

Kimmo Aho

ARVIOINNIN HAASTEET KÄYTÄNNÖN TULOKSELLISUUSARVIOINTIRAPORTOINNISSA

Johtamisen ja talouden tiedekunta
Kandidaatintutkielma
Tammikuu 2020

TIIVISTELMÄ

Kimmo Aho: Arvioinnin haasteet käytännön tuloksellisuusarviointiraportoinnissa
Kandidaatintutkielma
Tampereen yliopisto
Hallintotieteiden tutkinto-ohjelma
Tammikuu 2020

Minkä tahansa toiminnan kehittämiseen liittyy olennaisena osana sen arviointi. Arvioinnin avulla toiminnasta saadaan syvällisempi ymmärrys sekä tietoa mahdollisista puutteista. Kuitenkin on muistettava, että arviointi on lähtökohtaisesti ihmisen toteuttamaa, ja näin ollen siihen sisältyy valintojen tekemistä. Tämän tutkimuksen tavoitteena on selvittää, millaisia haasteita tuloksellisuusarvioinnin suorittamiseen liittyy julkishallinnossa, ja miten nämä haasteet näyttäytyvät arvioinnin raportoinnissa. Tutkimus on luonteeltaan laadullinen ja menetelmänä siinä käytettiin aineistolähtöistä sisällönanalyysiä.

Tutkimuksen analyysi toteutettiin analysoimalla Valtion tarkastusviraston verkkosivuillaan julkaisemia tuloksellisuustarkastuskertomuksia. Näistä kertomuksista aineistoon rajattiin kahdeksan raporttia vuosilta 2017–2019. Tämän lisäksi raportit pyrittiin valitsemaan siten, että aineistossa on mukana mahdollisimman laajasti eri hallinnonalojen toimintoja käsitteleviä tarkastuksia.

Aineiston analyysissä nousi esille kolme merkittävää haastetta, jotka esiintyvät läpi aineiston. Ensinnäkin arvioinnin käytettävissä olevassa tiedon määrässä sekä laadussa oltiin havaittu merkittäviä puutteita. Toiseksi haasteeksi nousi sopivan ajankohdan löytäminen arvioinnin toteuttamiselle. Kolmas haaste koskee arvioitavan kohteen kompleksisuutta ja sitä kautta vertailtavuuden puutetta.

Aineistosta löydetty arvioinnin haasteet noudattavat hyvin aiempaa tutkimuskirjallisuutta. Löydetty haasteet kuitenkin nousivat hyvin erilaisin painotuksin raporttien välillä. Tästä pääteltiin, ettei olemassa olevat mallit haasteista toimi suoraan arviointiin ilman arvioijan omaa tulkintaa.

Tutkimus on luonteeltaan hyvin yleistason kuvaus julkissektorin arvioinnin haasteista. Tulevaisuudessa tehtävässä arvioinnin haasteisiin keskittyvässä tutkimuksessa haasteiden näkymistä käytännön toiminnassa voidaan käsitellä syvemmin rajaamalla tarkastelun kohteeksi vain osa haasteista. Samoin löydettyjä haasteita voidaan tarkastella tietyn toimialan arvioinnin näkökulmasta.

Avainsanat: Tuloksellisuus, arviointi, mittaus, vaikuttavuus, taloudellisuus, tehokkuus

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

Sisällys

1 Johdanto	5
1.1 Johdatus aiheeseen	5
1.2 Tutkimustehtävä ja tavoitteet	7
1.3 Tutkimuksen eteneminen	7
2 Tuloksellisuuden arviointi.....	8
2.1 Tutkimuksen käsitteet.....	8
2.2 Teoreettinen viitekehys	9
2.2.1 Kolmen E:n malli.....	9
2.2.2 Päämies-agentti-malli.....	11
2.2.3 Arvioinnin funktiot	12
2.2.4 Arvioinnin haasteet	13
3 Tutkimuksen toteutus.....	14
3.1 Aineisto	14
3.2 Menetelmä	15
4 Analyysi	16
4.1 Analyysin rakenne.....	16
4.2 Tiedon puutteet.....	18
4.3 Arvioinnin ajankohta.....	20
4.4 Kohteen kompleksisuus.....	22
4.5 Muut haasteet	24
5 Johtopäätökset.....	25
5.1 Pohdintaa	25
5.2 Luotettavuus	27
5.3 Jatkotutkimus.....	28
Viitteet.....	29

Kuviot

1. Kolmen E:n malli panos-tuotos-asetelmassa.....	10
2. Arvioinnin funktiot kolmijakona.....	13
3. Julkishallinnon interventioiden arvioinnin haasteet.....	17

1 Johdanto

1.1 Johdatus aiheeseen

Tuloksellisuuden tarkastelu julkisissa organisaatioissa sekä interventioissa on ollut pinnalla tieteellisessä keskustelussa viimeistään 1990-luvulta lähtien, kun julkissektorin ajattelutavat muuttuivat tulosjohtamisen sekä uuden julkisjohtamisen (New Public Management) myötä. Uudessa julkisjohtamisessa on kyse muutosvoimasta, jossa entisen kaltaiset byrokraattiset, poliittiset ja hallinnolliset järjestelmät alettiin nähdä negatiivisina ja liian hitaina. Siihen liittyy myös vahvasti näkemys tarpeesta parantaa julkisten organisaatioiden tuloksellisuutta suhteessa yksityiseen sektoriin. Voidaan sanoa, että uusi julkisjohtaminen tarkoittaa myös pyrkimystä uudistaa julkisen hallinnon toiminnat yrity maailman oppien mukaisesti. Samalla tämä tarkoittaa julkisen ja yksityisen sektorin rajapinnan hämärtymistä. Suuntausta kuvaavia muutosvoimia ovat olleet esimerkiksi yhtiöittäminen, tulosohjaukseen siirtyminen sekä sääntelyn muuttaminen joustavammaksi.

Tuloksellisuus terminä alkoi yleistyä suomalaisessa keskustelussa 1980-luvulla, jolloin myös julkisten toimintojen tuloksellisuus alkoi esiintyä entistä useammin julkishallinnon asiakirjoissa. Esimerkkinä tätä muutosta kuvaavista merkittävistä hankkeista voidaan nostaa valtion liikelaitosuudistus. Tämä ja muut reformit syntyivät vahvasti OECD-maiden yleisen suunnanmuutoksen myötävaikutuksesta. (Meklin 2009, 35.)

Toisaalta on tärkeää tiedostaa myös uuteen julkisjohtamiseen kohdistuvan kritiikin merkitys. OECD:n julkaisusarjassa *Value for Money in Government* julkaistussa raportissa *Building on Basics* (2015, 19–22) kritiikki tiivistetään viiteen kohtaan. Ensinnäkin päätöksenteossa puutteellinen poliittinen ammattitaito voi johtaa politiikoihin, joita on käytännössä mahdoton toteuttaa. Toisaalta julkishallinnossa entistä joustavimmat toiminnan standardit ovat itse asiassa johtaneet aiempaa korkeampiin kustannuksiin. Kolmas kritiikin osa kiinnittää huomiota tulosbudjetoinnista johtuneeseen palvelun laadun sekä byrokratian heikkenemiseen. Tässä byrokratia nähdään positiivisessa valossa toimintaa varmentavana tekijänä. Raportin neljännessä kohdassa samoin tulosbudjetointi johtaa parlamentaarisen kontrollin sekä budjettidokumentaation luettavuuden heikkenemiseen. Viidennen ja viimeisen kohdan mukaan julkisten toimintojen ulkoistaminen itsessään johtaa palvelun laadun laskuun sekä kustannusten nousuun.

Yksi uuden julkisjohtamisen kantava ajatus on *value for money* -ajatusmalli. Sen mukaisesti kansalaisilla on veronmaksajina oikeus odottaa, että verorahoitteinen julkinen toiminta järjestetään mahdollisimman kustannustehokkaasti ja vaikuttavasti. Näin ollen tuotosten ja tuottavuuden mittaaminen nähtiin tarpeelliseksi myös julkissektorin toiminnassa. Toiminnasta tuotettu tuloksellisuusarviointi on yksi vastaus tähän tarpeeseen. Virtanen ja Vakkuri (2016, 16) määrittelevät tuloksellisuusarvioinnin ”pyrkimyksiksi edistää ja parantaa julkisen hallinnon kykyä tehdä kansalaisten hyvinvointia edistäviä ratkaisuja”. *Value for money* -ajattelussa tuloksellisuusarvioinnilla on suuri rooli nimenomaan toimijan tilivelvollisuuden varmistamisessa. Tässä tutkimuksessa käytetään arvioinnin ja tarkastuksen käsitteitä synonyymeinä.

Edellä mainittu *value for money* -malli korostuu, kun tuloksellisuusarviointi liitetään laajempaan organisaation suorituskyvyn hallinnan käsitteeseen. Siinä huomio kiinnittyy paitsi tuloksellisuusarvioinnin tuottamiseen myös sen hyödyntämiseen. Suorituskyvyn hallinnassa on pohjimmiltaan kyse päätöksenteon ja siitä seuraavien toimintojen rationaalisuuden edistämisestä. Jako arvioinnin tuottamisen ja hyödyntämisen välillä toistuukin johdonmukaisesti aiheesta tehdyssä tutkimuskirjallisuudessa.

Suomalaisessa tutkimuskirjallisuudessa tuloksellisuuden arviointia on lähestytty johdonmukaisesti tietyn alan tai organisaation kautta. Sen sijaan julkissektorin tuloksellisuutta kokonaisvaltaisesti käsitteleviä teoksia on hyvin rajatusti. Aiheen kattavimpia teoksia on Petri Virtasen ja Jarmo Vakkurin *Julkisen toiminnan tuloksellisuusarviointi* (2016), joka on hyvin teoriapainotteinen alan perusteos. Myös Ismo Lumijärvi on kirjoittanut laajemmin julkisen sektorin tuloksellisuudesta. Suosituttuja aihealueita tuloksellisuuden käsittelyyn ovat suomenkielisessä tutkimuksessa olleet hoito- sekä koulutusalojen tuloksellisuuden eri näkökulmat, että kuntien ja valtionhallinnon toiminnan tuloksellisuus (kts. esim. Oravakangas 2005, Hakkarainen 2012, Nikkilä 2015, Kortelahti 2019).

Myös englanninkielisessä kirjallisuudessa tuloksellisuutta on usein lähestytty jonkin organisaation tai tapauksen kautta. Varsinainen laajempi kirjallisuus käsittelee arviointia yleisesti teoreettisista näkökulmista, esimerkkinä Evert Vedungin *Public Policy and Program Evaluation* (2017).

Tämä tutkimus pyrkii asettautumaan aiheesta tehdyn alakohtaisen tutkimuksen sekä puhtaasti teoriaa käsittelevän kirjallisuuden väliin. Kun otetaan vielä huomioon tämän tutkimuksen keskittyminen arvioinnin haasteisiin, saadaan siitä toivottavasti lisäarvoa aiempaan tutkimuskirjallisuuteen nähden.

