustavuus ja tehokkuus, esteettinen ja minimalistinen muotoilu, käyttäjien auttaminen virheiden tunnistamisessa, diagnosoimisessa sekä niistä palautumisessa ja viimeisimpänä apu ja dokumentaatio (Nielsen, 2020).

Pelkästään esimerkiksi näiden Nielsenin heuristisen arvioinnin ohjeiden noudattaminen käytettävyysohjelmien löytämiseen ei kuitenkaan riitä, mikäli halutaan saavuttaa mahdollisimman korkea käytettävyyden taso. Thyvalikakath, Monaco, Thambuganipalle & Schleyer (2009) suorittivat Nielsenin heuristisen arvioinnin mukaista käytettävyydestausta neljälle eri hammashoidon potilastietojärjestelmälle ja näin todetuista empirisistä käytettävyysohjelmista löydettiin 39–64 %. Keskimäärin käytettävyysohjelmista löydettiin siis vain puolet, kuten Kuva 1 näyttää.

Number of usability problems found through usability testing	... predicted by heuristic violations		
		Specific (percentage)	General (percentage)	Total (percentage of usability problems)
System				
EagleSoft	60	20 (77%)	6 (23%)	26 (43%)
PracticeWorks	44	0 (0%)	17 (100%)	17 (39%)
SoftDent	86	5 (15%)	34 (87%)	39 (45%)
Dentrix	96	41 (67%)	20 (33%)	61 (64%)
Total	286	66 (46%)	77 (54%)	143 (50%)

Kuva 1. Käytettävyydestaustauksessa löydetty ongelmat verrattuna heuristisin arvioinnin menetelmin ennustettuihin (Thyvalikakath, ym., 2009)

Heuristiset arvioinnin menetelmät löysivät siis merkittävän määrän käytettävyysohjelmista, muttei tarpeeksi, jotta pelkästään siihen voitaisiin luottaa. Heuristinen arviointi ei myöskään ennustanut käytettävyysohjelmia kovin tarkasti, vaan pikemminkin vain toivat niiden potentiaaliset piirteet esille. Tutkimus ei myöskään valitettavasti tuonut ilmi sitä, mitkä heuristiset rikkeet tuottivat käytettävyysohjelmia todennäköisimmin. Olisi erittäin toivottavaa, että vakavat käytettävyysohjelmat huomattaisiin mahdollisimman aikaisessa vaiheessa kehitysprosessia, mutta tällä hetkellä tämän mahdollisuus heuristisen arvioinnin avulla on avoin kysymys. Johtopäätöksenä oli kuitenkin se, että eri metodien yhteiskäyttö on kaikista hyödyllisin tapa tutkia käytettävyyttä (Thyvalikakath, ym., 2009).

Yksi käytettävyydestaustauksen tapa on myös etänä tehty asynkroninen käytettävyydestaustaus. Tässä käyttäjät ja arvioijat ovat ajallisesti ja tilallisesti erillään toisistaan (Bruun, Gull, Hofmeister & Stage, 2009). Bruun ym. (2009) vertailivat tutkimuksessaan kolme eri asynkronista käytettävyydestaustatapaa keskenään, sekä laboratorio-olosuhteissa suoritettujen testauksen tuloksiin. Nämä kolme tapaa olivat käyttäjien raportoimat kriittiset tapaukset, foorumilla tapahtuva verkossa raportointi ja keskustelu, sekä päiväkirjamainen pitkäaikainen käyttäjäraportointi. Perinteinen laboratorioissa suoritettu käytettävyydestaustaus oli edellä mainituille tavoille ikään kuin vertailukohtana. Tutkimuksen lopputuloksena oli kuitenkin se, että asynkroniset menetelmät suoriutuivat käytettävyysohjelmien löytämisessä huomattavasti laboratorio-olosuhteissa tapahtuvaa testausta huomattavasti nopeammin. Päiväkirjamainen raportointi oli asynkronisista tavoista parhain, mutta senkin avulla löydettiin vain puolet niistä käytettävyysohjelmista, mitä laboratorio-

olosuhteissa löydettiin. Huomionarvoista on lisäksi se, että Bruunin ym. (2019) tutkimukseen osallistui 10 yliopisto-opiskelijaa, jotka voivat olla vielä tavallista tottuneempia arvioiden tekemiseen ja raportointiin. Asynkroniset menetelmät säästävät kuitenkin huomattavasti resursseja laboratorio-olosuhteisiin verrattuna.

Edellisten tutkimuksien pohjalta voidaan todeta, että käytettävyyden tutkimisessa juuri resurssit ovat iso muuttuja, joiden puutteellisuus voi vääristää tuloksia huomattavasti. Resurssien tarpeen lisäksi perinteiset metodit voivat tuottaa tuloksia, jotka ovat hyvin subjektisia (Kamińska, Zwoliński & Laska-Leśniewicz, 2022). Tarvittavien resurssien hillitsemiseksi on tutkittu jonkin verran käytettävyydetutkimuksen automatisoinnin mahdollisuutta. Kamińska ym. (2022) tutkivat tällaista mahdollisuutta virtuaalisessa todellisuudessa nimeten heidän tutkimuksensa aiheen pilottitutkimukseksi. He käyttivät käytettävyyden tutkimiseen ainoastaan objektiivista dataa tuottavia keinoja, kuten silmikkonäyttöjä, pään ja käsien liikkeiden seuranta, EEG:tä sekä videotallenteita. Näiden tukena ja vertailukohteena oli kyselyistä ja haastatteluista saadun subjektiivisen datan automaattinen analyysi. Hyvän käytettävyyden periaatteina he käyttivät jo aiemmin mainitun Jakob Nielsenin toista ohjetta hyvälle käytettävyydelle, joka koostuu viidestä elementistä; opittavuudesta, tehokkuudesta, muistettavuudesta, virheistä sekä tyytyväisyydestä. Automaattisilla testausmenetelmillä löydettiin maksimissaan 84,23 % käytettävyyssongelmista, kun tuloksia verrattiin kyselyistä ja haastatteluista löydettyihin ongelmiin. Tämä mainittu prosenttimäärä saatiin siis yhdistelemällä eri metodien löydöksiä, eikä käyttämällä yhtä tiettyä tapaa.

