

Sami Kollanus

Ohjelmistojen tarkastuskäytänteiden
puutteet ja ongelmat teoriassa
ja käytännössä



JYVÄSKYLÄ LICENTIATE THESES IN COMPUTING 4

Sami Kollanus

Ohjelmistojen tarkastuskäytänteiden puutteet
ja ongelmat teoriassa ja käytännössä



UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2006

Ohjelmistojen tarkastuskäytänteiden puutteet ja ongelmat teoriassa ja käytännössä

JYVÄSKYLÄ LICENTIATE THESES IN COMPUTING 4

Sami Kollanus

Ohjelmistojen tarkastuskäytänteiden puutteet
ja ongelmat teoriassa ja käytännössä



UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2006

Editor
Mauri Leppänen
Department of Computer Science and Information Systems, University of Jyväskylä

PDF ISBN 951-39-2398-3

ISBN 951-39-2436-X
ISSN 1795-9713

Copyright © 2006, by University of Jyväskylä

Jyväskylä University Printing House, Jyväskylä 2006

ABSTRACT

Kollanus, Sami

Weaknesses and problems in inspection practices in theory and practice

Jyväskylä: University of Jyväskylä, 2005, 106 p.

(Jyväskylä Licentiate Theses in Computing

ISSN 1795-9713; 4)

ISBN 951-39-2398-3

Software inspection is an already 30 years old method, which has been acknowledged as an important part of software engineering. However, we still know quite little about inspections in their real context. This thesis studies that area, especially weaknesses and problems in inspection practices. These issues are discussed both in theory and in practice with case studies from six software organizations. First, the weaknesses in the current inspection practices are studied by using the inspection capability maturity model (ICMM), which was constructed in the beginning of the research. Secondly, the study identifies known problems in inspection practices in the literature and the common problems in six case organizations. The major conclusion is that there is still much to improve in the current practices. In addition, even regularly used inspection practices may include problems that cause significant loss of inspection effectiveness.

Keywords: software inspection, peer review, practices, problems, process improvement

Tekijä Sami Kollanus
Tietojenkäsittelytieteiden laitos
PL 35 (Agora), 40014 Jyväskylän yliopisto,
sami.kollanus@jyu.fi

Ohjaajat Jussi Koskinen ja Markku Sakkinen
Tietojenkäsittelytieteiden laitos
Jyväskylän yliopisto

Tarkastajat Ilkka Tervonen
FT, professori
Oulun yliopisto

Jarmo Ahonen
FT, professori
Kuopion yliopisto

ESIPUHE

Tätä tutkimusta tehdessä on ollut mielenkiintoista havaita, kuinka ohjelmistoyritykset ovat olleet kiinnostuneita osallistumaan ja kehittämään omaa toimintaansa. Minulla ei ollut ennestään kontakteja tutkimukseen osallistuneisiin yrityksiin, mutta silti kaikki pyydetyt organisaatiot lähtivät innolla mukaan ja oma kokemukseni yhteistyöstä oli todella positiivinen. Haluankin tässä esittää kiitokseni erityisesti kaikille tutkimukseen osallistuneille organisaatioille ja niiden haastatteluihin osallistuneille työntekijöille. Haluan kiittää myös työn ohjaajia, Jussi Koskista ja Markku Sakkista, sekä monia muita, jotka ovat tukeneet minua matkan varrella.

Jyväskylässä 8.12.2005

Sami Kollanus

KUVIOT

| | | |
|----------|--|----|
| KUVIO 1: | Virheiden korjaamisesta aiheutuva suhteellinen kustannus projektin eri vaiheissa (Boehm 1981)..... | 12 |
| KUVIO 2: | Katselmointeja käsittelevien artikkelien määrä kymmenessä valitussa julkaisusarjoissa eri ajanjaksoina.. | 23 |
| KUVIO 3: | Eri teemojen osuus kaikista aineiston artikkeleista 5-vuotisjaksoina | 26 |
| KUVIO 4: | Tiivistelmä ICMM-mallin tasoista..... | 46 |

TAULUKOT

| | | |
|-------------|---|----|
| TAULUKKO 1: | Tutkimusongelmien ja tutkielman rakenteen suhde..... | 19 |
| TAULUKKO 2: | Katselmointeja käsittelevien artikkelien määrä eri julkaisuissa vuosina 1990-2004. | 23 |
| TAULUKKO 3: | Artikkelien luokittelu..... | 25 |
| TAULUKKO 4: | Käytänteiden toteutuminen kohdeorganisaatioissa..... | 60 |

SISÄLLYS

ABSTRACT

ESIPUHE

KUVIOT JA TAULUKOT

SISÄLLYS

| | | |
|-------|--|----|
| 1 | JOHDANTO..... | 11 |
| 1.1 | Tutkimuksen motivaatio..... | 13 |
| 1.1.1 | Tarkastukset on havaittu hyödyllisiksi | 13 |
| 1.1.2 | Tarkastuksia toteutetaan käytännössä heikosti | 14 |
| 1.1.3 | Aihe kiinnostaa tiedeyhteisöä edelleen..... | 14 |
| 1.1.4 | Tarkastusten soveltamiseen käytännössä tarvitaan lisää tutkimusta | 14 |
| 1.2 | Tutkimuskysymykset ja -menetelmät..... | 15 |
| 1.3 | Keskeiset käsitteet..... | 16 |
| 1.3.1 | Laatu käsitteenä | 16 |
| 1.3.2 | Katselmointi | 16 |
| 1.3.3 | Virhe | 17 |
| 1.3.4 | Dokumentti..... | 17 |
| 1.3.5 | Tarkastus..... | 18 |
| 1.3.6 | Tarkastusmenetelmän vaiheet..... | 18 |
| 1.4 | Tutkielman rakenne | 19 |
| 2 | KATSAUS KATSELMOINTEJA KÄSITTELEVÄÄN KIRJALLISUUTEEN | 20 |
| 2.1 | Kirjallisuusanalyysin toteutus | 21 |
| 2.2 | Kirjallisuusanalyysin tulosten kuvaus..... | 22 |
| 2.2.1 | Tutkimuksen määrän tarkastelu | 22 |
| 2.2.2 | Tutkimuksen luokittelu | 24 |
| 2.2.3 | Muita huomioita aineistosta | 27 |
| 2.3 | Tekninen näkökulma katselmointeihin..... | 28 |
| 2.3.1 | Lukutekniikat | 28 |
| 2.3.2 | Tehokkuustekijät katselmoinneissa | 31 |
| 2.3.3 | Prosessit..... | 35 |
| 2.3.4 | Sisältöön liittyvät erityiskysymykset..... | 36 |
| 2.4 | Johdon näkökulma | 36 |
| 2.4.1 | Katselmointien merkitys organisaatiolle..... | 37 |
| 2.4.2 | Muut johdon näkökulmaan liittyvät asiat | 37 |
| 2.5 | Muut teemat tutkimuksessa..... | 38 |
| 2.5.1 | Virheiden määrän arviointi..... | 38 |
| 2.5.2 | Työkalunäkökulma | 39 |
| 2.5.3 | Katselmointien kokonaisvaltainen tarkastelu | 40 |
| 2.6 | Yhteenveto | 41 |

| | | |
|--------|--|----|
| 3 | TARKASTUSKÄYTÄNTEIDEN KYPSYYSTASON MITTAAMINEN..... | 44 |
| 3.1 | Muita tarkastuskäytänteiden kypsyyksille..... | 44 |
| 3.2 | ICMM – Inspection Capability Maturity Model | 45 |
| 3.2.1 | Taso 2 – Suorittava..... | 46 |
| 3.2.2 | Taso 3 – Määritelty | 48 |
| 3.2.3 | Taso 4 – Hallittu | 50 |
| 3.2.4 | Taso 5 – Optimoiva..... | 52 |
| 3.3 | Yhteenveto kypsyyksien mittaamisesta | 53 |
| 4 | EMPIIRISEN TUTKIMUKSEN ESITTELY | 54 |
| 4.1 | Tutkimukseen osallistuneet organisaatiot | 54 |
| 4.2 | Tutkimuksen toteuttaminen..... | 55 |
| 4.3 | Huomioita valituista menetelmistä..... | 56 |
| 5 | KUINKA KATSELMOINNIT TOTEUTUVAT KÄYTÄNNÖSSÄ..... | 58 |
| 5.1 | Muita tutkimuksia | 58 |
| 5.2 | Tarkastusten toteutuminen kohdeorganisaatioissa | 60 |
| 5.2.1 | Yleiskuva tuloksista..... | 61 |
| 5.2.2 | Vaatimusten tarkastus | 62 |
| 5.2.3 | Suunnitelmien tarkastus..... | 62 |
| 5.2.4 | Tarkastusten johtajien koulutus | 62 |
| 5.2.5 | Testitapausten tarkastus | 63 |
| 5.2.6 | Koodin tarkastus..... | 63 |
| 5.2.7 | Määritelty tarkastusprosessi | 63 |
| 5.2.8 | Koulutus kaikille..... | 64 |
| 5.2.9 | Mukautettu tukimateriaali | 64 |
| 5.2.10 | Tiedon keräys ja käyttö | 64 |
| 5.2.11 | Organisaation politiikka | 65 |
| 5.2.12 | Määritellyt vastuutehtävät | 65 |
| 5.2.13 | Allokoidut resurssit..... | 65 |
| 5.3 | Tutkimustulosten analysointia | 66 |
| 5.4 | Yhteenveto | 67 |
| 6 | TARKASTUSTEN ONGELMAT..... | 69 |
| 6.1 | Tarkastusten ongelmat kirjallisuudessa | 70 |
| 6.1.1 | Palaverin aikataulutus voi aiheuttaa viivettä..... | 70 |
| 6.1.2 | Palaveri kuluttaa resursseja, mutta siellä ei löydetä paljon virheitä..... | 71 |
| 6.1.3 | Huono tukimateriaali..... | 72 |
| 6.1.4 | Osallistujat eivät ymmärrä tarkastusprosessia..... | 72 |
| 6.1.5 | Kritiikki kohdistetaan tekijään | 73 |
| 6.1.6 | Huono tarkastusten suunnittelu..... | 73 |
| 6.1.7 | Palaveri ajautuu epäoleellisiin asioihin..... | 73 |
| 6.1.8 | Huono valmistautuminen | 74 |
| 6.1.9 | Tarkastuksiin osallistuu vääriä ihmisiä..... | 74 |
| 6.1.10 | Roolit saattavat puuttua | 75 |

| | | |
|--------|--|-----|
| 6.1.11 | Huonolaatuinen dokumentti vie turhaan aikaa..... | 75 |
| 6.1.12 | Huono asenne..... | 75 |
| 6.1.13 | Liian paljon materiaalia | 76 |
| 6.1.14 | Tarkastuksiin ei ole varattu resursseja | 76 |
| 6.2 | Kokemukset yrityksissä | 76 |
| 6.2.1 | Palaverin aikataulutus voi aiheuttaa viivettä..... | 78 |
| 6.2.2 | Palaveri kuluttaa resursseja, mutta siellä ei löydetä paljon virheitä..... | 78 |
| 6.2.3 | Huono tukimateriaali..... | 78 |
| 6.2.4 | Osallistujat eivät ymmärrä tarkastusprosessia..... | 79 |
| 6.2.5 | Kritiikki kohdistetaan tekijään | 79 |
| 6.2.6 | Huono tarkastusten suunnittelu..... | 79 |
| 6.2.7 | Palaveri ajautuu epäoleellisiin asioihin..... | 80 |
| 6.2.8 | Huono valmistautuminen | 80 |
| 6.2.9 | Tarkastuksiin osallistuu väärä ihmisiä..... | 80 |
| 6.2.10 | Roolit saattavat puuttua | 81 |
| 6.2.11 | Huonolaatuinen dokumentti vie turhaan aikaa..... | 81 |
| 6.2.12 | Huono asenne..... | 81 |
| 6.2.13 | Liian paljon materiaalia | 82 |
| 6.2.14 | Tarkastuksiin ei ole varattu resursseja | 82 |
| 6.3 | Yhteenveto tarkastusten ongelmista..... | 83 |
| 7 | JOHTOPÄÄTÖKSET | 85 |
| | SUMMARY | 88 |
| | LÄHTEET | 90 |
| | LIITE 1 – TÄYDELLINEN LISTA KIRJALLISUUSANALYYSIIN SISÄLTÄYNEISTÄ ARTIKKELEISTA | 98 |
| | LIITE 2 – HAASTATTELURUNKO..... | 104 |

1 JOHDANTO

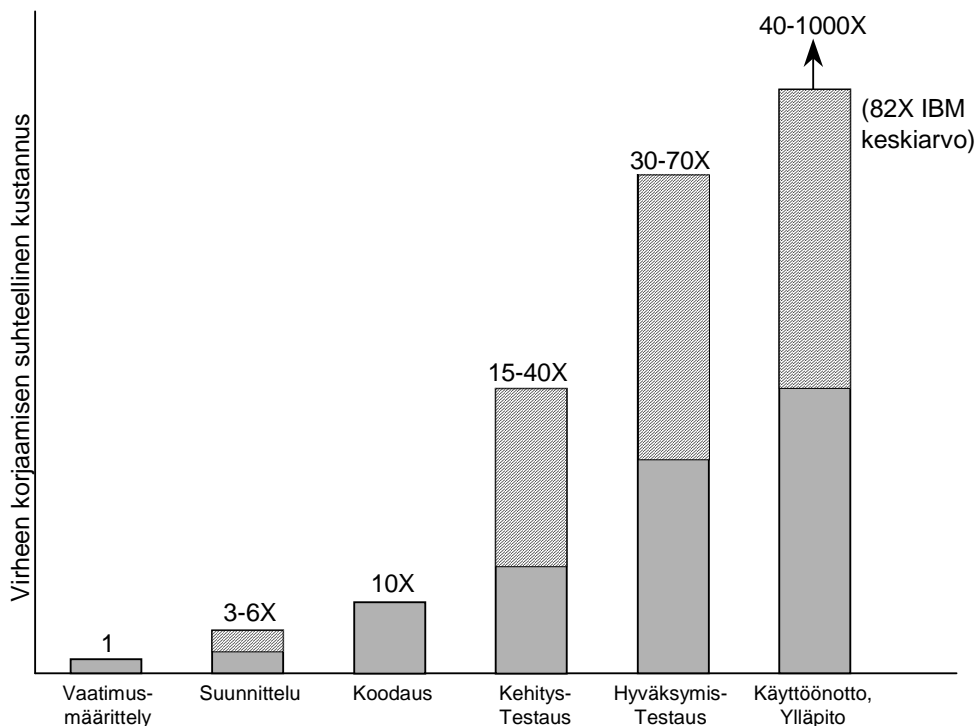
Michael Fagan (1976) julkaisi alkuperäisen IBM:n organisaatiossa kehitetyn *ohjelmistojen tarkastusmenetelmän* (*software inspection*) jo lähes 30 vuotta sitten. Hän raportoi silloin tarkastusten käyttöönoton parantaneen sekä ohjelmistojen laatua että samalla koko ohjelmistotuotantoprosessin tehokkuutta. Sen jälkeen tarkastus on laajasti tunnustettu merkittäväksi menetelmäksi ohjelmistotekniikan alalla. Esimerkiksi McConnell (2000) on arvostanut tarkastuksen merkittävimpien ohjelmistotekniikkaan vaikuttaneiden menetelmien joukkoon. Tarkastus on myös vakiinnuttanut paikkansa alan oppikirjoissa (esim. Sommerville 2001).

Ohjelmistojen laatuongelmat ovat varmasti tuttuja kaikille tavallisillekin tietokoneen käyttäjille. Nykypäivänä koko länsimainen yhteiskunta toimii erilaisten ohjelmistojen varassa. Niin liikenne, raha-asioiden hoito kuin monet kodinkoneetkin ovat riippuvaisia ohjelmistoista. Virheet kriittisimmissä ohjelmistoissa voivat johtaa jopa ihmishenkien menetykseen (esim. Neumann 2005). Kurinalaisilla tarkastuksilla voidaan Freedmanin ja Weinbergin (1990) mukaan saada karsittua virheiden määrä kymmenenteen osaan alkuperäisestä määrästä jo ennen testausvaihetta.

Usein testaus ja tarkastus asetetaan vastakkain, kun mietitään tarkastusten hyödyllisyyttä. Monilla ohjelmistoalan työntekijöillä vaikuttaa olevan pinttynyt käsitys siitä, että testaus on nopeampi ja tehokkaampi tapa hoitaa ohjelmiston laadunvarmistus. Joissakin tutkimuksissa on kuitenkin saatu jopa päinvastaisia tuloksia ja todettu koodin lukemisen olevan testausta tehokkaampi tapa löytää virheitä (Basili ja Selby 1987). Useat tutkimukset ovat todenneet koodin lukemisen ja testauksen löytävän tehokkaasti erityyppisiä virheitä (Roper ym. 1997). Lisäksi tarkastusten avulla voidaan valvoa toiminnallisten virheiden lisäksi myös muita laatutekijöitä, kuten koodin ylläpidettävyyttä ja dokumentaation laatua. Tutkimusten mukaan jopa yli 90 % ohjelmiston elinkaaren kustannuksista saattaa mennä ylläpitoon, joten ylläpidettävyyden huomioiminen ei ole aivan vähäinen asia ohjelmistojen kehittämistyössä (Erlikh 2000).

Tarkastusten ajatellaan usein vaikuttavan ainoastaan tuotoksen laatuun ja sen vuoksi ne saatetaan nähdä vain ylimääräisinä kustannuksina organisaatiolle. Jo Fagan (1976) alkuperäisessä tarkastusmenetelmän esittelevässä artikkelissaan

totesi, että tarkastusten käyttöönotto tehosti pilottiprojekteissa koko tuotanto-prosessia ja samalla säästi kustannuksia. Sen jälkeen useissa lähteissä on raportoitu tarkastusten avulla saavutetuista säästöistä. Esimerkiksi Gilb ja Graham (1993) esittelevät kirjassaan useita tällaisia tapauksia. Kustannussäästö perustuu siihen, että virhe on halvinta korjata mahdollisimman aikaisessa vaiheessa. Klassinen Boehmin (1981) tutkimus (kuvio 1) totesi esimerkiksi, että vaatimusmäärittelyssä tapahtuneen virheen korjaaminen toteutusvaiheessa tuli keskimäärin kymmenen kertaa kalliimmaksi verrattuna sen korjaamiseen määrittelyvaiheessa.



KUVIO 1 Virheiden korjaamisesta aiheutuva suhteellinen kustannus projektin eri vaiheissa (Boehm 1981)

Tarkastusten merkitys tuotannon tehokkuudelle on siis kaikkein keskeisin prosessin alkupäässä. Kuitenkin tarkastusten on useissa tapauksissa laskettu maksavan itsensä takaisin prosessin vaiheesta riippumatta. Esimerkiksi Bush (1990) laskee, että tarkastuksesta läpipäässeän virheen korjaaminen myöhemmässä vaiheessa tuotantoprosessissa maksoi JPL:n organisaatiossa keskimäärin 1700 dollaria. Jokaisessa tarkastuksessa löytyi keskimäärin 12 virhettä. Bushin mukaan arvioitu keskimääräisen tarkastuksen tuottama säästö kustannukset mukaan luettuna oli noin 25 000 dollaria.

Tarkastukset on siis todettu tärkeäksi ja taloudellisestikin kannattavaksi menetelmäksi ohjelmistotuotannossa. Kuitenkin aiempien tutkimusten perusteella tarkastusten käyttö ohjelmistoteollisuudessa vaikuttaa vielä puutteelliselta (Johnson 1998; Ciolkowski ym. 2003). Tämän tutkimuksen tarkoitus on kes-

kittyä tähän ongelma-kohtaan ja pohtia käytännössä tarkastusten kehittämistä ohjelmistoalan organisaatioissa.

Seuraavassa alakohdassa 1.1 esitellään tarkemmin tutkimusalueen merkittävyyttä ja lähtökohtia tämän tutkimuksen tekemiseen. Kohdassa 1.2 esitetään konkreettiset tutkimuskysymykset ja tutkimuksessa käytetyt menetelmät. Kohdassa 1.3 käydään lyhyesti läpi tämän tutkielman rakenne ja eri osien yhteys toisiinsa.

1.1 Tutkimuksen motivaatio

Tämä tutkimus käsittelee tarkastusten kehittämistä ohjelmistoalan organisaatioissa. Tässä kohdassa keskitytään perustelemaan tämän valitun tutkimusalueen merkittävyyttä sekä tiedeyhteisölle että käytännön ohjelmistotuotannolle. Seuraavat alakohdat käyvät läpi keskeisimmät tutkimusalueen merkittävyyttä tukevat argumentit.

1.1.1 Tarkastukset on havaittu hyödyllisiksi

Tarkastukset on havaittu tärkeäksi menetelmäksi ohjelmistotuotannossa sekä teoriassa, että käytännössä. Jo edellä johdannon aluksi esiteltiin hiukan tarkastusten merkitystä tuotannon laadulle, tehokkuudelle ja kustannuksille. Seuraavassa on listattu joitakin tutkimustuloksia keskeisimmistä näkökulmista tarkastuksen merkittävyyteen:

- Doolan (1992) laski, että jokainen tarkastus maksaa itsensä takaisin 30-kertaisesti.
- Grady ja Van Slack (1994) arvioivat HP:n säästäneen tarkastusten ansiosta 21,5 miljoonaa dollaria vuonna 1993.
- Russellin (1991) tutkimuksessa jokainen tarkastuksiin käytetty tunti säästi 33 tuntia ylläpitokustannuksia.
- Basili ja Selby (1987) totesivat koodin lukemisen olevan tehokkaampi tapa löytää jopa samoja virheitä kuin testaus.

Aiempaa kirjallisuutta tarkasteltaessa on mielenkiintoista, että sellaisia tutkimuksia ei juuri löydy, jotka kyseenalaistaisivat vakavasti tarkastusten merkityksen ohjelmistotuotannossa. Kirjallisuudessa esitetty kritiikki kohdistuu tavallisesti tarkastusten toteutustapaan. Esimerkiksi Votta (1993) aloitti artikkelillaan tiedeyhteisössä pitkään eläneen keskustelun perinteisen palaverin merkityksestä tarkastusprosessissa. Sen sijaan tarkastusten tarpeellisuutta yleensä kyseenalaistavia tutkimustuloksia ei kirjallisuudessa juuri löydy.

1.1.2 Tarkastuksia toteutetaan käytännössä heikosti

Tarkastusten merkitys tunnetaan hyvin yleisesti ohjelmistoalalla, mutta siitä huolimatta tarkastusten toteutus on käytännössä hyvin puutteellista. Tämä on tosin kirjallisuuden perusteella ennemminkin yleisesti tunnustettu tosiasia kuin systemaattisesti tutkittu ilmiö. Kuitenkin joitakin tätä tukevia viitteitä tutkimuksessa on.

Johnsonin (1998) raporttoimassa kyselyssä 90 ohjelmistoalan ammattilaisesta 80 % ilmoitti käyttävänsä työssään tarkastuksia epäsäännöllisesti tai ei ollenkaan. Ciolkowskin ym. (2003) toteuttamassa laajemmin katselmointien toteutusta käsittelevässä kyselyssä todettiin, että 42 % vastaajista katselmoi säännöllisesti vaatimusmäärittelydokumenteja ja 28 % ohjelmakoodia.

