



OULUN YLIOPISTO
UNIVERSITY of OULU

Käytettävyyden arvioinnin menetelmät pelikehityksessä

Oulun yliopisto
Tietojenkäsittelytieteiden laitos
LuK-tutkielma
Kimmo Martimo
26.4.2018

Tiivistelmä

Videopeliteollisuus on kasvanut nopeasti maailman suurimpien viihdealojen joukkoon. Pelien käytettävyys on pelaajien mielestä yksi tärkeimmistä pelin valintaan vaikuttavista tekijöistä, mutta tästä huolimatta pelien käytettävyystutkimus on lähtenyt kunnolla käyntiin vasta 2000-luvun puolella. Tutkimusala on niin nuori, ettei käsitteistäkään olla vielä päästy yhteisymmärrykseen. Tutkimusta tarvitaan, sillä pelialan sisällä kehitetty ja laajasti käytetty menetelmä, pelitestausta, on kallis ja sitä voi suorittaa vasta myöhemmissä vaiheissa pelinkehitystä.

Tässä tutkielmassa selvitettiin kirjallisuuskatsauksen avulla millaisia menetelmiä on kehitetty viihdealojen pelien käytettävyyden arviointiin sekä millaisia tuloksia niitä soveltamalla on saatu. Löydettiin, että pelikäytettävyyden tutkimus on keskittynyt hyötyohjelmien käytettävyydessä paljon käytetyn heuristisen arvioinnin muokkaamiseen peleihin soveltuvaksi. Heuristinen arviointi sopii näennäisesti hyvin peleihin, sillä se on halpa menetelmä ja sitä voi suorittaa aikaisessa vaiheessa pelinkehitystä. Kuitenkin kaikki kirjallisuuskatsauksella löydetyt menetelmät olivat kehitetty epäilyttävällä tavalla ja testattu puutteellisesti. Tutkitut menetelmät eivät ole siirtyneet käyttöön pelialalla ja selvästi tarvitaan lisätutkimusta, jotta voitaisiin kehittää parempia menetelmiä.

Avainsanat

käytettävyys, käytettävyysarviointi, heuristinen arviointi, videopelit, pelit

Sisällys

Tiivistelmä	2
Sisällys	3
1. Johdanto.....	4
2. Käytettävyys käsitteenä	5
2.1 Hyötyohjelmien käytettävyys	5
2.2 Pelien käytettävyyden konteksti	6
2.3 Käytettävyys ja pelattavuus	6
2.4 Pelien tyyllilajien vaikutus pelien käytettävyyden arviointiin.....	7
3. Tutkimusmenetelmät	8
4. Tutkimus.....	9
4.1 Pelitestausta	9
4.2 Heuristinen arviointi	9
4.2.1 Heuristic Evaluation of Playability (HEP)	10
4.2.2 Korhosen ja Koiviston mobiilipeliheuristiikat	12
4.2.3 Korhosen ja Koiviston usean pelaajan mobiilipeliheuristiikat.....	13
4.2.4 Pinellen ja muiden heuristiikat	14
4.2.5 Networked Game Heuristics (NGH)	16
4.2.6 PLAY-heuristiikat	17
4.2.7 GAP-heuristiikat.....	18
4.2.8 Hochleitnerin ja muiden heuristiikat	19
5. Keskustelu	21
6. Yhteenveto.....	23
Kirjallisuusviitteet.....	25

1. Johdanto

Videopeliteollisuus on liikevaihdolla mitattuna kasvanut nopeasti maailman suurimpien viihdealojen joukkoon. Pelialan maailmanlaajuisen liikevaihdon on arvioitu olleen 67 miljardia dollaria vuonna 2012. Vertailun vuoksi esimerkiksi elokuva-alan maailmanlaajuinen liikevaihto on arvioitu olleen 85 miljardia dollaria vuonna 2011. (Marchand & Hennig-Thurau, 2013.) Pelialan menestyksellä on siis suuri taloudellinen arvo.

Pelien käytettävyys on tärkeää pelaajille. Rajanen ja Marghescu (2006) tutkivat mitkä seikat vaikuttavat eniten pelaajien pelivalintoihin sekä mitä asioita pelaajat kokevat häiritseviksi pelaamisen aikana. Tutkimuksen mukaan käyttöliittymä ja pelattavuus ovat yksi tärkeimpiä asioita pelien valinnassa ja huono käyttöliittymä ja huono pelattavuus ovat yksi häiritsevimmistä asioista pelaamisen aikana. (Rajanen & Marghescu, 2006.) Pelin huonolla käytettävyydellä voi siis olla heikentävä vaikutus pelin kokonaislaatuun ja menestykseen (Pinelle, Wong, & Stach, 2008a; Rajanen & Rajanen, 2017).

Tästä huolimatta, käytettävyytutkimus ja pelit ovat kohdanneet toisensa merkittävässä määrin vasta 2000-luvun alussa. Peleillä ja perinteisillä hyötyohjelmilla onkin paljon annettavaa toisilleen mitä tulee käytettävyyteen. (Jørgensen, 2004.)

Tämän tutkielman tarkoitus on selvittää kirjallisuuskatsauksen avulla:

1. Millaisia menetelmiä on kehitetty viihdeellisten pelien käytettävyyden arviointiin.
2. Millaisia tuloksia menetelmiä soveltamalla on saatu.

Luvussa 2 esitellään käytettävyyden käsite sekä hyötyohjelmien, että pelien kontekstissa ja vertaillaan näiden ominaisuuksia.

Luvussa 3 kerrotaan tutkielmaan valittujen käytettävyyshälymenetelmien haku- ja valintaprosessista.

Luvussa 4 käydään läpi pelikäytettävyyden arviointiin kehitettyjä menetelmiä sekä niillä saatuja tuloksia.

Luvussa 5 käsitellään löydettyjä menetelmiä kokonaisuutena, arvioidaan niiden soveltuvuutta pelikäytettävyyden arviointiin sekä pohditaan niiden merkitystä pelialalle.

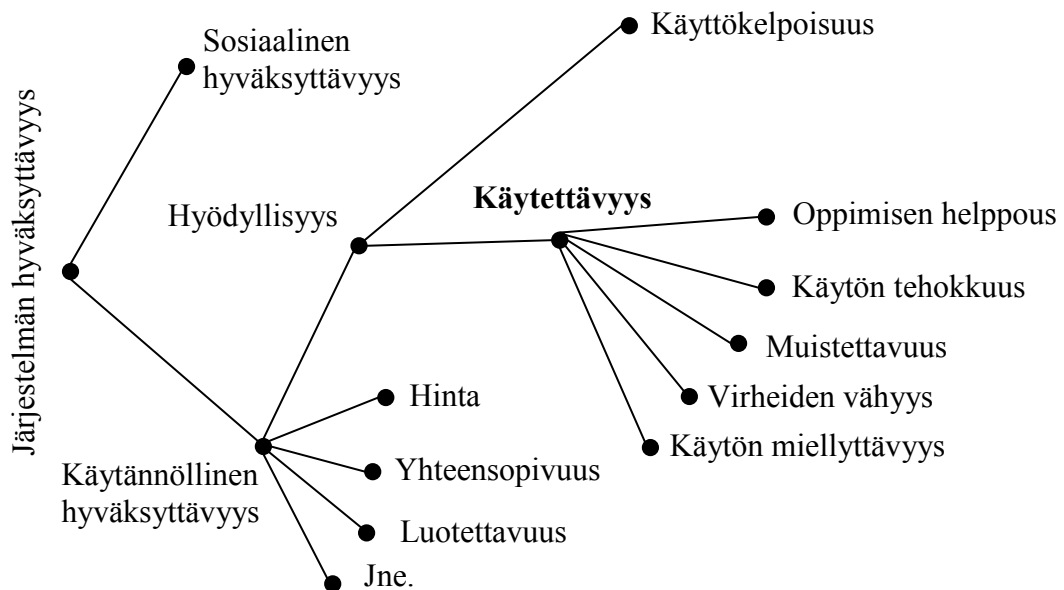
Luvussa 6 esitetään yhteenveto tutkielmasta ja esitetään suosituksia tulevalle tutkimukselle. Luku sisältää myös omaa pohdintaa aiheesta.

2. Käytettävyys käsitteenä

Käytettävyttä on tutkittu pitkään eri nimillä ja erilaisilla huomion kohteilla. 1940-luvulla tutkittiin ergonomiaa, jossa huomion kohteena oli ihmisten ja koneiden välinen vuorovaikutus. Tutkimuskohteena oli esimerkiksi miten koneen mittarit ja ohjaimet pitäisi asettaa. 1960-luvun lopussa aloitettiin tietojärjestelmien tutkiminen. Tietojärjestelmätieteessä tutkittiin periaatteessa samoja asioita kuin ergonomiassa, mutta kohde oli siirtynyt tietokoneisiin. Kuitenkin tietojärjestelmätieteessä tutkittiin käyttöliittymää laajempaa kokonaisuutta: toimintaa, työtä ja organisaatiota jossa se tapahtuu. Nykymuotoinen käyttöliittymätutkimus ja käytettävyystutkimus syntyi 1980-luvulla. (Oulasvirta, 2011.)

2.1 Hyötyohjelmien käytettävyys

Käytettävyden käsite hyötyohjelmille kytkeytyy laajempaan kokonaisuuteen, järjestelmän hyväksyttävyyteen. Järjestelmän hyväksyttävyyttä käsitellään tässä luvussa vain siltä osin kuin se on vaikuttanut käytettävyden lähtökohtiin.



Kuvio 1. Järjestelmän hyväksyttävyyteen liittyvät ominaisuudet (Nielsen, 1993, p. 25).

Kuviossa 1 nähdään miten käytettävyys liittyy laajempaan kokonaisuuteen, järjestelmän hyväksyttävyyteen. Laajemman kokonaisuuden vaikutus käytettävyteen ilmenee siten, että käytettävydessä lähdetään liikkeelle siitä oletuksesta, että käyttäjä haluaa saada järjestelmän avulla tehdyksi jonkin tehtävän (Nielsen, 1993, p. 24).

Nielsen määrittää käytettävyden koostuvan viidestä osa-alueesta:

- Oppimisen helppous. Järjestelmän tulisi olla helppo oppia, jotta käyttäjä pääsisi mahdollisimman nopeasti tekemään töitä järjestelmän avulla.