Tämä ei ole tarpeeksi siihen, että käytettävyydetestaus olisi voitu suorittaa ainoastaan automaattisilla testausmetodeilla, mutta se on hyvä lähtökohta lisätutkimukselle ja kertoo siitä, että automaattisilla testausmenetelmillä on potentiaalia käytettävyyden tutkimisessa. Kamińska ym. (2022) ehdottivat jatkotutkimuksiin yhdeksi parannukseksi sen, että lisädataa hankittaisiin eri tavoin, kuten infrapunaämpökuvauksella, galvaanisella ihovasteella, liikkeiden tallentamisella sekä katseenseurannalla. Yhdistämällä näitä edellä mainittuja tapoja automaattisiin menetelmiin, voitaisiin saada mahdolliseksi tilanne, jossa pääosa käyttäjätestauksesta suoritettaisiin automaattisesti. Tämä vaatisi kuitenkin paljon lisätutkimuksia ja testausta.

3.3 Katseenseurannan hyödyntäminen käytettävyyden parantamisessa

Aiemmat käsitellyt tutkimukset katseenseurannasta ovat korostaneet sen oikeanlaisen käytön tärkeyttä. Täytyy tietää, mitä katseenseurannan käytöllä halutaan saavuttaa ja onko sen käyttämisestä hyötyä juuri siinä, mitä tutkitaan. Bojko (2005) luetteli ohjeistuksessaan katseenseurannan hyödyntämiseen eri asioita, joihin katseenseurannan käyttö voi vastata. Katseenseurannan avulla voidaan tulkita käyttäjän päätöksentekoprosesseja, määrittää visuaalisen haun tehokkuutta, hakustrategioita sekä käyttäjien odotuksia, arvioida visuaalisen suunnittelun ja liiketoimintatavoitteiden välistä vastaavuutta sekä viimeisimpänä rajoittaa käytettävyyssongelmien syitä. Ennen testausta, testauksen aikana, sekä testauksen jälkeen tulee kuitenkin pitää mielessä tärkeitä huomioita.

Ennen testausta tulee suunnitella testauksen tavoitteet hyvin, valita oikeat tuloksien mittaustavat, käyttää korkealaatuisia ärsykeitä, joihin katse kiinnitetään sekä käyttää testauksessa henkilöitä, joita on helppo seurata. Tämä tarkoittaa sitä, että värikkäät

piilolinssit, silmämeikit ja paksusankaiset lasit on syytä jättää pois testauksesta. Testauksen aikana osallistujille tulee antaa tarkat ohjeet sekä välttää ”ääneen ajattelemista”. Testauksen jälkeen on vältettävä liian vaikeasti ymmärrettävän termistön käyttöä tuloksista puhuessa, varottava ettei visuaalisiin tuloksiin, kuten lämpökarttoihin luoteta liikaa, sekä tehdä yhdenmukainen johtopäätös tuloksista (Bojko, 2005).

Bojkon (2005) kertoi esimerkissään eri keinoista mitata tuloksia, sekä siitä, mitä käyttäjän tietynlainen käyttäytyminen kertoo:

- Se, monestiko käyttäjän katse kiinnittyy tietylle rajatulle alueelle, kertoo siitä, kuinka informatiivinen alue on ja kuinka kiinnostavana käyttäjä sen kokee. Suuri määrä alueeseen kohdistuneita katseita kertoo siitä, että kiinnostus aluetta kohtaan on korkea ja se on käyttäjälle informatiivinen.
- Se, kauanko käyttäjä katsoo tiettyä rajattua aluetta, kertoo siitä, kuinka selkeää ja tiheää alueella oleva informaatio on. Epäselvä informaatio tai suuri informaation määrä vaikeuttaa tiedon käsittelyä heijastuen alueelle pitkään kiinnittyvään katseeseen.
- Katseen polun monimutkaisuus, katseenkiinnityksien määrä sekä ensimmäisen katseen aika tietylle rajatulle alueelle kertovat kaikki asettelun tehokkuudesta. Huonosti mietitty käyttöliittymän ulkomuodon järjestely aiheuttaa tiedonhaun tehottomuutta, joka ilmenee pitkillä ja monimutkaisilla katseen poluilla, joissa palataan toistuvasti edellisiin kohtiin. Lisäksi tämä aiheuttaa suuria määriä katseen kohdistamisia eri paikkoihin sekä sitä, että ensimmäinen katse johonkin kestää kauan.
- Näkymässä käyntien määrä, alueelle katsovien käyttäjien prosenttiosuus sekä järjestys, jossa eri osia käydään läpi kertovat kaikki eri alueiden näkyvyydestä sekä siitä, kuinka tärkeäksi ne on koettu. Alueen hyvä näkyvyys sekä tärkeys heijastuu käyntien tiheyteen, alueelle katsovien käyttäjien prosenttiosuuteen sekä siihen, onko alue yksi ensimmäisistä kohdista, johon katse kiinnittyy.

Kun nämä mittarit on otettu huomioon, käyttäjätestauksen tuottamien lämpökarttojen sekä katseen polkujen pohjalta voidaan tehdä päätelmiä käyttöliittymän käytettävyydestä, ja käytettävyysongelmia voidaan identifioida. Kuva 2 näyttää sen, miten katseen polku piirtyy numeroituna käyttöliittymään.



Kuva 2. Katseen polku (Bojko, 2005)

Kuva 3 näyttää sen, miten lämpökartta muodostuu. Punaisemmat alueet ovat alueita, joihin katse on kohdistunut useammin.