Olemassa olevien tutkimustulosten perusteella voidaan siis sanoa että useimmilla yrityksillä ensimmäinen kynnyks on saada tarkastukset toimimaan säännöllisesti. Toinen kysymys on sitten tutkia sitä, kuinka hyvin säännöllisesti toteutetut tarkastukset toimivat käytännössä. Voidaan tehdä ainakin sellainen hypoteettinen oletus, että välttämättä kaikki tarkastuksia harjoittavat organisaatiot eivät toteuta niitä mahdollisimman tehokkaasti. Tässä oletuksessa on oikeastaan lähtökohta koko tälle tutkimukselle, joka tarkastelee tarkastusten toteuttamisen ongelmakohtia.

1.1.3 Aihe kiinnostaa tiedeyhteisöä edelleen

Käytännön yrityselämän näkökulman lisäksi on tutkimuksen kannalta merkittävää, että myös kansainvälinen tiedeyhteisö on kiinnostunut aihealueesta. Jotkut aihealueet saattavat olla tiedeyhteisössä ohimeneviä ilmiöitä ja menettää tietyn ajan jälkeen kiinnostavuutensa. Tarkastusten kohdalla tiedeyhteisön kiinnostus vaikuttaa kuitenkin osoittavan aivan päinvastaista kehityssuuntaa, jos kiinnostusta mitataan aihetta käsittelevien julkaisujen määrällä. Luvussa 2 esitetään yksityiskohtaisesti viimeiset 15 vuotta kattavan kirjallisuusanalyysin tulokset. Sen mukaan alan keskeisimmissä julkaisuissa on viimeisen viiden vuoden aikana julkaistu enemmän katselmoiteja käsitteleviä artikkeleita kuin koko 1990-luvulla yhteensä.

1.1.4 Tarkastusten soveltamiseen käytännössä tarvitaan lisää tutkimusta

Luvussa 2 esitettävän kirjallisuusanalyysin perusteella voidaan sanoa, että tarkastusten käytännön soveltamisesta ja siihen liittyvistä ongelmista on aiemmin tehty todella vähän tutkimusta. Valtaosa tarkastuksia käsittelevästä tutkimuksesta perustuu erilaisiin kontrolloituihin kokeisiin. Niiden perusteella tunnetaan jo kohtuullisesti monia tarkastuksen tehokkuuteen vaikuttavia tekijöitä teoriasa. Kirjallisuudessa on myös jonkin verran tapaustutkimuksia tarkastuksia onnistuneesti käyttävistä suuryrityksistä. Nämä raportit ovat usein aika pintapuolisia, eivät usein raportoi kohdatuista käytännön ongelmista ja antavat luonnollisesti tarkastuksiin vain suuryrityksen näkökulman.

Tarkastukset on siis käytännössä havaittu hyväksi menetelmäksi, mutta kuitenkin useissa organisaatioissa niitä ei käytetä laajasti. Lisäksi säännöllisestikin tarkastuksia toteuttavat organisaatiot saattavat toteuttaa niitä tehottomasti. Tässä on selvästi alue, johon kaivataan lisää ymmärrystä, jotta selvästi hyväksi havaittu menetelmä saadaan laajasti tehokkaaseen käyttöön ohjelmistoalan yrityksissä.

1.2 Tutkimuskysymykset ja -menetelmät

Tämä tutkimus keskittyy tarkastelemaan tarkastusten toteuttamista ohjelmisto- ja tuottavassa organisaatiossa. Tutkimuksen pääkysymys on:

Mitkä ovat keskeisimmät ongelmat ja puutteet tarkastusten toteuttamisessa?

Varsinaista tutkimusongelmaa lähestytään sekä teoriassa että käytännössä. Tutkimus on jaettu pienempiin osakokonaisuuksiin, jotka vastaavat seuraaviin kysymyksiin:

- Mitä aiempien tutkimusten perusteella tiedetään tarkastuksista?
- Mitä käytänteitä liittyy tarkastusten toteuttamiseen ja kuinka ne voidaan priorisoida kypsyystasomallin muotoon?
- Kuinka edellä määritellyt käytänteet toteutuvat käytännössä ohjelmistoja tuottavissa organisaatioissa?
- Millaisia käytännöllisiä ongelmia saattaa liittyä tarkastusten toteuttamiseen aiempien tutkimusten perusteella?
- Millaisia ongelmia esiintyy käytännössä ohjelmia tuottavissa organisaatioissa?

Motivaatio tutkimuksen taustalla on tarkastuskäytänteiden kehittämisen pohdiskelu. Tämä tausta näkyy tutkimuksessa siten, että näkökulma ongelmien tarkastelussa tähtää käytänteiden kehittämiseen, eikä tyydy ainoastaan kuvailemaan tämän hetkistä tilannetta. Varsinaiset tutkimustulokset tässä työssä tosin rajautuvat suurimmaksi osaksi kuvailemaan tämän hetkistä asioiden tilaa.

Tutkimukseen sisältyy tapaustutkimus, joka on toteutettu kuudessa suomalaisessa ohjelmistoja tuottavassa organisaatiossa. Organisaatiossa pyrittiin selvittämään haastatteluilla, miten katselmointeja niissä toteutetaan ja millaisia ongelmia katselmointeihin liittyy. Haastattelujen tuloksia on analysoitu pääasiassa kvalitatiivisesti pyrkien ymmärtämään mahdollisimman hyvin erilaisissa tilanteissa olevia organisaatioita. Keskeisin tutkimuksen tulos on hahmottaa kirjallisuuden ja tapaustutkimuksen avulla katselmointien toteutukseen liittyvää ongelmakenttää.

1.3 Keskeiset käsitteet

Tässä kohdassa pyritään määrittelemään ja rajaamaan tämän tutkimuksen kannalta keskeiset käsitteet. Katselmoiteihin liittyvän käsitteistön käyttö ei ole englanninkielisessäkään kirjallisuudessa aivan yhdenmukaista ja samoille asioille saattaa olla useita eri termejä. Suomenkielessä taas ei vakiintuneita käsitteitä ole juuri lainkaan, sillä alan kirjallisuus on lähes täysin englanninkielistä.

1.3.1 Laatu käsitteenä

Tarkastuksista puhuttaessa ollaan olennaisesti tekemisissä laadun käsitteen kanssa. Katselmoinnin pääasiallinen tehtävä organisaatiossa on yleensä laadunvarmistus tai laadun mittaaminen. Weinberg (1992) tiivistää laadun yksinkertaisesti: "Laatu on arvoa (*value*) jollekin henkilölle". Ohjelmistotuotannossa näkökulma laatuun ja laadunvarmistukseen voi tarkoittaa eri osapuolille hyvin eri asioita. Johdon näkökulmasta tehokas laadun tuottaminen voi ensisijaisesti tarkoittaa säästöä virheiden aiheuttamissa kustannuksissa tai samasta syystä tehostunutta tuotantoprosessia. Työntekijän näkökulmasta laatu voi olla suurimmaksi osaksi tuotosten laatua. Asiakkaalle se voi puolestaan merkitä laajempaa kokonaisuutta, kuten toiminnallisesti virheetöntä lopputuotetta, hyvää käytettävyyttä, laadukasta dokumentaatiota ja toimivaa palvelua.

Tässä tutkimuksessa ei ole tärkeää yrittää rajata laadun käsitettä tiettyyn näkökulmaan. Oleellista on sen sijaan ymmärtää, että laatu voi merkitä ohjelmistotuotannossa eri asioita. Se, mitä hyvällä laadulla tarkoitetaan, täytyy määrittellä tapauskohtaisesti riippuen siitä, mistä näkökulmasta asiaa halutaan tarkastella. Tarkastuksilla on joka tapauksessa keskeinen rooli laadunvarmistuksessa kaikista laadunäkökulmista katsottuna.

1.3.2 Katselmointi

Englanninkielessä *review* on kohtuullisen vakiintunut yleiskäsite, jonka IEEE:n standardi määrittelee seuraavasti (IEEE 1988):

Ohjelmiston osien tai projektin tilan arviointi (evaluation), jonka tarkoitus on tunnistaa tuotosten eroavuudet suunnitelmiin nähden sekä tuottaa kehitysehdotuksia. Tämä arviointi noudattaa formaalia prosessia (esimerkiksi johdon katselmointiprosessia, ohjelmiston tarkastusprosessia tai läpikäyntiprosessia).

Suomenkielessä tästä on käytetty sekä termiä *katselmointi* että *katselmus*. Tässä työssä edellä mainitun määritelmän vastineena käytetään termiä *katselmointi*. Katselmointi on siis laaja yleiskäsite, jonka alle kuuluu erilaisia katselmointimenetelmiä. Katselmoiteina voidaan pitää hyvin erilaisia tilaisuuksia, joita järjestetään monta erilaista tarkoitusta varten. Sisäisen laadunvarmistuksen lisäksi katselmoiteina voidaan pitää ulkoisia katselmoiteja asiakkaan kanssa, jotka ovat usein luonteeltaan enemmänkin hyväksymistilaisuuksia. Lisäksi laadun

varmistusta tai mittausta varten voidaan järjestää katselmointeja erillisen laatu-yksikön toimesta. Myös yrityksen johtoa varten järjestetyt lähinnä projektin tilaa arvioivat katselmoinnit ovat tavallisia.

Tässä tutkimuksessa keskitytään tarkastelemaan katselmointia ainoastaan laadunvarmistuksen näkökulmasta ja siksi käsitettä käytetään käytännössä hieman IEEE:n määritelmää suppeammassa merkityksessä. Kun tässä työssä käytetään sanaa katselmointi, halutaan korostaa sitä, että tarkoitetaan laajasti eri katselmointimenetelmiä, joiden pääasiallinen tarkoitus on etsiä virheitä jostain dokumentista. Lisäksi tässä tutkimuksessa näkökulma katselmoiteihin keskittyy prosessin sisäisiin katselmoiteihin ohjelmistotuotannossa. Katselmoinnit nähdään osana normaalia ohjelmistotuotantoprosessia, jossa katselmoinnit toteutetaan vertaiskatselmoiteina. Siis katselmoijina ovat pääosin eri rooleissa toimivat työtoverit, eivätkä esimerkiksi esimiehet tai ulkopuolinen laatuyksikkö.

1.3.3 Virhe

Englannin kielessä käytetään yleisesti sanoja *error*, *bug*, *fault*, *mistake* ja *defect* synonyymeina. Myös niiden määrittely tutkimuksissa vaihtelee tutkijan tai koulukunnan mukaan. Näiden sanojen suhteita toisiinsa on yrittänyt hahmottaa esimerkiksi United Kingdom Software Metrics Association (UKSMA 2002). Suomen kielessä vastaavia sanoja ovat virhe, vika, erhe, puutteellisuus ja epäkohta.

Katselmointien yhteydessä käytetään lähes poikkeuksetta englanninkielen sanaa *defect*. Suomenkielessä sen vastineena puolestaan käytetään vaihtelevasti sanoja vika, virhe, puutteellisuus ja epäkohta. Tässä tutkimuksessa käytetään englannin sanan *defect* suomenkielisenä vastineena yksinkertaisesti sanaa *virhe*. Tässä merkityksessään virhe voidaan ymmärtää tapauksesta riippuen poikkeamana ohjelmiston spesifikaatiosta tai käyttäjävaatimuksista.

Usein tarkastusten yhteydessä halutaan korostaa sitä, että prosessin aikana dokumentista esitetyt kommentit eivät ole virheitä ennen kuin ne todetaan todellisiksi virheiksi. Ennen tätä hyväksymistä englanninkielessä näistä kommenteista käytetään yleisesti sanoja *finding* tai *issue*. Tässä työssä näiden vastineena käytetään sanaa *löydös*.

1.3.4 Dokumentti

Usein sana *dokumentti* yhdistetään ainoastaan tekstimuotoisiin dokumentteihin tai korkeintaan suunnitteluvaiheen kaavioihin. Tässä työssä sanaa dokumentti käytetään hyvin laajasti tarkoittaen mitä tahansa ohjelmistotuotantoprosessissa syntyvää tuotosta. Tässä merkityksessään dokumentti voi siis viitata vaikkapa ohjelmakoodiin.

1.3.5 Tarkastus

Edellä katselmointikäsitteen määrittelyssä todettiin tarkastuksen olevan eräs katselmointimenetelmä. Alkuperäinen ohjelmistojen tarkastusmenetelmä kehitettiin IBM:n organisaatiossa 1970-luvun alussa (Fagan 1976). Kirjallisuudessa tästä alkuperäisestä menetelmästä puhutaan usein Faganin tarkastuksena. Toinen keskeinen lähde tarkastusmenetelmästä puhuttaessa on Gilbin ja Grahamin (1993) kirjoittama kirja. Siinä kirjoittajat ovat kehittäneet menetelmää jonkin verran, mutta prosessi on pääpiirteittäin samanlainen kuin Faganin alkuperäisessä menetelmässä. Usein kirjallisuudessa käsiteltäessä jotain tarkastusmenetelmän sovellettua toteutustapaa puhutaan *perinteisestä tarkastuksesta (traditional inspection)*. Silloin sillä tarkoitetaan Faganin (1976) ja Gilb ja Grahamin (1993) mukaista prosessia. Myös tässä työssä noudatetaan samaa periaatetta.

Kirjallisuudessa tarkastus käsitettä käytetään hiukan vaihtelevasti. Toiset tutkijat tarkoittavat tarkastuksella edellä mainittua perinteistä tarkastusta ja haluavat nimittää siitä hiukankin poikkeavia käytänteitä katselmoinneiksi. Tavallinen käytäntö kirjallisuudessa on kuitenkin käyttää termiä laajemmassa merkityksessä kattamaan laajasti kaikki alkuperäisestä menetelmästä tehdyt muunnokset. Toisaalta jotkut katselmointikäsitettä mieluummin käyttävät tutkijat määrittelevät katselmointiprosessin aivan yhdenmukaiseksi perinteisen tarkastuksen kanssa. Tässä tutkimuksessa tarkastusta käytetään laajana käsitteenä kattaen monenlaiset alkuperäisestä menetelmästä tehdyt sovellukset.

1.3.6 Tarkastusmenetelmän vaiheet

Englanninkielisessä kirjallisuudessa käytetään tarkastukseen liittyvistä vaiheista erilaisia nimiä taustalla olevan ”koulukunnan” mukaan. Nämä koulukunnat viittaavat tavallisesti joko Faganin (1976) alkuperäiseen tarkastusmenetelmään tai Gilb ja Grahamin (1993) vastaavaan esitykseen. Lisäksi esimerkiksi Laitenberger ja DeBaud (2000) käyttävät hiukan erilaisia termejä määrittäessään tarkastusprosessin kuvausta. Seuraavassa on listattu tarkastusprosessin vaiheet Faganin määrittelyn mukaisesti. Samalla on esitetty tutkimuksessa vaiheista käytettävät suomenkieliset termit.

- Aloituspalaveri (Overview)
- Valmistautuminen (Preparation)
- Tarkastuspalaveri (Inspection)
- Virheiden korjaus (Rework)
- Seuranta (Follow-up)

1.4 Tutkielman rakenne

Kohdassa 1.2 esitettiin tutkimusongelmat, joihin yleisesti tässä tutkimuksessa ja sen osakokonaisuuksissa pyritään vastaamaan. Taulukossa 1 esitetään, kuinka mainittujen tutkimusongelmien käsittely sijoittuu tutkielman eri osiin.

TAULUKKO 1 Tutkimusongelmien ja tutkielman rakenteen suhde

| | |
|--|--------|
| Mitä aiempien tutkimusten perusteella tiedetään tarkastuksista? | Luku 2 |
| Mitä käytänteitä liittyy tarkastusten toteuttamiseen ja kuinka ne voidaan priorisoida kypsyystasomallin muotoon? | Luku 3 |
| Kuinka edellä määritellyt käytänteet toteutuvat käytännössä ohjelmistoja tuottavissa organisaatioissa | Luku 5 |
| Millaisia ongelmia saattaa liittyä tarkastusten toteuttamiseen aiempien tutkimusten perusteella? | Luku 6 |
| Millaisia ongelmia esiintyy käytännössä ohjelmia tuottavissa organisaatioissa? | Luku 6 |

Luvussa 2 tehdään kattava katsaus katselmointeja käsittelevään kirjallisuuteen. Siinä tarkastellaan systemaattisesti katselmointeja käsittelevää tutkimusta viimeisen 15 aikana kymmenessä keskeisessä ohjelmistotekniikan alan julkaisussa. Luvussa luokitellaan tutkimus alakategorioihin ja esitellään tutkimusalan keskeisimmät tutkimustulokset sekä suuntausten kehittyminen tarkasteltuna ajanjaksona.

Luvussa 3 käsitellään tarkastusten kypsyystason mittaamista. Luvun keskeinen asia on hahmottaa tarkastuksiin liittyvät käytänteet ja priorisoida ne kypsyystasomallin muotoon. Pääosa luvusta keskittyy esittelemään alun perin artikkelissa Kollanus (2005a) julkaistua ICMM-mallia, joka on tarkoitettu välineeksi tarkastuskäytänteiden kypsyystason arviointiin ja kehittämiseen. ICMM on eräänlainen lähtökohta tälle tutkimukselle, sillä se kehitettiin tutkimuksen alkuvaiheessa empiirisen tutkimuksen apuvälineeksi.

Luvut 4,5 ja 6 esittelevät tutkimuksen empiiristä osuutta. Neljännessä luvussa kuvataan johdantona seuraaville luvuille, kuinka empiirinen tutkimus toteutettiin. Luvut 5 ja 6 käsittelevät molemmat tarkastuskäytänteisiin liittyviä puutteita ja ongelmia hiukan eri näkökulmasta. Luvussa 5 tarkastellaan, kuinka kohdeorganisaatioissa toteutetaan tarkastuksia ja millaisia puutteita käytänteissä havaittiin. Luvussa 6 puolestaan tarkastellaan artikkeliin Kollanus (2005b) pohjautuen tarkastuskäytänteisiin liittyviä ongelmia sekä aiemmassa kirjallisuudessa, että käytännössä tutkimuksen kohdeorganisaatioissa. Lopuksi luvussa 7 tehdään yhteenveto tutkimuksen keskeisistä johtopäätöksistä.

2 KATSAUS KATSELMOINTEJA KÄSITTELEVÄÄN KIRJALLISUUTEEN

Tässä luvussa luodaan systemaattinen katsaus katselmoinneista tehtyyn tutkimukseen. Luvun tarkoituksena on jäsentää alan aiempaa tutkimusta ja tarkastella siinä vaikuttaneita suuntauksia. Samalla tässä luvussa esitellään alan keskeisiä tutkimustuloksia.

Aiemman tutkimuksen tarkastelu perustuu yhdeksään keskeiseen ohjelmistotekniikan alalla toimivaa lehteen sekä alan merkittävimpään vuosittaiseen konferenssiin (International Conference on Software Engineering). Näistä lähteistä tarkastellaan systemaattisesti aikavälillä 1990-2004 julkaistuja katselmointeja koskevia artikkeleita.

Tämän tutkimuksen yleislinjasta poiketen tässä luvussa tarkastellaan yleisesti katselmointimenetelmiä käsittelevää tutkimusta, koska monessa tapauksessa luokittelu tarkastuksia (*inspection*) ja katselmointia (*review* tai *peer review*) yleisesti käsitteleviin artikkeleihin olisi ollut hankalaa. Eri tutkijat, jotka kuvaavat katselmointimenetelmän suunnilleen samalla tavalla, käyttävät eri käsitteitä. Suurin osa tutkijoista käyttää käsitettä tarkastus, mutta jotkut haluavat mieluummin puhua samasta asiasta yleisemmin katselmointeina. Sen vuoksi on siis selkeämpää käsitellä tässä luvussa yleisesti katselmointeja.

Seuraavassa kohdassa 2.1 esitellään, kuinka tässä luvussa toteutettu kirjallisuuden analysointi on toteutettu. Kohdassa 2.2 kuvaillaan aineistosta löytyneet yleiset trendit ja jäsenetään aiempi tutkimus tarkastelun kannalta mielekkäisiin luokkiin. Seuraavat kohdat 2.3-2.5 esittelevät tutkimuksen eri osa-alueita ja keskeisimpiä tutkimustuloksia edellä esitetyn luokituksen mukaisesti. Näissä kohdissa esitetyt tutkimustulokset perustuvat pääosin kirjallisuusanalyysin aineistoon, mutta täydennykseksi mukaan on otettu joitakin keskeisiä muita lähteitä, jotka selventävät tutkimuksen kokonaiskuva. Lopuksi kohdassa 2.6 tehdään yhteenveto tämän luvun keskeisimmistä asioista ja pohditaan samalla hieman tutkimuksen tulevaisuutta.

2.1 Kirjallisuusanalyysin toteutus

Kirjallisuuskatsauksen tarkastelun aluetta täytyi rajata mielekkäästi sekä ajan että tarkasteltavien julkaisujen suhteen. Aikaperspektiivissä tarkastelun kohteeksi oli saatava riittävän pitkä tarkasteluväli tutkimuksessa esiintyvien trendien tunnistamiseksi. Alustavan lähteiden tutkimisen jälkeen vaikutti siltä, että ennen 1990-lukua tarkastusmenetelmästä tai katselmoinneista yleensä on kirjoitettu hyvin vähän julkaisuja. Tästä syystä vuosi 1990 vaikutti hyvältä lähtökohdalta tutkimuksen aloittamiselle. Kirjallisuutta tarkasteltiin siis 15 vuoden aikavälillä vuosilta 1990-2004.

Suurin osa systemaattisen tarkastelun kohteeksi valituista julkaisusarjoista löytyi lähdeviitteiden perusteella. Ensimmäisenä käytiin läpi ne julkaisut, joissa julkaistuihin katselmoiteja käsitteleviin artikkeleihin on viitattu usein alan kirjallisuudessa. Lähdeviitteiden alustavan tarkastelun jälkeen vaikutti siltä, että katselmoineista oli kirjoitettu pääasiassa yleisesti ohjelmistotekniikan alalla toimivissa julkaisusarjoissa. Tämän johtopäätöksen tuloksena tarkasteltavaan aineistoon liitettiin vielä joitakin merkittäviä ohjelmistotekniikan alan lehtiä, jotka eivät olleet nousseet esille alustavassa tarkastelussa. Lehtien lisäksi tarkasteluun otettiin mukaan ohjelmistotekniikan alan johtavassa konferenssissa (International Conference on Software Engineering) julkaistut artikkelit.

Edellä kuvatun prosessin tuloksena tässä luvussa tarkasteltaviksi lähteiksi valikoituivat:

- IEEE Transactions on Software Engineering
- Journal of systems and software
- Empirical Software Engineering
- IEEE Software
- Information and Software Technology
- Journal of Software Testing, Verification and Reliability
- Communications of ACM
- ACM Transactions on Software Engineering and Methodology
- IEE Software Engineering Journal
- International Conference on Software Engineering

Valituista lähteistä käytiin aluksi sisällysluettelon ja abstraktien perusteella läpi kaikki vuosina 1990-2004 julkaistut artikkelit. Tämän alustavan tarkastelun perusteella valittiin kaikki katselmoiteihin liittyvät artikkelit tarkempaa käsitelyä varten. Seuraavassa vaiheessa alustavan listan sisältämät artikkelit käytiin tarkemmin läpi ja niistä kirjattiin muistiin joitakin tunnuspiirteitä, kuten tutkimusongelma ja käytetty tutkimusmenetelmä. Tässä vaiheessa tehtiin lopullinen rajaus aineistoon mukaan otettavista artikkeleista ja alustavaa lista karsiutui sen seurauksena jonkin verran. Seuraavassa on listattu joitakin esimerkkejä siitä, minkälaisin perustein katselmoiteihin läheisesti liittyviä artikkeleita rajattiin lopullisen aineiston ulkopuolelle.