1.2 Tutkimustehtävä ja tavoitteet

Tämän tutkimuksen tarkoituksena on hahmottaa julkisista interventioista tehtyihin arviointeihin liittyviä haasteita. Tutkimustyöllä on pyrkimys jäsenellä, millaiset tekijät vaikeuttavat arviointi-informaation tuottamista sekä millaiset tekijät rajoittavat tämän tiedon luotettavuutta ja käyttökelpoisuutta. Wanda Orlikowski (2002) on kuvannut, kuinka tiedon tuottamiseen organisaatioissa liittyy aina vahva inhimillinen panos. Sama koskee myös tuloksellisuusarvioinnin tuottamista. Arviointi on lähtökohtaisesti ihmisen tuottamaa tietoa, ja näin ollen siihen liittyy aina myös valintojen tekemistä. Tällä tutkimuksella pyritään selvittämään myös, millaisilla painotuksilla näitä valintoja on tehty. Orlikowskin (2002) artikkelin mukaisesti ihmisen toiminta ja tulkinnat näkyvät paitsi arviointitiedon tuottamisessa, myös sen hyödyntämisessä. Arviointitiedon hyödyntämisen haasteita ovat käsitelleet laajemmin esimerkiksi Smith (1995) ja Virtanen & Vakkuri (2016, 212–224). Tässä tutkimuksessa hyödyntämisen haasteet pyritään pitämään kuitenkin vain osana arvioinnin haasteiden kokonaisuutta.

Tutkimuksen tavoitteeksi voisi määritellä aiheesta löytyvän teorian ja käytännön raportoinnin tuomisen lähemmäksi toisiaan. Tämän tutkimuksen varsinainen tutkimuskysymys tiivistyy seuraavaan:

Miten arvioinnin haasteet näyttäytyvät käytännön tuloksellisuusarviointiraportoinnissa?

Tämän kysymyksen lisäksi tutkimuksessa selvitetään, miten hyvin Vedungin arvioinnin kahdeksan haasteen malli soveltuu tämänkaltaiseen ”meta-arviointiin”. Kyseinen malli kuvataan tarkemmin tässä tutkimusraportissa teoreettisen viitekehyksen yhteydessä.

Tutkimuksen aineistoon liittyen, tutkimuksen tarkoituksena ei ole käsitellä raportoinnin kohteita tai niihin liittyviä valintoja. Tutkimuksessa on kiinnostuttu vain varsinaisen raportoinnin sisällöstä. Analyysin yhteydessä esiin nousseet, tiettyjen alojen arviointiin liittyvät ominaispiirteet on pyritty jättämään sivuosaan analyysissä. Näitä voidaan käsitellä lyhyesti, mutta tämä ei ole tutkimuksen varsinainen fokus.

1.3 Tutkimuksen eteneminen

Tässä tutkimusraportissa on tähän mennessä kuvattu tutkimuksen aiheen lähtökohtia sekä sitä, miten tutkimus sijoittuu aiempaan tutkimuskirjallisuuteen. Lisäksi esiteltiin tutkimukselle asetetut tavoitteet

sekä varsinainen tutkimuskysymys. Raportin toisessa luvussa käydään läpi tutkimuksen tärkeimmät käsitteet sekä teoreettinen viitekehys. Kolmas pääluke esittelee tutkimuksen käsittelytavan sekä valitun tutkimusmenetelmän. Neljännessä pääluvussa siirrytään käsittelemään tutkimuksen aineistoa. Luvun alussa esitellään analyysin tulokset kuviona, jonka mukaisesti sitä seuraava varsinainen analyysin kuvaus rakentuu. Lopuksi tutkimusraportin viidennessä pääluvussa pohditaan yleisemällä tasolla arvioinnin haasteellisuutta sekä vaihtoehtoja sen toteuttamiselle. Lisäksi luvussa reflektoidaan tutkimuksen luotettavuutta sekä aiheita mahdolliselle jatkotutkimukselle.

2 Tuloksellisuuden arviointi

2.1 Tutkimuksen käsitteet

Tämän tutkimusraportin käsittelytavan kannalta tärkeitä käsitteitä ymmärtää ovat arviointi, julkiset interventiot arvioinnin kohteena sekä tuloksellisuus. Arvioinnille on vaikea esittää yhtä yleisesti kattavaa määritelmää, johtuen sen käytännönläheisestä luonteesta sekä vahvasta riippuvuudesta kulloisestakin toimintaympäristöstä. Scriven (1991) määrittää arvioinnin jonkin asian tai jonkin prosessin tuottaman suorituksen arvokkuuden, ansioiden ja arvon selvittämiseksi. Valtiovarainministeriön määritelmä arvioinnin tavoitteista on hyvin samankaltainen. Sen mukaan arvioinnin tavoite on selvittää arvioitavan kohteen ”arvo, arvokkuus tai merkitys päätöksentekoa varten”. Arvioinnissa on aina otettava huomioon sen sidos arvoihin, mikä johtaa siihen, että samaa kohdetta voidaan lähestyä erilaisista lähtökohdista ja erilaisin painotuksin. (Salminen 2005, 101.)

Colin Talbot (2005) jaottelee arviointia myös sen kohteen mukaan. Tämän jaottelun mukaisesti arvioinnin kohteena voi olla yksilö, organisaatio tai interventio. Tämä tutkimus keskittyy nimenomaan julkisiin interventioihin eli pääosin julkisella sektorilla toteutettaviin projekteihin, ohjelmiin sekä toimintapolitiikoihin.

Interventioiden arvioinnin erottaa organisaatioiden arvioinnista erityisesti oletus interventioiden ajallisesta ja toiminnallisesta rajallisuudesta. Interventioille on siis määriteltävissä hetki, jolloin siihen aletaan kohdentaa resursseja, samoin kuin hetki, jolloin resurssien kohdentaminen lopetetaan. Samoin interventiot ovat tavoitteiltaan rajatumpia verrattuna usein hyvinkin monitavoitteellisiin organisaatioihin. (Virtanen & Vakkuri 2016, 19.)

Tuloksellisuuden käsite itsessään voidaan nähdä monitulkintaiseksi. Suppeasti katsottuna tuloksellisuus voidaan nähdä synonyyminä tuottavuudelle, kun taas laajemmin katsottuna se kuvaa tuottavuuden lisäksi myös tehokkuutta, vaikuttavuutta sekä työn laatua (Lumijärvi 2009, 21). Tuloksellisuus voidaankin nähdä ikään kuin kattokäsitteenä erilaisille arvioinnin kohteille. Toisaalta Meklin (2009, 35) kuvaa tuloksellisuutta ylimpänä päämääränä, johon rationaalisesti toimiva julkishallinto pyrkii. Vaihtoehtoisesti tuloksellisuus voi olla myös periaate, jota noudatetaan rationaalisessa toiminnassa.

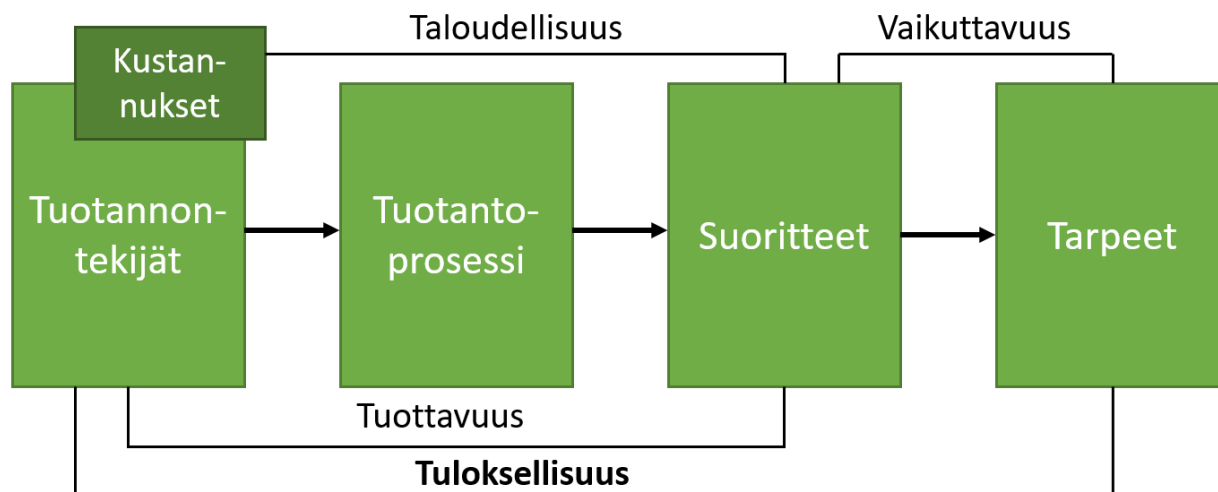
Meklin antaa myös esimerkkinä valtiovarainministeriön määritelmän tuloksellisuudelle ja sen alakäsitteille 1980-luvun lopulta. Siinä tuloksellisuus rinnastetaan tarkoittamaan onnistumista, ja sen alakäsitteiksi määritetään tuottavuus, taloudellisuus ja vaikuttavuus. Myös VTV:n verkkosivuilla tuloksellisuustarkastus määritellään samoilla termeillä: ”Tuloksellisuustarkastuksessa tarkastetaan valtion taloudenhoidon tarkoituksenmukaisuutta eli sitä, että valtion varoja käytetään *taloudellisesti, tehokkaasti ja vaikuttavasti*.” Molemmat näistä määritelmistä ovat melko suoria sovelluksia kolmen E:n mallista, jota kuvaillaan seuraavassa alaluvussa osana tämän tutkimuksen teoreettista viitekehystä.

2.2 Teoreettinen viitekehys

2.2.1 Kolmen E:n malli

Tämän tutkimuksen teoreettisessa viitekehyksessä tuloksellisuuden merkitystä kuvaamaan käytetään kolmen E:n mallia. Tämä malli on kansainvälisesti yleinen tapa jäsenellä nimenomaan julkissektorin toiminnan tuloksellisuutta. Virtanen ja Vakkuri (2016, 106) kuvaavat kuinka malli on pyrkimys laajentaa vajavaista yksityisen sektorin liiketaloudellista lähestymistapaa tuloksellisuuteen osaksi laajempaa yhteiskunnallista arviointia. Kolmen E:n mallin voidaan nähdä jalostuneen nykyiseen muotoonsa pitkän ajan kuluessa. Mallin mukaisia alakäsitteitä on käytetty jo 1960-luvulla ohjelmabudjetoinnin yhteydessä. (Meklin 2009, 36.)

Kolmen E:n mallin kirjaimet tulevat englannin kielen sanoista *economy* eli taloudellisuus, *efficiency* eli tuottavuus sekä *effectiveness* eli vaikuttavuus. Yleisin tapa (esim. Kortelahti 2019, Virtanen & Vakkuri 2016, Carter 1992) havainnollistaa mallia toiminnassa on liittää se seuraavaksi esiteltävän kuvion mukaiseen asetelmaan, jossa tietyistä panoksista eli tuotannontekijöistä muunnetaan tuotantoprosessissa haluttuja tuotoksia eli suoritteita. Lainaten Virtasen ja Vakkurin määritelmää kuviossa tarpeet voidaan nähdä tuotettujen suoritteiden eli esimerkiksi palveluiden kohteiksi. Tiettyä palvelua tuotetaan siis jonkin yhteiskunnallisen tarpeen tai ongelman ratkaisemiseksi.