Kuva 3. Lämpökartta (Bojko, 2005)

Idrees, Kraft, Winter, Kuchler, Baumeister, Reilly, Reichert & Pryss (2023) tutkivat etäkäyttöisen terveystalviteita tarjoavan alustan eSanon käytettävyyttä kymmenellä osallistujalla. Katseenseurannan käyttö oli heillä avainasemassa ja he käyttivät sen osalta mittareinaan kokonaismääräistä katseen kohdistumista millisekunneissa, keskimääräistä katseen kohdistumista millisekunneissa, sekä katseen kohdistumisien lukumäärää. Lisäksi katseenseurannan käyttö mahdollisti heidän tutkia käyttäjien muodostamaa katseen polkua. Katseenseurannan tuoman datan tulkitsemisen myötä heidän tekemiään havaintoja olivat se, että 40 % käyttäjistä käytti vähintään puolet ajasta kysymyksiin vastaamiseen, 40 % käyttäjistä ohittivat pitkiä tekstikappaleita joko osittain tai kokonaan, ainakin yhden käyttäjän keskittymistä vaikeutti muiden kysymyksien näkyminen, kesken käyttämisen ilmestyvät ilmoitukset johtivat joillakin käyttäjille niiden painamiseen, 30 % käyttäjistä eivät olleet tietoisia siitä, että he olivat suorittaneet jonkin osa-alueen jo, osa käyttäjistä jätti vahingossa vastaamatta joihinkin kysymyksiin ja he huomasivat tämän vasta, kun koittivat siirtyä seuraavalle sivulle järjestelmän ilmoittua tästä, puolet

käyttäjistä eivät löytäneet laajennettavia tekstiosioita, sekä viimeisimpänä se, että valikosta löytyviä tekstejä ei huomioitu ollenkaan.

Tällaiset edellä mainitut huomiot mahdollistavat sen, että käytettävyysongelmien löytäminen, sekä niiden korjaaminen on verrattain yksinkertaista. Esimerkiksi se, että iso osa käyttäjistä jättää pitkiä tekstikappaleita lukematta voitaisiin korjata jakamalla tekstiä pienempiin osiin. Idrees ym. (2023) luettelivat omiksi parannusehdotuksikseen katseenseurannasta saadun datan pohjalta kerralla näkyvien kysymyksien minimoinnin esimerkiksi yhteen kysymykseen kerrallaan, käyttäjän itse laajennettavien tekstikenttien poistamisen, käyttöliittymän ulkonäön optimoimisen, kuvien koon pienentämisen, jotteivat ne vie turhaan käyttäjien huomiota, sekä laajojen tekstikappaleiden korvaamisen esimerkiksi interaktiivisilla media elementeillä. Mahdollisimman käyttäjäystävällinen käyttöliittymä on erityisen tärkeää terveystalvueluita tarjoavissa järjestelmissä, sillä käyttäjien mielialat voivat vaihdella paljon.

Katseenseurannan käyttäjäystävällisyys on kehittynyt paljon. Nykyään käyttäjän katsetta voidaan seurata käyttäjän näkökulmasta huomaamattomasti, sekä ilman sitä, että käyttäjällä olisi yllään mitään lisälaitteita, sillä suurin osa kaupallisista katseenseurannan mahdollistavista järjestelmistä kuvaavat silmän liikkeitä etäältä. Tämä edistysaskel on ollut merkittävä, sillä lisälaitteiden pitäminen tai seurannasta aiheutuva häiriö voi vaikuttaa siihen, miten käyttäjä on vuorovaikutuksessa käyttöliittymän kanssa (Djamasbi, 2014). Djamasbin (2014) mukaan katseenseurannalla on potentiaalia tulla nettisivujen suunnittelun ja kehityksen alan standardiksi, mutta ennen sitä käyttäjäystävällisyyden tulee parantua myös tutkijan näkökulmasta, eikä vain testikäyttäjän. Katseenseurannan käyttö vaatii manuaalisen asennuksen ja kalibroinnin, mikäli sitä ei ole valmiiksi integroitu monitoriin, joka on vielä harvinaista. Monitoreihin integroidut järjestelmät ovat taasen vaikeampia hyödyntää kenttätutkimuksessa. Ongelmia tulee myös mobiilikkehityksessä, sillä mobiililaitteiden kanssa käytettynä katseenseurannan mahdollistava välineistö on painavaa ja työlästä laittaa valmiuteen. Data-analyysi mobiililaitteita varten vaatii myös suuren määrän manuaalista prosessointia, joka johtaa siihen, että katseenseurannan käyttö mobiiliselvölkksien käytettävyyden testaamista varten vaatii paljon työtä ja aikaa (Djamasbi, 2014).

Katseenseurannan mahdollistavat ohjelmistot ovat usein myös kalliita, sillä ne on toteutettu monimutkaisesti korkean tason ohjelmointikielillä. Yksi tapa laskea katseenseurannan kuluja olisikin toteuttaa ohjelmistot alemman tason ohjelmointikielillä, kuten C:llä tai C++:lla (Hiley, Redekopp & Fazel-Rezai, 2006). Toisaalta Hiley ym. (2006) toteuttivat tutkimuksensa katseenseurannan luotettavasta käyttöön otosta normaalilla tietokoneeseen liitettävällä web-kameralla, joka vaati muutenkin ylimääräistä muokkaamista, kuten sen, että kamera tunnistaisi infrapunavaloa. Näin ollen käytetty ohjelmisto on muutenkin monimutkaisempaa kuin käytettäessä kameroita, jotka on tehty katseenseurannaa varten. Hiley ym. (2006). Tutkimus on vanha ja katseenseurannassa käytettävä teknologia on kehittynyt paljon vuodesta 2006. Pääpiirteet ovat kuitenkin pysyneet samana ja vanhentunut laitteisto auttaa ymmärtämään käytettävän teknologian toimintaperiaatteita paremmin. Katseen sijainti laskettiin määrittämällä pupillin keskikohdan suhteellista sijaintia neljään silmämunaan heijastavaan valoon, jotka tulivat näytön reunoiissa sijaitsevista neljästä ledistä (Hiley ym., 2006). Silmään heijastuviin valopisteisiin pohjautuvan kalibroinnin luotettavuus on ymmärrettävästi kyseenalaista. Tilan, jossa testausta suoritetaan, täytyy olla valaistukseltaan ihanteellinen, eikä esimerkiksi heijastuksia saa sekoittaa näyttöön kiinnitettyjen ledien heijastuksiin.

