



# **Heuristinen evaluointi videopelien käytettävyyden arvioinnissa**

Oulun yliopisto  
Tietojenkäsittelytieteiden laitos  
LuK-tutkielma  
Urpo Lankinen  
28.3.2015

## Tiivistelmä

Heuristinen evaluointi on ohjelmistojen käytettävyyden arviointimenetelmä, jossa asiantuntijat vertaavat ohjelmistoa olemassa oleviin yleisesti tunnistettuihin periaatteisiin, eli heuristiikkoihin, erilaisien käytettävyysongelmien löytämiseksi. Tässä kirjallisuuskatsauksessa keskitytään videopelien asettamiin erityisiin haasteisiin käytettävyyssarvioinnin kannalta muihin ohjelmistoihin verrattuna, videopelejä koskeviin käytettävyyshauristiikkoihin ja niiden kehityshistoriaan.

### *Avainsanat*

videopelit, tietokonepelit, käytettävyys, heuristinen evaluointi, heuristiikat

### *Tutkielman ohjaaja*

Netta Iivari

## Alkusanat

Käytettävyys on ohjelmistokehityksessä vaikeasti lähestyttävä aihe. Tietojenkäsittely pohjautuu matematiikkaan ja logiikkaan, ja parhaat ohjelmistot pohjaavat todistetusti toimiviin algoritmeihin ja suunnitteluratkaisuihin. Moni aloittava ohjelmistokehittäjä kuitenkin havaitsee hyvin nopeasti, että ohjelmistokehityksestä ja järjestelmien suunnittelusta ei tule mitään ilman ihmistuntemusta – ja ihmisten käyttäytymistä ei ennusteta puhtaalla logiikalla. Tämä ei kuitenkaan tarkoita, etteikö ihmisten käyttäytymisessä olisi sääntöjä tai johdonmukaisuuksia.

Videopelit ovat ohjelmistoista usein haastavimpia: hyvän pelin tekemiseen tarvitaan paitsi hyvin suunniteltu ohjelmalogiikka myös taiteellisesti miellyttävä sisältö. Onko siis sääntöjä, joiden seuraamisella voidaan varmistaa lopputulokseksi hyvin toimiva peli, ilman että muista osatekijöistä joudutaan tinkimään? Kliinisen objektiivisesti ”hyväksi havaittuja” pelejä tuskin on olemassa, mutta ainakin pelin *toimivuuden* asteesta voidaan muodostaa jonkinlainen mielipide, jonka useimmat kehittäjät ja pelaajat voivat yksin tuumin allekirjoittaa, jottei huonon pelisuunnittelun historia jälleen kerran toista itseään. Tämä onkin heuristisen evaluoinnin suurin tarkoitus.

Työn avustamisessa erityiset kiitokset kuuluvat ohjaaja Netta Iivarille sekä Johdatus tutkimustyöhön -kurssin opettajalle Raija Haloselle – sekä kannustajina toimineille ystäville ja sukulaisille.

Urpo Lankinen

Oulu, 28. maaliskuuta, 2015

# Sisältö

1. Johdanto.....	5
2. Käytettävyyteen liittyviä määritelmiä.....	6
2.1 Käytettävyys ja käyttäjäkokemus.....	6
2.2 Hauskuus, pelattavuus ja pelikokemus.....	7
2.3 Halvat käytettävyyden arviointimenetelmät.....	7
2.4 Heuristiikat.....	8
2.5 Heuristinen evaluointi.....	9
3. Käytettävyys peleissä.....	10
3.1 Pelien erot muihin sovellusohjelmiin.....	10
3.2 Käytettävyyden arviointi videopelien kehitysprosessissa.....	10
3.3 Uusien heuristiikkojen löytämisen merkitys pelikehitykselle.....	12
3.4 Heuristiikkojen tarve ja käyttökokemukset.....	12
3.5 Heuristiikkojen kritiikki ja kysymys validiudesta.....	13
4. Heuristiikat ja heuristinen evaluointi videopelien kontekstissa.....	15
4.1 T.W. Malone: varhaisimmat peliheuristiikat.....	15
4.2 The 400 Project.....	16
4.3 Federoff: ensimmäiset modernit videopeliheuristiikat.....	16
4.4 Heuristic Evaluation of Playability (HEP) ja Heuristics of Playability (PLAY).....	17
4.5 Mobiilipelien heuristiikat.....	19
4.6 Pinelle ja kumppanit: peliarvostelut heuristiikkojen lähteenä.....	19
4.7 Kritiikkitilkitseminen.....	21
5. Yhteenveto.....	22

# 1. Johdanto

Käytettävyyden merkitys kasvaa tietotekniikan yleistyessä, ja nyky maailmassa niin käytettävyyden kuin videopelienkin merkitys on suuressa kasvussa. Jakob Nielsen luonnehti tilannetta näin: ”Henkilökohtaisten tietokoneiden vallankumous ja laitteiden hintojen putoaminen johtaa tietokoneiden leviämiseen laajemmille käyttäjryhmille, ja nämä käyttäjryhmät käyttävät tietokoneita suorittamaan entistä vaihtelevampia tehtäviä.” Nielsen näkee myös videopelit yhtenä käytettävyyden kehityksen veturina: ”Videopelit ja entistä paremmat henkilökohtaisten tietokoneiden ohjelmistot ovat todistaneet käyttäjille, että miellyttävien ja lähestyttävien käyttöliittymien tuottaminen on mahdollista, joten käyttäjät he eivät enää yhtä helposti hyväksy huonoja käyttöliittymiä.” (Nielsen, 1993, s. 8.)

Eri tyyppisten ohjelmistojen käytettävyyttä joudutaan usein arvioimaan eri tavoin. Videopelien luonteesta johtuen niiden käytettävyyttä joudutaan arvioimaan aivan eri lähtökohdista ja erilaisin menetelmin kuin muita ohjelmistoja; esimerkiksi videopelien elämyksellisestä, kokemuksellisesta ja yllättävästä luonteesta johtuen pelit eivät saavuta samanlaista yhtenäistä käytettävyyttä kuin sovellusohjelmat. (Livingston, Mandryk, & Stanley, 2010, s. 48; Pinelle, Wong, & Stach, 2008b, s. 129.) Lisäksi myös videopelien sisällä on suuria eroja: yksin- ja moninpelit asettavat omia haasteitaan (Korhonen & Koivisto, 2006, s. 15). Samoin eri peligenrejen välillä on suuria eroavaisuuksia, mutta toisaalta peligenrejen sisäinen yhtenäisyys houkuttaa kehittäjiä yhtenäistämään pelikokemusta (Livingston, Mandryk, & Stanley, 2010, s. 48).

Tutkimusongelmanani ja pääasiallisena kirjallisuuskatsauksen näkökulmana onkin videopelien käytettävyyden arviointimenetelmien suurimmat erot muihin ohjelmistoihin verrattuna. Sopivatko ohjelmistojen yleiset hyvän käytettävyyden suunnitteluperiaatteet myös pelien suunnitteluun, vai pitääkö pelien luonteesta johtuen ottaa huomioon erityisiä käytettävyyden näkökulmia videopelien suunnittelussa? Millaisia työkaluja pelikehittäjillä on pelien käytettävyyden arvioimiseksi?

Tässä kirjallisuuskatsauksessa esitellään pääasiassa *heuristista evaluointia*, joka on osoittautunut merkittäväksi pelien käytettävyyden arviointimenetelmäksi. Katsauksessa esitellään erilaisia tutkimuksissa löydettyjä heuristiikkoja, joita käytettävyydsiantuntijat voivat soveltaa pelien heuristiseen evaluointiin.

Suurin osa katsauksessa käytetyistä artikkeleista löydettiin ACM Digital Library (DL) -palvelusta. Artikkeleita haettiin mm. hakusanoilla ”game”/”video game”, ”usability”, ”heuristics”. Ensimmäisten lupaavien artikkelien jälkeen etsin ACM DL:stä, Google Scholarista ja muista tietokannoista kyseisissä artikkeleissa mainittuja lähteitä. Osa määritelmiä koskevista lähteistä on löydetty tutkimalla Wikipedian käytettävyydsartikkeleissa mainittuja lähteitä.

Luvussa 2 esitellään käytettävyyttä yleisesti ja käytettävyyteen liittyviä erilaisia käsitteitä. Luvussa 3 esitellään pelien käytettävyyden piirteitä ja pelien asettamia erityishaasteita käytettävyydelle. Luvussa 4 esitellään heuristista evaluointia pelien arviointimenetelmänä; luvussa esitellään erilaisia merkittäviä kirjallisuudessa mainittuja heuristiikkoja, joiden avulla pelien käytettävyyttä voidaan arvioida varhaisessa kehitysvaiheessa. Luvussa 5 on yhteenveto sisällöstä ja analyysiä eri menetelmistä.

## 2. Käytettävyyteen liittyviä määritelmiä

Tässä luvussa esellään käytettävyyteen liittyviä määritelmiä yleisellä tasolla.

### 2.1 Käytettävyys ja käyttäjäkokemus

ISO:n standardi 9241 käsittelee tietokonejärjestelmien ergonomiaa, ja sen osa 210, vahvistettu marraskuussa 2010, esittelee käyttäjäkeskeisen suunnittelun periaatteet (Suomen Standardoimisliitto SFS [SFS], 2010). Vuonna 2010 julkaistu ISO 9241-210 -standardi korvaa vanhan ISO 9241:n määritelmiin keskittyneen osan ISO 9241-11 vuodelta 1998 ja ISO 13407 -standardin vuodelta 1999. Standardin suomennos valmistui vuonna 2011.

ISO:n määritelmän mukaan käytettävyys on ”mitta, miten hyvin määrätyt käyttäjät voivat käyttää järjestelmää, tuotetta tai palvelua tietyssä käyttötilanteessa saavuttaakseen määritetyt tavoitteet tuloksellisesti, tehokkaasti ja tyytyväisinä” (SFS, 2010, s. 16).

Jakob Nielsen on yksi siteeratuimpia käytettävyyden asiantuntijoita. Nielsenin (1993, s. 26) mukaan käytettävyys on moniulotteinen tietokonejärjestelmän ominaisuus, johon kuuluu useita eri attribuutteja: opittavuus, tehokkuus, muistettavuus, virheettömyys ja tyytyväisyys (learnability, efficiency, memorability, errors, satisfaction).

- *Opittavuus* tarkoittaa, että käyttäjä voi omaksua tarvitsemansa järjestelmän ominaisuudet nopeasti.
- *Tehokkuus* tarkoittaa käyttäjän tekemien asioiden nopeaa suorittamista sen jälkeen kun hän on oppinut tarvittavat toimintatavat.
- *Muistettavuus* tarkoittaa käyttäjän aiemmin oppimien toimintojen nopeaa mieleen palautumista, jos käyttäjä ei ole käyttänyt järjestelmää tai sen tiettyä harvoin tarvittavaa osaa pitkään aikaan.
- *Virheettömyys* tarkoittaa, että käyttäjän normaalit toiminnot eivät johda helposti virhetilanteisiin, että käyttäjä ymmärtää järjestelmän antamat virheilmoitukset ja että käyttäjä voi helposti palautua virhetilanteista. Käyttäjän toimet eivät saa myöskään johtaa katastrofaalisiin virheisiin.
- *Tyytyväisyys* tarkoittaa, että käyttökokemus kokonaisuudessaan miellyttää käyttäjää.

Käytettävyys riippuu eri käyttäjistä ja käyttötilanteista, joten sen käytännönläheinen arviointi onnistuu pelkästään määrittelemällä käyttäjäryhmä ja nimenomaiset halutut tehtävät, ja pohjaamalla käytettävyytutkimus suureen joukkoon koehenkilöitä (Nielsen, 1993, s. 27).

