

AALTO-YLIOPISTO

TEKNILLINEN KORKEAKOULU

Informaatio- ja luonnontieteiden tiedekunta

Informaatioverkostojen tutkinto-ohjelma

KÄYTETTÄVYYSKYSELYT KÄYTTÄJÄKOKEMUKSEN TUTKIMISESSA

Marianne Tenhula

Kandidaatintyö, joka on jätetty opinnäytteenä tarkastettavaksi tekniikan kandidaatin tutkintoa varten

Espoo 28.11.2010

Ohjaaja: Johanna Viitanen

AALTO-YLIOPISTO TEKNILLINEN KORKEAKOULU PL 11000, 00076 Aalto http://www.aalto.fi	KANDIDAATINTYÖN TIIVISTELMÄ	
Työn tekijä: Marianne Tenhula		
Työn nimi: Käytettävyysskyselyt käyttäjäkokemuksen tutkimisessa		
Tiedekunta: Informaatio- ja luonnontieteiden tiedekunta		
Tutkinto-ohjelma: Informaatioverkostojen tutkinto-ohjelma		
Pääaine: Vuorovaikutus ja viestintä	Pääaineen koodi: T120-2	
Vastuuopettaja: Stina Immonen		
Ohjaaja: Johanna Viitanen		
<p>Sekä akatemian että teollisuuden aloilla on viime vuosina ymmärretty käyttäjäkokemuksen eli elämyksellisyyden, mielihyvän ja tyytyväisyyden merkitys tuotteiden ja järjestelmien käytössä. Tämä tutkielma käsittelee käyttäjäkokemusta käytettävyysskyselyiden näkökulmasta, ja tutkii, voidaanko käytettävyyden mittaamiseen suunniteltuja standardikyselyitä hyödyntää käyttäjäkokemuksen mittaamisessa. Työssä esitellään myös muutamia käyttäjäkokemuksen mittaamista varten suunniteltuja kyselyitä ja pohditaan kyselyiden haasteita käyttäjäkokemuksen tutkimisessa.</p> <p>Käytettävyysskyselyt ovat perinteinen osa käytettävyyssuunnittelua. Ne ovat verrattain nopeita hallinnoida, täyttää sekä analysoida. Standardoidut käytettävyysskyselyt ovat päteviä ja luotettavia arvioimaan tuotteen tai järjestelmän yleistä käytettävyyttä, ja erilaisia standardoituja kyselyitä on useita. Kyselyt ovat tehokkaita tapoja kerätä subjektiivista ja määrällistä tietoa käyttäjiltä, mutta yleisyytensä takia ne eivät useinkaan sovi ainoaksi käytettävyyden tutkimusmenetelmäksi.</p> <p>Käyttäjäkokemuksen arviointimenetelmiä on kehitetty viime vuosina reilusti, ja käytettävyysskyselyt ovat yksi hyvä keino kerätä subjektiivista tietoa käyttäjän tuntemuksista ja käyttökokemuksista. Käyttäjäkokemuksen ja sen tutkimuksen ollessa lähestymistapana melko tuore eivät kyselyt kuitenkaan vielä ole riittävän luotettavia käyttäjäkokemuksen mittareita.</p> <p>Olenneisinta kyselyiden luotettavuuden ja pätevyyden kannalta on, että kysely vastaa siihen, mitä mitataan, ja sen tulosten analysointia varten määritetään selkeät käyttäjäkokemuksen mittarit - mitkä ovat tuotteen tai järjestelmän hyvän käyttäjäkokemuksen kriteerit. Tarvitaan vielä rutkasti lisää tutkimusmateriaalia ja -tuloksia, jotta voidaan määritellä riittävän tarkkoja mittaristoja ja kriteereitä hyvälle käyttäjäkokemukselle. Näin kyselyistäkin saadaan tulevaisuudessa enemmän irti, kun voidaan tarkemmin määritellä, mitä mitataan.</p>		
Päiväys: 28.11.2010	Julkaisukieli: suomi	Sivumäärä: 33
Avainsanat: Käyttäjäkokemus, käytettävyysskysely, käyttäjäkokemuskysely, ISO 9421-210		

Sisällysluettelo

1. Johdanto.....	1
2. Näkökulmia käyttäjäkokemukseen.....	3
2.1 Käyttäjäkokemuksen määritelmä ISO 9241-210 mukaan.....	3
2.2 Lähtökohtia käyttäjäkokemuksen tutkimiseen	4
2.3 Käyttäjäkokemustutkimuksen erityispiirteitä.....	5
3. Käytettävyyskyselyt	7
3.1 Erilaisia kyselytyyppejä	7
3.1.1 Mobiilikyselyt	8
3.2 Standardikyselyt	9
3.3 Muokatut kyselyt.....	15
3.4 Kyselyiden käytön kulmakiviä.....	16
4. Käyttäjäkokemus ja käytettävyyskyselyt	18
4.1 Standardikyselyt ja käyttäjäkokemus	18
4.2 Käyttäjäkokemuskyselyt	20
4.2.1 AttrakDiff	20
4.2.2 Läsnaoloa tutkiva kysely	21
4.2.3 Lasten pelikokemusta tutkiva kysely	22
4.3 Kyselymenetelmien käytön haasteita	23
5. Johtopäätökset	26
Lähteet	29
Kirjallisuuslähteet.....	29
WWW-lähteet.....	32

1. Johdanto

Ajatellaan hetki matkapuhelinta. Matkapuhelin voidaan nähdä pieninä osina liukuhihnalta toiselle ja lopulta yksittäisenä tuotteena kaupan hyllyllä hukkumassa koreampiensa keskelle, jolloin se on vain tuote muiden joukossa. Varsinaisen kokemuksen puhelimesta saa vasta käyttötilanteessa ja siinä hetkessä, kun soittaa puhelun läheiselleen, lähettää tekstiviestin äidilleen, lukee työsähköpostit junamatkalla tai rentoutuu uuvuttavan päivän jälkeen matopelin parissa. Tuotteena matkapuhelin voi siis olla mitäänsanomaton ja yhdentekevä, vasta kokemus ja ympäristö antavat sille syvemmän merkityksen ja määrittävät mielipiteen ja asenteen matkapuhelinta kohtaan. Puhutaan matkapuhelimen käyttäjäkokemuksesta.

Esimerkin kaltainen käyttäjäkokemus on vuosikymmenen aikana noussut käytettävyytutkijoiden mielenkiinnon kohteeksi. Se korostaa, kuinka tärkeää on kerätä tietoa käyttäjän ajatuksista, mielipiteistä sekä tuntemuksista, ja tehdä tuotteita ja järjestelmiä ennen kaikkea käyttäjän tarpeisiin. Tämä tutkielma käsittelee käyttäjäkokemusta käytettävyysskyselyiden näkökulmasta, ja tutkii, voidaanko käytettävyyden mittaamiseen suunniteltuja standardikyselyitä hyödyntää käyttäjäkokemuksen mittaamisessa.

Tässä työssä käytettävyys määritellään ISO 9241-standardin mukaan tuloksellisuuden, tehokkuuden ja tyytyväisyyden kautta. Käytettävyys on alue jossa tuotteen käyttäjät saavuttavat tiettyjä tavoitteita tehokkaalla, tuloksellisella ja tyydyttävällä tavalla tietyssä käyttöympäristössä tai -kontekstissa.

Tutkielman toisessa luvussa pureudutaan lyhyesti käyttäjäkokemuksen käsitteeseen ja siihen, miksi käyttäjäkokemus ylipäätään on kiinnostava tutkimuksen kohde. Käyttäjäkokemuksen nousemista käytettävyysalan keskiöksi kuvaavat hyvin tutkijoiden ristiriitaiset näkemykset sen määrittelemisestä. Yhteistä eri määritelmille ovat käyttäjän tunteiden ja elämyksellisyyden huomioiminen, ja määrittelemisen tueksi on julkaistu ISO-standardi 9241-210, joka esitellään lyhyesti luvussa 2.

Kyselylomakkeet ovat yleinen tutkimuksissa käytetty tiedonkeruumenetelmä, ja käytettävyyskyselyissä niitä voidaan käyttää joko ainoana menetelmänä tai osana muita

menetelmiä. (Nielsen, 1993) Erilaisia standardikyselyitä on tehty helpottamaan käytettävyyden arviointia eri vaiheissa tuotekehitystä. Ne mittaavat tuotteiden ja järjestelmien yleisiä ominaisuuksia, ja niitä on kehitetty iteratiivisesti testaamalla ja muokkaamalla lomakkeita useaan kertaan. Näin voidaan varmistaa tulosten toistettavuus sekä luotettavuus, ja niitä voidaan vertailla muihin järjestelmiin. Kolmannessa luvussa esitellään yleisimpiä käytettävyysskyselyitä, minkälaisia asioita ne painottavat ja arvioidaan, kuinka hyvin ne sopivat käytettävyyden tutkimiseen, ja minkälaisia kokemuksellisia näkökulmia ne painottavat.

Nykypäivän tuotekehitysprojekteissa on tehokkuuden ja tuloksellisuuden lisäksi entistä enemmän tutkittu, miltä tietyn tuotteen käyttäminen tuntuu, tai mitä käyttäjä tuotteelta ja vuorovaikutukselta tuotteen kanssa toivoo. Neljännessä luvussa arvioidaan standardikyselyiden soveltumista käyttäjäkokemuksen tutkimiseen. Käyttäjäkokemuksen eri näkökulmia mittaamaan on kehitetty erilaisia kyselyitä, joista nostetaan esille muutama hieman eri ulottuvuuksia painottava menetelmä. Luvussa myös arvioidaan kyselyiden käyttöä käyttäjäkokemuksen tutkimisessa ylipäänsä. Viidennessä luvussa kootaan yhteen työn lopputuloksia ja pohdintoja tulevaisuuden käyttäjäkokemustutkimusta ajatellen.

Kandidaatintyön tarkoitus on luoda katsaus käytettävyyttä mittaaviin standardikyselyihin sekä niiden hyödyntämiseen käyttäjäkokemuksen tutkimisessa. Lisäksi työssä esitellään joitakin juuri käyttäjäkokemuksen mittaamista varten suunniteltuja kyselyitä ja pohditaan kyselyiden haasteita käyttäjäkokemuksen tutkimisessa. Työssä pyritään vastaamaan kysymyksiin

- Minkälaisia standardimallisia käytettävyysskyselyitä on kehitelty? Minkälaisia seikkoja ne painottavat? Voidaanko niitä hyödyntää käyttäjäkokemuksen mittaamisessa?
- Minkälaisia käyttäjäkokemuskyselyitä on kehitetty, ja mitä asioita ne painottavat?
- Mitä tulee ottaa huomioon käytettäessä kyselyitä käyttäjäkokemuksen arvioimiseen?

2. Näkökulmia käyttäjäkokemukseen

Vaikka käyttäjäkokemus on käsitteenä ja ilmiönä noussut tutkijoiden mielenkiinnon kohteeksi vasta viime aikoina, on käyttäjäkokemuksen määrittämisessä ollut jo 1980-luvun lopulla parikin edelläkävijää, joihin Hassenzahl ja Tractinsky viittaavat raportissaan (2006). Whiteside ja Wixon uskoivat jo vuonna 1987, että tärkeimpiä käytettävyyden mittareita eivät ole opittavuus tai tuottavuus, vaan käyttäjän järjestelmän tai tuotteen käyttöhetken esiin nostattamat tunteet. Vuotta myöhemmin Carrol ja Thomas haastoivat tutkijat ajattelemaan hauskuuden merkitystä järjestelmien suunnittelussa. Voi kuitenkin vuosikymmenen, ennen kuin tämä käsitys alettiin sisäistää ja ymmärtää, ja käyttäjäkokemus saada jalansijaa käytettävyyden rinnalla. Tästä syystä käyttäjäkokemusta koskeva tutkimusaineisto painottuu vielä paljolti käsitteen määrittämiseen ja ilmiön ymmärtämiseen. Luvussa 2.1 käsitellään käyttäjäkokemuksen määritelmää uusimman ISO 9241-210 -standardin kautta.