- Shull ym. (2000) käsittelevät artikkelissaan samoja lukemistekniikoita, joita on käytetty useissa tarkastuksia käsittelevissä tutkimuksissa. Heidän artikkelissaan käsitellään kuitenkin yleisesti ohjelman ymmärtämisestä (*program comprehension*) erilaisten lukutekniikoiden avulla. Tämä on selkeästi katselmoiteja tukevaa aineistoa, mutta jos tämä laskettaisiin mukaan, samalla täytyisi tähän sisällyttää kaikki ohjelman ymmärtämistä käsittelevä kirjallisuus.
- Chillarege ym. (1992) käsittelevät virheiden luokittelua. Tämä on myös selkeästi katselmoiteja tukevaa aineistoa, mutta kyseisen tutkimuksen näkökulma ei sijoitu erityisesti katselmointikontekstiin.
- Card (1990) hahmottelee laadun parantamista kokonaisuutena. Katselmoiteja pidetään artikkelissa vain yhtenä hyvin pienenä osana kokonaisuutta. Tällaiset laadun hallintaa kokonaisuutena käsittelevät jutut jätettiin lopullisen aineiston rajauksen ulkopuolelle.
- Vuonna 2003 lehdissä IEEE Transactions on Software Engineering (TSE) ja IEEE Software julkaistiin samaan aikaan ilmestyneet erikoisnumerot, joissa esitettiin samat tutkimukset eri kohderyhmille suunnattuna. TSE lehden artikkelit oli suunnattu tiedeyhteisölle ja IEEE Software lehden artikkelit ohjelmistoalan ammattilaisille. Pällekkäisyyden vuoksi aineistoon otettiin mukaan ainoastaan TSE lehden versiot.

Tarkemman tarkastelun jälkeen kaikki lopulliseen aineistoon valikoituneet artikkelit jaettiin luokkiin pääasiallisen tutkimusongelman mukaisesti. Luokittelussa ei noudatettu etukäteen päätettyä jäsenystä, vaan se tehtiin aineistolähtöisesti. Luokittelua kuvataan tarkemmin seuraavassa kohdassa.

2.2 Kirjallisuusanalyysin tulosten kuvaus

Tässä kohdassa esitellään kirjallisuusanalyysin pohjalta löydettyjä keskeisiä trendejä muutamasta eri näkökulmasta. Aluksi tarkastellaan katselmoiteja käsittelevän aineiston määrää eri julkaisuissa ja eri aikajaksoina. Seuraavaksi luokitellaan tutkimus eri teemoihin ja esitellään tutkimuksen sisällöllisten trendien kehitystä eri ajanjaksoina. Lopuksi käsitellään vielä joitakin muita näkökulmia, jotka kuvaavat katselmointien tutkimusta.

2.2.1 Tutkimuksen määrän tarkastelu

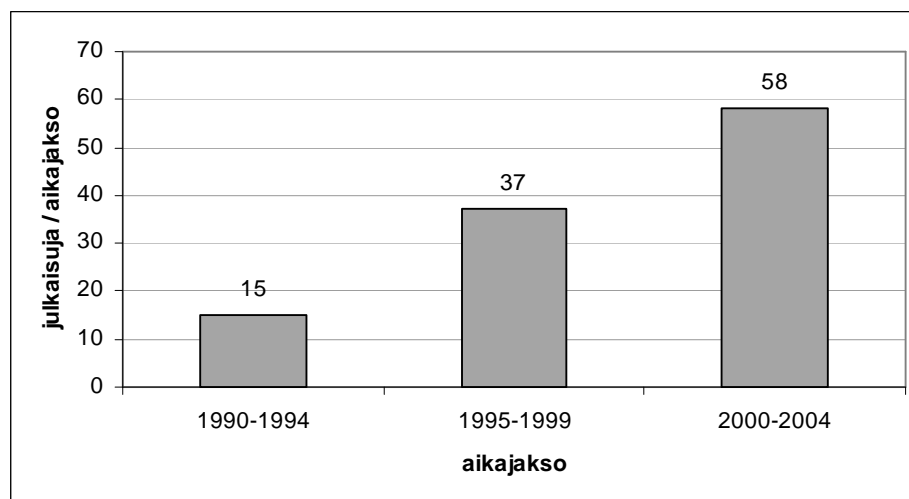
Tarkastellaan aluksi aineiston artikkeleita määrällisesti. Taulukko 2 esittää kustakin julkaisusta löytyneiden katselmoiteja käsittelevien artikkeleiden kokonaismäärä sekä vuodet, jotka kyseisen julkaisun tarkastelu kattaa. Julkaisut on järjestetty niistä löytyneiden katselmoiteja käsittelevien artikkeleiden määrän mukaan. Valintakriteerit täyttäviä artikkeleita löytyi kaiken kaikkiaan 110 kap-

palette. Taulukosta voidaan nähdä, että hyvin suuri osa niistä keskittyy muutamain keskeiseen julkaisuun.

TAULUKKO 2 Katselmointeja käsittelevien artikkelien määrä eri julkaisuissa vuosina 1990-2004.

| Julkaisu | Vuodet | Määrä |
|---|------------------|------------|
| International Conference on Software Engineering | 1990-2004 | 22 |
| IEEE Transactions on Software Engineering | 1990-2004 | 18 |
| Journal of Systems and Software | 1990-2004 | 17 |
| Empirical Software Engineering | 1997-2004 | 16 |
| IEEE Software | 1990-2004 | 14 |
| Information and Software Technology | 1990-2004 | 11 |
| Journal of Software Testing, Verification and Reliability | 1996-2004 | 6 |
| Communications of ACM | 1990-2004 | 4 |
| ACM Transactions on Software Engineering and Methodology | 1992-2004 | 2 |
| IEE Software Engineering Journal | 1990-1996 | 0 |
| Artikkeleita yhteensä | 1990-2004 | 110 |

Aineistoa tarkasteltaessa aikaperspektiivillä voidaan havaita, että katselmointeja käsittelevän tutkimuksen määrä on jatkuvasti kasvanut. Ennen vuotta 1990 aiheesta on vain harvoja satunnaisia julkaisuja ja 1990-luvun alkupuolellakin julkaisujen määrä oli vielä vähäinen. Aiheen tutkimus on oikeastaan todella lähtenyt liikkeelle 1990-luvun puolenvälin paikkeilla ja on siitä asti jatkanut kasvua viime vuosiin asti. Tämä voidaan todeta kuvioista 2, joka esittää, miten aineiston 110 artikkelia jakaantuvat ajallisesti viisivuotiskausina kuvattuna.



KUVIO 2 Katselmointeja käsittelevien artikkelien määrä kymmenessä valitussa julkaisusarjoissa eri ajanjaksoina

Osa julkaisuista ei ole ilmestynyt koko aikaa tarkastellulla aikavälillä 1990-2004. Jonkin verran tutkimuksen määrän kehittymisestä voi selittyä julkaisujen mää-

rän hienoisella kasvamisella. Varsinkin Empirical Software Engineering oli aineistossa eräs merkittävistä lähteistä ja se on ilmestynyt vasta vuodesta 1996 alkaen. Kokonaisuudessa tällä ei kuitenkaan voida selittää kovin paljon. Monien keskeisten näkökulmien tutkiminen vaikuttaa joka tapauksessa alkaneen vasta 1990-luvun puolessa välissä tai sen jälkeen.

2.2.2 Tutkimuksen luokittelu

Tässä yhteydessä tutkituissa ainoastaan Laitenberger ja DeBaud (2000) ovat yrittäneet aiemmin jäsentää katselmointeja käsittelevää kirjallisuutta kokonaisuutena. Heidän esittämänsä jäsenitys on useassa kohdassa edelleen toimiva, mutta siinä on käytännössä kaksi keskeistä puutetta. Heidän aineistonsa on kerätty ainakin pääosin ennen vuotta 1998. Siihen mennessä oli julkaistu ainoastaan noin 30 prosenttia tässä luvussa esitellystä aineistosta. Tutkimuksen painopiste on sen jälkeen muuttunut jonkin verran. Toiseksi Laitenberger ja DeBaud luokittelivat tutkimuksessaan samoja artikkeleita useampaan eri luokkaan ja tämä toimintatapa näkyy luokituksessa. Aineiston määrällisen tarkastelun kannalta on kuvaavampaa, jos yksi artikkeli luokitellaan ainoastaan yhteen luokkaan, johon se ensisijaisesti kuuluu. Laitenbergerin ja DeBaudin luokittelu ei aivan täysin toimi tässä tarkoituksessa.

Edellä mainituista syistä johtuen oli mielekäästä jäsentää katselmointeja käsittelevä tutkimus uudella tavalla, joka kuvaa mahdollisimman hyvin tässä luvussa käsiteltyä aineistoa. Artikkelit jaettiin yksiulotteisesti tutkimusongelmien mukaisesti eri teemoihin. Yksi artikkeli kuuluu siis ainoastaan yhteen luokkaan ensisijaisen tutkimusongelman mukaisesti luokiteltuna.

Aineistosta erottui selvästi yhdeksän erillistä teemaa. Eri teemat aineistossa olivat melko selvät, sillä luokittelun ulkopuolelle jäi lopulta vain viisi artikkelia. Taulukko 3 kuvaa eri teemoja ja niihin sijoittuneiden artikkelien määrää. Osa teemoista on siinä jäsenetty vielä korkeamman tason luokkiin. Näissä ylemmän tason luokissa on soveltuvilta osin käytetty Laitenbergerin ja DeBaudin (2000) jo ennestään alalla tunnettua jäsenitystä. Heidän luokittelumallistaan on otettu teknisen näkökulman ja johdon näkökulman käsitteet sekä se, että työkalunäkökulmaa pidetään erillisenä omana luokkana. Kahta heidän malliinsa kuuluvaa yläluokkaa (organisaationäkökulma ja arviointinäkökulma) ei käytetty, sillä tässä toteutetulla luokittelutavalla niihin ei saatu esiintymiä. Lisäksi Laitenbergerin ja DeBaudin tutkimuksen jälkeen virheiden määrän arvioinnista on tullut oma merkittävä tutkimusalueensa, joka muodostaa tässä oman luokkansa.

Taulukosta 3 voidaan nähdä, että tekninen näkökulma katselmointeihin on selvästi keskeisin tutkimuskohde, sillä siihen sijoittui puolet kaikista artikkeleista. Tekninen näkökulma pitää sisällään erilaisten katselmointiprosessien kuvaukset, erilaiset dokumenttien lukutekniikat, tarkastusten tehokkuuteen vaikuttavat prosessin sisäiset tekijät ja sisältöön liittyvät erityiskysymykset. Näistä eri lukutekniikat ja tehokkuustekijät katselmoinneissa olivat lisäksi kaksi suurinta yksittäistä teemaa tutkitussa aineistossa.

TAULUKKO 3 Artikkelien luokittelu

| | | |
|--|---|-------------------|
| Tekninen näkökulma | Lukutekniikat 21 (19 %) Tehokkuustekijät katselmoineissa 21 (19 %) Prosessit 8 (7 %) Sisältöön liittyvät erityiskysymykset 5 (5 %) | 55 (50 %) |
| Johdon näkökulma | Katselmointien merkitys organisaatiolle 7 (6 %) Muut aiheet johdon näkökulmasta 8 (7 %) | 15 (14 %) |
| Työkalu-näkökulma | | 13 (12 %) |
| Virheiden määrän arviointi | | 18 (16 %) |
| Kokonaisvaltainen katsaus alaan | | 4 (4 %) |
| Luokittelemattomat | | 5 (5 %) |
| ARTIKKELIT YHTEENSÄ | | 110 (100%) |

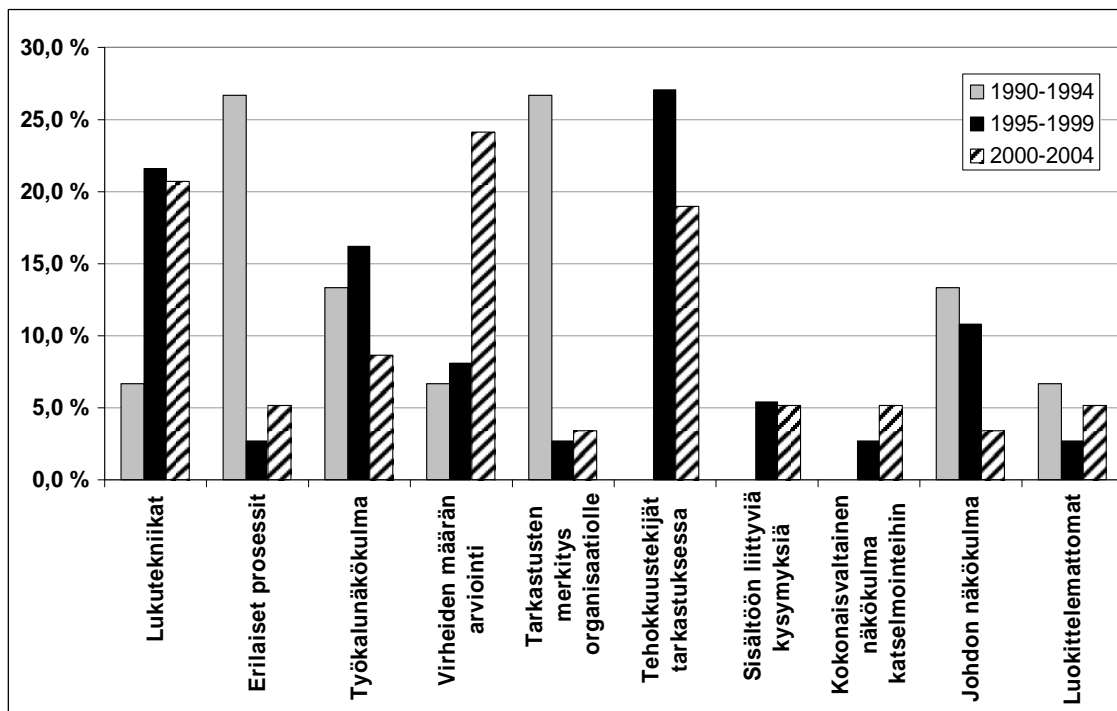
Tässä esitellystä luokittelusta on hyvä huomioida, että se ei kerro aivan koko totuutta tarkastellusta aineistosta. Useissa tekniseen näkökulmaan keskittyvissä tutkimuksissa tuotetaan samalla johdon kannalta kiinnostavaa tietoa, mutta tässä ne on kuitenkin ensisijaisen tutkimusongelman mukaisesti sijoitettu tekniseen näkökulmaan kuuluviksi. Tutkimus pitää siis kokonaisuutena sisällään enemmän johdon näkökulmaan liittyvää tietoa kuin tuosta 14 prosentin osuudesta voisi päätellä. Tällaisten osuuksien laskeminen vaatisi kuitenkin huomattavasti monimutkaisempaa ja vaikeammin hahmotettavaa luokittelutapaa.

Teemojen jakaminen juuri kuvatuunlaiseksi ei ole aivan itsestään selvää, vaan luokittelussa noudatettiin tiettyjä periaatteita mahdollisimman tasapainoisen luokittelun aikaansaamiseksi. Esimerkiksi lukutekniikat olisi hyvin voitu pitää loogisesti tehokkuustekijöihin kuuluvana tutkimuksena. Sitä koskevaa tutkimusta on kuitenkin tehty yhtä paljon kuin muita tehokkuustekijöitä koskevaa tutkimusta yhteensä ja siitä syystä se oli mielekästä erottaa omaksi teemakseen. Seuraavassa listassa on kuvattu keskeisimmät luokittelussa noudatetut periaatteet.

- Jotkut artikkelit käsittelivät useampia eri luokkiin kuuluvia teemoja. Tällaisessa tapauksessa luotettiin siihen, että artikkelin otsikko kuvaa tärkeintä tutkimusongelmaa ja artikkeli luokiteltiin sen mukaisesti.
- Joissakin työkaluja koskevissa artikkeleissa pääpaino on enemmän prosessissa kuin ohjelman teknisessä puolessa. Kaikki työkaluja käsittelevät artikkelit sijoitettiin kuitenkin hieman edellisestä kohdasta poiketen työkalukategoriaan
- Lukutekniikoita käsittelevät artikkelit luokiteltiin omaan teemaan, vaikka ne voitaisiin loogisesti sijoittaa myös tehokkuustekijöihin.

Artikkelien luokittelua on mielekästä tarkastella myös aikaperspektiivillä. Kuvio 3 esittää eri tutkimusteemojen osuutta kaikesta katselmoineista käsittelevästä tutkimuksesta eri viiden vuoden ajanjaksoina. Harmaa palkki jokaisessa teemassa kuvaa aikaväliä 1990-1994, musta palkki aikaväliä 1995-1999 ja raidalli-

nen palkki aikaväliä 2000-2004. Kuvasta voidaan nähdä, mitkä ovat olleet katselmointeja käsittelevän tutkimuksen keskeisiä teemoja eri ajanjaksoina.



KUVIO 3 Eri teemojen osuus kaikista aineiston artikkeleista 5-vuotisjaksoina

Eri aikajaksoina tapahtuneet muutokset kuvaavat osin tutkimusalan luonnollista kehitystä. Tutkimuksen alkuvaihe (1990-1994) keskittyi suurelta osin erilaisiin katselmointiprosesseihin ja siihen, kuinka katselmoinnit vaikuttavat organisaation toimintaan. Myös muista näkökulmista ensimmäiset ajatukset nousivat esiin tämän ajanjakson aikana, mutta varsinaisesti aktiivinen tutkimus painottuu myöhemmälle ajalle.

Seuraavan viisivuotiskauden (1995-1999) aikana katselmointeja käsittelevä tutkimus sai monelta osin nykyisen muotonsa. Tutkimuksesta katoaa lähes täysin erilaisten katselmointiprosessien tarkastelu ja se alkoi perehtyä enemmän katselmointiprosessin sisäisiin asioihin. Tässä vaiheessa ryhdyttiin aktiivisesti etsimään erilaisia prosessin sisäisiä tekijöitä, joita kehittämällä voidaan parantaa katselmointien tehokkuutta. Erilaiset virheiden etsimiseen käytettävät lukutekniikat ja muut tehokkuustekijät nousivat hallitseviksi alueiksi katselmointien tutkimisessa.

Viimeisen tarkastellun viisivuotiskauden (2000-2004) aikana keskeinen muutos edelliseen on virheiden määrän arviointiin keskittyvän tutkimuksen kasvu. Tästä näkee selvästi sen, kuinka joidenkin tutkijoiden mielenkiinto voi ohjata merkittävästi tutkimuksen trendejä alalla. Virheiden määrän arvioinnissa ei ole vielä kukaan päästy käytännön kannalta kovin merkittäviin tuloksiin eikä sitä voida muutenkaan käytännön soveltamisen kannalta pitää kovin merkittävänä osan katselmointien tutkimusta. Tästä huolimatta se on herättänyt tutki-

joissa sellaista mielenkiintoa, että se on ollut 2000-luvulla katselmointien tutkimuksen aktiivisin teema.

2.2.3 Muita huomioita aineistosta

Kirjallisuusanalyysin aineistosta on huomioitava, että se ei ole kovin suuri suhteutettuna tarkasteltuun aikaan. Tärkeimmistä suuntauksista voidaan saada edellisessä alakohdassa esitellyn luokittelun perusteella selvä kuva, mutta pienemmissä luokissa artikkeleiden määrä on aika pieni.

Eräs selvä huomio on, että tutkimuksen sisältöön vaikuttaa selvästi joidenkin aktiivisten tutkijoiden henkilökohtainen mielenkiinto. Joissakin tapauksissa (esim. virheiden määrä) puolet tiettyyn teemaan luokitelluista artikkeleista on parin aktiivisen tutkimusryhmän tuotosta. Jo tästä syystä tutkimuksen sisällöllinen jakauma saattaa näyttää erilaiselta tutkijapopulaation vaihtuessa.

Edellisessä alakohdassa esitelty luokittelu oli tehty ainoastaan yksiulotteisesti pääasiallisen tutkimusongelman mukaisesti. Aiemmin puhuttiin jo siitä, että tällaisen luokittelun seurauksena esimerkiksi johdon näkökulman huomiomisesta alan tutkimuksessa ei välttämättä saada todenmukaista kuvaa. Tämän korjaamiseksi luokittelua voisi kehittää siten, että jokaisesta artikkelista pyritään arvioimaan, missä määrin eri teemat ovat siinä edustettuna. Sillä tavalla voitaisiin päästä tarkempaan kuvaukseen, mutta toisaalta se olisi huomattavasti monimutkaisempi tapa toteuttaa ja lisäksi siinä painottuisi entistä enemmän luokittelijan oma tulkinta.

Tässä esitetyssä luokittelussa artikkeleita on luokiteltu ainoastaan tutkimusongelman mukaisesti, mutta tutkimusta voidaan tarkastella myös muiden luokitteluperusteiden mukaan. Tässä analysoiduista artikkeleista esimerkiksi 75 prosenttia raportoi jonkinlaisen empiirisen tutkimuksen tuloksia. Empiirisestä tutkimuksesta 59 prosenttia on toteutettu kontrolloiduissa kokeissa ja vastaavasti 41 prosenttia pohjautuu todellisen työelämän kokemuksiin. Kontrolloiduista kokeista yli 80 prosenttia on toteutettu opiskelijoille ja lopuissa tutkimuksissa koehenkilöinä on toiminut alan ammattilaisia. Näistä tunnuslukuista voidaan päätellä, että empiirisellä tutkimuksella on alalla vahva asema. Tyypiesimerkki alan tutkimuksesta on kontrolloitu koe, joka on toteutettu ohjelmistoalan yliopisto-opiskelijoille.

Tässä luvussa on säännönmukaisesti käytetty käsitettä katselmointi (*review*), koska tutkittavan aineiston rajaaminen oli selkeämpää, kun mukaan otettiin kaikki katselmointeja käsittelevät artikkelit. Tutkimus painottuu kuitenkin hyvin selvästi tarkastusmenetelmän (*inspection*) tutkimiseen. Artikkeleista yli 80 prosenttia käyttää käsitettä tarkastus. Lisäksi suurin osa lopuistakin artikkeleista määrittelee katselmointiprosessin käytännössä tarkastusmenetelmän kaltaiseksi. Ainoastaan aivan muutama artikkeli käsittelee selvästi tarkastuksesta poikkeavaa katselmointia.

2.3 Tekninen näkökulma katselmoiteihin

Tekniseen näkökulmaan luokiteltiin sellaiset tutkimukset, jotka jollain tavalla vastaavat kysymykseen ”Miten katselmoiteja tulisi toteuttaa?”. Tästä poikkeuksena on työkalunäkökulma, joka ottaa osittain kantaa myös samankaltaiseen kysymyksenasetteluun, mutta sitä päätettiin käsitellä tässä luokittelussa erillisenä luokkana. Tekninen näkökulma on ollut alan tutkimuksessa kaikkein laajin niin, että siihen sijoittui luokittelussa puolet kaikesta alan tutkimuksesta. Seuraavissa alakohdissa käsitellään keskeisimpiä tekniseen näkökulmaan liittyviä tutkimustuloksia.

2.3.1 Lukutekniikat

Lukutekniikat olivat erillisistä teemoista aktiivisin tutkimusalue yhdessä erilaisten tehokkuustekijöiden kanssa. Sitä käsitelti 19 % koko aineiston tutkimuksista. Lukutekniikoilla tarkoitetaan tässä erilaisia tapoja etsiä virheitä katselmoitavista dokumenteista. Näillä tekniikoilla viitataan poikkeuksesta yksilökeskeiseen näkökulmaan. Taustalla on siis yhteinen oletus siitä, että virheiden etsiminen tapahtuu itsenäisesti valmistautumisvaiheessa ja mahdollisesti järjestettävä palaveri painottuu löydösten kirjaamiseen. Tutkijoiden keskuudessa ei ole syntynyt mielenkiintoa kehittää virheiden etsimiseen menetelmiä, jossa toimijana olisi yksilön sijaan ryhmä. Toisaalta kuitenkin monet menetelmät pitävät sisällään ajatuksen työnjaosta, vaikka eivät pidäkään virheiden etsimistä ryhmäprosessina.