Käytettävyys on osa järjestelmien hyväksyttävyyden (system acceptability) kenttää. Käytettävyys edesauttaa ohjelman hyödyllisyyttä (usefulness), jonka osana on myös ohjelmistojen yleinen soveltuvuus tehtävään (utility). Hyödyllisyys on yksi

järjestelmien käytännöllisistä hyväksyttävyysskriteereistä (practical acceptability), joihin kuuluvat myös hinta, luotettavuus ja muut samantapaiset kriteerit. Käytännöllisen hyväksyttävyyden lisäksi toinen järjestelmien hyväksyttävyysskriteeri on sosiaalinen hyväksyttavuus. (Nielsen, 1993, s. 24-25.)

Käytettävyys on myös osa laajempaa käyttäjäkokemuksen (user experience) käsitettä. ISO määrittelee käyttäjäkokemukseksi ”henkilön havainnot ja vasteet, jotka ovat seurausta tuotteen, järjestelmän tai palvelun käytöstä ja/tai ennakoidusta käytöstä”. Lisähuomautuksen mukaan ”käytettävyys käyttäjien henkilökohtaisten tavoitteiden näkökulmasta tulkittuna voi sisältää sen tyyppisiä aisti- ja tunnenäkökulmia, joita tyyppillisesti liitetään käyttäjäkokemukseen. Käytettävyyskriteereitä voidaan käyttää arvioimaan käyttäjäkokemuksen joitakin näkökulmia.” (SFS, 2010, s. 16.)

Useimmat tässä kirjallisuuskatsauksessa mainitut artikkelit eivät anna pelien käytettävyydelle erityistä määritelmää. Pinelle, Wong ja Stach (2008a, s. 1453; 2008b, s. 130) määrittelevät pelien käytettävyyden ”määräksi, jossa pelaaja voi oppia, kontrolloida ja ymmärtää peliä”.

## 2.2 Hauskuus, pelattavuus ja pelikokemus

Varhaisissa pelien käytettävyyden arvioinnissa käsiteltiin hyvin epämääräisesti määriteltävää pelien ”hauskuutta”. Esimerkiksi Malone (1982) ei johdannossaan keskity käytettävyyteen vaan nimenomaan miellyttävyyteen: pelejä pidetään mielenkiintoa herättävinä (”captivating”) ja hauskoina (”fun”), ohjelmien käyttöliittymiä miellyttävänä käyttää (”enjoyable to use”). Federoff (2002) mainitsee jo työn otsikossaan hauskuuden (”fun”), mutta antaa kuitenkin kirjallisuuskatsauksessaan kattavan kuvauksen pelien hauskuuden aikaisemmista määritelmistä (Federoff, 2002, s. 8-10).

Yleisessä kielenkäytössä pelien arvioinnissa käytetään usein termiä pelattavuus (playability). Pelien käytettävyyden tutkimuksessa termillä on useita eri määritelmiä, jotka usein ovat merkitykseltään laajempia kuin pelkkä käytettävyys. Esimerkiksi Karvonen (2005, s. 7) vertaa pelattavuutta ja käytettävyyttä käsitteinä: ”Pelien kohdalla käytettävyyttä parempi termi näyttääkin olevan pelattavuus (playability). *Pelattavuudella* viitataan pelin hyvyyteen kokonaisvaltaisesti, mutta käsitteenä se on kuitenkin vielä epämääräinen ja jäsentymätön.” Karvonen käsittelee pro gradu -työssään pelattavuutta ja pelikokemusta käsitteenä hyvin monelta suunnalta.

Desurvire, Caplan ja Toth (2004) eivät sinänsä anna artikkelissaan suoraa määritelmää pelattavuudelle, siitäkin huolimatta että ”playability” on mainittu jo otsikossa. Nacke ja kumppanit (2009) esittävät, että pelattavuuden arviointi käsittelee yleensä pelin onnistuneisuutta pelisuunnittelun näkökulmasta. Toinen näkökulma on pelikokemus (player experience), joka vertautuu yleisesti käyttäjäkokemukseen. ”Pelattavuus arvioi pelejä suunnittelun parantamiseksi, kun taas pelikokemuksen menetelmät arvioivat pelaajia pelaamisen parantamiseksi”. Toisaalta ”hyvä pelattavuus on edellytys pelikokemuksen arvioimiseksi. Pelin suunnittelussa ei saisi olla ongelmia, jotka menevät yksityisen pelikokemuksen eteen.”

## 2.3 Halvat käytettävyyden arviointimenetelmät

Nielsenin johdolla käytettävyyden arviointiin tuotiin halvat käytettävyyden arviointimenetelmät (*discount usability*), joiden etuna on, että ohjelmistokehittäjät voivat oppia uusia heuristiikkoja ja niiden soveltamista nopeasti esimerkiksi kirjallisuudesta tai lyhyistä seminaareista (Nielsen, 1994, s. 25). Näihin halpisiin

käytettävyyden arviointitekniikoihin kuuluvat heuristisen evaluoinnin lisäksi käyttöskenaarioiden arviointi paperiprototyyppien avulla ja yksinkertaistettu, ääneen ajatteluun pohjaava pienen ryhmän käyttäjätestaus (Nielsen, 1993, s. 17-20). Heuristisen evaluoinnin ja muiden halpojen käytettävyyden arviointimenetelmien ”ei ole tarkoituskaan tuottaa 'täydellisiä' ratkaisuja, eikä löytää kaikkia mahdollisia käyttöliittymien suunnitteluvirheitä” (Nielsen, 1993, s. 160).

Nielsen esitteli halvat käytettävyyden arviointimenetelmät ensimmäisen kerran konferenssiartikkelissa syyskuussa 1989, vaihtoehdoksi – muttei korvaajaksi – ohjelmistoalalla vallalla olleelle kattavalle mutta erittäin työläälle ja kalliille kvantitatiiviselle käytettävyyden arvioinnille. Vuonna 2009, 20 vuotta tämän alkusysäyksen jälkeen, Nielsen valitteli halpojen käytettävyyden arviointimenetelmien laiminlyöntiä ohjelmistoteollisuudessa – esimerkiksi Nielsenin mukaan heuristiikkojen ja käyttöliittymien suunnitteluohjeistojen merkitys on kasvanut, mutta niitä ei edelleenkään sovelleta tarpeeksi ahkerasti. (Nielsen, 2009.)

## 2.4 Heuristiikat

Nielsenin mukaan *heuristiikat* ovat ”tunnistettuja käyttöliittymien suunnitteluperiaatteita” (Nielsen, 1993, s. 155; Nielsen, 1994, s. 26) ja ”yleisiä sääntöjä, jotka näyttävät kuvaavan käyttöliittymien yhteisiä ominaisuuksia” (Nielsen, 1994, s. 28).

Nielsen (1993, s. 19) mainitsee, että käytössä olevia erilaisia käytettävyyshauristiikkoja ja suunnitteluperiaatteita on tuhansia, ja luettelee kymmenen perushauristiikkaa, joiden avulla voidaan löytää suurin osa yleisistä ohjelmistojen virheistä. Nielsen tuotti luettelon Rolf Molichin kanssa vuonna 1990. Päivitetty luettelo (Nielsen, 1994, s. 30) pohjautui Nielsenin jatkotutkimukseen 249 käytettävyyshauristuksesta.

- Yksinkertainen ja luonnollinen vuorovaikutus [1993] / esteettinen ja minimalistinen suunnittelu [1994]: Ohjelman pitää kertoa käyttäjälle kaikki mitä hänen tarvitsee tietää käyttötilanteessa, muttei hämmentää häntä liiallisella ja vaikeasti ymmärrettävällä tiedolla.
- Puhu käyttäjän kieltä [1993] / järjestelmän ja tosimaailman yhtenäisyys [1994]: Ohjelman käyttämien termien ja ilmaisujen pitää olla käyttäjän ymmärrettävissä, ja käyttöliittymässä pitää välttää käyttäjän asiantuntemuksen ulkopuolelta tulevia termejä ja ilmaisuja.
- Minimoi muistettavien asioiden määrä [1993] / tunnistaminen muistamisen sijaan [1994]: Käyttäjän ei pitäisi tarvita muistaa itse asioita työvaiheesta toiseen, vaan ohjeiden ja aikaisempien tietojen pitää olla suoraan käytettävissä oikeassa työvaiheessa.
- Yhdenmukaisuus [1993] / yhdenmukaisuus ja standardit [1994]: Käyttäjän pitää pystyä luottamaan, että komennot ja ohjelman termit tarkoittavat joka tilanteessa samaa. Ohjelman pitää noudattaa käyttöjärjestelmän käytäntöjä.
- Palaute [1993] / ohjelmiston tilanteen selkeys [1994]: Käyttäjän pitää aina tietää, mitä ohjelma kulloinkin tekee, ja kuulla välittömästi palautetta virheistä.
- Selkeät poistumistiet [1993] / käyttäjän valta ja vapaus [1994]: Jos käyttäjä päättää perua toiminnon, hänen pitää päästä helposti pois aikaisemmasta tilanteesta. Ohjelman tulee tukea kumo- ja tee uudelleen -toimintoja.



- Oikotiet [1993] / joustavuus ja käytön tehokkuus [1994]: Aloittelijoille tulee antaa yksinkertainen käyttöliittymä. Edistyneille käyttäjille tulee tarjota oikoteitä ja pikatoimintoja, ja mahdollisuus muokata niitä oman työnteon helpottamiseksi.
- Hyvät virheilmoitukset [1993] / auta käyttäjää tunnistamaan, diagnosoimaan ja palautumaan virheistä [1994]: Selkeäkieliset ilmoitukset jotka kertovat virheen yksityiskohdat tarkasti ja ehdottavat ratkaisuja. Pelkkien virhekoodien käyttöä tulee välttää.
- Virheiden ennaltaehkäisy: Ohjelmiston selkeä rakenne ja selkeät käyttötavat ennaltaehkäisevät virhetilanteita. Käyttäjää pitää varoittaa ennen peruuttamattomien muutoksien tekemistä.
- Ohjeet ja dokumentaatio: Parhaiten suunniteltuja ohjelmistoja voi käyttää ilman opastusta, mutta laadukas, helposti selattava ja helposti haettava ohjeisto helpottaa käyttöä ongelmatilanteissa.

Nielsenin vuoden 1994 heuristiikoista ja niiden historiasta on saatavilla myös verkkoartikkeli (Nielsen, 1995).

Heuristiikoilla on myös yhteys käyttöliittymien *suunnittelusääntöihin* (guidelines) ja tätä kautta myös käyttöliittymiä koskeviin *standardeihin*. Erilaiset käyttöliittymien suunnittelusäännöt voivat toimia pohjana heuristiikoille. (Nielsen, 1993, s. 92.)

## 2.5 Heuristinen evaluointi

Nielsen (1993, s. 155) kuvailee vapaamuotoista heuristista evaluointia yksinkertaisimmallaan prosessiksi, jossa käyttäjät tutkivat käyttöliittymiä ja kertovat, mikä käyttöliittymässä on hyvää ja huonoa. Muodollinen käytettävyyssarvioinnissa käytettävä heuristinen evaluointi on puolestaan prosessi, jossa tarkoituksena on systemaattisesti havaita, mitkä osat käyttöliittymästä ovat tai eivät ole olemassa olevien heuristiikkojen mukaisia.