Mikäli puutteelliset resurssit estävät kalliiden katseenseurannan mahdollistavien toteutuksien käytön, katseenseurantaan voidaan nykyään käyttää älypuhelimia ja perinteisiä WEB-kameroita. Perinteisen WEB-kameran käyttö vaatii yleensä kuitenkin erityisen ohjelmiston, jotta visuaalinen data saadaan muutettua käytettävyydestä kannalta oleelliseksi. Tähän on kuitenkin olemassa avoimen lähdekoodin ohjelmistoja. Tällaiset budjettiratkaisut eivät ole yhtä tehokkaita kuin kalliimmat ratkaisut, mutta ne tarjoavat hyvän kompromissin hinnan ja suorituskyvyn välillä (Novák, Masner, Benda, Šimek & Merunka, 2023).

3.4 Katseenseurannan käyttö muiden mittausten kanssa

Katseenseurantaa ei tulisi käyttää ainoana keinona käytettävyyden tutkimisessa, vaan sen oikeaoppinen käyttö voi tuoda mittavaa lisäarvoa käytettynä muiden käytettävyydestutkimuksen keinojen kanssa. Yksi esimerkki olisi käyttää katseenseurantaa yhdessä heuristisen arvioinnin menetelmien, sekä käyttäjähaastatteluiden kanssa. Näin sellaisetkin käytettävyysongelmat, joita katseenseurannan avulla ei löydetä tulevat löydettyksi todennäköisemmin. Katseenseurannan avulla itsessään ei voida esimerkiksi tulkita kaikkia käyttäjien tunteita luotettavasti, esimerkiksi turhautumista tai tyytyväisyyttä. Katseenseurannan avulla ei myöskään voida olettaa löytyvän käytettävyysongelmia, joiden konteksti ei liity varsinaisesti käyttöliittymään katsomiseen.

Vaikka katseenseuranta oli suuressa roolissa Idreesin ym. (2023) suorittamassa verkkopohjaisen terveystalviteita tarjoaman järjestelmän käytettävyydestutkimuksessa, siihen oli yhdistetty muita käytettävyydestutkimuksen keinoja mahdollisimman monien käytettävyysongelmien löytämiseksi. Käytettyjä mittareita olivat System Usability Scalen mukaiset valmiit kysymykset käytettävyydestä, tutkijoiden itse laatimat järjestelmäkohtaiset kysymykset, tiettyihin paikkoihin kohdistetun katseen kesto, tiettyihin paikkoihin katseen kohdistamisen lukumäärä, katseen liikkeen analyysi tekstejä lukiessa, käyttäjien rohkaiseminen ajattelemaan ääneen järjestelmää käyttäessä ja näiden ajatuksien tulkitseminen, sekä viimeisimpänä testauksen jälkeen suoritettu vapaamuotoinen haastattelu, jonka aikana käyttäjät voivat tuoda omia huomioitaan esille. Idreesin ym. (2023) johtopäätökset tutkimuksesta tukivat sitä, että eri käytettävyydestutkimuksen menetelmien yhdistäminen on kriittistä, jotta saadaan syvempi ymmärrys siihen, miten käyttäjä on vuorovaikutuksessa järjestelmän kanssa.

Katseenseurannasta saadun datan hyödynnettävyyden tehostamiseksi sen käyttöä voi myös tukea erilaisten yhtäaikaisten lisämittausten avulla. Falkowska, Sobecki & Pietrzak (2016) tutkivat sitä, korreloiko silmän kohdistaminen johonkin tiettyyn käyttöjärjestelmän osaan jonkin tietyn tunteen kokemiseen käyttämällä EEG:tä ja katseenseurantaa testihenkilöillä. Tutkimuksen tavoitteena oli siis selvittää se, voidaanko katseenseurannan ja EEG:n yhdistämisellä vetää parempia johtopäätöksiä järjestelmien käytettävyydestä. Tutkimuksessa käytetty EEG mahdollistaa tämän tunnistamalla käyttäjien turhautuneisuus, lyhyen aikavälin innostuneisuus, pitkän aikavälin innostuneisuus, tyyneys ja rauhallisuus, sekä tehtäviin keskittyminen. Testikäyttäjien kohteena oli käyttää netissä olevaa sovellusta, joka toimii toiminnanohjausjärjestelmänä netissä oleville kaupoille. Sovellus piti kirjaa yritysten myynnistä reaaliajassa sekä tunnisti asiakkaiden mieltymyksiä.