- Käytön tehokkuus. Järjestelmän tulisi mahdollistaa korkea tuottavuuden taso käyttäjän opittua järjestelmän käytön.
- Muistettavuus. Järjestelmän tulisi olla helposti muistettavissa, jotta käyttäjä voisi jatkaa järjestelmän käyttöä tauon jälkeen opettelematta kaikkea uudelleen.
- Virheiden vähyys. Järjestelmän pitäisi olla sellainen, että käyttäjät tekisivät mahdollisimman vähän virheitä ja tehdyistä virheistä voidaan palautua helposti. Tuhoisien virheiden tekeminen ei pitäisi olla mahdollista.
- Käytön miellyttävyys. Järjestelmän pitäisi olla miellyttävä käyttää, jolloin käyttäjät ovat tyytyväisiä sen käyttöön.

Hyötyohjelmien käytettävyydessä pyritään maksimoimaan oppimisen helppous ja muistettavuus pyrkimällä yhdenmukaisuuteen mahdollisimman suuressa määrässä käyttöliittymäratkaisuja. Mikäli käyttäjä on jo oppinut toisia hyötyohjelmia käyttäessään käyttämään tietynlaista käyttöliittymäratkaisua, sitä ei tarvitse käyttäjälle opettaa.

2.2 Pelien käytettävyyden konteksti

Käytettävyyden käsite pelien kontekstissa eroaa olennaisesti hyötyohjelmien käytettävyydestä. Hyötyohjelmia käytetään pääasiassa tuottavaan työhön ja pelejä käytetään viihtymiseen (Pagulayan ja muut, 2008). Tällöin voidaan ajatella, että hyötyohjelmien käytettävyyden lähtökohtaan vaihdetaan tehtävän tilalle viihtyminen. Toisin kuin jotain tehtävää, viihtymistä ei voida ajatella asiana joka tulee valmiiksi tai tehdyksi. Peliä voidaan pelata hyvin pitkään pääsemättä mihinkään erityiseen tavoitteeseen, mutta samalla pelaaja on viihtynyt koko ajan. Hyötyohjelmien käytettävyyden ideaali käyttäjän pääsemisestä tavoitteeseensa mahdollisimman nopeasti ei siis päde pelien kohdalla.

Hyötyohjelmien käytössä käyttäjän tekemien virheiden määrä halutaan mahdollisimman vähäiseksi. Peleissä taas pelaajan odotetaan tekevän virheitä, koska pelit suunnitellaan haastamaan pelaajat kehittämään uusia taitoja pelissä edetäkseen (Pinelle ja muut, 2008a). Koska haaste kuuluu peleihin niin olennaisena osana, pelit eivät voi olla mahdollisimman helppoja. Hyötyohjelmien suunnittelun tavoitteena on ohjelma jota on helppo oppia ja käyttää, mutta pelien suunnittelun tavoitteena on peli jonka pelaamista on helppo oppia, mutta jonka pelaamisessa on vaikea kehittyä mestariksi (Malone, 1982). Korhosen (2010) mukaan pelien käytettävyyssarvioinnin tarkoituksena onkin vähentää saavutusten tiellä olevien esteiden sijaan hauskuuden tiellä olevia esteitä.

Pelit eroavat hyötyohjelmista myös käyttöliittymän innovatiivisuudessa. Pelikehittäjillä on suuri motivaatio tehdä innovatiivisia suunnitteluratkaisuja peleihinsä, jotta ne erottuisivat muista peleistä. Tämän vuoksi myös pelien käyttöliittymät eivät ole samalla tavalla yhteneväisiä kuten hyötyohjelmat pyrkivät olemaan. Tämä innovatiivisuus on johtanut siihen, että eri pelien välillä on hyvin suuria eroja käyttöliittymissä. (Pinelle, Wong, & Stach, 2008b.)

2.3 Käytettävyys ja pelattavuus

Pelien käytettävyyden arviointiin liittyvistä pelien ominaisuuksista ja käsitteistä ei olla päästy tutkijoiden piirissä yhteisymmärrykseen. Koska pelaajan tavoite pelien pelaamisessa on viihtyminen ja pelin viihdyttävyyteen vaikuttavat hyvin monet asiat, pelikäytettävyyden määrittäminen on hankalaa.

Desurviren, Caplanin ja Tothin (2004) mukaan pelin käytettävyyden arvioinnissa pitää mennä pitemmälle kuin pelkän pelin käyttöliittymän käytettävyyden arviointiin. Heidän

mukaansa koko pelikokemus täytyy ottaa arvioinnin piiriin. Näihin lukeutuvat myös pelin mekaniikat ja tarina. Näitä seikkoja he kutsuvat pelin pelattavuuden arvioinniksi. Pinelle ja muut (2008a) kritisoivat tätä määritelmää liian laajaksi sekä liiaksi pelin nautittavuuteen keskittyväksi. Heidän mukaansa pelikäytettävyyden määritelmä on "pelaajan kyvyn aste pelin oppimiseen, ohjaamiseen sekä ymmärtämiseen". Pelin viihdyttävyyden, osallisuus sekä tarina ovat asioita, jotka ovat tiukasti yhteydessä taiteellisiin seikkoihin kuten käsikirjoitukseen ja musiikkiin sekä teknisiin seikkoihin kuten graafiseen laatuun ja suorituskykyongelmiin.

Federoffin (2002, p. 20) mielestä pelien käytettävyyttä ei voida arvioida ottamatta kontekstia huomioon. Hänen mielestään käytettävyyden arviointiin ei riitä pelkästään käyttöliittymän ja pelimekaniikan arviointi, vaan käytettävyyden arviointiin tulee sisällyttää myös pelin pelattavuus. Hänen mukaansa kaikki nämä kolme seikkaa vaikuttavat pelaajan tyytyväisyyteen joka on kaikkein tärkein mittari pelien käytettävyyden arviointiin.

2.4 Pelien tyyllilajien vaikutus pelien käytettävyyden arviointiin

Pinelle ja muut (2008b) esittävät, että vaikka yksittäisten pelien väliset erot käyttöliittymissä ovat hyvin suuret, eri tyyllilajien sisällä olevilla peleillä on havaittavissa samanlaisuuksia. Tämän johdosta he ehdottavatkin, että tyyllilajeilla voidaan jakaa pelit luokkiin joiden avulla voidaan syventää tietämystä käytettävyyden menetelmien soveltamisesta eri tyyllilajin peleihin.

Tutkiakseen tätä he valitsivat kuusi pelien tyyllilajia ja kehittivät kaksitoista yleisluontoista käytettävyysohjelmaluokkaa joita peleissä yleensä on. Sitten he hakivat jokaiselle kuudelle tyyllilajille 18 ammattimaisten peliarvostelijoiden tekemää peliarvostelua. He etsivät arvosteluista merkkejä käytettävyysohjelmissa ja lajittelivat ne keksimiensä käytettävyysohjelmaluokkien mukaan. Tutkimuksen tuloksena löydettiin eri tyyllilajeille ominaisia käytettävyysohjelmaluokkia. Tutkijoiden mukaan pelien tyyllilajien huomioiminen on hyödyllistä, mutta ei täysin ongelmatonta. Liiallinen keskittyminen useimmin esiintyviin ongelmiin saattaa johtaa vähemmän esiintyvien, mutta silti merkittävien, ongelmien huomiotta jättämiseen. Lisäksi monet pelit eivät sovi yhteen tyyllilajiin, vaan ne sisältävät elementtejä useista tyyllilajeista, jolloin yhden tyyllilajin ongelmiin keskittyminen johtaa ongelmien huomiotta jättämiseen. (Pinelle ja muut, 2008b.)

Livingston, Mandryk ja Stanley (2010) kehittivät Pinellen ja muiden (2008b) tulosten pohjalta menetelmän, jonka avulla pelistä voidaan tehdä arvostelijankestävä (engl. *criticproof*). Arvostelijankestävyys perustuu peleistä löydettyjen käytettävyysohjelmissa asetamiseen tärkeysjärjestykseen arvioijan subjektiivisen vakavuusarvion, pelin tyyllilajin sekä ongelmien esiintymistiheyden peliarvosteluissa pohjalta.

Menetelmässä määritellään aluksi pelin ja sen alaosien (esim. minipelit) osuus tyyllilajeihin prosenttiosuutena. Näistä osuuksista ja Pinellen ja muiden (2008b) tuloksista saadaan laskemalla käytettävyysohjelmaluokille uusi pelikohtainen painokerroin. Näiden lisäksi käytetään arvioijan subjektiiviseen käytettävyysohjelmissa vakavuusarviota taulukkoa, jolla saadaan vakavuusarviota vastaava numerokerroin. Lopulta pelikohtainen painokertoimen ja arvioijan vakavuusarvion numerokertoimen avulla lasketaan lopullinen tärkeysnumero jonka avulla löydetty käytettävyysohjelmat voidaan asettaa tärkeysjärjestykseen.

3. Tutkimusmenetelmät

Käsiteltävät pelikäytettävyydestutkimukset haettiin Oulun yliopiston tarjoamalla Nelli-portaalilla kaikista tietojenkäsittelyalan tietokannoista sekä Google Scholarilla. Hakusanoina käytettiin termien "game", "games", "videogame", "videogames", "computer game", "computer games", "usability" sekä "playability" yhdistelmiä ja muunnoksia. Löydetyistä tutkimuksista valittiin ne, jotka täyttivät seuraavat kriteerit:

- Tutkimus vaikutti nopealla katsauksella asialliselta.
- Tutkimus oli julkaistu kohtuullisen tasokkaassa jurnaalissa. Tämän rajauksen poistamat tutkimukset olivat yleensä myös nopealla vilkaisulla epäilyttävän oloisia.
- Tutkimuksen kohteena oli viihdekäyttöön tarkoitettut pelit. Tällä rajattiin pois ns. hyötypelit (engl. *serious game*) kuten oppimispelit ja mainospelit.
- Tutkimuksessa esiteltiin pelikäytettävyyden arviointiin joko uusi menetelmä tai muokkaus vanhaan menetelmään.
- Tutkimus oli tarpeeksi uusi. Tämän kriteerin tarkoituksena oli poistaa joukosta sellaiset menetelmät, jotka on kehitetty nykyisin erittäin epäsuosituille pelityypeille kuten esimerkiksi pelihallipelit (engl. *arcade game*).
- Tutkimuksessa esitelty menetelmä sopi tai väitettiin sopivan käytettävyyden arviointiin pelin kehitysvaiheessa.
- Tutkimuksessa esiteltyä menetelmää oli testattu empiirisesti.

Näiden lisäksi mukaan otettiin myös peliteollisuuden sisällä kehitetty menetelmä, pelitestaus. Pelitestauksen mukanaolo on tärkeää, sillä monien pelikäytettävyyden arviointimenetelmien kehitys on lähtenyt liikkeelle pelitestauksen puutteista. Pelitestausta käytetään myös usein menetelmien testauksen vertailukohteena.

