

Johanna Mustaniemi

KÄYTETTÄVYYDEN ARVIOINTIMENETELMÄT

Tämä teos on lisensoitu Creative Commons Nimi mainittava 1.0 Suomi lisenssillä. Nähdäksesi lisenssin vieraile <http://creativecommons.org/licenses/by/1.0/fi/> sivulla tai lähetä kirje Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

Tietojärjestelmätieteen
kandidaatintutkielma
24.3.2009

Jyväskylän yliopisto
Tietojenkäsittelytieteiden laitos
Jyväskylä

TIIVISTELMÄ

Mustaniemi, Johanna Emilia

Tietojärjestelmätieteen kandidaatintutkielma / Johanna Mustaniemi

Jyväskylä: Jyväskylän yliopisto, 2009

47 s.

Kandidaatintutkielma

Tässä tutkielmassa tarkastellaan käytettävyyden arviointimenetelmiä. Työssä esitellään keskeisimmät menetelmät, joilla tuotteiden käytettävyyttä voidaan arvioida ja parantaa. Tutkielman päämääränä on selkeyttää eri arviointimenetelmiä sekä niiden välisiä eroja ja selvittää, mitkä ovat menetelmien vahvuudet ja heikkoudet. Arviointimenetelmiin perehdytään aiempien tutkimusten ja kirjallisuuden pohjalta.

Käytettävyyden arviointimetodit voidaan jakaa kahteen pääryhmään, asiantuntija-arviointeihin sekä empiirisiin käyttäjätesteihin. Asiantuntija-arvioinneissa arviointi suoritetaan ilman käyttäjiä ja tällaisia menetelmiä ovat muun muassa heuristinen arviointi sekä kognitiivinen läpikäynti. Empiiriset käyttäjätestit ovat kokeellisia ja käyttäjät osallistuvat arviointiprosessiin. Käyttäjätestien yleisin menetelmä on käytettävyytestaus.

Tutkielman perusteella voidaan havaita, että erilaisia käytettävyyden arviointimenetelmiä on olemassa lukuisia. Menetelmät eroavat huomattavasti toisistaan, ja kullakin menetelmällä on omat vahvuutensa sekä puutteensa. Eri menetelmät soveltuvatkin erilaisiin tilanteisiin ja niillä voidaan paljastaa hyvin erityyppisiä käytettävyysongelmia. Jotta arviointi olisi mahdollisimman tuloksellista, on suositeltavaa käyttää erilaisia menetelmiä rinnakkain.

AVAINSANAT: Käytettävyys, käytettävyyden arviointi, arviointimenetelmä, asiantuntija-arviointi, käyttäjätesti

SISÄLLYSLUETTELO

1 JOHDANTO	4
2 KÄYTETTÄVYYDEN TAUSTAA.....	7
3 KÄYTETTÄVYYDEN MÄÄRITELMÄ.....	8
3.1 Käytettävyys ISO-standardin mukaan	8
3.2 Käytettävyys Nielsenin mukaan	10
3.3 Käytettävyyden osa-alueet Hyysalon mukaan	12
4 ARVIOINTIMENETELMÄT	15
4.1 Asiantuntija-arvioinnit.....	17
4.1.1 Heuristinen arviointi	19
4.1.2 Kognitiivinen läpikäynti	22
4.1.3 Muita menetelmiä	23
4.2 Empiiriset käyttäjätestit	26
4.2.1 Käytettävyystestaus	29
4.2.2 Muita menetelmiä	33
4.2.3 Tiedonkeruumenetelmiä.....	36
5 YHTEENVETO	39
LÄHDELUETTELO	42

1 JOHDANTO

Käytettävyydeltään hyvä tuote luo käyttäjälle mieluisamman käyttökokemuksen, ja käyttäjä on tyytyväinen tuotteeseen. Joudumme kuitenkin niin kotona kuin töissäkin käyttämään jatkuvasti tuotteita, joiden käytettävyys on huonoa. Tämä aiheuttaa muun muassa turhautumista, stressiä sekä ajan tuhlaantumista. (Sinkkonen, Kuoppala, Parkkinen & Vastamäki 2006, 271)

Käytettävyys (engl. usability) kuvaa yleisesti ottaen, kuinka helppoa tuotetta on käyttää (McNamara & Kirakowski 2005). Käytettävyyden voidaankin sanoa olevan jokaisen tuotteen laatuominaisuus, jolla ilmaistaan, kuinka helppoa ja tehokasta tuotteen käyttö on (Sinkkonen 2004).

Käytettävyyteen tulee kiinnittää huomioita niin fyysisen laitteen kuin tietokoneohjelmistonkin suunnittelussa. Viime vuosina myös useat yritykset ovat huomanneet käytettävyyden merkityksen, ja käytettävyys on tullut yhä olennaisemmaksi osaksi sekä suunniteltaessa uusia tuotteita että kehitettäessä olemassa olevia. Nielsen (1993, 2-3) on todennut, että käytettävyyteen panostamisella voidaan saavuttaa huomattavia kustannussäästöjä, vaikka kulujen pieneneminen voi välillä olla epäsuoraa ja vaikeasti mitattavissa. Hänen mukaansa säästöä syntyy esimerkiksi tuotteen käytön oppimisen nopeutuessa ja virheiden sekä teknisen tuen tarpeen vähentyessä. Lisäksi tuotteen käytön on todettu tehostuvan ja eri ominaisuuksien käyttöasteen kasvavan (Sinkkonen ym. 2006, 280). Käytettävyyteen panostaminen lisää myös tuotteen menestymismahdollisuuksia (Väyrynen, Nevala & Päivinen 2004, 8) ja tuo yritykselle huomattavaa kilpailuetua (Nielsen 1993, 4).

Käytettävyydestä oikeiden käyttäjien kanssa on oikein toteutettuna eräs luotettavimmista keinoista arvioida ja parantaa tuotteen käytettävyyttä (Riihiahho 1998). Käytettävyydestä onkin tunnetuin ja useimmiten

hyödynnetty käytettävyyden arviointimenetelmä (Nielsen 1995). Käytettävyydestä ei kuitenkaan aina ole tehokkain menetelmä käytettävyyden arviointiin. Se voidaan kokea esimerkiksi liian hitaaksi ja kalliiksi menetelmäksi tai tuotteen suunnittelu on niin alkuvaiheessa, ettei käytettävyydestä tekeminen käyttäjien kanssa ole mahdollista. (Korvenranta 2005) Käytettävyyttä voidaan kuitenkin arvioida ja kehittää useiden muidenkin menetelmien avulla.

Tämän tutkielman tarkoituksena on esitellä lukijoille keskeisimmät käytettävyyden arviointimenetelmät. Arviointimenetelmien kirjo on monipuolinen, eikä suomalaisessa kirjallisuudessa ole juurikaan esitelty eri menetelmiä kokoavasti yhdessä. Lisäksi useat suomenkieliset käytettävyyttä käsittelevät teokset keskittyvät useimmiten tarkastelemaan ainoastaan käytettävyydestä, unohtaen täysin muut arviointimenetelmät.

Tutkielman keskeisenä tavoitteena on selvittää, mitä vahvuuksia ja heikkouksia erilaisilla menetelmillä on. Menetelmien vahvuuksia ja heikkouksia tarkastellaan ensisijaisesti pääryhmien välillä. Tutkielmassa pyritään myös selventämään termien käyttöä, sillä joitakin aiheeseen liittyviä termejä ja käsitteitä käytetään ajoittain ristiriitaisesti päällekkäin. Esimerkiksi käyttäjätesti sekä käytettävyydestä ovat termejä, joiden käyttö on usein sekavaa.

Tässä tutkielmassa puhutaan yleisesti ottaen tuotteen käytettävyydestä. *Tuotteella* (engl. product) voidaan tarkoittaa niin fyysistä laitetta, käyttöliittymää tai vaikkapa ohjelmistoa. Tutkielmassa keskitytään tarkastelemaan erityisesti graafisen käyttöliittymän käytettävyyttä sekä käytettävyyden arviointia. Useimmat esitetyt menetelmät sopivat kuitenkin myös muunlaisten tuotteiden käytettävyyden arviointiin.

Käyttöliittymä (engl. user interface) tarkoittaa suppeasti määriteltynä tietokonesovelluksen osaa, jolla käyttäjä on yhteydessä sovelluksen tai laitteen kanssa (Kalimo 1996). Käyttöliittymä on siis se osa, jonka käyttäjä näkee ja jonka

kautta hän käyttää sovellusta. *Graafinen käyttöliittymä* (engl. graphical user interface, GUI) on ikkunoista, valikoista sekä erilaisista kuvakkeista koostuva käyttöliittymä (Kalimo 1996). Siinä komennot toteutetaan muun muassa erilaisissa valikoissa esimerkiksi hiirtä tai näppäimistöä hyödyntäen (Paananen 2001, 435).

Luvussa kaksi käsitellään lyhyesti käytettävyyden historiaa käyttöliittymien näkökulmasta. Lisäksi luvussa tutustutaan ihmisen ja tietokoneen vuorovaikutukseen, joka on käytettävyyteen läheisesti liittyvä monitieteellinen tieteenala.

Ennen kuin käytettävyyttä voidaan arvioida, on oleellista määritellä, mitä käytettävyydellä oikeastaan tarkoitetaan ja millaisista ominaisuuksista se muodostuu. Luvussa kolme käytettävyyttä tarkastellaan ISO-standardin, Jakob Nielsenin sekä suomalaisen FT Sampsa Hyysalon mukaan.

Luvussa neljä perehdytään käytettävyyden arviointimenetelmiin. Arviointimenetelmät jaetaan kahteen pääryhmään, asiantuntija-arviointeihin sekä empiirisiin käyttäjätesteihin. Asiantuntija-arvioinneista käsitellään tarkemmin heuristinen arviointi sekä kognitiivinen läpikäynti ja kokeellisista käyttäjätesteistä käytettävyydestä.

Koska käytettävyyden arviointimenetelmistä on tehty lukuisia muunnelmia, ei tässä tutkielmassa ole mahdollista esitellä kaikkia olemassa olevia arviointimenetelmiä ja niiden variaatioita. Tutkielmaan on kuitenkin pyritty keräämään kaikki yleisimmät menetelmät. On myös hyvä huomioda, että uusia muunnoksia syntyy jatkuvasti käytettävyydsiantuntijoiden soveltaessa olemassa olevia menetelmiä erilaisiin arviointitilanteisiin.

2 KÄYTETTÄVYYDEN TAUSTAA

Ensimmäiset käyttöliittymien käytettävyyttä koskevat tieteelliset tutkimukset toteutettiin jo 1950- ja 60-luvuilla. Tuolloin tutkijat yliopistoissa ja alan yrityksissä aloittivat käyttöliittymiä koskevat tutkimustyöt, joilla on ollut myöhemmin erittäin keskeinen asema käyttöliittymiä kehitettäessä. Kyseiset tutkimukset käsittelivät muun muassa suoran manipuloinnin sekä ikkunoiden hyödyntämistä käyttöliittymissä. (Myers 1998)

Käytettävyyteen liittyviä ongelmia alettiin tutkia aktiivisemmin 1960- ja 1970-lukujen vaihteessa, jolloin puhuttiin käyttäjäystävällisyydestä (engl. user friendly). Alan asiantuntijat totesivat kuitenkin kyseisen termin soveltumattomaksi ja alettiin puhua muun muassa ihmisen ja koneen vuorovaikutuksesta (engl. human-computer interaction, HCI) sekä käyttäjäkeskeisestä suunnittelusta (engl. user-centered design, UCD). (Nielsen 1993, 23)

Ihmisen ja koneen vuorovaikutus on monitieteellinen tieteenala, joka tutkii vuorovaikutteisten tietojärjestelmien suunnittelua, arviointia sekä toteutusta. HCI:n tavoitteena on kehittää järjestelmien käytettävyyttä, jotta käyttäjät voivat suorittaa toimintonsa tuloksellisesti, tehokkaasti sekä turvallisesti. (Preece, Rogers, Sharp, Benyon ym. 1994)

Ihmisen ja koneen vuorovaikutusta sekä käytettävyyttä pidetään usein lähes samana asiana, mutta osa alan asiantuntijoista näkee, että käsitteillä olisi myös oleellisia eroavaisuuksia. Muun muassa Sinkkonen ym. (2006, 18) esittävät, että käytettävyys huomioi ihminen osana organisaatiota ja tahtovana toimijana, kun HCI ei ota tätä näkökulmaa useinkaan huomioon. Preece ym. (1994, 14) puolestaan näkevät käytettävyyden pikemminkin HCI:n avainkäsitteenä eli osana ihmisen ja koneen vuorovaikutusta.

3 KÄYTETTÄVYYDEN MÄÄRITELMÄ

Käytettävyyttä on tutkittu tieteellisesti useiden vuosikymmenten ajan ja sille on esitetty useita eri määritelmiä. Käytettävyys on sellaisenaan suhteellisen abstrakti ominaisuus, joten se puretaan määrittelyissä usein konkreettisiksi osakokonaisuuksiksi (Sinkkonen 2004).

Kaksi useimmiten käytettyä määritelmää ovat kansainvälisen standardointiorganisaatio ISO:n (International Organization for Standardization) määritelmä sekä Jakob Nielsenin käytettävyyden määritelmä, joissa molemmissa käytettävyys jaetaan pienemmiksi ominaisuuksiksi eli attribuuteiksi. Uudempaa näkökulmaa edustaa muun muassa Sampsa Hyysalon määritelmä, jossa käytettävyys jaetaan kuuteen osa-alueeseen.