Testikäyttäjien tehtävänä oli selvittää sovelluksen avulla verkkosivustolla tehtyjen käyntien määrän ero vuosien 2014 ja 2015 välillä, selvittää useimmiten käytetty verkkosivulle saapumisen lähde, antaa tiedot vuoden 2014 tuloista ja tarkistaa käyttäjät,

joiden sivuilla oli suurin osuus tuloista, sekä tarkistaa se, missä kuussa tehtiin vähiten tilauksia vuonna 2015 ja kertoa tämän kyseisen kuukauden tilausten keskimäärä. Tutkimuksen pohjalta voitiin vetää johtopäätös siihen, että korrelaatiota löytyi. Esimerkiksi käyttöliittymän tiettyyn osaan monta kertaa katseella palaaminen aiheutti yleisesti turhautumista käyttäjissä. Muita korrelaatioita oli lyhyen aikavälin innostuneisuuden ja katseen kohdistumisen keston positiivinen korrelaatio, tehtävään keskittymisen ja ensimmäisen katseen kohdistumisen positiivinen korrelaatio, pitkän aikavälin innostuneisuuden ja katseen kohdistumisen keston positiivinen korrelaatio, turhautuneisuuden ja ensimmäisen vilkaisuun kestävä ajan positiivinen korrelaatio, turhautumisen ja katseen kohdistumisen keston negatiivinen korrelaatio, lyhyen aikavälin innostuneisuuden ja alueelle katsotun ajan lyhenemisen korrelaatio, sekä viimeisimpänä käyttäjän tyyneyden ja rauhallisuuden vaikutus siihen, että tietylle alueelle katsotaan lyhyemmän aikaa (Falkowska ym, 2016).

Joseph & Muruges (2020) puolestaan tutkivat sitä, voisiko katseenseurannan avulla potentiaalisesti mitata käyttöjärjestelmien käytön kognitiivista kuormitusta, joka voi puolestaan antaa dataa helppokäyttöisyyden toteutumisesta. Tavoitteena oli tarjota helppo lähtökohta tutkijoille, interaktiivisen teknologian suunnittelijoille ja avustavan teknologian ammattilaisille, jotka haluavat syventää näkemystä katseenseurannan mittareista sekä niiden merkityksestä. Tämä saavutettaisiin tunnistamalla eri mitattavia asioita.

Katseenseurannan mittareiksi tunnistettiin lämpökartta sekä katseen polku, joiden avulla voidaan tutkia katseen kohdistamisen kestoa, katseen kohdistamisen määriä, katseen kohdistamisen lopullisia määriä, katseen kohdistamisen tiheyden jakautumista, tietylle rajatulle alueelle kohdistuneiden katseiden määrää, tiettyyn kohtaan kohdistuneiden katseiden määrää jaettuna kokonaisuudella, aikaa ensimmäiseen kertaan kun jotain katsotaan, keskiarvoa katseen kohdistumisen ajasta, katseen kohdistumisen tiheyttä, katseen kohdistumisen määrää jaettuna sekunnissa, tietyssä paikassa viipymisen aikaa, katsetta, katseen esimerkkejä, sekä sitä, kuinka usein huomio vaihtuu toisaalle. Lisäksi muodostetusta katseen polusta voidaan tulkita sen kestoa, pituutta, tiheyttä, suuntaa, suhdetta yleisiin polkuihin, sekä sakkadien ja fiksaatioiden suhdetta (Joseph & Muruges, 2020).

Kognitiivisessa tehtävänälyysissä mittareiksi määritettiin tehtävän suorittamiseen käytettyä aikaa sekä tehtävän suorittamiseen käytettyjen vaiheiden määrää. Kognitiivista kuormitusta pystyttiin mittaamaan myös sakkadien, pupillin koon muutoksien ja silmien räpyttelyn perusteella. Sakkadien mittareiksi tunnistettiin sakkadien pituus, niiden määrä, liikkeen laajuus, niiden regressio sekä niiden kesto. Pupillin koon mittarina on sen koon vaihtelut, jotka voivat kertoa ylipäätään kognitiivisesta kuormasta. Silmien räpyttelyn mittareiksi tunnistettiin räpyttelyn määrä verrattuna keskiarvoon, aika jonka silmä on kiinni, nopeus, jolla silmä suljetaan ja avataan, sekä räpyttelyn viive. Erityisen nopea räpyttely voi implikoida korkeaa mentaalista ponnistelua, samoin kuin erityisen hidas räpyttely. Heidän yhteenvetonsa oli se, että katseenseurantaa voitaisiin mahdollisesti käyttää myös käyttöjärjestelmien kognitiivisen kuormituksen mittaamiseen, sillä eri mittaamistapoja saatiin määriteltyä. Tämä mahdollistaisi vielä monipuolisemman lisäarvon saavuttamisen katseenseurannan käytön avulla (Joseph & Muruges, 2020).

4. Pohdinta

Tämän tutkimuksen tulokset ovat yhdenmukaisia muiden katseenseurantaa tutkivien modernien tutkimuksien kanssa, jotka ovat päätyneet siihen lopputulokseen, että katseenseurannan oikeanlainen käyttö parantaa käyttöliittymien käytettävyyden arviointia mahdollistamalla erilaisen käyttäjän käyttäytymisen tutkimistavan muihin menetelmiin nähden. Katseenseurannan käyttö itsessään ei kuitenkaan riitä käytettävyyden varmistamiseen, vaan paras hyöty saadaan, kun se on yksi käytettävyydestutkimuksen tapa muiden joukossa. Tutkimus on osoittanut myös sen, että mitä useampaa eri käytettävyydestutkimuksen tapaa käytetään, sitä enemmän käytettävyyssongelmia löydetään.

Jokainen projekti ei kuitenkaan missään nimessä voi käyttää kaikkia tunnettuja käytettävyyden tutkimistapoja käyttöliittymiensä käytettävyytensä varmistamiseen, vaan käytetyt tavat tulisi valita tarpeen sekä resurssien mukaan. Kärjistettynä esimerkkinä ääniohjatussa käyttöliittymässä katseenseurannan käyttö toisi varmasti niin vähän lisäarvoa, että resursseja menisi turhaan hukkaan. Tämän vuoksi projektissa olisi mielestäni hyvä olla mukana käytettävyyden asiantuntijoita joko kehitysyrityksestä, tai kokonaan ulkopuolisesta firmasta, jotka olisivat vastuussa käytettävyydestaustuksen suunnittelusta. Asiantuntijuus on myös erityisen tärkeää katseenseurantaa käytettäessä, jotta osataan etsiä oikeanlaista dataa oikeanlaisilla menetelmillä. Näen itse asian niin, että jokaisessa käytettävyydestutkimuksessa tulisi olla ainakin testikäyttäjille suunnatut testattavalle järjestelmälle erikseen laaditut kysymykset, jotka voidaan kohdistaa järjestelmälle ominaisiin piirteisiin. Jos käytetään valmista kysymyspohjaa, moni juuri käsittelyssä olevaa järjestelmää koskevat piirteet voivat jäädä käsittelemättä.