Kuvio 1: Kolmen E:n malli panos-tuotos-asetelmassa (Mukaien Virtanen & Vakkuri 2016)

Mallissa taloudellisuus kuvaa sitä, kuinka paljon haluttujen suoritteiden aikaansaaminen kustantaa. Taloudellisuutta tarkasteltaessa on siis kiinnostuttu tuotannon tekijöistä nimenomaan rahamääräisinä kustannuksina. Havainnollistaen taloudellisuus voi siis tarkoittaa esimerkiksi sitä, miten paljon maksaa yhden hoitopäivän järjestäminen sairaalassa. Erityisesti tarkasteltaessa julkista toimintaa, ei taloudellisuuden arvioinnissa voida unohtaa laadun merkitystä. Tällöin voisi olla tarkoituksenmukaista määritellä taloudellisuus esimerkiksi tietyn tasoisten suoritteiden tuottamiseksi mahdollisimman alhaisin kustannuksin. (Carter 1992, 37.)

Tuottavuuden käsite mallissa kuvaa tuotannon tekijöiden sekä suoritteiden välistä suhdetta. Se eroaa taloudellisuudesta siinä, että tuottavuudessa tarkastellaan nimenomaan reaalitalouden yksiköitä, esimerkiksi henkilön määrää tai tehtyjä työtunteja (Meklin 2009, 36). Aiempaan sairaalaan liittyvään esimerkkiin viitaten tuottavuutta tarkasteltaessa voitaisiin olla siis kiinnostuneita, kuinka monta työtuntia yhden potilaan hoitamiseen käytetään. Kun halutaan parantaa toiminnan tuottavuutta, on huomattava, että se voidaan tehdä joko lisäämällä suoritteiden määrää suhteessa tiettyyn tuotannon tekijöiden määrään tai vaihtoehtoisesti tuottamalla sama määrä suoritteita pienemmällä määrällä tuotannon tekijöitä.

Mallin kolmesta E:stä vaikuttavuus aiheuttaa ehkä eniten hämmennystä ja on vaikein määritellä (Carter 1992, 38). Samalla se voidaan kuitenkin nähdä tärkeimmäksi tavoitteeksi julkishallinnon toiminnan suhteen (Virtanen & Vakkuri 2016, 114). Yksinkertaistaen voidaan todeta vaikuttavuuden kuvaavan sitä, miten hyvin toiminnassa tuotetut suoritteet vastaavat tunnistettuihin ja osaltaan myös tunnistamattomiin tarpeisiin. Vaikuttavuudessa ei siis olla niinkään kiinnostuttu esimerkiksi tuotettujen palveluiden määrästä, vaan niiden tuottamista vaikutuksista. Havainnollistaen

vaikuttavuus voi tarkoittaa esimerkiksi sitä, miten sairaalan toiminta parantaa alueellista kansalaisten terveyden tilaa. Arvioinnin kannalta vaikuttavuuden käsittelyä vaikeuttavat esimerkiksi palvelujen monipuolisuus sekä pitkän aikavälin kustannusten ja vaikutusten välinen epäselvyys (Ulfes ym. 2010, 46).

Kolmen E:n mallia on kasvavissa määrin kritisoitu siitä, ettei se ota huomioon sen kolmen alakäsitteen ulkopuolelle jääviä tekijöitä. Onkin ehdotettu, että mallia laajennettaisiin neljännellä E:llä, *ethics*. Tämä eettinen ulottuvuus mallissa ottaisi huomioon esimerkiksi organisaation sosiaaliset suhteet sekä työympäristön. (Bringselius 2018.) Myös muita tekijöitä on ehdotettu neljänneksi E:ksi laajentamaan mallia. Gildenius ja Janse van Rensburg (2017) ehdottavat eettisyyden lisäksi ympäristön (*environment*) tai oikeudenmukaisuuden (*equity*) mukaan ottamista malliin.

2.2.2 Päämies–agentti-malli

Arvioitavan kohteen sekä arvioinnin tilaajan välistä suhdetta voi tarkastella päämies–agentti-mallia hyödyntäen. Tätä mallia käytetään usein kuvaamaan tilannetta, jossa yksi toimija on delegoinut jonkin tehtävän suorittamisen toiselle toimijalle (Eisenhardt 1989). Nämä kaksi toimijaa muodostavat teorian mukaiset agentin ja päämiehen roolit. Petersenin (1993, 278) mukaan päämiehen ja agentin välinen riippuvuus syntyy, kun agentilla on tehtävään paremmin soveltuvat taidot tai jos päämies kokee pystyvänsä käyttämään aikansa tehokkaammin muihin tehtäviin.

Päämies–agentti-malliin liittyvät vahvasti oletukset päämiehen ja agentin välisistä tavoiteristiriidoista sekä tiedon epäsymmetrisyydestä. Tavoiteristiriidassa on kyse siitä, kun agentti toimii omien etujensa mukaisesti päämiehen etujen kustannuksella. Tiedon epäsymmetrisyydellä taas tarkoitetaan oletusta, jonka mukaan agentilla on usein enemmän tietoa omista motivaatioistaan sekä tehtävän suorittamisesta. Epäsymmetria syntyy, kun tämä tieto ei jakaudu tasaisesti agentin ja päämiehen välillä. (Kivistö 2009.)

Agentin toiminnasta suoritettun arvioinnin voidaan nähdä olevan keino hallita tavoitteiden ristiriitoja sekä pienentää tiedon epäsymmetriaa. Päämies saa arvioinnista tietoa agentin toiminnasta esimerkiksi tulospohjaisia mittareita tarkasteltaessa. Syvällisemmässä arvioinnissa tietoa saadaan mahdollisesti myös agentin tavoitteista päämiehen tavoitteisiin nähden. Petersen (1993) kuvaa myös, kuinka päämies–agentti-mallin mukaisessa asetelmassa toiminnasta syntyneet tuotokset syntyvät agentin vaikutuksen lisäksi ulkoisten tekijöiden vaikutuksesta. Päämies tai agentti eivät pysty omalla toiminnallaan vaikuttamaan näihin tekijöihin, joten tuotoksiin liittyy aina tietynlainen epävarmuus.

2.2.3 Arvioinnin funktiot

Tutkimuskirjallisuudessa tuloksellisuusarvioinnille määritellään usein kolme funktiota (Arvidsson 1986; Niiranen 2008; Virtanen & Vakkuri 2016). Tässä tutkimuksessa mukaillaan Vedungin (2017, luku 6) versiota tästä kolmijaosta. Ensinnäkin arvioinnin tuottamaa informaatiota voidaan hyödyntää päämies-agentti-mallin mukaisesti agentin toiminnan legitimitetin vahvistamiseen. Tämän tutkimuksen valossa kyseessä voisi olla tietyn viraston toiminnan tarkoituksenmukaisuuden vahvistaminen eduskunnalle. Vedung (2017) esittelee tästä agentin tilivelvollisuudesta mielenkiintoisen seikan: suurin osa hallinnollisista toimijoista toimii samaan aikaan sekä agentteina ylemmille toimijoille, että päämiehinä alemmille toimijoille. Arvioinnin kannalta tämä on mielenkiintoinen huomio, sillä se voi johtaa siihen, että myös päämiehet välttelevät tekemästä liian kriittisiä arviointia peläten sen johtavan lopulta päämiehen oman toiminnan arviointiin.

Toisena käyttötarkoituksena organisaation johto hyödyntää arviointitietoa toiminnan ohjaamisessa. Tarkoituksena on siis käyttää arvioinnista saatua dataa esimerkiksi toiminnan kustannustehokkuuden parantamiseen tai sen kohdistamiseen paremmin asiakkaalle. Virtasen ja Vakkurin (2016) listauksessa tuloksellisuusarvioinnin hyödyntämistavoista ohjaamista vastaava funktio kuvataan tuloksellisuudesta palkitsemisena ja rankaisemisena. He kuvaavat tämänkaltaisen ohjauksen linkittyvän 1980-luvulla tapahtuneeseen julkishallinnon uudistamiseen entistä tulostavoitteellisempaan suuntaan.

Kolmantena funktiona arvioinnille on tiedon kerääminen toiminnasta. Tässä arvioinnin tarkoituksena on ymmärtää paremmin arvioinnin kohdetta ja sen toimintaa. Tiedon keräämisen voidaan nähdä liittyvän päätöksenteon rationaalisuuteen perustumiseen sekä tämän edistämiseen (Virtanen & Vakkuri 2016). Tämä arvioinnin funktio voidaan ymmärtää kuitenkin enemmän akateemisena kuin käytännönläheisenä toimintona (Vedung 2017).

Funktio	Tilivelvollisuus	Ohjaaminen	Tiedon lisääminen
Käyttötarkoitus	Toiminnan legitimitietin vahvistaminen	Toiminnan suunnan ohjaus sekä tehostaminen	Yleinen tiedon lisääminen sekä ymmärtäminen
Kohderyhmä	Päämies / viiteryhmät	Organisaation johto	Tieteellinen yhteisö / muut kiinnostuneet

Kuvio 2: Arvioinnin funktiot kolmijakona

2.2.4 Arvioinnin haasteet

Tässä tutkimuksessa aineiston raportteja käsitellään pohjaten vahvasti Evert Vedungin kirjassa *Public policy and program evaluation* (2017, 93–99) esittämään arvioinnin kahdeksan haasteen malliin. Vedungin mukaan juuri nämä arvioinnin haasteet antavat arvioinnille sen identiteetin verrattuna tieteelliseen tutkimuksen tekoon. Mallin mukaiset kahdeksan haastetta esitetään seuraavina kysymyksinä.

1. Mihin tehdyllä arvioinnilla pyritään? Onko arvioinnin tilaajalla tai tekijällä pyrkimys aidosti selventää arvioinnin kohdetta vai onko taustalla esimerkiksi tilivelvollisuuteen liittyviä kysymyksiä?
2. Kenen tulisi suorittaa arviointi, ja miten se tulisi organisoida?
3. Miten arvioinnin kohde on määritelty ja rajattu? Tämä kysymys on relevantti erityisesti VTV:n raportoinnin kohdalla, ottaen huomioon tarkastusviraston käsittelemien arvioinnin kohteiden lukuisuuden sekä erilaisuuden.
4. Miten arvioinnin kohde on toteutettu? Millaisin valinnoin on päädytty esimerkiksi jonkin toimintaohjelman vireillepanosta lopullisiin tuotoksiin?
5. Miten arvioinnin kohteen tuotokset ja vaikutukset on määritelty? Tämä määrittely voi olla arvioinnin kohteesta riippuen hyvinkin vaikeaa. Tuotosten arviointiin liittyy vahvasti myös arvioinnin ajallisen keston rajallisuus.
6. Millaisia epävarmuustekijöitä arvioinnin kohteeseen sisältyy, ja miten ne vaikuttavat sen arviointikelpoisuuteen?

7. Millä perusteilla arviointia tehdään? Millä tasolla arvioinnin kohteen voidaan sanoa onnistuneen, ja mitkä ovat sen varsinaiset saavutukset?
8. Miten suoritettua arviointia hyödynnetään sen valmistuttua?