Nielsen (1993, s. 155-160; 1994, s. 25-35) kuvaa heuristisen evaluoinnin suoritustavan peruspiirteet: arvioinnin suorittaa evaluoija (evaluator), tavallisimmin useampi evaluoija yhdessä; kattava heuristinen evaluointi vaatii monen evaluoijan yhteistyötä, koska eri evaluoijilla on taipumuksena havaita erilaisia virheitä. Erona esimerkiksi ääneen ajatteluun pohjaavaan käytettävyydestäukseen, jossa normaalia käyttäjäkuntaa edustavien käyttäjien käytöstä, reaktioita ja ajatusprosessia havainnoidaan mahdollisimman luonnollisesti, evaluoinnissa voi olla osallisena myös havainnoija (observer) eli koevastaava (experimenter), joka voi ohjeistaa ja avustaa evaluoijia. Tämä voi olla esimerkiksi tarpeen, jos prototyyppi ei toimi oikein tai se on tavalla tai toisella puutteellinen. Evaluoijien kuitenkin tulee suorittaa evaluointi yksitellen, jottei heidän palautteensa vaikuta toisien evaluoijien havaintoihin ja havainnointiprosessiin. Evaluoijat tutkivat käyttöliittymää yksi elementti ja heuristiikka kerrallaan, useimmiten monta kertaa, ja raportoivat mitä heuristiikkoja ei ole sovellettu oikein. Nämä havainnot kootaan lopuksi yhteen kirjalliseksi raportiksi, tai evaluoija voi kertoa niistä suullisesti koetilanteen havainnoijalle. Yksi evaluointikierron voidaan suorittaa muutamassa tunnissa.

### 3. Käytettävyys peleissä

Tässä luvussa esitellään mitä erityisiä haasteita pelit asettavat käytettävyyden arvioinnille, käytettävyydestä osana pelien kehitysprosessia, ja miten pelien käytettävyyden arviointi poikkeaa sovellusohjelmien käytettävyyden arvioinnista. Lisäksi tutkitaan pelien käytettävyyden arvioinnin hyödyllisyyttä pelikehitykselle.

#### 3.1 Pelien erot muihin sovellusohjelmiin

Yksi pelien käytettävyyden ongelmista on, että videopelit eroavat monessa määrin tavallisista sovellusohjelmista. Jo Malone (1982, s. 65) teki eron ”lelujen” ja ”työkalujen” välillä: ”leluja” käytetään itseisarvon vuoksi ilman päämäärää, ”työkaluja” ulkoisien päämäärien saavuttamiseksi.

Federoff (2002, s. 5) esimerkiksi toteaa ISO:n käytettävyyden määritelmästä seuraavaa: ”ISO 9241-11:n käytettävyyden määritelmään kuuluu kolme itsenäistä määrettä: tuloksellisuus, tehokkuus ja tyytyväisyys. Videopelien käytettävyyden tapauksessa tuloksellisuus ja tehokkuus ovat toissijaisia tyytyväisyyteen verrattuna. Kuluttajat ostavat ohjelmistoja suorittaakseen tehtäviään tarpeidensa mukaan, mutta peli ostetaan vapaaehtoisesti puhtaasti viihdearvon takia. Jos peli ei ole hauska pelata, se ei menesty markkinoilla.”

Pinelle, Wong ja Stach (2008b, s. 129) huomauttavat, että peleillä on tiettyjä hyvin perustavanlaatuisia eroja sovellusohjelmiin ja sovellusohjelmien suunnitteluperiaatteisiin: ”Keskittyminen uutuuteen yhdenmukaisuuden sijaan aiheuttaa sen, että pelit eivät yleensä käytä ikkunakäyttöliittymiä ja tavanomaisia käyttöliittymäkirjastoja, jotka määräävät normaalien sovellusten ulkoasun ja toiminnallisuuden (*look and feel*). Tämä on johtanut huomattavaan pelien käyttöliittymien kirjavuuteen, jossa voidaan havaita hämmästyttäviä eroja eri suunnittelujen välillä niin visuaalisissa kuin interaktiivisissakin näkökulmissa.”

Samoin Pinelle, Wong ja Stach (2008a, s. 1454) toteavat eri käytettävyyсарviointimenetelmistä: ”Tekniikat kuten kognitiivinen läpikäynti, pluralistinen läpikäynti ja tehtäväanalyysi pohjautuvat osittain olettamukselle, että käyttäjät haluavat suorittaa sovelluksella ennalta määrättyjä tehtäviä. Toimintosarjojen ajatus ei välttämättä kuitenkaan ole hyödyllinen peleissä, koska ihmiset pelaavat niitä eri tavoin riippuen strategiasta ja motivaatiosta. Samoin jotkin pelit kehottavat pelaajia pelin päämäärättömään tutkimiseen, joten pelaajien käyttöliittymän käytössä voi olla huomattavia eroja.”

#### 3.2 Käytettävyyden arviointi videopelien kehitysprosessissa

Monentyypinen testaus on merkittävä osa videopelien kehitysprosessia. Lahti (2014) käsittelee pro gradu -työnsä viidennessä luvussa pelien testausprosessia lähinnä laadunvalvonnan näkökulmasta, ja näkee pelien laadun avaintekijänä koko julkaisuprosessissa: huonolaatuiset pelijulkaisut kustautuvat kehittäjille taloudellisina tappioina ja imagon tahriutumisenä. (Lahti, 2014, s. 21.)

Käytettävyydestä on vain yksi ohjelmistojen testauksen muoto. Schell (2014, s. 434) erottaa neljä pelejä koskevaa testausprosessia:

- Kohderyhmättestaus (focus groups) jossa kerätään pelien kohderyhmään kuuluvilta pelaajilta mielipiteitä lähinnä liiketoiminta- ja markkinointinäkökulmasta.
- Laadunvalvontatästäus (QA testing) jossa etsitään peleissä olevia ohjelmointivirheitä.
- Käytettävyydestäus (usability testing), jossa keskitytään pelien käytettävyyteen.
- Pelitestäus (playtesting), jossa keskitytään pelin sisältöön.

Schellin määritelmän mukaan pelitestäus keskittyy pelien suunnitteluun taiteellisesta näkökulmasta, ja siihen liittyy kysymyksiä kuten ”mikä pelissä on hauskinda”, ”mikä pelissä on tylsintä” ja ”ovatko pelaajat koskaan hämmentyneitä pelin aikana” (Schell, 2014, s. 436). Samoin pelitestäuksen suorittajiksi Schell suosittelee kehittäjiä, kehittäjien ystäviä, hyviä pelaajia ja ”nenäliinoja” (tissue testers) – ”kertakäyttöisiä” pelaajia jotka kokeilevat peliä tietämättä siitä mitään etukäteen ja raportoivat ensivaikutelmistaan (Schell, 2014, s. 437-438). Levy ja Novak (viitattu lähteessä Lahti, 2014, s. 21-22) esittävät, kuinka laadunvalvontatästäuksessa puolestaan tarvitaan erillinen työryhmä: pelin tuottajan alaiseen tiimiin kuuluu usein laadunvalvontajohtaja (QA manager), päätestäaja (lead tester), mahdollinen apulaispätestäaja (floor lead) sekä varsinaiset testajat.

Erilaisia testausmenetelmiä käytetään pakostakin eri osissa kehitysprosessia. Schultz ja Bryant (viitattu lähteessä Lahti, 2014, s. 27-31) sekä Chandler (viitattu lähteessä Lahti, 2014, s. 27-31) esittelevät pelitestäuksen laadunvalvonnallisia menetelmiä eri kehitysvaiheissa (esituotanto, alfaversion, beetaversion, lopullinen versio ja julkaisun jälkeiset versiot).

Aivan kuten laadunvalvonnankin tapauksessa, myös eri käytettävyyden arviointimenetelmät sopivat eri tavoin eri osiin kehitysprosessia, ja organisaatioon vaaditaan sitä varten erityistä henkilökuntaa – heuristisen evaluoinnin tapauksessa evaluoijana toimivat asiantuntijat ovat usein erillään kehittäjistä (Nielsen, 1993, s. 155-160; 1994, s. 25-35).

Heuristinen evaluointi soveltuu pääasiassa varhaiseen kehitysvaiheeseen, jossa se on osoittautunut suosituksi. Pelikehityksessä käytetään monia muita käytettävyyden arviointimenetelmiä, varsinkin myöhemmissä kehitysvaiheissa. Heuristista evaluointia ei usein voida käyttää yksinomaisten käytettävyyden arviointimenetelmänä pelien kehitysprosessissa. Tästä rajoituksesta huolimatta sillä on etuja muihin menetelmiin: ”Pelitestäus on yksi yleisimmistä pelien suunnitteluongelmien löytötavoista, mutta tämä menetelmä vaatii pelattavissa olevan prototyypin, joka on valmiina vasta myöhemmissä pelikehityksen vaiheissa. [...] Toisin kuin pelitestäuksen tapauksessa, [käytettävyyden arviointimenetelmät] eivät tarvitse käyttäjien panosta ja niitä voidaan käyttää varhaisen mockupien kuten myös toimivien prototyyppien kanssa.” (Pinelle, Wong, & Stach, 2008b, s. 130.)

Muut arviointimenetelmät tukevat kuitenkin heuristiikkoihin nojaavaa arviointia. Esimerkiksi Desurvire, Caplan ja Toth (2004, s. 1512) toteavat yhteenvetonaan: ”HEP on hyödyllinen varhaisessa pelinkehityksen vaiheessa [...] Kuitenkin joudumme nojaamaan käyttäjätestäukseen, koska riippumatta siitä kuinka hyvin ymmärrämme

pelien pelaajia ja ihmisiä, heidän käytöksensä on silti ennalta arvaamatonta. Havainnoituamme pelaajien käyttäytymistä saamme spesifistä tietämystä suunnitteluongelmien ratkaisemiseksi.”

### 3.3 Uusien heuristiikkojen löytämisen merkitys pelikehitykselle

Nacke ja kumppanit (2009) mainitsevat, että heuristinen evaluointi on yksi pelattavuuden parantamisen työkaluista. Vaikka esimerkiksi Nielsen (1993) havaitsi useita yleiskäyttöisiä heuristiikkoja, pelitutkijat ovat huomanneet jo muiden muassa edellä mainituista syistä eri tyyppisten sovelluksien tarvitsevan erilaisia heuristiikkoja. Tutkimuksen historian tilaa kuvaa se, että Nielsen ei luvussa 1 mainitsemaani esimerkkiä lukuun ottamatta mainitse vuoden 1993 kirjassaan videopelejä erikseen.

Desurvire, Caplan ja Toth (2004, s. 1509) mainitsevat, että Nielsenin (1994) kohteena olivat pääasiassa sovellusohjelmat, ja että vastaavien heuristiikkojen kehittäminen nimenomaan pelituotantoa varten on tarpeen. Artikkelitoteaa, että heuristiikkoja on kehitetty niin peliteollisuudessa kuin käytettävyyssiantuntijoidenkin keskuudessa. Samoin Pinelle, Wong ja Stach (2008a, s. 1454) toteavat Nielsenin heuristiikoista: ”Ne viittaavat normaaleihin käyttöliittymäkonsepteihin, kuten dialogeihin, kumoamiseen ja uudelleen tekemiseen, ja virheiden estämiseen. Näistä käsitteistä on kuitenkin vähän hyötyä pelien kontekstissa.”

Livingston, Mandryk ja Stanley (2010, s. 48) toteavat: ”Heuristiset periaatteet on usein suunniteltu yleistettäväksi erilaisille sovelluksille, mutta ne eivät ota huomioon pelityyppien, pelien ohjaustapojen ja pelien käyttöliittymien erilaisuuksia. Pelien kirjavuus on osoitus luovuudesta, mutta johtaa myös suureen määrään käytettävyyssvaatimuksia, joten tietyt heuristiikat eivät välttämättä sovellu kaikkiin peleihin.”