Kujalan (2010) mukaan käyttäjäkokemuksen tavoitteena on tuoda käytettävyyden rinnalle käsite, joka laajentaa käytettävyydelle ominaista tehokkuutta painottavaa lähestymistapaa ja tuo esille elämyksellisyyttä ja ihmisen tuntemusten painotusta. Kiinnostus käyttäjäkokemuksen tutkimiseen vaihtelee alasta ja tutkimuksen tavoitteista riippuen. Näkökulmia ja lähtökohtia tutkia käyttäjäkokemusta esitellään luvussa 2.2. Luvussa 2.3 käydään lyhyesti läpi käyttäjäkokemustutkimuksen erityispiirteitä.

2.1 Käyttäjäkokemuksen määritelmä ISO 9241-210 mukaan

Käyttäjäkokemus painottaa aikaisempaa HCI-tutkimusta enemmän käyttäjän positiivisia tuntemuksia tuotteen tai järjestelmän käytöstä (Hassenzahl & Trackinzy, 2006). Siinä missä perinteinen käytettävyydetutkimus pyrkii poistamaan tuotteista ja järjestelmistä virheet ja negatiivisia, turhautumista ja ärtymystä aiheuttavat piirteet, käyttäjäkokemus on erityisen kiinnostunut niistä seikoista, jotka järjestelmän tai tuotteen käytössä aiheuttavat käyttäjälle onnistumisen kokemuksia ja mielihyvää. Käyttäjäkokemuksen kokonaisvaltaisuutta kutsutaan myös elämyksellisyydeksi.

Tutkijat ja alan ammattilaiset ovat edelleen hieman eri mieltä käyttäjäkokemuksen määrittelemisestä. Pätevien mittareiden puuttuminen on yksi suurimmista käyttäjäkokemuksen tutkimista koskevista ongelmista. ISO-standardin sisällyttämä

määritelmä käyttäjäkokemuksesta antaa hieman paremman kehysten ja lähtökohdan tutkimukselle, vaikka määritelmää kritisoidaan edelleen sen laajuudesta (Jokela, 2010).

Uusimman ISO-standardin 9241-210 (ISO/IEC 2010) mukaan käyttäjäkokemuksella tarkoitetaan henkilön havaintoja ja reaktioita, jotka aiheutuvat tuotteen, järjestelmän tai palvelun käytöstä tai odotetusta käytöstä. Se käsittää käytön aikana sekä sitä ennen ja jälkeen ilmenevät käyttäjän tunteet, uskomukset, mieltymykset, havainnot, fyysiset ja psyykkiset reaktiot, käytöksen sekä saavutukset. Käyttäjäkokemus muodostuu useiden erilaisten tekijöiden, kuten järjestelmän tuotekuvan, esillepanon, toiminnallisuuden, järjestelmän suorituskyvyn sekä kapasiteetin seurauksena. Siihen vaikuttavat myös käyttäjän sisäinen ja fyysinen tila, aiemmat kokemukset, asenteet, taidot, persoonallisuus sekä käyttökonteksti. Mikäli käytettävyyttä tarkastellaan käyttäjän henkilökohtaisista tavoitteista käsin, voi siihen sisältyä havaintoihin ja tunteisiin liittyviä näkökulmia, jotka liitetään käyttäjäkokemukseen. Tällöin samoja kriteereitä, joita käytettävyyden tutkimuksessa hyödynnetään, voidaan käyttää käyttäjäkokemuksen arvioimiseen.

2.2 Lähtökohtia käyttäjäkokemuksen tutkimiseen

Käyttäjäkokemuksen tutkimiseen liittyy erilaisia näkökulmia ja motiiveita riippuen, tutkitaanko sitä teollisuudessa ja kaupallista tarkoitusta varten vai akateemisella, tieteellisellä taholla. Akateeminen näkemys on kiinnostunut ymmärtämään käyttäjäkokemuksen mittaamista ilmiönä (Roto et al. 2010a), ja tutkijat pyrkivät löytämään arviointimenetelmiä, jotka tuottavat luotettavaa, pätevää ja kattavaa tietoa käyttäjäkokemuksesta pohjautuen teoreettiselle tutkimukselle. Teollisuudessa tavoitteena taas on parantaa tuotteen käyttäjäkokemusta, jotta se tuottaisi mahdollisimman paljon voittoa. (Väätäjä & Roto, 2009) Tästä syystä teollisuuden nopeatempoisissa tuotekehitysprosessissa on erityisen tärkeää löytää kustannustehokkaita arviointimenetelmiä, jotka ovat myös nopeita toteuttaa, ja joiden tulokset ovat analysoitavissa mahdollisimman pian arvioinnin tekemisestä. Kyselyt ovat tästä syystä erittäin suosittuja osana tuotekehitysprosessia.

Teollisuudessa tehdään useita pieniä arviointeja siinä missä akateemisella puolella keskitytään tekemään suurempia ja hyvin suunniteltuja tutkimuksia mahdollisimman kattavan tiedon keräämiseksi. Akateemiset tutkimukset käsittävät kontrolloituja

muuttujia ja tutkimusympäristöjä, jossa projektin toteuttajat ovat pääosassa käytettävyyssiantuntijoita, kokeneita tutkijoita. Teollisuuden arviointitapaukset liittyvät todellisen elämän tilanteisiin ja tuotesuunnittelu tai -kehitysryhmässä on usein vain 1-2 käytettävyyssiantuntijaa. Usein yritykset myös ulkoistavat käytettävyysoasaamisen ulkopuolisille tahoille. (Roto et al. 2010)

2.3 Käyttäjäkokeumustutkimuksen erityispiirteitä

Käyttäjäkokeuksen tutkiminen on käytettävyyden arviointia pidempiaikaisempaa: käytettävyyden tutkiminen keskittyy enemmän toiminnan aikana tai välittömästi sen jälkeen tehtävään arviointiin. Käyttäjäkokeuksen kannalta on mielenkiintoista, miltä esimerkiksi vuoden käytössä olleen kännykän käyttö on tuntunut. Hyvän käyttäjäkokeuksen on myös nähty olevan suoraan verrannollinen käyttäjäuskollisuuteen, ja juuri tämän takia pitkäaikaisen käyttäjäkokeuksen tutkiminen on erityisen kiinnostavaa etenkin teollisuuden näkökulmasta (Roto et al. 2010a). Mikä saa ihmisen pysymään esimerkiksi tietyssä pankissa tai ostamaan tiettyä pesuainetta vuosi toisensa jälkeen, vaikka välillä kokeekin pettymyksen hetkiä tuotetta tai järjestelmää käyttäessään?

Olenneinta käyttäjäkokeuksen tutkimisessa on selvittää, miltä järjestelmän tai tuotteen käyttäminen henkilöstä tuntuu. Alana käyttäjäkokeuksen tutkiminen on hyvin alkutekijöissään, ja tehokkaita tutkimusmenetelmiä kehitetään jatkuvasti. Ei ole olemassa yhtä yleispätevää menetelmää, vaan jokaisen tutkimuksen kohdalla on yksilöllisesti mietittävä, minkä menetelmän avulla päästään tutkimustavoitteisiin parhaiten. (Kujala, 2010) Menetelmät ovat pitkälti samanlaisia kuin käytettävyydenkin tutkimuksessa, mutta tutkimuksen pääpaino keskittyy tunteisiin ja kokemukseen järjestelmän toiminnallisuuden ja tehokkuuden sijasta. Tässä tutkielmassa käsitellään kyselymenetelmiä käyttäjäkokeuksen tutkimisessa. Muita menetelmiä ovat esimerkiksi haastattelut, laboratoriokokeet, kenttätutkimukset, asiantuntija-analyysit ja käyttäjän havainnointi.

Hassenzahl ja Tractinsky (2006) määrittelevät onnistuneen tutkimuksen lähtökohdaksi aineettomien tarpeiden ymmärtämisen ja määrittämisen. On selvitettävä, miten nämä korreloivat tuotteen laatuun, ja mitkä tuotteen ominaisuudet vastaavat kutakin tarvetta. Sama asia tuli myös ilmi syksyllä 2010 järjestetyssä SIGHCI-seminaarissa, jossa

painotettiin mitattavan asian määrittämisen tärkeyttä ennen tutkimusten käynnistämistä. Miten voidaan arvioida käyttäjäkokemuksen hyvyyttä ja laatua, mikä on hyvän käyttäjäkokemuksen mittari? Tutkimusta suunniteltaessa tulisi siis asettaa tietyt tavoitteet ja rajat, joiden toteuduttua käyttäjäkokemusta voidaan pitää hyvänä. Tämä on paljon helpommin sanottu kuin tehty, sillä subjektiivisten asioiden ja niinkin yksilöllisten ja intiimien kuin tunteiden arvioiminen ja skaalaaminen on hyvin haasteellista.

Roto & Rautava (2008) ovat tutkineet käyttäjäkokemuksen elementtejä ja yrityksen arvolutausta, ja esittävät raportissaan neljä käyttäjäkokemuksen mittariehdotusta. Näitä ovat hyödyllisyys, käytettävyys, sosiaalinen arvo sekä nautinnollisuus. Heidän tavoitteenaan on ollut määritellä yleisiä käyttäjäkokemuksen mittareita hyvän käyttäjäkokemuksen suunnittelua varten, erityisesti Nokian matkapuhelimia silmällä pitäen. Tutkimuksen mukaan näitä neljää elementtiä ja niiden mittareita arvioimalla, tutkimalla ja kehittämällä päästään lähelle hyvää käyttäjäkokemussuunnittelua. Suunniteltavasta tuotteesta tai järjestelmästä riippuu aina yksilöllisesti, mitä elementeistä halutaan painottaa erityisesti. Esimerkiksi pelaamiseen tarkoitettujen käyttöliittymien ei tarvitse välttämättä olla kaikista helppokäyttöisimpiä, tärkeämpää on mielihyvän ja mukaansatempaavuuden painottaminen.

3. Käytettävyyskyselyt

Sekä Nielsenin (1993) että Kirakowskin (2000) mukaan kyselyt ovat erittäin hyvä keino mitata käyttäjän subjektiivisia näkemyksiä ja kokemuksia tuotteesta tai järjestelmästä. Kyselyiden avulla saadaan selville objektiivisesti vaikeasti mitattavia asioita kuten käyttäjän tyytyväisyyttä ja mahdollista turhautumista laitteeseen. Kyselyä voidaan pitää luotettavana, mikäli se toistettaessa samanlaisessa ympäristössä samankaltaisten käyttäjien testaamana antaa samoja tuloksia.

Luvussa 3.1 käsitellään erilaisia kyselytyyppejä, ja luku 3.1.1 nostaa esiin nykytutkimuksen kannalta erityisen kiinnostavan kyselytyypin - mobiilikyselyn. UsabilityNet-sivustolle (2006) on kerätty yleisimpiä standardoituja kyselymenetelmiä, jotka ovat kehittyneet arviointi- ja muokkauskierteen seurauksena. Luvussa 3.2 käsitellään näistä kyselyistä käytettävyysalan tutkimuksessa paljon huomiota saaneita SUMI:a, SUS:ia, USE:a, CSUQ:ta sekä WAMMI:a. Luku 3.2 käsittelee lyhyesti muokatun tai kokonaan itse tehdyn kyselyn kannalta olennaisia asioita ja luku 3.3 nostaa esiin kyselyiden käytön soveltuvuutta käytettävyyden tutkimiseen.

Edellisessä luvussa avattiin käyttäjäkokemuksen määritelmää ja tutkimuksen syitä, ja tässä luvussa käytettävyyskyselyitä pyritään analysoimaan erityisesti käyttäjäkokemuksen näkökulmasta, ja löytämään kyselyistä kokemuksellisuutta sekä käyttäjän tunteita mittaavia elementtejä.