4. Tutkimus

Hyötyohjelmien käytettävyyden arviointiin on kehitetty useita menetelmiä kuten ääneen ajattelu, fokusryhmien käyttö, haastattelut, heuristiikat (myös nimellä expert review) ja kognitiivinen läpikäynti (Genise, 2002). Kuitenkin monet näistä eivät kelpaa pelien arviointiin, sillä ne perustuvat etukäteen määriteltyihin tehtävänkulkuihin tai ovat liian keskittyneitä hyötyohjelmatyyppeihin käyttöliittymiin (Pinelle ja muut, 2008a).

4.1 Pelitestausta

Pelitestausta (engl. *playtesting*) on pelialalla perinteisesti käytetty menetelmä. Pelitestauksessa testauksen kohteena on pelattava prototyyppi tai lähes valmis peli (Pinelle ja muut, 2008a). Testaajat ovat tavallisia pelaajia, eivät käytettävyyssiantuntijoita. Testaajat pelaavat peliä tai prototyyppiä ja heidän toimintaansa seurataan. Testausta seurattaessa huomataan suurimmat ongelmat, kuten suurimmat väärinkäsitykset ja asiat joita testaajat tekevät väärin. Lisäksi testaajia pyydetään ilmaisemaan mikäli jokin asia häiritsee heitä, jolloin saadaan tietoon piilevämmät ongelmat.

Pelitestauksessa on paljon huonoja ominaisuuksia. Koska pelitestausta vaatii vähintään sellaisen prototyypin jota voi pystyä pelaamaan tai lähes valmiin pelin, pelitestausta voi suorittaa vasta kohtuullisen myöhään pelinkehitysprosessissa. Näin ollen testauksessa löydettyjen ongelmien korjaus on kalliimpaa ja hankalampaa. (Korhonen, Paavilainen, & Saarenpää, 2009.) Lisäksi pelitestausta vaatii paljon testaajia ja jokaista testaajaa varten tarkkailijan, mikä tekee itse testauksesta kallista (Federoff, 2002, p. 21).

4.2 Heuristinen arviointi

Heuristinen arviointi on käytettävyyden arvioinnin menetelmä, jossa käyttäjien seuraamisen sijaan käytetään käytettävyyssarvioijia. Käytettävyyssarvioijilla on käytössään heuristiikat. Heuristiikat ovat tuoteluokalle yhteinen lista tunnistettuja yleisluontoisia käytettävyyssperiaatteita. Heuristista arviointia suorittavat arvioijat käyvät läpi käyttöliittymää ja tarkastavat käyttöliittymän olevan heuristiikkojen mukainen. Löydetyt ongelmat kirjataan ylös. Heuristiikoilla löydetyistä ongelmista ei voi suoraan johtaa tietoa siitä, mitä ongelman korjaamiseksi tulisi tehdä. (Nielsen, 1994, p. 26.)

Heuristinen arviointi on halpa käytettävyyden arvioinnin menetelmä. Tämä johtuu siitä, että se on yksinkertaisuutensa vuoksi helppo opetella sekä nopea toteuttaa. Helppoutensa vuoksi heuristisen arvioinnin tekemisen kynnyks on matala. Arvioinnin kohteena olevan käyttöliittymän ei tarvitse olla toimiva, sillä heuristista arviointia voidaan suorittaa esimerkiksi paperille piirretyille käyttöliittymähahmotelmille. Heuristisen arvioinnin tuloksia voi parantaa lisäämällä arvioijien määrää. Heuristisella arvioinnilla ei ole kuitenkaan mahdollista löytää kaikkia käytettävyyso ongelmia. (Nielsen, 1994, p. 25.)

Pinellen ja muiden (2008a) mukaan heuristinen arviointi sopii pelien käytettävyyden arviointiin muita yleisiä hyötyohjelmien käytettävyyssmenetelmiä paremmin sillä heuristinen arviointi ei tee oletuksia tehtävien rakenteesta ja on muokattavissa

erikoisalojen käyttöön. Koska pelit eroavat niin suuresti hyötyohjelmista, pelien arviointiin tarvitaan erityisesti pelien arviointiin kehitetyt heuristiikat (Korhonen ja muut, 2009). Heuristisella arvioinnilla on pelikontekstissa erikoinen ominaisuus, sillä heuristiikan rikkominen ei välttämättä aina ole merkki käytettävyysongelmasta. Pelikehittäjät voivat tietoisesti tehdä esimerkiksi viholliset ja ystävät hankalaksi erottaa toisistaan, jotta peli on haastavampi. (Pinelle ja muut, 2009.)

Heuristiikat ovatkin tutkituimpia pelikäytettävyyden arviointikeinoja. Monet tutkijat ovat kehittäneet omia heuristiikkojaan tai muokanneet Nielsenin hyötyohjelmille tarkoitettuja heuristiikkoja soveltumaan enemmän peleihin.

Seuraavissa kappaleissa esitellään löydetty tutkimukset kronologisessa järjestyksessä, sillä monet heuristiikoista on kehitetty aiemmissa heuristiikoissa koettujen ongelmien johdosta.

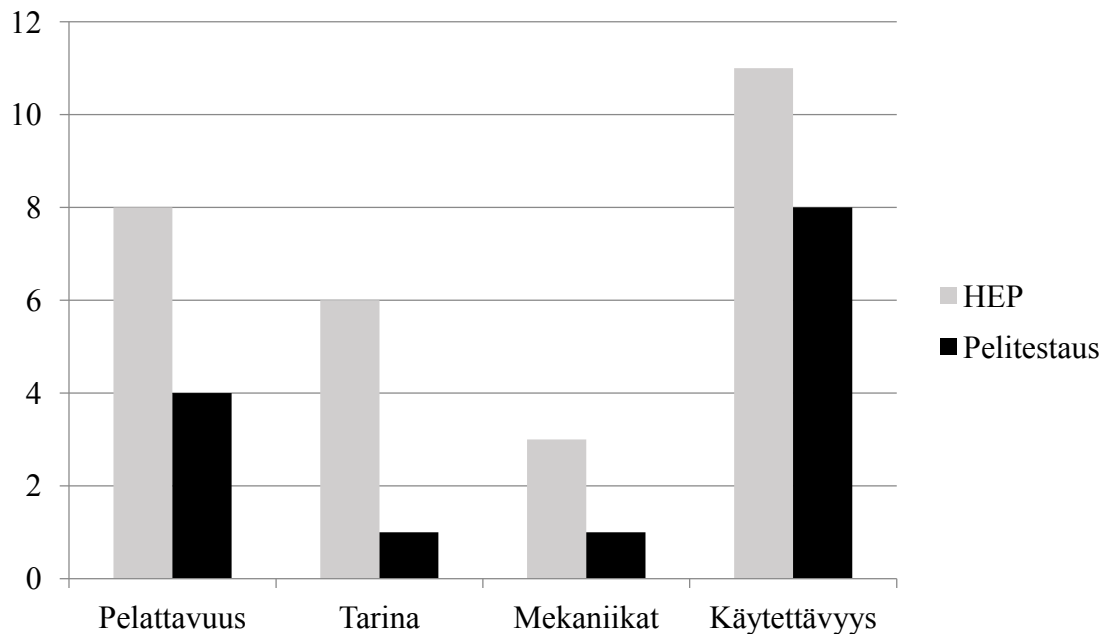
4.2.1 Heuristic Evaluation of Playability (HEP)

Desurvire ja muut (2004) kehittivät heuristiikkakokoelman jota he kutsuvat nimellä Heuristic Evaluation of Playability (HEP). HEP-menetelmän heuristiikat ovat jaettu neljään kategoriaan:

- Pelattavuus: pelin voittamiseksi päihitettävät haasteet.
- Tarina: pelin juoni ja hahmokehitys.
- Mekaniikat: säännöt jotka luovat rakenteen vuorovaikutukseen ympäristön kanssa.
- Käytettävyys: käyttöliittymä ja peliohjaimet.

Pelattavuus-kategoriassa on 16 heuristiikkaa, tarina-kategoriassa 8, mekaniikka-kategoriassa 7 ja käytettävyys-kategoriassa 12. Yhteensä heuristiikkoja on 43 kappaletta.

Kehitettyjä heuristiikkoja testattiin pelinkehitysprosessissa aikaisessa vaiheessa olevalla pelillä. Pelistä oli tehty prototyyppi, joka koostui navigoitavista ruudunkaappauskuvista. HEP-menetelmää verrattiin samalle peliprototyypille tehtyyn pelitestaukseen.



Kuvio 2. Vertailu pelitestauksessa ja HEP-menetelmällä löydettyjen ongelmien määrästä (Desurvire ja muut, 2004).

Tuloksena prototyypistä löytyi yhteensä 26 uniikkia ongelmaa. Näistä ongelmista 12 löytyi sekä HEP-menetelmällä, että pelitestauksella. HEP-menetelmällä löydettiin 14 ongelmaa joita ei löydetty pelitestauksella ja pelitestauksella löydettiin kaksi ongelmaa joita ei löydetty HEP-menetelmällä. HEP-menetelmällä ja pelitestauksella löydettiin erilaisia ongelmia (kuvio 2). Pelitestauksella löydettiin pelikohtaisia ongelmia kuten haasteeseen ja terminologiaan liittyviä ongelmia kun taas HEP-menetelmän avulla löydettiin yleisempiä ongelmia.

Pinelle ja muut (2008a) kritisoivat HEP-menetelmää liiallisesta keskittymisestä nautinnollisiin seikkoihin jolloin HEP-menetelmän heuristiikat eivät sovellu hyvin pelien käytettävyyden arviointiin. Paavilaisen (2010) mukaan HEP-menetelmää ei vahvistettu tarpeeksi hyvin, sillä testauksessa oli vain yksi arvioija käyttämässä HEP-menetelmää ja vain neljä testaajaa pelitestauksessa. Lisäksi HEP-menetelmää oli testattu vain yhdellä ei-pelattavalla prototyypillä. Näiden ongelmien johdosta Paavilainen kritisoikin tutkijoiden lausuntoa HEP-menetelmän hyödyllisyydestä arveluttavaksi. Paavilaisen mukaan osa heuristiikoista on liian epäselviä, osa liian subjektiivisia ja osa sopivampia suunnitteluvaatimuksiksi kuin arviointiin.