### 3.4 Heuristiikkojen tarve ja käyttökokemukset

Uusien kehitysmenetelmien löytäminen ei luonnollisestikaan sinällään riitä, vaan pelien kehittäjät tarvitsevat myös jonkinlaista osoitusta uusien menetelmien hyödyllisyydestä. Laitinen (2006) tutki heuristista evaluointia osana Shadowgrounds -pelin kehitysprosessia. Asiantuntijat evaluoivat pelin käytettävyyttä laboratoriossa, ja pelin kehittäjät pystyivät seuraamaan evaluointia. Tutkimukseen kuului myös laboratoriossa järjestetty koe, jossa sovellettiin käytettävyytestausmenetelmiä. Kehittäjät saivat evaluoijilta palautetta peliin liittyen, ja kuukauden kuluttua kehittäjiltä kysyttiin, miten palaute vaikutti pelin jatkokehitykseen. 43% havaituista ongelmista oli uusia, ja 74% oli kehittäjien mukaan sen tyyppisiä ongelmia, että he joko eivät välttämättä olisi niitä löytäneet ilman asiantuntijoiden apua, tai ongelmien löytäminen olisi voinut kestää pitempään.

Peliä evaluoitiin kuten mitä hyvänsä muuta ohjelmaa: tutkimuksessa ei käytetty nimenomaan videopelien evaluointiin tarkoitettuja heuristiikkoja, vaan Nielsenin (1993) perusheuristiikkoja ja evaluoijien omaa käytettävyyssiantuntemusta. Evaluoijia ei myöskään käsketty keskittymään nimenomaan videopelien ominaispiirteisiin.

Mielenkiintoisena havaintona Laitisen tutkimuksessa oli, että evaluoijien videopelien pelaamistiedoilla ja -taidoilla ei ollut merkitystä: pelejä harrastamattomat käytettävyyssiantuntijat löysivät yhtä tehokkaasti kehittäjien mielestä merkittäviä ongelmia kuin pelien ja käytettävyyden ”kaksoisiantuntijat”. Tästä on täysin päinvastaista kokemusta sovellusohjelmien puolella: Nielsen (1993, s. 161) viittasi

aiempiin tutkimuksiin, joissa sovellusalueen ja käytettävyyden kaksoisasiantuntijat löysivät ohjelmistojen ongelmia huomattavasti tehokkaammin kuin puhtaat käytettävyydsiantuntijat.

Yksi suurista ongelmista heuristiikkojen kehittämisessä on uusien heuristiikkojen käytännönläheisyys ja validius. Laitinen (2006, s. 65) keskittyi erityisesti ilmivaliditeettiin (*face validity*): ”Yksi mahdollinen syy käytettävyydmetodien hylkäämiseen voi olla, että pelikehittäjät, pelien tuottajat ja markkinointiosasto voivat joko epäillä käytettävyysevaluointimenetelmien hyödyllisyyttä tai olla tietämättömiä niistä. [...] Jos kehittäjät luulevat että käytettävyyden asiantuntija-arviointi ja testaus ei tuota uutta ja hyödyllistä tietoa, he eivät käytä näitä menetelmiä. Sama koskee myös menetelmillä löydettyjä ongelmia. Jos pelikehittäjät eivät pidä löydettyjä ongelmia uskottavina, he eivät korjaa niitä.” Laitinen esitti, että heuristiikkojen soveltuvuuden ja käyttökelpoisuuden arviointi voi olla hyvä ilmivaliditeetin mittari. Laitinen kuvaa heuristiikkojen ilmivaliditeettia ”karkeaksi” arvoksi, mutta kuitenkin arvoksi joka on hyvin kiinnostava jatkotutkimuksia silmällä pitäen.

Myös Paavilainen, Alha ja Korhonen (2012) kokeilivat heuristiikkojen toimivuutta käytännössä. He viittasivat Paavilaisen (2010) aiempaan tutkimukseen sosiaalisen median parissa, joskaan artikkeli ei käsitellyt Paavilaisen aiempaa luonnosta sosiaalisen median pelien heuristiikoiksi. Tutkimuksen kohteena olivat noviisievaluoiijat ja olemassa olevien yleisheuristiikkojen soveltuminen suppeaan kohdealueeseen – toisin sanoen yleisien pelien suunnitteluheuristiikkojen pätevyys erityisesti sosiaalisen median peleissä. Tekijöiden mukaan yleisluontoisten heuristiikkojen soveltaminen sosiaalisen median peleihin voi auttaa löytämään myös erityisesti sosiaalisen median pelejä koskevia ongelmia. (Paavilainen, Alha, & Korhonen, 2012, s. 336-337.) Paavilaisen aikaisempia sosiaalisen median pelien heuristiikkoja ei käytetty, koska ne eivät sellaisenaan soveltuneet noviiseille. Uusi heuristiikkajoukko pohjautuikin Korhosen ja Koiviston malliin (Korhonen & Koivisto, 2006, 2007), pois lukien mobiiliheuristiikat. (Paavilainen, Alha, & Korhonen, 2012, s. 338-339.) Tutkimuksen noviisievaluoiijina toimivat yliopisto-opiskelijat, ja lisäksi tutkimukseen osallistui käytettävyydsiantuntijoista koostuva ”metaevaluoiijien” ryhmä. Asiantuntijoiden johtopäätökset poikkesivat usein noviisien johtopäätöksistä.

Tässä tutkimuksessa havaittiin, että yleiset heuristiikat pätevät myös sosiaalisen median peleihin. Esimerkiksi evaluoiijat kokivat sosiaalisen median peleissä yleiset popup-mainokset lähinnä ärsyttävinä; tämä soti heuristiikkaa ”navigointi on johdonmukaista, loogista ja minimalistista” vastaan (Paavilainen, Alha, & Korhonen, 2012, s. 343-344). Samoin pelistä sai sitä enemmän irti, mitä useampi pelaajan ystävistä pelasi peliä; ”Peli mahdollistaa myös yksinäisen/omatoimisen pelaamisen” -heuristiikka soti tätä vastaan, koska on mahdollista, että kaikkien pelaajien ystävät eivät kiinnostu pelistä, joten pelikokemus jää vajaaksi (Paavilainen, Alha, & Korhonen, 2012, s. 346).

### 3.5 Heuristiikkojen kritiikki ja kysymys validiudesta

Pelien arviointi tieteellisin menetelmin on usein vaikeaa. Esimerkiksi Karvonen (2005, s. 7) toteaa pelattavuudesta ja aikaisemmista kirjallisuudessa mainituista heuristiikoista: ”Pelattavuutta mittaavia kriteeristöjä on tähän mennessä kehitetty vain jokunen [...] ja niidenkin käytännön sovellettavuus on kyseenalaista.”

Nielsen (1993, s. 165-170) esittää kaksi käytettävyydestä yleistä koskevaa metodologista ongelmaa: luotettavuus (reliability) ja validius (validity). Luotettavuus koskee testauksen tiedonkeruun tuottaman tiedon tarkkuutta: yksittäisten koehenkilöiden välillä on huomattavia eroja, joten pienestä otannasta ei voida tehdä

tilastollisesti merkittäviä johtopäätöksiä. Esimerkkinä Nielsen mainitsee tehtävien suoritusajat, jotka voivat vaihdella hyvinkin paljon henkilöstä riippuen, joten pienempien suoritusaikojen käyttäminen hyvän käytettävyyden mittarina ei ole aina mielekästä. Validius on suurempi kysymys, johon korkeimmassa mielessä liittyy koko tutkimuksen relevanssi – mitataanko tutkimuksella yleensäkin käytettävyyden kannalta merkittäviä asioita? – ja toisaalta suurempi kysymys siitä, onko alkuperäisillä oletuksilla todellisuudessa yhteyttä lopputuloksiin. Ajatuskokeena Nielsen mainitsee tutkimuksen, jossa vertaillaan tekstipohjaista ja graafista käyttöliittymää: ensimmäinen naiivi johtopäätös voi olla, että alkuperäisen oletuksen mukaisesti tekstipohjainen käyttöliittymä on käyttäjien mielestä huonompi kuin graafinen käyttöliittymä, vaikka todellisuudessa käyttöliittymän ulkoasulla ei ole merkittävää vaikutusta käyttäjien mielipiteeseen ja sen sijaan käyttäjät kokevat tärkeäksi sen, että suuriresoluutioiseen graafiseen näyttöön mahtuu enemmän käyttötilanteessa tarvittavaa tietoa.

Heuristista evaluointia koskevissa artikkeleissa keskeinen validiuden mittari on usein heuristiikkojen soveltuvuus käytäntöön myös alkuperäisen heuristiikkojen keräämistilanteen ulkopuolella. Esimerkiksi Pinelle, Wong ja Stach (2008a, s. 1459) validoivat keräämiään heuristiikkoja evaluoimalla peliä, joka ei kuulunut heuristiikkojen keräysvaiheessa tutkittuihin peleihin.

Artikkeleissa heuristiikat esitetään yleensä yksinkertaisesti opeteltavana ja sovellettavana taulukkona. Uusien heuristiikkojen tunnistamisen jälkeen niiden soveltuvuus ja käytännöllinen hyöty pelinkehitykselle pitää myös osoittaa. Tämä oli esimerkiksi Desurviren, Caplanin ja Tothin (2004, s. 1509) mukaan avainasemassa: ”Huolimatta peliheuristiikkojen kehityksestä, ennen tätä projektia pelien pelattavuuden heuristiikkoja ei ole koottu kattavaan luetteloon eikä niitä ole verifioitu millään tavoin.”

Paavilainen (2010), osana uusien sosiaalisen median pelien heuristiikkojen kehitystä, esitti kattavan vertailun monista aikaisemmista pelien käytettävyysheuristiikoista. Hän huomauttaa, että heuristiikkojen laatu on parantunut vuosien kuluessa. Paavilaisen arvion mukaan monet aiemmat heuristiikkalistat eivät juuri sovi Nielsenin halpojen arviointimenetelmien joukkoon, koska heuristiikat ovat joskus epäselviä, ja liian suuri heuristiikkojen joukko voi olla vaikeasti opittava ja muistettava. Paavilainen myös huomauttaa monen heuristiikkajoukon kohdalla, että kyseisiä malleja ei ole välttämättä validoitu riittävästi tutkimuksien julkaisuaikaan. Paavilaisen kriittisen kirjallisuuskatsauksen tarkoituksena oli pohjustaa uusien heuristiikkojen kehittämistä sosiaalisessa mediassa pelattavia pelejä varten. Ainoat tätä edeltävät Paavilaisen löytämät sosiaalisen median pelien heuristiikat oli käsitelty Aki Järvisen ja Tony Ventricon artikkeleissa. Näistä Paavilainen muodosti 10 alustavaa heuristiikkaa sosiaalisen median peleille.

## 4. Heuristiikat ja heuristinen evaluointi videopelien kontekstissa

Tässä luvussa esitellään merkittävimpiä kirjallisuudessa esiteltyjä videopelejä koskevia heuristiikkoja. Esimerkit on koottu aikajärjestykseen.

### 4.1 T.W. Malone: varhaisimmat peliheuristiikat

Malone (1982) mainitaan useissa tässä katsauksessa olevissa artikkeleissa varhaisimpana yrityksenä tuottaa nimenomaan pelejä koskevia heuristiikkoja. Esimerkiksi Federoff (2002, s. 5) toteaa: ”Ohjelmistoille on löydetty heuristiikkoja (Nielsen), mutta sikäli kuin havaintoni pitävät paikkansa, vain T.W. Malone on yrittänyt kehittää heuristiikkoja nimenomaan pelien omalaatuista ohjelmistokategoriaa varten.” Malonen vuoden 1982 artikkeliin on viitattu monissa muissakin katsauksen artikkeleissa, ja Federoff viittaa myös Malonen vuonna 1980 valmistuneeseen väitöskirjaan.