3.1 Erilaisia kyselytyyppejä

Peruskyselytyyppejä on kolmenlaisia. Faktatietoa keräävät kyselyt keräävät perustietoa käyttäjistä, kuten ikä, sukupuoli, koulutus jne. Kyselyiden alussa on usein faktatietoa keräävä osuus, jolla kartoitetaan perustietoja osallistujista. Mieliopidekyselyt keskittyvät käyttäjien mielipiteiden keräämiseen, eikä kysymyksiin ole oikeaa tai väärää vastausta. Asennetta tutkivilla kyselyillä pyritään selvittämään vastaajan asenteita ja sisäistä palautetta eri elämän tapahtumiin ja tilanteisiin. Esimerkiksi sosiaalista mediaa tutkivia asennekyselyitä on jo tehty monia. Tällaisia kyselyitä voidaan kutsua myös tyytyväisyyskyselyiksi. (Kirakowski, 2000)

Kyselyillä voidaan kerätä niin laadullista kuin määrällistäkin tietoa riippuen kyselyn tyypistä. Kyselyillä on mahdollista kerätä kokoavaa tietoa, jonka avulla tuotetta tai järjestelmää voidaan vertailla, sekä muodollista tietoa yksityiskohtaisemman

informaation saamiseksi tuotteen tai järjestelmän ominaisuuksista. Kyselyt voivat olla paperimuotoisia tai verkossa täytettäviä, ja niissä voi olla sekä avoimia että suljettuja kysymyksiä, tai ne voivat noudattaa tarkistuslistan mallia. (Väätäjä & Roto, 2009)

Avoimiin kysymyksiin käyttäjällä on mahdollisuus kirjoittaa omin sanoin kokemuksiaan laitteen tai järjestelmän käytöstä, jolloin saadaan kerättyä laadullista tietoa. Avoimet kysymykset ovat kuitenkin raskaita analysoida ja usein niitä tehokkaampia ovat suljetut kysymykset, joihin käyttäjän tulee vastata valmiiksi annetuista vaihtoehdoista omia tuntemuksiaan parhaiten kuvaava vaihtoehto. Suljetut kyselyt ovat hyvin yleisiä käytettävyysskyselyissä, sillä ne ovat kaikista kustannustehokkaimpia subjektiivisen määrällisen tiedon keräämiseen. (Nielsen 1993)

Tutkimuksen tavoitteet ja tutkimusstrategia määrittelevät, sopivatko kyselyt osaksi käytettävyystudkimusta. Vaikka valmiita kyselyitä onkin kehitelty ja standardoitu käytettävyystudkimusta varten, on silti hyvä harkita tutkimuskohtaisesti, soveltuuko jo kehitetty kysely hyvin tutkimuksen tavoitteisiin. Valmiita kyselyitä voi myös muokata oman tutkimuksen erikoisluonteeseen sopivaksi, tai kyselyn voi laatia kokonaan itse. (Nielsen, 1993)

3.1.1 Mobiilikyselyt

Matkapuhelimet ja muut mobiililaitteet ovat nykypäivänä olennainen osa kuluttajien elämää sekä työn, vapaa-ajan että viihteen muodossa. Käyttäjäkokemus- ja käytettävyysskyselyt taas ovat tärkeä osa tuotekehitystä, ja kustannustehokkaita sekä luotettavia kyselytyyppejä kehitetään jatkuvasti. Erityisesti käyttäjäkokemuksen tutkimisessa olisi tärkeää, että kyselyyn vastaamisympäristö olisi mahdollisimman lähellä käyttäjän omaa luonnollista käyttöympäristöä ja -tilannetta. Tämän mahdollistavat erilaiset mobiili- ja verkkokyselyt, jotka ovat tehokkaan tiedonkeruutavan uusimpia muotoja.

Mobiilikyselyillä tarkoitetaan kyselyitä, joihin vastataan mobiilivälineellä, kuten matkapuhelimella (esimerkiksi kosketusnäyttöpuhelimet). Niiden avulla saadaan kerättyä käyttäjiltä tietoa entistä vaivattomammin ja edullisemmin, ja tulosten analysointi on nopeaa. Kyselyissä on kuitenkin kiinnitettävä huomiota kyselyn selkeyteen ja yksinkertaiseen rakenteeseen, ja otettava huomioon kyselyn ulkoasu sekä konteksti. (Väätäjä ja Roto, 2010)

Väätäjä & Roton (2009) mukaan mobiilikyselyiden tulisi olla riittävän lyhyitä, jotta niiden täyttäminen olisi nopeaa ja vaivatonta ja vastaajan mielenkiinto sekä motivaatio pysyisivät yllä koko kyselyn ajan. Käyttäjällä tulisi myös olla mahdollisuus vastata kyselyyn mobiilivälillä ilman internet-yhteyttä. Tekstiviestein suoritettavat kyselyt ovat käyttäjän kannalta mutkattomia, sillä matkapuhelin kulkee yleensä mukana kaikkialle, ja reagoiminen kyselyyn tapahtuu parhaimmillaan heti viestin saavuttua. Verkossa olevat kyselyt taas rajaavat verkon ulottumattomissa olevat käyttäjät pois vastaajajoukosta.

Mobiilikyselyille on olennaista hieman tiivistää alkuperäisen kyselyn kysymysmateriaalia tai jakaa kysely useisiin osiin. (Väätäjä ja Roto, 2009) Näin tehtäessä jotain olennaista tietoa saattaa jäädä kyselystä pois, ja toisaalta myös vastaajien vastausprosentin merkitys kasvaa. Tulokset ymmärrettävästi vääristyvät vähemmän, jos jättää vastaamatta yhteen kysymykseen, kun kysymyksiä on kaikkiaan 100, kuin jos kysymyksiä on vain 10.

Verkkokyselyiden avulla saadaan kerättyä käyttäjätietoa hyvin nopeasti, ja kehittyvän yhteiskunnan mobiilisovellusten myötä matkapuhelimiin lähetettävät kyselyt parantavat tätä tiedonkeruumenetelmää entisestään. (Väätäjä & Roto, 2009) Verkkokysely on mobiilikyselyn tavoin helppo kohdistaa tietylle kohderyhmälle, sillä linkin kyselyyn tai itse kyselyn voi lähettää esimerkiksi sähköpostilla tai tekstiviestillä matkapuhelimeen juuri halutuille henkilöille tai ryhmille.

Mobiilikyselyillä on paikkansa tulevaisuudessa etenkin teollisuuden alalla, sillä tuotekehitysprosessissa on usein tärkeää saada määrällistä tietoa mahdollisimman pian. Mobiilikyselyiden erityispiirteitä on vielä kuitenkin tutkittava runsaasti, jotta voidaan kehittää luotettavia ja käyttäjäystävällisiä kyselyitä juuri mobiilikäyttöliittymiin sopiviksi, ja saada luotettavia kyselytuloksia.

3.2 Standardikyselyt

Standardikyselyt ovat pitkän toistuvan kehitystyön, testaamisen ja muokkaamisen tuloksia, jotka ovat vakiinnuttaneet asemansa käytettävyyden arviointimenetelmänä. Niitä voidaan pitää luotettavina ja pätevinä oikeaan ja hyvin määriteltyyn tarkoitukseen käytettynä.

Standardikyselyt ovat objektiivisiä menetelmiä, ja ne on suunniteltu erityisesti käytettävyyssiantuntijoiden itsenäiseen käyttöön. Standardikyselyt estävät keksimästä pyörää uudelleen – jos hyvä ja luotettava kysely on jo kehitetty, sitä kannattaa hyödyntää. (Lewis, 1993)

SUMI¹ on Kirakowskin (1994) kehittämä kysely, jonka avulla voidaan arvioida ohjelmistotuotteen tai prototyypin käytön laatua sekä tunnistaa käytettävyyssvikoja ennen tuotteen viemistä markkinoille. SUMI:sta on saatavilla myös suomenkielinen versio, ja lomakkeita aloituspaketteineen voi ostaa ammattikäyttöön. SUMI-lomake sisältää 50 väittämää, joiden vastausvaihtoehdot ovat ”samaa mieltä”, ”eri mieltä” tai ”en tiedä”. SUMI-kyselystä saatuja tuloksia voidaan verrata markkinoiden yleiseen tilanteeseen ja muihin järjestelmiin (Vanhala 2005).

SUMI:ssa ohjelmiston käytettävyys jaetaan viiteen ulottuvuuteen, joista jokaista kohti lomakkeessa on kymmenen kysymystä. Ulottuvuudet ovat tehokkuus, tunne, avuliaisuus, hallinta sekä opittavuus. Tämän lisäksi lomakkeella mitataan yleistä käytettävyyttä, johon koostetaan näistä 50 väittämästä 25. (Kirakowski, 2000)

Parhaita tuloksia SUMI-kyselystä saadaan, jos kyselyyn vastaa keskimäärin 30 henkilöä verkkoversiossa tai 12 paperilla. Paperikyselyssä on etuna, että vastaajia voi valikoida verkkovastauksia paremmin, ja vastauksien voi rajata tarkemmin, eikä ennalta-arvaamattomia häiriötekijöitä ole, kuten verkon kautta vastaamisessa mahdolliset keskeytykset, samanaikainen musiikin kuuntelu tai puhelimesta puhuminen. (Kirakowski, 1994)

SUMI soveltuu tehokkaasti uusien tuotteiden arviointiin tuotekehitysprosessin monissa vaiheissa ja vertailemaan tuotteen eri versioita. SUMI:n avulla asetetaan tavoitteita käytön laadulle ja seurataan laadun kehittymistä ja näin tuodaan esille käyttöliittymän hyviä ja huonoja puolia. (Kirakowski, 1994)

Ruy (2005) vertailee kyselymenetelmiä tutkimuksessaan ja arvioi SUMI:n yhdeksi pätevimmistä ja luotettavimmista käyttäjätyytyväisyyttä ja järjestelmän käytön laatua mittaavista kyselyistä. SUMI:n etu muihin käytettävyyden mittauskeinoihin on menetelmän helppoudessa sekä verrattaisessa edullisuudessa niin tutkijoiden kuin

¹ Software Usability Measurement Inventory

osallistujienkin näkökulmasta. SUMI:n tehokas tulosten kerääminen ja analysoiminen on erityisesti ripeän tuotekehitysprosessin kannalta positiivinen asia.

Toisaalta SUMI:a voidaan hyödyntää vasta melko myöhäisessä vaiheessa tuotekehitystä, sillä kunkin järjestelmäversion tulisi olla testattavissa ja saatavilla koetilanteessa, joten SUMI:n käyttöä varten vaaditaan ainakin prototyyppi suunniteltavasta järjestelmästä. SUMI:n avulla ei myöskään saada selville kuin yleisiä järjestelmän toimivuuteen ja käytettävyyteen liittyviä seikkoja, ja ongelmien tarkempaa kuvausta ja paikallistamista varten on käytettävä toisenlaisia menetelmiä. (Ryu, 2005)

SUMI:n 50 väittämän kokoelma mittaa kattavasti eri käytettävyyden osa-alueita, ja ottaa myös huomioon ihmisen kokemukset tunteita käsittelevissä kysymyksissään.

SUS² sisältää kymmenen väittämän kokoelman, johon vastaaja reagoi valitsemalla eniten omaa näkemystään lähinnä olevan kohdan viisiportaisesta Likert-asteikosta. Vastapareina asteikossa ovat ”vahvasti samaa mieltä” sekä ”vahvasti eri mieltä”. SUS on jo pitkään käytössä ollut tiedonkeruumenetelmä tuotteen, järjestelmän tai palvelun käytettävyydestä, ja sen kehitti John Brooke vuonna 1986 (Finstad, K. 2006). Saadut vastaukset analysoidaan siten, että väittämille annetut pisteet lasketaan yhteen, jolloin saadaan vertailtava luku välillä 0-100, ja sitä kautta arvio järjestelmän tai tuotteen käytettävyydestä. (Brooke, 1996)

SUS:ia voidaan käyttää monipuolisesti tuotteiden ja järjestelmien, kuten internet-sivujen, matkapuhelinten käyttöliittymien sekä erilaisten vuorovaikutteisten sovellusten käytettävyyden mittaamiseen. Sen 0-100 -pisteytysasteikko on intuitiivisesti ymmärrettävissä, ja kyselyyn vastaaminen sekä tulosten laskeminen on rajatun kysymysmäärän takia nopeaa ja helppoa. SUS:ia ei ole myöskään patentoitu, joten sen käyttö on kustannustehokasta. (Ryu, 2005) Helppo hallinnoitavuus ja yksinkertainen pisteytys tekevät SUS:ista suosituksen arviointimenetelmän erityisesti tuotekehityksessä muiden menetelmien tukena (Finstad, K. 2006). Tullis ja Stetson (2004) ovat myös osoittaneet tutkimuksessaan SUS:in olevan muita vakiintuneita kyselyitä (esim. QUIS sekä CSUQ) luotettavampi internet-sivujen yleisen käytettävyyden arvioinnissa.