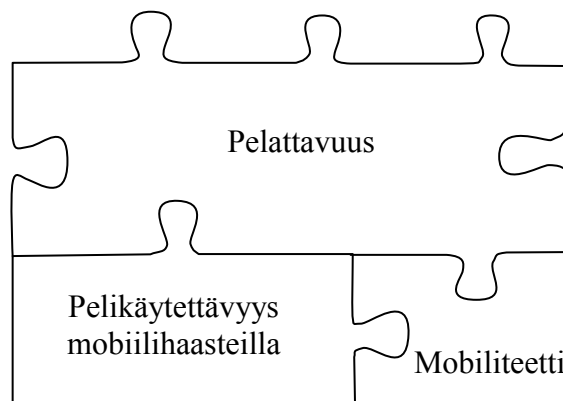
Korhonen ja muut (2009) testasivat HEP-menetelmää käytännössä toteuttamalla neljällä arvioijalla arvion valmiista mobiilipelistä HEP-menetelmää käyttäen. Arvioijat valittivat HEP-menetelmän sisältävä liian suuren määrän heuristiikkoja ja osan heuristiikoista olleen liian vaikeita käyttää arviointiin. Arvioijien mukaan HEP-menetelmän kategorioista ei ollut apua, sillä osa heuristiikoista koettiin olevan väärässä kategoriassa tai liian päällekkäisiä toisten heuristiikkojen kanssa. Lisäksi osa heuristiikkojen kuvauksista olivat arvioijien mielestä liian yksityiskohtaisia ja osa liian ylimalkaisia. Ylipäätään arvioijien mielestä oli hankala arvioida pelin nautittavuuteen liittyviä seikkoja arvioinnin aikana, sillä arviointitilanne esti pelistä nauttimisen.

4.2.2 Korhosen ja Koiviston mobiilipeliheuristiikat

Korhonen ja Koivisto (2006) kehittivät älypuhelimella pelattaville mobiilipeleille sopivat heuristiikat. Korhosen ja Koiviston mukaan mobiilipelit tarvitsevat omat heuristiikkansa mobiiliuden tuottamien erityisseikkojen vuoksi. Tällaisia seikkoja ovat esimerkiksi pelaamisen keskeytyminen esimerkiksi puhelimen soimisen tai bussin saapumisen vuoksi. Lisäksi mobiilipelejä pelataan myös ulkona, jolloin ympäristö vaikuttaa suuresti peliolosuhteisiin. Pelejä myös pelataan yleensä vain lyhyen hetken. (Korhonen & Koivisto, 2006.)

Tutkijoiden mukaan aikaisemmin kehitetyt heuristiikat, kuten Desurviren ja muiden (2004) kehittämä HEP-menetelmä, eivät sovi mobiilipeleihin. Aikaisemmat heuristiikat eivät ottaneet kantaa mobiiliuden mukana tuomiin seikkoihin, ne olivat huonosti selitettyjä ja päällekkäisten heuristiikkojen vuoksi moniselitteisiä.

Tutkijat aloittivat heuristiikkojen kehittämisen tekemisen tutkimalla älypuhelimia ja niiden käyttöä, Nielsenin hyötyohjelmille kehittämiä heuristiikkoja sekä pelinkehityksen yleisiä ohjenuoria. Näiden pohjalta he kehittivät alustavat yksitoista heuristiikkaa. Näitä heuristiikkoja testattiin prototyyppiasteella olevalle mobiilipelille neljällä arvioijalla. Pelistä löytyi 61 käytettävyyssongelmaa, mutta niistä 16:lle ei löydetty vastaavaa heuristiikkaa. Näitä ongelmia analysoimalla kehitettiin 18 uutta heuristiikkaa, jotka yhdistettiin jo olemassa oleviin heuristiikkoihin. Sen jälkeen koko 29 heuristiikan kokoelmaa arvioitiin pelinkehittäjien kanssa ja heuristiikkoja muokattiin pelinkehittäjien kommenttien pohjalta.



Kuvio 3. Heuristiikkamoduulit (Korhonen & Koivisto, 2006).

Lopulliset heuristiikat jakautuvat kolmeen osaan: pelikäytettävyyteen, mobiliteettiin ja pelattavuuteen. Pelikäytettävyyteen sisältyi 12 heuristiikkaa, mobiliteettiin 3 ja pelattavuuteen 14. Heuristiikkojen on tarkoitus olla modulaarisia, jolloin niitä voidaan käyttää erikseen (kuvio 3). Pelikäytettävyys- ja pelattavuusmoduulit ovatkin tarkoitettu yhteisiksi kaikille peleille eikä pelkästään mobiilipelien käytettävyyden arviointiin. Heuristiikkoihin voidaan myös kytkeä muita moduuleja, kuten esimerkiksi monen pelaajan peleille kehitettyjä heuristiikkoja. Lopullisia heuristiikkoja arvioitiin viidellä prototyyppiasteella olevalla pelillä ja 2-4 arvioijalla. Pelit edustivat monia erilaisia tyyllilajeja ja kohderyhmiä. Mukana oli sekä yhden, että useamman pelaajan pelejä.

Lopullisilla heuristiikoilla löydettiin 235 käytettävyyssongelmaa. Arvioinnin aikana löydettiin helposti pelikäytettävyyteen ja mobiliteettiin kuuluvien heuristiikkojen rikkomuksia, mutta pelattavuuteen liittyvä ongelmia oli hankalampi löytää. Usean

pelaajan peleillä löydettiin sellaisia käytettävyysoongelmia, joihin ei löytynyt vastaavaa heuristiikkaa.

Paavolainen (2010) kehuu heuristiikkojen testausta riittäväksi ja heuristiikkoja selkeiksi, yhdenmukaisiksi ja helpoiksi ymmärtää. Heuristiikkojen vahvistusta olisi voinut lujittaa tekemällä pelitestausta ja vertaamalla tuloksia. Muutamat heuristiikat vaikuttivat kuitenkin olevan päällekkäisiä otsikkotasolla.

Korhonen ja muut (2009) testasivat heuristiikkoja teettämällä neljällä arvioijalla arvion valmiista mobiilipelistä. Arvioijien mukaan heuristiikkojen kuvauksien taso oli johdonmukainen, mutta ongelmia aiheutti heuristiikkojen yksityiskohtaisten kuvausten sijainti erillisessä dokumentissa. Lisäksi kuvaukset olivat liian pitkiä, jolloin niiden lukemiseen meni liian pitkään aikaa. Osa heuristiikoista oli arvioijien mukaan päällekkäisiä.

Korhonen (2011) testasi heuristiikkoja varhaisessa kehitysvaiheessa olevalla mobiilipelillä. Testauksessa oli mukana 18 yliopisto-opiskelijaa, jotka olivat kaikki käyneet perustason käytettävyysskurssin. Tutkimuksessa oli kaksi vaihetta. Ensimmäisessä vaiheessa arvioijat arvioivat peliä heuristiikkojen avulla. Toisessa vaiheessa arvioijille annettiin käytettävyyssiantuntijan samasta pelistä laatima käytettävyysongelmalista ja arvioijien tehtävänä oli valita ongelmia vastaavat heuristiikat.

Ensimmäisessä vaiheessa arvioijat löysivät yhteensä 250 ongelmaa, joista 47 hylättiin käytettävyyteen liittymättöminä virheraportteina. Hyväksytyistä ongelmista valittiin 17 tarkempaan tarkasteluun. Yksittäiset arvioijat olivat löytäneet ongelmille vastaavat heuristiikat, sillä ongelmille valittujen heuristiikkojen keskiarvo oli 1,08 heuristiikkaa ongelmalle. Kuitenkin ryhmänä arvioijat olivat valinneet samalle ongelmalle useita eri heuristiikkoja, mediaaniarvona neljä heuristiikkaa ongelmalle. Käytettävyyssiantuntijan valintoihin verrattuna arvioijat olivat valinneet samat heuristiikat keskimäärin 35 %:ssa tapauksista.

Toisessa vaiheessa arvioijilla oli selvästi hankalampaa valita ongelmaa vastaava heuristiikka, sillä arvioijat olivat valinneet keskimäärin 1,21 heuristiikkaa ongelmalle. Arvioijat eivät pystyneet valitsemaan ongelmaa vastaavaa heuristiikkaa 23:ssa tapauksessa. Käytettävyyssiantuntijan valintoihin verrattuna arvioijat olivat valinneet samat heuristiikat keskimäärin 48 %:ssa tapauksista.

Arvioijia pyydettiin arvioimaan heuristiikkoja subjektiivisesti. Heille ei kerrottu aikaisemmista tuloksista, joten kaikki arvioivat tuloksia pelkästään omien kokemuksiansa pohjalta. Arvioijien mielestä heuristiikat olivat kohtuullisen ymmärrettäviä ja helppokäyttöisiä, vaikka tulokset näyttävät muuta. (Korhonen, 2011.)

4.2.3 Korhosen ja Koiviston usean pelaajan mobiilipeliheuristiikat

Korhonen ja Koivisto tunnistivat kehittämistään mobiilipeliheuristiikoista aukon usean pelaajan mobiilipelien osalta (Korhonen & Koivisto, 2006). Näin ollen he kehittivätkin aiempiin heuristiikkoihinsa laajennoksen usean pelaajan mobiilipelejä varten. Heuristiikkojen kehitys aloitettiin tekemällä kirjallisuuskatsaus jossa etsittiin monen pelaajan peleihin liittyviä erityisheitteitä. Näiden pohjalta keksittiin kuusi alustavaa heuristiikkaa. Alustavia heuristiikkoja testattiin kolmella prototyyppiasteella olevalla mobiilipelillä, jotka edustivat eri tyyllilajeja. Lisäksi kahdelle pelille tehtiin pelitestaustyylinen kenttätutkimus, jossa kerättiin pelaajien kommentteja peleistä. Kaikille

kuudelle alustavalle heuristiikalle löydettiin vastaavia ongelmia. Testauksen pohjalta löydettiin tarve kahdelle uudelle heuristiikalle. Heuristiikat yhdistettiin ja saatiin lopulliset kahdeksan heuristiikkaa.

Lopullisia heuristiikkoja testattiin epämuodollisesti arvioimalla niiden avulla kuusi PC-peliä. Pelit edustivat useita tyyllilajeja. Tutkijoiden mukaan epämuodollisen testauksen tulosten perusteella voidaan sanoa, että heuristiikat ovat hyödylliset ja niitä on myös mahdollista soveltaa muihinkin kuin mobiilipeleihin. He kuitenkin myöntävät että heuristiikkojen vahvistamiseksi tarvitaan vielä lisää testausta.