3.1 Käytettävyys ISO-standardin mukaan

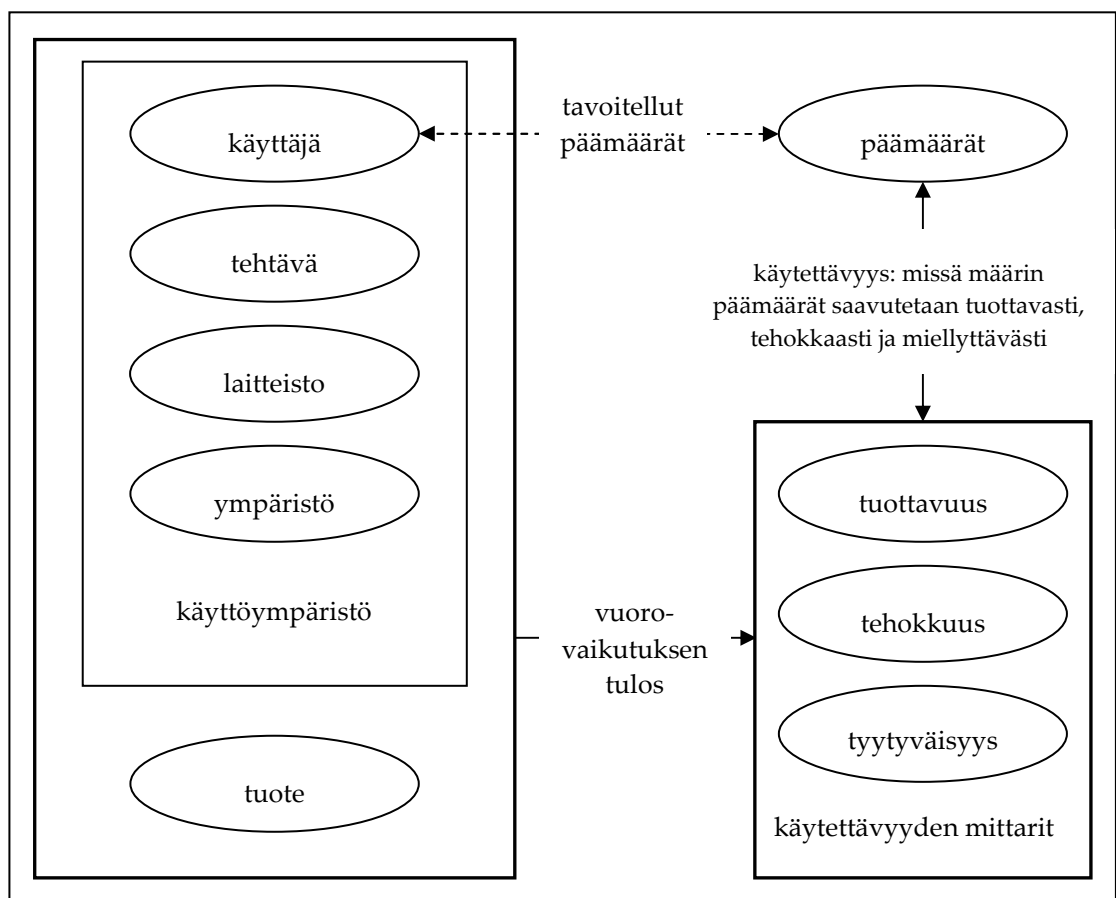
ISO 9241-standardi käsittelee näyttöpäätteillä tehtävän toimistotyön ergonomisia vaatimuksia. Standardin 11. osassa määritellään käytettävyys ja sen arviointi. Standardin mukaan käytettävyydellä tarkoitetaan sitä, kuinka tuottavasti, tehokkaasti ja miellyttävästi tietyt käyttäjät voivat käyttää tuotetta määriteltujen tavoitteiden saavuttamiseksi tietyssä käyttöympäristössä. Kyseisessä määritelmässä käytettävyydelle annetaan siis kolme keskeistä attribuuttia:

- tuottavuus (engl. effectiveness)
- tehokkuus (engl. efficiency)
- miellyttävyys (engl. satisfaction). (ISO 1998)

Tuottavuudella tarkoitetaan sitä, missä määrin ja kuinka täydellisesti käyttäjät saavuttavat määritellyt tavoitteensa tuotteen avulla. Kyseisestä attribuutista käytetään usein myös suomennuksia tarkkuus sekä vaikuttavuus. *Tehokkuus* kuvaa paljonko resursseja käytetään suhteessa tavoitteiden saavuttamiseen

tuottavasti. *Miellyttävyys* eli tyytyväisyys puolestaan ilmentää tuotteen käytön hyväksyttävyyttä sekä mukavuutta. (ISO 1998)

ISO standardi (1998) esittelee myös käytettävyyden kehikon, jossa määritellään käytettävyyteen liittyvät elementit sekä niiden väliset suhteet (KUVA 1). Standardin mukaan on oleellista tunnistaa päämäärät sekä määritellä mittarit ja käyttöympäristön elementit osatekijöihin, jotta käytettävyys voidaan määritellä ja jotta sitä voidaan arvioida.



KUVA 1. Käytettävyyden rakenne ISO-standardin (1998) mukaan (termit suomennettu).

3.2 Käytettävyys Nielsenin mukaan

Toinen yleisesti käytetty määritelmä on Jakob Nielsenin vuonna 1993 luoma käytettävyyden määritelmä. Nielsenin määritelmässä käytettävyydelle nimetään useampia attribuutteja kuin ISO-standardissa. Nielsen jättää kuitenkin lähes kokonaan huomioimatta tuottavuuden eli sen, kuinka hyvin käyttäjät saavuttavat päämääränsä tuotetta käyttämällä. Tästä huolimatta Nielsenin määritelmää voidaan pitää selkeämpänä ja kattavampana kuin ISO:n määritelmää.

Nielsenin (1993, 26) mukaan käytettävyys sisältää viisi laatukomponenttia.

Hänen esittämät käytettävyyden attribuutit ovat:

- opittavuus (engl. learnability)
- tehokkuus (engl. efficiency)
- muistettavuus (engl. memorability)
- virheettömyys (engl. errors)
- tyytyväisyys (engl. satisfaction).

Opittavuus tarkoittaa, että tuotteen käytön tulee olla helppo oppia. Tällöin käyttäjä voi välittömästi aloittaa tuotteen käytön ja saada tuloksia aikaiseksi. Opittavuutta voidaan pitää yhtenä keskeisempänä käytettävyyden attribuuttina ja se on myös yksi helpoin tekijä käytettävyyden mittaamiseen. Opittavuutta tarkastellessa on tärkeää huomioida, että oppiminen on useimmiten alussa nopeaa, mutta tasaantuu jatkossa, kun tuotteen käyttö hiljalleen opitaan. Oppimisen mittaamisessa voidaan tarkastella aikaa, joka kuluu siihen, kun käyttäjä saavuttaa etukäteen määritellyn työtehon. (Nielsen 1993, 27–29)

Tehokkuus käsittelee sitä, kuinka korkeaan tuottavuuteen käyttäjän on mahdollista päästä sen jälkeen, kun hän on saavuttanut tietyn oppimistason tuotteen käytössä. Oppimistason määrittelyyn on tilanteesta riippuen erilaisia tarkastelutapoja. Joissain tapauksissa tarkastellaan esimerkiksi sitä, kuinka

kauan käyttäjä on tuotetta käyttänyt tai millä tasolla hänen oppimisensa on verrattaessa saavutettua oppimistasoa Nielsenin oppimiskäyrään (ks. Nielsen 1993, 28–31). Tehokkuutta voidaan arvioida esimerkiksi mittaamalla aika, joka menee määritellyn tehtävän suorittamiseen. (Nielsen 1993, 30–31)

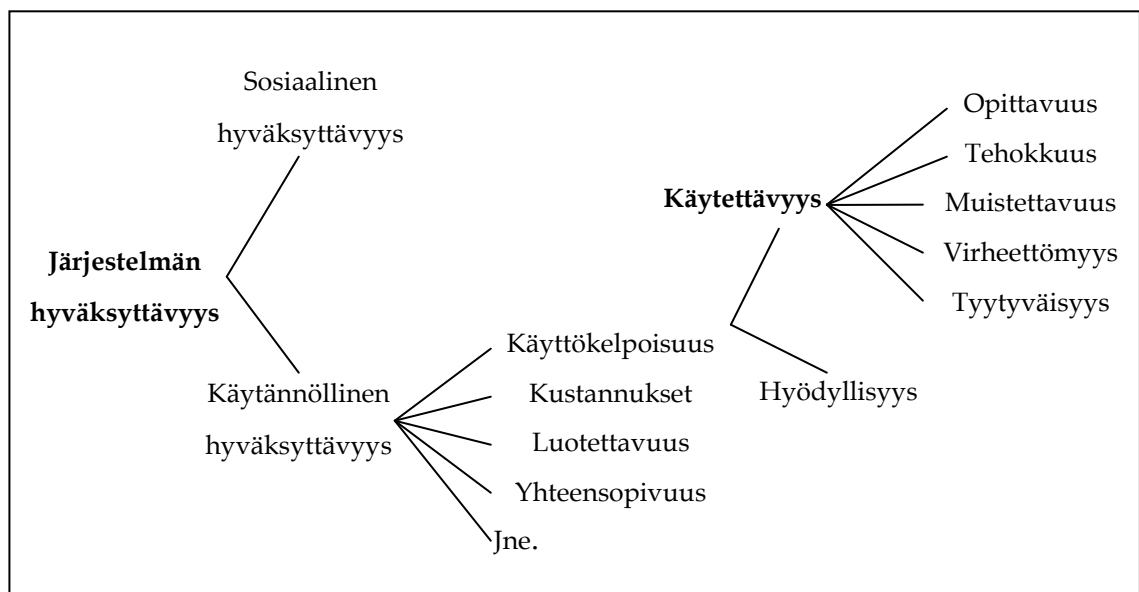
Muistettavuus tarkoittaa, että tuotteen käytön tulisi olla helppo muistaa. Tällöin käyttäjän ei tarvitse opetella käyttöä uudelleen, vaikkei hän olisikaan käyttänyt tuotetta hetkeen. Muistettavuus korostuu erityisesti tarkastellessa satunnaisia käyttäjiä, mutta se koskee myös esimerkiksi lomalta palaavia työntekijöitä. Käytettävyyttä arvioidessa muistettavuutta tarkastellaan harvemmin kuin muita käytettävyyden attribuutteja. (Nielsen 1993)

Virheettömyys merkitsee, että käyttäjän tulisi tehdä mahdollisimman vähän virheitä käyttäessään tuotetta. Virhe määritellään yleensä toiminnoksi, joka ei johda haluttuun lopputulokseen. Virheettömyyttä voidaan arvioida esimerkiksi laskemalla käyttäjän tekemät virheelliset toiminnot. Lisäksi tulee huomioida virheen vakavuus arvioitaessa tuotteen käytettävyyttä. (Nielsen 1993, 32–33)

Tyytyväisyys tarkoittaa, että tuotteen käytön tulee olla miellyttävää käyttäjälle. Käyttäjien tulee siis pitää tuotteesta ja heidän tulee kokea tyytyväisyyttä käyttäessään sitä. Tyytyväisyyden mittaaminen on useimmiten tehokkainta toteuttaa erilaisilla kyselykaavakkeilla. Eräs suosittu menetelmä on Likert-asteikon käyttäminen, jossa käyttäjä arvioi erilaisten väitteiden paikkansapitävyyttä omien kokemustensa pohjalta. (Nielsen 1993, 33–35)

On muistettava, että suunnittelu- ja kehitystyössä tulee olla valmis myös kompromisseihin, sillä kaikkia tekijöitä ei aina voida saada hyväksi yhtä aikaa. Esimerkiksi opittavuuden ja tehokkuuden välillä joudutaan usein joustamaan toisen osa-alueen haitaksi. Aina tulisi kuitenkin pyrkiä lopputulokseen, jossa kaikki käytettävyyden komponentit huomioidaan riittävästi ja käytettävyyden osa-alueiden vaatimukset saadaan toteutettua. (Nielsen 1993, 40–42)

Nielsen (1993) huomauttaa käytettävyyden olevan kuitenkin osa suurempaa kokonaisuutta, järjestelmän hyväksyttävyyttä (KUVA 2). *Järjestelmän hyväksyttävyyys* (engl. system acceptability) kuvaa sitä, onko järjestelmä kokonaisuudessaan tarpeeksi hyvä toteuttaakseen käyttäjän kaikki tarpeet ja vaatimukset (Nielsen 1993, 24). Tässä tutkielmassa keskitytään kuitenkin käytettävyyteen sinällään, eikä muita järjestelmän hyväksyttävyyden osa-alueita tarkastella enempää.



KUVA 2. Järjestelmän hyväksyttävyyden osatekijät Nielsenin (1993, 25) mukaan (termit suomennettu).

3.3 Käytettävyyden osa-alueet Hyysalon mukaan

Yksi uudemmissa käytettävyyden määritelmistä on suomalaisen FT Sampsa Hyysalon kuvaus käytettävyyden osa-alueista. Hyysalo tarkastelee käytettävyyttä hieman eri näkökulmasta kuin edellä ISO-standardi ja Nielsen. Hyysalon lähestymistavan voidaan katsoa olevan käytännönläheisempi ja yksityiskohtaisempi, mikä tuo uutta ja erilaista näkökantaa käytettävyyden osa-alueiden määrittelyyn. Vaikka Hyysalo on keskittynyt käsittelemään käytettävyyttä pääasiassa fyysisten laitteiden näkökulmasta, ovat hänen

esittämät osa-alueet sovellettavissa erittäin hyvin myös muiden tuotteiden, kuten käyttöliittymien, käytettävyyteen.