² System Usability Scale

SUS:in kymmenen väittämää sisältää huomattavasti viittauksia käyttäjän tuntemuksiin ja kokemuksellisiin näkökulmiin, joten sen avulla saadaan myös viitteitä käyttäjäkokemuksesta. Väittämissä viitataan esimerkiksi ”Tuote tuntui minusta helppokäyttöiseltä” tai ”Tunsin itseni varmaksi käyttäessäni tuotetta”. SUS:in avulla voidaan tutkia myös asenteita.

Toisaalta SUS:in avulla ei SUMI:n tavoin saada selville kuin yleisiä arvioita tuotteen tai järjestelmän käytettävyydestä, eikä yksityiskohtaisempi ongelmien osoittaminen ja analysointi ole mahdollista. SUS:issa tulee vastaan myös muitakin käytettävyysskyselyitä horjuttava ongelma – mitä yksittäinen saatu tulos tarkoittaa käytännössä, miten numeeriset arvot käännetään käytettävyyden kriteereiksi? (Bangor et al. 2009) Voidaanko esimerkiksi sanoa, että järjestelmä, joka sai SUS-kyselyn tuloksena arvioksi 50, on käytettävyydeltään hyvä?

Kyselylomakkeiden termit tulisi valita mahdollisimman yleisesti ymmärrettäviksi ja välttää harvinaisia sanoja. Finstad (2006) löysi tutkimuksessaan SUS-lomakkeesta termin, jonka vain syntyperäiset englannin kielen puhujat ovat ongelmitta ymmärtäneet. 8. väitteessä käytetään adjektiivia *cumbersome* (hankala, vaivalloinen) seuraavasti: ”*I found the system to be very cumbersome to use*”. Tämä on vaikuttanut negatiivisesti lomakkeen luotettavuuteen ja pätevyyteen vastaajien ollessa monikansallisia. Finstadin ehdotus ongelman korjaamiseksi on korvata termi sanalla *awkward*.

QUIS³ kehitettiin Marylandin Yliopistossa HCI-laboratoriossa. Se on standardoitu kyselymenetelmä, joka mittaa käyttäjän ja vuorovaikutteisten tietokonejärjestelmien välistä yhteyttä käyttäjän näkökulmasta. (Chin et al. 1993)

Kyselyn tarkoitus on olla tukena järjestelmien suunnittelussa ja arvioimisessa, tarjota tutkijoille pätevä työkalu subjektiivisen näkökulman mittaamiseen sekä toimia käytettävyysslaboratorioissa testityökaluna. QUIS on lisensoitu ja siitä on saatavilla lyhyt sekä pidempi versio niin paperisena kuin internet-kyselyn muodossakin. (Norman et al.)

QUIS-kotisivulla (Norman et al.) selvitetään kyselyn lähestymistapaa käytettävyyteen. QUIS tutkii käytettävyyttä ja sen arviointia keskimäärin viiden ulottuvuuden kautta

³ Questionnaire for User Interaction Satisfaction)

riippuen kyselyn versiosta. Ulottuvuudet ovat ohjelmiston yleisvaikutelma, näyttö, terminologia ja järjestelmän tiedot, oppiminen sekä järjestelmän kyvykkyys. Kyselyyn vastaaja arvioi lomakkeen kymmenportaisen asteikon avulla järjestelmää koskevia väitteitä ja sen ominaisuuksia.

QUIS käsittelee melko yleisesti käytettävyyttä ja sen näkökulmia, joten sitä voidaan käyttää monipuolisesti erilaisten järjestelmien arvioimiseen. Sitä voidaan pitää melko luotettavana ja pätevänä, sillä kyselyä on kehitetty jo seitsemän version verran. QUIS 7.0 on kaupallinen, ja ottaa viiden ulottuvuuden lisäksi huomioon opastuksen onnistuneisuuden sekä ohjelman asennuksen helppouden. Lisäksi kaksi eripituista versiota antavat mahdollisuuden valita joustavammin juuri tarkoitukseen sopivan kyselymuodon. Toisaalta lisensointi ja kaupallistuminen tarkoittavat myös kustannusten nousua – akateemiseen käyttöön tarkoitettu lomake maksaa 300 dollaria ja kaupallinen käyttö 1000 dollaria (Vanhala 2005).

QUIS on kuitenkin esimerkiksi SUS:ia yksityiskohtaisempi jo kysymysten määrän takia, ja keskittyy vähemmän kokemuksellisiin näkökulmiin. Siinä missä SUS arvioi kokonaisuutta, QUIS jakaa tutkittavat alueet kategorioihin ja pyrkii löytämään käytettävyysongelmia ja mielipiteitä tarkemmin niiden alta. Toisaalta järjestelmän opittavuuteen sekä kokonaisarvioon liittyvät kysymykset voidaan nähdä kokemuksellisuutta enemmän painottaviksi, joten QUIS-kyselyssäkin on käyttäjäkokemusta mittaavia piirteitä.

USE⁴-käytettävyysskyselyn kolme keskeisintä elementtiä ovat hyödyllisyys, tyytyväisyys ja helppokäyttöisyys. Kyselyä kehitettäessä tarkoitus on ollut luoda lyhyt kysely, jota voidaan käyttää käytettävyyden tärkeimpien ulottuvuuksien selvittämiseen ja näiden arvioimiseen. Kyselyä voi käyttää ohjelmistojen, laitteistojen ja palvelujen arvioimiseen ja sen avulla tuotteita voi vertailla keskenään. Kysely rakentuu 7-kohtaiselle Likertin asteikolle, jossa käyttäjät vastaavat tuntemustensa mukaan välillä ”vahvasti eri mieltä” ja ”vahvasti samaa mieltä”. (Lund, 1998)

Lundin (1998) faktorianalyysin mukaan jokainen tehty tutkimus osoitti, että käyttäjät arvioivat tuotteita ensisijaisesti kolmen ulottuvuuden kautta, joista USE muodostuu. Helppokäyttöisyys ja hyödyllisyys koettiin toisiinsa vastaavasti vaikuttavina, eli mitä

⁴ Usefulness, Satisfaction and Ease of Use

helppokäyttöisemmäksi tuote koettiin, sitä hyödyllisempi se myös oli ja päinvastoin. Helppokäyttöisyys voidaan vielä nähdä jaettavaksi kahteen osaan – käytön helppouteen sekä oppimisen helppouteen. Lund löysi myös näyttöä estetiikan ja hauskuuden merkityksestä tuotteen käytettävyyteen. Esteettisyys, mielihyvä ja viihtyvyys liitetään vahvasti käyttäjäkokemukseen, joten USE-kysely ottaa huomioon kokemuksellisen näkökulman erityisesti näihin liittyvillä kysymyksillään.

CSUQ⁵ on kenttätutkimukseen suunnattu käytettävyysskysely. Siinä käytetään 7-portaista Likert-asteikkoa, ja käytettävyyttä tarkastellaan kolmen pääulottuvuuden kautta. Nämä tarkastelukulmat ovat järjestelmän hyödyllisyys, tiedon laatu sekä käyttöliittymän laatu (Vanhala 2005). Pääulottuvuuksiensa kautta CSUQ on käyttäjätyytyväisyyttä kokonaisuutena mittaava kyselymenetelmä.

Lewisin (1993) tekemän tutkimuksen mukaan CSUQ soveltuu erittäin hyvin laboratorion ulkopuolelle tehtäviin tutkimuksiin. Lisäksi Lewis korostaa käyttäjätyytyväisyyden mittaamisen tärkeyttä erilaisissa käytettävyyttä koskevissa tilanteissa. Standardoitu CSUQ on luotettava ja sen tuottamia arvoja voidaan helposti verrata toisten järjestelmien ja laitteistojen saamiin tuloksiin. Lewis (2001) palasi omaan tutkimukseensa ja toisti aiemman kokeensa varmistaakseen CSUQ:n luotettavuuden ja pätevyyden. Tulokset vastasivat mainiosti aikaisemmin tehtyä tutkimusta, joten CSUQ:n voidaan katsoa olevan edelleen pätevä käyttäjätyytyväisyyden mittaamiseen tarkoitettu kysely.

Käyttäjätyytyväisyyteen liittyy helppokäyttöisyyden lisäksi vahvasti, miltä käyttäjästä tuntuu käyttäen tuotetta, ja minkälainen kokonaisvaltainen kokemus tuotteen ja käyttäjän vuorovaikutuksesta syntyy. Näin CSUQ:stakin on löydettävissä kokemuksellisia näkökulmia, joiden tutkiminen on käyttäjäkokemuksen kannalta todella tärkeää.

WAMMI⁶ on verkkosivustojen käytettävyyden arviointiin laadittu lomake. SUMI:n tavoin myös WAMMI on kaupallinen ja standardoitu tuote, ja siitä on tehty suomenkielinen versio (Vanhala 2005). WAMMI on pätevä ja kustannustehokas menetelmä internet-järjestelmän käytettävyyden mittaamiseen. Kysely jakautuu kuuteen

⁵ Computer System Usability Questionnaire

⁶ Web site Analysis and MeasureMent Inventory

kategoriaan, jotka ovat kiinnostavuus, kontrolloitavuus, tehokkuus, avuliaisuus, opittavuus sekä yleinen käytettävyys. (Moore et al. 2001)

WAMMI-kyselyn avulla internet-sivuston käytettävyyttä saadaan arvioitua nopeasti ja helposti, ja se on standardointinsa ansiosta hyvin suosittu menetelmä verkkosivujen käytettävyyden arviointiin. UsabilitNet-sivustolla (2006) on esimerkiksi mahdollisuus täyttää WAMMI-kysely.

Eri maiden väliset erot käyttäjien reaktioissa ja tyytyväisyyksissä vaihtelevat suuresti samalla kun verkkomarkkinat jatkavat voimakasta kasvuaan. WAMMI:n kehittäjät ovat ymmärtäneet äidinkielellä täytetyn kyselyn tärkeyden, ja WAMMI onkin saatavissa lähes 15 kielellä, ja tarjontaa kehitetään koko ajan. (WAMMI, 2010)

WAMMI on esimerkki rajatummasta käytettävyyskyselystä, sen avulla voidaan tutkia ainoastaan verkkosivujen käytettävyyttä. Sen suosio on osoittanut rajauksen onnistuneen mainiosti ja kyselylle olleen tarvetta. WAMMI ottaa huomioon teknisten asioiden lisäksi kokemuksellisia ja käyttäjän tuntemuksiin liittyviä asioita kuten turhautumista, tyytyväisyyttä ja käytön mukavuutta.

3.3 Muokatut kyselyt

Mikäli standardoidusta kyselypatteristosta ei löydä omiin tarpeisiinsa ja tutkimukseensa soveltuvaa kyselyä, voi kyselyn myös suunnitella kokonaan tai osittain itse (Kirakowski, 2000). Hyvin suunniteltu kysely vastaa standardoituja, yleisempiä käytettävyysteemoja kuvaavia kyselyitä paremmin tutkimusongelmaan, ja antaa kenties yksityiskohtaisempia selvityksiä järjestelmän tai tuotteen käytettävyydestä.

Kirakowski (2000) kuvaa kyselyjen suunnittelu- ja muokkaamisprosessia aikaa, kärsivällisyyttä ja rahaa vaativaksi työksi. Mikäli suunnittelee kyselyn itse, on tärkeää tutustua alan tutkimuksiin ja standardikyselyihin, jotta saisi omasta kyselystä mahdollisimman luotettavan ja pätevän. Omaakin kyselyä tulisi arvioida, kehittää ja tarkastella kriittisesti. Kirakowski kehottaakin harkitsemaan tarkkaan, olisiko omaan tarkoitukseen soveltuvaa standardoitua kyselyä jo tehty.