Alkuperäisiin ajatuksiin tarkastusmenetelmästä kuuluu olennaisena osana *tarkistuslistojen* käyttö virheiden etsinnässä (Fagan 1976). Tarkistuslistojen idea on koota kokemustieto listaksi tavallisimpia dokumentissa esiintyviä virheitä. Ciolkowskin ym. (2003) kyselytutkimuksessa puolet vastaajista ilmoitti käyttävänsä tarkistuslistoja. Lisäksi 35 % ilmoitti noudattavansa *ad hoc* periaatetta eli toimivansa ilman mitään tukimateriaalia.

Tarkistuslistoja käytettäessä tavallisesti kaikki katselmoijat käyvät saman listan läpi. Parnas ja Weis (1987) esittivät jo 80-luvulla kritiikkiä tähän ja esittivät ajatuksia katselmoitien organisoimiseksi uudella tavalla siten, että katselmoijille jaetaan tehokkuuden parantamiseksi erilaiset roolit. Toinen heidän keskeinen ajatuksensa oli antaa katselmoijalle perinteistä tarkistuslistapohjaista virheiden etsimistä aktiivisempi rooli. Monet myöhemmin esitetyt erilaiset sovellukset lukutekniikoista pohjautuvat näihin Parnasin ja Weisin esiin nostamiin ajatuksiin.

Erilaisten vaihtoehtoisten lukutekniikoiden tutkimus perustuu suurimaksi osaksi empiirisille kokeille, joissa niitä verrataan tarkistuslistojen käyttöön tai *ad hoc* menetelmään. Vaihtoehtoisia lukumenetelmiä on kehitetty tavallisesti tarkistuslistojen käytöstä esitetyn kritiikin pohjalta. Laitenberg ja DeBaud (2000) esittävät aiemman kirjallisuuden pohjalta seuraavat kritiikin pääkohdat:

- Tarkistuslistat ovat usein liian yleisiä eli niitä ei ole riittävästi mukautettu ympäristöön, jossa niitä käytetään.
- Konkreettinen ohjeistus tarkistuslistojen käytöstä usein puuttuu.
- Jos tarkistuslistat perustuvat aiempiin kokemuksiin, on niiden perusteella vaikea havaita uudenlaisia virheitä.

Aktiivinen erilaisten lukutekniikoiden tutkimus alkoi Porterin ja Vottan (1994) julkaisemasta artikkelista, jossa he esittelivät ajatuksen *skenaariopohjaisesta lukumenetelmästä (scenario-based reading)*. Skenaariopohjainen menetelmä perustuu nimensä mukaisesti skenaarioihin, joissa opastetaan tarkastajia yksityiskohtaisemmin kuin tavallisissa tarkistuslistan kysymyksissä. Toiseksi katselmoijille annettiin erilaiset skenaariot, jolloin he keskittyivät erityyppisiin virheisiin.

Porter ja Votta (1994) toteuttivat opiskelijoilla kokeen, jossa vertailtiin skenaariopohjaista, tarkistuslistapohjaista ja ad hoc -menettelyyn perustuvaa vaatimusmäärittelyn tarkastusta toisiinsa. Johtopäätöksenä he totesivat skenaariopohjaisen tarkastuksen muita tehokkaammaksi ja kyvykkäämmäksi löytää virheitä. Tarkistuslistojen käytöllä ja ad hoc -menettelyllä ei todettu olevan merkitsevää eroa.

Edellä kuvattua tutkimusta on myöhemmin tehty useaan otteeseen. Porter ym. (1995) esittävät saman kokeen toistettuna uudella opiskelijajoukolla ja saivat samanlaisia tuloksia. Myöhemmin Porter ja Votta (1998) toistivat saman kokeen vielä opiskelijoiden sijaan ohjelmistoalan ammattilaisilla. Tulokset vahvistivat edelleen alkuperäisiä johtopäätöksiä, mutta menetelmien välisen eron todettiin olevan pienempi kuin opiskelijakokeissa. Myös Miller ym. (1998) ovat saaneet opiskelijakokeessa tukea oletukselle, että skenaariot toimivat tarkistuslistoja tehokkaammin. Alkuperäistä tutkimusta toistettaessa on saatu myös toisenlaisia tuloksia. Fusaro ym. (1997) ja Sandahl ym. (1998) eivät todenneet omissa opiskelijakokeissaan merkittävää eroa näiden menetelmien välillä. He eivät kuitenkaan saaneet aivan vastakkaisia tuloksia, joten kokonaisuutena voidaan ajatella jonkinlaista tukea löytyvän skenaarioiden hyödyllisyydelle. Siihen vaikuttavista tekijöistä ei kuitenkaan voida näiden tutkimusten perusteella esittää kovin pitkälle meneviä johtopäätöksiä.

Basili ym. (1996) määrittelevät skenaariopohjaisen tekniikan eräänlaiseksi yläkäsitteeksi, joka jakaantuu erilaisiin sovelluksiin. Edellä kuvattua Porterin ja Vottan menetelmää he nimittävät *virhelähtöiseksi lukemiseksi (defect-based reading)*. He itse esittelevät artikkelissaan *perspektiivipohjaisen lukutekniikan*, joka on toisenlainen sovellus skenaariopohjaisesta lukemisesta. Perspektiivipohjaisessa lukumenetelmässä luodaan katselmoijalle aktiivisempi rooli virheiden etsinnässä. Basili ym. tarkastelevat myös vaatimusten katselmointia ja menetelmä on kehitetty siihen soveltuvaksi. Perusajatus on, että katselmoija luo ensin omaa rooliaan vastaavan dokumentin ja tarkastelee vaatimuksia sen pohjalta. Esimerkiksi testaajan roolissa toimiva katselmoija luo ensin vaatimusten pohjalta testisuunnitelman ja tarkastelee sitten katselmoitavaa vaatimusdokumenttia sen pohjalta.

Basili ym. (1996) tutkivat ammattilaisilla toteutetussa kokeessa perspektiivipohjaista lukemista verrattuna totuttuun tapaan lukea dokumenttia. Heidän alkuperäinen tavoitteensa perspektiivipohjaisen menetelmän kehittämiseksi oli saada dokumentissa esiintyvät virheet paremmin katettua. Kokeessaan he totesivat tämän tavoitteen toteutuvan, mutta sen lisäksi myös yksilöiden löytämä virheiden määrä oli suurempi käytettäessä perspektiivipohjaista menetelmää. Myöhemmin perspektiivipohjaista menetelmää on sovellettu koodin (Laitenberger ja DeBaud 1997) ja suunnitelmien (Laitenberger ym. 2000) tarkastukseen. Myös näissä tutkimuksissa on saatu rohkaisevia tuloksia. Mainittujen näkökulmien lisäksi Laitenberger ja DeBaud (1997) totesivat koodin katselmointia käsittelevässä tutkimuksessaan, että perspektiivipohjainen menetelmä vaikutti tasaavan katselmoijien kokemuseron vaikutusta löydettyjen virheiden määrään. Perspektiivipohjaisinkin lukutekniikan käytöstä on tosin saatu tuloksia, jotka eivät tue sen paremmuutta tarkistuslistoihin verrattuna (Sabaliauskaite ym. 2003).

Kaikki muutkin kirjallisuudessa esitetyt lukutekniikat ovat jollakin tavalla skenaariopohjaiseen ajatteluun pohjautuvia. Monissa menetelmissä keskeinen tekijä on katselmoijan kognitiivinen aktivointi niin, että hän joutuu aktiivisesti työskentelemään dokumentin kanssa. Seuraavassa on lyhyesti listattu muita tutkitussa aineistossa käsiteltyjä menetelmiä:

- Thelin ym. (2001, 2003 ja 2004b) esittelevät *käyttäjälähtöisen* lukumenetelmän (*usage-based reading*), jota tutkitaan suunnitelmien katselmoinnissa. Ajatuksena on ottaa pohjaksi vaatimusmäärittelyvaiheessa luodut käyttötapaukset ja tarkastella suunnitteludokumenttia niiden avulla. Keskeinen pyrkimys tässä menetelmässä on keskittyä katselmoineissa käyttäjän kannalta keskeisiin toiminnallisiin virheisiin.
- Dunsmore ym. (2001, 2002 ja 2003) esittelevät *abstraktiopohjaisen* menetelmän, jota on tutkimuksessa sovellettu koodin katselmointiin. Menetelmässä katselmoija luo koodia lukiessaan siitä korkeamman abstraktiotason kuvauksen. Tarkoituksena on varmistaa, että katselmoija on ymmärtänyt lukemansa koodin.
- Kelly ja Shepard (2004) esittelevät artikkelissaan *tehtäväkeskeisen* tarkastusmenetelmän (*task-directed inspection*), jota käytetään koodin katselmointiin. He puhuvat tarkastusmenetelmästä, koska menetelmä pitää sisällään myös joitakin ajatuksia tarkastuksen organisoimisesta, ei ainoastaan koodin lukemisesta. Keskeinen ajatus menetelmässä on samankaltainen kuin abstraktiopohjaisessa lukemisessa. Kelly ja Shepard vain vaativat katselmoijaa tuottamaan enemmän korkeamman tason kuvausta lukemastaan koodista. Katselmoijaa vaaditaan kuvaamaan erikseen ohjelmassa käytettyä tietoa, suorituslogiikkaa ja sen yhteyksiä suunnitteludokumenttaatioon.

Yhteenvedonä erilaisista lukutekniikoista voidaan todeta, että erilaisilla skenaariopohjaisilla menetelmillä on saatu lupaavia tuloksia verrattuna tarkistuslis-

toihin. Tulokset eivät ole kuitenkaan täysin yhdenmukaisia ja useimpia esitettyjä menetelmiä ei ole toistettu muiden tutkijoiden toimesta. Tuloksiin vaikuttaneista menetelmien taustalla olevista tekijöistä ei siis voida olemassa olevan tutkimuksen valossa tehdä suuria johtopäätöksiä. Lisäksi tutkimusta on tehty tähän mennessä ainoastaan verraten jotakin vaihtoehtoisista menetelmistä tarkistuslistapohjaiseen tai ad hoc menetelmään. Muita tässä esiteltyjä erilaisia menetelmiä ei ole vielä paljon verrattu toisiinsa.

2.3.2 Tehokkuustekijät katselmoinneissa

Erilaiset tehokkuustekijät katselmoinneissa ovat toinen suuri yksittäinen teema. Tähän kategoriaan luokiteltiin 19 % kaikista aineiston artikkeleista. Tehokkuustekijöillä viitataan tässä erilaisiin prosessin sisäisiin tekijöihin, jotka vaikuttavat katselmointien tehokkuuteen. Yleisemmällä tasolla katselmointien vaikutusta ohjelmistotuotannon tehokkuuteen tarkastellaan puolestaan johdon näkökulman yhteydessä. Oikeastaan edellä esiteltyt lukutekniikat voitaisiin sisällyttää tähän, sillä perustelu niiden kehittämiseksi on toiminnan tehostaminen. Sitä käsiteltiin kuitenkin edellä omana teemanaan ja siitä syystä sitä ei käsitellä enää tässä yhteydessä.

Tehokkuus ei ole tässä yhteydessä yksiselitteinen asia. Keskeisimmät tehokkuuteen liittyvät vakiintuneet mitat ovat löytyneiden virheiden määrä (*efficacy*) ja kustannustehokkuus (*efficiency*). Useissa tutkimuksissa pelkän virheiden määrän mittaamisen on katsottu riittävän. Käytännön kannalta se on kuitenkin usein riittämätön mittari, sillä resursseja ei voida ajatella olevan käytettävissä rajattomasti. Siksi suurin osa tehokkuutta arvioivaa tutkimusta pyrkii optimaaliseen kustannus-hyöty -suhteeseen eli siihen, että kulutetaan mahdollisimman vähän resursseja löytynyttä virhettä kohden. Sekään ei ole ongelmaton näkökulma, sillä vaikkapa jättämällä katselmointi tekemättä kokonaan käytetään varmasti vähiten aikaa löytynyttä virhettä kohden.

Katselmointeihin liittyvien tehokkuustekijöiden tutkimuksen alkukohtana voidaan pitää Vottan (1993) artikkelia, jossa hän asettaa kyseenalaiseksi palaverin merkityksen perinteisessä tarkastusprosessissa. Sen jälkeen katselmointien tehokkuuteen vaikuttavia tekijöitä on tutkittu useasta eri näkökulmasta. Seuraavassa on listattu kirjallisuusanalyysin aineiston pohjalta tärkeimpiä tehokkuuteen vaikuttavia tekijöitä:

- Yksilön suoritus
- Palaveri
- Valmistautuminen
- Materiaalin määrä
- Tiimin koko
- Koulutus
- Roolit

Yksilön suoritus

Sauer ym. (2000) tutkivat katselmointeja teoreettisesti käyttäytymisteorian avulla ja esittävät sen valossa, että yksilön asiantuntemus on keskeisin katselmointien tehokkuuteen vaikuttava tekijä. Knight ja Myers (1993) puolestaan totesivat tutkimuksessaan kokemuksen käytetystä ohjelmointikielestä vaikuttavan virheiden löytämiseen koodin tarkastuksessa. Myös useissa muissa tutkimuksissa on saatu tuloksia, joiden pohjalta on ajateltu yksilön olevan tehokkuutta selittävä tekijä (esim. Porter ym. 1997).

Palaveri

Palaverin merkitys tarkastusprosessissa on ollut Vottan (1993) artikkelista lähtien eräs suosituimmista aiheista katselmointien toteutuksessa. Votta vertasi tutkimuksessaan normaalia tarkastusprosessia sellaiseen sovellutukseen, jossa palaveria ei järjestetty, vaan valmistautumisvaiheessa löydetyt virheet koottiin muulla tavoin. Hän totesi ilman palaveria toteutetuissa tarkastuksissa löytyneen jopa enemmän virheitä. Tämän löydöksen perusteella hän kritisoi palaverin järjestämistä hyödyttömäksi ja ainoastaan resursseja kuluttavaksi elementiksi tarkastusprosessissa. Palaverin eduksi hän kuitenkin laski sen, että siellä löydöksistä suodattui pois suuri osa nk. vääristä virheistä (*false positives*). Väärillä virheillä tarkoitetaan sellaisia valmistautumisvaiheessa havaittuja löydöksiä, jotka myöhemmin paljastuvat perusteettomiksi.

Myöhemmin on toteutettu useita tutkimuksia, joissa on vahvistettu Vottan (1993) näkemyksiä siitä, että palaverin järjestäminen ei ole tehokasta virheiden löytämisen kannalta (esim. Porter ja Johnson 1997; Miller ym. 1998; Bianchi ym. 2001; Sabaliauskaite ym. 2004). Keskeinen tulos näissä tutkimuksissa on ollut se, että palaverit eivät tehostaneet virheiden löytymistä verrattuna ilman palaveria järjestettyyn katselmointiin. Johtopäätöksenä palaverin on todettu aiheuttavan prosessissa lähinnä ylimääräisiä kustannuksia ja useissa tutkimuksissa suositellaan perinteisen palaverin korvaamista jollain vaihtoehtoisella organisointitavalla.

Palaverien merkityksestä ei kuitenkaan olla kirjallisuudessa aivan yhtä mieltä. Kitchenham ym. (2002) pitävät tutkimusmenetelmiä käsittelevässä artikkelissaan huonona esimerkkinä johtopäätösten tekemisestä sellaisia tutkimuksia, joissa katselmointipalaveria pidettiin merkityksettömänä. Heidän mukaansa pelkästään virheiden löytämisen tehokkuutta tarkastelemalla ei voida pitää palaveria turhana, sillä palaverilla on todettu olevan monia muitakin hyviä puolia. D'Astous ja Robillard (2000) puhuvat tutkimuksessaan samasta asiasta korostaessaan palaverin monenlaista merkitystä. Lisäksi Ebert ym. (2001) ovat saaneet jopa monista muista tutkimuksista poikkeavia tuloksia tehokkuusnäkökulmaankin. He totesivat tutkimuksessaan fyysisesti palaveriin koontuvat tiimit tehokkaammiksi löytämään virheitä kuin tiimit, jotka eivät pitäneet palaveria.

Palaverien merkityksen tutkimisessa on siis oikeastaan kaksi eri linjaa. Toiset tutkijat keskittyvät tehokkuusnäkökulmaan ja tavallisesti kritisoivat pa-

laverin merkitystä. Toiset tutkijat puolestaan korostavat palaverin muunlaista merkitystä. Palaveri on eri tutkimuksissa nähty esimerkiksi oppimisen ja tiedon jakamisen paikkana (Johnson 1998). Myöskään kaikki tehokkuusnäkökulmasta palaverin merkitystä kritisoivat tutkimukset eivät suosittele johtopäätöksensä palaverista luopumista. Johnson ja Tjahjono (1998) esimerkiksi suosittelevat palaveria tarkastuksia käyttöönotettaessa. Työntekijöiden kokemuksen karttuessa he suosittelevat sitten tehokkuustekijöihin vedoten palaverista luopumista.

Valmistautuminen

Valmistautuminen on Gilbin ja Grahamin (1993) mukaan tarkastusprosessin merkittävin vaihe tarkastusten tehokkuuden kannalta. Tätä pidetään ilmeisesti niin itsestään selvänä asiana, että sitä ei ole juuri tutkimuksessa kyseenalaistettu, kuten esimerkiksi edellä käsiteltyä. Joitakin tätä ajatusta vahvistavia tuloksia on syntynyt erilaisten tutkimusten sivutuotteena. Laitenberger ym. (2002) havaitsivat tutkimuksessaan DaimlerChryslerilla, että virheitä löytyi sitä enemmän, mitä enemmän valmistautumiseen käytettiin aikaa. Myös Christensson ym. (1990) havaitsivat tutkimuksessaan korrelaation valmistautumisen ja löytyneiden virheiden määrän välillä.

Materiaalin määrä

Gilb ja Grahamin (1993) mukaan on tärkeää tarkastuksen tehokkuudelle, että valmistautumiseen käytetään riittävästi aikaa jokaista tarkastettavan materiaalin sivua kohden. Heidän suosituksensa ajankäyttöön esimerkiksi koodin tarkastuksessa on noin tunti sivua kohden. Faganin (1986) suositus ohjelmakoodin tarkastukselle oli puolestaan 125 riviä tunnissa. Jos aikaa ei käytetä riittävästi, jää Gilbin ja Grahamin mukaan helposti löytämättä syventymistä vaativat merkittävät virheet, joiden löytämisessä määrämuotoinen tarkastus kurinalaisesti toteutettuna on parhaimmillaan.

Nämä mainitut suositukset saavat jonkinlaista tukea Dunsmoren (2000) tutkimuksesta, jossa hän selvitti optimaalista materiaalin määrää koodin tarkastuksessa tehokkuuden näkökulmasta. Hän päätyi suosittelemaan materiaalin määräksi 200 riviä tunnissa. Käytännössä tilanne vaikuttaa kuitenkin olevan toinen. Seaman ja Basili (1998) totesivat, että tuntia kohden tarkastettiin noin 60 sivua ohjelmakoodia. Myös Bourgeois (1996) viittaa artikkelissaan samankaltaisiin havaintoihin.

Tiimin koko

Joissakin tutkimuksissa on pyritty etsimään tehokkuuden kannalta optimaalista tiimin kokoa. Porter ym. (1997) selvittivät kokeilla muutaman eri tekijän vaikutusta tarkastusten tehokkuuteen. Heidän tutkimuksessaan todettiin muun muassa, että tiimin koon kasvattaminen kahdesta neljään tarkastajaan ei merkittävästi parantanut tarkastusten tehokkuutta. Laitenberger ym. (2002) totesivat kuitenkin tutkimuksessaan kahden tarkastajan toimivan merkittävästi tehok-

kaammin kuin yksinäisen tarkastajan tarkastaja. Kaksi tarkastajaa on siis tavallinen käsitys tiimin optimikoosta näiden tutkimusten mukaan. Näissä tutkimuksissa tosin tehokkuudella tarkoitetaan mahdollisimman pientä ajankäyttöä löydettyä virhettä kohden. Jos tehokkuudella ajatellaan ainoastaan löydettyjen virheiden määrää, voidaan tiimiä kasvattamalla päästä jonkin verran parempiin tuloksiin.

Koulutus

Sopivan koulutuksen on joissakin tutkimuksissa todettu parantavan virheiden löytämistä katselmoinneissa. Ebert ym. (2001) totesivat tutkimuksessaan koulutuksen parantaneen katselmointien tehokkuutta. Rifkin ja Deimel (1994) tekivät saman havainnon jo aiemmin vertaillen erityyppisten koulutusten vaikutusta toimintaan. Heidän mukaansa merkittävä tekijä koulutuksessa oli käytännön kouluttaminen virheiden etsintään prosessitasoon keskittyvän koulutuksen sijasta. Lisäksi koulutuksen merkitys on tietysti itsessään looginen johtopäätös siitä, että katselmoijien yksilöllisen suorituksen on todettu olevan merkittävin tehokkuuteen vaikuttava yksittäinen tekijä.

Roolit

Erilaisten roolien merkitystä tarkastusprosessille on yritetty selvittää joissakin tutkimuksissa. Suurin osa roolien merkityksen tutkimisesta painottuu erilaisiin lukutekniikoihin, joita käsiteltiin jo edellisessä alakohdassa. Useissa eri lukutekniikoita käsittelevissä tutkimuksissa tarkistuslistojen käyttöä verrataan johonkin toisenlaiseen lukutekniikkaan. Ero tarkistuslistojen käyttöön on tavallisesti sekä lukutekniikassa että kehittyneemmässä tiimin sisäisessä työnjaossa. Sen vuoksi näistä tutkimuksista on vaikea erottaa sitä, mikä osuus havaitusta tehokkuuden parantumisesta on peräisin lukutekniikasta, mikä taas työn jakamisella eri rooleihin. Voidaan ehkä olettaa, että molemmilla on merkitystä.

Land ym. (2000) ovat tutkineet rooleja hiukan eri merkityksessä. Kun lukutekniikoiden yhteydessä rooleilla tarkoitettiin työnjakoa virheiden etsinnässä, Landin ym. tutkimuksessa tutkitaan prosessitason roolien vaikutusta kooditarkastuksiin. Prosessitason rooleilla tarkoitetaan prosessin toteuttamiseen liittyviä rooleja, kuten tarkastuksen johtaja. Land ym. eivät todenneet tällaisten organisointiin liittyvien prosessitason roolien vaikuttavan tarkastusten tehokkuuteen.

Muita huomioita

Porter ym. (1998) tutkivat samalla kertaa useamman eri tekijän vaikutusta tarkastusten tehokkuudelle. He muuttivat kokeissaan tiimin kokoa, katselmointikierrosten määrää ja sitä, korjattiinko virheet ennen mahdollista toista kierrosta. He totesivat, että tällaisilla prosessitason muutoksilla ei ollut juurikaan merkitystä tarkastusten tehokkuuteen. Biffel ym. (2001) vahvistavat omassa tutkimuksessaan, että katselmointikierrosten havaittiin laskevan merkittävästi tehokkuutta, kun mitattiin löydettyä virhettä kohden käytettyä aikaa. He tekivät kui-

tenkin lisäksi laskelman, jossa oletettiin jokaisen katselmoinnissa löytymättömän virheen aiheuttavan tietyn määrän kustannuksia myöhemmässä vaiheessa. Tämän laskelman mukaan toinenkin katselmointikierros maksaa itsensä takaisin säästyneenä työnä.