Vedung jakaa mallin haasteet kolmeen eri tasoon, riippuen siitä, mihin arviointiprosessin vaiheeseen mikin haaste sijoittuu. Ensimmäinen ja toinen haaste käsittelevät kohteen arviointia itsessään. Haasteet 3–7 taas käsittelevät arvioinnin kohteena olevaa interventiota. Mallin kahdeksas haaste käsittelee sitä, miten kohteesta tehtyä arviointia hyödynnetään sen valmistuttua. Tätä kahdeksatta kysymystä voidaan peilata edellisessä alaluvussa esitettyihin arvioinnin funktioihin.

3 Tutkimuksen toteutus

3.1 Aineisto

Aineistona tutkimuksessa käytettiin Valtion tarkastusviraston (myöhemmin VTV) verkkosivuilta löytyviä tuloksellisuustarkastuskertomuksia. Tätä aineistoa tuettiin lisäksi lausunnoilla, joita tarkastuskohteet ovat antaneet kertomuksien luonnoksista. VTV määrittelee itsensä riippumattomaksi valtion taloudenhoidon ammattitarkastajaksi sekä finanssipolitiikan ja vaali- ja puoluerahoituksen valvojaksi. VTV:n tehtävät määritellään tarkemmin Suomen perustuslaissa. Tiivistettynä sen perustehtäviin kuuluvat kuitenkin valtionalouden hoidon laillisuuden ja tuloksellisuuden tarkastaminen sekä eduskunnan taloudenhoidosta saaman informaation paikkansapitävyys (Valtiovarainministeriö).

VTV julkaisee vuosittain noin 15–20 tuloksellisuusarviointikertomusta. Tähän tutkimukseen aineistoksi rajattiin näistä kertomuksista kahdeksan raporttia. Rajaus tehtiin kahden tekijän perusteella. Ensimmäinen rajaustekijä on raporttien aikajänne siten, että tutkimukseen mukaan otetut raportit ovat vuosilta 2017–2019. Tällä rajauksella varmistettiin tutkimuksen aineiston ajankohtaisuus. Toinen rajaustekijä tulee VTV:n verkkosivuilta löytyvistä neljästä vaikuttavuusalueesta. Nämä alueet ovat *kestävä julkinen talous, kestävä hallinnon kehittäminen, hyvinvoiva ja turvallinen yhteiskunta* sekä *toimiva tiedonhallinta*. Tutkimuksen aineistoon on pyritty ottamaan jokaisesta mainitusta alueesta raportteja, jotta saadaan laaja kattaus yhteiskunnallisesta tuloksellisuusraportoinnista. Tällä valinnalla pyrittiin myös välttämään sellainen tilanne, että

tutkimukseen päätyisi vain tietylle aihealueelle tyypillisiä arvioinnin haasteita. Tutkimuksen aineiston raportit ovat julkaisujärjestyksessä otsikoitu seuraavasti:

1. Innovaatiostrategian toteuttaminen julkisissa hankinnoissa, 2. Sähköisten palvelujen toimintavarmuuden ohjaus, 3. Valtio cleantech-strategian toimeenpano, 4. Kotoutumiskoulutus, 5. Valtioneuvoston biotalousstrategian valmistelu, 6. Omaisuuden elinkaaren hallinnan ohjeet, 7. Ohjauksen vaikutus ensihoitopalvelun toimivuuteen, 8. Avoin tieto ja viranomaistiedon hyödyntäminen

3.2 Menetelmä

Tämän tutkimuksen ensisijainen pyrkimys on selvittää, millaisia haasteita käytännön arviointiraportoinnissa esiintyy sekä millaisia valintoja niissä ollaan tehty. Nämä kysymykset ovat lähtökohtaisesti laadullisia, joten myös tutkimusmenetelmät muodostuivat kvalitatiivisiksi. Laadullisessa tutkimuksessa ei ole tarkoituksena niinkään muodostaa yleistettäviä tuloksia kohteesta, vaan muodostaa syvällisempää ymmärrystä tutkimuksen aiheesta. Tutkimuksen aineistoa käsitellään aineistolähtöisen sisällönanalyysin kautta. Tuomen ja Sarajärven (2009, luku 4) mukaan aineistolähtöinen laadullinen analyysi etenee yksinkertaisten seuraavasti. Ensin aineisto pilkotaan tutkimuksen kannalta merkityksellisiksi osiksi, joille sen jälkeen annetaan merkitys käsitteellistämisen kautta ja jotka lopuksi järjestellään uudelleen loogiseksi kokonaisuudeksi perustuen tutkijan tulkintaan aineistosta. Tämä prosessi on tämän tutkimuksen kontekstissa kuvattu seuraavissa kappaleissa. Tämän lisäksi tutkimuksessa on tarkoituksena teorialähtöistä sisällönanalyysiä lainaten haastaa Vedungin kahdeksan haasteen mallia. Analyysiä aloittaessa huomattiin, ettei malli suoraan soveltunut tämän tutkimuksen aineistoon rakenteeltaan, joten tutkimuksen kannalta hedelmällisempää oli lähestyä aineistoa nimenomaan aineistolähtöisesti.

Aineiston analyysi aloitettiin aineistolähtöisen sisällönanalyysin mukaisesti perehtymällä raporteihin ja poimimalla tekstistä ilmauksia ja lauseita jotka viittasivat arvioinnin suorittamista vaikeuttaviin tekijöihin. Tätä vaihetta kuvataan usein (esim. Tuomi & Sarajärvi 2009) laadullisessa tutkimuksessa aineiston pelkistämiseksi tai redusoinniksi. Koska tämän tutkimuksen aineisto ei suoraan käsitellyt varsinaisia arvioinnin haasteita, vaan niiden löytäminen vaati aineiston tulkintaa, pyrittiin nämä maininnat poimimaan tekstistä kokonaisten lauseiden tai kappaleiden mukana. Näin varmistuttiin siitä, että maininnat ovat oikeassa kontekstissa tutkimuskysymykseen nähden. Jokaisesta raportista muodostettiin tällä menetelmällä erillinen lista haasteisiin viittaavista ilmauksista.

Tämän jälkeen raporteista muodostettuja listoja vertailtiin keskenään ja pyrittiin tätä kautta löytämään niiden välisiä samankaltaisuuksia sekä eroja. Laadullisen tutkimuksen termein havaintoja siis klusteroitiin omiksi ryhmikseen. Alasuutari (2001, 35) korostaa pelkistämisestä saatujen havaintojen tulkintaa tutkimuksen oman kontekstin ja viitekehyksen kautta. Myös tässä tutkimuksessa tulkinta toteutettiin nimenomaan arvioinnin haasteiden näkökulmasta. Löydettyjen samankaltaisuuksien perusteella muodostettiin haasteita kuvaavia alaluokkia, jotka ylittivät tässä vaiheessa jo raporttien väliset rajat. Muodostettujen alaluokkien voidaan ajatella kuvaavan analyysissä löydettyjä yksittäisiä arvioinnin haasteita.

Analyysiä tehdessä huomattiin nopeasti näiden alaluokkien muodostuvan kolmeksi erilliseksi ryhmäksi, joista muodostui analyysin yläluokat. Näiden kolmen yläluokan ulkopuolelle jäi kuitenkin kaksi ”irralista” haastetta, jolle ei varsinaisesti löydetty yhdistävää tekijää muiden yläluokkien tapaan. Nämä haasteet koottiin yhteen neljänneksi luokaksi. Yläluokkien muodostaminen voidaan nähdä tässä tutkimuksessa aineiston abstrahoinniksi, eli klusteroinnin kautta saatujen ryhmien käsitteellistämiseksi tutkimuksen käyttöön. Analyysin varsinainen tulos sekä yllä kuvattu analyysin rakenne on pyritty kuvaamaan neljänneksen luvun alussa esitetyssä kuviossa.

Aineiston analyysin tulkinnallinen puoli toteutettiin peilaten vahvasti Vedungin (2017, 93–99) arvioinnin kahdeksan haasteen malliin, joka on esitelty tämän tutkimusraportin teoreettisessa viitekehyyksessä. Muodostaessa tutkimuksen ala- ja yläluokkia pyrittiin samalla pohtimaan, miten mallin mukaiset haasteet esiintyvät suhteessa aineistosta löydettyihin haasteisiin. Aineistoa analysoidessa kiinnitettiin myös huomiota mahdollisiin haasteisiin, joita ei esiintynyt Vedungin mallissa, tai mallin haasteita, jotka eivät syystä tai toisesta esiintyneet aineistossa. Näin tutkimuksen analyysiä toteuttaessa samalla analysoitiin teorialähtöistä sisällönanalyysiä lainaten Vedungin mallin soveltuvuutta tämän kaltaiseen meta-analyysiin.

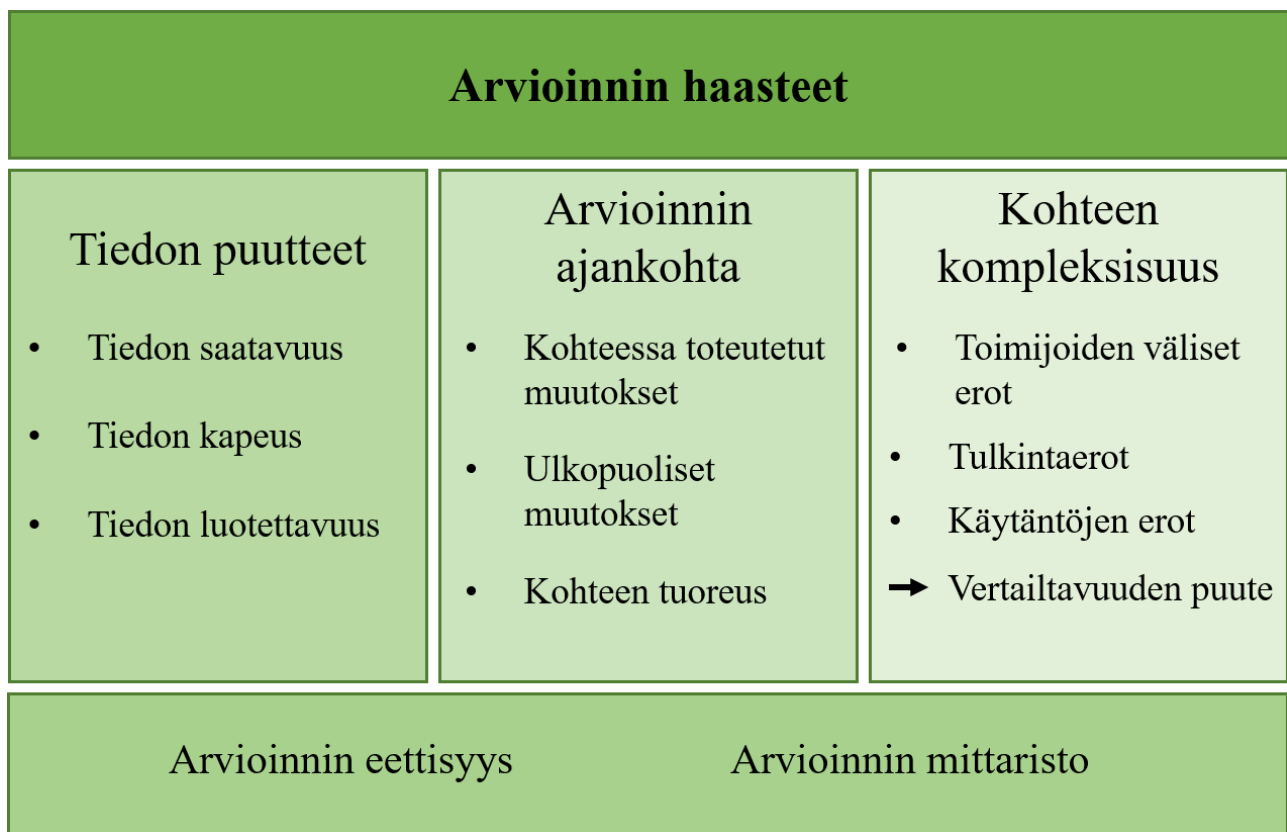
4 Analyysi

4.1 Analyysin rakenne

Tässä luvussa käsitellään tutkimuksen aineiston varsinainen analyysi. Analyysin esittely etenee mukailien alla olevaa kuviota, joka samalla tiivistää analyysin tulokset. Kuvion kolme vierekkäistä osiota kuvaavat analyysissä muodostettuja kolmea yläluokkaa, joiden alle aineistossa yleisesti esiintyvät haasteet on luokiteltu. Tämän lisäksi yläluokkien alapuolella kuvioon on luokiteltu

”irtonaisia” haasteita, joita ei koettu tarkoituksenmukaiseksi yhdistää samalla tavalla yläluokaksi ottaen huomioon niiden vähäisen esiintymisen aineistossa. Kyseiset haasteet kuitenkin koettiin merkityksellisiksi tämän tutkimuksen tutkimuskysymyksen kannalta. Kokonaisuudessaan kuvion tarkoituksena on esitellä julkissektorin arvioinnin yleisiä haasteita tämän tutkimuksen tuloksina.