Kyselyitä voi myös muokata itse vastaamaan paremmin tutkittavaa asiaa. Standardoituihin käyttäjäkokemusta mittaaviin kyselyihin on esimerkiksi olemassa

valmiiksi vaihtoehtoja, joista on mahdollista jättää joitakin osia pois kyselystä kokonaan tai korvata kohtia oman tutkimuksen kannalta olennaisilla seikoilla. (WAMMI, 2010)

3.4 Kyselyiden käytön kulmakiviä

Kirakowskin (2000) mukaan luotettavan ja pätevän kyselyn avulla voidaan kerätä tietoa ja saada palautetta käyttäjän näkökulmasta, joten kyselyillä on tärkeä paikkansa tuotekehitysprojekteissa. Usein kyselyt ovat melko nopeita ja kustannustehokkaita, ja niiden avulla voidaan kerätä paljon määrällistä tietoa. Toisaalta useimmat kyselyt kertovat vain käyttäjän reaktion kokemastaan tilanteesta. Kysely voi kertoa, että tuotteessa on jotakin vikaa, mutta sen avulla ei voi tarkemmin määritellä, mikä vika oikeastaan on. (Kirakowski, 2000) Hyvin suunnitellulla, testatulla ja muokatulla kyselyllä voidaan toki päästä lähelle tällaista tilannetta.

Standardimuotoisia käytettävyyškyselyitä on vertailtu keskenään, jotta saataisiin selville yhdenmukaisuuksia ja kyselyiden vertailtavuutta. Tutkimuksissa on kuitenkin tullut esille (Wehsung & Naumann, 2008), että esimerkiksi SUS ja SUMI eivät ole kovin vertailukelpoisia keskenään, eivätkä ne ole kovin yksimielisiä tuloksissaan järjestelmän käytettävyydestä. Tämä johtunee suurelta osin siitä, että myös standardikyselyt painottavat hieman eri näkökulmia ja määrittelevät käytettävyyden mittarit ja mitattavat asiat toisistaan poikkeavilla tavoilla.

Wehsung & Naumann (2008) esittävät tutkimuksessaan, jonka avulla pyrittiin selvittämään eri käytettävyyškyselyiden eroja multimodaaleja järjestelmiä arvioitaessa, että SUMI:n tulokset ovat eniten ristiriidassa muiden kyselyiden tulosten kanssa. Vertailtavat kyselyt olivat SUS, SUMI, AttrakDiff⁷ sekä SASSI⁸. SUMI:n tulokset olivat vahvasti yhteisymmärryksessä muiden kyselyiden tulosten kanssa ainoastaan mitattaessa tunteisiin perustuvia puolia. Vertailussa SUS oli sen sijaan melko johdonmukainen. Kyselyistä ei myöskään voitu määritellä yksiselitteisesti parhaiten käytettävyyttä arvioivaa kyselyä.

⁷ AttrakDiff on käyttäjäkokemusta tutkiva kysely, jota käsitellään luvussa 4.2.1

⁸ SASSI (Subjective Assessment of Speech System Interface) on kyselymenetelmä, jonka avulla tutkitaan puheentunnistusta varten suunniteltujen käyttöliittymien käytettävyyttä. Lisätietoa Hone, K.. A Tool for The Subjective Assessment of Speech System Interface. Saatavissa: <http://www.brunel.ac.uk/~csstksh/sassi.html>. [Viitattu 27.11.2010]

Kyselytutkimuksilla saadaan kerättyä määrällistä tietoa nopeasti ja helposti suuria määriä, mutta kuinka monta käyttäjää riittää varmistamaan kyselyn luotettavuuden? Tullis ja Stetson (2004) pureutuivat tutkimuksessaan itse kehittelemänsä kyselymenetelmän arviointiin ja jo vakiintuneiden kyselyiden, SUS:in, QUIS:in ja CSUQ:n luotettavuuteen. Kyselyillä pyrittiin arvioimaan kahta internet-sivustoa sekä tutkimaan, kuinka pieni käyttäjien määrä tuottaa riittävän luotettavan tuloksen kyselystä. Tutkimuksessa päädyttiin siihen, että kohtuullisen luotettavan kyselyn voi tehdä jopa 12-14 koehenkilöllä. Yleisesti tutkijat kuitenkin suosivat vähintään 20 koehenkilön otoksia. Tutkimuksen mukaan kyselyistä yksinkertaisin SUS saavutti kaikista luotettavimmat tulokset pienillä koehenkilömäärillä. SUS:in tarkastelupaino on järjestelmän tai tuotteen kokonaiskäytettävyydessä, hyödyllisyydessä ja helppokäyttöisyydessä, eikä tarkemmin eriteltynä esimerkiksi visuaalisessa ilmeessä tai sisällön jäsentelyssä, kuten useissa muissa käytettävyysskyselyissä, joten sitä voidaan pitää luotettavana yleisen käytettävyyden ja kokonaisuuden subjektiivisen näkökulman arviointimenetelmänä.

Kyselyiden suurin ongelma on käytettävyyden mittaamisessa yleisestikin tunnistettu - minkä kriteereiden mukaan tuotteen käytettävyyttä ja erityisesti subjektiivista kokemusta voi kutsua hyväksi? Lund (1998) tiedosti tarpeen standardoiduille käytettävyyden mittareille, ja hänen mukaansa mittareiden avulla voidaan paremmin ja realistisemmin määritellä, mitä hyvään käytettävyyteen vaaditaan. On tärkeää löytää olennaisimmat suunnittelulementit, jotka tukevat hyvää käytettävyyttä. Standardikyselyiden tavoite on saada yhdenmukaisia mittausmenetelmiä, joiden avulla tuotteita olisi mahdollista arvioida luotettavasti sekä vertailla markkinoilla oleviin vastaaviin tuotteisiin.

Kuten yllä esitetyistä tutkimustuloksista nähdään, eivät standardikyselytkään aina ole täysin yhdenmukaisia. Standardikyselyiden avulla hallitaan kuitenkin hyvin tulosten laadukkuutta ja luotettavuutta, sillä ne ovat käyneet läpi monta arviointi- ja parannusprosessia (SIGHI, 2010). Omien tai muokattujen kyselyiden suunnittelemisessa ja hallinnoimisessa on aina riskinsä, ja vaikka niiden avulla voitaisiin saada juuri tapaukseen sopivaa yksilöllistä ja yksityiskohtaista tietoa, eivät tulokset välttämättä ole lainkaan yleistettävissä, niiden hallinnointi ja analysointi on vaikeaa tai lähes mahdotonta, ja pahimmassa tapauksessa kysely ei mittaa sitä, mitä halutaan.

4. Käyttäjäkokemus ja käytettävyysskyselyt

Käyttäjäkokemus, sen arviointi ja tiedonkeräysmenetelmät ovat juuri nyt käytettävyyssalalla erittäin kiinnostava aihe. Teollisuudessakin on ymmärretty käyttäjäkokemuksen merkitys muun muassa asiakasuskollisuuden tekijänä, joten motivaatiota aiheen tutkimiseen riittää. Kyselyt ovat menetelmänä erittäin tehokas subjektiivisen määrällisen tiedon kerääjä, ja niitä käytetään paljon käyttäjäkokemuksen arvioimiseen.

Myös standardikyselyitä on hyödynnetty käyttäjäkokemuksen mittaamisessa. Luku 4.1 tuo esiin standardikyselyiden soveltamisen näkökulmia sekä haasteita. Erityisesti käyttäjäkokemuksen arvioimiseen kehitettyjä kyselyitä käsitellään luvussa 4.2. Luvussa 4.3 käsitellään haasteita ja asioita, joita tulisi ottaa huomioon kyselymenetelmiä käytettäessä käyttäjäkokemuksen mittaamiseen.

4.1 Standardikyselyt ja käyttäjäkokemus

ISO-standardin mukaan käytettävyyden kriteereitä ja menetelmiä voidaan käyttää myös käyttäjäkokemuksen arvioimiseen, mikäli käytettävyyttä tarkastellaan käyttäjän henkilökohtaisista tavoitteista käsin, ja siihen liittyy havaintoihin ja tunteisiin pureutuvia näkökulmia. (ISO/IEC 9241-210, 2010) Luvussa 3.2 käytiin läpi yleisimpiä standardikyselyitä, ja niistä löydettiin elämyksellisyyttä ja käyttäjän tunteita mittaavia kohtia, joten niitä on mahdollista joiltain osin hyödyntää käyttäjäkokemuksen tutkimisessa.

Kuten luvussa 3.2 on mainittu, SUS on yleispätevä ja hyvä järjestelmän tai laitteen kokonaiskäytettävyyden mittari, ja sitä voidaan soveltaa myös käyttäjäkokemuksen mittaamiseen. Finstad'in (2010) artikkelin mukaan SUS ei kuitenkaan yksinään ole täysin riittävä käyttäjäkokemusta mittaamaan, sillä sen avulla ei saada selkeää kokonaiskuvausta käyttäjäkokemuksesta. Sen sijaan se voi Finstadin mukaan toimia hyvänä lähtökohtana tuotteen tai palvelun käyttäjäkokemuksen arvioimiselle, ja siitä saatujen tulosten avulla voidaan tehokkaammin valita seuraava, tarkemmin käyttäjäkokemusta mittaava menetelmä. Finstad kehittelikin SUS:in pohjalta vielä lyhyemmän ja paremmin sekä ISO-standardin että käyttäjäkokemuksen huomioon

ottavan kyselyn, UMUX:n⁹, joka lopulta sekin keskittyy liikaa käytettävyyden subjektiiviseen arvioimiseen.

Wechsung & Naumann (2008) raportoivat tutkimuksestaan, jossa vertailivat SUMI:a, SUS:ia ja AttrakDiffi:a multimodaalisten järjestelmien käytettävyyden arvioinnissa. Näiden kyselyiden tulokset eivät olleet yleisesti vertailukelpoiset keskenään, sillä kukin kysely painottaa hieman eri näkökulmia käytettävyydestä. Käyttäjäkokemuksen kannalta huomattavaa kuitenkin on, että tutkimuksen mukaan sekä AttrakDiffin että SUMI:n kokemuksellisuutta ja elämyksellisyyttä mittaavat puolet olivat vertailukelpoisia keskenään. Tämä tutkimustulos osoittaisi, että SUMI:n kokemukselliset väittämät sopivat hyvin käyttäjäkokemuksen tutkimiseen. Samanlaisia viitteitä antaa Frokjaerin et al. (2000) tutkimus, jonka mukaan SUMI:n avulla voidaan mitata käyttäjätyytyväisyyttä, sillä kysymysasettelut ovat asenteita mittaavia ja arvioivia. Lisää tutkimustuloksia kuitenkin tarvitaan, jotta voidaan vahvistaa AttrakDiff:in ja SUMI:n positiivinen vertailukelpoisuus.

Tutkielmassa käsitellyt standardikyselyt ottavat huomioon käyttäjäkokemuksen kannalta olennaisia näkökulmia, mutta niiden painotuserot vaihtelevat merkittävästi. Esimerkiksi yksi SUMI:n kuudesta ulottuvuudesta painottaa käyttäjätyytyväisyyttä ja kokemuksellisia аспекteja, kun SUS:in 10 väittämästä lähes kaikki liittyvät jollain tavalla kokemuksellisuuteen ja käyttäjän tunteisiin järjestelmää käytettäessä. WAMMI:n väittämiin mahtuu tarkempaa tietoa sekä järjestelmästä tarkemmin että tuntemuksista sen käyttöä kohtaan, jolloin kokemuksellisuus jää pienempään osaan, kun USE:n kohdalla yksi kolmesta ulottuvuudesta, tyytyväisyys, liittyy selvästi elämyksellisyyteen. Painotuseroista johtuen kyselyt eivät ole keskenään kovin vertailukelpoisia käyttäjäkokemuksen tutkimisessa.