Porterin ym. (1998) mukaan syitä katselmointien tehokkuuden vaihteluun täytyy etsiä muista tekijöistä, kuten katselmoitavasta dokumentista ja katselmoijista. Jälkimmäinen vahvistaa omalta osaltaan yksilön henkilökohtaisen suorituksen vaikutusta tehokkuuteen. Dokumentin vaikutusta puolestaan ei ole paljon tutkittu. Christenson ym. (1990) ovat tosin todenneet dokumentin kompleksisuuden vaikuttavan tarkastusten kykyyn löytää virheitä.

2.3.3 Prosessit

Erilaisten katselmointiprosessien tutkimus on jäänyt tässä tutkimusaineistossa melko pieneksi. Erilaiset prosessit edustavat lähinnä tutkimuksen alkuaikaa Faganin (1976) tarkastusprosessin esityksestä 1990-luvun alkupuolelle asti. Myöhemmässä vaiheessa katselmointien tutkimus on vakiintunut lähes täysin tutkimaan tarkastusprosessin sisäisiä tekijöitä, ei kehittämään enää prosessitason variaatioita. Kaikkiaan tähän kategoriaan luokiteltiin kahdeksan artikkelia.

Faganin jälkeen Parnas ja Weis (1987) esittelivät 1980-luvulla oman menetelmänsä (*active design review*), jota käsiteltiin hiukan jo erilaisten lukumenetelmien yhteydessä. Saman vuosikymmenen lopulla Bisant ja Lyle (1989) esittivät kahden henkilön tarkastusmenetelmän pohjautuen usein esitettyyn kritiikkiin, jonka mukaan perinteinen tarkastus vie kohtuuttoman paljon resursseja. Myöhemmin Kusumoto ym. (1998) ovat tutkineet tämänkaltaisen kahden henkilön tarkastusta empiirisesti ja saaneet jonkinlaista tukea sen toimivuudelle käytännössä.

Myöhemmässä vaiheessa on esitetty enää joitakin prosessitasolla erilaisia katselmointimenetelmiä. Martin ja Tsai (1990) puhuvat n-kertaisesta tarkastuksesta, joka on kehitetty erityisen kriittisiä sovelluksia varten. Siinä keskeisenä ajatuksena on, että sama dokumentti katselmoidaan usean eri tiimin voimin ja löydetään siten suurempi osuus virheistä.

Knight ja Myers (1993) ovat esittäneet useampaan vaiheeseen perustuvan koodin katselmointia varten kehitetyn tarkastusmenetelmän (*phased inspection*). Siinä erilaisia asioita katselmoidaan eri vaiheissa eri ihmisten toimesta. Esimerkiksi kieli, ulkoasu ja toiminnallisten virheiden etsintä tapahtuvat prosessin eri vaiheissa.

Eräs tuore näkökulma tarkastusten organisointiin on Thelinin ym. (2004a) esittämä ajatus otokseen perustuvasta tarkastuksesta, joka voidaan tehdä, kun ei ole aikaa tarkastaa kaikkia dokumentteja. Siinä ensin tehdään esitarkastusvaihe, jossa heidän suosituksensa mukaan yksi tarkastaja käy läpi 20-30% dokumenteista. Tulosten perusteella arvioidaan, mitkä dokumentit sisältävät eniten virheitä ja priorisoidaan ne varsinaista tarkastusta varten arvioidun virhemäärän mukaan.

Harvoista erilaisista prosesseista mikään ei ole erityisesti herättänyt tutkijoiden mielenkiintoa, vaan tutkimus on keskittynyt suurimmaksi osaksi tarkas-

tusprosessin toteuttamisen variaatioihin. Keskeisin tarkastusprosessin organisointiin liittyvä tekijä tutkimuksessa on ollut palaverin organisointi, jota käsiteltiin jo tehokkuusseikkojen yhteydessä edellisessä alakohdassa.

Useissa eri tutkimuksissa on päädytty suosittelemaan resurssien optimaaliseen hyödyntämiseen vedoten jotain organisoitimenetelmää, jossa kevennetään jotenkin palaverin vaatimaa resurssien tarvetta. Esimerkiksi Sauer ym. (2000) päätyivät tehokkuusseikkoja kartoittavassa tutkimuksessaan esittämään menetelmää, jossa perinteisen palaverin sijaan ainoastaan pari asiantuntijaa käy läpi valmistautumisvaiheessa havaitut löydökset. Kelly ja Shepard (2004) esittävät tarkastuksesta sovelluksen, jossa on täysin luovuttu palaverista. Heidän mallissaan tarkastuksissa pidetään kuitenkin aloituspalaveri.

Eräs jonkin verran mielenkiintoa tutkimuksessa herättänyt näkökulma tarkastusprosessiin on työkalutuki. Siihenkin liittyvässä tutkimuksessa keskeinen teema on ollut kritiikki perinteistä palaveria kohtaan ja sen työkalutuettu organisointi. Työkaluihin liittyviä näkökulmia käsitellään erikseen kohdassa 2.5.

2.3.4 Sisältöön liittyvät erityiskysymykset

Sisältöön liittyvillä erityiskysymyksillä tarkoitetaan tässä sellaisia näkökulmia, jotka ottavat kantaa tarkastusten sisällöllisiin kysymyksiin. Ne vastaavat tavalisesti kysymykseen ”mitä asioita dokumenteista tulisi tarkastaa?”. Tähän kategoriaan luokiteltuja artikkeleita oli ainoastaan viisi kappaletta.

Macdonald ym. (1996) ja Dunsmore ym. (2000) tutkivat artikkeleissaan oliopohjaisen ohjelmakoodin tarkastamiseen liittyviä erityispiirteitä. Chernak (1996) esittää mallin tarkistuslistan systemaattista kehittämistä varten. Traore ja Aredo (2003) pohtivat artikkelissaan automaattisen verifiointin ja katselmointien välistä yhteyttä. De Almeida ym. (2003) puolestaan esittelevät kokemukseen pohjautuen parhaita käytänteitä koodin tarkistuksessa.

2.4 Johdon näkökulma

Johdon näkökulmaan keskittyvät tutkimukset ovat varmasti käytännön soveltamisen kannalta hyödyllistä aineistoa, mutta jostain syystä tähän keskittyvä tutkimus ei ole saanut aikaan sen mukaista mielenkiintoa. Johdon näkökulmaan luokiteltiin yllättävän pieni osa (14 %) aineiston sisältämästä tutkimuksesta. Toisaalta myös muihin tutkimuksiin sisältyy useita johdon kannalta mielenkiintoisia näkökulmia.

Johdon näkökulmaan luokitelluista artikkeleista ei löytynyt muuta selvästi yhtenäistä teemaa kuin katselmointin merkitys organisaatiolle. Tästä syystä kaikki muut johdon näkökulmaa käsittelevät tutkimukset luokiteltiin oman teeman alle. Artikkelien määrä näiden kahden kategorian kesken jakaantui melko tasan. Seitsemän artikkelia käsitteli selvästi tarkastusten merkitystä or-

ganisaatiolle ja kahdeksan artikkelia tarkasteli katselmointeja muilla tavoin johdon näkökulmasta.

2.4.1 Katselmointien merkitys organisaatiolle

Tämä kategoria painottuu selvästi katselmointien tutkimuksen alkuaikoihin ja tyypillinen artikkeli on raportti jonkin yrityksen kokemuksista. Bush (1990) kuvaa sitä, kuinka JPL:n organisaatiossa on laskettu tarkastusten hyötyä organisaatiolle. Hänen laskelmiensa mukaan jokaisen löydetyn virheen korjaaminen myöhemmin prosessissa maksaisi keksimäärin 1700 dollaria ja siten jokainen tarkastus säästää keskimäärin 25 000 dollaria. Russel (1991) on todennut jokaisen tarkastuksiin käytetyn tunnin säästävän 33 tuntia ylläpitotyötä. Doolan (1992) on todennut, että jokainen tarkastuksiin käytetty tunti maksaa itsensä takaisin 30-kertaisesti. Myös Grady ja Van Slack (1994) raportoivat samankaltaisia tuloksia Hewlett Packardin kokemuksista.

Eräs tähän kategoriaan luokiteltu näkökulma on katselmointien vertaaminen testaukseen. Basili ja Selby (1987) totesivat kokeessa, että koodia luettaessa löydettiin virheitä enemmän ja tehokkaammin kuin eri testausmenetelmillä. So ym. (2002) eivät todenneet omassa vastaavassa tutkimuksessaan eroa löydettyjen virheiden määrässä, mutta havaitsivat koodin lukemisen kuluttavan vähiten resursseja. Roper ym. (1997) eivät puolestaan omassa tutkimuksessaan havainneet merkittävää eroa eri koodin lukemisen ja testausmenetelmien välillä. Roper ym. (1997) viittaavat artikkelissaan useisiin tutkimuksiin, joissa on saatu tästä keskenään ristiriitaisia tuloksia.

Koodin lukemisen tehokkuudesta verrattuna testaukseen ei siis ole aiemman tutkimuksen valossa saatu vakuuttavaa näyttöä puoleen tai toiseen. Sen sijaan keskeinen johtopäätös useissa tutkimuksissa on, että eri testausmenetelmät ja koodin lukeminen ovat vahvoja erityyppisten virheiden löytämisessä (esim. Roper ym. 1997). Johtopäätöksenä suositellaan tavallisesti menetelmien yhdistämistä.

Eräs erillinen tähän kategoriaan luokiteltu artikkeli on Müllerin (2004) tutkimus, jossa hän vertailee pariohjelmointia ja katselmointia toisiinsa. Kokeessa verrattiin yksinäisen ohjelmoijan katselmointien tukemana tekemää työtä pariohjelmointiin. Lopputuloksen laadussa ei havaittu merkittävää eroa, mutta pariohjelmointi oli jonkin verran kalliimpaa.

2.4.2 Muut johdon näkökulmaan liittyvät asiat

Tähän kategoriaan luokiteltiin muut johdon näkökulmaa ensisijaisesti käsittelevät artikkelit, joita löytyi aineistosta kaikkiaan kahdeksan kappaletta. Tyypillisiä aiheita olivat katselmointien kehittäminen, kustannuslaskenta ja niiden yhteys muuhun toimintaan.

Tässä kategoriassa on useampia artikkeleita, jotka käsittelevät jotenkin metriikoita. Barnard ja Price (1994) käsittelevät jo 1990-luvun alkupuolella artikkelissaan erilaisia katselmointimetriikoita. Madachy (1996) esittelee puolestaan mallin, jonka pohjalta voidaan arvioida tarkastusten vaikutusta tuotanto-

prosessiin. Briand ym. (1998) ovat omassa tutkimuksessaan kehittäneet mallin erityisesti tarkastusten kustannusten laskemista varten.

Chatzigeorgiou ja Antoniadis (2003) pohtivat artikkelissaan katselmointien aikataulutusta projektissa. He totesivat tutkimuksessaan, että tarkastusten ja muiden tehtävien vaatimalla työkuormalla oli tapana kasautua projekteissa samoille viikoille. Kusumoto ym. (1996) esittävät mallin, jonka mukaan katselmointien resursointi voidaan organisoida.

Jakob ja Pillai (2003) käsittelevät artikkelissaan tilastolliseen virheiden seurantaan perustuvaa menetelmää, jonka avulla voidaan kehittää sekä ohjelmointityötä, että koodin katselmointia. Chaar ym. (1993) käsittelevät puolestaan yhdessä testausta ja tarkastuksia, mutta myös heidän näkökulmansa painottuu toiminnan kehittämiseen.

Edellisten lisäksi Jalote ja Haragopal (1998) nostavat esiin hiukan erilaisen näkökulman tässä kategoriassa. He käsittelevät organisaation muutosvastarintaa tarkastusten käyttöönoton yhteydessä ja esittävät artikkelissaan mallin käyttöönoton toteuttamiseksi. Lähtökohta tässä mallissa on muutosvastarinnan ehkäisy.

2.5 Muut teemat tutkimuksessa

Tähän kohtaan on koottu yhteisesti lyhyt esittely muista kirjallisuusanalyysissä muodostuneista luokista. Näitä teemoja ovat tässä virheiden määrän arviointi, työkalunäkökulma ja katselmoiteja jollakin lailla kokonaisvaltaisesti käsittelevät tutkimukset. Näihin teemoihin luokiteltiin yhteensä 32 % aineiston artikkeleista.

2.5.1 Virheiden määrän arviointi

Tähän kategoriaan kuuluvat sellaiset tutkimukset, jotka pyrkivät arvioimaan katselmoinnin jälkeen dokumenttiin jääneiden virheiden määrää. Lähtökohtana tälle tutkimukselle esitetään tavallisesti se, että tämä tieto on hyödyllinen projektinhallinnan kannalta, kun tehdään päätöksiä jatkotoimenpiteistä katselmoinnin jälkeen.

Virheiden määrän arviointi oli eräs suurista teemoista koko kirjallisuusanalyysin aineistossa. Sitä käsittelee yhteensä 18 artikkelia, mikä on 16 prosenttia koko aineistosta. Lisäksi sen osuus on viime vuosina ollut kasvussa niin, että se on ollut 2000-luvulla jopa suosituin tutkimusaihe katselmointien tutkimuksessa.

Virheiden määrän arvioinnin jonkinlaisena lähtökohtana voidaan pitää Eickin ym. (1992) artikkelia, jossa on sovellettu ekologiassa tunnettua toistopyydyystysotantaa (capture-recapture sampling) virheiden määrän arviointiin. Mallin periaate on se, että löytyneiden virheiden perusteella yritetään päätellä dokumentissa olevien virheiden kokonaismäärä. Suurin osa virheiden määrän arviointia käsittelevästä tutkimuksesta yrittää tämän lähtökohdan pohjalta löy-

tää mahdollisimman luotettavan laskentamallin virheiden määrän arvioimiseksi.

Edellä kuvattu lähestymistapa edellyttää, että katselmointien toimintaa organisaatiossa tunnetaan ennestään. Esimerkiksi Padberg (2002) esittää, kuinka katselmoineista kerättävästä historiatiedosta voidaan muodostaa eräänlainen yritysکوhtainen profiili katselmoinnissa löydetyistä ja löytämättömistä virheistä. Virheiden määrän arviointiin käytettävässä laskentamallissa tämä profiili edustaa tiettyjä vakioita ja katselmoinnissa löydetyt virheet muuttujia. Petersson ym. (2004) ovat luoneet kattavan katsauksen tämän alueen siihen mennessä tehdystä tutkimuksesta. Useissa tutkimuksissa erilaisia laskentamenetelmiä on pidetty lupaavina, mutta ne on todettu vielä käytännön soveltamisen kannalta liian epätarkkoiksi.

Virheiden määrän tutkimus on viime vuosina jakaantunut kahteen eri osaan. Mainittuja laskentamenetelmiä kutsutaan *objektiiviseksi arvioinniksi*. Niiden rinnalle on syntynyt tutkimusta, jossa on tutkittu myös katselmointiin osallistuneiden ihmisten *subjektiivista arviointia*. Tällaisissa tutkimuksissa osallistujia pyydetään yksinkertaisesti arvioimaan, kuinka suuri osa dokumentin virheistä katselmoinnissa löydettiin.

Biffel ja Grossman (2001) esittävät, että tutkitut objektiiviset arviointimenetelmät olivat aika epätarkkoja, varsinkin kun virheiden löytymisprosentti oli pieni. Objektiivinen menetelmä vaikutti toimivan kohtuullisen hyvin joissain olosuhteissa. Biffel ja Grossman pitävät keskeisenä kysymyksenä sitä, milloin tuloksiin voi luottaa. He päätyivät suosittelemaan, että objektiivisiä menetelmiä pitäisi täydentää subjektiivisella arvioinnilla.

Subjektiivista arviointia ovat tutkineet muun muassa El Emam ym. (2000), Yin ym. (2004) ja Thelin (2004). El Emam ym. (2000) tutkivat subjektiivisen arviointi tarkkuutta ammattilaisilla tehdyssä kokeessa ja totesivat mediaanivirheen olleen lähellä nollaa. Yin ym. (2004) tekivät vastaavan kokeen opiskelijoilla ja saivat samankaltaisia tuloksia. Näiden kokeiden perusteella siis katselmoijat onnistuvat subjektiivisessä arvioinnissa tavallisesti kohtuullisen hyvin. Käytännön soveltamisen kannalta arvioiden luotettavuus on kuitenkin heikko. Vähintäänkin pitäisi pystyä luotettavasti arvioimaan, milloin arvioihin voi luottaa ja milloin ei. Thelin (2004) vertailee artikkelissaan objektiivista ja subjektiivista arviointia keskenään ja päätyy pitämään objektiivista näistä luotettavampana.

2.5.2 Työkalunäkökulma

Tähän kategoriaan luokiteltiin aineistossa kaikki tietokoneavusteiseen katselmointiin liittyvät artikkelit. Työkalunäkökulma on selvästi myös eräs tutkijoita kiinnostava näkökulma, sillä siihen luokiteltiin 13 artikkelia eli 12 % kirjallisuusanalyysin aineistosta. Eriytynyt mielenkiinto työkaluihin kohdistui tämän aineiston valossa 1990-luvun loppupuolella internetin kehityksen myötä. Tavallinen näkökulma tässä kategoriassa on tutkia työkalutuettua hajautettua katselmointia.

Työkaluihin keskittyvä tutkimus tutkitussa aineistossa alkoi 1990-luvun alkupuolella, jolloin Mashayekhi ym. (1993) esittelivät tarkastusten hajauttamis-

ta tukevan työkalunsa. Johnson (1994) esitti myös kehityksen alkuvaiheessa samankaltaisia ajatuksia työkalutuesta. Keskeinen ajatus Johnsonin ajatuksissa on työkalulla mahdollistettu asynkroninen katselmointi, jonka ansiosta katselmoijien ei tarvitse olla samaan aikaan samassa paikassa. Johnsonin mukaan tällä saavutetaan kustannusten säästöä ja katselmointien joustavuuden parantuminen.

Erilaisia työkaluja ovat esitelleet aineistossa myös Tervonen (1996), Macdonald ja Miller (1997 ja 1999) sekä Perry ym. (2002). Yhteistä näille on se, että kaikissa mainituissa työkaluissa pääpaino on työn prosessitasoisessa tukemisessa. Aineistossa tästä poikkeaa selkeästi Andersonin ym. (2003) työkalu, joka keskittyy katselmoitavan materiaalin havainnollistamiseen. Hedberg (2004) on koonnut tässä käsiteltyä aineistoa kattavamman koosteen erilaisista katselmointityökaluista ja niiden ominaisuuksista.

Osa työkaluja käsittelevästä tutkimuksesta on painottunut työkalun tekni-
sten ominaisuuksien sijasta sen tukeman prosessin tutkimiseen. Esimerkiksi Macdonald ja Miller (1998) vertaavat opiskelijakokeessa työkalutuettua tarkastusta perinteiseen tarkastukseen. Työkalutuettu tarkastus toteutettiin synkronisesti eli katselmoijat osallistuivat samaan aikaan yhteiseen ”palaveriin”. Ainoa erotus toiseen ryhmään oli, että he keskustelivat ainoastaan tietokoneen näppäimistön välityksellä. Macdonald ja Miller eivät todenneen virheiden löytymisessä eroa ryhmien välillä ja pitivät tätä tulosta rohkaisevana työkalun käyttöä ajatellen.

Tyran ja George (2002) tekivät vastaavanlaisen kokeen ja totesivat työkaluryhmän löytäneen enemmän virheitä kuin perinteiseen palaveriin osallistuneet. He selittivät tulosta monenlaisilla keskustelun luonteeseen liittyvillä syillä. Perinteisessä palaverissa esimerkiksi keskustelu harhautui enemmän sivupo-
luille ja useammin keskustelua dominoi joku vahva persoona. Näiden tekijöiden selitettiin ehkäisevän uusien virheiden löytymistä tai jopa valmistautumisvaiheessa löydettyjen virheiden esille tuomista.

Työkalunäkökulmasta tehty tutkimus liittyy, kuten edellä mainituissa tutkimuksissa, usein perinteisen palaverin merkitykseen katselmointiprosessissa. Perpich ym. (1997) tutkivat asynkronista hajautettua tarkastusta tukevan työkalun käyttöönottoa kohdeyrityksessä (Lucent Technologies). He painottivat tuloksissaan sitä, kuinka löydettyjen virheiden määrä pysyi ennallaan, mutta kustannukset olivat aiempaa pienemmät. Stein ym. (1997) saivat hyvin paljon näitä havaintoja tukevia tuloksia. He havaitsivat kuitenkin omassa tutkimuksessaan, että asynkroninen tarkastus ei löytänyt hyvin tietyntyyppisiä virheitä. Lisäksi he painottavat palaverilla olevan muutakin merkitystä kuin virheiden etsiminen eivätkä he suosittele perinteisen tarkastuksen korvaamista asynkronisella.

2.5.3 Katselmointien kokonaisvaltainen tarkastelu

Tähän kategoriaan luokiteltiin artikkelit, jotka tarkastelevat katselmointien tutkimusta tai käytäntöä laajasta perspektiivistä. Aineiston artikkeleista ainoastaan neljä kuului tähän luokkaan, mutta näkökulma poikkeaa sen verran muista, että niitä käsitellään tässä erikseen. Tähän luokkaan kuuluvien artikkelien vähyyttä

selittänee se, että alan historia on kohtuullisen lyhyt ja toisaalta tämän tyyppisen tutkimuksen tekeminen on usein aikaa vievää ja kallista.

Tarkasteltu aineisto sisälsi kaksi tarkastusmenetelmän tutkimusta tarkastelevaa kirjallisuuskatsausta. Tässä luvussa käsiteltiin jo aiemmin Laitenbergin ja DeBaudin (2000) julkaisemaa kirjallisuusanalyysia, jossa he pyrkivät jäsentämään kokonaisvaltaisesti ennen vuotta 1998 tehtyä tarkastuksiin liittyvää tutkimusta. Toisen samantyyppisen tutkimuksen ovat julkaisseet Aurum ym. (2002). Siinä tarkastellaan tarkastusmenetelmän kehittymistä 25 vuoden aikana. He keskittyvät katsauksessaan tarkastusmenetelmään ja sen sovelluksiin, eivätkä käsittele alan tutkimusta niin kokonaisvaltaisesti kuin Laitenberg ja DeBaud.

Ciolkowski ym. (2003) ovat julkaisseet aineiston ainoan tutkimuksen, jossa on vakavasti yritetty kartoittaa katselmointien toteuttamista ohjelmistoyrityksissä. Heidän tutkimuksessaan noin 30-40 % tutkituista yrityksistä katselmoi säännöllisesti dokumentteja hiukan dokumenttityypistä riippuen. Vaatimusdokumentteja katselmoidaan hiukan säännöllisemmin kuin koodia. Toinen pienimuotoinen yritys kartoittaa katselmointien toteuttamista käytännössä on Johnsonin (1998) raportoima vapaamuotoinen kysely. Siinä 80 % vastanneista ohjelmistoalan ammattilaisista ilmoitti käyttävänsä tarkastuksia työssään epäsäännöllisesti tai ei ollenkaan. Varsinainen sisältö Johnsonin artikkelissa on kuitenkin sen pohtiminen, kuinka siihen mennessä tehdyn tutkimuksen valossa tarkastuksia tulisi kehittää tulevaisuudessa.