Näen katseenseurannan itse sellaisena asiana, jonka käyttö käytettävyyden tutkimisessa tulee kasvamaan katseenseurannan teknologian kehityksen mukana. Laitteista tulee pienempiä, tarkempia, halvempia ja helppokäyttöisempiä madaltaen kynnyistä niiden hyödyntämiseen. Yksi tärkeimmistä tulevaisuuden kehityssuunnista katseenseurannalle onkin Djamasbin (2014) mukaan se, että helpommat ja intuitiivisemmat käyttöliittymät helpottaisivat tiedon keräämistä, validointia ja analysointia. Tämä koskee etenkin mobiililaitteita, sillä kuten aiemmin todettu, niistä kerätyn datan käsittely on erityisen työlästä. Myös tekoälyn rooli tulee kasvamaan katseenseurannan hyödyntämisessä, kun saatua dataa voidaan prosessoida automaattisesti tekoälyn avulla. Tekoälyn suurempi hyödyntäminen voisi myös tehdä matalamman tason katseenseurannan toteutuksista luotettavampia, jolloin resurssipula ei tarkoita sitä, että katseenseurantaa tulisi pitää pienemmässä roolissa.

Teknologian kehittyessä katseenseurantaa käsittelevistä tutkimuksista tulee laadukkaampia. Helppokäyttöinen, kompakti ja halpa katseenseuranta mahdollistaa sen, että tutkimuksien resursseihin mahtuu enemmän testihenkilöitä laajemmalla skaalalla. Kaikissa tämän tutkimuksen lähteissä olevissa tutkimuksissa testihenkilöitä olisi voinut olla enemmän ja niitä olisi voinut olla erilaisemmista ihmisryhmistä. Tutkimuksissa testihenkilöinä käytettiin pääasiassa nuorempia ihmisiä samankaltaisista taustoista, joten iän tai kulttuurien tuomia eroja ei voitu ottaa huomioon. Tämä vaikuttaa tuloksien luotettavuuteen ja hyödyllisyyteen varsinkin silloin, jos testihenkilöt eivät edusta suurinta osaa todellisista loppukäyttäjistä. Olisi mielenkiintoista nähdä tutkimuksia erityisesti siitä, eroaako eri kulttuuritaustasta tulevien ihmisten antamat tulokset toisistaan jotenkin tavoilla, jotka olisivat yleistettävissä. En myöskään nähnyt aiemmissä tutkimuksissa mitään mainintoja siitä, piirtyykö katseen polku eri tavoilla ihmisillä, jotka lukevat

pääasiallisesti tekstiä esimerkiksi arabiaksi tai hepreaksi. Näissä kielissä lukeminen tapahtuu oikealta vasemmalle, toisin kuin esimerkiksi suomen- tai englanninkielisessä tekstissä, joissa tekstin lukeminen tapahtuu vasemmalta oikealle. Suuri osa tutkimuksista tutkivat myös aihettaan ensikertaa, selittäen testihenkilöiden pienen määrän. Lähitulevaisuudessa mikäli samoja aiheita tutkitaan uudelleen, voisi olettaa, että testihenkilöiden määrää kasvatettaisiin.

Ikä vaikuttaa visuaaliseen käyttäytymiseen nettisivuja katsoessa. Kun verrattiin vanhemman ikäryhmän visuaalista käyttäytymistä erilaisia nettisivuja katsoessa nuoremman ikäryhmän käyttäytymiseen katseenseurannan avulla, vanhempi ikäryhmä kiinnitti katseensa pidemmäksi aikaa nettisivujen eri kohtiin sivuja katsoessaan. Nuorempi ikäryhmä suosi huomattavasti vähemmän nettisivuja, joissa on suurempi määrä komponentteja. Joissain tapauksissa vanhemman ikäryhmän katseen kattoivat myös laajemman alueen. Vanhempien ikäryhmien käyttäytyminen vaikutti siis kärsivällisemmältä, kun taas nuorempien ikäryhmien käyttäytyminen vaikutti kärsimättömämmältä (Djamasbi, Siegel, Skorinko & Tullis, 2011). Vanhempien ikäryhmien käyttö testihenkilöinä on erityisen tärkeää suomalaisissa järjestelmissä, sillä vanhemmat ikäryhmät tulevat kasvamaan suhteessa nuorempiin. Tämä tarkoittaa sitä, että kotimaisista loppukäyttäjistä yhä isompi osa tulee olemaan vanhempia, riippuen toki järjestelmästä. Myös erilaisten sairauksien voi olettaa vaikuttavan silmien liikkeisiin, joka olisi hyvä ottaa huomioon etenkin terveystalvueluita koskevisu järjestelmissä.

5. Johtopäätökset

Katseenseurannan käyttö käytettävyydestestauksen tukena voi siis tuoda merkittävää lisäarvoa testaukseen, mutta sen käyttöön liittyy myös riskejä, jotka olisi hyvä ottaa huomioon jo ennen testauksen aloittamista. Laitteita ja ohjelmistoja tulee käyttää oikein ja huolellisesti, sekä täytyy tietää, minkälaista dataa halutaan, jotta sitä voi hyödyntää. Yleinen virhe katseenseurannan käytössä onkin se, että katseenseurannan avulla dataa kerätään paljon, eikä lopuksi edes tiedetä, miten kaikkea kerättyä dataa voidaan hyödyntää (Bojko, 2005). Tällaiset riskit saadaan minimoitua ottamalla käytettävyydestutkimuksen suunnitteluun mukaan sellaisia asiantuntijoita, jotka ovat todella perehtyneet katseenseurannan hyödyntämiseen.