Hyysalo (2006) on esittänyt käytettävyyden koostuvan kuudesta eri osa-alueesta. Nämä osa-alueet ovat:

- toimintojen vastaavuus
- ryhmittely
- liikkuminen
- tuotteen vastaavuus käyttäjän tottumuksiin
- graafinen suunnittelu ja värit
- nimeäminen ja symbolit.

Toimintojen vastaavuus tarkoittaa, että tuotteen toimintojen tulee vastata siihen, mitä käyttäjät todella pyrkivät tuotteella tekemään. On erittäin yleistä, että tuotteeseen on sisällytetty paljon tarpeettomia elementtejä, mutta samalla siitä puuttuu käyttäjän kannalta tärkeitä toimintoja. (Hyysalo 2006, 159)

Ryhmittely sisältää toimintojen sekä kenttien ryhmittelyn. Tuotteet, niin fyysiset laitteet kuin ohjelmistotkin, koostuvat useimmiten eri osioista sekä toiminnoista ja näiden oikeanlainen ryhmittely on yksi oleellisimpia käytettävyyteen liittyvistä tekijöistä. Jotta ryhmittelyssä onnistutaan, tulee tärkeimmät asiat tuoda parhaiten näkyville ja samaan kokonaisuuteen liittyvät toiminnot ryhmitellä yhtenäisesti. Lisäksi on tärkeää sijoittaa ryhmät siten, että käyttäjä voi siirtyä luontevasti ryhmästä toiseen. (Hyysalo 2006, 159)

Liikkuminen käsittää tuotteen osien sisällä ja osasta toiseen liikkumisen. Tuotteen tulisi kertoa aina missä paikassa ja tilassa ollaan, mihin käyttäjän tulisi edetä päästäkseen päämääräänsä ja milloin käyttäjä on onnistuneesti saanut tuotteen tekemään jonkin toiminnon. Lisäksi tuotteen tulisi ilmaista käyttäjälle, että on turvallista kokeilla eri toimintoja eli käyttäjän on mahdollista peruuttaa tekemänsä toimenpiteet. (Hyysalo 2006, 160)

Tuotteen vastaavuus käyttäjän tottumuksiin sisältää myös käyttäjän kokemukset aiemmista vastaavista tuotteista. Kyseinen osa-alue on yksi keskeisimmistä käytettävyyteen vaikuttavista tekijöistä, mutta valitettavan usein kyseinen näkökulma unohdetaan. Oleellinen merkitys on sillä, miksi käyttäjät tuotteen todella mieltävät, sillä käyttäjät suhtautuvat uusiin tuotteisiin aiempien kokemustensa perusteella. (Hyysalo 2006, 160)

Graafisella suunnittelulla ja värityksellä voidaan myös parantaa käytettävyyttä. Niiden avulla voidaan esimerkiksi ryhmitellä ja erotella toimintoja. (Hyysalo 2006, 160)

Nimeäminen ja symbolit ovat usein suunnittelijoiden kompastuskivinä. Suunnittelijat erehtyvät usein kuvittelemaan, että pelkästään toimintojen nimeämisellä ja ohjeiden kirjoittamisella he onnistuvat luomaan käytettävyydeltään hyvän tuotteen. Nimeämisellä ja symboleiden suunnittelemisella voidaan kuitenkin korkeintaan paikata muita puutteita käytettävyydessä, kun taas huonolla nimeämisellä saadaan useimmiten aikaiseksi vain lisää ongelmia. Hyvien termien tuleekin aina vastata kyseisen kohderyhmän käyttäjien käsityksiä toiminnoista. (Hyysalo 2006, 161)

Koska Hyysalon esittämät käytettävyyden osa-alueet ovat huomattavasti käytännönläheisempiä sekä konkreettisempia kuin ISO:n ja Nielsenin määritelmät, ei niitä voi juurikaan verrata keskenään. Hyysalon määritelmässä voidaan kuitenkin nähdä olevan paljon yhdenmukaisuuksia erilaisiin heuristiikkalistoihin, joita käsitellään tarkemmin luvussa 4.1.1 Heuristinen arviointi.

4 ARVIOINTIMENETELMÄT

Käytettävyyden arviointimenetelmiä (engl. usability evaluation methods) on kehitetty lukuisia ja myös niiden luokittelusta on esitetty erilaisia jaotteluja. Riihiahon (2000b) mukaan menetelmät voidaan jakaa kahteen pääryhmään, asiantuntija-arviointeihin sekä empiirisiin käyttäjätesteihin. Jako määritellään sen perusteella, osallistuvatko käyttäjät käytettävyyden arviointiprosessiin.

Asiantuntija-arvioinneissa tuotteen käytettävyyttä arvioivat asiantuntijat, eivätkä käyttäjät ole mukana arvioinnissa (Korvenranta 2005). Empiirisissä eli kokeellisissa käyttäjätesteissä käytettävyyttä arvioidaan testaamalla tuotetta oikeiden käyttäjien kanssa (Nielsen 1995), jolloin käyttäjät ovat keskeisessä osassa arviointiprosessissa.

Käytettävyyden arviointimenetelmiä voidaan ryhmitellä myös muilla perusteilla. Esimerkiksi Nielsen (1995) on esittänyt menetelmien ryhmittelyä neljään eri luokkaan. Nämä luokat ovat automaattinen (engl. automatic), empiirinen (engl. empirical), muodollinen (engl. formal) sekä epämuodollinen (engl. informal) arviointi. Automaattisessa arvioinnissa käyttöliittymää arvioidaan erityisillä arviointiohjelmistoilla, empiirisessä arvioinnissa todellisten käyttäjien kanssa ja muodollisessa arvioinnissa hyödynnetään matemaattisia malleja sekä kaavoja (Nielsen 1995). Epämuodollinen arviointi perustuu sääntöihin eli heuristiikkoihin sekä arvioijien kokemuksiin ja tietämykseen. Epämuodollisella arvioinnilla tarkoitetaan siis nimenomaan asiantuntija-arviointeja. (Mack & Nielsen 1994, 2)

Nielsen (1995) kritisoi kuitenkin automaattisia sekä muodollisia arviointimenetelmiä, eivätkä ne hänen mielestään ole toimivia menetelmiä käytettävyyden arviointiin. Tässä työssä keskitytäänkin Riihiahon (2000b) esittämään jaotteluun arviointimenetelmien pääryhmistä. Kyseisen jaottelun

voidaan katsoa vastaavan myös Nielsenin ryhmittelyä, kun toimimattomiksi todetut automaattiset ja muodolliset menetelmät jätetään huomioimatta.

Oikean menetelmän valinta voi kuitenkin usein olla ongelmallista, koska eri arviointimenetelmiä on olemassa lukuisia ja jokaisella menetelmällä on hieman erilainen tarkoitus (Rubin 1994, 30). Menetelmien luokittelut helpottavat osaltaan menetelmän valintaa, mutta arvioinnin organisoijan tulee olla hyvin perillä eri menetelmistä ja niiden valintakriteereistä, jotta hän osaisi valita oikeanlaisen menetelmän kuhunkin tilanteeseen (Bolchini & Garzotto 2008).

Menetelmän valintaan vaikuttavat useat eri tekijät. On oleellista miettiä, millaisiin käytettävyystekijöihin arvioinnissa halutaan erityisesti kiinnittää huomiota, sillä jokaisella menetelmällä on erilainen päämäärä (Rubin 1994, 26). Jos arvioinnin pääasiallisena tavoitteena on esimerkiksi arvioida tuotteen käytön opittavuutta, menetelmän tulee olla eri, kuin jos tarkastelun kohteena on käytön tehokkuus. Menetelmän valintaan vaikuttaa myös arvioinnin tavoite: onko arvioinnin tarkoituksena kehittää jo olemassa olevaa tuotetta käytettävämmäksi vai arvioidaanko uutta ideaa. Greenberg ja Buxton (2008) korostavatkin, että arviointimenetelmä täytyy valita siten, että se on sopiva ja tarkoituksenmukainen selvitettävään ongelmaan tai tutkimuskysymykseen nähden – muutoin arvioinnin tekeminen on tuloksetonta.

Myös käytössä olevat resurssit vaikuttavat menetelmän valintaan. Oleellisimmat resurssit ovat useimmiten aika sekä raha. Lisäksi asiantuntijoiden ja mahdollisten testikäyttäjien saatavuus sekä tilat ja välineet tulee huomioida. Menetelmän valintaan vaikuttaa usein myös arvioitavan tuotteen kehitysaste. Mikäli tuotteesta ei ole kehitetty edes prototyyppiä, voi arviointi olla lähes mahdotonta loppukäyttäjien kanssa. (Bødker & Madsen 1998; Nielsen 1993, 225; Rubin 1994, 26)

Jokaisella arviointimenetelmällä on omat vahvuutensa ja heikkoutensa. Tutkimuksissa on todettu, että asiantuntija-arvioinnin menetelmillä voidaan

löytää hyvinkin erilaisia ongelmia kuin empiirisillä käyttäjätesteillä. Käyttäjätesteillä voidaan puolestaan päästä erilaisiin tuloksiin kuin asiantuntija-arvioinneilla. Usein onkin suositeltavaa valita useampi menetelmä ja yhdistellä eri menetelmiä, jotta arviointi olisi mahdollisimman tehokasta sekä tuloksellista. (Nielsen 1995)

4.1 Asiantuntija-arvioinnit

Asiantuntija-arvioinnit (engl. usability inspection methods) ovat käytettävyyden tutkimusmenetelmiä, jolla voidaan kartoittaa tuotteen käytettävyyttä (Ranne 2005). Käytettävyyden arviointi perustuu asiantuntijan tai asiantuntijaryhmän arvioon tuotteesta sekä sen käytöstä (Korvenranta 2005). Erilaisia asiantuntija-arvioinnin menetelmiä ovat muun muassa kognitiivinen läpikäynti, standardikatselmukset (Mack & Nielsen 1994, 5-6) ja heuristinen arviointi, joka on tunnetuin asiantuntija-arvioinnin menetelmä (Riihiaho 1998).

Käytettävyyteen ja sen arviointiin tulee kiinnittää huomiota koko tuotekehitysprosessin ajan, ja asiantuntija-arviointimenetelmien avulla on mahdollista paneutua käytettävyyteen jo varhaisessa vaiheessa (Riihiaho 1998). Tyypillisesti asiantuntija-arvioinnin menetelmiä hyödynnetäänkin tuotteen suunnitteluvaiheessa (Nielsen 1995), mutta niitä voidaan käyttää myös tuotteen elinkaaren muissa vaiheissa, esimerkiksi arvioitaessa valmiin tuotteen käytettävyyttä. Tällöin tuloksia hyödynnetään kehitystyössä. (Korvenranta 2005)

Asiantuntija-arvioinnit ovat vapaamuotoisia arviointimenetelmiä, joilla käytettävyyttä voidaan arvioida helposti ja taloudellisesti (Nielsen 1995). Asiantuntija-arvioinnit ovat myös nopeita oppia sekä toteuttaa, sillä ne eivät vaadi juurikaan etukäteisvalmistelua ja niiden avulla tuotteen käytettävyys voidaan arvioida jopa yhdessä päivässä (Korvenranta 2005).

Asiantuntija-arvioinnin menetelmiä on kuitenkin kritisoitu siitä, että niiden avulla löydetään useimmiten vähemmän merkityksellisiä käytettävyyssongelmia, jolloin todelliset ongelmat voivat jäädä havaitsematta. Menetelmiä on arvosteltu myös niin sanotuista vääristä hälytyksistä (engl. false-alarms), joilla tarkoitetaan erittäin vähäpätöisiksi luokiteltavien ongelmien tai täysin virheellisten asioiden raportointia. (Molich & Jeffries 2003) Lisäksi arvioijien kokemuksilla on havaittu olevan huomattava vaikutus arviointiin ja arvioinnista saataviin tuloksiin (Hertzum & Jacobsen 2003).

Riihiaho (1998) onkin todennut, että useimmiten käytettävyyttä arvioidaan luotettavimmin oikeiden käyttäjien kanssa, sillä nämä tuovat arviointiin mukaan tietämyksen muun muassa työtehtävistään sekä työtavoistaan. Voidaankin sanoa, että juuri se, mikä tekee asiantuntija-arvioinneista niin kustannustehokkaita menetelmiä, on samalla niiden suurin heikkous.

Myös Cockton ja Woolrych (2002) ovat kritisoineet voimakkaasti asiantuntija-arviointimenetelmiä kutsuen niitä halpakäytettävyyssmetodeiksi (engl. discount usability methods). Heidän mielestään kyseisiä menetelmiä ei tulisi käyttää lainkaan sellaisenaan virhealittiutensa vuoksi. Halpametodi-termin on tosiasiaassa esittänyt Nielsen jo useita vuosia aikaisemmin. Nielsen (1994) on kuitenkin todennut, että halvasta maineestaan huolimatta menetelmien avulla on mahdollista löytää huomattava määrä käytettävyyssongelmia. Kaiken lisäksi ongelmat voidaan löytää jo hyvin varhaisessa kehitysvaiheessa, ja esimerkiksi heuristisen arvioinnin on todettu olevan erittäin tehokas käytettävyyden arviointimenetelmä (Nielsen 1994).