Malone esittelee, miten yksinkertaiseen tikanheittopeliin voidaan rakentaa erilaisia käyttöliittymiä. Suunnittelu alkoi abstraktista numeronarvauspelistä. Ensimmäinen versio ilmoitti yksinkertaisesti, oliko pelaajan arvaus oikein vai väärin. Seuraavissa versioissa peliin lisättiin hyödyllistä palautetta: peli antaa pisteitä oikeista arvauksista, ja lisäksi ilmoittaa, oliko numero oikea, liian suuri tai liian pieni. Viimeisissä versioissa peliin lisättiin konkreettisempaa ja mielikuvituksellisempaa sisältöä fantasiaelementtien muodossa. Peli muuttui tikanheittopeliksi, jossa numeroarvauksia visualisoidaan tikanheittoanimaatioilla ja pelaaja koettaa heittää tikkaa maalitaululla oleviin ilmapalloihin. Peliin luotiin myös sirkusilmapiiriä musiikin avulla. Kaikkiaan pelistä tehtiin 8 eri versiota, ja eri versioiden hauskuutta arvioitiin vertailemalla, kuinka kauaksi aikaa koululaiset innostuvat pelaamaan peliä verrattuna kaikissa kokeissa samana pysyneeseen hirsipuupeliin.

Malone ei esitä heuristiikkojaan suoraan käytettävissä olevana luettelona kuten monet myöhemmät tutkijat. Hän arvioi pelien hauskuutta kolmen eri perusteen avulla: haaste (challenge), fantasia (fantasy) ja mielenkiintoisuus (curiosity).

*Haasteena* peleissä toimii pelin päämäärällisyys ja pelaajan kannalta epävarma lopputulos. Päämäärällisyyteen liittyy myös esimerkiksi pisteiden laskenta ja edistymisen seuraaminen. Epävarmaan lopputulokseen vaikuttavat myös vaikeustason valinta, käyttöliittymän kasvava monimutkaisuus, ja monipuolinen pisteytysmekanismi.

*Fantasiaa* peleissä edustaa tunteellisesti kiinnostavat asiat, esimerkiksi grafiikka ja musiikki, ja pelissä käytettävät metaforat (esimerkiksi pelissä näkyvät tikat ja maaleina olevat ilmapallot). Malone havaitsi, että fantasian lisääminen peliin kasvatti pelin hauskuutta muita elementtejä enemmän.

*Mielenkiintoisuutta* edustaa käyttöliittymästä saatavan tiedon optimaalinen kompleksisuus ja käyttäjälle välittyvän tiedon selkeä rakenne. Optimaaliseen kompleksisuuteen liittyy pelin audiovisuaalisen sisällön palveleminen sekä tiedonvälityksen että fantasian tarpeita, samoin pelin satunnaisuuden ja huumorin

pysyminen sopivalla tasolla ilman että kumpikaan vaikeuttaa pelin ymmärtämistä tai arvioimista. Tiedon rakenteen selkeyteen liittyy pelin pelaajalle tarjoaman tiedon määrä – pelaajalle ei saa jäädä tiedon suhteen epäselvyyksiä, tiedon pitää olla selkeää ja pelin ei myöskään tule tarjota liikaa tietoa.

Rajoitteena myöhemmät tutkijat, esimerkiksi Federoff (2002, s. 5), ovat maininneet Malonen keskittyneen tutkimuksessaan lähinnä opetuspeleihin. Toisaalta Federoff toteaa, että Malonen tutkimuksen kohteena oli nimenomaan pelien hauskuus, joten rajoittuminen opetuspeleihin ei haitanne heuristiikkojen yleistämistä.

## 4.2 The 400 Project

Pelinkehittäjä Hal Barwood aloitti ”The 400 Project” -projektin (Falstein & Barwood, n.d.) vuonna 2001, ja seuraavana vuonna projektiin liittyi toinen pelikehittäjä Noah Falstein.

Tekijät aloittivat kokoamaan 400 pelinkehitystä koskevan nyrkkisäännön kokoelmaa, koska 400 sääntöä tuntui heistä riittävältä. Vuonna 2001 Barwood esitteli ensimmäiset 4 sääntöä. Sääntöjä on sen jälkeen kerätty tekijöiltä itseltään ja muilta pelinkehittäjiltä. Viimeksi luetteloa on päivitetty vuonna 2006 ja luettelossa on nykyisin 112 sääntöä; verkkosivulta ei käy ilmi projektin nykytilannetta, mutta viimeisimpien tietojen mukaan projekti jatkuu edelleen. Vertaisarvioitujen julkaisujen ja tiedekonferenssien sijaan tekijät ovat kommentoineet luetteloa ja sen edistymistä Game Developers Conference -kehittäjäkonferenssissa vuosina 2001, 2002 ja 2006.

Vertaisarvioinnin puutteesta ja arkipäiväisen käytännönläheisestä lähestymistavasta huolimatta luettelo on kuitenkin mainittu ainakin muutamalla sanalla useissa heuristiikkoja koskevissa tiedeartikkeleissa varhaisena heuristiikkojen kokoamisyrittäksenä. Näistä maininnoista esimerkkeinä Desurvire, Caplan ja Toth (2004, s. 1509), Desurvire ja Wiberg (2009, s. 558) sekä Paavilainen (2010, s. 57).

Barwoodin Finite Arts -verkkosivulta ladattavassa Excel-taulukossa on esitelty itse nyrkkisääntö, lyhyt alle 250 sanan selitys, soveltuvuusalue (esimerkiksi ”Basic”, ”Meta”, ”Feedback”, ”Production”...) ja säännön keksijä.

## 4.3 Federoff: ensimmäiset modernit videopeliheuristiikat

Federoff (2002) teki pro gradu -työnään yhden varhaisimmista nimenomaan videopelejä koskevista heuristiikkamalleista. Paavilainen (2010, s. 57) kutsuu Federoffin työtä ”ensimmäiseksi moderniksi videopelispesifiseksi heuristiikkamalliksi, johtuen sen rakenteesta ja suunnittelumetodista (kirjallisuuskatsaus ja empiirinen tutkimus)”.

Federoff tutki viikon ajan viisihenkisen kehitystiimin työskentelyä ensimmäisen kehityskuukauden loppupuolella, jolloin pelistä oli valmiina ensimmäisiä prototyyppjejä. Hän esittelee kirjallisuuskatsauksessa aikaisempia heuristiikkoja, ja esimerkiksi pohtii seikkaperäisesti miten hyvin Nielsenin (1994) perusheuristiikat soveltuvat pelikehitykseen (Federoff, 2002, s. 16-20).

Loppuyhteenvedon Federoff (2002, s. 41-42) luetteli 40 heuristiikkaa, jotka oli jaoteltu kolmeen pääkategoriaan:

- Käyttöliittymä (game interface). Esimerkkejä: ”Käyttöliittymä on mahdollisimman vähän pelin tiellä” ja ”Valikkotasojen määrä tulee minimoida”.



- Mekaniikat (game mechanics). Esimerkkejä: ”Mekaniikat tuntuvat luonnollisilta ja pelihahmojen liike ja paino tuntuvat oikeilta” ja ”Käyttäjän toiminta kuvastuu välittömästi pelin antamana palautteena”.
- Sisältö (game play). Esimerkkejä: ”Pelin lopputulos tulisi olla yllättävä” ja ”Pelin tulisi olla reilu”.

Kaksi heuristiikkaa olivat näiden kategorioiden rajoilla (Taiteen soveltuminen toimintoihin – ”game interface and play”, pelaajan auttaminen nopeaan peliin siirtymiseen – ”game mechanics and play”).

Paavilainen (2010, s. 58) kutsuu Federoffin mallia ”kyseenalaiseksi” ja ”puolueelliseksi”, koska metodologia muistuttaa yksittäistä tapaustutkimusta – kohteena oli vain yksi peliprojekti ja yksi pieni pelinkehitysryhmä. Lisäksi Paavilaisen mukaan Federoff ei pyri yleistämään mallia. Paavilaisen mukaan laajempi tutkimus olisi kasvattanut tuloksien validiutta. Saman rajoituksen havaitsee Federoff (2002, s. 41) itsekkin: ”Uudet tapaustutkimukset muissa pelitaloissa tarjoaisivat mahdollisuuden vertailuun tässä tutkimuksessa kerättyyn dataan. [...] Viimeisenä ehdotuksena jatkotutkimukselle on verifioida heuristiikkaluettelo. Palaute muilta pelikehitykseen osallistuneilta henkilöiltä voisi edesauttaa tätä.”

#### 4.4 Heuristic Evaluation of Playability (HEP) ja Heuristics of Playability (PLAY)

Heuristic Evaluation of Playability (Desurvire, Caplan, & Toth, 2004) on yksi usein siteeratuista malleista. Alkuperäinen malli esitettiin CHI '04 -konferenssissa, ja julkaistiin vuonna 2009 päivitettyssä muodossa nimellä ”Heuristics of Playability (PLAY)” (Desurvire & Wiberg, 2009).

Alkuperäisessä HEP-mallissa (Desurvire, Caplan, & Toth, 2004) heuristiikat löydettiin ”nykyisestä kirjallisuudesta” ja käytettävyyssiantuntijat arvioivat ja luokittelivat ne. Heuristiikkojen validointia varten tutkimuksessa kehitettiin Flash-peliprototyyppi, jossa oli pelkästään pelin ruutukuvia eikä varsinaista pelisisältöä; tekijöiden mukaan tämä vastaa normaalia heuristiikkojen käyttötilannetta, jossa heuristista evaluointia sovelletaan varhaisessa pelin kehitysvaiheessa. Evaluoinnin lopputulosta verrattiin prototyypin käyttäjätestauksesta saatuun palautteeseen.

HEP luettelee 43 heuristiikkaa, jotka on jaettu neljään kategoriaan:

- Haasteet ja pelikokemus (*game play*), 16 heuristiikkaa. Esimerkkejä: ”Pelaajan rasitus minimoidaan vaihtelemalla aktiviteetteja ja rytmitystä pelin aikana” (1), ”Pelaaja ei saa kokea tulleeensa rangaistuksi jatkuvasta samassa kohdassa epäonnistumisessa” (11), ”Haasteen tulee olla positiivinen eikä negatiivinen kokemus, jonka johdosta pelaajan tulee inspiroitua jatkamaan eikä lopettamaan” (16)
- Tarina, pelimaailma ja hahmot (*game story*), 8 heuristiikkaa. Esimerkkejä: ”Pelaaja ymmärtää tarinan yhtenä kokonaisvaltaisena visiona” (1), ”Pelaaja kuluttaa aikaa miettimällä mahdollisia tarinan loppuratkaisuja” (3), ”Hahmot kiinnostavat pelaajaa, koska he ovat pelaajan kaltaisia, ovat luonteeltaan kiinnostavia, ja kehittyvät toiminnan edetessä” (8)
- Mekaniikat (*mechanics*), 7 heuristiikkaa. Esimerkkejä: ”Mekaniikoilla ja ohjaimilla tehtävillä toiminnoilla on selkeät ja opittavat pelissä tapahtuvat

vasteet” (4), ”Pelin oppimista nopeutetaan seuraamalla peliteollisuuden trendejä ja vastaamalla pelaajien odotuksiin” (5).

- Käytettävyys (*usability*), 12 heuristiikkaa. Esimerkkejä: ”Pelin aloittaminen ja lopettaminen on helppoa, ja peli voidaan tallentaa eri tilanteissa” (2), ”Pelaaja kokee pelivalikot osana peliä” (4), ”Pelin pelaamiseen ei tarvita ohjekirjaa” (8)

Validoinnissa heuristisen evaluoinnin ja käyttäjätestauksen palaute analysoitiin. Heuristinen evaluointi löysi lukumäärällisesti enemmän ongelmia, mutta käyttäjätestaus löysi enemmän pelispesifisiä ongelmia. Osa eri menetelmillä löydettyistä ongelmista oli samoja.