Paavolaisen (2010) mukaan heuristiikkojen testaus on riittävää sekä heuristiikat ovat selkeät, yhdenmukaiset ja helppo ymmärtää. Heuristiikkojen vahvistusta olisi voinut lujittaa testaamalla niitä pelitestaukseen verraten. Lisäksi muutaman heuristiikan lyhyt kuvaus vaikutti päällekkäiseltä toisen heuristiikan kanssa.

4.2.4 Pinellen ja muiden heuristiikat

Pinelle ja muut (2008a) kritisoivat aiempia Federoffin (2002) sekä Desurviren ja muiden (2004) heuristiikkoja liiallisesta keskittymisestä pelin nautinnollisuuteen, joten he kehittivät oman heuristiikkakokoelman joka keskittyy enemmän käytettävyyteen. He etsivät GameSpot-verkkosivuilta haetuista 108:sta peliarvostelusta käytettävyyso ongelmia ja kehittivät heuristiikkansa näiden ongelmien pohjalta. Tutkijat eivät löytäneet käytettävyyso ongelmia peleistä jotka olivat saaneet korkeat arvosanat peliarvosteluissa, joten he rajasivat valitsemansa peliarvostelut peleihin, jotka olivat saaneet arvosteluasteikolla 0-10 arvosanaksi 8 tai alle. Arvostelut jaettiin kolmelle tutkijalle ja jokainen kehitti itsenäisesti löytämiensä käytettävyyso ongelmien pohjalta ongelmakategoriat ja jaotteli ongelmat niihin. Lopuksi tutkijat yhdistivät löytämänsä ongelmakategoriat ja muokkasivat niitä kunnes päädyttiin kahteentoista kategoriaan. Ongelmakategorioiden pohjalta kehitettiin kymmenen heuristiikkaa joilla vastaavan kategorian ongelmat voidaan välttää. Puuttuvat kaksi ongelmakategoriaa johtuvat siitä, että eräät kaksi heuristiikkaa kattavat kumpikin kaksi ongelmakategoriaa.

Taulukko 1. Heuristiikoilla löydetty ongelmat kategorioittain (Pinelle ja muut, 2008a).

Heuristiikka	Ongelmien määrä
1. Johdonmukaisuus	4
2. Muokattavuus	2
3. Ennustettavuus	4
4. Asianmukaiset näkymät	3
5. Ei-pelattavan sisällön ohitus	0
6. Syötekartoitukset	8
7. Ohjattavuus	4
8. Pelin tila	10
9. Ohjeistus ja tuki	12
10. Visuaaliset edustukset	17

Tutkijat testasivat heuristiikkoja palkkaamalla viisi arvioijaa, joilla kaikilla oli taustaa pelien pelaamisesta sekä käytettävyyso arviointien tekemisestä. Arvioijat arvioivat valmiin pelin kokeiluversiota. Kyseinen peli ei esiintynyt heuristiikkojen kehityksessä käytetyissä peliarvosteluissa. Arvioijat löysivät yhteensä 45 ongelmaa (taulukko 1).

Arvioijien mielestä heuristiikat saivat heidät huomaamaan ongelmia jotka olisivat muuten jääneet huomaamatta. Arvioijilla ei ollut ongelmia heuristiikkojen ymmärtämisessä ja heuristiikkojen soveltuvuudesta peleihin arvioijat olivat varmoja.

Tutkijoiden mukaan heidän menetelmänsä vaatii vielä lisää testausta suuremmalla arvioijamäärällä ja eri tyyllilajien peleillä. Koska heuristiikat kehitettiin yhden pelaajan PC-pelien pohjalta, tutkijat eivät ole varmoja kuinka hyvin heuristiikat pätevät muunlaisiin peleihin kuten konsolipeleihin, monen pelaajan peleihin ja hyötypeleihin. Koska heuristiikat keskittyvät pelkästään käytettävyyteen ja pelin nautintoon liittyvät seikat ovat tärkeitä pelin menestykselle, tutkijat ehdottavat heuristiikkoja käytettäväksi yhdessä muiden heuristiikkojen kuten Federoffin (2002) heuristiikkojen ja HEP-menetelmän heuristiikkojen kanssa.

Paavilaisen (2010) mukaan heuristiikkojen kehityksessä käytettyjen käytettävyyso Ongelmien etsiminen peliarvosteluista on mielenkiintoista mutta myös ongelmallista, sillä peliarvostelijat eivät ole käytettävyyсарvioijia. Lisäksi peliarvostelijoiden ja käytettävyyсарvioijien tavoitteet ja kohderyhmä eroavat toisistaan huomattavasti. Paavilainen esittää, että tämä ilmenee peliarvosteluista löydettyjen käytettävyyso Ongelmien alhaisessa määrässä. Koeffel ja muut (2010) kritisoivat GameSpot-verkkosivuston käyttöä arvostelujen lähteenä sillä GameSpotin on epäilty erottaneen peliarvostelijan, koska kyseinen arvostelija oli kirjoittanut negatiivisen peliarvostelun GameSpotin sponsorin julkaisemasta pelistä. Paavilainen (2010) jatkaa kritisoimalla heuristiikkojen testauksen tulosten tilastollista voimaa heikoksi. Hän kehuu heuristiikkojen selkeyttä ja selityksiä. Heuristiikat ovat ymmärrettävissä ilman muita dokumentteja ja heuristiikkojen määrä ei ole liian suuri.

Korhonen (2011) testasi heuristiikkoja varhaisessa kehitysvaiheessa olevalla mobiilipelillä. Testauksessa oli mukana 18 yliopisto-opiskelijaa, jotka olivat kaikki käyneet perustason käytettävyyso Kurssin. Tutkimuksessa oli kaksi vaihetta. Ensimmäisessä vaiheessa arvioijat arvioivat peliä heuristiikkojen avulla. Toisessa vaiheessa arvioijille annettiin käytettävyyso Asiantuntijan samasta pelistä laatima käytettävyyso Ongelmalista ja arvioijien tehtävänä oli valita ongelmia vastaavat heuristiikat.

Ensimmäisessä vaiheessa arvioijat löysivät yhteensä 230 ongelmaa, joista 59 hylättiin käytettävyyteen liittymättöminä virheraportteina. Hyväksytyistä ongelmista valittiin 17 tarkempaan tarkasteluun. Yksittäiset arvioijat olivat löytäneet ongelmille vastaavat heuristiikat, sillä arvioijat olivat valinneet keskimäärin 1,08 heuristiikkaa ongelmalle. Kuitenkin ryhmänä arvioijat olivat valinneet samalle ongelmalle useita eri heuristiikkoja, mediaaniarvona kolme heuristiikkaa ongelmalle. Käytettävyyso Asiantuntijan valintoihin verrattuna arvioijat olivat valinneet samat heuristiikat keskimäärin 33 %:ssa tapauksista.

Toisessa vaiheessa arvioijilla oli hankalampaa valita ongelmaa vastaava heuristiikka, sillä arvioijat olivat valinneet keskimäärin 1,13 heuristiikkaa ongelmalle. Arvioijat eivät pystyneet valitsemaan ongelmaa vastaavaa heuristiikkaa 89:ssä tapauksessa. Käytettävyyso Asiantuntijan valintoihin verrattuna arvioijat olivat valinneet samat heuristiikat keskimäärin 41 %:ssa tapauksista.

Arvioijia pyydettiin arvioimaan heuristiikkoja subjektiivisesti. Heille ei kerrottu aikaisemmista tuloksista, joten kaikki arvioivat tuloksia pelkästään omien kokemuksiansa pohjalta. Arvioijien mielestä heuristiikat olivat kohtuullisen ymmärrettäviä ja helppokäyttöisiä, vaikka tulokset näyttävät muuta. (Korhonen, 2011.)

4.2.5 Networked Game Heuristics (NGH)

Pinelle ja muut (2008a) totesivat tutkimuksessaan tarpeen erityisesti usean pelaajan peleille kehitetyille heuristiikoille. Pinelle, Wong, Stach ja Gutwin (2009) kehittivät tällaiset heuristiikat. Heidän mukaansa usean pelaajan peleissä on samat käytettävyyshuolet kuin yhden pelaajan peleissä, mutta usean pelaajan peleissä on lisäksi omat käytettävyyshuolensa. Usean pelaajan pelien täytyy tukea usean pelaajan välistä kommunikaatiota, koordinaatiota, tietoisuutta sekä sosiaalisia vuorovaikutuksia. Usean pelaajan pelikäytettävyyden määritelmänä he käyttävät seuraavaa: "pelaajan kyvyn aste onnistuneeseen vuorovaikutukseen muiden pelaajien kanssa ja heidän toimintansa ymmärtämiseen".

Pinelle ja muut (2009) kehittivät heuristiikkansa samalla tavalla kuin Pinelle ja muut (2008a). He hakivat GameSpot- ja GameSpy-verkkosivuilta 382 usean pelaajan pelin arvostelua ja etsivät niistä käytettävyysoongelmia jotka liittyivät pelin usean pelaajan aspekteihin. Ongelmia löytyi 213:sta arvostelusta. Niistä poistettiin vielä ne käytettävyysongelmat jotka johtuivat pelisuunnittelusta ja ulkoisista seikoista kuten verkkoviiveestä ja pelipalvelinten saatavuudesta. Tämän jälkeen arvosteluja oli jäljellä 123 kappaletta. Arvostelut jaettiin kolmen tutkijan kesken ja jokainen heistä kehitti omat ongelmakategoriat ja jaottelivat löydetyt ongelmat näihin kategorioihin. Lopulta kaikkien kolmen kategoriat yhdistettiin ja muokattiin lopulliseksi kymmeneksi ongelmakategoriaksi. Jokaista ongelmakategoriaa kohden kehitettiin heuristiikka jonka avulla kyseisen ongelmakategorian ongelmia voitaisiin välttää. Syntyneitä heuristiikkoja kutsutaan nimellä Networked Game Heuristics (NGH).

NGH-heuristiikkoja testattiin kahdella pelillä kymmenen arvioijan voimin. Vertailun vuoksi arvioijat jaettiin kahteen viiden hengen ryhmään. Ensimmäinen ryhmä käytti arviointiin ensin NGH-heuristiikkoja ja sitten Bakerin (2002) groupware-heuristiikkoja. Toinen ryhmä käytti heuristiikkoja päinvastaisessa järjestyksessä. Valitut kaksi peliä olivat molemmat avoimen lähdekoodin pelejä, sillä tutkijoiden mukaan avoimen lähdekoodin peleissä on samanlaisia suunnitteluvikoja kuin kaupallisten pelien pelattavissa prototyypeissä.