Mitä asiantuntijalla tarkoitetaan, vaihtelee paljon menetelmäkohtaisesti, sillä useat menetelmät eivät vaadi esimerkiksi tuotteen alaan tai käytettävyyteen liittyvää asiantuntijuutta. Useimmat asiantuntija-arvioinnin menetelmistä ovat niin helppoja ja nopeita oppia, että esimerkiksi tuotteiden suunnittelijat voivat suorittaa arvoinnit helposti (Nielsen 1995). Arvioijat voivatkin olla esimerkiksi

käytettävyyden asiantuntijoita, kyseisen tuotteen asiantuntijoita tai tarkasteltavan alan asiantuntijakonsultteja (Mack & Nielsen 1994, 1). Nielsen (1995) on kuitenkin todennut, että mikäli arvioinnin suorittaa tai siihen osallistuu käytettävyydsiantuntija, saavutetaan useimmiten parempia tuloksia.

4.1.1 Heuristinen arviointi

Heuristinen arviointi (engl. heuristic evaluation) on käytettävyyden arviointimenetelmä, jossa yksi tai useampi asiantuntija tarkastelee järjestelmän käyttöliittymää tutkien, kuinka hyvin se täyttää tietyt käytettävyydsperiaatteet eli heuristiikat (Dumas & Redish 1999, 65). *Heuristiikat* (engl. heuristics) ovat lista sääntöjä ja ohjeita, joita käyttöliittymän tulisi noudattaa ollakseen hyvä käytettävyydeltään (Kuutti 2003, 47).

Heuristista arviointia ei välttämättä pidetä aina yhtä luotettava arviointimenetelmänä kuin käytettävyyden arviointia oikeiden käyttäjien kanssa. Menetelmällä on kuitenkin omat vahvuutensa, joiden vuoksi se on yksi käytetyimpiä arviointimenetelmiä. Ensinnäkin heuristinen arviointi voidaan suorittaa jo tuotteen varhaisessa kehitysvaiheessa (Nielsen & Molich 1990), mikä on harvoin mahdollista oikeiden käyttäjien kanssa. Kun havaitut käytettävyydsongelmat korjataan mahdollisimman aikaisessa vaiheessa, voidaan saavuttaa huomattavia säästöjä (Kuutti 2003, 48). Heuristinen arviointi on myös halpa ja nopea toteuttaa, sillä varsinaisia etukäteisjärjestelyjä ei juurikaan tarvita. Lisäksi heuristinen arviointi on intuitiivinen ja siihen on helppo motivoida osallistujia. (Nielsen & Molich 1990)

Heuristisessa arvioinnissa asiantuntijan määritelmä vaihtelee suuresti, sillä arvioinnin voi periaatteessa suorittaa myös henkilö, jolla ei ole ennestään kokemusta käytettävyydestä tai sovellusalueesta (Kuutti 2003, 49). Arviointi ei kuitenkaan ole tällöin kovin tehokasta, sillä on todettu, että tällainen henkilö löytää vain noin 22 % käytettävyydsongelmista. Mikäli arvioijalla on vahva

tietämys käytettävyydestä, löydettävien käytettävyyssongelmien määrä lähes kaksinkertaistuu, ja arviointi on huomattavasti tehokkaampaa. (Nielsen 1993, 161) Tehokkaimpiin lopputuloksiin päästään, mikäli arvioija on sekä sovellusalueen että käytettävyyden asiantuntija (Riihiaho 1996). Tällöin yksi arvioija voi löytää jopa 60 % käytettävyyssongelmista (Nielsen 1993, 161). Tällaisia niin sanottuja tuplaeksperttejä (Kuutti 2003, 49) on kuitenkin usein vaikea löytää arvioijiksi (Riihiaho 1996). On kuitenkin suositeltavinta, että arvioijiksi pyritään saamaan henkilöitä, joilla on ennestään vahva tietämys käytettävyydestä (Nielsen 1993, 162).

Nielsen (1993, 156) on todennut, että yksi arvioija pystyy löytämään keskimäärin noin 35 % käytettävyyssongelmista. Tämän vuoksi hän suosittelee, että heuristisessa arvioinnissa tulisi käyttää keskimäärin viittä, mutta vähintään kolmea eri arvioijaa. Nielsen perustelee tätä sillä, että eri arvioijat löytävät erilaisia käytettävyyssongelmia ja useampaa arvioijaa käytettäessä pystytään saavuttamaan huomattavasti parempia tuloksia.

Mikäli heuristiseen arviointiin osallistuu useampia arvioijia, tutustuvat he kukin yksitellen tuotteeseen. Kun kaikki arvioitsijat ovat arvioineet tuotetta yksin, keskustelevat he havainnoista keskenään ja tekevät löydöistään yhteenvedon. (Nielsen 1994, 26) Heuristisen arvioinnin lopputuloksena syntyy lista käytettävyyssongelmista, joita arvioitsijat ovat havainneet. Arvioijien tulisi määritellä ongelmat tarkasti ja listata niistä jokainen yksitellen, jotta lopputuloksista olisi mahdollisimman paljon hyötyä tuotekehityksessä. (Nielsen 1994, 31)

Ensimmäiset heuristiikat kehitettiin 1980-luvun lopulla, jolloin ne olivat useimmiten laajoja sääntökokoelmia, sisältäen jopa useita satoja yksityiskohtaisia ohjeita. Nämä listat todettiin kuitenkin epäkäytännöllisiksi laajuutensa vuoksi ja kevyemmät heuristiikkalistat alkoivat saavuttaa suosiota. (Kuutti 2003, 47) Myös Nielsen ja Molich (1990) huomasivat kevyempien

listojen vahvuuden ja kehittivät heuristiikkalistan, jossa he kuvailivat yhdeksän käytettävyyden peruseriaa. Kyseinen heuristiikkalista on hyvin pitkälti sama kuin tänä päivänä käytetyin niin sanottu Nielsenin Kymmenen heuristiikan lista (TAULUKKO 1.) (Kuutti 2003, 49; Sinkkonen ym. 2006, 206).

1. Järjestelmän tilan näkyvyys (engl. Visibility of system status)
2. Yhtenevyys järjestelmän ja tosielämän välillä (engl. Match between system and the real world)
3. Käyttäjän kontrolli ja vapaus (engl. User control and freedom)
4. Yhteneväisyys ja standardit (engl. Consistency and standards)
5. Virheiden estäminen (engl. Error prevention)
6. Enemmän tunnistaminen kuin muistaminen (engl. Recognition rather than recall)
7. Käytön joustavuus ja tehokkuus (engl. Flexibility and efficiency of use)
8. Esteettinen ja minimalistinen suunnittelu (engl. Aesthetic and minimalist design)
9. Auta käyttäjää tunnistamaan, määrittelemään ja toipumaan virhetilanteista (engl. Help users recognize, diagnose, and recover from errors)
10. Opastus ja dokumentaatio (engl. Help and documentation)

TAULUKKO 1. Nielsenin (1994, 30) Kymmenen heuristiikan lista (suomennettu).

Heuristiikkalistoja on kehitetty useita muitakin kuin edellä esitelty Nielsenin lista. Osa alan asiantuntijoista on täydentänyt Nielsenin listaa ja osa on luonut täysin uusia heuristiikkalistoja. Esimerkiksi Shneidermanin Kahdeksan kultaista sääntöä dialogin suunnitteluun on varsin tunnettu heuristiikkalista. Lisäksi on olemassa yleisiä käytettävyydsstandardeja ja myös useat yritykset ovat luoneet omia tyyliohjeistuksiaan, joita käytetään heuristisissa arvioinneissa. (Korvenranta 2005)

Heuristista arviointia on sen laajasta suosiosta huolimatta myös kritisoitu runsaasti. Muun muassa Nielsen (1994, 32) on todennut, ettei menetelmän avulla mahdollisesti löydetä kaikkia käytettävyysoongelmia. Arvioinnin laatu ja saavutettavat tulokset ovat myös erittäin riippuvaisia asiantuntijoiden taidoista sekä kokemuksista (Po, Howard, Vetere & Skov 2004). Kritiikkiä on esitetty myös siitä, ettei heuristinen arviointi tarjoa selkeää järjestelmällistä tapaa käytettävyysongelmien korjaamiselle. Tosin löydetyt käytettävyysongelmat on useimmiten helppo ja selkeä korjata arvioinnissa laaditun raportin ja heuristiikkalistassa esitettyjen käytettävyyssperiaatteiden avulla. (Nielsen 1994, 31)

Heuristista arviointia käytetään joissain artikkeleissa ja tutkimuksissa synonyymina asiantuntija-arvioinnille (Korvenranta 2005). Termien välillä nähdään kuitenkin tässä tutkielmassa selkeä ero asiantuntija-arvioinnin edustaessa kokonaista arviointimenetelmien joukkoa ja heuristisen arvioinnin ollessa nimitys yhdelle yksittäiselle menetelmälle.

4.1.2 Kognitiivinen läpikäynti

Kognitiivinen läpikäynti (engl. cognitive walkthrough) on toinen yleisesti tunnettu asiantuntija-arvioinnin menetelmä. Kognitiivinen läpikäynti eroaa kuitenkin huomattavasti heuristisesta arvioinnista, sillä kyseisessä menetelmässä arviointi kohdistuu erityisesti oppimisen helppouteen (Ranne 2005), kun heuristinen arviointi tarkastelee käytettävyyttä huomattavasti laajemmalla tasolla. Kognitiivinen läpikäynti perustuu oppimisen teoriaan, jossa uuden tuotteen käytön oppimisen uskotaan tapahtuvan luonnollisimmin tutkimalla ja kokeilemalla uutta tuotetta (Wharton, Rieman, Lewis & Polson 1994, 105–107).

Kognitiivisessa läpikäynnissä arvioija suorittaa sovelluksen tukemat tehtävät läpi, tarkastellen samalla kuinka hyvin tuotteen käyttöliittymä tukee käyttäjää

oppimisessa, tavoitteiden asettelussa sekä tarvittavien toimintojen löytymisessä (Riihiaho 1996). Menetelmä voidaan jakaa viiteen eri vaiheeseen, jotka ovat esiselvitysten tekeminen, arvioitsijoiden kokoaminen, kognitiivisen läpikäynnin suorittaminen, tulosten kirjaaminen sekä havaittujen ongelmakohtien korjaaminen (Wharton ym. 1994, 106).

Kognitiivisen läpikäynnin voi suorittaa joko yksittäinen henkilö tai ryhmä, joka voi koostua esimerkiksi suunnittelijoista, insinööreistä sekä eri alojen asiantuntijoista. Jotta arviointi olisi tehokasta, tulisi ainakin yhden arvioijan olla perehtynyt käytettävyyteen. (Wharton ym. 1994, 109) Riihiaho (1996) on todennut, että menetelmän tehokas käyttö vaatii jonkin verran tietämystä myös psykologian teorioista sekä sanastosta.

Menetelmän kehittäjät havaitsivat kognitiivisen läpikäynnin tehokkaaksi arviointimenetelmäksi erityisesti tilanteisiin, joissa tuotteen käytettävyyttä ei ole mahdollista arvioida valmiilla prototyypillä (Dumas & Redish 1999, 68). Menetelmä soveltuukin käytettäväksi erittäin hyvin jo kehitysprosessin alkuvaiheessa, sillä arviointi voidaan suorittaa helposti esimerkiksi pelkän järjestelmäkuvauksen tai alkeellisen prototyypin avulla (Wharton ym. 1994, 108).

Menetelmää kritisoidaan kuitenkin siitä, että se unohtaa muut käytettävyyden osatekijät keskittyessään tarkastelemaan opittavuutta (Wharton ym. 1994, 108). Menetelmä ei olekaan kovin tehokas ja luotettava käytettävyyden arviointimenetelmä, jos sitä käytetään ainoana menetelmänä, vaan se soveltuu paremminkin täydentämään muita arviointimenetelmiä.

4.1.3 Muita menetelmiä

Heuristisen arvioinnin ja kognitiivisen läpikäynnin lisäksi on kehitetty useita muitakin käytettävyyden arviointimenetelmiä, joita suunnittelijat ja

asiantuntijat voivat hyödyntää kehittäessään tuotteita käytettävämmiksi. Tällaisia menetelmiä ovat Mackin ja Nielsenin (1994, 5-6) mukaan:

- muodollinen käytettävyystarkistus (engl. formal usability inspection)
- piirreläpikäynti (engl. feature inspection)
- yhtenäisyysarviointi (engl. consistency inspection)
- standardikatselmus (engl. standards inspection)
- suositukseen perustuva arviointi (engl. guideline review).

Tutkijoiden luokitukset siitä, mihin pääryhmään menetelmät kuuluvat, ovat harvoin täysin yhtenäisiä. Esimerkiksi suositukseen perustuva arviointi sijoitetaan joissakin luokitteluissa käytettävyyden automaattisen arvioinnin menetelmiin (esim. Heimonen 2005). Toinen menetelmä, jonka luokittelusta on usein erilaisia näkemyksiä, on ryhmäläpikäynti. Nielsenin (1994) ja useat muut alan asiantuntijat sijoittavat ryhmäläpikäynnin yhdeksi asiantuntija-arvioinnin menetelmäksi, vaikka kyseisessä menetelmässä käyttäjät osallistuvat oleellisesti arviointiprosessiin. Nykyään menetelmän katsotaan kuuluvan pikemminkin empiirisiin käyttäjätesteihin (Riihiaho 2000a, 8), joiden yhteydessä se esitellään tässä tutkielmassa.