Paavilainen (2010, s. 58) kritisoi vuoden 2004 versiota HEP-mallista epäselväksi kehitysmetodologian kannalta: konferenssiartikkelissa heuristiikat koonneita asiantuntijoita eikä heidän menetelmiään esitellä, eikä artikkelista käy ilmi mitkä heuristiikat saatiin kirjallisuuslähteistä ja mitkä ovat asiantuntijoiden muokkaamia. Paavilainen pitää myös pieneen käyttäjäjoukkoon ja yhteen peliprototyyppiin nojaavaa validointia riittämättömänä. Samoin Paavilaisen mielestä jotkin heuristiikat ovat hyvin subjektiivisia ja täten vaikeita käyttää evaluointiprosessissa (esimerkkinä edellä mainittu ”hahmot kiinnostavat pelaajaa”, game story 8), ja niiden tarkkuus ja sovellettavuus eri tyyppeihin vaihtelee.

Vuoden 2009 PLAY-mallissa (Desurvire & Wiberg, 2009, s. 558) tekijät toteavatkin, että HEP osoittautui sovellusalueeltaan hyvin suppeaksi: ”HEP oli hyödyllinen mutta vain rajatuissa tapauksissa. Pelien kenttä on laaja, ja siihen sisältyy pelien genret, toimitus, parannukset ja uusien pelien kehitys. [...] [PLAY] luotiin erityisesti auttamaan pelikehittäjiä kaikissa suunnitteluprosessin vaiheissa, varsinkin alun konseptivaiheessa jolloin pelisuunnitelman muutokset ovat halvempia. [...] Toisin kuin HEP, PLAY ottaa huomioon että pelien suunnittelu on taidetta siinä missä tiedettäkin.”

Heuristics of Playability eli PLAY -mallin (Desurvire & Wiberg, 2009) tarkoituksena oli täydentää HEP-mallin puutteita. Tutkimukseen valikoitiin pelejä hakemalla hyviä ja huonoja arvosanoja saaneita pelejä Metacritic -sivustolta. PLAY keskittyy kolmeen peligenreen: toimintaseikkailuihin, ensimmäisen persoonan ammutapeleihin ja reaaliaikaisiin strategiapeleihin. Tutkimus toteutettiin kyselynä pelimessuilla: pelit valikoitiin alustavasti arvosanojen perusteella, ja kyseisiä pelejä pelanneilta messukävijöiltä kysyttiin mielipiteitä kyseisistä peleistä. Kysymyksiä oli kahta tyyppiä: kaikkia pelejä koskevia kysymyksiä ja genren pelejä koskevia kysymyksiä.

PLAY-artikkeli luettelee heuristiikat, jossa oli tilastollisesti merkittäviä eroja hyviä ja huonoja arvosteluja saaneiden pelien välillä.

1. Haasteet ja pelikokemus (game play): Jatkamisen mielekkyys (5 heuristiikkaa), haaste, strategia ja rytmitys (6 heuristiikkaa), pelimaailman yhtenäisyys (2 heuristiikkaa), päämäärät (3 heuristiikkaa), pelaajien ja pelityylien eroavaisuuksien huomioon ottaminen (4 heuristiikkaa), pelaajien mielikuva päätäntävällästä (2 heuristiikkaa)
2. Tunteelliset ja viihdyttävät seikat (coolness/entertainment/humor/emotional immersion): pelaajan tunteellinen side pelihahmoihin ja -maailmaan (1 heuristiikka), pelin mielenkiintoisuus ja viihdyttävyys (1 heuristiikka), huumorin sopivuus (1 heuristiikka), pelaajan uppoutuminen peliin (1 heuristiikka)
3. Käytettävyys ja mekaniikat (usability/game mechanics): dokumentointi ja tutoriaalit (2 heuristiikkaa), pelitilanne ja pisteytys (4 heuristiikkaa), palaute (2

heuristiikkaa), terminologia (3 heuristiikkaa), pelaajan kokema rasitus (2 heuristiikkaa), pelin käyttöliittymän ulkonäkö (4 heuristiikkaa), navigointi (1 heuristiikka), virheiden estäminen (5 heuristiikkaa), tarinan immerssiivisyys (1 heuristiikka)

## 4.5 Mobiilipelien heuristiikat

Korhonen ja Koivisto (2006) esittelivät heuristiikkamallin mobiilipelejä varten. Seuraavana vuonna he julkaisivat myös heuristiikkamallin monen pelaajan mobiilipelejä varten (Korhonen & Koivisto, 2007). Tekijöiden mukaan älypuhelimien ja kannettavien pelikonsolien peleissä on omia erityispiirteitään, jotka tulee ottaa huomioon heuristiikoissa.

Korhosen ja Koiviston malli on luonteeltaan ”modulaarinen”: mallissa on kolme osaluuetta, pelin sisältö (gameplay), käytettävyys (usability) ja mobiiliominaisuudet (mobility), ja näitä voidaan soveltaa erikseen. (Paavilainen (2010) huomauttaa, että Korhonen ja Koivisto käyttävät pelin sisältöä useimmiten kuvaavan termin ”game play” sijaan termiä ”gameplay”.) Mallin eri osia voidaan soveltaa eri vaiheissa kehitystä: Pelin sisältöä voidaan käsitellä jo varhaisissa kehitysvaiheissa, ja mobiili- ja käytettävyyskysymyksiä voidaan soveltaa toimivien prototyyppien valmistuttua (Korhonen & Koivisto, 2006, s. 10). Paavilainen, Alha ja Korhonen (2012, s. 339) huomauttavat myös, että Korhosen ja Koiviston artikkeleita on myös kommentoitu useita kertoja myöhemmissä artikkeleissa.

Alkuperäinen malli pohjasi 11 heuristiikkaan, joiden pohjana oli kirjallisuuskatsaus, aikaisemmat heuristiikat ja mobiililaitteiden ja niiden käyttötilanteiden analyysi. Jatkokehitys ja validointi tapahtui soveltamalla heuristista evaluointia mobiilipeliin, ja tutkimalla millaisia asiantuntijoiden löytämiä ongelmia malli ei kattanut. Uusien heuristiikkajoukkojen jälkeen validointi tapahtui evaluoimalla viittä eri peliä. (Korhonen & Koivisto, 2006, s. 10-13.)

Valmiissa mallissa oli 12 käytettävyyteen, 3 mobiiliominaisuuksiin ja 14 sisältöön liittyvää heuristiikkaa. Mobiiliheuristiikat olivat ”Peli pitää pystyä aloittamaan ja lopettamaan nopeasti”, ”Peli ottaa ympäristön huomioon” ja ”Pelin keskeytykset hoidetaan sopivalla tavalla”. (Korhonen & Koivisto, 2006, s. 13-14.)

Vuoden 2007 mallissa esiteltiin 8 heuristiikkaa mobiiliverkkopelejä varten, sekä selitystä verkkopelien asettamista erityishaasteista. Uusia heuristiikkoja olivat muun muassa ”Peli antaa tietoja muista pelaajista”, ”Peli mahdollistaa myös yksinäisen/omatoimisen pelaamisen” ja ”Pelisuunnittelu piilottaa verkon vaikutukset”. (Korhonen & Koivisto, 2007, s. 32-33.)

## 4.6 Pinelle ja kumppanit: peliarvostelut heuristiikkojen lähteenä

Pinelle, Wong ja Stach (2008a, 2008b) tutkivat pelien käytettävyyttä peliarvostelujen näkökulmasta. Kumpikin vuoden 2008 konferenssijulkaisuista pohjautuivat samaan yhteensä 108 GameSpot -sivustolta valikoidun arvostelun joukkoon, 18 arvostelua 6 eri peligenrestä. (Pinelle, Wong, & Stach, 2008a, s. 1455-1456; Pinelle, Wong, & Stach, 2008b, s. 130.)

Ensimmäisen tutkimuksen metodologia oli kolmiosainen. Kustakin arvostelusta koodattiin peliarvostelijoiden havaitsemat virheet. Toisessa vaiheessa ongelmat

jaoteltiin kategorioihin. Kolmannessa vaiheessa kategorioista muodostettiin varsinaiset heuristiikat. (Pinelle, Wong, & Stach, 2008a, s. 1455-1456.)

Lopputuloksena oli 10 heuristiikkaa selityksineen: Yhdenmukaiset reaktiot käyttäjän toimintaan, pelin asetusten muokkaus, tietokoneen ohjaamien yksiköiden käyttäytymisen yhdenmukaisuus, pelaajan näkökulmat kaikkiin tilanteisiin, toistuvien epäinteraktiivisten kohtauksien ohittaminen, ohjaintoimintojen muokkaus, hallittavat ohjaustavat, pelaajan pitäminen pelitilanteen tasalla, pelaajan ohjeistus ja selkeä ja yhdenmukainen visuaalinen palaute (Pinelle, Wong, & Stach, 2008a, s. 1458).

Heuristiikkojen validointi tapahtui antamalla joukolle evaluoijia tehtäväksi evaluoida demoversiota pelistä jonka arvostelu ei kuulunut alkuperäiseen arvostelujoukkoon ja kommentoida peliä ja itse heuristiikkajoukon toimivuutta (Pinelle, Wong, & Stach, 2008a, s. 1459).

Yhtenä mahdollisena ongelmana ja rajoituksena tekijät mainitsevat suuren näkökulmaeron tutkimuksen ja peliarvostelujen välillä. ”Arvostelut kirjoittaneet toimittajat eivät ole käytettävyyssiantuntijoita eivätkä pelikritiikissään keskittyneet nimenomaan käytettävyysoongelmiin. Arvostelut on kirjoitettu pelaajien hyödyksi, ja niissä kerrotaan suunnitteluongelmista ja teknisistä ongelmista, jotka saattavat yleensä haitata pelien käytettävyyttä. Uskomme että arvosteluissa paneudutaan seikkaperäisesti suurimpiin käytettävyysoongelmiin, mutta hienovaraisemmat ongelmat on voitu jättää huomiotta.” Samoin tekijät näkivät rajoituksina keskittymisen nimenomaan käytettävyysoongelmiin, eikä pelikokemusta haittaaviin ongelmiin. (Pinelle, Wong, & Stach, 2008a, s. 1461.)

Toinen artikkeli analysoi peligenrejä erikseen. Tutkijat havaitsivat, että eri peligenreissä on selkeästi havaittavissa erilaisia ”ongelmakuvioita”: tietyntyyppiset ongelmat esiintyivät useammin tietyissä peligenreissä. Tekijät pystyivät muodostamaan havainnoistaan ongelmaprofiilin kullekin genrelle. Esimerkiksi urheilupeleissä ohjaustavat olivat usein ongelmallisia, ja strategiapeleissä tietotulva ja mikrohallinnan tarve ovat suuria ongelmakohtia. (Pinelle, Wong, & Stach, 2008b, s. 133-134.) Johtopäätöksenä tekijät huomauttavat, että peligenret tulee ottaa huomioon pelisuunnitteluprosessissa, ja artikkelissa esitetyt genreprofiilit voivat olla hyödyksi pelien käytettävyyttä arvioitaessa. Toisaalta tekijät eivät suosittele keskittymään yksinomaan genrepohjaiseen arviointiin, koska monet nykyiset pelit rikkovat genererajoja. (Pinelle, Wong, & Stach, 2008b, s. 135.)