Näyttää siis on, että standardikyselyitä voidaan hyödyntää myös tunteita mittaavissa tutkimuksissa, mutta tarvitaan lisää tutkimustietoa aiheesta, jotta tulevaisuudessa olisi mahdollista hyödyntää kyselyitä luotettavammin.

⁹ Usability Metric for User Experience

4.2 Käyttäjäkokeuskyselyt

Erilaisia käyttäjäkokeuskyselyitä on kerätty All About UX -sivustolle, jossa esitellään 16 käyttäjäkokeuksen tutkimiseen tarkoitettua kyselyä tai asteikkoa. Tässä luvussa käydään läpi lyhyesti AttrakDiff-kysely (4.2.1), joka painottaa käytännöllisyyden ja mielihyvän näkökulmia, läsnäolokysely PQ (4.2.2), joka mittaa läsnäolon voimakkuutta ja merkitystä käyttäjäkokeuksessa sekä yksi käyttäjäkokeuksen yleisimmistä tutkimuksen kohteista, pelimaailma ja lapsen pelikokemusta mittaava KidsGEQ (4.2.3). Näiden lyhyiden katsausten avulla pyritään antamaan tiivistä viitettä siitä, kuinka monipuolisia asioita käyttäjäkokeuskyselyillä voidaan mitata.

4.2.1 AttrakDiff

AttrakDiff on Hassenzahlin et al. (2010) suunnittelema kysely, joka painottaa mielihyvän sekä käytännölläisyyden laadun mittaamista vuorovaikutteisissa tuotteissa ja on saatavilla vapaasti internetissä. AttrakDiffin suunnittelijoiden tutkimuksen mukaan tuotteen tai järjestelmän käytöllä tavoitellaan mielihyvää sekä käytännölläisyyttä ja ne myös nähdään toisistaan itsenäisinä. Molemmat tekijät vaikuttavat olennaisesti tuotteen kiinnostavuuteen.

AttrakDiff-kyselyssä on vastakohta-adjektiiveja, joista käyttäjän on valittava viisiasteikkoisella skaalalla parhaiten omaa kokemustaan kuvaava kohta. AttrakDiffin kehittäjä on UI Design GmbH, ja kysely on alun perin käännetty saksan kielestä. AttrakDiff Basicin avulla käyttäjä voi vastata kyselyyn verkossa. Tuloksena kyselystä tulee graafinen yhteenveto, joka sijoittaa tuotteen tai järjestelmän käyttäjäkokeuksen mielihyvän ja käytännön koordinaatistoon. (Roto et al. 2010a)

AttrakDiffin avulla voi tehdä kolmenlaisia kyselyitä ja arviointeja. Tuotetta voi arvioida yksinään, vertailla tilannetta ennen kehittelyä sekä sen jälkeen tai vertailla kahta eri tuotetta keskenään. (Hassenzahl et al. 2010) Kuva 1 esittää saman tuotteen vertailua ennen muutoksia sekä sen jälkeen. Mitä lähemmäksi oikeaa yläreunaa tuloslaatikko sijoittuu, sitä paremmaksi AttrakDiffin avulla arvioitu käyttäjäkokeus arvioidaan.



Kuva 1: Ennen-jälkeen -tilannetta vertailevan AttrakDiff-kyselyn tulos mielihyvä-käytäntö -koordinaatistossa (Hassenzahl et al. 2010) (Hassenzahl, 2010)

AttrakDiffia voidaan käyttää niin kentällä, laboratoriossa, kuin verkossakin. Kysely on saavuttanut standardiaseman, eli sitä voi melko luotettavasti käyttää käyttäjäkokemuksen mittaamisessa. AttrakDiff tuottaa tuloksena arvosanan, joka ei kenties vielä yksinään ole selkeästi analysoitavissa, mutta jota voidaan vertailla tehokkaasti muiden tuotteiden tai järjestelmien saamaan arvoon samasta kyselystä. AttrakDiff ei tuota laadullista tietoa, joten sillä ei voi kovin hyvin selvittää yksityiskohtaisia ongelmia järjestelmässä tai tuotteessa, mutta se antaa viitteitä siitä, miltä käyttäjistä tuntuu käyttää tuotetta tai järjestelmää kokonaisuudessaan, ja mitkä osa-alueet eivät täytä toiveita.

AttrakDiff-kyselyä on kritisoitu sanojen merkitysten muuttumisesta käännettäessä. Luotettavimmat käytettävyysskyselyn tulokset saadaan, kun henkilö vastaa kyselyyn omalla äidinkielellään, jolloin tämäkin kysely, toimiakseen Suomessa, tulisi kääntää suomen kielelle. Tämä taas tarkoittaa lisää epäonnistumisen mahdollisuuksia ja sanojen merkitysten muuttumista, kun jo kerran käännettyä aletaan vielä kääntää (Roto et al. 2010a). Englannissa on myös todettu olevan paremmin tähän kyselyyn soveltuvia termejä, ja suomenkielisiä vastaavia on haastavaa löytää. Kehitys- ja arviointityö jatkuu kuitenkin koko ajan.

4.2.2 Läsnaoloa tutkiva kysely

Läsnaololla tarkoitetaan sitä, että henkilö kokee olevansa jossakin paikassa tai ympäristössä, vaikka sijaitsisikin fyysisesti muualla. Läsnaolon elementtejä ovat uppoutuminen, osallisuuden tunne sekä valikoiva tarkkaavaisuus. Witmer & Singer (1998) ovat kehittäneet läsnaoloa koskevan kyselyn, PQ, joka mittaa

virtuaaliympäristössä koetun läsnäolon ja mahdollisten sitä edistävien tekijöiden vaikutusta kokemuksen voimakkuuteen. PQ käyttää 7-kohtaista asteikkoa, joka koostuu vastakohtapareista sekä keskellä olevasta neutraalista tekijästä. Kysymyksiä on yhteensä 32. Käyttäjä merkitsee rastin siihen kohtaan asteikkoa, joka vastaa parhaiten käyttäjän kokemusta. Asteikon käyttäminen on erittäin hyödyllistä, mikäli se mittaa tarkalleen sitä, mitä sen on tarkoitus mitata ja tekee sen hyvin.

PQ:ta on tutkittu melko paljon, ja sitä voidaan pitää verrattain luotettavana kyselymenetelmänä. Usoh et al. (2000) ovat tutkineet kahden erilaisen läsnäoloa mittaava kyselyn käyttöä sekä virtuaali-, että todellisessa ympäristössä, ja saaneet näyttöä siitä, että PQ mittaa verrattain hyvin myös todellista ympäristöä.

Käyttäjäkokemuksen mittaamiseen PQ on siis erityisen hyvä virtuaaliympäristöissä, ja jatkotutkimuksen kannalta kiinnostavaa on esimerkiksi selvittää oppimisen ja suoriutumisen vaikutusta läsnäolon kokemukseen (Witmer & Singer, 1998). Uppoutumisen käsite on tutkimuksissa melko uusi, ja se on havaittu erittäin tärkeäksi osaksi etenkin vuorovaikutteisessa virtuaaliympäristössä. Se tekee kokemuksesta merkityksellisemmän ja läsnäolosta tiiviimmän. Tulevaisuudessa voisi olla myös kiinnostavaa tutkia, sopivatko tietynlaista uppoutumiskokemusta edistävät ympäristöt paremmin oppimiselle, esimerkiksi kokonaisvaltaista käyttäjäkokemusta tavoitteleva SecondLife-virtuaalimaailma.¹⁰

4.2.3 Lasten pelikokemusta tutkiva kysely

Pelikokemukset ovat yksi käytettävyyden tutkituimmista aiheista. Peliaiheet tutkimukset keskittyvät lähinnä tutkimaan pelaamisen positiivisia että negatiivisia vaikutuksia, kuten sosiaalista kanssakäymistä ja itsensä kehittämistä (esimerkiksi Singstar Playstationille) tai aggressiivista käyttäytymistä (väkivaltaiset tietokonepelit). Lasten kokemuksia digitaalisten pelien pelaamisesta ei ole tutkittu vielä kovin laajasti. Poels et al. (2008) ovat kehittäneet lasten pelikokemusta tutkivan kyselyn, KidsGEQ¹¹, joka on tarkoitettu 8-12 -vuotiaille lapsille, ja sen tavoitteena on mitata pelikokemuksen

¹⁰ SecondLife on virtuaaliympäristö, jossa voi luoda itselleen Avatarin, joka luo kontakteja, osallistuu luentoihin, harrastaa, oppii ja havainnoi virtuaalimaailmassa. SecondLife on suosittu palvelu, jolla on yli 14 miljoonaa käyttäjää. Lisätietoa: <http://secondlife.com/whatis/?lang=en-US>. [Viitattu 28.11.2010]

¹¹ the Kids Game Experience Questionnaire

aikana havaittuja kokemuksia ja tuntemuksia. Lapsi voi täyttää kyselyn itsenäisesti ja nopeasti, eikä se häiritse pelin pelaamista. Kysely perustuu aikuisille suunnattuun GEQ-pelikokemuksekyselyyn, mutta sen muoto ja sanavalinnat on muokattu paremmin lasten ymmärrettäviksi. 21-kohtainen kysely jaksaa pelikokemuksen 7 ulottuvuuteen, jotka käsittävät upottautumisen, jännityksen, kokemukset kyvyistä, negatiiviset ja positiiviset puolet, haasteellisuuden sekä mukaansatempaavuuden.

Poels et al. (2008) testasivat kyselyä 31 11-13-vuotiaalla lapsella, jotka pelasivat aartenmetsästyspeliä langattomin mobiiliohjaimin. Jokaisen pelin jälkeen lapset täyttivät KidsGEQ:n, ja selvisi, että kysely kohdat ymmärrettiin hyvin, mutta kysely kokonaisuudessaan koettiin hieman liian pitkäksi. Luotettavuusanalyysin mukaan neljä ulottuvuudesta olivat kohtuullisen luotettavia, mutta jännityksen sekä positiivisten ja negatiivisten elementtien kohdalla luotettavuus oli melko matala. Kyselyn tulevaisuuden kannalta on tärkeää tehdä lisää tutkimus- ja kehitystyötä luotettavuuden parantamiseksi.

4.3 Kyselymenetelmien käytön haasteita

Kenties suurin haaste kyselyiden hyödyntämisessä käyttäjäkokemuksen mittaamisessa on Vääätäjän ja Roton (2009) mukaan tutkijoiden ristiriitainen ja monimuotoinen käyttäjäkokemuksen määritelmä. Käytettävyyden ja käyttäjäkokemuksen suhde on edelleen kiistelty aihe, ja vaikka uusin standardi tuo hieman selkeyttä ja antaa rajaviivoja käyttäjäkokemuksen käsitteeseen, on sitäkin kritisoitu hieman liian laajaksi ollakseen selkeä ja yksiselitteinen (Jokela, 2010). Tästäkin syystä monet käytettävyysekyselyt painottuvat edelleen erityisesti käytettävyyden mittaamiseen, eikä elämyksellisiä näkökulmia ole otettu kunnolla huomioon. Jotta kyselyjä voitaisiin pitää pätevinä ja luotettavina, olisi päästävä yksimielisyyteen tai vähintään selkeyteen käyttäjäkokemuksen käsitteestä.

Ollakseen päteviä ja luotettavia akateemiset kyselyt tukeutuvat ja pohjaavat vankasti teoreettiseen tietoon. Eri tutkijoilla ja siten myös kyselyillä on usein eri tarkastelu- ja painotuskulmat käyttäjäkokemukseen, joten on erityisen tärkeää ymmärtää kunkin kyselyn olennaisin tarkoitus ja lähtökohdat, sekä mitä kyselyllä mitataan. (Vääätäjä ja Roto, 2009) Toisaalta kyselyiden monipuolisuus ja erilaaisuus mahdollistavat myös erilaisten kyselyiden yhdistelemisen omiin tarkoituksiin sopivaksi.