Formaali eli *muodollinen käytettävyystarkistus* on kuusivaiheinen arviointimenetelmä, joka on yhdistelmä heuristista arviointia sekä yksinkertaistettua kognitiivista läpikäyntiä (Nielsen 1995). Käytettävyyden arviointi tapahtuu useimmiten ryhmässä, jossa jokaisella osallistujalla on tarkka ennalta määritelty rooli (Mack & Nielsen 1994, 6). Menetelmässä siis yhdistetään yksilön ja ryhmän arviointi (Nielsen 1995). Sopivaksi arvioijamääräksi muodollisessa käytettävyystarkistuksessa suositellaan kolmea henkeä (Faulkner 2000, 186). Menetelmää ei kuitenkaan tule sekoittaa Mackin ja Nielsenin (1994, 2) esittämään muodolliseen arviointiin (engl. formal evaluation), joksi he kutsuvat yhtä käytettävyyden arviointimenetelmien pääryhmää (ks. luku 4. Arviointimenetelmät).

Piirreläpikäynti eli ominaisuuksien arviointi keskittyy tarkastelemaan tuotteen toimintoja. Arviointi voi koskea esimerkiksi sitä, vastaavatko toiminnot oikean loppukäyttäjän tarpeita. (Mack & Nielsen 1994, 6). Arvioinnissa käydään vaihe vaiheelta läpi tuotteen eri toiminnot sekä toimintosarjat kiinnittäen huomiota erityisesti yleisimmin käytettyihin toimintoihin ja pitkiin toimintasarjoihin sekä siihen, sisältyykö sarjoihin tavanomaisuudesta poikkeavia tai erityisosaamista vaativia toimintoja (Nielsen 1995).

Yhdenmukaisuuden tarkistamisessa eli *yhtenäisyysarvioinnissa* arvioitsijat käyvät tuotteen läpi tarkastellen, että tuote on esimerkiksi toiminnoiltaan ja terminologialtaan kauttaaltaan yhdenmukainen (Faulkner 2000, 180). Yhdenmukaisuutta arvioitaessa tulee huomioida myös muiden vastaavien tuotteiden toimintojen yhdenmukaisuus (Mack & Nielsen 1994, 5-6). Esimerkiksi käyttöliittymien pikakomentojen tulee noudattaa yleisesti käytettäviä yhdistelmiä (Faulkner 2000, 180).

Standardien tarkistus eli *standardikatselmus* on menetelmä, jossa eri standardien asiantuntijat arvioivat, kuinka hyvin tuote noudattaa standardeja (Mack & Nielsen 1994, 6). Standardit ovat usein esimerkiksi jonkin yrityksen itse luomia sääntöjä, eivätkä ne täten ole yleisesti julkaistuja tai noudatettavia suunnitteluperiaatteita (Faulkner 2000, 181). Kyseiset standardit ovatkin usein luotu lisäämän tuotteen vastaavanlaisuutta muihin jo markkinoilla oleviin tuotteisiin verrattuna, ja menetelmän tavoitteena on osaltaan tarkastella, kuinka hyvin tuote noudattaa muiden vastaavien tuotteiden standardeja (Mack & Nielsen 1994, 6).

Tarkistuslistaan eli *suositukseen perustuvassa arvioinnissa* asiantuntijat tarkastelevat kuinka hyvin tuote noudattaa yleisesti määriteltyjä suuntaviivoja. Menetelmä eroaa standardikatselmuksesta siten, että arvioinnin pohjana olevat kriteerilistat ovat yleisesti julkaistuja laajoja sääntökokoelmia käytettävyyksikriteereistä (Faulkner 2000, 179, 181). Menetelmän käyttö vaatii

kuitenkin erittäin syvää asiantuntemusta, sillä listat voivat sisältää jopa tuhat erillistä kohtaa. Tästä syystä menetelmää käytetäänkin erittäin harvoin. (Mack & Nielsen 1994, 5)

Edellä esitellyistä menetelmistä ainoastaan yhtenäisyysarviointi on täysin ryhmässä tehtävään arviointiin perustuva menetelmä, sillä useimmissa menetelmissä varsinainen arviointi suoritetaan yhden arvioijan toimesta kerrallaan. Useissa menetelmissä lopullinen arviointi kuitenkin koostuu useamman arvioijan eli arviointiryhmän yhteisistä tuloksista. (Nielsen 1995)

4.2 Empiiriset käyttäjätestit

Kokeellisissa eli *empiirisissä käyttäjätesteissä* (engl. empirical user tests) todelliset käyttäjät osallistuvat käytettävyyden arviointiin (Mack & Nielsen 1994, 2). Käyttäjätesteissä arviointi tapahtuu useimmiten havainnoimalla käyttäjää, joka suorittaa tuotteella ennalta määritellyjä testitehtäviä. Tällaisia käytettävyyden arviointimenetelmiä ovat esimerkiksi käytettävyydestaus sekä ryhmäläpikäynti. (Riihiaho 2000a, 8)

Käyttäjätestit ovat erinomaisia menetelmiä kun tuotteen käytettävyyttä halutaan arvioida todellisissa käyttötilanteissa oikeiden käyttäjien kanssa. Testitilanteet voidaan järjestää joko erillisissä laboratorioissa tai todellisissa kenttäolosuhteissa, joilla molemmilla on omat vahvuutensa ja heikkoutensa. (Riihiaho 2000a, 8-9) Laboratorioissa koetilanne on helpompi järjestää ja tallentaa. Tällöin datan kerääminen on myös vaivattomampaa. Käyttäjätestien järjestäminen laboratorio-olosuhteissa vaatii kuitenkin huomattavasti enemmän resursseja kuin kenttäolosuhteissa. (Dumas & Redish 1999, 92-93) Kentällä olosuhteet ovat puolestaan aidoimmillaan ja havainnointi voidaan suorittaa ilman käyttäjän ja havainnoijan vuorovaikutusta, mutta tällöin ulkopuolisen havainnointi voi myös häiritä käyttäjää (Riihiaho 2000a, 8-9).

Kokeiden järjestäminen käyttäjien kanssa vaatii enemmän resursseja kuin asiantuntija-arvioinnit. Arvioinnin suorittaminen oikeiden käyttäjien kanssa vie useimmiten huomattavasti enemmän aika- ja raharesursseja sekä vaatii enemmän asiantuntijuutta. (Nielsen 1993, 17) Käyttäjätestien heikkoutena on myös, että ne vaativat tavallisesti vähintään toimivan prototyypin olemassaoloa, jotta arviointi voitaisiin suorittaa. (Riihiahho 2000a, 11) Tästä syystä käyttäjätestit soveltuvat harvoin arviointimenetelmiksi tuotteen elinkaaren alkuvaiheessa. Kokeelliset käyttäjien kanssa suoritettavat arvoinnit ovat suuremmista resurssivaateista huolimatta keskeisimpiä käytettävyyden arviointimenetelmiä, sillä niiden avulla voidaan saada suoraa tietoa siitä, kuinka ihmiset todellisuudessa käyttävät tuotteita ja millaisia ongelmia he kohtaavat aidoissa käyttötilanteissa (Nielsen 1993, 165).

Käyttäjätestaukset voidaan jakaa kahteen eri tyyppiin, kehitystesteihin sekä kokoaviin arviointeihin. *Kehitystestin* (engl. formative evaluation) tavoitteena on havaita mahdolliset käytettävyysongelmat ja sitä käytetään usein apuna iteratiivisessa suunnittelussa. *Kokoavassa arvioinnissa* (engl. summative evaluation) pyritään puolestaan arvioimaan tuotteen kokonaisvaltaista laatua. Tällöin arvioidaan esimerkiksi, mikä vaihtoehto olisi toimivin ratkaisu kyseiseen tilanteeseen. (Nielsen 1993, 170)

Nielsen (1993, 165) painottaa kaikessa testauksessa reliabiliteetin ja validiteetin tärkeyttä. *Reliabiliteetilla* (engl. reliability) tarkoitetaan tutkimustulosten luotettavuutta eli testi tulee olla toistettavissa samoin tuloksin. Käyttäjätesteissä reliabiliteettia heikentää muun muassa se, että testikäyttäjien välillä on useimmiten huomattavia yksilöllisiä eroja. Tällöin koehenkilöiden valinta vaikuttaa suuresti tuloksiin. (Nielsen 1993, 165–166) *Validiteetti* (engl. validity) eli pätevyys merkitsee, että mitataan oikeaa asiaa. Käyttäjätesteissä tulee olla huolellinen, että testi mittaa käytettävyyttä juuri siten, miten tutkija on tarkoittanut. Ongelmia aiheutuu, mikäli tutkija ei huomioi esimerkiksi

käyttäjien sosiaalisia vaikutuksia tuloksiin tai testitehtävät ovat jo alun pitäen laadittu virheellisesti. (Nielsen 1993, 169)

Käyttäjätestejä järjestettäessä myös etiikan merkitys korostuu, sillä arviointiin osallistuu todellisia käyttäjiä, ihmisiä, joiden tunteet ja hyvinvointi tulee ensisijaisesti huomioida testitilanteen aikana. Kokeen järjestäjän velvollisuutena onkin varmistaa, että testikäyttäjä tuntee olonsa miellyttäväksi niin kokeen aikana kuin sen jälkeen. Testikäyttäjiä ei esimerkiksi suositella kutsuttavan koehenkilöiksi, jotta heille ei syntyisi mielikuvaa, että he ovat itse testauksen kohteena. (Nielsen 1993, 181–182) Käyttäjälle tulee korostaa, että tarkoituksena on testata tuotetta, eikä käyttäjää, ja käyttäjä voi missä vaiheessa tahansa keskeyttää testin. Testin aikana käyttäjälle ei tule viestiä mahdollisista virheistä tai hitaudesta. Tulosten säilyttäminen luottamuksellisena ja tämän kertominen käyttäjälle on myös erittäin oleellista. (Nielsen 1993, 184) Koetta suunniteltaessa ja järjestäessä tulee ehdottomasti huomioida myös testikäyttäjien kulttuurierot sekä erityisryhmien ominaisuudet (Burmeister 2000a, 2000b).

Käyttäjätesti (engl. user test), käytettävyydesti (engl. usability test) sekä käytettävyydestaus (engl. usability testing) ovat termejä, joiden käyttö on usein hyvin kirjavaa eri lähteissä. Termiä käytettävyydestaus käytetään esimerkiksi joissain lähteissä synonyyminä käyttäjätestille, jolla kuitenkin yleisesti tarkoitetaan yhtä arviointimenetelmien pääryhmää (Riihiahho 2000a, 7), kuten edellä esiteltiin. Käytettävyydestausta käytetään usein virheellisesti myös kuvaamaan mitä tahansa käytettävyyden arviointimenetelmää, vaikka käytettävyydestaus on nimenomaan tietty käytettävyyden arviointimenetelmä (Rubin 1994, 25). Termejä käytettävyydesti ja käytettävyydestaus voidaan puolestaan käyttää toistensa synonyymeinä. Tässä työssä arviointimenetelmästä kokonaisuudessaan käytetään termiä käytettävyydestaus ja varsinaista testitilannetta kutsutaan käytettävyydestiksi.

4.2.1 Käytettävyytestaus

Käytettävyytestauksen (engl. usability testing), kuten muidenkin käytettävyyden arviointimenetelmien, perimmäisenä tarkoituksena on kehittää testattavan tuotteen käytettävyyttä (Dumas & Redish 1999, 22). Käytettävyytestauksessa käytettävyyden arviointi perustuu siihen, kuinka hyvin käyttäjät pystyvät käyttämään testattavaa tuotetta käyttöympäristössään. (Hyysalo 2006, 155)

Käytettävyytestaus on erinomainen menetelmä, kun tuotteen käytettävyydestä halutaan saada tietoa varsinaisen loppukäyttäjän näkökulmasta. Se on myös oikein toteutettuna eräs tehokkaimmista keinoista parantaa tuotteen käytettävyyttä. Käytettävyytestaus onkin yleisimmin hyödynnetty käytettävyyden arviointimenetelmä (Nielsen 1995).

Käytettävyytestissä pyritään luomaan mahdollisimman aito käyttötilanne, jossa testihenkilö suorittaa ennalta suunniteltuja testitehtäviä. Käytettävyytestauksen avulla on siten mahdollista arvioida objektiivisesti tuotteen toimivuutta sekä havainnoida kontrolloiduissa olosuhteissa, kuinka käyttäjät todella toimivat käyttäessään tuotetta. (Dumas & Redish 1999)

Käytettävyytestauksen toteuttaminen voidaan Sinkkosen ym. (2006, 280–281) mukaan jakaa kolmeen päävaiheeseen, joihin sisältyy useita osa-alueita. Nämä vaiheet ovat:

- testin suunnittelu
- testin suorittaminen
- analysointi ja raportointi.

Käytettävyytestauksen toteuttaminen aloitetaan testauksen suunnittelulla ja suunnitelmasta tulisi laatia aina kirjallinen testaussuunnitelma (Nielsen 1993, 170). Huolella tehty suunnitelma on äärimmäisen tärkeä, jotta testauksesta onnistuttaisiin saamaan totuudenmukaisia sekä hyödyllisiä tuloksia. Suunnittelun aluksi on tärkeää määritellä tutkimusongelma ja muotoilla

tutkimuskysymykset, sillä testauksen tavoitteet ja päämäärät vaikuttavat huomattavasti siihen, millaiseksi testaus suunnitellaan (Nielsen 1993, 170).