Seuraava heuristiikkaluettelo, NGH eli Networked Game Heuristics (Pinelle, Wong, Stach, & Gutwin, 2009) keskittyi verkkopeleihin, ja pohjautui 328 peliarvosteluun GameSpot- ja GameSpy -verkkosivustoilta. Metodologia oli pääpiirteittäin sama kuin vuoden 2008 ensimmäisessä tutkimuksessa. Sinällään tutkimus ei korvannut vuonna 2008 havaittuja heuristiikkoja, vaan keskittyi nimenomaan verkkopeleihin. Osa heuristiikoista oli uusia, osa pohjautui edelliseen tutkimukseen; yhtenä lähteenä olivat esimerkiksi Korhosen ja Koiviston (2007) verkkopelejä koskevat heuristiikat.

Jälleen kerran tuloksena oli 10 heuristiikkaa selityksineen: Istuntojen hallinta, joustava moninpelien löytäminen, sopivat kommunikaatiotyökalut, tuki koordinaatiolle, hyödyllinen tilannetieto kanssapelaajista, avatarien tunnistettavuus, aloittelijoiden alkuavustus, sosiaalisen interaktion tuki, pelin aiheuttamien interaktioiden hidastuksien minimointi ja häiriköiden hallintamahdollisuus (Pinelle, Wong, Stach, & Gutwin, 2009, s. 174). Validointi tapahtui samaan tapaan kuin vuoden 2008 ensimmäisessä tutkimuksessa: tällä kertaa evaluoijat tutkivat kahta avoimen lähdekoodin verkkopeliä (Pinelle, Wong, Stach, & Gutwin, 2009, s. 175).

Vuoden 2008 tutkimuksissa havaittuja heuristiikkoja validoitiin pelaamalla pelien demoversioita, joten ei-asiantuntijoiden ja populaarilähteiden käyttäminen tutkimuksen lähtökohtana ei välttämättä ole ongelmallista. Toisaalta esimerkiksi Paavilainen huomautti, että vain yhteen tai kahteen peliarvostelusivustoon nojaaminen voi tehdä näistä tutkimuksista puolueellisia. Hän mainitsee, että GameSpot ei välttämättä ole luotettava peliarvostelujen lähde, koska pelijulkaisijoilla vaikuttaa olevan mainostuksensa kautta arvostelujen sisältöön määräysvaltaa, viitaten Jeff Gerstmannin erottamiseen ja toimittajien joukkoyhtymiseen vuonna 2007. (Paavilainen, 2010, s. 59-60.) Toisaalta Pinelle, Wong ja Stach (2008a, s. 1455) havaitsivat jo pilottitutkimuksessa, että yli 8 pistettä GameSpotissa saaneet peliarvostelut olivat heidän kannaltaan hyödyttömiä, koska arvostelijat eivät kertoneet kohtaamistaan käytettävyysongelmista.

#### 4.7 Kritiikkitilkitseminen

Livingston, Mandryk ja Stanley (2010) julkaisivat konferenssiartikkelin pelien ”kriiikkitilkittämisestä” (Critic Proofing) – kriitikoilta tulevan palautteen sisällyttämisestä pelien kehitysprosessiin kokonaisvaltaisesti.

Tekijät ottavat uuden näkökulman heuristiseen evaluointiin. Uusien heuristiikkojen sijaan tekijät kyseenalaistavatkin itse heuristisen evaluoinnin prosessin. Niin palautteen priorisointi kuin genrejen merkitys ovat tekijöiden mukaan huonolla tasolla perinteisessä heuristisessa evaluoinnissa. Otsikossa mainittu ”kriiikkotilkittäminen” juontuu Biowaren Mass Effect 2 -projektista, jossa sarjan ensimmäisestä osasta saatu palaute oli tärkeässä osassa uuden pelin kehittämisessä.

Tekijät esittelevät matemaattisen mallin, jonka avulla eri ongelmille voidaan antaa prioriteetti ja kuinka nämä ongelmat voidaan uudelleenpriorisoida genrejen avulla. Esimerkkinä tekijät mainitsevat kuinka toimintapeleissä ohjausongelma, jolle on annettu alhainen prioriteetti, voi helposti nousta mallin avulla korkeammalle kuin korkeasti priorisoitu ohjeiden ongelma, koska toimintapeleissä toimiva ohjaus on avainasemassa.

Menetelmän validoimiseksi tekijät esittelevät Capsized -pelin keskeneräisestä versiosta saadun kritiikin soveltamisen jatkokehityksessä. Evaluoijat löysivät 28 ongelmaa, jotka priorisoitiin mallin mukaan. Pelin kehittäjien mukaan priorisointi helpotti ongelmien organisointia ja soveltui kehittäjien käyttöön huomattavasti paremmin kuin subjektiivisesti lajitellut havainnot.

## 5. Yhteenveto

Kirjallisuuskatsauksen tarkoituksena oli tutkia asettavatko videopelit erityisiä haasteita käytettävyyden arvioinnille, ja tähän liittyen tutkia eri kirjallisuustietokantojen avulla videopelien käytettävyyden arviointimenetelmien nykytilaa.

Ensimmäinen tutkimuskysymys: *sopivatko ohjelmistojen yleiset hyvän käytettävyyden suunnitteluperiaatteet myös pelien suunnitteluun, vai pitääkö pelien luonteesta johtuen ottaa huomioon erityisiä käytettävyyšnäkökulmia videopelien suunnittelussa?* Ohjelmistojen yleisessä arvioinnissa käytettävät heuristiikat antavat hyvän perustan käytettävyyden arvioinnille, mutta pelien erityispiirteet huomioon ottavat heuristiikat antavat tarkempaa ja käyttökelpoisempaa tietoa pelikehittäjille.

Toinen tutkimuskysymys: *millaisia työkaluja pelikehittäjillä on pelien käytettävyyden arvioimiseksi?* Heuristinen evaluointi on selkeästi yksi toimivimmista käytettävyyden arviointimenetelmistä varhaisessa ohjelmien kehitysvaiheessa, mutta muitakin arviointimenetelmiä on olemassa ja ne täydentävät myös heuristisen evaluoinnin puutteita muissa kehitysvaiheissa.

Heuristinen evaluointi on erittäin suosittu menetelmä sovellusohjelmistojen kehittäjien keskuudessa, ja se on muotoutumassa myös merkittäväksi työkaluksi pelinkehittäjien keskuudessa. Pelikehityksen puolella heuristiikkojen historia on monella tapaa mielenkiintoinen: heuristiikkojen kattavuusalue on vaihdellut ja eri tutkijat ovat jaotelleet heuristiikkakategoriansa hieman eri tavalla. Kehityksen myötä heuristiikkaluettelot ovat myös supistuneet – esimerkiksi HEP:n 43 heuristiikasta (Desurvire, Caplan, & Toth, 2004) ollaan päästy tuoreempiin 10 heuristiikan luetteloihin (Pinelle, Wong, & Stach, 2008a).

Yksi tutkimuskysymyksistäni oli ohjelmistojen yleisien suunnitteluperiaatteiden ja videopelien suunnitteluperiaatteiden eroavaisuus. Hyötyohjelmilla ja peleillä on suuria eroja ja pelispesifisten heuristiikkojen kehitys on hyödyllistä. Jopa pelien välillä on suuria eroja; esimerkiksi Pinelle, Wong ja Stach (2008b) havaitsivat genrejen selkeät eroavaisuudet. Kirjallisuuden valossa on selvää, että jo yleiset ohjelmistojen arviointimenetelmät soveltuvat myös videopelien käytettävyyden arvioitiin sellaisenaan, kuten Laitinen (2006) havaitsi – mutta toisaalta nimenomaan videopelejä varten kehitetyt heuristiikat ja heuristisen evaluoinnin menetelmät auttavat havaitsemaan käytettävyysongelmia tehokkaasti, kuten kävi ilmi esimerkiksi Livingstonin, Mandrykin ja Stanley'n (2010) sekä Paavilaisen, Alhan ja Korhosen (2012) tutkimuksissa.

Tutkimusalan nuoruudesta johtuen ei näytä siltä, että Nielsenin (1994) heuristiikkaluettelon tapaista yhtenäistä, yleispätevää ja yleisesti tunnettua heuristiikkaluetteloa nimenomaan videopelejä varten olisi kehitetty, mutta jo esimerkiksi Korhosen ja Koiviston (2006, 2007) heuristiikkoja koskeva jatkotutkimus, josta Paavilaisen, Alhan ja Korhosen (2012) tutkimus on esimerkkinä, osoittaa, että pelejä koskevien heuristiikkojen tutkimus etenee hyvää vauhtia. Heuristiikkojen kriittinen vertailu, heuristisen evaluoinnin käyttökokemuksien kerääminen pelikehittäjiltä sekä muu samansuuntainen jatkotutkimus lieneekin avainasemassa heuristisen evaluoinnin jatkokehittämisessä.

Uudet heuristiikkaluettelot siteeraavat aikaisempia tutkimuksia ja mainitsevat osittain samoja heuristiikkoja; esimerkiksi Pinelle, Wong, Stach ja Gutwin (2009) mainitsivat samoja heuristiikkoja kuin Korhonen ja Koivisto (2007). Koska videopelejä koskevat heuristiikat ovat edelleen vahvan kehityksen vaiheessa, voi jopa kysyä, kannattaako etsiä yhteisiä ja eroavia piirteitä eri heuristiikkaluetteloiden välillä vaiko tarkastella heuristiikkaluetteloita yhtenä ajan myötä kehittyvänä luettelona, joka kiertää tutkijalta tutkijalle kirjallisuuden kautta. Nielsenkään ei luonut kuuluisia heuristiikkojaan tyhjästä, vaan ne pohjasivat aiempaan tutkimukseen. Videopelejä koskevien heuristiikkojen luettelot ovat tulleet ja menneet sitten Federoffin vuoden 2002 tutkimusten, mutta Nielsenin vuoden 1994 heuristiikkoja siteerataan sovellusohjelmien kehityksessä nykypäivänäkin samassa muodossa; kenties samantapaisen peliheuristiikkojen kultaisen standardin kehittäminen pitäisi olla tulevan tutkimuksen suuntauksena.

Tärkeämpi kysymys onkin kuinka tarkkoja peliheuristiikkojen oikeastaan tulisi olla. Ehkäpä Nielsenin heuristiikkojen suosion syynä on niiden yksinkertaisuus ja yleispätevyys. Kuten Paavilainen (2010) vihjaa, liian pitkät heuristiikkaluettelot voivat olla haitaksi – Nielsen yksinkertaisti heuristiikkaluetteloaan tarkoituksella.

Yksi silmiinpistävimmistä asioista heuristiikkaluetteloissa on niiden kattavuus ja limittyvyys. Pelit ovat kokonaisvaltaisia elämyksiä, joten voidaan kysyä, kuinka mielekästä on arvioida pelkäästään pelien käytettävyyttä, vai onko mielekkäämpää arvioida koko käyttäjäkokemusta. Havaitut heuristiikat eivät kata pelkäästään puhtaasti käytettävyyteen liittyviä asioita, vaan käytettävyyden on usein kytköksissä pelimekaniikkaan ja varsinaiseen pelisisältöön. Yksi mielenkiintoisimmista tässä katsauksessa olevista esimerkeistä käytettävyyden merkityksestä on Federoffin tutkimuksessa. Haastatellessaan kehittäjiä Federoff huomasi, että käytettävyyden terminä oli kehittäjille jopa tuntematon, ja kehittäjät halusivat keskittyä enemmän pelin hauskuuteen liittyviin asioihin, ja näistä keskusteleminen avasi keskustelua paremmin kuin puhtaasti käytettävyyteen keskittyminen. (Federoff, 2002, s. 29-30.) Jo näin varhaisissa tutkimuksissa on selkeästi tiedostettu, että pelien luonteesta johtuen heuristiikkojen on syytä kattaa pelien sisältöä ja toimivuutta laajalta alueelta myös käyttäjäkokemuksen osalta. Pinelle, Wong ja Stach (2008a, s. 1461) totesivat, että huolimatta heidän puhtaasti käytettävyyteen liittyvästä näkökulmasta, erilaisia hauskuuteen ja mielenkiintoisuuteen liittyviä kysymyksiä ei tule unohtaa. Toisaalta Korhonen ja Koivisto (2006, s. 10) havaitsivat, että aiemmissa tutkimuksissa kategoriat olivat osittain mielivaltaisia siten, että pelin sisältöön ja pelin mekaniikoihin liittyvät kategoriat voitiin helposti yhdistää – näitä asioita ei oikeastaan ollut tarpeen käsitellä erikseen. Vielä vuosikymmenien tutkimuksenkin jälkeen videopelit ovat tutkimusalue, jossa terminologia ei ole aivan tarkkaa ja rajanvedot ovat joskus haastavia – ja joskus pitää kysyä, tarvittiinko aiempia rajanvetoja lainkaan.