2.6 Yhteenveto

Tässä luvussa on tarkasteltu systemaattisesti katselmointeja käsittelevää kirjallisuutta vuosilta 1990-2004. Tutkimuksen kohteena oli kymmenen keskeistä ohjelmistotekniikan alan julkaisusarjaa, joista katselmointeja käsitteleviä artikkeleita löytyi yhteensä 110 kappaletta. Ennen 1990-lukua katselmointeja käsitteleviä tutkimuksia on julkaistu vain muutamia.

Tarkastellulla aikavälillä katselmointien tutkimus on kasvanut voimakkaasti aina viime vuosiin asti. Katselmoinnit siis kiinnostavat tutkijoita edelleen ja jopa enenevässä määrin. Katselmointeja käsittelevää tutkimusta tarkasteltaessa viisivuotisjaksoina, voidaan havaita seuraavanlaisia muita trendejä:

- Vahvimmin tutkittua aluetta ovat dokumenttien lukutekniikat ja erilaiset katselmointien tehokkuuteen vaikuttavat tekijät.
- Virheiden määrän arviointi on viime vuosina noussut jopa suosituimmaksi tutkimuskohteeksi.
- Tutkimuksen alkuvaiheelle tunnusomaista olivat erilaisten katselmointimenetelmien esittely ja katselmointien merkityksen arviointi kohdeyrityksissä. Myöhemmin 1990-luvun puolenvälin jälkeen tällaisia tutkimuksia on tehty hyvin vähän.

- Tutkimuksen näkökulma on melko vakiintunut tutkimaan tarkastusmenetelmää. Alkuaikojen jälkeen tutkimus on keskittynyt suureksi osaksi erilaisiin tarkastusmenetelmän erityiskysymyksiin, kuten esimerkiksi palaverin merkityksen arviointiin.

Täytyy muistaa, että aineiston kokonaismäärä ei ole kovin suuri, joten kovin luotettavia johtopäätöksiä tutkittujen ilmiöiden yleisyydestä ei voida tehdä. Aineistossa oli kaikkiaan 110 artikkelia, jotka edustavat hyvin erilaisia näkökulmia katselmointeihin. Näistä näkökulmista Porterin ja Vottan (1994) skenaariopohjaista lukumenetelmää käsittelevä tutkimus on ainoa, jota on vakavasti yritetty toistaa useampaan kertaan. Sen lisäksi useiden eri yritysten kokemuksista kertovien raporttien mukaan voidaan ajatella, että katselmoinneista on yleisesti hyötyä yrityksille. Alalla tarvitaan kuitenkin vielä paljon tutkimusta ennen kuin katselmointien toimintaa voidaan todella ymmärtää todellisessa kontekstissa. Vertailukohdaksi voidaan ottaa vaikkapa psykologian tutkimus, jossa jotain yksittäistä ilmiötä on voitu tutkia sadoissa keskenään vertailukelpoisissa kokeissa.

Tehdyn tutkimuksen perusteella on hankala arvioida tutkimuksen suuntaa tulevaisuudessa. Suuntaan vaikuttaa selvästi tutkijoiden henkilökohtainen kiinnostus. Esimerkiksi virheiden määrän arviointi on viime vuosina ollut katselmointien tutkimuksessa suosituin yksittäinen aihe, vaikka todennäköisesti harva tutkija pitää sitä käytännön soveltamisen kannalta kaikkein merkittävimpänä tutkimuskohteena. Toinen tutkimuksen suuntaan vaikuttava asia ovat käytännön mahdollisuudet toteuttaa tutkimusta. Vaikka tutkijat ovat yhtä mieltä siitä, että opiskelijakokeiden tuloksia ei välttämättä voida luotettavasti yleistää todelliseen kontekstiin, perustuu merkittävä osa alan tutkimuksesta opiskelijakokeisiin. Tutkimuksen suuntaa tulevaisuudessa on siis näistä tekijöistä lähtien hankala arvioida, mutta tehdyn tutkimuksen pohjalta voidaan nähdä jotain siitä, mihin tutkimusta olisi käytännön soveltamisen kannalta hyvä suunnata.

Parnas ja Lawford (2003) pohtivat tarkastusten tulevaisuutta ja nostavat siitä esiin tarpeen tarkentaa prosessia, huomioida reaaliaikajärjestelmien erityispiirteitä ja kehittää työkalutukea. Prosessin tarkentamisella he tarkoittavat lähinnä erilaisten lukutekniikoiden kehittämistä. Siihen voidaan yhtä hyvin yhdistää erilaisten tehokkuustekijöiden tutkimus. Näiden asioiden tutkiminen vaatii varmasti jatkossa lisää tarkennusta, jotta tarkastukset voidaan suunnitella toimimaan mahdollisimman tehokkaasti. Selvä puute näiden asioiden tutkimuksessa on niiden ymmärtäminen käytännön kontekstissa. Yleisesti tarkastuksia käsittelevässä tutkimuksessa on pyritty ymmärtämään, mitkä tekijät tarkastusprosessissa vaikuttavat tehokkuuteen rajallisessa koeympäristössä. Tulevaisuudessa on kuitenkin tarve ymmärtää, mitkä ympäristötekijät vaikuttavat tehokkuuteen todellisessa erilaisissa toimintaympäristöissä. Sitä tunnetaan tois-
taiseksi huonosti.

Toinen Parnasin ja Lawfordin (2003) mainitsema näkökulma tarkastusten tulevaisuuteen oli nykyisin yleisten reaaliaikajärjestelmien huomioiminen. He ottavat esimerkiksi Xun (2003) artikkelin, jossa hän esittää menetelmän, jolla voidaan parantaa reaaliaikajärjestelmien tarkastettavuutta suunnitteluvaiheessa.

Tällaisten sisällöllisten erityiskysymysten huomioimiselle yleensä tulee varmastikin tarvetta tulevaisuudessa. Kun prosessitasolla joko käytännössä tai tutkimuksessa on päästy tyydyttävälle tasolle, voidaan tehokkuuden parantamista hakea sisällön ymmärtämistä kehittämällä. Tällaista tutkimusta aineistossa on melko vähän. Harvoista tutkimuksista esimerkkeinä ovat Macdonald ym. (1996) ja Dunsmore ym. (2000), jotka ovat tutkineet oliopohjaisen koodin tarkastamisen erityispiirteitä.

Kolmas asia, jonka Parnas ja Lawford (2003) tuovat tulevaisuudennäkymissä esille, on työkalutuki. He käyttävät esimerkkinä Andersonin ym. (2003) työkalua, jossa avustetaan käyttäjää virheiden löytämisessä ja sitä pyritään sitä kautta tehostamaan tarkastusprosessia. Tässä luvussa esitellyn kirjallisuusanalyysin valossa tutkijoiden kiinnostus työkalunäkökulmaan on hiukan hiipunut 2000-luvulla. Kuitenkin työkalutuki on asia, jolle on varmasti tarvetta tulevaisuudessa.

Yhteenvetona katselmointien tutkimuksesta voidaan sanoa, että siihen liittyvät osa-alueet ovat viimeisen kymmenen vuoden aikana alkaneet hahmottua. Tutkimus on kuitenkin aika alussa ja sitä tarvitaan vielä paljon ennen kuin katselmoiteja voidaan tietämyksen pohjalta tehokkaasti soveltaa erilaisissa toimintaympäristöissä.

3 TARKASTUSKÄYTÄNTEIDEN KYPSYYSTASON MITTAAMINEN

Tässä luvussa esitellään tarkastuskäytänteiden kypsyystason mittaamista. Kapale painottuu suureksi osaksi tämän tutkimuksen yhteydessä kehitettyyn ICMM-malliin, koska sillä on oleellinen osa myös tutkimuksen empiirisen osan toteuttamisessa. Tutkimuksen alussa oli tarve saada työkalu, joka toimisi viitekehyksenä organisaatioita vertailtaessa ja joka voisi toimia apuvälineenä tarkastuskäytänteiden kehittämisessä. Aiemmasta tutkimuksesta ei kuitenkaan löytynyt tähän tarkoitukseen soveltuvaa valmista mallia ja siitä syystä ratkaisuksi päädyttiin kehittämään ICMM.

Seuraavassa kohdassa 3.1 esitellään lyhyesti tarkastuskäytänteiden kypyyden arviointiin aiemmin esitetyjä malleja. Kohdassa 3.2 esitetään artikkeliin Kollanus (2005a) pohjautuen yksityiskohtaisesti ICMM-malli ja sen tausta. Lopuksi kohdassa esitetään yhteenveto keskeisimmistä tässä luvussa tehdyistä johtopäätöksistä.

3.1 Muita tarkastuskäytänteiden kypsyysmalleja

Tarkastuskäytänteiden kypsyystason mittaamista ei ole kovinkaan paljon aiemmin tutkittu. Joitakin sen suuntaisia esityksiä kirjallisuudessa on kuitenkin olemassa, mutta niistä jokaisessa on omat ongelmansa tämän tutkimuksen kannalta.

Testausprosessien kypsyyttä arvioivia malleja on olemassa useita ja osaan niistä on sisällytetty myös katselmointeihin liittyviä käytänteitä. Esimerkiksi Burnstein ym. (1999) ottavat katselmointeja jonkin verran huomioon omassa mallissaan. Heidän mallinsa on rakenteeltaan sellainen, että katselmointeja ei voida tarkastella siinä omana prosessinaan. Ericsonin ym. (1997) TIM-malli on ainoa kirjallisuudessa tunnettu testausprosessien kypsyystasomalli, jossa katselmointeja voidaan tarkastella erillisenä prosessina. Sen ongelma on kuitenkin mallin rajallisuus katselmointien kannalta relevanttien tekijöiden kuvaamisessa.

Toiseksi TIM mallin käyttämisen tueksi ei ole olemassa riittävää dokumentaatiota, eikä mallissa vaadituille kriteereille ole esitetty juurikaan perusteluja..

Eräs harvoja tarkastusten kypsyystason arvioimiseen keskittyviä malleja on esitetty Gradyn ja Van Slackin (1994) artikkelissa. Malli on kehitetty Hewlett Packardin organisaatiossa ja sen taustalla ovat organisaation omat kokemukset tarkastusten käytöstä pitkällä aikavälillä. Mallin heikkous onkin siinä, että se kuvaa paremmin HP:n kaltaisen jättikokoisen yrityksen tarkastuskäytänteiden käyttöönottoa kuin sitä, kuinka tällä hetkellä tarkastuksia kehittävän organisaation kannattaisi toimia. Tämä Gradyn ja Van Slackin malli sisältää kuitenkin useita hyviä huomioita ja se on omalta osaltaan vaikuttanut merkittävästi seuraavassa kohdassa esiteltävän ICMM-mallin kehittämiseen.

Toisen tarkastuskäytänteiden kypsyyttä arvioivan mallin ovat kehittäneet Tervonen ym. (2001). Mallia ovat kehittäneet edelleen Harjuma ym. (2004). Heidän esityksessään on määritelty 29 erilaista indikaattoria, jotka ovat kypsyyden määrittäviä tekijöitä 12 eri prosessialueella. Mallissa on hyvin hahmotettu tarkastuskäytänteisiin liittyviä prosesseja ja sen vuoksi se voi olla hyvä apuväline tarkastusten kehittämisessä organisaatiossa. Mallissa eri prosessialueita ei ole priorisoitu, eikä se perustu kypsyytasoihin samalla tavoin kuin esimerkiksi Gradyn ja Van Slackin (1994) malli. Sen vuoksi Harjumaan ym. esittämän mallin avulla tässä tutkimuksessa tärkeässä osassa oleva organisaatioiden vertaileminen on hankalaa.

Aiemmasta tutkimuksesta ei vaikuttanut löytyvän sellaista mallia, joka olisi ollut käyttökelpoinen ja soveltunut tämän tutkimuksen tavoitteisiin. Siitä syystä ennen tutkimuksen empiiristä osuutta kehitettiin aiempiin tutkimustuloksiin pohjautuva oma tarkastuskäytänteiden kypsyyttä arvioiva malli. Seuraavaksi käsitellään tätä mallia.

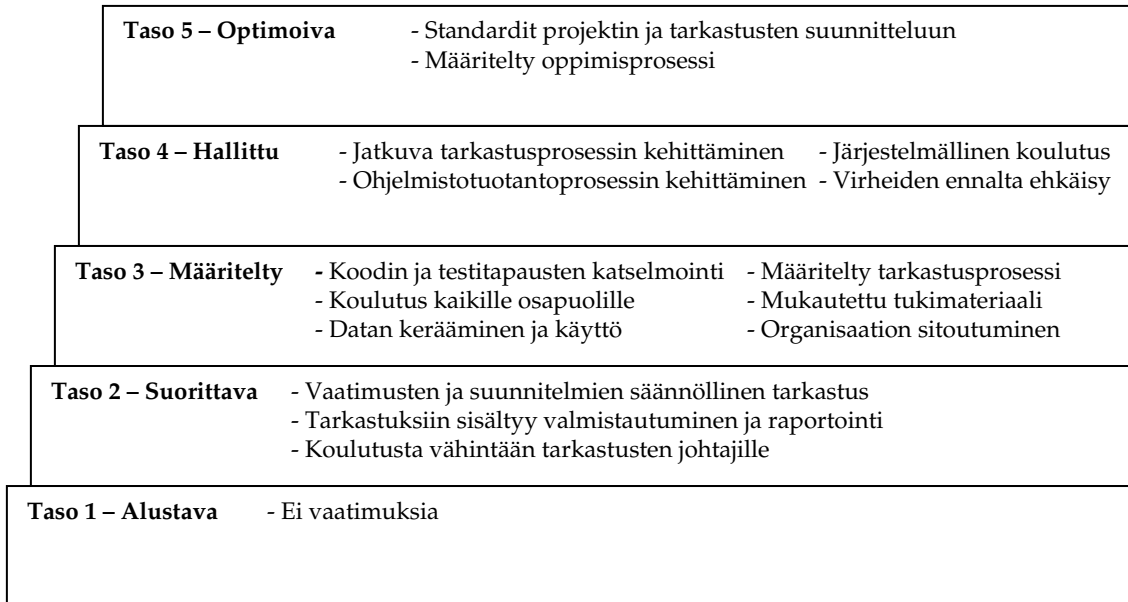
3.2 ICMM – Inspection Capability Maturity Model

ICMM on malli, joka on kehitetty mittaamaan tarkastusten kypsyystasoa ohjelmistoja tuottavassa organisaatiossa. Kypsyystason mittaamisen lisäksi ICMM:n tarkoitus on olla kehys, jonka avulla voidaan kehittää organisaation tarkastuskäytänteitä. Malli ei kerro, kuinka käytänteitä tulisi kehittää, mutta voi auttaa tunnistamaan organisaation tarkastuskäytänteiden kypsyystasoa vastaavia miellekkäitä kehittämiskohteita.

ICMM:n esikuvana on käytetty koko ohjelmistotuotantoprosessin kypsyyttä mittaavaa Software Engineering Institutin CMMI-mallia (SEI 2002). Ajatus kypsyytasoihin noudattaa samaa periaatetta. Tasojen jakaminen viiteen portaaseen vaikuttaa luonnolliselta tavalta esittää organisaation kypsyyttä. Esimerkiksi testauksen kypsyyttä mittaava TIM (Ericson ym. 1997) sekä tarkastusten kypsyyttä kuvaava Gradyn ja Van Slackin (1994) esitys ovat päätyneet samankaltaiseen jaotteluun.

Jokaisella tasolla on tietty määrä vaatimuksia, joista kaikki täytyy täyttää ennen kuin organisaatio saavuttaa tason. Siis organisaatio voi täyttää esimer-

kiksi lähes kaikki kolmannen tason vaatimukset, mutta se on silti tasolla yksi, jos ei täytä kaikkia toisen tason vaatimuksia. Perusajatuksen lisäksi ICMM:n yksittäiset tasot muistuttavat periaatteeltaan CMMI:n vastaavia tasoja. ICMM:n eri tasot on kuvattu tiivistetysti kuviossa 4.



KUVIO 4 Tiivistelmä ICMM-mallin tasoista

Ensimmäinen taso on lähtötaso, jolla ei ole mitään vaatimuksia. Toisella tasolla oleva organisaatio toteuttaa ainakin joitakin tarkastuksia säännöllisesti. Kolmannella tasolla olevalla organisaatiolla on hyvin määritelty prosessi tarkastuksia varten ja se toimii prosessin mukaisesti. Neljäs taso tuo mukaan tarkastusten määrätietoisuuden kehittämisen ja ajatuksen virheiden ennaltaehkäisystä. Viides taso keskittyy tietoiseen kustannustehokkuuden parantamiseen. Seuraavaksi käydään yksityiskohtaisesti jokainen näistä tasoista läpi.

3.2.1 Taso 2 – Suorittava

Tasolla kaksi organisaatio yksinkertaisesti toteuttaa tarkastuksia säännöllisesti. Tarkastuksia täytyy toteuttaa jokaisessa projektissa, mutta organisaatiolla ei tarvitse olla dokumentoitua prosessia tarkastuksia varten. Tarkastukset tällä tasolla voivat siis olla melko heikosti organisoitua toimintaa ja ne eivät välttämättä toimi kovin tehokkaasti. Kuitenkin tälläkin tasolla tarkastuksilta vaaditaan joitakin kaikkein oleellisimpia ominaisuuksia.

Vaatimukset tällä tasolla ovat melko matalat. Tämä oli tietoinen valinta ICMM-mallin kehittämisvaiheessa. Johnsonin (1998) tutkimuksessa 80 % osallistujista käytti tarkastuksia epäsäännöllisesti tai ei ollenkaan. Vaikuttaa siis siltä, että useimmilla organisaatioilla on edelleen suuri työ päästä tarkastusten järjestämisessä säännöllisyyteen.

Suorittavalla tasolla organisaatiolta vaaditaan seuraavia käytänteitä:

- Vaatimusmäärittelyn ja tärkeimpien suunnitteludokumenttien tarkastukset
- Tarkastukseen sisältyvä valmistautuminen ja raportointi
- Koulutusta vähintään tarkastusten johtajille

Vaatimusmäärittely ja tärkeimmät dokumentit on tarkastettava jokaisessa projektissa. On hyvin tunnettua, että virheiden löytäminen ja korjaaminen maksaa sitä enemmän, mitä myöhemmässä vaiheessa ne löydetään. Klassisen Boehmin (1981) tutkimuksessa todettiin, että testausvaiheessa löytynyt vaatimukseen liittyvä virhe maksoi kymmeniä kertoja enemmän korjata kuin vaatimusten määrittelyvaiheessa. Alkuvaiheen dokumentaatio on siksi kaikkein tärkeintä tarkastaa ja siitä organisaation tulisi aloittaa tarkastusten käyttöönotto organisaatiossa.

Se, mitä tarkoittaa tärkeimmät suunnitteludokumentit, riippuu projektista ja organisaatiosta. Tärkeimpiin suunnitteludokumentteihin kuuluu tavallisesti ainakin arkkitehtuurisuunnitelmat ja joitakin muita suunnitteludokumentteja. Tällä tasolla ei ole kuitenkaan mielekäästä vaatia jokaisen pienen komponentin suunnitelmien tarkastamista.

Tarkastukseen täytyy sisältyä valmistautuminen ja raportointi. Tällä tasolla tarkastusprosessin ei tarvitse vielä olla dokumentoitu, eikä sen tarvitse tiukasti noudattaa formaalia, esimerkiksi Faganin (1976) mukaista, tarkastusprosessia. Kuitenkin tarkastusten täytyy tärkeimmiltä osiltaan muistuttaa tarkastusta ja siksi toimintaan täytyy sisältyä vähintään valmistautuminen ja raportointi. Esimerkiksi vapaamuotoinen läpikäynti ei täytä näitä kriteereitä. Kuitenkin esimerkiksi sähköpostin välityksellä organisoitu tarkastus täyttää kriteerit, jos siitä tehdään raportti.

On yleisesti tunnustettua, että valmistautuminen on tarkastuksen tärkeimpiä asioita. Tarkastuksen tehokkuuden kannalta on ensisijaisen tärkeää, että osallistujat ovat hyvin valmistautuneita. Gilbin ja Grahamin (1993) arvio on, että optimaalisessa tilanteessa tulisi 80 % virheistä löytyä valmistautumisvaiheessa ja loput 20 % tarkastuspalaverissa. He toteavat myös, että tarkastusmenetelmän tehokkuus epämuodollisiin katselmointeihin verrattuna tulee esiin vasta silloin, kun tarkastajat käyttävät valmistautumiseen riittävästi aikaa. Muussa tapauksessa löydetään vain helposti löydettäviä virheitä, jotka löydetäisiin todennäköisesti epämuodollisessa katselmoinnissakin.

Tällä tasolla raportoinnille ei ole tiukkoja vaatimuksia. Kuitenkin jokaisesta tarkastuksesta täytyy tehdä raportti, joka sisältää tiedot vähintään tarkastetusta dokumentista, osallistujista ja tarkastuksen ajasta. Myös lista tarkastuksessa löydetyistä virheistä tulee sisältyä raporttiin. Pääasia on, että raportoinnin perusteella voi seurata tarkastusten toteutumista.

Tarkastusten johtajat täytyy kouluttaa. Tämä on myös eräs tärkeimmistä asioista tarkastuksessa. Tarkastuksen johtajalla on vastuu tarkastusprosessista, jo-

ten vähintään hänellä täytyy olla hyvä käsitys tarkastuksen merkityksestä tuotantoprosessissa. Lisäksi hänellä täytyy olla tietoa keskeisistä tarkastuksen tehokkaaseen toteuttamiseen vaikuttavista tekijöistä, kuten siitä, miten tarkastuspalaveri tulisi organisoida..

3.2.2 Taso 3 – Määritelty

Määritelty taso vaati organisaatiolta hyvin määriteltyä prosessia tarkastuksia varten. Määritellyn prosessin täytyy pitää sisällään tarkastuksen tehokkuuden kannalta keskeisimmät elementit. Tämä taso ottaa huomioon myös CMMI-mallin siten, että tällä tasolla oleva organisaatio täyttää samalla CMMI:n katselmointeihin liittyvät vaatimukset. CMMI-mallissa nämä vaatimukset ovat myös kolmannella tasolla verifiointia ja validointia käsittelevillä prosessialueilla.

Määritellyllä tasolla vaaditut prosessialueet ovat:

- Koodin ja testitapausten katselmointi
- Määritelty tarkastusprosessi
- Koulutus kaikille osapuolille
- Mukautettu tukimateriaali
- Datat kerääminen ja käyttö
- Organisaation sitoutuminen

Koodi ja testitapaukset täytyy tarkastaa jokaisessa projektissa. Ei ole välttämättä tarpeen tarkastaa muodollisesti jokaista koodinpätkää, mutta joka tapauksessa tärkeimmät koodin osat ja testitapaukset täytyy tarkastaa. Lisäksi tarkastukset tulee sisällyttää projektisuunnitelmaan. Jokaisen projektin täytyy miettiä näiden vähimmäisvaatimusten lisäksi, mitä dokumentteja sen on tarpeen tarkastaa muodollisesti. Organisaatiolla voi myös olla määriteltynä yleinen politiikka, joka määrää, mitkä dokumentit projektissa täytyy tarkastaa.

Tarkastusprosessi täytyy olla määritelty ja dokumentoitu. Määritellyn prosessin täytyy suunnitella seurata alkuperäisen tarkastusprosessin (Fagan 1976 tai Gilb ja Graham 1993) vaiheita. Prosessilta vaaditut vaiheet ovat suunnittelu, valmistautuminen, kokoaminen ja seuranta. Lisäksi prosessikuvauksen tulisi sisältää alku- ja loppukriteerit erilaisille tuotoksille. Nämä mainitut elementit ovat erittäin tärkeitä tarkastuksen tehokkuuden kannalta. Esimerkiksi Gilb ja Graham (1993) ovat esittäneet yksityiskohtaisia perusteluja eri vaiheiden merkityksestä tarkastusprosessissa.