Käyttäjäkokeemusta käsittelevä kirjallisuus vilisee käyttäjäkokeemukseen vaikuttavia asioita käyttäjän tunteista, käyttökontekstista ja järjestelmästä tilannesidonaisuuteen. Kyselyitä käytettäessä käyttäjäkokeuksen arviointiin voikin olla haasteellista ymmärtää näiden eri tekijöiden suhde ja vaikutus toisiinsa, jolloin kyselyiden analysoinnissa virheiden riski kasvaa. (Väätäjä ja Roto, 2009)

Väätäjän ja Roton (2009) mukaan käyttäjäkokeemus on tilanneriippuvaista, ja käyttäjäkokeuksen kannalta olennaista olisi, että kysely voitaisiin täyttää mahdollisimman lähellä todellista käyttötilannetta, jotta mittaustulokset olisivat mahdollisimman todenmukaisia ja päteviä. Mobiilikyselyt ja niiden kehittäminen ovat tästä näkökulmasta katsottuna erittäin edistyksellisiä.

Käyttäjäkokeuksen arviointia varten tarvitaan määritellyt, selkeät mittarit, joiden mukaan voidaan arvioida käyttäjäkokeemusta ja verrata sitä muihin. Erityisesti tämä ongelma heräsi SIGCHI-syysseminaarissa (2010). Mitkä ovat luotettavat ja hyväksyttävät mittarit hyvälle käyttökokeukselle, mitä lopulta kertoo esimerkiksi AttrakDiffistä tai SUS:ista tuloksena saatu arvo? Mitä arvoa voidaan pitää riittävän hyvänä?

Jotta kyselyä hyödynnettäisiin tuotekehityksessä mahdollisimman tehokkaasti, tulisi kyselyn avulla selvittää entistä tarkemmin, mitkä tekijät aiheuttavat järjestelmässä ongelmia tai mitkä ovat erityisen ansioituneita tarkoituksessaan. Näin kyselyillä selviäisi yksityiskohtaisemmin käyttäjäkokeuksen ja käytettävyyden ongelmia, ja niihin olisi mahdollista paneutua tuotekehityksessä. Yleispätevä ja kokoava näkökulma antaa yleiskatsauksen laitteen tai järjestelmän käyttäjäkokeuksesta, mutta ei vielä riitä tarkempaan analyysiin.

On tärkeää, että tehtyjä kyselyitä jaetaan julkisesti, oli kyse sitten akateemisesta tutkimuksesta tai teollisuuden alan kaupallisesta kyselystä, jotta niistä olisi mahdollista oppia ja hyötyä. Standardoidut kyselyt ovat käyneet läpi toistuvan arviointi- ja parannusprosessin, joten kehityksen ja laajan testaamisen kautta saadaan vielä aikaan käyttäjäkokeuksellekin pätevä ja luotettava kyselypatteristo eri tutkimustilanteita varten. All About UX –sivusto (2010) on hyvä alku.

Kyselyiden hyödyllisyyttä on tutkittu muun muassa kyselyiden hallinnoinnin sekä vastaajien taustan ja kokemuksen kautta. Root & Draperin tutkimuksessa pohdittiin

(1983) käyttäjien järjestelmä- tai tuotetuntemuksen vaikutusta kyselyn tulokseen. Mikäli käyttäjät eivät olleet käyttäneet järjestelmää tai sovellusta aikaisemmin, avoimet ja yleiset kysymykset eivät vastanneet kovin hyvin käytettävyyden arviointia todellisessa tilanteessa. Oli myös turhaa kysyä käyttäjiltä mielipidettä jonkin tietyn asian muuttamiseen, mikäli heille ei annettu mahdollisuutta testata järjestelmää muutoksen kera. Näin ei saatu mitattua todellista tyytyväisyyttä muutokseen, ainoastaan tyytyväisyyttä järjestelmään ylipäänsä.

Luvussa 3.2 esitettiin erilaisia standardimallisia käytettävyyškyselyitä. Kuten Ruy & Smith-Jackson (2006) tuovat artikkelissaan esille, standardikyselyt, kuten SUMI ja QUIS, ovat liian yleisiä mittaamaan tarkkoja käytettävyydongelmia tuotteissa tai järjestelmissä. Niiden avulla päästään tarkempiin tuloksiin ainoastaan käyttötilanteen ja loppukäyttäjien tarkalla määrittelyllä, sekä ymmärtämällä ja kuvaamalla tarkasti, mitä järjestelmästä halutaan arvioida. WAMMI:n kaltaisia rajatumpia ja tarkempia kyselyitä tarvitaan.

Brooke (1996) kuvailee raportissaan, että käytettävyydelle ei ole olemassa ehdottomia kriteerejä, sillä tuotteen tai järjestelmän käytettävyys on aina käyttötilanteeseen sidottu, jolloin käytettävyyden kannalta olennaisinta on tuotteen tai järjestelmän sopivuus tarkoitukseensa. Tätä samaa voisi soveltaa käyttäjäkokemuksen mittaamiseen standardikyselyiden avulla - olennaisinta on, että oli kysely sitten standardi tai itse kehitetty, se sopii tarkoitukseensa, ja mittaa sitä, mitä halutaan.

Käyttäjäkokemuskyselyt painottavat erityisen hyvin ja usein lähes pelkästään käyttäjän tunteisiin ja kokemuksiin liittyviä seikkoja, mutta tämäkään ei tunnu miellyttävän kaikkia tutkijoita. Jokela (2010) arvostelee kirjassaan liian keskittynyt tutkimusta pelkästään käyttäjäkokemuksen ympärille. Hän painottaa, että käyttäjäkokemus on tärkeä ymmärtää ja hyväksyä osana käyttäjätestausta ja -tutkimusta, mutta sen kustannuksella ei saisi unohtaa järjestelmän tai tuotteen teknisempää käytettävyyttä. Käyttäjäkokemuskyselyiden tuloksia analysoitaessa tulisi pitää mielessä kyselyiden subjektiivisuus ja tunnepitoisuus. On muistettava, että käyttäjät eivät ole järjestelmän asiantuntijoita, eivätkä he myöskään aina osaa tunnistaa kokemiaan tunteita, jolloin kyselyn vastaukset voivat vääristyä. Jokela muistuttaakin alan paradoksaalisesta sanonnasta, joka kehottaa kuuntelemaan käyttäjiä, mutta jättämään huomiotta heidän kommenttinsa.

5. Johtopäätökset

Tässä tutkielmassa on perehdytty käytettävyysskyselyihin käyttäjäkokemuksen mittaamisessa. Yleisimpiä käytettävyyden mittaamiseen suunniteltuja standardoituja kyselyitä on esitelty sekä pohdittu niiden sopimista tarkoitukseensa ja soveltumista käyttäjäkokemuksen tutkimiseen. Lisäksi on esitelty muutama käyttäjäkokemuksen tutkimista varten suunniteltu kysely käyttäjäkokemuksen ja sen tutkimuksen laajuuden konkretisoimiseksi.

Vaikka Nielsen (1993) ei mainitse vielä kirjassaan käyttäjäkokemusta, puhuu hän kuitenkin subjektiivisista tunteista ja niiden mittaamisen tärkeydestä, joita käyttäjäkokemus käsittää. Nähdään siis, että käyttäjän tunteita on otettu huomioon jo aiemmissa tutkimuksissa ja teorioissa, mutta viime vuosikymmenen aikana niiden merkitys on vasta korostunut huomattavasti.

Käyttäjäkokemuksen tutkimisen kannalta haastavaa on erityisesti, että ihmisten tarpeet ja vaatimukset ovat yksilöllisiä ja liittyvät hyvin vahvasti tunteisiin, ja käyttäjäkokemuksen määritelmä on yhä tutkijoiden tiukan kritiikin alla. Kaikkien miellyttäminen on mahdotonta. Tulevaisuudessa olisi siis löydettävä järjestelmistä ja tuotteista juuri kohderyhmälle yleistettävissä olevia mittareita, joiden avulla hyvä käyttäjäkokemus määriteltäisiin. Tämä ongelma on yleisesti tiedostettu käyttäjäkokemusta käsittelevässä kirjallisuudessa, ja mittareiden kehittämisen eteen tullaan tulevaisuudessa varmasti tekemään aktiivista tutkimustyötä.

Kyselyt ovat erittäin tehokas subjektiivista määrällistä tietoa keräävä menetelmä, ja sopivat hyvin käyttäjäkokemuksen arvioimiseen erityisesti osana muita menetelmiä. Eri kyselyt painottavat kuitenkin eri näkökulmia ja määrittävät käytettävyyden kannalta olennaiset seikat eri tavalla, jolloin tulosten vertailtavuus eri kyselyiden välillä on vaikeaa. Tulevaisuuden tutkimuksessa olisi näin ollen kysyntää tarkemmalle vertailevalle tutkimukselle eri kyselyiden välillä, jolloin painotuserot osattaisiin ottaa paremmin huomioon tulosten analysoinnissa.

Ruy & Smith-Jackson (2006) painottavat myös, että tietopohjan kehittämisen edistämiseksi tulisi tehdä enemmän analyyttisiä tutkimuksia alan asiantuntijoilta tai käytettävyydestäustilanteista. Niin kauan, kuin tutkimustulokset ovat melko satunnaisia, hieman eri aiheita painottavia tutkimuksia, ei voida yleistää oikein mitään

käytettävyysskyselyiden pätevydestä ja soveltuvuudesta käyttäjäkokemuksen mittaamiseen. Positiivista näyttöä kuitenkin on, erityisesti SUMI:n ja SUS:in kohdalla, mutta tämän väittämän tueksi tarvitaan vielä rutkasti tutkimusmateriaalia.

Vaikka standardikyselyt ovat luotettavia ja päteviä, ja erityisesti SUS ja SUMI painottavat yleistä käytettävyyttä sekä tunnepuolta, ovat ne kuitenkin niin yleisiä menetelmiä, ettei niiden avulla ole mahdollista kerätä yksityiskohtaista tietoa tuotteen tai järjestelmän eri osien luomasta käyttäjäkokemuksesta. Standardikyselyt eivät ehkä niinkään ole käyttäjäkokemuksen mittaamiseen yksinään soveltuva menetelmä, mutta sopivat sitä paremmin muiden menetelmien tueksi.

Käyttäjäkokemuskyselyt, kuten AttrakDiff, on suunniteltu nimenomaan käyttäjän tunteita ja elämyksellisiä näkökulmia mittaamaan, joten niiden avulla saadaan standardikyselyitä tarkempaa tietoa siitä, miltä laitteen tai järjestelmän käyttäminen tuntuu. Toisaalta moniakaan näistä kyselyistä ei voida pitää vielä kovin pätevinä, sillä ne ovat vasta kehitysvaiheensa alkutaipaleella. Kehitys- ja arviointityötä tehdään jatkuvasti, ja käyttäjäkokemuksen arviointi kyselymenetelmin tulee varmasti kasvattamaan suosiotaan erityisesti tuotekehityksen piirissä.

Standardoiduissa kyselyissä on kuitenkin sen verran potentiaalia, ja niiden eteen on tehty niin paljon työtä, että niitä kannattaa ehdottomasti hyödyntää ja jalostaa tulevaisuuden käyttäjäkokemustutkimuksessa. Standardikyselyiden muokkaaminen entistä kokemuksellisempia elementtejä painottavampaan suuntaan voisi olla paikallaan, ja esimerkiksi AttrakDiffin ja SUMI:n näkemyksiä yhdistäviä kyselyitä voisi testata käyttäjäkokemuksen tarkemmassa mittaamisessa.

Inhimilliset käyttäjät ovat mutkikkaita tutkittavia. Käyttäjät eivät aina itsekään osaa eritellä kokemiaan tunteita, saattavat antaa enemmän painoarvoa käyttötilanteen loppuvaiheen kokemuksille tai hävetä omia kykyjään esimerkiksi pelitilanteissa, ja vääristellä tuloksia tästä syystä. Tulevaisuudessa voitaisiin tutkia mahdollisuutta liittää käytettävyysskyselyiden yhteyteen pieni käyttäjän haastattelu tai käyttäjän havainnointia, jolloin kyselyiden tulokset eivät olisi ainut tulos pohja, johon nojataan, kun tehdään esimerkiksi päätöksiä tuotesuunnittelun seuraavista vaiheista.