Kirjallisuudessa dokumentoidun validoinnin ja käytännön testauksen perustella heuristinen evaluointi on tehokas menetelmä varhaisessa ohjelmistojen kehitysvaiheessa. Heuristiikkojen avulla voidaan estää kalliiden virheiden tekeminen tulevaisuudessa: ne soveltuvat varsin hyvin pelikehitykseen ennen kuin varsinaisia prototyyppijä edes aletaan rakentaa.

Nielsenin kuvaamat halvat käytettävyyden arviointimenetelmät sopivat hyvin moderniin ohjelmistokehitykseen. ”Heuristisen evaluoinnin päämäärä on löytää käyttöliittymien käytettävyysongelmia, jotta ne voidaan korjata osana iteratiivista suunnitteluprosessia” (Nielsen, 1993, s. 155). ”[Heuristinen evaluointi] sopii ketteränä ja kustannustehokkaana laadullisena menetelmänä hyvin sosiaalisten pelien nopeaan kehityssykliin” (Paavilainen, Alha, & Korhonen, 2012, s. 349). Halpoja käytettävyyden arviointimenetelmiä voidaan soveltaa myös hyvin erikokoisten kehitystiimien tuotteisiin, mikä on teollisuuden nykytilassa tärkeää – nykyisin pelejä kehittävät

muutaman hengen indie-kehitystiimit, suuret monikansalliset pelitalot, ja kaikki näiden välillä.

Heuristiikat soveltuvat paitsi varsinaiseen asiantuntijoiden suorittamaan heuristiseen evaluointiin, myös pelien kehittäjien ja suunnittelijoiden mielessä pidettäviksi säännöstoiksi. Halvat käytettävyyden arviointimenetelmät myös demokratisoivat käytettävyydsarviointia: ”Myös muut kuin asiantuntijat voivat löytää heuristiikkojen avulla monia suunnitteluongelmia, ja suurin osa lopuista ongelmista voidaan löytää ääneen ajatteluun pohjaavalla yksinkertaisella testauksella” (Nielsen, 1993, s. 20).

Yksi henkilökohtaisesti yllättävimmistä – vaiko sittenkään – havainnoista on koko heuristiikkoja koskeva tutkimusparadigma. Suurin osa lähdemateriaalista on luonteeltaan laadullista tutkimusta, pohjautuen käyttäjien mielipiteiden kokoamiseen eri lähteistä ja käyttäjien ja asiantuntijoiden haastatteluihin. Nielsen (2009) ehdotti laadullisten, käyttäjäkokemuksia keräävien käytettävyydenmenetelmien toimivan halvemmin ja tehokkaammin kuin numeroarviointiin ja eksakteihin tilastollisiin menetelmiin pohjaavat määrälliset menetelmät. Kirjallisuudessa dokumentoiduissa tapauksissa nämä menetelmät ovat selkeästi osoittaneet erittäin tehokkaiksi, ja pelinkehittäjiltä tullut palaute on ollut positiivista. Koin tämän osittain yllättävänä, koska useimmiten ohjelmistokehitys nojaa matemaattis-loogiseen pohjaan – mutta toisaalta *en* kokenut tätä yllättävänä, koska käytettävyys nojaa hyvin pitkälti ihmisten käyttäytymisen dokumentointiin. On selvää, että ohjelmistokehitys kokonaisuudessaan vaatii kumpaakin tutkimuksen lähestymistapaa. Laadullinen tutkimus on monella tapaa käyttökelpoisin tapa tutkia käyttäjien käyttäytymistä, mutta kuten esimerkiksi Paavilainen (2010) osoittaa, heuristiikkojen historia ollut validiuden osalta varsin kirjavaa. Jo Nielsen (1993, s. 169-170) nosti esiin käytettävyydetutkimuksen validiuden ja relevanssin haasteet.

Heuristiikkojen jatkotutkimukset ja käytännön soveltaminen ovat kuitenkin osoittaneet, että toimivien ja hyödyllisten heuristiikkojen kehittäminen on mahdollista – ja heuristiikat toimivat myös käytännössä. Ehkäpä tämänsuuntaiset tulokset kannustavat pelinkehittäjiä käyttämään tämäntapaisia käytettävyyden arviointimenetelmiä useammin. Varsinkin indie-pelien puolella näkee vieläkin skeptisyyttä ”pehmeämpien” tieteenalojen suuntaan, ja jo kirjallisuuden valossa tämä on aiheetonta.

Tämä kirjallisuuskatsaus on pakostakin hyvin suppea – jo käsillä olevasta materiaalista löytyi runsaasti materiaalia ja viittauksia muihin tutkimuksiin ja tutkimuksia arvioineisiin artikkeleihin, joita ei ollut mahdollisuuksia käsitellä tässä kirjallisuuskatsauksessa. Monissa artikkeleissa mainittiin kirjallisuuslähteitä, joita ei ollut saatavilla tai jotka löytyivät liian myöhään. Samoin aiemmin kirjallisuuskatsaukseen suunniteltu muiden videopelien käytettävyyden arviointimenetelmien esittely jäi heuristiikkoja koskevan materiaalin runsauden takia käsittelemättä; monet muut käytettävyyden arviointimenetelmät soveltuvat myös erinomaisesti pelikehitykseen.

Itseäni kiinnostavana jatkotutkimuksen aiheena olisi pelien välillä ja pelisarjojen sisällä tapahtuneet muutokset, mahdollisesti käytettävyyden näkökulmasta. Livingston, Mandryk ja Stanley (2010) sivusivat tätä aihetta pelien jatkokehityksen näkökulmasta, mutta artikkeli ei keskittynyt aiheeseen kovin kattavasti käytännön kannalta. Pelisarjojen sisällä tapahtunut kehitys on mielenkiintoista, ja varsinkin fanien keskuudessa käydään vilkasta keskustelua eri pelien paremmuudesta ja toimivuudesta. Näihin keskusteluihin kaippaa usein tieteellistä näkökulmaa.



## Lähteet

- 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* (pp. 1509–1512). New York, NY, USA: ACM. doi:10.1145/985921.986102
- Desurvire, H., & Wiberg, C. (2009). Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration. In A. A. Ozok & P. Zaphiris (Eds.), *Online Communities and Social Computing* (Vol. 5621, pp. 557–566). Springer Berlin Heidelberg. doi:10.1007/978-3-642-02774-1\_60
- Falstein, N., & Barwood, H. (n.d.). The 400 Project. Retrieved March 8, 2015, from <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* (Master's Thesis). Department of Telecommunications, Indiana University, Indiana. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.8294&rep=rep1&type=pdf>
- Karvonen, J. T. (2005). *Mobilipelien pelattavuuden arviointi* (Pro gradu -tutkielma). Jyväskylän yliopisto, Jyväskylä. Haettu [https://jyx.jyu.fi/dspace/bitstream/handle/123456789/12471/URN\\_NBN\\_fi\\_jyu-%20%20%20%20%20%202006151.pdf?sequence=1](https://jyx.jyu.fi/dspace/bitstream/handle/123456789/12471/URN_NBN_fi_jyu-%20%20%20%20%20%202006151.pdf?sequence=1)
- Korhonen, H., & Koivisto, E. M. I. (2006). Playability Heuristics for Mobile Games. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services* (pp. 9–16). New York, NY, USA: ACM. doi:10.1145/1152215.1152218
- Korhonen, H., & Koivisto, E. M. I. (2007). Playability Heuristics for Mobile Multi-player Games. In *Proceedings of the 2Nd International Conference on Digital Interactive Media in Entertainment and Arts* (pp. 28–35). New York, NY, USA: ACM. doi:10.1145/1306813.1306828
- Lahti, M. (2014). *Game Testing in Finnish Game Companies* (Master's Thesis). Department of Media Technology, Aalto University, Espoo. Retrieved from [https://aaltodoc.aalto.fi/bitstream/handle/123456789/15193/master\\_Lahti\\_Mikko\\_2015.pdf?sequence=1](https://aaltodoc.aalto.fi/bitstream/handle/123456789/15193/master_Lahti_Mikko_2015.pdf?sequence=1)
- Laitinen, S. (2006). Do usability expert evaluation and test provide novel and useful data for game development? *Journal of Usability Studies*, 2006(2), 64–75. Retrieved from [http://www.uxpajournal.org/wp-content/uploads/pdf/usability\\_game\\_development.pdf](http://www.uxpajournal.org/wp-content/uploads/pdf/usability_game_development.pdf)
- 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* (pp. 48–55). New York, NY, USA: ACM. doi:10.1145/1920778.1920786
- 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* (pp. 63–68). New York, NY, USA: ACM.  
doi:10.1145/800049.801756

- Nacke, L. E., Drachen, A., Kuikkaniemi, K., Niesenhaus, J., Korhonen, H. J., van den Hoogen, W. M., Poels, K., IJsselsteijn, W. A., de Kort, Y. A. W. (2009). Playability and Player Experience Research. In *Proceedings of DiGRA 2009*.
- Nielsen, J. (1993). *Usability Engineering*. AP Professional.
- Nielsen, J. (1994). Heuristic Evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods*. John Wiley & Sons, Inc.
- Nielsen, J. (1995, January 1). 10 Heuristics for User Interface Design. Retrieved March 10, 2015, from <http://www.nngroup.com/articles/ten-usability-heuristics/>
- Nielsen, J. (2009, September 14). Discount Usability: 20 Years. Retrieved February 16, 2015, from <http://www.nngroup.com/articles/discount-usability-20-years/>
- 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* (pp. 56–65). New York, NY, USA: ACM. doi:10.1145/1920778.1920787
- Paavilainen, J., Alha, K., & Korhonen, H. (2012). Exploring Playability of Social Network Games. In A. Nijholt, T. Romão, & D. Reidsma (Eds.), *Advances in Computer Entertainment* (Vol. 7624, pp. 336–351). Springer Berlin Heidelberg. doi:10.1007/978-3-642-34292-9\_24
- 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* (pp. 1453–1462). New York, NY, USA: ACM. doi:10.1145/1357054.1357282
- 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* (pp. 129–136). New York, NY, USA: ACM. doi:10.1145/1496984.1497006
- 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* (pp. 169–178). New York, NY, USA: ACM. doi:10.1145/1531674.1531700
- Schell, J. (2014). Good Games Are Created through Playtesting. In *The Art of Game Design* (pp. 433–448). A K Peters/CRC Press. doi:10.1201/b17723-28
- Suomen Standardoimisliitto SFS. (2010). *SFS-EN ISO 9421-210: Ihmisen ja järjestelmän vuorovaikutuksen ergonomia. Osa 210: Vuorovaikutteisten järjestelmien käyttäjäkeskeinen suunnittelu*.