Suunnittelu sisältää tarkastuksen aikataulutuksen, tiimin valinnan ja roolien määrittämisen osallistujille. Suunnitteluvaiheessa tulisi myös määrittellä erityiset aloitus- ja lopetuskriteerit tarkastukselle. Lisäksi tarkastuksen tukimateriaali, kuten tarkistuslistat, tulisi suunnitteluvaiheessa tarpeen mukaan muokata tarkastukseen soveltuviksi.

Jonkinlainen valmistautuminen tarkastusprosessissa vaadittiin jo edellisellä tasolla, mutta siihen liittyen ei vaadittu dokumentoituja ohjeita. Tällä tasolla

prosessikuvauksen tulee pitää sisällään ohjeet valmistautumiseen käytettävästä ajasta. On tärkeää, että osallistujat ovat hyvin valmistautuneita ennen kokoamisvaihetta, mutta on myös tärkeää, että resursseja ei käytetä liikaa.

Perinteinen tarkastusprosessi (Gilbin ja Grahamin 1993 mukaan) pitää sisällään palaverin, jonka ensisijainen tarkoitus on koota osallistujilta virheet, jotka on löydetty valmistautumisen aikana. Kirjallisuudessa on esitetty jonkin verran kritiikkiä tämän kaltaisen palaverin tarpeellisuudesta (Votta 1993; Johnson ja Tjahjono 1998). Siksi palaveri ei ole ehdoton vaatimus tarkastusprosessissa, vaan virheiden kokoamiseen voidaan käyttää myös muita keinoja. Kuitenkin määritellyn tavan olisi hyvä kannustaa osallistujia jonkinlaiseen vuorovaikutukseen. Esimerkiksi jonkinlainen virtuaalinen tarkastus (Harjumaa ym. 2001) on eräs vaihtoehto palaverin järjestämiselle.

Prosessikuvauksen täytyy kertoa, kuinka korjausten seuranta on suoritettava. Määritys voi sisältää useampia toimintamalleja erilaisia tilanteita varten. Esimerkiksi hyvin kriittisen dokumentin kohdalla voidaan järjestää uusi tarkastus korjattua dokumenttia varten. Toisaalta joissakin tilanteissa vastuu korjausten seurannasta voidaan antaa tarkastuksen johtajalle ja joskus seurantaa ei välttämättä vaadita ollenkaan.

Kaikille tarkastuksiin liittyville osapuolille on järjestettävä vähintään johdantotasoinen koulutus. Faganin (1986) mukaan sekä esimiehet, tarkastusten johtajat että muut tarkastuksiin osallistuvat tulisi kouluttaa ennen tarkastusten käyttöönottoa organisaatiossa. Hän suorastaan varoittaa artikkelissaan että, tarkastusten aloittaminen ilman koulutusta saattaa johtaa menetelmän tehokkuuden menettämiseen ja turhautumiseen. Hän kertoo myös, että hänellä on käytännön kokemusta tämänkaltaisista yrityksistä.

Tarkastusten johtajien kouluttaminen vaadittiin jo edellisellä tasolla. Tällä tasolla vaaditaan koulutuksen järjestämistä kaikille tarkastuksiin liittyville osapuolille. Esimiesten tulee ymmärtää vähintään perusteet tarkastuksista ja ennen kaikkea tarkastusten hyödyt heidän näkökulmastaan katsottuna. Tarkastuksiin osallistujien tulee ymmärtää tarkastusten merkitys sekä heidän oma roolinsa tarkastusprosessissa.

Organisaatiolla tulee olla käytössä sen toimintaan mukautettu tarkastusta tukeva materiaali. Organisaatiolla tulisi olla sekä dokumenttien kirjoittamista ohjeistavat standardit että ohjeet, jotka ohjaavat osallistujien toimintaa tarkastuksissa. Eri rooleille voi olla tarkastuksessa erilaiset ohjeet. Lisäksi organisaatiolla tulee olla käytössään jonkinlainen tukimateriaali, kuten tarkistuslistoja (Gilb ja Graham 1993) tai skenaarioita (Basili ym. 1996), joka ohjeistaa tarkastajia virheiden löytämisessä.

Tarkistuslistoja on kritisoitu siitä, että usein niitä ei ole riittävästi mukautettu ympäristöön, jossa niitä käytetään (Laitenberger ja DeBaud 2000). Toinen kritiikin kohde tavallisessa tarkistuslistaa käyttävässä tarkastuksessa on tarkastajien tekemän työn päällekkäisyys silloin, kun kaikki tarkastajat käyvät läpi saman pitkän tarkistuslistan (Porter ja Votta 1998). Materiaalin tulisi siis olla

huolellisesti mukautettu sitä käyttävään organisaatioon ja lisäksi erilaisen tukimateriaalin luominen eri rooleille on hyvin suositeltavaa.

Tarkastuksista täytyy kerätä tietoa ja kerättyä tietoa täytyy sekä analysoida, että käyttää. Kerättävällä tiedolla tarkoitetaan tässä tarkastuksista tehtäviä raportteja, jotka tavallisesti sisältävät ainakin tiedot tarkastetusta dokumentista, projektista, osallistujista ja löytyneiden virheiden määrän. Löytyneet virheet tulisi analysoida ja luokitella esimerkiksi karkeasti merkittäviin (*major*) ja pieniin (*minor*) virheisiin (katso Fagan 1976 tai Gilb ja Graham 1993). Tätä kerättyä tietoa tulisi käyttää tarkastusten todellisen toteutumisen tarkkailuun ja valvontaan. Tätä tietoa tulisi hyödyntää myös tarkastusprosessin kehittämiseen.

Organisaatiossa täytyy olla merkkejä organisaatiotasoisesta sitoutumisesta tarkastuksiin. Tässä vaadituilla näkyvillä merkeillä tarkoitetaan resurssien varaamista, organisaation politiikkaa ja vastuuhenkilöiden varaamista. On erittäin tärkeää, että tarkastuksille on varattu riittävästi resursseja. Resursseja tarvitaan tarkastusten järjestämisen lisäksi niiden kehittämiseen. Resurssien varaaminen tulisi olla osa projektin suunnittelua. Jos tarkastuksille ei ole selkeästi varattu aikaa työsuunnitelmiin, työntekijöillä ei saata olla aikaa niihin osallistumiseen tai ainakaan valmistautumiseen. Tarkastusten kehittämisen kohdalla on vielä todennäköisempää, että se jää kokonaan tekemättä, ellei siihen ole erikseen varattu resursseja.

CMMI (SEI 2002) mainitsee jokaisella prosessialueella organisaation politiikan ja vastuuhenkilöiden asettamisen osana prosessien käyttöönottoa. Ne ovat tärkeitä viestejä johdolta aihealueen merkityksestä organisaation toiminnalle. Karl Wiegertsin (2002) mukaan organisaatiolla tulisi lisäksi olla kirjalliset odotukset ja tavoitteet. Nämä ovat näkyviä merkkejä johdon sitoutumisesta toimintaan.

3.2.3 Taso 4 – Hallittu

Halittu taso keskittyy tarkastusprosessin kehittämiseen. Tällä tasolla oleva organisaatio kehittää jatkuvasti tarkastusprosessiaan tehokkaammaksi löytämään virheitä ja käyttää samalla tarkastuksia välineenä koko ohjelmistotuotantoprosessin kehittämisessä. Tämä taso ottaa myös mukaan näkökulman virheiden ennalta ehkäisystä pelkän jo syntyneiden virheiden löytämisen ja korjaamisen sijasta.

Hallitulla tasolla vaaditut prosessialueet ovat:

- Jatkuva tarkastusprosessin kehittäminen
- Virheitä ennalta ehkäisevä toiminta
- Järjestelmällinen koulutusohjelma
- Ohjelmistotuotantoprosessin kehittäminen

Jatkuvaan tarkastusprosessin kehittämiseen liittyvät toimenpiteet täytyy olla määriteltynä, dokumentoituna ja oikeasti käytössä. Tämä tarkastusprosessin kehittäminen pitää sisällään tarkastuskäytänteiden, tukimateriaalin ja koulutuksen kehittämisen. Organisaatiossa täytyy olla määriteltynä, milloin ja mitä kehittämistoimenpiteitä on tehtävä. Kehittämistoiminnalle täytyy olla myös asetettuna vastuuhenkilöt.

Porter ym. (1998) totesivat tutkimuksessaan, että tarkastusprosessiin tehdyillä muutoksilla ei vaikuttanut olevan merkittävää vaikutusta virheiden löytymisen tehokkuuteen. Kuitenkin joillakin prosessin muutoksilla oli huomattavaa vaikutusta tarkastusten aikatauluihin. Siksi tarkastusprosessiin keskittyvässä kehittämisessä tulisi keskittyä ehkäisemään aikataulutukseen liittyviä viiveitä.

Tarkastusta tukevan materiaalin kehittäminen tulisi kohdistaa parantamaan tarkastajan suorituskykyä virheiden etsimisessä. Porter ja Votta (1998) tutkivat erilaisia menetelmiä virheiden löytämiseen ja totesivat tutkimuksessaan skenaariot tavallisia tarkistuslistoja paremmaksi työkaluksi virheiden löytämisessä. Tutkimuksen mukaan on siis selvää, että organisaatio voi parantaa virheiden löytymistä tarkastuksissa kehittämällä tukimateriaalia. Tukimateriaalin kehittäminen tulisi perustua organisaation kokemukseen ja tarkastuksista kerättyyn tietoon. Myös tarkastuskoulutusta tulisi kehittää keskittyen tarkastajien henkilökohtaisen suorituskyvyn parantamiseen.

Virheiden ennalta ehkäisevät käytänteet täytyy olla määriteltynä ja käytössä. Sekä Fagan (1986) että Gilb (2000) ovat yhtä mieltä siitä, että tarkastuksia ei tulisi käyttää ainoastaan virheiden etsimiseen, vaan myös niiden ennaltaehkäisyyn. Se vaatii systemaattista kokemusten ja tarkastuksista kerätyn tiedon käyttämistä kehitystyön tukemisessa. Tämä toiminta on yhteydessä jatkuvaan tarkastusten kehittämiseen ja samoja käytänteitä voidaan hyödyntää virheiden ennaltaehkäisyä tukevan materiaalin kehittämisessä.

Organisaatiossa voi olla työryhmä, joka säännöllisesti kokoontuu kehittämään materiaalia kertyneiden kokemusten ja tarkastuksista kerätyn tiedon pohjalta. Käytännössä virheitä ennalta ehkäisevä materiaali voi tarkoittaa työhohjeita, sääntöjä ja listoja tyypillisimmistä virheistä eri työvaiheissa. Tämä tietämys tulisi sisällyttää myös uusien työntekijöiden kouluttamiseen.

Organisaatiossa täytyy olla systemaattinen koulutusohjelma. Tämä tarkoittaa käytännössä sitä, että organisaatiolla on suunnitelma siitä, millaista tarkastuksiin liittyvää koulutusta ja milloin tarjotaan eri henkilöstöryhmille. On suositeltavaa, että järjestetään perehdyttämiskoulutuksen lisäksi jonkinlaista vuosittaista koulutusta. Johtotasolle, tarkastusten johtajille ja muille työntekijöille tulisi olla määriteltynä erilaiset koulutusohjelmat.

Systemaattinen koulutus on tärkeää niin johdon sitoutumisen kuin tarkastuksen tehokkuudenkin kannalta. Rifkin ja Deimel (1994) havaitsivat tutkimuksessaan, että tarkastajan kyky lukea dokumentaatiota on tärkein yksittäinen tekijä tarkastusten tehokkuudessa. He tekivät tämän johtopäätöksen vertaillen erilaisia tarkastuskoulutuksia eräässä kohdeyrityksessä. Dokumentin lu-

kemisen kehittämiseen keskittynyt koulutus paransi merkittävästi tarkastuksissa löydettyjen virheiden määrää, kun taas tarkastukseen prosessitasolla keskittynyt koulutus kontrolliryhmässä ei tuottanut samanlaista tulosta. Siis tästä voidaan päätellä, että hyvin suunnitellun koulutuksen avulla voidaan parantaa tarkastajien kykyä löytää virheitä.

Tarkastuksia täytyy hyödyntää koko ohjelmistotuotantoprosessin kehittämisessä. Tarkastukset voivat tarjota arvokasta tietoa koko ohjelmistotuotantoprosessin kehittämistä varten. Tarkastuksista kerättävää tietoa analysoimalla voidaan saada tuotettua hyödyllisiä kehittämis ehdotuksia. Lisäksi Gilb ja Graham (1993) ehdottavat, että tarkastuspalaverin lopuksi tulisi varata aikaa tuotantoprosessiin liittyviä kehittämis ehdotuksia varten.

3.2.4 Taso 5 – Optimoiva

Keskeinen asia optimoivalla tasolla on kustannustehokkuus. Tällä tasolla organisaatiolla on hyvät tarkastuskäytänteet, jotka on kehitetty vahvasti kokemukseen perustuen. Organisaatio on määritellyt prosessin tarkastusten kehittämistoimintaa varten sekä prosessimuutosten käyttöönottoa varten. Tällä tasolla organisaatiolla on myös kokemukseen perustuvia standardeja sille, kuinka tarkastuksia tulee järjestää erilaisissa olosuhteissa.

Optimoivalla tasolla vaaditut prosessialueet ovat:

- Standardit projektin ja tarkastusten suunnitteluun
- Määritelty oppimisprosessi

Organisaatiolla tulee olla standardit tarkastusten suunnitteluun. Erilaiset projektit voivat muodostaa tarkastusten kannalta erilaiset olosuhteet. Tässä standardilla tarkoitetaan sen määrittämistä, miten tarkastuksia tulee järjestää missäkin olosuhteissa. Standardit voivat pitää sisällään erilaisia tarkastuksen toteutustapoja ja suosituksia niiden käytöstä erilaisissa olosuhteissa. Nämä standardit tulisi luoda historiatietoon perustuen ja niiden lähtökohtana tulee olla kustannustehokkuus.

Gilb ja Graham (1993) esittävät, että tavallisesti tarkastukset vievät noin 10-15 prosenttia ohjelmistokehityksen kokonaisbudjetista. Se on niin huomattava summa, että huolimatta tarkastuksen selkeistä hyödyistä, rahan käyttö täytyy harkita huolellisesti. Monissa organisaatioissa on hyvin erilaisia projekteja, joilla on myös erilaiset laatuvaatimukset. Tällä tasolla organisaation tulee olla tietoinen tarkastuksen kustannuksista erityyppisissä projekteissa. Sen pohjalta voidaan luoda esimerkiksi erilaisen laatutason täyttävät tarkastusohjelmat erilaisille projekteille.

Organisaatiolla tulee olla dokumentoitu oppimisprosessi. Oppimisprosessilla tarkoitetaan tässä sitä, kuinka prosessin kehittäminen toteutetaan käytännössä ja saadaan muutokset tehokkaasti käyttöön. Jo edellisellä tasolla vaadittiin jatkuvaa prosessien kehittämistä ja siihen liittyvien toimenpiteiden määrittelyä.

Tällä tasolla vaaditaan lisäksi, että organisaatio toteuttaa suunnitelmallisesti muutosten käyttöönnoton. Nämä määritykset pitävät sisällään sen, kuinka muutokset dokumentoidaan ja kuinka työntekijät koulutetaan uusiin käytänteisiin. Tällä oppimisprosessilla tulee olla vastuhenkilö, joka vastaa muutosten toteutuksesta määritysten mukaisesti.

3.3 Yhteenveto kypsyytason mittaamisesta

Tässä luvussa on käsitelty tarkastuskäytänteiden kypsyytason mittaamista. Luvun aluksi esiteltiin lyhyesti aiempaa aihetta käsittelevää kirjallisuutta, jota on olemassa hyvin niukasti. Siinä yhteydessä todettiin, ettei niissä ollut tämän tutkimuksen tarkoitukseen sopivaa mallia. Siitä syystä tämä luku keskittyi pääosin esittelemään tutkimusta varten luotua ICMM-mallia. Se on malli, jonka esikuvan on käytetty koko ohjelmistotuotantoprosessin kypsyyttä mittaavaa CMMI (SEI 2002) mallia. Mallin tarkoitus on mahdollistaa organisaation tarkastuskäytänteiden kypsyyden arviointi sekä toimia apuvälineenä tarkastuskäytänteiden kehittämisessä.

ICMM-malli, joka tässä luvussa on esitelty, luotiin tutkimuksen pohjaksi puhtaasti kirjallisuudessa esitettyjen tutkimustulosten pohjalta. Se on siis eräänlainen ensimmäinen kokeiluversio mallista ja siksi siihen ei pidä tässä vaiheessa asennoitua valmiina mittarina. Tässä tutkimuksessa sen rooli on toimia ennemminkin viitekehystenä, jonka pohjalta eri kohdeorganisaatioiden käytänteitä analysoidaan ja verrataan toisiinsa.

Yleisesti tämänkaltainen kypsyytason mittaaminen on ongelmallista, koska siinä ei kyetä arvioimaan määriteltujen käytänteiden tehokkuutta. Kypsyytasonmallin antama ohjeistus on niin korkealla tasolla, että organisaatio voi määritellä ja käyttää huonoimmassa tapauksessa todella huonosti toimivia käytänteitä ja silti täyttää kaikki vaikkapa mallin kolmannella tasolla vaaditut kriteerit.

Eräs ICMM-mallin ongelmakohtista on neljännen ja viidennen tason kuvausten arviointi. Tällainen malli pitää sisällään oletuksen siitä, että organisaation tarkastustoiminnan tulisi olla sitä tehokkaampaa, mitä korkeammalla tasolla organisaatio on. Kolmanteen tasoon asti vaadittujen käytänteiden vaikutuksesta tarkastusten tehokkuuteen löytyy jonkinlaisia viitteitä tarkastusmenetelmää käsittelevistä tutkimuksista, mutta ylemmille tasoille ei tällaista kirjallisuuden tukea ole vielä olemassa. Neljännen ja viidennen tason määrittelyt perustuvat oletukseen, että käytänteet, jotka CMMI (SEI 2002) vaatii yleisemmin vastaavilla tasoilla, toimivat myös tarkastusten kontekstissa.

ICMM-malli toimii siis tällaisenaan hyvin tässä tutkimuksessa, jossa sitä käytetään vain viitekehystenä, joka pitää sisällään keskeiset tarkastuksiin liittyvät käytänteet. Kuitenkin sen käyttö konkreettisenä työvälineenä kypsyyden mittaamisessa tai tarkastuskäytänteiden kehittämisessä vaatii vielä merkittävästi jatkokehitystä ja empiiristä arviointia.

4 EMPIIRISEN TUTKIMUKSEN ESITTELY

Merkittävä osa tästä tutkielmasta perustuu tapaustutkimukseen, jossa tarkastellaan kuutta suomalaista ohjelmistoja tuottavaa organisaatiota haastattelujen avulla. Empiirisellä tutkimuksella oli kolme erillistä tavoitetta:

- Selvittää, kuinka tarkastuksia ja vapaamuotoisempia katselmointeja toteutetaan suomalaisissa ohjelmistoja tuottavissa organisaatioissa
- Tunnistaa merkittävimmät puutteet tarkastuskäytännöissä
- Tunnistaa keskeisimmät katselmointien toteuttamiseen liittyvät ongelmat

Tässä luvussa kuvaillaan, kuinka empiirinen tutkimus toteutettiin. Varsinaiset empiiriset tutkimustulokset esitellään myöhemmissä luvuissa.

4.1 Tutkimukseen osallistuneet organisaatiot

Tutkimukseen osallistui kuusi suomalaista ohjelmistoja tuottavaa organisaatiota. Mukaan valikoitui sellaisia organisaatioita, joilla oli ennestään kontakteja korkeakoulumaailmaan. Tämä tausta saattaa tarkoittaa sitä, että mukana olleet organisaatiot ovat poikkeuksellisen kiinnostuneita toimintansa kehittämisestä, eivätkä ehkä edusta aivan keskimääräistä organisaatiota ohjelmistoalalla. Kuitenkaan tarkoitus ei ollut valita ainoastaan ennako-oletusten mukaan parhaita tai suurimpia organisaatioita, vaan organisaatiot aineistossa ovat hyvinkin erityyppisiä. Ainoastaan aivan pienet organisaatiot jätettiin tietoisesti tutkimuksen ulkopuolelle.

Aineiston kuusi organisaatiota ovat viidestä eri ohjelmistoyrityksestä. Suurempien yritysten kohdalla käsite organisaatio viittaa siis tarkastelun kohteena olleeseen yksikköön, ei koko yritykseen. Yhdestä yrityksestä otettiin vertailun vuoksi mukaan kaksi erillistä yksikköä. Yritysten koko vaihtelee alle sa-

dasta tuhansiin työntekijöihin. Kuusi organisaatiota jakaantuu yrityksen koon mukaan seuraavasti:

- Kaksi organisaatiota alle 100 työntekijän yrityksistä
- Kaksi muuta organisaatiota alle 500 työntekijän yrityksistä
- Loput kaksi organisaatiota ovat yli 500 työntekijän yrityksistä

Tutkimuksen kohdeorganisaatioista neljä tekee pääasiassa tyypillistä projekti-liiketoimintaa ulkopuolisille asiakkaille. Yksi organisaatio tekee yrityksen omaa tuotekehitystä. Yksi organisaatio on puolestaan suuremman teollisuusyrityksen ohjelmistoyksikkö, jonka tuotokset ovat tavallisesti yrityksen sisäisiä projekteja ja päätyvät osaksi laajempaa järjestelmää.

Eräs tavoite organisaatioiden valinnassa oli saada otokseen ”tavallisia” ohjelmistoja tuottavia organisaatioita. Olemassa olevissa tutkimuksissa on yleensäkin melko vähän ymmärrystä tarkastusten käytännön soveltamiseen liittyvistä kysymyksistä, mutta lisäksi tutkimusta on tehty lähinnä jättikokoisten yritysten (HP, IBM, DaimlerChrysler) näkökulmasta. Tällaiset suuret yritykset ovat tehneet yleensä vuosikymmeniä määrätietoista prosessien kehittämistä ja ovat lähtökohdiltaan erilaisessa tilanteessa kuin alan keskiverto yritykset. Tässä tutkimuksessa tarkastuskäytänteiden puutteita ja ongelmia tutkittaessa on tarkoitus saada näkökulma ensisijaisesti tavallisesta alan organisaatiosta, ei alan ”huippuyksiköistä”.

4.2 Tutkimuksen toteuttaminen

Tutkimusaineisto kerättiin haastattelujen avulla. Jokaisessa organisaatiossa toteutettiin kolme haastattelua eli haastateltavia oli yhteensä 18. Ensisijaisesti haastateltaviksi pyydettiin edustajia seuraavista rooleista: laatupäällikkö, projektipäällikkö ja joku suoritustason työntekijä. Useimmissa organisaatioissa pyyntöä vastaavat roolien edustajat löytyivät ongelmitta. Joissakin organisaatioissa ei ollut varsinaista laatupäällikköä, mutta vastaavanlainen rooli kyllä löytyi. Suoritustason työntekijä oli yhdessä tapauksessa testaaja ja kaikissa muissa tapauksissa ohjelmistosuunnittelija, joka tekee käytännössä sekä ohjelmistojen teknistä suunnittelua että toteutusta.