Suunnittelussa tulee paneutua myös testiympäristön sekä osallistujien valintaan (Nielsen 1993, 171). Huomiota tulee kiinnittää siihen, millä perusteilla ja miten testikäyttäjät valitaan ja kuinka monta heitä tarvitaan (Sinkkonen ym. 2006, 283–284). Nielsen ja Landauer (1993) ovat todenneet, että jo 3-5 testikäyttäjää on riittävä määrä käytettävyydestestauksessa. Rubin (1994, 119) painottaa lisäksi, että testitulokset ovat paikkansapitäviä ainoastaan silloin, kun testikäyttäjät ovat tuotteen tyypillisiä loppukäyttäjää tai vähintään mahdollisimman lähellä heitä. Osallistujien valinnassa tulee huomioida myös se, ketkä testin järjestäjistä osallistuvat testitilanteeseen ja mitkä ovat heidän roolinsa (Nielsen 1993, 171).

Suunnittelussa tulee myös määrittellä, miten käytettävyyttä tullaan mittaamaan ja millaista tietoa testin aikana kerätään. Lisäksi testitehtävät tulee määrittellä yksityiskohtaisesti (Dumas & Redish 1999; Nielsen 1993).

Testitehtävien aikana voidaan hyödyntää useita käytettävyyden arviointi- ja tiedonkeruumenetelmiä, jotka tulee esitellä tutkimussuunnitelmassa (Sinkkonen ym. 2006, 285–287). Useimmin hyödynnettyjä menetelmiä ovat havainnointi sekä ääneen ajattelu. Näiden lisäksi tietoa voidaan kerätä esimerkiksi videokuvaamalla testitilanne. (Nielsen 1993) Käytettävyydestestauksessa hyödynnettäviä tiedonkeruu- sekä arviointimenetelmiä esitellään tarkemmin luvussa 4.2.2 Muita menetelmiä sekä 4.2.3 Tiedonkeruumenetelmiä.

Käytettävyydestestauksen suunnitteluvaihe päättyy pilottitestiin, jolloin testataan varsinaisen testin toimivuutta ja jonka perusteella tehdään tarpeelliset muutokset koejärjestelyihin. Pilottitestissä pyritään toimimaan samoin kuin varsinaisessa testaustilanteessa, mutta siitä saatavaa dataa ei analysoida eikä sitä myöskään sisällytetä lopullisiin tuloksiin. (Dumas & Redish 1999, 264) Pilottitestin järjestäminen on äärimmäisen tärkeää, jotta käytettävyydestistä

löydettäisiin mahdolliset virheet ja ongelmakohdat ennen oikeita testejä. Testikäyttäjien saaminen käytettävyytsteihin on yleensä työlästä ja tämän vuoksi onkin merkittävää, että yhtään testausta ei jouduta hylkäämään sen vuoksi, että koejärjestelyt tai testitehtävät olisivat puutteellisia. (Rubin 1994, 225)

Käytettävyytstestauksen toisena vaiheena on varsinaisen testin suorittaminen. Suoritusvaihe alkaa useimmiten testitilan sekä testin esittelyllä, jolloin testikäyttäjälle annetaan tarvittavat taustatiedot sekä ohjeistus kokeen kulusta. Testitilanteen tulee olla mahdollisimman luonnollinen ja testikäyttäjän tulee tuntee olonsa miellyttäväksi. (Dumas & Redish 1999, 276) Tämä on tärkeää, jotta saadaan kerättyä totuudenmukaista tietoa, sillä testitulokset voivat vääristyä esimerkiksi testihenkilön hermostuneisuuden takia (Kuutti 2003, 74).

Usein ennen varsinaisia testitehtäviä käyttäjää pyydetään täyttämään alkukysely, jolla kerätään taustatietoja testikäyttäjistä. Tyypillisiä taustatietoja ovat esimerkiksi testikäyttäjän ikä, ammatti sekä osaamistaso testin kohdealueelta. (Sinkkonen ym. 2006, 289) Alkukysely on voitu suorittaa myös ennakkoon, esimerkiksi testikäyttäjien valinnan yhteydessä tai sähköpostitse (Rubin 1994, 151).

Mikäli testikäyttäjällä ei ole kysyttävää kokeenkulusta, siirrytään testiosioon, jossa testikäyttäjä suorittaa ennalta määritellyjä tehtäviä. Jos testitehtäviä on useampia, on ne hyvä antaa testihenkilölle yksi kerrallaan. On myös suositeltavaa, että testitehtävät annetaan testihenkilölle sekä suullisesti että kirjallisena. (Sinkkonen ym. 2006, 289)

Jo suunnitteluvaiheessa tulee päättää, missä määrin kokeenjärjestäjä voi osallistua ja mahdollisesti neuvoa testikäyttäjää kokeenkulun aikana. Tämä on erityisen tärkeää, jotta eri testihenkilöiden kanssa säilytetään yhtenäinen linja ja testeistä saadaan vertailukelpoisia tuloksia esimerkiksi suoritusaikojen suhteen (Kuutti 2003,75). Varsinaisen testin aikana kokeenjärjestäjien tulisi vain

tarkkailla tilannetta ja olla mahdollisimman huomaamattomia (Rubin 1994, 231). On suositeltavaa, että kokeenpitäjä on mahdollisimman vähän vuorovaikutuksessa testihenkilön kanssa eikä hän missään nimessä saa tuoda omia mielipiteitään esille arvioitavasta tuotteesta. Kokeenvetäjän täytyy myös varoa osoittamasta testihenkilölle milloin tämä toimii tuotteen toimintojen kannalta oikein tai väärin testaustilanteessa. On kuitenkin suositeltavaa neuvoa käyttäjää tilanteissa, joissa käyttäjä ei pysty ratkaisemaan tehtävää useista yrityksistä huolimatta. Tällöin voidaan välttää koehenkilön ärsyyntyminen koetilanteeseen ja testaus saadaan vietyä onnistuneesti läpi. (Nielsen 1993, 190)

Testitehtävien jälkeen testikäyttäjää pyydetään usein täyttämään kyselylomake ja suoritetaan loppuhaastattelu. Haastattelu voi olla sekä epämuodollista keskustelua että ennalta laadittujen kysymysten läpikäyntiä. (Kuutti 2003, 75; Rubin 1994, 233) Koetilanteen päättyessä on tärkeää muistaa kiittää testihenkilöä avusta sekä varmistaa, ettei testikäyttäjälle jäänyt epäselvyyksiä koetehtäviin tai varsinaiseen koetilanteeseen liittyen (Nielsen 1993, 184–185).

Viimeisenä ja usein eniten aikaa vievänä vaiheena on kerätyn aineiston analysointi sekä tulosten muodostaminen ja raportointi. Useimmiten aineistoa on monessa eri muodossa, mikä hidastaa analysoinnin aloittamista, sillä kaikki kerätty tieto täytyy muokata helposti käsiteltävään muotoon. Esimerkiksi muistiinpanot sekä ääninauhat tulee kirjoittaa puhtaaksi ja numeerinen aineisto syöttää koneelle. (Kuutti 2003, 78–79) Kun raakamateriaali on työstetty helpommin käsiteltävään muotoon, voidaan suorittaa varsinainen tulosten analysointi. On suositeltavaa priorisoida ongelmat ja aloittaa analysointi sellaisista tehtävistä, joissa testihenkilöillä oli eniten vaikeuksia. (Rubin 1994, 274) Käytettävyydestauksien, kuten useiden muidenkin menetelmien, analysoinnissa, hyödynnetään yleisesti Nielsenin (1993, 103) esittämää viisiportaista käytettävyysohjelmien luokitteluasteikkoa:

0 = Kyseessä ei ole käytettävyysohjelma

1 = Kosmeettinen käytettävyysohjelma – korjataan, mikäli ylimääräistä aikaa

- 2 = Vähäinen käytettävyysoongelma - korjaamisella matala prioriteetti
- 3 = Merkittävä käytettävyysoongelma - korjaamisella korkea prioriteetti
- 4 = Katastrofaalinen käytettävyysoongelma - välttämätöntä korjata ennen kuin tuote julkaistaan

Analysoinnin jälkeen tuloksista kirjoitetaan raportti, jossa esitetään tutkimus sekä saadut tutkimustulokset ja toimenpidesuositukset (Rubin 1994, 289–290). Useimmiten tutkimuksen tulokset esitetään vielä henkilökohtaisesti muun muassa tuotesuunnittelusta ja -kehityksestä vastaaville henkilöille (Sinkkonen ym. 2006, 292).

Vaikka käytettävyydestä pidetäänkin yhtenä tehokkaimpana keinona löytää käytettävyysongelmat, ei sen käyttö useinkaan ole mahdollista. Yksi rajoittavimmista tekijöistä on, että käytettävyydestäuksen järjestäminen vaatii huomattavasti enemmän resursseja, kuin useat muut menetelmät. Käytettävyydestäuksen järjestäminen vaatii aina sopivien testihenkilöiden löytämistä, mikä voi usein olla aikaavievää sekä kallista. Lisäksi testäuksen järjestäminen, suorittaminen sekä tulosten analysointi on hidasta ja käytettävyydestäuksesta aiheutuu huomattavasti enemmän kustannuksia, kuin esimerkiksi useimmista asiantuntija-arvioinnin menetelmistä. (Nielsen 1993; 1995) Käytettävyydestäusta kritisoidaankin usein sen kalleuden sekä aikaavievyyden vuoksi. Käytettävyydestäusta ei kuitenkaan ikinä tulisi jättää hyödyntämättä pelkästään sen aiheuttamien kustannusten vuoksi. Kun testäuksen hyödyllisyyttä verrataan kustannuksiin, voidaan todeta, että käytettävyydestäuksella saavutettavat säästöt ovat pitkällä aikavälillä sen aiheuttamia kustannuksia suuremmat. (Dumas & Redish 1999, 17)

4.2.2 Muita menetelmiä

Käytettävyydestäuksen rinnalle on kehitetty lukuisia muitakin empiirisiä käyttäjätestimenetelmiä, joilla tuotteen käytettävyyttä voidaan arvioida ja

testata todellisten käyttäjien kanssa. Useimmat menetelmistä on muokattu alkuperäisestä käytettävyydestä (Riihiaho 2000b). Muita käyttäjätestauksen menetelmiä ovat muun muassa:

- ryhmäläpikäynti (engl. pluralistic walkthrough)
- paritestausta (engl. constructive interaction, paired-user testing)
- vapaa läpikäynti (engl. informal walkthrough, exploratory testing)
- visuaalinen läpikäynti (engl. visual walkthrough)
- tilannesidonnainen läpikäynti (engl. contextual walkthrough) (Riihiaho 2000a, 8, 2000b).

Ryhmäläpikäynti eli pluralistinen läpikäynti on yksi harvoja käyttäjätestimenetelmiä, jonka käyttö ei edellytä vähintään valmista prototyyppiä tuotteesta (Bias 1991). Menetelmässä arviointilanteeseen osallistuu todellisia käyttäjiä, tuotteen kehittäjiä sekä käytettävyyden asiantuntijoita, jotka kaikki samaistuvat käyttäjän rooliin. Järjestelmä ja sen toiminnot esitetään osallistujille kuva kuvalta samassa järjestyksessä, kuin ne näkyisivät oikeassakin tuotteessa. Kuvat nähdessään osallistujat kirjoittavat ylös, mitä toimintoja he tekisivät kunkin kuvan kohdalla. Viimeisessä vaiheessa osallistujat keskustelevat ryhmässä ratkaisusta, että käyttäjät esittävät omat mielipiteensä ensin. (Hollingsed & Novick 2007) Koska tuotetta arvioidaan paperiversion avulla, soveltuu menetelmä lähinnä opittavuuden arviointiin (Riihiaho 2000b). Ryhmäläpikäynti yhdistää asiantuntija-arviointien sekä käyttäjätestien ominaisuuksia. Kyseisen menetelmän luokittelusta ollaankin kahta eri mieltä osan tutkijoista sijoittaessa sen asiantuntija-arviointeihin (ks. esim. Nielsen 1995).

Paritestauksessa kaksi käyttäjää suorittaa yhdessä ennalta laadittuja tehtäviä keskustellen niistä samalla toisilleen. On todettu, että ääneen ajattelu tapahtuu paljon luonnollisemmin ja tuloksettaammin, kun testikäyttäjät saavat kertoa ajatuksiaan vertaiselleen. Paritestausta onkin suosituin käyttäjätestimenetelmä, jossa arviointiin osallistuu samaan aikaan useampi käyttäjä. (Haak, Jong &

Schellens 2006) Kahler, Kensing ja Muller (2000) kehottavat huomioimaan käyttäjäparien valinnoissa, että käyttäjien eroavaisuudet esimerkiksi asiantuntevuudessa tai puhetavassa voivat haitata testin onnistumista. Tämän vuoksi onkin suositeltavaa, että käyttäjät ovat tietämykseltään sekä taidoiltaan suhteellisen tasavertaisia parinsa kanssa. Paritestauksen on todettu soveltuvan niin opittavuuden, muistettavuuden kuin virhetilanteiden määrän arviointiin (Riihiaho 2000b).