Katseenseurannan käyttö käytettävyydestestauksessa edellyttää myös riittäviä resursseja niin laitteiden ja ohjelmistojen, kuin testaajien ja asiantuntijoiden kannalta. Etenkin pienissä organisaatioissa olisi hyvä siis pohtia tarkkaan, voidaanko katseenseurantaa ottaa käytettävyydestestaukseen mukaan siten, että varmistutaan siitä, että lisäarvoa testaukseen saadaan. Mikäli resurssit ovat verrattain pienet, mutta katseenseurantaa halutaan silti käyttää, käytettävyydestutkimuksessa voidaan hyödyntää alemman tason toteutuksia katseenseurannasta. Tällöin tulee kuitenkin muistaa, että kompromisseja tehdään niin tehokkuuden kuin tulosten luotettavuuden kannalta, joten katseenseurannan roolia ei tulisi pitää yhtä isona kuin projekteissa, joissa käytetään korkeamman tason katseenseurantaa.

Tutkimuksien tulokset osoittavat sitä, että oikein käytettynä katseenseurannan avulla saadaan uusia tärkeitä näkökulmia käytettävyysongelmien löytämiseksi etenkin, jos halutaan saada tietoa siitä, mihin käyttäjien huomio kiinnittyy käyttöliittymää katsoessa ja mihin asioihin kiinnitetään ensiksi huomiota. On tärkeä pitää mielessä silti se, että parhaan tuloksen saavuttamiseksi katseenseuranta on vain yksi menetelmä käytettävyyden parantamiseksi muiden menetelmien joukossa, eikä sen tulisi korvata muita menetelmiä. Katseenseuranta on siis ennen kaikkea tapa täydentää muita menetelmiä ja saada lisätietoa käyttäjien käyttäytymisestä ja reaktioista käyttöliittymää käyttäessään. Se, kuinka isossa roolissa katseenseurannan käyttö on, tulisi riippua testattavan käyttöliittymän ominaisuuksista, sekä käytössä olevista resursseista.

Katseenseurannan käytöstä saadaan suurempaa lisäarvoa myös silloin, jos katseenseurannan käytön aikana käytetään muita mittareita, jotka tukevat itsessään katseenseurannasta saadun datan tulkittamista. Tästä hyvä esimerkki on EEG:n käyttö, joka mahdollistaa käyttäjän tunnetilojen arvioinnin. Ilman EEG:tä katseenseurannan avulla voidaan saada tietoa siitä, että käyttäjät katsovat jotain kuvaa kauemmin kuin muita. EEG:n avulla voitaisiin kuitenkin mahdollisesti selvittää se, miksi kyseistä kuvaa katsotaan niin kauan. Kuva voisi esimerkiksi olla joidenkin käyttäjien mielestä aiheeseen sopimaton aiheuttaen turhautumista.

Novák, ym. (2023) tutkivatkin katseenseurannan käytön tärkeyttä eri käytettävyydestutkimuksissa, joissa katseenseurantaa hyödynnettiin. Heidän löytämässään ja valitsemisessaan tutkimuksissaan 50 %:ssa katseenseurantaa käytettiin tärkeimpänä, tai yhtenä tärkeimmistä tavoista käyttäjän kokemuksen arviointiin. 24 %:ssa katseenseurantaa käytettiin toissijaisena keinona käytettävyyden arviointiin yleensä muun heuristiikan, vuorovaikutustietojen, haastatteluiden, suoritusasteiden sekä muiden käytettävyydestestauksen keinojen kanssa. 26 %:ssa tutkimuksista ainoastaan kerrottiin katseenseurannan toimintatavoista sekä sen tarjoamista mahdollisuuksista. Heidän

löydöksiin pohjalta voisi siis olettaa, että oikein toteutettu katseenseurannan hyödyntäminen tuo todellista lisäarvoa suurelle osalle käytettävyystudkimuksista sekä sen, että katseenseurantaa on mahdollista käyttää ensisijaisena lähestymistapana tietyissä tapauksissa käytettävyystudkimuksessa.

Tulevaisuudessa katseenseurannan hyödyntämisen kynnyks tulee madaltumaan entisestään, kun katseenseurannan käytöstä tulee helppokäyttöisempää sekä halvempää. Katseenseurannalla on potentiaalia tulla jopa käytettävyystudkimuksen standardikäytännöksi. Lähitulevaisuudessa myös tekoälyn hyödyntäminen katseenseurannan käyttämisessä tulee lisääntymään.

Tutkimuksen havaittuja rajoituksia oli se, että lähteiden testihenkilöt olivat homogeenisiä ja määrät suhteellisen pieniä. Jatkotutkimuksiin olisi suositeltavaa ottaa laajemmalla skaalalla testihenkilöitä, mikäli resurssit ja teknologia sen mahdollistavat. Useassa tässä kirjallisuuskatsauksessa käsitellyssä tutkimuksessa puhuttiin pilottitutkimuksesta, sillä katseenseurannasta löytyvien tutkimusten määrä ja laatu ovat puutteellisia verrattaessa moneen muuhun käytettävyystudkimuksen osa-alueeseen. Aiheeseen on kuitenkin alettu keskittymään viime vuosina enemmän. Jos esimerkiksi Google Scholarista etsii artikkeleita, joissa ”eye tracking” löytyy artikkelin otsikosta vuosien 2010–2012 välillä, tuloksia on noin 1840. Kun sama haku tehdään vuosien 2020–2022 välillä, tuloksia löytyy noin 4580.