Analyysiä viedään eteenpäin käsittelemällä ensin jokaista kuvion kolmesta yläluokasta. Samalla löydettyjä haasteita pyritään peilaamaan tutkimuksen teoriapohjaan, erityisesti aiemmin esiteltyyn Vedungin malliin. Kolmen yläluokan jälkeen analyysissä käsitellään lyhyesti myös loput irralliset haasteet omana alalukunaan. Analyysin huomioita havainnollistetaan lainauksilla aineiston raporteista. Analyysin läpinäkyvyyden varmistamiseksi lainausten perään on merkitty raportin numerolla, mistä raportista kyseessä oleva lainaus on peräisin. Raportit on luettavuuden kannalta numeroitu niiden julkaisujärjestyksen mukaan, mikä löytyy kuvattuna tämän tutkimusraportin toisesta pääluvusta.



Kuvio 3: Julkishallinnon interventtioiden arvioinnin haasteet

Ennen analyysiin siirtymistä on todettava, että haasteiden määrä sekä monipuolisuus vaihtelivat huomattavasti sen suhteen, miten ne raporteissa tuodaan esille. Aineistosta ei pystytty kuitenkaan yleistämään syytä tälle vaihtelulle esimerkiksi tietyn toimialan tai tarkastuksen aihepiirin suhteen. Tästä huolimatta jokaisesta raportista saatiin analysoitua tämän tutkimuksen kiinnostuksen mukaisia arvioinnin haasteita, joten jokainen raporteista oli tarkoituksenmukaista pitää mukana aineistossa.

4.2 Tiedon puutteet

Jo aineiston analyysin alkuvaiheessa havaittiin, että merkittävä osa raporttien haasteista liittyi tiedon hankintaan, sekä sen käyttökelpoisuuteen arvioinnin toteutuksessa. Tietoon liittyvät erilaiset haasteet nousivat esille jokaisessa raportissa. Yksi arviointia varten kerättävän tiedon hankintaa vaikeuttava tekijä on, ettei kohdeorganisaatiossa ole seurattu ja raportoitu toiminnan etenemistä. Raporteista nousi esille tätä selittäviksi tekijöiksi esimerkiksi hankkeen puutteellinen suunnittelu sen alkuvaiheessa, sekä toiminnan raportoinnin vastuiden epäselvyys. Nämä haasteet nousevat esille varsinkin intervention *jälkeen* tehdyissä arvioinneissa. Raportissa 1 tuodaankin esille tarve tulevan arvioinnin integroimiselle hankkeeseen jo sen laatimisvaiheessa.

”Tarkastuksessa haastateltujen mukaan strategian toimeenpanon seuranta oli vähäistä. Seuranta-asiakirjoja ei tarkastuksessa löytynyt.” –Raportti 1

Interventioiden *taloudellisuuteen* liittyvät tiedon rajoitteet nousivat aineistosta selkeästi erilaisena verrattuna muihin tuloksellisuuden osa-alueisiin. Jokaisessa raportissa, jossa hankkeen talousinformaatiota käsiteltiin, tietoa oli löydettävissä. Silti osassa raporteista talousinformaatio ei ollut riittävän laajaa varsinaisen taloudellisuuden arvioimiseen. Aineistosta olikin tulkittavissa, että usein talousinformaatiota tuotetaan organisaatioissa esimerkiksi lainsäädännön lähtökohdista toiminnan arvioimisen ja kehittämisen sijaan.

”[...]hankintojen taloudellisuutta ei voitu luotettavasti arvioida käytettävissä olleen tiedon perusteella.” –Raportti 2

Toisaalta, vaikka raportointia olisi hankkeen kulkua olisi raportoitu, voi näin tuotetun tiedon hankkiminen arviointiin jälkikäteen olla hankalaa. Joissain kohteissa tuotettu tieto oli ikään kuin ”kadonnut” organisaatiosta. Kyseessä olevista raporteista välittyi kuva, ettei arviointia oltu välttämättä nähty todennäköisenä tai merkittävänä tekijänä intervention jo valmistuttua, eikä tiedon säilymistä oltu tästä johtuen huomattu varmistaa.

”[...]kaikki asiakirjat oli koottu yhteen ja viety ministeriön arkistoon. Kukaan ei kuitenkaan tiennyt asiakirjojen olinpaikkaa.” –Raportti 1

”[...]kuljetuksista ei tällä hetkellä saada luotettavaa valtakunnallista tietoa, vaan tiedot täytyy poimia eri rekistereistä.” –Raportti 5

Yllämainittuja tiedon hankintaan liittyviä haasteita yhdistävä tekijä on arvioitavien organisaatioiden hajautuneisuus sekä pirstaloituneisuus. Useista aineiston raporteista nousi selvästi esille, että arvioitava interventio oli järjestetty monen eri tahon yhteistyössä. Tämä nousi monessa raportissa esille toiminnan vastuuketjujen ja tehtävänjaon epäselvyytenä, mikä osaltaan vaikeutti arviointiin tarvittavan tiedon löytämistä.

Toinen tietoon liittyvä merkittävä tekijä on sen käytettävyys arvioinnin kannalta. Tiedon käytettävyys on tässä yhteydessä kaksijakoinen ongelma. Ensinnäkin tuotettu tieto voi olla liian kapea-alaista tai epätarkkaa arvioinnin tarpeisiin. Toisaalta tuotettu tieto voi olla arvioinnin kannalta epäluotettavaa, eikä sitä pystytä varmentamaan. Näistä jälkimmäinen vaihtoehto korostui aineistossa erityisesti haastatteluiden avulla saadussa informaatiossa. Aineiston raporteista oli selkeästi havaittavissa tietynlainen epäluottamus, jos tarkastuksessa oltiin jouduttu tukeutumaan pelkkään haastatteluaineistoon. Haastattelut näyttäytyvät aineistossa toissijaisina lähteinä, ja mitattavan, määrällisen datan puuttuessa nämä haasteet tuotiin selvemmin esille myös arvioinnin raportointiin. Tämänkaltainen tilanne on tulkittavissa päämies–agentti-mallissa esiintyvän tiedon epäsymmetrian kautta. Arvioijalla ei ole varmuutta haastateltavien todellisista motiiveista, joten tarkastuksessa turvaudutaan mieluummin ”kovaan” dataan, jossa ei ole mukana inhimillisiä epävarmuustekijöitä.

”Koska selvitys perustui kapeaan tilastoaineistoon ja on tehty melko pian uusien mallien käyttöönoton jälkeen, se ei ole riittävä arviointi uusien mallien toimivuudesta.” –Raportti 2

”[...]Palvelujen käytöstä ei kuitenkaan saatu tarkastuksessa seurantaan perustuvaa tietoa.” –Raportti 2

”Nyt ei esimerkiksi joistain suunnitteludokumenteista voitu varmuudella sanoa, oliko niitä ylipäättään edes tehty.” –Raportti 1

Saatavilla olevan tiedon laajuus sekä puutteet näyttäytyivät aineistossa tarkastuksen rajauksia ohjaavina tekijöinä. Joissain raporteissa tarkastuksen alkuperäistä tavoitetta oltiin jouduttu muokkaamaan, ja osa tavoitteista oli jouduttu jättämään kokonaan pois. Tarkastuksen tavoitteiden uudelleenasettelussa on havaittavissa yhteys Vedungin mallin ensimmäiseen arvioinnin haasteeseen. Mikä on ollut käynnistetyn arvioinnin alkuperäinen tarkoitus? Tässä yhteydessä onkin tarkoituksenmukaista miettiä myös Vedungin mallin kahdeksannen haasteen mukaisesti, mihin tarpeeseen arvioinnilla pyritään vastaamaan sekä miten arviointia tullaan hyödyntämään sen

valmistuttua. Voidaan kyseenalaistaa, miten mielekästä arviointia on jatkaa, jos sen tuloksina saatava informaatio kertoo merkittävästi eri kohteista kuin oli sitä aloittaessa tarkoituksena.

Vedungin mallin toinen haaste kyseenalaistaa arvioinnin tekijän roolin: tulisiko arviointi suorittaa ulkopuolisen tekemänä vai itsearviointina organisaatiossa. Tämä kysymyksenasettelu ei suoraan istu VTV:n toimintaan ottaen huomioon sen roolin ulkopuolisena tarkastajana, mutta yllä esitettyjen tiedon puutteiden perusteella voidaan silti miettiä arvioinnin toteuttamista uudelleen. Tulisiko julkisten interventioiden arviointikulttuuria ohjata itsearvioivampaan suuntaan? Tällä saattaisi olla esimerkiksi positiivisia vaikutuksia toiminnan seuraamisen ja tiedon arkistoinnin suhteen.

4.3 Arvioinnin ajankohta

Tietoon liittyvien haasteiden lisäksi toiseksi merkittäväksi haasteeksi aineistosta nousi suoritettujen tarkastusten ajankohta. Erityisesti arvioinnin kohteena olevaan interventioon kohdistuvat muutokset hankaloittavat sopivan ajankohdan valitsemista. Aineistosta ei erityisesti selvinnyt, miten VTV on ottanut huomioon esimerkiksi tarkastuksen tavoitteet suhteessa sen toteuttamisajankohtaan.