Kyselyn suunnittelemisessa itse on huomioitava, että luotettavan ja pätevän kyselyn suunnittelemien vie hyvin paljon aikaa ja resursseja, ja lopputulos ei siltikään ole

välttämättä vertailukelpoinen tai vastaa lopulta mitattavaan ongelmaan. Kehitettyjä kyselyitä ja niiden luotettavuuden parantamista edistää kyselyjen dokumentointi, arviointi sekä yleinen jakaminen esimerkiksi verkossa. All About UX -sivusto on lupaava esimerkki käyttäjäkokemuksen tutkimusmenetelmäpatteriston luomisesta, mutta se vaatii aktiivista yhteistyötä ja laajalti resursseja koko käytettävyyssiantuntijaverkostossa.

Väättäjä & Roto (2010) ovat kehittäneet mobiilin pilottikyselyn käyttäjäkokemuksen tutkimiseen, joka on saanut kiitettävän vastaanoton. Mobiilikyselyt ovat tätä päivää ja tulevaisuutta, ja niiden suunnittelu nostaa esiin jälleen uudenlaisia haasteita. Kuten Väättäjä & Roto raportissaan mainitsevat, on tehtävä vielä paljon tutkimus- ja kehitystyötä, jotta mobiilikyselyt vastaisivat standardikyselyiden luotettavuuteen ja tehokkuuteen. Työ on kuitenkin vaivan arvoista, sillä toimivat mobiilikyselyt keräävät tulevaisuudessa entistä nopeammin ja enemmän määrällistä tietoa käyttäjäkokemuksesta todellisessa käyttöympäristössä.

Vaikka käyttäjäkokemus on tärkeä ja mielenkiintoinen tutkimusala, on myös muistettava, että käyttäjäkokemuksen mittaaminen ei voi olla ainut tuotteen tai järjestelmän hyvyden tai pätevyyden mittari. Voi käydä myös niin, että tuotteesta, josta käyttäjillä on tutkitusti hyvä käyttäjäkokemus, ei tulekaan hittiä markkinoilla, sillä siitä puuttuu jotakin teknisesti olennaista tai sen markkinointi ei ole onnistunut. Tuotekehitysprosessissa on siis tärkeää pitää käyttäjä ja käyttäjäkokemus mielessä, mutta panostettava siihen järkevissä määrin resursseja.

Käyttäjäkokemus ja sen arviointimenetelmät on alana vielä melko tuore, ja elää nyt suurta murrosvaihettaan. Tarvitaan vielä rutkasti lisää tutkimusmateriaalia ja -tuloksia, jotta voidaan määritellä riittävän tarkkoja mittaristoja ja kriteerejä hyvälle käyttäjäkokemukselle. Näin kyselyistäkin saadaan tulevaisuudessa enemmän irti, kun voidaan tarkemmin määritellä, mitä mitataan.

Lähteet

Kirjallisuuskäsitteet

Bangor, A., Kortum, P. & Miller, J. (2009) Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*. Vol. 4:3. S. 114-123. Saatavissa:

http://www.upassoc.org/upa_publications/jus/2009may/JUS_Bangor_May2009.pdf

[Viitattu 24.10.2010]

van Boven, L. & Gilovich, T. (2003) To do or to have? That is the question. *Journal of Personality and Social Psychology*. Vol. 85:6. S.1193-1202. DOI: 10.1037/0022-3514.85.6.1193.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceeding of ACM CHI'88*, Washington, DC. S. 213-218.

Finstad, K. (2006) The System Usability Scale and Non-Native English Speakers. *Journal of Usability Studies*. Vol 1:4. S.185-188. Saatavissa:

http://www.upassoc.org/upa_publications/jus/2006_august/finstad_sus_non_native_speakers.pdf [Viitattu 24.10.2010]

Finstad, K. (2010) The Usability Metric for User Experience. *Interacting with Computers*. S.323-327. DOI: 10.1016/j.intcom.2010.04.004.

Frokaer, E., Hertzum, M & Hornbaek, K. (2000) Measuring Usability: Are Effectiveness, Efficiency and Satisfaction Really Correlated? *Proceedings of the ACM CHI 2000 Conference on Human Factors in Computing Systems*. ACM Press, New York. S 345-352.

Hassenzahl, M. & Tractinsky, N. (2006) User experience – a research agenda. *Behaviour & Information Technology*. ISSN: 1362-3001. Vol 25:2. S.91-97. DOI: 10.1080/01449290500330331.

ISO/IEC (2010). 9241-210 Ergonomics of human-system interaction - Part 210: Human-centered design for interactive systems. ISO/IEC 9241-210:2010(E).

- Jokela, T.** (2010) Navigoi oikein käytettävyyden vesillä. Opas käytettävyysohjattuun vuorovaikutussuunnitteluun. Sähköinen versio Väylä-Yhtiöt Oy:n 2010 kustantamasta painetusta versiosta. ISBN: 978-952-5823-27-1.
- Kujala, S.** (2010) Mitä käyttäjäkokemus on ja miten sitä voidaan tutkia? SIGCHI Finland ry:n Syysseminaari 10.11.2010. Yle, Helsinki.
- Lewis, J.** (1993) IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. Technical Report 54.786. IBM Corporation.S. 17-20. Saatavissa: <http://drjim.0catch.com/usabqtr.pdf> [Viitattu 13.11.2010]
- Lewis, J.** (2001) Psychometric Evaluation of the CSUQ Using Data from Five Years of Usability Studies. IBM Voice Systems. Florida. S. 7-23. Saatavissa: <http://www-course.cs.york.ac.uk/rmh/csuqhist-ral.pdf> [Viitattu 13.11.2010]
- Moore, A., Parr, G., Logan, M., Neely, H., Roesner, D. & Durer, U.** (2001) Developing a European internet and kiosk-based health information system. Journal of Medical Internet Disease. DOI:10.2196/jmir.3.1.e6
- Nielsen, J.** (1993) Usability Engineering. Academic Press. A Harcourt Science and Technology Company. ISSN: 0-12-518406-9.
- Poels, K., IJsselsteijn, W. & de Kort, Y.** (2008). Development of the Kids Game Experience Questionnaire. Poster presented at the Meaningful Play Conference, East Lansing, USA. Saatavissa: [http://www.gamexplab.nl/includes/pages/publications/posters/Poels_2008_Meaningfull Play_poster.pdf](http://www.gamexplab.nl/includes/pages/publications/posters/Poels_2008_Meaningfull_Play_poster.pdf) [Viitattu: 28.11.2010]
- Root, R. & Draper, S.** (1983) Questionnaires as a Software Evaluation Tool. CHI'83 Proceedings. UCSD HMI Project. Institute for Cognitive Science C-015. University of California, San Diego.
- Roto, V. & Rautava, M.** (2008) User Experience and Brand Promise. International Engagability & Design Conference (Idec4), in conjunction with NordiCHI'08 conference. 19.10.2008, Lund, Ruotsi.

Roto, V., Vermeeren, A. & Väänänen-Vainio-Mattila K. (2010a). User eXperience Evaluation Methods. Which Method to Choose? Tutoriaali. SIGCHI Finland ry:n Syysseminaari 10.11.2010. Yle, Helsinki.

Ryu, S. (2005) Development of Usability Questionnaires for Electronic Mobile Products and Decision Making Methods. Faculty of Virginia Polytechnic Institute and State University. Virginia. S. 8-18. Saatavissa: [Viitattu 13.11.2010]

Ryu, S. & Smith-Jackson, T. (2006) Reliability and Validity of the Mobile Phone Usability Questionnaire (MPUQ). Journal of Usability Studies. Vol 2:1. S. 39-53. Saatavissa: http://www.upassoc.org/upa_publications/jus/2006_november/ryu_smith-jackson_mobile_phone_questionnaire.pdf [Viitattu 27.11.2010]

Tullis, T. & Stetson, J. (2004) A Comparison of Questionnaires for Assessing Website Usability. UPA 2004 Presentation. Human Interface Design Department, Dicelity Center for Applied Technology, Boston. Saatavissa: <http://home.comcast.net/~tomtullis/publications/UPA2004TullisStetson.pdf> [Viitattu 24.10.2010]

Usoh, M., Catena, E., Arman, S. & Slater, M. (2000) Using Presence Questionnaires in Reality. MIT Press Journal. Vol 9:5. S. 497-503. DOI:10.1162/105474600566989

Vanhala, T. (2005) Kyselylomakkeet käytettävyytutkimuksessa. Ovaska, S., Aula, A & Marjaranta, P. (toim). Käytettävyytutkimuksen menetelmät. Tampereen yliopisto, Tietojenkäsittelytieteiden laitos B-2005-1. S. 17-36.

Vuolle, M., Kallio, T., Kulju, M., Tiainen, M., Vainio, T., & Wigelius, H. (2008) Developing a questionnaire for evaluating mobile business service experience. The proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services. MobileHCI 2008. S.53-62.

Väättäjä, H. & Roto, V. (2009) Questionnaires in User Experience Evaluation. UXEM Workshop in conjunction with INTERACT 2009. Saatavissa: <http://delivery.acm.org/10.1145/1760000/1753985/p3361-vaataja.pdf?key1=1753985&key2=1667137821&coll=portal&dl=ACM,ACM&CFID=106126253&CFTOKEN=36014902> [Viitattu 14.11.2010]

Vääätäjä, H. & Roto, V. (2010) Mobile Questionnaires for User Experience Evaluation. CHI 2010: Work-in-Progress. ACM 978-1-60558-930-5/10/04.

Wechsung, I & Naumann, A. (2008) Evaluation Methods for Multimodal Systems: A Comparison of Standardized Usability Questionnaires. André, E. ja muut (toim.). Deutsche Telekom Laboratories, TU Berlin. S. 276-284. Saatavissa: <http://www.springerlink.com/content/xj4478p37740l55j/> [Viitattu 24.10.2010]

Witmer, B. & Singer, J. (1998) Measuring Presence in Virtual Environments: A Presence Questionnaire. U.S. Army Research Institute for the Behavioral and Social Sciences. Presence, Vol 7:3. S. 225–240. Saatavissa: <http://mitpress.mit.edu/journals/PRES/ps00734.pdf> [Viitattu 28.11.2010]

WWW-lähteet

Brooke, J. (1996) SUS – A quick and dirty usability scale. Redhatch Consulting Ltd. Saatavissa: <http://hell.meiert.org/core/pdf/sus.pdf?/> [Viitattu 27.11.2010]

Hassenzahl, M., Burmester, M. & Koller, F. (2010) AttrakDiff™ - A Service Provided by User Interface Design GmbH. Internet-sivusto. Saatavissa: <http://www.attrakdiff.de/en/Home/> [Viitattu 14.11.2010]

Kirakowski, J. (1994) What is SUMI?. Saatavissa: <http://sumi.ucc.ie/whatis.html>. [Viitattu 26.11.2010]

Kirakowski, J. (2000) Questionnaires in Usability Engineering – A List of Frequently Asked Questions (3rd Ed.). Human Factors Research Group, Cork, Ireland. Saatavissa: <http://www.ucc.ie/hfrg/resources/qfaq1.html> [Viitattu 24.10.2010]

Lund, A. (1998) USE Questionnaire Resource Page. Introduction to the USE Questionnaire. Saatavissa: <http://usesurvey.com/> [Viitattu 24.10.2010]

Norman, K & Shneiderman, B. & Harper, B. QUIS: The Questionnaire for User Interaction Satisfaction. Human-Computer Interaction Lab. Saatavilla: <http://www.cs.umd.edu/hcil/quis/> [Viitattu 25.10.2010]

Roto, V., Lee, M., Pihkala, K., Castro, B., Vermeeren, A., Law, E., Väänänen-Vainio-Mattila, K., Hoonhout, J. & Obrist, M. (2010b). All About UX - Information

for user experience professionals. Internet-sivusto. Saatavissa:
<http://www.allaboutux.org/>. [Viitattu 14.11.2010]

UsabilityNet. (2006) Questionnaire resources. Saatavissa:
http://www.usabilitynet.org/tools/r_questionnaire.htm [Viitattu 24.10.2010]

WAMMI. (2010) Internet-sivusto. Saatavissa: <http://www.wammi.com/index.html>.
[Viitattu 14.11.2010]