Vapaa läpikäynti on käytettävyyden arviointimenetelmä, jossa käyttäjä testaa tuotetta vapaasti ilman erillistä ohjeistusta tai ennalta määriteltyjä testitehtäviä. Testi tulee kuitenkin suunnitella yhtä huolella kuin muutkin käyttäjätestimenetelmät, sillä myös vapaa läpikäynti vaatii tarkan toimintasuunnitelman ja päämäärien määrittelyn, jotta menetelmä olisi tehokas ja tuloksellinen. Menetelmän tehokkuus on riippuvainen myös testaajan tiedoista, taidoista sekä kokemuksista. (Itkonen & Rautiainen 2005) Ohjaajan on tunnettava testattava tuote tarkasti, vaikkakin hänen pääasiallinen tehtävänsä on tarkkailla käyttäjää testin aikana. Menetelmä vaatii lähes valmiin tuotteen tai toimivan prototyypin olemassaoloa, mikä rajoittaa menetelmän käyttöä. (Sinkkonen ym. 2006, 287)

Visuaalisen läpikäynnin menetelmä on kehitetty, jotta käytettävyyden arvioinnissa voitaisiin kerätä tietoa käyttäjän perspektiivistä, kuinka käyttäjä tulkitsee tuotetta ja sen toimintoja. Menetelmää käytetään useimmiten täydentämään perinteistä käytettävyydestausta sekä ääneen ajattelua. Visuaalisessa läpikäynnissä kokeenjärjestäjä pyytää käyttäjää tarkastelemaan tuotetta, kuvailemaan näkemäänsä ja pohtimaan toimintojen tarkoitusta. Tavoitteena on ymmärtää muun muassa kuinka käyttäjä tarkastelee tuotetta, mihin hän kiinnittää huomion ja mitkä termit ovat luontevia käyttäjälle (Riihiaho 2000a, 83).

Tilannesidonnainen läpikäynti on muunnelma tilannesidonnaisesta haastattelusta (engl. contextual inquiry) (Riihiahho 2000a, 84), jossa seurataan käyttäjien työtä todellisissa oloissa ja haastatellaan heitä työhön liittyen (Raven & Flanders 1996). Tilannesidonnainen läpikäynti eroaa tilannesidonnaisesta haastattelusta siten, että läpikäynnissä havainnoitsijoina toimivat käytettävyyssiantuntijat, kun haastattelussa huomioita tekevät suunnittelijat. Menetelmää hyödynnetään muun muassa tilanteissa, joissa tehtävän keskeyttäminen tai ääneen ajattelu eivät ole mahdollisia. Kerätyn aineiston analysointi tapahtuu samoin kuin perinteisessä käytettävyytestauksessa. (Riihiahho 2000a, 83–84)

Käytettävyyden arviointimenetelmiä on kehitetty myös erityisiä kohderyhmiä ajatellen. Esimerkiksi *vertaisopetus* (engl. peer tutoring) on kehitetty lasten kanssa tehtävää käytettävyyden arviointia varten. Vertaisopetuksessa lapsi tutustuu aluksi testattavaan tuotteeseen itsenäisesti, minkä jälkeen hän opettaa tuotteen käyttöä toiselle lapselle arvioijan havainnoidessa tilannetta samalla sivusta. (Höysniemi, Hämäläinen & Turkki 2003)

4.2.3 Tiedonkeruumenetelmiä

Varsinaisten arviointimenetelmien lisäksi on olemassa useita tekniikoita, erilaisia tiedonkeruumenetelmiä, joita käytettävyyden arviointimenetelmissä voidaan hyödyntää (Nielsen 1993, 207). Useimmiten käytettyjä tiedonkeruumenetelmiä ovat muun muassa:

- havainnointi (engl. observation)
- ääneen ajattelu (engl. thinking aloud)
- haastattelu (engl. interview)
- käyttäjäkysely (engl. questionnaire, survey) (Faulkner 2000, 30; Nielsen 1993).

Se, onko menetelmä itsenäinen arviointimenetelmä vai tiedonkeruumenetelmä, on usein epäselvää, eikä tähän ole esitetty tarkkaa määrittelyä. Osa alan

asiantuntijoista pitääkin joitakin edellä mainituista tekniikoista itsenäisinä arviointimenetelminä. Eräs tällaisista menetelmistä on muun muassa ääneen ajattelu (vrt. esim. Nielsen 1993, 195 ja Riihiaho 2000a, 23)

Havainnointia hyödynnetään yleisesti esimerkiksi käytettävyydestauksen yhteydessä (Dumas & Redish 1999, 30). Tällöin kokeenjärjestäjä pysyttelee testitehtävien aikana taka-alalla tarkkaillen testihenkilön vuorovaikutusta testattavaan tuotteeseen. Testaajan tehtävänä on myös kirjata havainnot sekä mahdolliset käyttäjän kommentit ylös, mikä helpottaa jatkossa aineiston käsittelyä (Dumas & Redish 1999, 24). Havainnointia voidaan käyttää myös omana menetelmänä (Nielsen 1993, 207), jolloin aineiston kerääminen perustuu ainoastaan tarkkailijan havaintoihin.

Myös *ääneen ajattelua* käytetään laajasti käytettävyystudkimuksessa. Menetelmässä testikäyttäjää pyydetään ajattelemaan ääneen samalla kun hän suorittaa tehtävää. Menetelmää voidaan käyttää niin itsenäisenä menetelmänä, jolloin käyttäjä suorittaa normaaleja työtehtäviään, kuin käyttäjätestauksien yhteydessä, jolloin tehtävät ovat ennalta laadittuja testitehtäviä. Tavoitteena on selvittää ongelmakohdat sekä syyt ongelmien taustalla. (Ramey, Boren, Cuddihy, Dumas ym. 2006)

Eri tutkimusalojen kehittämiä *haastattelumenetelmiä* hyödynnetään paljon käytettävyyden arvioinnissa. Haastattelumuodot voivat vaihdella vapaamuotoisesta keskustelusta strukturoituun haastatteluun, jossa seurataan tarkasti ennalta laadittuja kysymyksiä. Käyttäjätestimenetelmien yhteydessä on useimmiten tehokkainta käyttää niin sanottua puolistrukturoitua eli teemahaastattelua. Puolistrukturoidussa haastattelussa kokeenpitäjällä on valmiina muutamia suuntaa antavia kysymyksiä, mutta haastattelussa voidaan käsitellä myös muita kysymyksiä (Faulkner 2000, 42–43). Nielsen (1993, 209) onkin todennut, että käytettävyyttä voidaan usein tutkia ja kehittää yksinkertaisesti esittämällä käyttäjälle kysymyksiä.

Erilaiset *käyttäjäkyselyt* ovat myös erinomainen tapa kerätä käytettävyystietoa suoraan käyttäjiltä. Kyselyt voivat sisältää sekä avoimia kysymyksiä, joihin vastataan omin sanoin, että strukturoituja monivalintakysymyksiä, joissa vastaus valitaan valmiista vaihtoehdoista. (Nielsen 1993, 211-212) Kyselyt sisältävät myös usein erilaisia skaala- eli asteikkokysymyksiä, joissa vastaaja ilmaisee mielipiteensä valmiin asteikon avulla. Eräs tunnetuimmista on Likert-asteikko. (Faulkner 2000, 37-38)

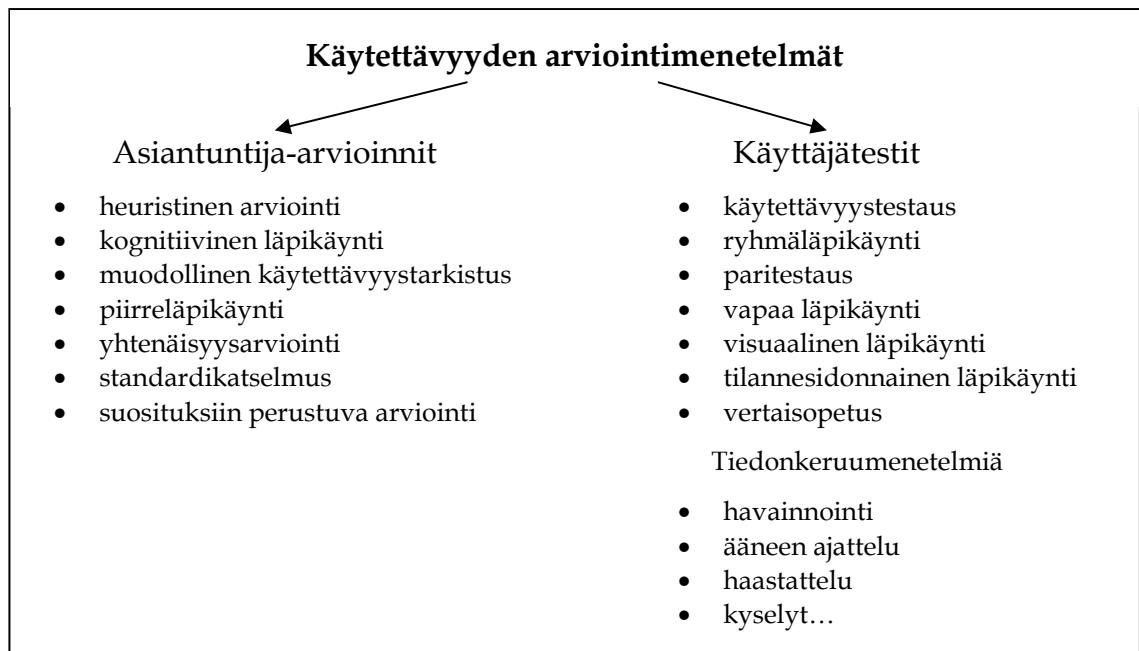
Kyselyä pidetään usein helppona ja nopeana keinona kerätä tietoa, jonka valmisteluun ei tarvitse paneutua yhtä huolella kuin esimerkiksi käytettävyydestä suunnitteluun. Onnistuneen kyselyn laatiminen vaatii kuitenkin huolellisen suunnittelun. Ennen kysymysten laatimista tulee olla selvillä, mikä kyselyn päämääränä on, ja kysymykset tulee kohdentaa siten, että päämäärä saavutetaan mahdollisimman hyvin. Kysymysten ja mahdollisten vastausvaihtoehtojen tulee olla selkeitä ja kysymysten järjestys tulee harkita tarkkaan. Kyselyitä laatiessa tulee kiinnittää huomiota myös siihen, kuinka vastaaja saadaan motivoitua vastaamaan kyselyyn. (Wilson 2007)

Tiedonkeruussa voidaan hyödyntää myös lukuisia muita menetelmiä sekä tekniikoita. Käyttäjätetit voidaan esimerkiksi videoida tai äänitallentaa tiedonkeruun ja analysoinnin helpottamiseksi. Myös pelkkä kuvaruudunkaappaus on mahdollista ohjelmistotuotteita arvioitaessa. Käytettävyyslaboratoriossa voidaan hyödyntää myös kehittyneempiä tekniikoita, jolloin testihenkilön silmänliikkeiden, pulssin tai aivojen sähköisen toiminnan seuraaminen on mahdollista. Mittausten perusteella voidaan havainnoida psykologisia reaktioita ja hyödyntää tätä tietoa testitulosten analysoinnissa. Myös automaattinen tiedon kerääminen lokiin ohjelmistotuotteiden käytettävyyttä arvioitaessa on yleistynyt. (Kuutti 2003; Nielsen 1993)

5 YHTEENVETO

Tässä tutkielmassa on tarkasteltu erilaisia käytettävyyden arviointimenetelmiä. Työssä on keskitytty esittelemään keskeisimmät arviointimenetelmät, sillä menetelmiä ja niiden muunnelmia on olemassa niin useita, ettei kaikkia menetelmiä ole mahdollista esitellä yhdessä tutkielmassa. Työssä on myös pyritty selkeyttämään suomenkielisen termistön käyttöä. Käsiteltäviin aiheisiin on perehdytty aiempien tutkimusten ja kirjallisuuden pohjalta.

Tutkielma on pohjautunut Riihiahon (2000b) esittämään käytettävyyden arviointimenetelmien ryhmittelyyn, jossa menetelmät jaetaan kahteen pääryhmään, asiantuntija-arviointeihin sekä käyttäjätesteihin. Riihiahon (2000a; 2000b) esittämää luokittelua arviointimenetelmistä on täydennetty yleisimmillä arviointimenetelmillä (KUVA 3). Tutkielmassa on käsitelty tarkemmin molempien pääryhmien suosituimmat menetelmät, heuristinen arviointi sekä käytettävyydestaus.



KUVA 3. Käytettävyyden arviointi- sekä tiedonkeruumenetelmiä Riihiahoa (2000a, 8) mukailten.

Tutkielman keskeisenä tavoitteena on ollut menetelmien vahvuuksien sekä heikkouksien selvittäminen (TAULUKKO 2). Tutkielman perusteella menetelmillä on omat selkeät vahvuutensa, mutta myös puutteensa. Tutkielmassa havaittiin myös, että samat tekijät esiintyvät usein yhtä aikaa sekä menetelmien vahvuuksina että heikkouksina.