”Strategian päättymisestä oli kulunut niin vähän aikaa, että sen kaikki vaikutukset eivät olleet vielä näkyvissä.” –Raportti 1

”[...]strategioiden vaikuttavuutta ei ollut mahdollista arvioida, koska aineistoa arviointiin ei ollut riittävästi saatavilla näiden strategioiden uutuuden takia.” –Raportti 6

Aineiston perusteella myös tarkastuksen ajankohta vaikuttaa sen rajauksiin. Esimerkiksi raportissa 4 kuvatun tarkastuksen aikana Euroopan unioni oli antanut aiheita koskevan uuden direktiivin, joka vaikutti merkittävästi arvioinnin lopputuloksiin. Tässä tapauksessa tarkastusluonnoksesta annetuissa lausunnoissa arviointia kritisoidaan siitä, ettei suosituksia voitu pitää tarkoituksenmukaisina ennen kuin asiasta annettuun uuteen direktiiviin oltiin perehdytty. Samassa lausunnossa myös kritisoidaan epäselvyyttä siitä, kumpaan direktiiviin suosituksissa viitataan. Kaiken kaikkiaan tarkastuksen aikana kohteen ulkopuolelta annetut muutokset näyttävät suuren haasteena arvioinnille.

”Kaikkien direktiivien huomioon ottaminen olisi pirstaloittanut ja laajentanut tarkastusta varsinkin, koska tarkastuksen meneillään ollessa tietovarantojen avaamisen alueella tapahtui paljon...”

–Raportti 4

Yllä esitetyn kaltainen, arvioinnin kohteen antama palaute arvioinnista luo mielenkiintoisen asetelman arvioinnin haasteiden näkökulmasta. Miten arvioijan tulisi suhtautua kohteesta itsestään lähtöisin olevaan palautteeseen? Toisaalta palautteessa on selvä mahdollisuus objektiivisuuden

puutteeseen, mutta toisaalta kohteessa on epäilemättä kattava tietämys aiheesta. VTV on tarkastuksissaan johdonmukaisesti ottanut huomioon lausunnoissa esiin tulleet teknisluontoiset korjaukset, mutta varsinaiseen sisältöön puuttuva kritiikki on käsitelty tilanteen mukaan. Täten voidaan ajatella VTV:n toimivan Vedungin mallin kahdeksannen haasteen mukaisesti, jossa kiinnitetään huomio siihen, ketä varten arviointia tehdään ja mihin tarkoitukseen sitä tullaan sen valmistuttua hyödyntämään. Yksi mielenkiintoinen näkökulma on, miten arvioinnin kohteesta oleva kritiikki otetaan huomioon arvioinneissa, joita hyödynnetään toiminnan legitimitietin vahvistamiseen verrattuna esimerkiksi toiminnan ohjausta varten suoritettaviin arviointeihin.

Myös arvioinnin kohteen sisällä tehtävät muutokset vaikeuttavat pysyvien ja ajantasaisten suositusten antamista. Useissa kohtaa raporteissa annetuissa suosituksissa oli jouduttu toteamaan tilanteen luultavasti muuttuvan parempaan suuntaan käynnissä olevan hankkeen valmistuttua tai suunnitteilla olevan järjestelmän käyttöönoton jälkeen. Joidenkin hankkeiden kohdalla tarkastukseen ei oltu saatu käyttöön tulevan hankkeen suunniteltuja vaikutuksia, mikä osaltaan hankaloitti tulevan tilanteen arviointia. Suuri osa muutoksista tuntui liittyvän esimerkiksi tietokantojen parantamiseen sekä muihin tietoteknisiin ratkaisuihin. Tältä osin muutokset arvioinnin haasteena voisikin liittää laajempaan muutostrendiin kuten julkishallinnon digitalisoitumiseen.

”Sen jälkeen, kun tarkastusvirasto otti tarkastusaiheen tarkasteltavakseen ja käsiteltäväkseen, useita viranomaistiedon hyödyntämiseen liittyviä hankkeita on päättynyt ja käynnistetty.”

–Raportti 4

Arvioinnin kohteeseen kohdistuvien muutosten lisäksi aineistosta nousi yksittäisiä mainintoja myös muista tarkastuksen ajankohtaan liittyvistä tekijöistä. Kohteena olevan intervention tuoreus vaikeutti selvästi pidempiaikaisten vaikutusten hahmottamista, kuten seuraavassa kappaleessa tarkemmin kuvataan. Tämän lisäksi yhteiskunnassa tapahtuvat toimintaympäristön muutokset monimutkaistavat suositusten laatimista. Esimerkiksi raportissa 5 ikärakenteen vanheneminen tunnistettiin hoitopalvelujen tarvetta muuttavaksi tekijäksi. Tämän kaltaiset yhteiskunnassa luontaisesti tapahtuvat suuret muutokset pakottavat arvioitsijan peilaamaan arvioinnin kohdetta sen toimintaympäristöä vasten. Kohdetta ei siis voida erottaa omaan ”kuplaansa”. Toisaalta tähän liittyen myös muiden toimijoiden vireillä olevat hankkeet voivat luoda päällekkäisyyttä arvioinnin kohteena olevan hankkeen vaikutusten kanssa. Tällaisessa tilanteessa voi olla vaikea kohdentaa, mitkä tulokset ovat juuri kyseisen hankkeen tuottamia.

Tuloksellisuuden kannalta interventioiden *vaikuttavuus* jäi usein tarkastuksissa vähälle huomiolle. Osaltaan tämä on seurausta seurantatiedon puutteesta, mutta erityisesti tarkastuksen

toteuttamisajankohta näytti vaikeuttavan kohteena olevan hankkeen vaikuttavuuden arviointia. Jos kohteena oleva interventio on liian tuore tai siihen ollaan lähitulevaisuudessa toteuttamassa muutoksia, ei tarkastuksessa voida tehdä luotettavasti ajantasaista arviointia sen vaikutuksista. Raportissa 1 mainitaan myös tarkastellun strategian lyhytikäisyys sen vaikuttavuuden hahmottamista vaikeuttavana tekijänä. Tutkimuskirjallisuuden pohjalta vaikuttavuuden arvioinnin vähäisyys ei tullut yllätyksenä. Esimerkiksi Virtanen ja Vakkuri (2016) kuvaavat vaikuttavuuden olevan vaikeimmin hahmotettava tuloksellisuuden osa-alueista.

Mietittäessä arvioinnin ajankohtaa haasteena, on havaittavissa yhteys ainakin Vedungin mallin toiseen haasteeseen. Mikä merkitys ajankohdalla on mietittäessä arvioinnin toteuttamisen järjestelyitä? Yllä esitetyt haasteet selvästi kuvaavat kuinka arvioinnin toteuttamisen ajankohta määrittelee paitsi arvioinnin rajauksia myös siitä syntyviä johtopäätöksiä. Toisaalta jos ajankohtaa käsitellään *vaikutusten* arvioinnin kautta, on havaittavissa yhteys myös Vedungin mallin viidenteen sekä kuudenteen haasteeseen, jotka koskevat arvioinnin kohteena olevan intervention tulosten ja vaikutusten määrittelyä sekä muiden yhteiskunnassa tapahtuvien muutosten erottamista arvioinnin kohteesta. Vaikuttavuuden määrittely vaatii arvioijalta selvästi omaa tulkintaa, sillä osassa raporteista tuli esille, ettei halutut vaikutukset olleet selvillä arvioitavassa organisaatiossakaan. Osaltaan tämä arvioijan tulkinnan tarve liittyy seuraavaksi käsiteltävään kohteen monimutkaisuuteen.

4.4 Kohteen kompleksisuus

Kolmas merkittävä haaste aineiston raporteissa vaikutti olevan arvioinnin kohteena olevan intervention kompleksisuus, ja siitä seuraava vertailtavuuden puute. Kohteen kompleksisuudella tarkoitetaan tässä yhteydessä esimerkiksi sitä, että hanke on suoritettu monen eri toimijan yhteistyössä. Täten arvioitaessa hankkeen toteutusta, on otettava huomioon myös mukana olleiden toimijoiden väliset erot sekä poikkeavat käytännöt. Toisaalta kompleksisuutta esiintyy myös arvioitaessa kohdetta, joka koostuu useasta eri osasta.

Ensinnäkin arvioinnin suorittaminen monimutkaistuu selvästi, jos sen kohdetta toteutetaan eri tavoin eri organisaatioissa. Esimerkiksi voidaan nostaa raportti 2, jossa arvioinnin kohteena oli kotoutumiskoulutus. Tässä tapauksessa arvioinnin haasteeksi nousi se, että eri toimijoiden järjestämien koulutusten välillä oli merkittäviä eroja niiden ajallisessa kestossa. Tämä yhdistettynä taloudellisen informaation puutteisiin vaikeuttaa selvästi eri toimijoiden välistä vertailtavuutta.

Myös mukana olevien organisaatioiden väliset tulkintaerot luovat eroja niiden välille. Tulkintaeroja oli havaittavissa niin käsitteiden määrittelyssä kuin toiminnan tavoitteissa. On helppo nähdä kuinka

eri toimijoiden järjestämät interventiot lähtevät helposti erkanemaan toisistaan, jos niiden tavoitteet eivät ole yhteneväiset. Aineiston perusteella ongelmana tuntui olevan se, että erilaisia ohjeita annettiin usealta tasolta, eivätkä ohjeet aina olleet yhdenmukaisia. Vedungin mallin viidennen haasteen mukaisesti arvioinnissa tulisikin pystyä määrittelemään sen kohteena olevan intervention tavoitteet. Jos tavoitteita ei ole suoraan löydettävissä kohteesta, jää niiden määrittely arvioijan oman tulkinnan varaan.

”Strategioiden laadinnassa olisi tärkeää huolehtia siitä, että eri osapuolet ymmärtävät käsitteet, vision ja tavoitteet mahdollisimman yhteneväisesti.” –Raportti 7

”Jotta määrällisiä tavoitteita voitaisiin asettaa, tulisi pystyä määrittelemään innovatiivisen hankinnan käsite.” –Raportti 8

Toisaalta kompleksisuutta aiheuttavat erot toimijoiden välisissä käytännöissä. Eri toimijat saattavat käyttää esimerkiksi erilaisia standardeja samankin asian toteuttamisessa. Eräessä tarkastuksessa (raportti 7) oltiin huomattu laaditun strategian jääneen etäiseksi käytännön työn kannalta. Tämän seurauksena eri toiminnot olivat muodostuneet henkilösidonnaisiksi, mikä osaltaan vaikeuttaa niiden vertailua eri organisaatioiden välillä.

Käytössä olevien erilaisten standardien lisäksi arvioinnissa yleiskuvan hahmottamista vaikeutti se, että organisaatioissa oli käytössä organisaation sisällä muodostuneita ”hyviä käytäntöjä”. Nämä käytännöt eivät varsinaisesti perustu mihinkään julkiseen ohjeeseen tai standardiin, joten ne voivat poiketa eri organisaatioiden välillä merkittävästi eikä yhteneväisyyksiä toiminnassa ole helppo nähdä. Tämän kaltaiset organisaatioiden sisäiset käytännöt pakottavat arvioitsijan tekemään valinnan mahdollisissa suosituksissaan. Arvioinnin kannalta toiminnan optimaalinen mitattavuus sekä vertailtavuus tarkoittaisivat mahdollisimman yhdenmukaisia käytäntöjä eri organisaatioiden välillä. Kuitenkin voi olla, että juurikin nämä organisaatioiden omat käytännöt ovat omiaan edistämään niiden tuloksellisuutta eivätkä ne välttämättä toimisi muissa organisaatioissa. Virtanen ja Vakkuri (2016, luku 6) ovat käsitelleet tätä ongelmaa organisaatioiden *tylsistymisenä*. Tieto tulevasta arvioinnista voi heikentää organisaatioissa tapahtuvaa innovointia, tai se voi keskittyä palvelemaan arvioinnin tarpeita tuloksellisuuden tarpeiden sijaan.