Taulukko 2. NGH-heuristiikoilla löydetyt ongelmat kategorioittain (Pinelle ja muut, 2009).

Heuristiikka	Ongelmien määrä
1. Istunnonhallinta	7
2. Pelinmuodostus	8
3. Kommunikaatio	8
4. Koordinaatio	12
5. Tietoisuus	8
6. Ruumiillistumat	5
7. Valmennus	10
8. Sosiaaliset mahdollisuudet	4
9. Vuorovaikutuksen aikaskaala	0
10. Huijaaminen ja epämieluisa käyttäytyminen	6

Arvioijat löysivät yhteensä 49 käytettävyysongelmaa Bakerin heuristiikoilla ja 67 NGH-heuristiikoilla (taulukko 2). Arvioijien mielestä Bakerin heuristiikat eivät olleet sopivia pelien käytettävyyсарviointiin kun taas NGH-heuristiikat soveltuivat hyvin. Joidenkin arvioijien mielestä NGH-heuristiikkojen joukossa oli sellaisia heuristiikkoja jotka eivät päde kaikkiin usean pelaajan peleihin, mutta silti heuristiikat eivät kattaneet kaikkia ongelmia mitä usean pelaajan peleihin liittyy. Toisaalta toisten arvioijien mielestä NGH-heuristiikoilla löydettiin ongelmia, joita he eivät olisi muuten huomanneet. Yhdeksän arvioijaa kymmenestä piti NGH-heuristiikkoja Bakerin heuristiikkoja parempina.

Tutkijoiden mukaan NGH-heuristiikkoja on mahdollista käyttää samankaltaisten usean pelaajan pelien arviointiin. On kuitenkin mahdollista, että heuristiikat eivät päde peleihin jotka ovat hyvin erilaisia verrattuna niihin peleihin joiden peliarvostelujen pohjalta heuristiikat luotiin. Spesifisesti heuristiikat eivät todennäköisesti päde sellaisiin peleihin joissa pelaajat pelaavat peliä jaetulla näytöllä.

Koska heuristiikat ottavat huomioon vain usean pelaajan pelien käytettävyyshuolet, NGH-heuristiikkojen lisäksi on käytettävä yksinpeleille kehitettyjä heuristiikkoja. On myös mahdollista, että pelissä tarkoituksella rikotaan heuristiikkoja, jotta saadaan luotua haastetta pelaajalle. Tutkijat väittävät NGH-heuristiikkojen edustavan niitä virheitä jotka jäävät yleensä huomaamatta pelinkehityksessä, koska heuristiikat kehitettiin peliarvostelujen pohjalta. He kuitenkin myöntävät, että peliarvostelijat eivät ole käytettävyyssasiantuntijoita, joten on mahdollista että heuristiikat ovat osittain puutteellisia. (Pinelle ja muut, 2009.)

Paavolaisen (2010) kritiikki NGH-heuristiikkoja kohtaan ovat suurelta osin samat kuin hänen kritiikkinsä Pinellen ja muiden (2008a) heuristiikkoja kohtaan sillä molemmat heuristiikat ovat kehitetty peliarvostelujen pohjalta. Paavolaisen mukaan löydös NGH-heuristiikkojen paremmuudesta verrattuna Bakerin heuristiikkoihin oli aksiomaattinen ja NGH-heuristiikkoja olisikin kannattanut verrata Korhosen ja Koiviston (2007) heuristiikkoihin. Myös NGH-heuristiikkojen tulosten tilastollista voimaa Paavolainen kuvailee heikoksi.

4.2.6 PLAY-heuristiikat

Desurvire ja Wiberg (2009) kehittivät aiempien HEP-heuristiikkojen (Desurvire ja muut, 2004) pohjalta parannellut heuristiikat joita he kutsuvat PLAY-heuristiikoiksi. Heidän mukaansa HEP oli hyödyllinen vain liian rajoitetuissa olosuhteissa ja liian myöhäisessä vaiheessa pelinkehitystä. Tämän vuoksi PLAY-heuristiikat on kehitetty olemaan yleinen perusta jota voidaan muokata pelikohtaisesti. PLAY-heuristiikat ovat suunniteltu käytettäväksi varsinkin pelikehityksen alkuvaiheessa, jolloin muutosten tekeminen on halvempaa, mutta heuristiikkoja voi käyttää läpi suunnitteluprosessin.

Tutkijat kehittivät HEP-heuristiikkojen pohjalta uudet heuristiikat, joiden kohteena olivat kolme pelien tyylilajia: reaaliaikainen strategia, toiminta-seikkailu sekä ensimmäisen persoonan räiskintä. Näitä heuristiikkoja muokattiin yhteistyössä pelialalla työskentelevien pelikehittäjien kanssa. Näin muodostettuja heuristiikkoja oli 116 kappaletta ja ne jaettiin seuraaviin kategorioihin:

- Pelattavuus
- Taidon kehittyminen
- Ohjeistus
- Strategia & haaste

- Pelin/tarinan mukaansatempaavuus
- Siistiys
- Käytettävyys/pelimekaniikat
- Ohjain/näppäimistö

Heuristiikkoja testaukseen kehitettiin oma kyselylomake jokaiselle kolmelle kohteena olleelle tyylilajille. Kyselylomake sisälsi myös kaksi listaa peleistä. Ensimmäinen lista sisälsi pelejä, jotka olivat saaneet peliarvosteluissa Metacritic-sivustolla keskiarvosanaksi yli 80 pistettä. Toinen lista sisälsi samalta sivustolta pelejä, jotka olivat saaneet keskiarvosanaksi alle 50 pistettä. Tutkijat värväsivät kyselylle vastaajia pelikonferenssissa. Vastaajat valittiin sen perusteella, että he olivat pelanneet mitä tahansa peliä molemmilta listoilta.

Hyväksytyjä vastauksia kyselyyn saatiin 52 kappaletta. Kyselyn tulokset analysoitiin tilastollisilla menetelmillä ja löydettiin 48 heuristiikkaa joilla oli eroja hyvien ja huonojen pelien välillä. Nämä 48 heuristiikkaa valittiin lopullisiksi PLAY-heuristiikoiksi. Tutkijat väittävät että PLAY-heuristiikkoja on käytetty pelinkehityksessä kahden vuoden ajan ja PLAY-heuristiikkojen käytöstä on hyviä kokemuksia. Tutkijat lupasivat julkaisevansa tulevaisuudessa tutkimuksissa lisää PLAY-heuristiikkojen testausta, mutta tällaisia tutkimuksia ei löydetty tätä tutkielmaa tehtäessä.

4.2.7 GAP-heuristiikat

Desurviren ja Wibergin (2015) mukaan peliala on siirtänyt keskittymistään paljon pelaavista satunnaisiin pelaajiin (engl. *casual gamer*) ja tämän vuoksi pelin lähestyttävyyden (engl. *approachability*) on muuttunut tärkeämmäksi. Lähestyttävyydellä tarkoitetaan pelin alkuvaiheessa tapahtuvaa pelimekaniikan opettelua sekä pelaajan motivointia pelin pelaamisen jatkamiseen. Tutkijoiden mukaan satunnaiset pelaajat tarvitsevat enemmän ohjausta pelien pelaamiseen, sillä heillä ei usein ole kovin paljoa aikaisempaa kokemusta peleistä.

Lähestyttävyyden lähtökohdista tutkijat lähtivät kehittämään sen arviointiin sopivia uusia heuristiikkoja. Kehityksen tuloksena syntyneet Game Approachability Principles (GAP) -heuristiikat kehitettiin aiempien PLAY-heuristiikkojen (Desurvire & Wiberg, 2009) sekä oppimistutkimuksen pohjalta. GAP-heuristiikat sisältävät 11 pääkategoriaa, joista yksi ovat PLAY-heuristiikat.

Heuristiikkoja testattiin neljälle pelille tehdyllä heuristisella arvioinnilla ja tuloksia verrattiin samoille peleille tehdyllä pelitestauksella. Pelitestaukseen osallistui yhteensä 32 pelaajaa ja jokaista peliä pelasi 6-12 pelaajaa. Suurin osa pelaajista oli nuoria miehiä. Noin puolet pelaajista luokiteltiin satunnaisiksi pelaajiksi. Pelitestaus suoritettiin ääneen ajattelu -menetelmällä yhden pelaajan ja yhden tutkijan ollessa paikalla. Lisäksi pelaajilta kysyttiin kysymyksiä. Tutkija kirjasi ylös pelaajan suoriutumisen ja kommentit sekä luokitteli nämä positiivisiksi ja negatiivisiksi kokemuksiksi. Tämän jälkeen löydökset luokiteltiin GAP-heuristiikkojen kategorioihin. Molemmilla tavoilla löydetty ongelmat jaoteltiin lähestyttävyyteen ja pelattavuuteen liittyviksi ongelmiksi ja niitä vertailtiin keskenään. Lopulta ongelmien kuvauksia analysoitiin samankaltaisuuksien ja eroavaisuuksien löytämiseksi kuvausten yksityiskohtaisuudessa.

Pelitestauksella löydettiin 22 käytettävyysongelmaa ja heuristisella arvioinnilla 43. Heuristisella arvioinnilla löydettiin enemmän ja laajemmalti lähestyttävyyden ongelmia kun taas pelitestauksella löydettiin enemmän pelattavuusongelmia. Ongelmien

kuvausten analyysissä pelitestauksella löydettyjen ongelmien kuvaukset olivat yksityiskohtaisempia kuin heuristisella arvioinnilla löydettyjen ongelmien kuvaukset. Tutkijoiden mukaan tämä ero johtuu todennäköisesti siitä, että GAP-heuristiikoilla löydetään yleiset ongelma-alueet kun taas pelitestauksella löydetään ongelma-alueiden sisältämät yksittäiset ongelmat. Tutkijoiden mukaan GAP-heuristiikat toimivatkin parhaiten pelitestauksen kanssa käytettynä niin, että GAP-heuristiikkoja käytetään tarkastuslistana suunnitteluvaiheessa ja pelitestauksella varmennetaan ja hiotaan suunnitelma loppuun.

4.2.8 Hochleitnerin ja muiden heuristiikat

Koeffel ja muut (2010) kehittivät useiden aikaisempien heuristiikkojen pohjalta 29 heuristiikan kokoelman. Myöhemmän tutkimuksen pohjalta näistä heuristiikoista löydettiin ristiriitoja ja käytön hankaluutta (Hochleitner, Hochleitner, Graf & Tscheligi, 2015). Näin ollen Hochleitner ja muut (2015) kehittivät Koeffelin ja muiden heuristiikkojen pohjalta uudet heuristiikat tarkoituksenaan tehdä niistä helppokäyttöisemmät, kattavammat ja paremmin eri tyylilajit huomioon ottavat. Heuristiikkojen kehityksessä on myös otettu vaikutteita käyttäjäkokemuksen (engl. *UX, user experience*) käsitteestä. Lopputuloksena syntyi 49 heuristiikan kokoelma jaettuna kahteen pääkategoriaan ja 12 alakategoriaan (taulukko 3).