Asiantuntija-arvioinnit	Empiiriset käyttäjätestit
<p style="text-align: center;">Vahvuudet</p> <ul style="list-style-type: none"> + Käyttäjät eivät osallistu arviointiin → menetelmät eivät vaadi suuria resursseja → taloudellisuus + Menetelmät vapaamuotoisia + Nopeita ja helppoja oppia sekä suorittaa + Eivät vaadi juurikaan etukäteisvalmistelua + Voidaan suorittaa varhaisessa vaiheessa → huomattavat säästöt + Menetelmillä mahdollista löytää useita käytettävyysongelmia + Eivät välttämättä vaadi asiantuntijuutta + Osallistujien saaminen usein helppoa <p style="text-align: center;">Heikkoudet</p> <ul style="list-style-type: none"> - Käyttäjät eivät osallistu arviointiin → tulokset eivät yhtä luotettavia kuin käyttäjätesteissä - Arvioijien kokemuksilla huomattava vaikutus tuloksiin - Löydetään usein merkityksettömiä ongelmia - Ns. väärät hälytykset yleisiä - Eivät välttämättä vaadi asiantuntijuutta → luotettavuus? 	<p style="text-align: center;">Vahvuudet</p> <ul style="list-style-type: none"> + Käyttäjät osallistuvat arviointiin → arvioinnissa mukana loppukäyttäjän näkökulma → tulokset useimmiten luotettavampia kuin asiantuntija-arvioinneissa + Käyttäjät tuovat mukanaan tietämyksen työtehtävistä sekä -tavoista + Mahdollisuus arvioida tuotetta todellisissa käyttötilanteissa → saadaan tietoa, kuinka tuotetta todellisudessa käytetään + Saavutettavat säästöt poikkeuksetta huomattavasti suuremmat kuin kustannukset <p style="text-align: center;">Heikkoudet</p> <ul style="list-style-type: none"> - Käyttäjät osallistuvat arviointiin → vaativat paljon resursseja - Hitaita ja kalliita suorittaa - Vaativat useimmiten vähintään valmiin prototyypin olemassaoloa → arviointi harvoin mahdollista suorittaa tuotteen varhaisessa kehitysvaiheessa - Vaativat järjestäjältä vahvaa asiantuntijuutta - Testihenkilöiden saaminen usein erittäin haastavaa

TAULUKKO 2. Käytettävyyden arviointimenetelmien vahvuudet ja heikkoudet pääryhmittäin.

Asiantuntija-arvioinnit ja käyttäjätestit eivät kuitenkaan koskaan ole toisiaan korvaavia, saati toistensa kanssa kilpailevia menetelmiä, sillä eri menetelmien avulla voidaan löytää hyvin erityyppisiä käytettävyysoongelmia. Ei myöskään voida todeta, että jokin menetelmä sopisi kaikkiin tilanteisiin tai olisi selkeästi muita menetelmiä parempi. Onkin suositeltavaa, että erilaisia menetelmiä käytetään rinnakkain, jolloin voidaan päästä parempiin lopputuloksiin kuin yksittäistä menetelmää käyttämällä.

Käytettävyyden arviointia ei tulisi koskaan jättää suorittamatta esimerkiksi resursseihin tai osaamattomuuteen vedoten, sillä tuotteen hyvä käytettävyys on ensisijaisen tärkeää. On myös todettu, että kun käytettävyyden arviointi suoritetaan asianmukaisesti, ovat saavutettavat säästöt lähes poikkeuksetta kustannuksia suuremmat.

Jotta käytettävyyttä voitaisiin arvioida tehokkaasti, tulee käytettävyyden peruseriaatteet olla selvillä. Arvioinnin organisoijan tulee myös tuntea eri menetelmät, sillä käytettävien arviointimenetelmien valinta voi usein olla haasteellista. Menetelmän valintaan vaikuttavat useat tekijät, mutta valinnassa tulee ensisijaisesti huomioida arvioinnille asetetut tavoitteet. Muita valintakriteereitä ovat esimerkiksi menetelmän nopeus, tuotteen kehitysvaihe ja menetelmällä todennäköisesti löydettävien käytettävyysongelmien määrä ja laatu. Erilaiset resurssivaatimukset, esimerkiksi budjetti, aika sekä osaaminen, rajoittavat usein menetelmien valintaa. Arviointimenetelmien suuri kirjo kuitenkin takaa, että jokaiseen tilanteeseen on mahdollista löytää sopiva menetelmä.

LÄHDELUETTELO

- Bias R. 1991. Interface-Walkthroughs: Efficient Collaborative Testing. *Software, IEEE* 8 (5), 94-95.
- Bolchini D. & Garzotto F. 2008. Quality and Potential for Adoption of Usability Evaluation Methods: An Empirical Study on Mile. *Journal of Web Engineering* 7(4), 299-317.
- Bødker S. & Madsen K. H. 1998. Methods & Tools: Context: An Active Choice in Usability Work. *ACM Interactions* 5(4), 17-25.
- Burmeister O. K. 2000a. Usability Testing: Revisiting Informed Consent Procedures for Testing Internet Sites. *ACM International Conference Proceeding Series; Vol. 7, Selected Papers From the Second Australian Institute Conference on Computer Ethics*. Darlinghurst, Australia: Australian Computer Society, Inc, 3-9.
- Burmeister O. K. 2000b. HCI Professionalism: Ethical Concerns in Usability Engineering. *ACM International Conference Proceeding Series; Vol. 7, Selected Papers from the Second Australian Institute Conference on Computer ethics*. Darlinghurst, Australia: Australian Computer Society, Inc, 11-17.
- Cockton G. & Woolrych A. 2002. Sale Must End: Should Discount Methods Be Cleared Off HCI's Shelves? *ACM Interactions* 9(5), 13-18.
- Dumas J. S. & Redish J.C. 1999. *A Practical Guide to Usability Testing*. Exeter: Intellect Books.
- Faulkner X. 2000. *Usability Engineering*. Chippenham, Wiltshire: Palgrave.

- Greenberg S. & Buxton B. 2008. Usability Evaluation Considered Harmful (Some of the Time). Teoksessa Czerwinski M. & Tan D. (toim.) Conference on Human Factors in Computing Systems, Proceeding of the Twenty-sixth Annual SIGCHI Conference on Human Factors in Computing Systems. Florence, Italy, April 5-10. New York: ACM, 111-120.
- Haak M. van den, Jong M. de & Schellens P.J. 2006. Constructive Interaction: An Analysis of Verbal Interaction in a Usability Setting. Professional Communication, IEEE Transactions on, 49 (4), 311-324.
- Heimonen T. 2005. Käytettävyyden automaattinen arviointi. Teoksessa Ovaska S., Aula A. & Majaranta P. (toim.) Käytettävyydestutkimuksen menetelmät. Tampereen yliopisto, Tietojenkäsittelytieteiden laitos B-2005-1, 169-186.
- Hertzum M. & Jacobsen N.E. 2003. The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods. International Journal of Human-Computer Interaction 15(1), 183 - 204.
- Hollingsed T. & Novick D.G. 2007. Usability Inspection Methods After 15 Years of Research and Practice. SIGDOC '07: Proceedings of the 25th Annual ACM International Conference on Design of Communication. El Paso, Texas, USA, October 22-24, 2007. New York: ACM, 249-255.
- Hyysalo S. 2006. Käyttäjätieto ja käyttäjätutkimuksen menetelmät. Helsinki: Edita Prima Oy.
- Höysniemi J., Hämäläinen P. & Turkki L. 2003. Using Peer Tutoring in Evaluating the Usability of a Physically Interactive Computer Game with Children. Interacting with Computers 15 (2), 203-225.
- ISO 1998. ISO 9241-11:1998 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability. International Organization for Standardization.

- Itkonen J. & Rautiainen K. 2005. Exploratory Testing: A Multiple Case Study. ISESE 2005: Proceedings of the International Symposium on Empirical Software Engineering, November 17-18. IEEE Computer Society, 84-93.
- Kahler H., Kensing F. & Muller M. 2000. Methods & Tools: Constructive Interaction and Collaborative Work: Introducing a Method for Testing Collaborative Systems. *ACM Interactions* 7 (3), 27-34.
- Kalimo A. 1996. Johdanto: Käytettävyyden merkitys. Teoksessa Kalimo A. (toim.) Graafisen käyttöliittymän suunnittelu - Opas ohjelmistojen käytettävyyteen. Jyväskylä: Tietotekniikan kehittämiskeskus TIEKE ry, 7-11.
- Korvenranta H. 2005. Asiantuntija-arvioinnit. Teoksessa Ovaska S., Aula A. & Majaranta P. (toim.) Käytettävyydestutkimuksen menetelmät. Tampereen yliopisto, Tietojenkäsittelytieteiden laitos B-2005-1, 111-124.
- Kuutti W. 2003. Käytettävyys, suunnittelu ja arviointi. Saarijärvi: Talentum Media Oy.
- Mack R.L. & Nielsen J. 1994. Executive Summary. Teoksessa Nielsen J. & Mack R.L. (toim.) Usability Inspection Methods. New York: John Wiley & Sons, Inc, 1-23.
- McNamara N. & Kirakowski J. 2005. Defining Usability: Quality of Use or Quality of Experience? Teoksessa Professional Communication Conference, 2005. IPCC 2005. Proceedings. International. Limerick, Ireland, July 10-13. IEEE, 200-204.
- Molich R. & Jeffries R. 2003. Comparative Expert Reviews. Teoksessa Cockton G. & Korhonen P. (toim.), CHI '03 Extended Abstracts on Human Factors in Computing Systems. Ft. Lauderdale, Florida, April 5-10. New York: ACM, 1060-1061.

- Myers B.A. 1998. A Brief History of Human-Computer Interaction Technology. New York: ACM Interactions 5(2), 44-54.
- Nielsen J. 1993. Usability Engineering. New York: Academic press.
- Nielsen J. 1994. Heuristic Evaluation. Teoksessa Nielsen J. & Mack R.L. (toim.) Usability Inspection Methods. New York: John Wiley & Sons, Inc, 25-62.
- Nielsen J. 1995. Usability Inspection Methods. Teoksessa Katz I., Mack R. & Marks L. (toim.) CHI '95 Conference Companion on Human Factors in Computing Systems. Denver, Colorado, May 07-11. New York: ACM Press, 377-378.
- Nielsen J. & Landauer T.K 1993. A Mathematical Model of the Finding of Usability Problems. Conference on Human Factors in Computing Systems, Proceedings of ACM INTERCHI'93 Conference. Amsterdam, The Netherlands, 24-29 April, 206-213.
- Nielsen J. & Molich R. 1990. Heuristic evaluation of user interfaces. Teoksessa Chew J. C. & Whiteside J. (toim.) Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people. Seattle, Washington, April 01-05. New York: ACM Press, 249-256.
- Paananen J. 2001. Tietotekniikan peruskirja. Porvoo: Docendo Finland Oy.
- Po S., Howard S., Vetere F. & Skov M.B. 2004. Heuristic Evaluation and Mobile Usability: Bridging the Realism Gap. Teoksessa Brewster S. & Dunlop M. (toim.) Mobile Human-Computer Interaction - Mobile HCI 2004, Lecture Notes in Computer Science, Vol. 3160, 49-60.
- Preece J., Rogers Y., Sharp H., Benyon D., Holland S. & Carey T. 1994. Human-Computer Interaction. Harlow: Addison-Wesley.

- Ramey J., Boren T., Cuddihy E., Dumas J., Guan Z., Haak M.J. van den & De Jong M.D.T. 2006. Does Think Aloud Work?: How Do We Know?. Teoksessa Olson G.M. & Jeffries R. (toim.), CHI '06 Extended Abstracts on Human Factors in Computing Systems. Montréal, Québec, April 22 - 27. New York: ACM, 45-48.
- Ranne S. 2005. Kognitiivinen läpikäynti. Teoksessa Ovaska S., Aula A. & Majaranta P. (toim.) Käytettävyystutkimuksen menetelmät. Tampereen yliopisto, Tietojenkäsittelytieteiden laitos B-2005-1, 125-140.
- Raven M.E. & Flanders A. 1996. Using Contextual Inquiry to Learn About Your Audiences. ACM SIGDOC Asterisk Journal of Computer Documentation 20 (1), 1-13.
- Riihiaho S. 1996. Asiantuntijoiden suorittama arviointi. Teoksessa Kalimo A. (toim.) Graafisen käyttöliittymän suunnittelu - Opas ohjelmistojen käytettävyyteen. Jyväskylä: Tietotekniikan kehittämiskeskus TIEKE ry, 108-113.
- Riihiaho S. 1998. Käytettävyyden arviointi ilman käyttäjiä. *Systemityö* 5(4), 4-6. Systemityöyhdistys SYTYKE ry.
- Riihiaho S. 2000a. Experiences with Usability Evaluation Methods. Teknillinen korkeakoulu, tietotekniikan lisensiaattityö.
- Riihiaho S. 2000b. Käytettävyystestauksen muunnelmia. Teoksessa Pantzar E. (toim.) Informaatio, tieto ja tietoyhteiskunta. Suomen Akatemian Tiedon tutkimusohjelman raportteja, 4/2000. Tampere: Tampereen yliopisto, 223-230.
- Rubin J. 1994. Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. New York: John Wiley & Sons, Inc.

- Sinkkonen I. 2004. Käyttöliittymät ja käytettävyys [online]. Adage Oy [viitattu 28.10.2008]. Saatavilla [www-muodossa <http://www.adage.fi/artikkelit/kayttoliittymat_kaytettavyys.html >](http://www.adage.fi/artikkelit/kayttoliittymat_kaytettavyys.html).
- Sinkkonen I., Kuoppala H., Parkkinen J. & Vastamäki R. 2006. Käytettävyyden psykologia. Helsinki: Edita Prima Oy.
- Väyrynen S., Nevala N. & Päivinen M. 2004. Ergonomia ja käytettävyys suunnittelussa. Teknologiateollisuuden julkaisuja 4/2004. Tampere: Teknologiateollisuus ry.
- Wharton C., Rieman J., Lewis C. & Polson P. 1994. The Cognitive Walkthrough Method: A Practitioner's Guide. Teoksessa Nielsen J. & Mack R.L. (toim.) Usability Inspection Methods. New York: John Wiley & Sons, Inc, 105-140.
- Wilson C.E. 2007. Designing Useful and Usable Questionnaires: You Can't Just "Throw a Questionnaire Together". ACM Interactions 14 (3), 48-49, 63.