”Muissa vastauksissa lueteltiin myös hyviä käytäntöjä, jotka ovat muotoutuneet organisaatioiden omassa toiminnassa.” –Raportti 3

”Hyvien käytäntöjen kirjo on suuri, mikä vaikeuttaa niiden ja ohjeiden vertailua toisiinsa.” –Raportti 3

Aiemmin tässä luvussa kuvattiin miten organisaatioiden sekä interventioiden pirstaloituneisuus tunnistettiin aineistossa vastuuketjujen hahmottamista vaikeuttavaksi tekijäksi. Kuten tämän tutkimuksen teoreettista viitekehystä (arvioinnin funktiot) kuvattaessa todettiin, julkishallinnossa suurin osa toimijoista ovat sekä agentteja että päämiehiä. Jatkossa toteutettavissa tutkimuksissa voisi olla mielenkiintoista käsitellä tarkemmin, miten tämän kaltaiset monimutkaiset ja päällekkäiset vastuuroolit näyttäytyvät arvioinnin näkökulmasta. Jos päämiehet todella toimivat alaistensa arvioinnissa sen mukaan, miten heitä itseään tullaan jatkossa arvioimaan, kumoo tämä koko arvioinnin tarkoituksen toiminnan tuloksellisuuden parantamisena.

Kokonaisuudessaan arvioinnin kohteen kompleksisuus haasteena peilautuu hyvin Vedungin mallin kolmanteen sekä neljänteen haasteeseen. Olisiko tarkoituksenmukaisempaa pyrkiä rajaamaan arviointi tiettyyn osaan interventiosta, jotta vältetään tai ainakin minimoidaan yllä kuvatut ongelmat? Toisaalta kuten edellisessä alaluvussa huomattiin, arvioinnin kohteen rajaus toimintaympäristöstään ei luultavasti anna täydellistä kuvaa sen toiminnasta. On myös otettava huomioon, että intervention järjestämisessä tehtyjen valintojen taustalla olevien syiden merkitys korostuu sitä enemmän, mitä monitasoisemmin se on toteutettu. Ovatko hanketta toteuttamaan valitut tahot määrättyneet niiden hyödyllisyyden mukaan vai mahdollisesti poliittisen prosessin kautta?

4.5 Muut haasteet

Tähän mennessä analyysissä on tunnistettu kolme haastetta, jotka toistuvat läpi aineiston. Nämä haasteet ovat siis tietoon liittyvät puutteet, arvioinnin ajankohta, sekä arvioinnin kohteen kompleksisuus. Näiden lisäksi aineistosta löydettiin pienempiä haasteita, jotka eivät esiintyneet aineistossa johdonmukaisesti, mutta olivat silti merkityksellisiä tämän tutkimuksen kannalta.

Osasta raporteista nousi esille haasteeksi arvioinnin mittariston valitseminen. Ensinnäkin organisaatiossa saatettiin tuottaa näennäisesti valmiiksi sopivaa tietoa toiminnan mittariksi, mutta esimerkiksi haastatteluaineistossa kyseisen tiedon tarkoituksenmukaisuus kyseenalaistettiin. Tämä luo tarkastukseen ristiriidan sen suhteen, miten luotettavina valmiita mittareita voidaan pitää. Toisaalta mittariston suunnittelemista vaikeutti toiminnalle asetettujen tavoitteiden tulkinnanvaraisuus. Kuten VTV on raportissaan todennut, epäselvät tai tulkinnanvaraiset haasteet luovat ongelman paitsi toiminnan mittaamiselle myös tavoitteiden toteutumisen todentamiselle. Mittariston suunnittelun haasteissa on selvä yhteys Vedungin mallin seitsemänten haasteeseen. Kyseisessä haasteessa kyseenalaistetaan millä meriiteillä interventiota arvioidaan ja millä perustein sen voidaan sanoa olevan onnistunut.

”Liian tulkinnanvaraisia tavoitteita on vaikeaa tai mahdotonta mitata tai todentaa.” –Raportti 8

Raportista 1 annetuissa lausunnoissa oli nostettu esille, kuinka arviointiraportin julkaiseminen mahdollisesti vaikuttaisi intervention toteutustiimiin. Lausunnossa mainitaan tiimin pieni koko, jolloin raportin teksti saattaisi henkilöityä tiimin yksilöihin. Tällainen asetelma tuo kiinnostavan eettisen näkökulman arviointiin. Kuinka arvioijan tulisi ottaa huomioon arvioinnin tulosten julkaisemisen vaikutukset kohteeseen ja erityisesti sen toteuttaneisiin henkilöihin? Tässä kyseisessä tapauksessa VTV oli todennut, että koska tiimillä oli selvä vastuu ohjelman toteutuksesta ja eikä toimeenpanossa oltu onnistuttu, ei raportin julkaisematta jättämiselle nähty riittäviä syitä. Vaikkei tämä tapaus suoraan sovellu esimerkiksi kolmen E:n malliin kohdistettuun tarpeeseen eettisille näkökulmille, molemmissa tapauksissa on havaittavissa tietynlainen kysyntä ”pehmeämmille” arvoille kovan tuloksellisuustiedon lisäksi. On myös huomattava, että esimerkiksi Vedungin mallissa ei oteta ollenkaan huomioon arvioinnin eettisyyteen liittyviä haasteita.

5 Johtopäätökset

5.1 Pohdintaa

Tämän tutkimuksen tavoitteena oli kartoittaa, millaisia haasteita julkissektorin interventioiden arvioinnissa esiintyy, sekä miten ne näyttäytyvät käytännön arviointiraportoinnissa. Tutkimus onnistui luomaan suhteellisen hyvän yleiskuvan valitusta aineistosta nousevista haasteista.

Suuri osa julkissektorin toiminnan arviointiin liittyvistä haasteista tuntuvat olevan lähtöisin suhtautumisesta arviointiin kohdeorganisaatioissa. Arviointia ei nähdä olennaisena osana toimintaa sen suunnitteluvaiheessa, mistä seuraa arvioinnin kannalta hyödyllisen tiedon ”hukkumista” toiminnan yhteydessä. Lisäksi kun arviointia ei ole integroitu toimintaan sen suunnittelussa, ei tietynkaltaista informaatiota välttämättä osata kerätä lainkaan toiminnan kuluessa. Olipa toteutetun arvioinnin käyttötarkoituksena legitimizeetin vahvistus, toiminnan ohjaus tai yleisen tiedon lisääminen, olisi johdonmukaista integroida arviointi entistä tiiviimmin toimintaan jo sen suunnitteluvaiheessa.

Mietittäessä tuloksellisuusarvioinnin toteuttamista kolmen E:n mallin kautta, aineiston perusteella näyttäisi siltä, että toiminnan seurannasta saatava tieto usein painottuu tuloksellisuusketjun alkupäähän. Toisin sanoen tietoa saadaan esimerkiksi intervention taloudellisuudesta ja tehokkuudesta, mutta pitkän aikavälin vaikuttavuudesta sekä tavoitteista kertova informaatio jää

helposti piiloon. Tämä oli aiemman tutkimuskirjallisuuden pohjalta odotettavaa, sillä kuten tässäkin tutkimusraportissa on jo mainittu, toiminnan vaikuttavuus on vaikeimmin hahmotettava ja käsiteltävä tuloksellisuuden aspekteista. Vaikka vaikuttavuuden mittauksesta tuskin koskaan on haittaa arvioinnille, on sen tarkoituksenmukaisuutta mietittävä aina arvioinnin tavoitteisiin nähden. Kuten Vedung (2017, 96) toteaa, vaikuttavuuden tarkastelussa on tavoiteltavaa seurata kohdetta pitkällä aikavälillä. Tällaisen tarkastuksen toteuttaminen on kuitenkin hyvin resursseja vievää, joten arvioinnissa usein keskitytään vain tavoiteltuihin osa-alueisiin.

On myös tarkoituksenmukaista miettiä, miten ja milloin arviointi tulisi järjestää kohteeseen nähden. Aineiston perusteella arviointia voisi helpottaa, jos se toteutettaisiin ikään kuin väliarviointina, eli jo intervention vielä käynnissä ollessa. Tällä voitaisiin ainakin osittain estää tiedon hukkimista ja mahdollisesti vielä vaikuttaa toiminnasta kerättävään tietoon. Samalla väliarvioinnilla on vielä mahdollista myös paremmin vaikuttaa toiminnan lopputuloksiin. Kaikissa tapauksissa tämä ei tietenkään ole mahdollista, ja joissain tapauksissa myöhemmin tehtävä jälkiarviointi on parempi ratkaisu. Jälkiarviointia tehtäessä on myös mietittävä, miten intervention arviointi ajoitetaan sen suhteen, kuinka pitkän aikavälin vaikutuksista halutaan saada tietoa. Kokonaisuudessaan arvioinnin ajoitus lienee varsin tapauskohtainen tekijä, johon vaikuttaa vahvasti arvioinnin tarkoitettu hyödyntämistapa.

Tutkimuksen analyysissä löydettyistä arvioinnin haasteista on löydettävissä selvä yhteys tutkimuksen teoriapohjassa esiteltyyn Vedungin arvioinnin kahdeksan haasteen malliin. Lähes kaikki mallin haasteet olivat löydettävissä aineistosta jossain muotoa. On kuitenkin huomattava, ettei malli sovellu täysin valmiiksi annettuun arviointiympäristöön. Esimerkiksi tämän tutkimuksen aineiston tapauksessa arvioinnin suorittajana oli VTV, joten kysymystä siitä, kenen tulisi suorittaa arviointi voidaan pitää jossain määrin epäolennaisena. Lisäksi mallin soveltaminen aineistoon vaatii tutkijalta tietynlaista tulkintaa arvioinnin kontekstin suhteen, joten mallia ei voida pitää suoraan annettuna tarkastuslistana arvioijan käyttöön. Vaikka löydetty haasteet esiintyivät kattavasti läpi aineiston, raporttien välillä oli hyvin vaihtelevia painotuksia niiden välillä riippuen arvioinnin kohteista. Näin ollen voidaan todeta, että on vaikea luoda täydellistä valmista ”karttaa” arvioinnin haasteista, ja tämänkaltaiset olemassa olevat mallit tulisi ottaa aina tulkinnanvaraisina ohjeistuksina.

Vedungin arvioinnin kahdeksan haasteen mallissa ei oteta ollenkaan huomioon itse arvioinnin eettisyyttä. Tämän pohjalta on huomattavaa, ettei malli käsittele lainkaan ”pehmeisiin” arvoihin keskittyviä haasteita. Koska arvioinnin eettinen näkökulma ja sen sosiaaliset seuraukset nousivat pinnalle tämän tutkimuksen rajatussa aineistossa, on hyvin mahdollista, että ne ovat löydettävissä laajemminkin yksityisen ja julkisen sektorin toiminnasta tehdyssä arvioinnissa. Tähän tutkimukseen

valitun teoriapohjan kautta on mielenkiintoista pohtia myös arvioinnin sosiaalisten seurausten mahdollista linkkiä esimerkiksi päämies-agentti-malliin. Aineiston perusteella nämä arvioinnin seuraukset intervention toteuttajille on selvästi tiedostettu kohdeorganisaatioissa. Onko tällä mahdollisesti vaikutusta organisaation käyttäytymiseen ja suhtautumiseen arviointia kohtaan?