Taulukko 3. Heuristiikkojen yhteenveto (Hochleitner ja muut, 2015).

Pääkategoria	Alakategoria	Sisältö	Heuristiikkojen lukumäärä
Pelattavuus / tarina	Tavoitteet	Kokonaistavoite, lyhyempiaikaiset tavoitteet	2
	Motivaatio	Pelaamisen jatkamiselle tärkeät asiat	7
	Haaste	Vaikeustaso, pelaajan taidot	7
	Oppiminen	Apu, opastus, virhetilanteet	4
	Hallinta	Pelimaailmaan vaikuttaminen	6
	Johdonmukaisuus	Ennustettavat ja pysyvät vaikutukset	2
	Tarina	Mukaansatempaavuus, kertomuksellisuus	3
Käyttöliittymä	Palaute	Pelitulanteen ja tapahtumien näkyvyys	6
	Ulkonäkö	Esineiden näkyvyys ja käyttötarkoitusta vastaava ulkonäkö	2
	Vuorovaikutus	Pelaajan syöte	3
	Muokattavuus	Mahdollisuus muokata peliä käyttäjän	2
	Valikot ja käyttöliittymä	Valikkojen ja käyttöliittymäelementtien tarkoituksenmukaisuus	5

Heuristiikkoja testattiin tekemällä kuudelle indie-pelille heuristinen arviointi ja vertaamalla sen tuloksia samojen pelien peliarvosteluista saamiin arvosanoihin. Indie-pelit ovat pelejä, jotka julkaistaan pelinkehittäjän toimesta ilman erillisen julkaisijan käyttöä. Arvosanat ovat haettu Metacritic-palvelusta, joka laskee pelille keskiarvosanan useista peliarvosteluista. Pelit edustivat kolmea eri tyyllilajia ja ne valittiin siten, että niiden keskiarvosanat asettuivat suurelle arvovälille. Tutkijoiden mukaan heuristinen arviointi sopii indie-peleihin alhaisten kustannuksen vuoksi. Yhteys käyttäjäkokemukseen tulee tutkijoiden mukaan myös tässä esille, sillä heidän mukaansa peliarvostelut ovat itse asiassa käyttäjäkokemuksen arviointia.

Heuristinen arviointi toteutettiin kolmella arvioijalla, joilla oli vaihteleva määrä kokemusta pelaamisesta sekä käytettävyydestä. Arvioijat tutustuivat heuristiikkoihin ennalta, pelasivat jokaista peliä tasan 30 minuuttia ja kirjasivat ylös löytämänsä käytettävyysongelmat. Arvioijat arvioivat jokaisen löytämänsä käytettävyysongelman Nielsenin (1994) viisiportaisella arviointiasteikolla. Lisäksi arvioijat arvioivat pelejä heuristiikkakohtaisesti antaen niille arvosanan sen perusteella kuinka hyvin peli noudatti heuristiikkoja.

Tuloksina saatuja löydettyjen käytettävyysongelmien määrää sekä heuristiikkakohtaisia arviointeja vertailtiin pelien Metacritic-keskiarvosanoihin tilastollisin menetelmin. Käytettävyysongelmien määrän ja Metacritic-keskiarvosanojen välillä havaittiin voimakas negatiivinen korrelaatio. Heuristiikkakohtaisten arviointien ja Metacritic-keskiarvosanojen välillä havaittiin voimakas positiivinen korrelaatio. Tämän pohjalta tutkijat toteavat heuristisen arvioinnin ja käyttäjäkokemuksen välisen yhteyden olemassaolon. Tutkijat myöntävät, että tulosten tarkkuutta heikentävät pelin arvosanan käyttö peliarvosteluista kokotekstin sijaan sekä lyhyt 30 minuutin arviointijakso. (Koeffel ja muut, 2010.) Heuristiikkojen suuri lukumäärä on myös ongelmallista, sillä Paavilaisen (2010) mukaan jo 40 heuristiikkaa on liian paljon.

5. Keskustelu

Pelitestausta lukuun ottamatta kaikki löydetty pelikäytettävyyssmenetelmät perustuivat heuristiseen arviointiin. Heuristisen arvioinnin soveltaminen pelien käytettävyyden arviointiin ei kuitenkaan ole ongelmatonta. Heuristisen arvioinnin alkuperäinen idea on ollut tehdä arviointia mahdollisimman halvalla ja mahdollisimman aikaisessa vaiheessa jotta muutosten tekeminen olisi halpaa (Nielsen, 1994, p. 25). Korhosen ja muiden (2009) tutkimuksessa arvioijat valittivat pelinautinnon olevan hankalaa arvioida, sillä arviointitilanne pakotti pelaamaan peliä eri tavalla kuin tavallisessa pelaustilanteessa. Kyseinen arviointitilanne oli kuitenkin pelinautinnon arvioimiseen optimaalinen, sillä arvioinnin kohteena ollut peli oli julkaistu ja täysin valmis. Mikäli valmiin pelin pelinautintoa on hankala arvioida, miten hankalaa se on tehdä suunnitteludokumenttien pohjalta? Tätä kysymystä ei käsitelty löydettyissä tutkimuksissa ollenkaan, sillä kaikkia menetelmiä testattiin myöhemmässä kehitysvaiheessa olevalla pelillä.

Kaikki löydetty menetelmät olivat kehitetty etsimällä käytettävyyssongelmia enemmän tai vähemmän valmiista peleistä. Osa menetelmistä oli kehitetty jopa peliarvosteluista etsittyjen ongelmien pohjalta. Pinelle ja muut (2009) väittävät peliarvosteluista käytettävyyssongelmien etsimisen olevan hyvä asia, sillä näin kehitetyillä heuristiikoilla löydetään ne ongelmat, jotka usein jäävät peleihin. Näin ollen heuristiikkojen tehtäväksi jää löytää ne ongelmat, jotka jäivät löytämättä pelitestauksella. Tässä on ilmeinen ristiriita, sillä heuristisen arvioinnin koko idea on tehdä sitä halvalla mahdollisimman aikaisin, siis ennen pelitestausta. Näin ollen kallista pelitestausta voitaisiin tehdä vähemmän, koska sillä tarvitsisi löytää vain heuristisella arvioinnilla löytymättä jäävät ongelmat. Heuristinen arviointi ei selvästikään ole yksin käytettynä riittävä keino pelikäytettävyyden arviointiin. Korhosen (2010) tutkimuksessa 40 % käytettävyyssongelmista jäivät löytämättä heuristisella arvioinnilla ja Desurviren ja muiden (2004) tutkimuksessa heuristisella arvioinnilla löydettiin pelitestaukseen verrattuna erilaisia ongelmia. Useiden arviointikeinojen käyttäminen tietysti tarkoittaa, että heuristisen arvioinnin käyttö ei tuo niin suurta kustannus- ja nopeusetua kuin se voisi yksin käytettynä tuottaa.

Heuristiikkojen sopivuus ylipäänsä pelien käytettävyyden arviointiin ei ole selvää. Alkuperäisten Nielsenin heuristiikkojen ajatus oli olla kohtuullisen helposti muistettava lista jonka avulla arvioija pystyy tarkastelemaan arvioimaansa järjestelmää eri näkökulmista. Pelit ovat kuitenkin monimuotoisempia kuin hyötyohjelmat, joten Nielsenin heuristiikkojen kokoa vastaavia heuristiikkakokoelmia kritisoitiin liian ylimalkaisiksi kun taas useampia heuristiikkoja sisältäviä ja siten tarkempia heuristiikkakokoelmia kritisoitiin liian laajoiksi. Siis samalla kun heuristiikkojen määrä ja siten tarkkuus kasvavat, niiden käyttökelpoisuus vähenee. Pelikontekstissa heuristiikkoja voidaan myös rikkoa tarkoituksella (Pinelle ja muut, 2009) mikä tekee arvioinnista hankalampaa. Näitä seikkoja ei kuitenkaan ilmeisesti nähdä ongelmana. Esimerkiksi Zhu, Zhao, Fang ja Moser (2017) näkivät ongelmana aiemmissa tutkimuksissa aineistona käytettyjen peliarvostelujen määrän vähyden. Ratkaisuna he käsitelivät yli 800 tuhatta peliarvostelua koneellisesti tehden sanastoanalyysiä arvostelujen adjektiivien ja substantiivien perusteella. Sanastoanalyysin löytämistä yhteyksistä kehitettiin 90 heuristiikkaa, mikä on löydettyistä tutkimuksista selvästi suurin määrä.

Korhonen ja muut (2009) testasivat Korhosen ja Koivoston (2006) sekä Desurviren ja muiden (2004) heuristiikkoja neljällä kahden arvioijan ryhmällä. Kolmella arvioijalla oli aikaisempaa kokemusta heuristisesta arvioinnista. Tulokset olivat yllättäviä, sillä 75 % löydettyistä ongelmista olivat vain yhden ryhmän löytämiä ja vain 2 % sellaisia, jotka kaikki ryhmät olivat löytäneet. Korhonen (2011) antoi kahdelle noviisiarvioijista koostuneelle ryhmälle listan pelikäytettävyysongelmia ja pyysi heitä osoittamaan jokaiselle ongelmalle riktun heuristiikan. Vertailussa käytettävyyssiantuntijan valintoihin paljastui, että noviisiarvioijat olivat osoittaneet saman heuristiikan vain 33-35 %:ssa tapauksista. Heuristisen arvioinnin suorittaminen vaatii siis asiantuntijuutta.