Tutkimustavat myös vaihtelivat paljon eri tutkimuksissa, joka on ongelmallista etenkin, kun katseenseurannan käytön isona riskinä on se, ettei tiedetä mitä pitäisi tutkia ja miten, jotta lisäarvoa saataisiin. Jatkotutkimukset voisivat siis keskittyä myös omien katseenseurannalla tehdyn käytettävyystudkimuksen standardien luomiseen ja vertailemiseen, jotta toimivaksi todistettuja menetelmiä voitaisiin koostaa yhteen. Ongelmana on myös erilaisten toteutuksien käyttö katseenseurannan mahdollistamiseen, jotka vaihtelevat tehokkuudeltaan ja luotettavuudeltaan.

Lähteet

Antonenko, P., Celepkolu, M., Fieldman, A., Fieldman, E., Jimenez, Y. & Wang, J. (2018). Exploring Relationships Between Eye Tracking and Traditional Usability Testing Data. *International Journal of Human-Computer Interaction*, 1–12. doi:10.1080/10447318.2018.1464

Bastien, J. M. C. (2010). Usability testing: a review of some methodological and technical aspects of the method. *International Journal of Medical Informatics*, 79(4), e18–e23. doi:10.1016/j.ijmedinf.2008.12.0

Benel, D. C. R., Ottens, D., & Horst, R. (1991). Use of an Eyetracking System in the Usability Laboratory. *Proceedings of the Human Factors Society Annual Meeting*, 35(6), 461–465. doi:10.1518/107118191786754969

Bojko, A. (2005). Eye Tracking in User Experience Testing: How to Make the Most of It. *Proceedings of the 14th Annual Conference of the Usability Professionals' Association (UPA)*. Montreal Canada

Bruun, A., Gull, P., Hofmeister, L., & Stage, J. (2009). Let your users do the testing. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*. doi:10.1145/1518701.1518948

Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. *International Journal of Psychophysiology*. doi:10.1016/j.ijpsycho.2020.05

Chandra, S., Sharma, G., Malhotra, S., Jha, D., & Mittal, A. P. (2015). Eye tracking based human computer interaction: Applications and their uses. *2015 International Conference on Man and Machine Interfacing (MAMI)*. doi:10.1109/mami.2015.7456615

Djamasbi, S. (2014). Eye Tracking and Web Experience. *AIS Transactions on Human-Computer Interaction*, 6(2), 37-54. Viitattu 14.3.2023, saatavilla: <https://aisel.aisnet.org/thci/vol6/iss2/2>

Djamasbi, S., Siegel, M., Skorinko, J., & Tullis, T. (2011). Online Viewing and Aesthetic Preferences of Generation Y and the Baby Boom Generation: Testing User Web Site Experience Through Eye Tracking. *International Journal of Electronic Commerce*, 15(4), 121–158. doi:10.2753/jec1086-4415150404

Falkowska, J., Sobiecki, J., Pietrzak, M. (2016). Eye Tracking Usability Testing Enhanced with EEG Analysis. In: Marcus, A. (eds) *Design, User Experience, and Usability: Design Thinking and Methods. DUXU 2016. Lecture Notes in Computer Science()*, vol 9746. Springer, Cham. https://doi.org/10.1007/978-3-319-40409-7_38

Hiley, J. B., Redekopp, A. H., & Fazel-Rezai, R. (2006). A Low Cost Human Computer Interface based on Eye Tracking. *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. doi:10.1109/iembs.2006.260774

Idrees, A., Kraft, R., Winter, M., Kuchler, A., Baumeister, H., Reilly, R., Reichert, M. & Pryss, R. (2023). Exploring the usability of an internet-based intervention and its providing eHealth platform in an eye-tracking study. *J Ambient Intell Human Comput* **14**, 9621–9636 <https://doi.org/10.1007/s12652-023-04635-4>

Joseph, A. & Muruges, R. (2020) Potential Eye Tracking Metrics and Indicators to Measure Cognitive Load in Human-Computer Interaction Research. *Journal of Scientific Research*, Volume 64, Issue 1, 2020. doi:10.37398/JSR.2020.640137

Kamińska, D., Zwoliński, G., & Laska-Leśniewicz, A. (2022). Usability Testing of Virtual Reality Applications—The Pilot Study. *Sensors*, 22(4), 1342. MDPI AG. doi:<http://dx.doi.org/10.3390/s22041342K>

Kessner, M., Wood, J., Dillon, R. F., & West, R. L. (2001). On the reliability of usability testing. *CHI '01 Extended Abstracts on Human Factors in Computing Systems - CHI '01*. doi:10.1145/634067.634127

Koester, T., Brøsted, J. E., Jakobsen, J. J., Malmros, H. P., & Andreasen, N. K. (2017). The Use of Eye-Tracking in Usability Testing of Medical Devices. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 6(1), 192–199. doi:10.1177/2327857917061042

Nielsen, J. (2020). 10 Usability Heuristics for User Interface Design. Viitattu 14.3.2023, saatavilla: <https://www.nngroup.com/articles/ten-usability-heuristics/>

Novák, J., Masner, J., Benda, P., Šimek, P. & Merunka, V. (2023) Eye Tracking, Usability, and User Experience: A Systematic Review, *International Journal of Human-Computer Interaction*, DOI: [10.1080/10447318.2023.2221600](https://doi.org/10.1080/10447318.2023.2221600)

Schnipke, S. K., & Todd, M. W. (2000). Trials and tribulations of using an eye-tracking system. *CHI '00 Extended Abstracts on Human Factors in Computing Systems - CHI '00*. doi:10.1145/633292.633452

Thyvalikakath TP, Monaco V, Thambuganipalle H, Schleyer T. (2009). Comparative study of heuristic evaluation and usability testing methods. *Stud Health Technol Inform*. 2009;143:322-7. PMID: 19380955; PMCID: PMC27366