Tämän tutkimuksen asetelmassa ei päästy kiinni tuloksellisuustiedon manipuloinnin sekä tiedolla pelaamisen ongelmiin. Esimerkiksi Virtanen ja Vakkuri (2016) kuvaavat kuinka päätöksentekijät saattavat tietyissä tilanteissa altistua houkutukselle tulkita tietoa itselleen edullisella tavalla. Toisaalta omien päämäärien kannalta haitallisesta tiedosta voidaan vaieta tai jättää se kokonaan huomioimatta päätöksenteossa. Nämä tekijät liittyvät kuitenkin olennaisesti tässä tutkimuksessa löydettyihin tiedon puutteisiin. Jotta tuotetulla tuloksellisuustiedolla kyetään tuottamaan mahdollisimman kattavaa ja luotettavaa tietoa päätöksentekijöille, on pidettävä huoli, että myös tarkastus perustuu käyttökelpoiselle tietopohjalle.

5.2 Luotettavuus

Tieteellisen tutkimuksen yhtenä keskeisimmistä tavoitteista voidaan pitää mahdollisimman luotettavan tiedon tuottamista. Tutkimuksen luotettavuuden arviointia lähestytään usein tarkastelemalla sen reliabiliteettia sekä validiteettia (Saaranen-Kauppinen & Puusniekka 2006). Tutkimuksen reliabiliteetilla tarkoitetaan mittausten ja havaintojen pysyvyyttä eri olosuhteissa sekä eri aikoina. Lisäksi sillä kuvataan tulosten toistettavuutta. Validiteetilla taas tarkoitetaan sitä, ovatko tutkimuksen toteutus ja tulokset päteviä ja oikein toteutettuja. Saaranen-Kauppisen ja Puusniekan mukaan validiteetti on näistä käsitteistä selvästi johdonmukaisempi arvioitaessa nimenomaan laadullista tutkimusta.

Validiteetin näkökulmasta voidaan kysyä, kuinka hyvin tutkimuksen aineisto kuvaa julkishallinnon arvioinnin haasteita. Tähän liittyen on huomattava, että kaikki tutkimuksen aineistona käytettävistä raporteista ovat Valtion tarkastusviraston toteuttamia ja julkaisemia. Aineistoa kerätessä kuitenkin todettiin raporttien antavan riittävän kuvan yhteiskunnallisesta tuloksellisuusraportoinnista ottaen huomioon VTV:n toimenkuvan julkishallinnon toiminnan tarkastajana. Ilmiön kuvaamisen kattavuutta pyrittiin tutkimuksessa myös parantamaan ottamalla aineistoon mukaan raportteja mahdollisimman laajasti eri toimialoilta.

Analyysin osalta on huomattava, että osassa aineiston raporteissa mainitaan esitettyjen haasteiden tiedostaminen organisaatioissa sekä aiotut toimenpiteet haasteiden korjaamiseksi. Tämän tutkimuksen mittakaavassa ja aikajänteessä ei kuitenkaan voitu käsitellä näitä korjaamiseen tähtäviä hankkeita,

joten haasteet esitellään sellaisina kuin ne raporteista välittyvät. Toinen analyysiä koskeva tekijä koskee varsinaisia haasteita. Osa haasteista olivat sellaisia, että ne tuotiin raportissa esille organisaatiosta tarkastuksessa löydettyinä haasteina, mutta samalla ne olivat nähtävissä myös tarkastuksen toteuttamista hankaloittavina tekijöinä. Analyysiä tehdessä tämä tekijä pyrittiin minimoimaan peilaamalla jokaista haastetta kyseessä olevan raportin tavoitteisiin nähden.

Koska tutkimuksen aineistona käytettävät raportit ovat julkista tietoa, päätettiin analyysin yhteydessä merkitä sitaattien yhteyteen kunkin sitaatin lähteenä toimivat raportit. Tällä valinnalla pyrittiin varmistamaan tutkimuksen luotettavuus sekä läpinäkyvyys.

5.3 Jatkotutkimus

Tämä tutkimus on luonteeltaan hyvin yleistason kuvaus julkissektorin arvioinnin haasteista. Analyysin tulokset antavat suhteellisen hyvän kuvan aineistosta löytyneistä haasteista, mutta niiden merkitystä kohdeorganisaatioissa ei päästy tarkastelemaan tämän tutkimuksen rajauksissa. Tutkimuksen pohjalta tehdyssä tulevassa tutkimuksessa voidaan keskittyä esimerkiksi valikoidusti yksittäisiin haasteisiin, ja tarkastella miten ne vaikuttavat arvioinnin kohteessa. Samoin haasteita voidaan käsitellä rajattuna tietystä toimialasta tehdyn arvioinnin näkökulmasta.

Tutkimuksen kaltaista arvioinnin haasteiden kartoitusta voidaan toteuttaa myös eri aineistolla. Tämän tutkimuksen analyysissä löydetyt haasteet kuitenkin esiintyivät niin vahvasti läpi aineiston, että ne ovat todennäköisesti löydettävissä ainakin jossain muodossa aineistosta riippumatta. Lisäksi toisenlaisella tutkimusasetelmalla voidaan tarkastella, miten haasteet näyttäytyvät arviointiprosessin aikana. Onko tarkastuksen käynnissä ollessa eteen tulevilla haasteilla vaikutusta arvioinnin toteuttamistapaan sekä tavoitteisiin? Lisäksi voisi olla mielenkiintoista pohtia haasteita tarkemmin tuloksellisuusketjun kautta. Ohjaavatko haasteet tarkastusta ketjun alkupään tehokkuuteen ja taloudellisuuteen pitkän aikavälin vaikuttavuuden sijaan?

Viitteet

- Alasuutari, P. (2001). *Johdatus yhteiskuntatutkimukseen*. Helsinki: Gaudeamus.
- Arvidsson, G. (1986). *Performance Evaluation*. Teoksessa Kaufmann, F., Majone, G., & Ostrom, V. *Guidance, control, and evaluation in the public sector: The Bielefeld interdisciplinary project*. Berlin: de Gruyter, 625–43.
- Bringselius L. (2018). *Efficiency, economy and effectiveness - but what about ethics? Supreme audit institutions at a critical juncture*. *Public Money & Management*, 38(2), 105–110. doi:10.1080/09540962.2018.1407137.
- Carter, N., Day, P., & Klein, R. (1995). *How organisations measure success : The use of performance indicators in government*. London, New York: Routledge. doi:10.4324/9780203208663.
- Eisenhardt, K. M. (1989). *Agency theory: An assessment and review*. *The Academy of Management Review*, 14(1), 57–74. doi:10.2307/258191.
- Gildenhuis, C. E., & Janse van Rensburg, J. O. (2017). *The fourth E of performance auditing*. *Southern African Journal of Accountability & Auditing Research*, 19, 117–127. Retrieved from <http://search.ebscohost.com>.
- Hakkarainen, H. (2012). *Ulkoisen radikaalisädehoidon tuloksellisuus eturauhassyöpöpotilailla TAYS:Ssa*. Tampereen yliopisto. Pro gradu -tutkielma.
- Kivistö, J. (2009). *Tehottomat lintsarit? Valtion ja yliopistojen suhde agenttiteorian näkökulmasta*. Teoksessa J. Vakkuri (toim.) *Paras mahdollinen julkishallinto? Tehokkuuden monet tulkinnat*. Helsinki: Gaudeamus. 51–72.
- Kortelahti, O. (2019). *Digitalisaatio ja tuloksellisuus valtionhallinnossa*. Tampereen yliopisto. Pro gradu -tutkielma.
- Lumijärvi, I. (2009). *Johtamisen vaikutus organisaation tuloksellisuuteen: Mikä on johtamisen vaikutus organisaation korkean tuloksellisuuden synnyttämisessä ja miten vaikutus ilmenee?*. Tampere: Tampereen yliopistopaino.
- Meklin, P. (2009). *Muutuuko mikään? Tuloksellisuuden käsitteen monitulkintaisuus julkishallinnossa*. Teoksessa J. Vakkuri (toim.) *Paras mahdollinen julkishallinto? Tehokkuuden monet tulkinnat*. Helsinki: Gaudeamus. 31–50.
- Niiranen, V. (2008). *The Many Purposes of Performance Evaluation*. *International Journal of Public Administration*, 31:10–11, 1208–1222, doi: 10.1080/01900690801973303.
- Nikkilä, E. (2015). *Tilaaajaosaaminen ja tuloksellisuus kunnan sosiaali- ja terveydenhuollon palveluissa: Case Hämeenlinna–Tampere*. Tampereen yliopisto. Pro gradu -tutkielma.
- OECD (2015). *Building on Basics, Value for Money in Government*, Paris: OECD Publishing. <http://dx.doi.org/10.1787/9789264235052-en>.
- Oravakangas, A. (2005). *Koulun tuloksellisuus?: Filosofisia valituksia koulun tuloksellisuuden problematiikkaan suomalaisessa yhteiskunnassa*. Jyväskylän yliopisto. Väitöskirja.

- Orlikowski, W. J. (2002). *Knowing in practice: enacting a collective capability in distributed organizing*. *Organization Science*, Vol. 13 No. 3, 249–73.
- Petersen, T. (1993). *The Economics of Organization. The principal–Agent Relationship*. *Acta Sociologica* 36, 277–293.
- Saaranen-Kauppinen, A., & Puusniekka, A. (2006). *Kvalimotiv - menetelmä opetuksen tietovaranto*. (verkkojulkaisu) Saatavilla <https://www.fsd.uta.fi/menetelmaopetus/kvali/index.html>
- Salminen, M. (2005). *Tulosohjauksen käsikirja*. Helsinki: Valtiovarainministeriö.
- Scriven, M. (1991). *Evaluation thesaurus (4. ed.)*. Newbury Park: Sage.
- Smith P. (1995). *Outcome-related performance indicators and organisational control in the public sector*. Teoksessa Holloway J., Lewis J., Mallory G. *Performance Measurement and Evaluation*. London: Sage, 192–216.
- Talbot, C. (2005). *Performance Management*. Teoksessa Ferlie, E., Lynn, L. E., Jr & Pollitt, C. *The Oxford handbook of public management*. Oxford: Oxford University Press.
- Tuomi, J. & Sarajärvi, A. (2018). *Laadullinen tutkimus ja sisällönanalyysi (Uudistettu laitos)*. Helsinki: Tammi.
- Ulfves, A., Lääperi, R., Rautiainen, A. & Närhi, K. (2010). *Aikuissosiaalityöhön kiinteästi liittyvien aktivointi- ja työllistämispalvelujen kustannuksia ja vaikutuksia*. Koske.
- Vakkuri, J. (toim.) (2009). *Paras mahdollinen julkishallinto?: Tehokkuuden monet tulkinnat*. Helsinki: Gaudeamus.
- Vedung, E. (2017). *Public policy and program evaluation*. London: Routledge.
- Virtanen, P. & Vakkuri, J. (2016). *Julkisen toiminnan tuloksellisuusarviointi*. Helsinki: Tietosanoma.