Pelikäytettävyyden tutkimus ei ole vielä vaikuttanut pelialan käytäntöihin. Rajasen ja Nissisen (2015) kyselytutkimuksessa pohjoiseurooppalaisissa peliyhtiöissä käytettiin hyvin paljon pelitestausta ja siihen verrattavissa olevia menetelmiä ja vain hyvin vähän heuristista arviointia. Selvästi myöskään tieto tutkituista menetelmistä ei ole saavuttanut peliyhtiöitä: kyselyyn vastanneista suurin osa ei ollut tietoisia heuristisesta arvioinnista ja osa vastaajista jopa piti heuristista arviointia virheellisesti liian aikaa vievänä tai kalliina keinona. (Rajanen & Nissinen, 2015). Samankaltaisia tuloksia on saatu kyselytutkimuksella myös pohjoisamerikkalaisista peliyhtiöistä (Rajanen & Rajanen, 2018.) Monet heuristiseen arviointiin perehtyneistä vastaajista pitivät kirjallisuudessa esitettyjä heuristiikkoja huonoina ja pitivät peliyhtiöiden itse räätälöimiä heuristiikkoja välttämättömänä (Rajanen & Nissinen, 2015; Rajanen & Rajanen, 2018.) Pelialan sisällä onkin tapahtunut liikehdintää tähän suuntaan. Falstein ja Barwood (2006) ovat kehittäneet vuodesta 2001 listaa pelisuunnittelun nyrkkisäännöistä nimellä "The 400 Project". Paavilaisen (2010) mukaan tätä kokoelmaa nyrkkisääntöjä voidaan pitää eräänlaisina heuristiikkoina, vaikka Falstein ja Barwood eivät niitä sillä nimellä kutsukaan.

6. Yhteenveto

Tässä tutkielmassa oli tarkoitus selvittää kirjallisuuskatsauksen avulla:

1. Millaisia menetelmiä on kehitetty viihteellisten pelien käytettävyyden arviointiin.
2. Millaisia tuloksia menetelmiä soveltamalla on saatu.

Kirjallisuuskatsauksessa löydettiin yksi pelialan sisällä kehitetty menetelmä, pelitestausta sekä useita heuristiseen arviointiin perustuvia menetelmiä. Kaikissa löydettyistä tutkimuksista toistui sama kaava: aiemmin kehitetyt heuristiikat todettiin huonoiksi tai puutteellisiksi, kehitettiin omat heuristiikat ja väitettiin niiden sopivan todella hyvin pelien käytettävyyden arviointiin. Suurin osa heuristiikoista oli kehitetty menetelmillä, joita kuvailisin epäilyttäväksi, kuten esimerkiksi peliarvosteluista käytettävyysongelmien etsiminen.

Paavilainen (2010) toteaa kaikkien arvioimiensa menetelmien kehitystavan ja testauksen olevan enemmän tai vähemmän puutteellisesti suoritettua. Näitä päätelmiä tukee myös muiden tutkijoiden löydökset menetelmien testaamisesta. Toisin kuin menetelmän kehittäjät väittävät, kaikista menetelmistä raportoitiin useita ongelmia. Monet menetelmistä sisältävät myös erittäin suuren määrän heuristiikkoja, mikä on todettu ongelmalliseksi.

Rajasen ja Nissisen (2015) sekä Rajasen ja Rajasen (2018) mukaan tutkitut menetelmät eivät ole siirtyneet ainakaan pohjoiseurooppalaisten tai pohjoisamerikkalaisten peliyhtiöiden käyttöön. Olettaisiin, että tilanne on myös muualla maailmassa sama. Mielestäni tämä ei ole ongelma, sillä mikään menetelmä ei ole riittävän laadukas laajamittaiseen käyttöön pelialalla.

Tämän tutkielman kirjallisuuskatsaus rajoitettiin vain viihteellisiin peleihin, mikä jätti kirjallisuuskatsauksesta pois muun muassa opetukselliset ja mainokselliset pelit. Lisäksi tutkielmassa rajoitettiin pois tutkimukset, jotka käsittelivät nykyisin epäsuosittuja pelityylejä, kuten pelihallipelejä.

Lisätutkimusta tarvitaan selvästi, sillä tutkimuksella ei ole ollut juuri vaikutusta käytännössä. Peliala on nopeasti kasvanut yhdeksi suurimmista viihdealoista, mutta pelien käytettävyyttä arvioidaan edelleen samoin vanhojen menetelmien. Lisäksi pelikäytettävyyden tutkiminen on selvästi rajoittunut perinteisen käytettävyyden heuristiikkojen muokkaamiseen ja soveltamiseen. Hyötyohjelmien käytettävyyden arviointiin on kehitetty paljon enemmän menetelmiä, eikä ole näyttöä siitä, että heuristinen arviointi olisi ainoa tai paras menetelmä pelikäytettävyyden arviointiin.

Tulevissa tutkimuksissa voisikin yrittää muokata muita hyötyohjelmien käytettävyyden menetelmiä peleihin soveltuviksi. Näiden lisäksi voisi olla arvokasta yrittää kehittää peleille täysin omia käytettävyyden menetelmiä, koska pelien käytettävyyden lähtökohdat ovat niin erilaiset perinteiseen käytettävyyteen. Lisää tutkimusta tarvitaan myös pelikäytettävyyden käsitteiden määrittämiseen ja pelien tyyllilajien, pelaamiseen käytettyjen alustojen ja välineiden sekä peliympäristön vaikutukseen pelikäytettävyyden arviointiin. Mielestäni tulevassa tutkimuksessa pitäisi myös pyrkiä tiiviimpään

yhteistyöhön pelinkehittäjien kanssa. Näin menetelmiä ei tarvitsisi kehittää liian monen välikäden kautta kulkeneiden tietojen, kuten peliarvostelujen, pohjalta. Lisäksi yhteistyö saattaisi johtaa menetelmien laajempaan ymmärtämiseen ja käyttöön pelialalla.

Kirjallisuusviitteet

- Baker, K., Greenberg, S., & Gutwin, C. (2002). Empirical development of a heuristic evaluation methodology for shared workspace groupware. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, 96-105. ACM.
- Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In *CHI'04 extended abstracts on Human factors in computing systems*, 1509-1512. ACM.
- Desurvire, H., & Wiberg, C. (2009). Game usability heuristics (PLAY) for evaluating and designing better games: The next iteration. In *International Conference on Online Communities and Social Computing*, 557-566. Berlin: Springer.
- Desurvire, H., & Wiberg, C. (2015). User Experience Design for Inexperienced Gamers: GAP—Game Approachability Principles. In *Game User Experience Evaluation* (pp. 169-186). Cham: Springer.
- Falstein, N., Barwood, H. (2006). *The 400 Project*. Lainattu 10.3.2018, saatavilla: <http://www.finitearts.com/Pages/400page.html>
- Federoff, M. A. (2002). *Heuristics and usability guidelines for the creation and evaluation of fun in video games* (Doctoral dissertation, Indiana University).
- Genise, P. (2002). Usability evaluation: methods and techniques: Version 2.0. *University of Texas*.
- Hochleitner, C., Hochleitner, W., Graf, C., & Tscheligi, M. (2015). A heuristic framework for evaluating user experience in games. In *Game User Experience Evaluation* (pp. 187-206). Cham: Springer.
- Jørgensen, A. H. (2004, October). Marrying HCI/Usability and computer games: a preliminary look. In *Proceedings of the third Nordic conference on Human-computer interaction*, 393-396. ACM.
- Koeffel, C., Hochleitner, W., Leitner, J., Haller, M., Geven, A., & Tscheligi, M. (2010). Using heuristics to evaluate the overall user experience of video games and advanced interaction games. In Bernhaupt, R. (Ed.), *Evaluating user experience in games* (pp. 233-256). London: Springer.
- Korhonen, H. (2010). Comparison of playtesting and expert review methods in mobile game evaluation. In *Proceedings of the 3rd International Conference on Fun and Games*, 18-27. ACM.
- Korhonen, H. (2011). The explanatory power of playability heuristics. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, 40. ACM.

- Korhonen, H., & Koivisto, E. M. (2006). Playability heuristics for mobile games. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, 9-16. ACM.
- Korhonen, H., & Koivisto, E. M. (2007). Playability heuristics for mobile multi-player games. In *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts*, 28-35. ACM.
- Korhonen, H., Paavilainen, J., & Saarenpää, H. (2009). Expert review method in game evaluations: comparison of two playability heuristic sets. In *Proceedings of the 13th international MindTrek conference: Everyday life in the ubiquitous era*, 74-81. ACM.
- Livingston, I. J., Mandryk, R. L., & Stanley, K. G. (2010). Critic-proofing: how using critic reviews and game genres can refine heuristic evaluations. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, 48-55. ACM.
- Malone, T. W. (1982). Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the 1982 conference on Human factors in computing systems*, 63-68. ACM.
- Marchand, A., & Hennig-Thurau, T. (2013). Value creation in the video game industry: Industry economics, consumer benefits, and research opportunities. *Journal of Interactive Marketing*, 27(3), 141-157.
- Nielsen, J. (1993). *Usability engineering*. Boston: Academic Press.
- Nielsen, J. (1994). Heuristic Evaluation. In Mack, R. L., & Nielsen, J. (Eds.), *Usability inspection methods* (pp. 25-62). New York: Wiley & Sons.
- Oulasvirta, A. (2011). Mitä on ihmisen ja tietokoneen vuorovaikutus?. In Oulasvirta, A. (Ed.), *Ihmisen ja tietokoneen vuorovaikutus* (pp. 13-42). Helsinki: Gaudeamus Helsinki University Press.
- Paavilainen, J. (2010). Critical review on video game evaluation heuristics: social games perspective. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, 56-65. ACM.
- Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2008). *User-centered design in games*. Boca Raton: CRC Press.
- Pinelle, D., Wong, N., & Stach, T. (2008a). Heuristic evaluation for games: usability principles for video game design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1453-1462. ACM.
- Pinelle, D., Wong, N., & Stach, T. (2008b). Using genres to customize usability evaluations of video games. In *Proceedings of the 2008 conference on future play: Research, play, share*, 129-136. ACM.
- Pinelle, D., Wong, N., Stach, T., & Gutwin, C. (2009). Usability heuristics for networked multiplayer games. In *Proceedings of the ACM 2009 international conference on Supporting group work*, 169-178. ACM.

- Rajanen, M., & Marghescu, D. (2006). The impact of game usability to player attitude. In *Proceedings of 29th Information Systems Research Seminar In Scandinavia, Helsingoer, Denmark*, 1-17.
- Rajanen, M., & Nissinen, J. (2015). A survey of game usability practices in Northern European game companies. In *Association for Information Systems* (No. 6).
- Rajanen, M., & Rajanen, D. (2017). Usability benefits in gamification. In *Proceedings of the 1st GamiFIN Conference*, 87-95.
- Rajanen, M., & Rajanen, D. (2018). Heuristic evaluation in game and gamification development. In *Proceedings of the 2nd GamiFIN Conference*.
- Zhu, M., Zhao, F., Fang, X., & Moser, C. (2017). Developing Playability Heuristics Based on Nouns and Adjectives from Online Game Reviews. *International Journal of Human-Computer Interaction*, 33(3), 241-253.