Tutkimuksessa käytettiin termiä katselmointi, sillä se vaikutti olevan tarkastusta tutumpi käsite. Tutkimuksen aluksi haastateltavia pyydettiin määrittelemään omin sanoin, mitä katselmointi tarkoittaa. Sitten määritelmästä keskusteltiin ja lopputuloksena katselmointi määriteltiin yhdessä, jotta sillä tarkoitetaan haastattelussa samaa asiaa. Keskeiset asiat katselmoinnin määrittelyssä olivat, että se on virheiden etsimistä jostain dokumentista ja sen tekijä on joku muu kuin tuotoksen tekijä. Myöhemmin varsinaisissa tarkastusten toteuttamisesta kartoittavissa kysymyksissä tarkennettiin joitakin tarkastukseen kuuluvia kriteerejä.

Katselmointikäsitteen yhteisen määrittelyn jälkeen jokaista haastateltavaa pyydettiin arvioimaan asteikolla 1-5 (5 on paras), kuinka hyvin katselmoinnit toteutuvat hänen omassa organisaatiossaan. Tällaisen subjektiivisen arvion tarkoitus oli testata haastateltavan ennakkoasennetta ennen kuin haastattelun aikana käyty keskustelu mahdollisesti johdattelisi ajatuksia.

Haastattelun varsinainen sisältö jakaantui kahteen eri osaan tutkimuksen tutkimuskysymysten mukaisesti. Jokaisessa haastattelussa samat kysymykset kysyttiin samassa järjestyksessä, jotta vastaukset olisivat mahdollisimman hyvin vertailukelpoisia keskenään. Haastattelun runko kysymyksineen on esitetty liitteessä 2. Liitteessä on esitetty ne kysymykset, jotka lopulta vakiintuivat haastatteluihin. Alun perin kysymyksiä oli enemmän esimerkiksi ICM-mallin ylempiin tasoihin liittyen, mutta ne karsiutuivat käytännössä haastatteluista pois, koska ne olivat hyvin etäisiä kohdeorganisaatioiden käytännön näkökulmasta.

Ensimmäisen osan tarkoitus oli kartoittaa tarkastuskäytänteiden toteutumista organisaatioissa. Sen kysymykset oli laadittu luvussa 3 esiteltävän ICM-mallin pohjalta ja tarkoitus oli selvittää, kuinka mallin sisältämiä käytänteitä toteutetaan. Osa tämän osan kysymyksistä oli tiukasti strukturoituja ja osa oli avoimia kysymyksiä. Tarkoitus ei ollut ainoastaan selvittää, täyttyykö mallin asettama kriteeri, vaan selvittää myös, millä tavalla käytänteitä toteutetaan organisaatiossa käytännössä. Esimerkiksi tarkastusten raportoinnista kysyttäessä selvitettiin mitä ja millä tavalla raportoidaan, eikä ainoastaan sitä tehdäänkö raportointia. Tarkastusten toteuttamisen ohella selvitettiin, kuinka usein niiden lisäksi tehdään vapaamuotoisempia katselmoiteja. Lisäksi tässä yhteydessä kysyttiin projektiliiketoimintaan painottuvilta organisaatioilta, kuinka aktiivisesti asiakkaat osallistuvat tarkastuksiin.

Haastattelun toisessa osassa tutkittiin tarkastusten toteuttamiseen liittyviä ongelmia. Sen osan alussa haastateltavilta kysyttiin avoimella kysymyksellä, mitkä ovat heidän mielestään keskeisimmät katselmoiteihin liittyvät ongelmat heidän organisaatiossaan. Sen jälkeen heille esitettiin luvussa 6 käsiteltäviä mahdollisia tarkastuksiin liittyviä ongelmia. Jokaisen esitetyn ongelman kohdalla heiltä kysyttiin, kuinka tutulta se tuntuu omassa organisaatiossa. Vastaus pyydettiin asteikolla 1-5 (5 tarkoittaa erittäin tuttua).

4.3 Huomioita valituista menetelmistä

Katselmoiteihin liittyvä yleinen tieto vaikutti olevan organisaatioissa niin rajallista, että käsitteet aiheuttivat jonkin verran hankaluuksia. Haastateltavat tunsivat toimintatavat ja käsitteet luonnollisesti omassa organisaatiossa, mutta ne vaikuttivat olevan monessa tapauksessa vaihtelevia organisaatioiden välillä. Tästä näkökulmasta haastattelu oli onnistunut valinta tutkimusmenetelmäksi, sillä haastattelussa tutkijalla on mahdollisuus varmistaa kysymysten yhdenmukainen tulkinta.

Tutkimuksen alkuvaiheessa harkittiin oheismenetelmänä kyselyä, jonka avulla olisi valituista organisaatioista saatu suurempi otos vastauksia. Edellä mainittujen käsiteongelmien vuoksi kysely vaikutti kuitenkin mahdottomalta toteuttaa luotettavasti tässä tapauksessa. Tulevaisuudessa kysely voisi ehkä sopia menetelmäksi, jos sen avulla tutkittaisiin jotain paljon rajoitetumpaa ilmiötä.

Tarkastusten ongelmia kartoittavat kysymykset olivat useissa kohdissa hankalia tulkita yhdenmukaisesti niin, että eri määrällisiä vastauksia voitaisiin luotettavasti verrata toisiinsa. Vastaukseen vaikutti sekä esitetyn ongelman esiintymistäajuus että kokemus sen merkittävydestä. Tavallinen tilanne oli, että haastateltavan mielestä ongelma toteutuu aika usein, mutta ei aiheuta toteutuessaan kovinkaan merkittäviä ongelmia. Haastatteluissa pyrittiin vastauksena saamaan näiden tekijöiden keskiarvo. Jos esimerkiksi esiintymistäajuuden puolesta vastaus olisi asteikolla (1-5) neljä ja merkittävyyden puolesta kaksi, merkittiin lopulliseksi vastaukseksi kolme. Hyvä puoli tässä asiassa oli tulkin-takysymyksen aiheuttama keskustelu, joka luonnollisella tavalla tuotti taustatekijöitä kuvailevaa laadullista aineistoa.

Haastattelujen alkuosassa käytiin tavallisesti melko runsasta keskustelua tarkastuskäytänteiden toteutumiseen liittyvistä asioista. Suureksi osaksi keskustelu seurasi siitä, että haastateltavien ymmärrys kysymyksissä käytettävistä käsitteistä ei ollut yhdenmukaista ja niiden taustaa täytyi selittää haastattelun kuluessa. Tällä haastattelun alkuosan keskustelulla saattoi olla jonkin verran johdattelevaa vaikutusta tarkastusten ongelmia käsitteleviin kysymyksiin. Vaikutus on tuskin kuitenkaan ollut kovin merkittävää tavallisimmiksi tunnistettujen ongelmien kohdalla. Esimerkiksi valmistautumiseen liittyviin ongelmiin monet viittasivat jo aivan haastattelun alussa.

5 KUINKA KATSELMOINNIT TOTEUTUVAT KÄYTÄNNÖSSÄ

Tässä luvussa käsitellään sitä, kuinka tarkastukset toimivat käytännön ohjelmistotuotannossa ja minkälaisia puutteita olemassa olevat käytänteet pitävät sisällään. Tarkastusten toteutusta käytännössä tarkastelevia tutkimuksia on raportoitu todella vähän. Lisäksi suurin osa olemassa olevista raporteista keskittyy kuvaamaan tarkastusten onnistumista jossain suuressa ja edistyneessä organisaatiossa. Tarkastusten tutkimuksesta vaikuttaa lähes kokonaan puuttuvan ”tavallisen yrityksen” näkökulma. Sitä yritetään tuoda esiin tässä luvussa pääasiassa kuuden kohdeorganisaation kokemusten kautta.

Seuraavaksi kohdassa 6.1 on lyhyesti esitelty aiemmin raportoituja tutkimuksia, joissa on tutkittu tarkastusten toteutumista ohjelmistoalalla. Sen jälkeen kohdassa 6.2 esitellään tarkastusten toteutumista kuudessa kohdeorganisaatiossa. Kohdassa 6.3 analysoidaan saatuja tutkimustuloksia ja lopuksi kohdassa 6.4 vedetään yhteen johtopäätöksiä tämän luvun sisältämistä asioista.

5.1 Muita tutkimuksia

Yleisesti tarkastusten toteuttamista ohjelmistoalalla tarkastelevia luotettavia tutkimuksia ei ole tehty. Tässä kohdassa esitellään lyhyesti kaksi kyselytutkimusta (Ciolkowski ym. 2003; Johnson 1998), joissa on jollakin tavalla aihealuetta kartoitettu.

Ciolkowski ym. (2003) tekivät vuonna 2002 kyselytutkimuksen katselmointien toteutumisesta ohjelmistoalalla. He kutsuivat henkilökohtaisten suhteidensa kautta ihmisiä vastaamaan kyselyyn. Vastauspyyntöjä lähetettiin eri maihin 865 kappaletta ja vastauksia 226. Tutkimus oli kansainvälinen, vaikkakin tutkimuksessa luonnollisesti korostui tutkijoiden oma maantieteellinen sijainti. Puolet vastauksista tuli Saksasta ja lisäksi 33 % muualta Euroopasta. 12 % vastauksista saatiin Pohjois-Amerikasta ja loput 5 % muualta maailmasta.

Ciolkowskin ym. (2003) tutkimuksen ensisijainen tarkoitus oli kartoittaa, millaisia katselmointimenetelmiä teollisuudessa käytetään, ei yrittää selvittää sitä, kuinka suuri osa yrityksistä toteuttaa katselmointeja. He itsekin toteavat, että tutkimuksen otos ei olisi kovin luotettava määrien arvioimiseen. Vastausprosentti oli aika alhainen ja tutkimukseen osallistuminen oli yrityksille vapaaehtoista. On siis oletettavaa, että tutkimukseen osallistuivat ne, jotka ennestään toteuttivat katselmointeja tai joilla oli muuten erityistä kiinnostusta laadunvarmistukseen. Lisäksi tutkimuksen kutsutut henkilöt valittiin henkilökohtaisten suhteiden avulla. On oletettavaa, että tutkijoiden henkilökohtaiset suhteet teollisuuteen ovat muodostuneet suurelta osin aiemman tutkimusyhteistyön kautta. Myös sen vuoksi otokseen ovat valikoituneet poikkeuksellisen hyvin vertailussa suoriutuvat yritykset.

Ciolkowski ym. (2003) kysyivät ensimmäisenä tarkoitusta katselmointien järjestämiseen. Vastanneista 73 % piti laadunvarmistusta erittäin tärkeänä, 52 % piti projektin statuksen arviointia tärkeänä ja 54 % ja 56 % näkivät katselmointien merkityksenä standardien toteutumisen kontrolloinnin. Keskeisimpinä esteinä katselmointien toteutumiselle nähtiin olevan aikapaine (75 %), kustannukset (56 %) ja koulutuksen puute (50 %). Vain 12 % sanoi saaneensa koulutusta katselmoinneista.

Vaatimuksia katselmoi säännöllisesti 42 %, suunnitelmia 40 % ja koodia 28 % vastanneista. Ainoastaan 40 % vastanneista sanoo tekevänsä itsenäistä virheiden etsimistä. Niistä, jotka tekevät itsenäisen virheiden etsimisen, 35 % toimii ad-hoc periaatteella, 50 % käyttää tarkistuslistaa ja loput muita tekniikoita. Vain 40 % vastanneista pitää säännöllisesti palaverin. Puolet vastaajista näkee palaverin tarkoituksena virheiden etsimisen, 57 % etukäteen löydettyjen virheiden keräämisen. Yrityksistä 60 % kerää jotakin tietoa katselmoinneista. Käytännössä esimerkiksi 36 % laskee virheiden määrän ja 20 % kerää tietoa tehokkuudesta.

Ciolkowskin ym. (2003) lisäksi Johnson (1998) on raportoinut vapaamuotoisesta kyselystä, jossa tutkittiin tarkastusten toteuttamista. Kysely toteutettiin USENET-verkoston avulla ja siinä kysyttiin ohjelmistoalan ammattilaisilta, tekevätkö he tarkastuksia säännöllisesti työssään. Kyselyyn saatiin 90 vastaajaa, joista 80 % ilmoitti heidän organisaationsa käyttävän tarkastuksia epäsäännöllisesti tai ei ollenkaan.

Tässä Johnsonin (1998) tutkimuksessa tutkimusaineistoa voidaan pitää vielä epäluotettavampana kuin Ciolkowskin ym. (2003) kyselyssä, sillä tutkimusotosta ei ole voitu millään tavalla kontrolloida. Molemmissa tutkimuksissa voidaan pitää todennäköisenä, että otokseen on valikoitunut sellainen joukko organisaatioita, joka ennestään toteuttaa tarkastuksia keskimääräistä paremmin tai on muuten erityisen kiinnostunut aiheesta. Siksi näitä molempia tutkimuksia voidaan pitää ainakin vahvana viitteenä siitä, että tarkastuksia toteutetaan säännöllisesti vain osassa yrityksistä.

Kyselytutkimusta menetelmänä katselmointien tutkimisessa voi pitää vähintäänkin haasteellisena. Tässä tutkimuksessa haastatteluissakin oli useissa tapauksissa hankaluuksia käsitteiden kanssa. Ongelmia käsitteiden kanssa tuli siitä huolimatta, että useimmat haastateltavat olivat tottuneet säännöllisiin kat-

selmointeihin ja haastattelussa on aina mahdollisuus selventää käsitteiden merkitystä. Haastateltavat tulkitsivat käsitteitä eri tavoin omasta kontekstistaan käsin. Siitä syystä tässä tutkimuksessa pitäydyttiin haastatteluissa, jotta organisaatioista saatiin keskenään vertailukelpoisia tuloksia.

5.2 Tarkastusten toteutuminen kohdeorganisaatioissa

Tässä kohdassa esitellään tarkastuskäytänteitä ja niiden toteutumista kuudessa kohdeorganisaatioissa. Tulokset perustuvat haastatteluihin, joiden toteutus on kuvattu luvussa 4. Haastateltaville esitettiin kysymyksiä, jotka kartoittivat tarkastusten ja vapaamuotoisempien katselmointien toteuttamista heidän omassa organisaatiossaan. Käytännössä haastatteluissa käytettiin tarkastuksen sijasta suunnilleen samassa merkityksessä termiä katselmointi, sillä se oli haastateltaville tutumpi käsite.

Haastatteluissa esitetyt kysymykset pohjautuivat suurelta osin luvussa 5. esiteltyyn ICMM-malliin, jota käytetään myös viitekehyksenä tulosten esittelyssä. Käytännössä organisaatioista löytyneet käytänteet sijoittuivat ICMM:n tasoille kaksi ja kolme. Taulukko 3 pitää sisällään tiivistettynä sen, miten organisaatiot haastattelujen perusteella toteuttavat näillä tasoilla vaadittuja käytänteitä. Musta pallo tarkoittaa sitä, että kriteeri täyttyy. Valkoinen tarkoittaa sitä, että kriteeri täyttyy vain osittain ja viiva tarkoittaa, ettei tällaista käytännettä ole lainkaan organisaation toiminnassa.

Haastateltavia pyydettiin aivan haastattelun aluksi subjektiivisesti arvioimaan, kuinka hyvin katselmoiteja toteutetaan heidän organisaatiossaan. Arviointi tehtiin asteikolla 1-5, jossa viisi on tarkoittaa parasta suoriutumista. Taulukossa 3 viimeisellä rivillä on organisaatiokohtainen keskiarvo haastateltavien subjektiivisista arvioista.

TAULUKKO 4 Käytänteiden toteutuminen kohdeorganisaatioissa

| Käytänne/ Organisaatio | O1 | O2 | O3 | O4 | O5 | O6 |
|---------------------------------|-----|-----|-----|----|----|-----|
| Taso 3 | | | | | | |
| Allokoidut resurssit | - | ○ | - | - | - | - |
| Vastuutehtävien määrittely | ○ | ○ | ○ | ○ | ○ | ○ |
| Organisaation politiikka | - | ● | ● | ● | ● | - |
| Tiedon keräys ja käyttö | - | - | - | - | - | - |
| Mukautettu tukimateriaali | - | ○ | ○ | - | - | - |
| Koulutusta kaikille | - | - | - | - | - | - |
| Määritelty prosessi | - | ● | ● | ● | ○ | ● |
| Koodin tarkastus | - | ○ | - | ○ | ○ | - |
| Testitapausten tarkastus | ○ | ● | ○ | ● | ● | ○ |
| Taso 2 | | | | | | |
| Tarkastuksen johtajien koulutus | - | - | - | - | - | - |
| Suunnitelmien tarkastus | ○ | ● | ○ | ● | ● | ○ |
| Vaatimusten tarkastus | ○ | ● | ● | ● | ● | ○ |
| Oma arvio / keskiarvo | 2,8 | 2,8 | 3,3 | 3 | 2 | 2,7 |

Seuraavat alakohdissa käydään läpi yleiskuva tutkimuksen tuloksista sekä kaikkien ICMM-mallin toisella ja kolmannella tasolla vaadittujen käytänteiden toteutuminen kohdeorganisaatioissa. Joissakin kohdissa tuloksia on verrattu Ciolkowskin ym. (2003) tutkimukseen. Tätä vertailua ei voi ottaa kovin vakavasti, sillä Ciolkowski ym. eivät raportoi tarkasti heidän kyselyssään esitettyjä kysymyksiä. Luotettavaa vertailua ei siis voida tehdä, mutta joitakin mielenkiintoisia näkökulmia vertailun avulla voidaan pohtia.

5.2.1 Yleiskuva tuloksista

Tarkastuskäytänteet kohdeorganisaatioissa vaihtelivat huomattavasti. Parhaassa tapauksessa tarkastusprosessi on hyvin määritelty ja useimmat dokumentit tarkastetaan jokaisessa projektissa. Toisessa ääripäässä parilla kohdeorganisaatiolla on vielä tehtävää siinä, että tarkastukset saadaan säännölliseen käyttöön. Taulukko 3 osoittaa kuitenkin, että parhaassakin yrityksessä täyttyi vain puolet ICMM-mallin toisen ja kolmannen tason asettamista kriteereistä. Käytännössä tämä tarkoittaa, että tarkastukset voivat olla säännöllisesti käytössä, mutta niiden tehokkuudessa saattaa olla toivomisen varaa. Tehokkuutta ei tosin voitu tosin empiirisesti tässä yhteydessä tutkia, sillä yksikään kohdeorganisaatioista ei käytä siihen tarvittavia metriikoita. Oletus perustuu aiempiin tutkimuksiin, joiden pohjalta ICMM-mallin kriteerit on kehitetty.

Taulukko 3 ei kerro koko totuutta organisaatioiden käytänteistä. Esimerkiksi organisaatioilla O4 ja O5 on taulukossa lähes identtiset profiilit, koska ne täyttävät samat kriteerit. Kuitenkin haastateltavat organisaatioissa O4 olivat oman organisaationsa käytänteisiin vähän tyytyväisempiä ja haastattelujen perusteella heidän käytänteensä vaikuttivat paremmin toimivilta kuin organisaatioissa O5.

Haastateltavilta siis kysyttiin aluksi, kuinka hyvin he arvioivat katselmointien toimivan heidän omassa organisaatiossaan. Mielenkiintoinen havainto oli, että näissä subjektiivisissa arvioissa oli melko vähän eroja riippumatta siitä, kuinka hyvin ne vaikuttivat toimivan käytännössä. Esimerkiksi haastateltavat organisaatioissa O1 ja O2 antoivat omalle organisaatiolleen keskimäärin saman arvosanan. Kuitenkin taulukossa 3 tehdyn vertailun perusteella ne olivat tutkimuksen kaksi ääripäätä tarkastuskäytänteiden toteutumisessa.

Haastateltavien arvioihin voi olla monenlaisia vaikuttavia tekijöitä. Jotkut haastateltavat saattavat arvostaa hyvin vapaita ja epämuodollisia käytänteitä, jotka eivät täytä lainkaan ICMM-mallin kriteereitä. Yleisellä tietämyksellä vaikutti myös olevan merkitystä arvioihin. Monessa tapauksessa kohtuullisen hyvä tieto tarkastusten potentiaalista sai haastateltavat asennoitumaan kriittisemmin nykyisiin käytänteisiin. Joka tapauksessa kaikki kohdeorganisaatiot näkivät huomattavasti kehitettävää omissa käytänteissään riippumatta käytänteiden tämän hetkisestä kypsyystasosta. Kehittämiskohteet vain ovat erilaisia eri kehitysvaiheissa olevissa organisaatioissa.

5.2.2 Vaatimusten tarkastus

Vaatimusmäärittelyt vaikuttivat olevan parhaiten tarkastettuja dokumentteja kohdeorganisaatioissa. Neljässä organisaatioissa vaatimukset tarkastettiin jokaisessa projektissa. Kriteerinä tarkastukselle tässä oli, että tilaisuuteen valmistaudutaan ja siitä tehdään jonkinlainen raportti (ks. ICMM-mallin kuvaus 5. luvussa). Myös kahdessa muussa organisaatioissa vaatimusmäärittelyjä tarkastetaan, mutta ei kuitenkaan jokaisessa projektissa. Silloinkin, kun varsinaista tarkastusta ei tehdä, pyydetään vaatimusmäärittelyyn yleensä vapaamuotoisemmin kommentteja vaikkapa sähköpostin välityksellä.

Ciolkowski ym. (2003) totesivat kyselytutkimuksessaan, että 42 % vastaajista ilmoitti katselmoivansa vaatimusdokumentteja säännöllisesti omassa organisaatioissaan. Tämän tutkimuksen kuudessa kohdeorganisaatioissa suhde on korkeampi. Oletettavasti ainakin viidessä organisaatioissa haastateltavat olisivat sanoneet katselmoivansa vaatimusmäärittelyt säännöllisesti.

5.2.3 Suunnitelmien tarkastus

Haastattelujen perusteella ainakin joitakin suunnitteludokumentteja katselmoidaan jokaisessa kohdeorganisaatioissa, mutta vain kolmessa suunnitelmia tarkastetaan muodollisesti jokaisessa projektissa. Haastatteluissa kysyttiin erikseen eri tarkkuustason suunnitelmien tarkastuksesta. Arkkitehtuuritason suunnitelmat tarkastetaan useimmin, kun taas hyvin teknisen tason suunnitelmia tarkastetaan harvemmin. Joissakin organisaatioissa kovin yksityiskohtaisia teknisen tason suunnitelmia ei edes tehdä.

Ciolkowskin ym. (2003) tutkimuksessa 40 % tutkimukseen osallistuneista ilmoitti katselmoivansa suunnitelmia säännöllisesti. Tämän tutkimuksen kuudesta organisaatiosta todennäköisesti viisi olisi tähänkin kohtaan ilmoittanut katselmoivansa joitakin suunnitelmia säännöllisesti. Ciolkowski ym. eivät erottelleet kyselyssään erilaisia suunnitelmia, joten on vaikea tietää, mitä vastaajat ajattelivat tässä kohdassa. Jotkut haastateltavat tässä tutkimuksessa eivät pitäneet arkkitehtuurikuvausta suunnitteludokumenttina ennen kuin siitä keskusteltiin. Syynä oli se, että organisaation prosessissa arkkitehtuurisuunnittelua pidetään määrittelyvaiheeseen kuuluvana tuotoksena. Joka tapauksessa tämän tutkimuksen kohdeorganisaatiot vaikuttavat katselmoivan suunnitelmia säännöllisemmin kuin Ciolkowskin ym. kyselyyn vastanneet organisaatiot.

5.2.4 Tarkastusten johtajien koulutus

Yhdessäkään kohdeorganisaatioissa ei järjestetä minkäänlaista katselmointeihin liittyvää koulutusta. Toiseksi, organisaatioissa ei ole varsinaisesti tällaista roolia kuin tarkastuksen johtaja. Tarkastuksen organisoinnin hoitaa tavallisimmin tarkastettavan tuotoksen tekijä. Palaverissa puolestaan tukeudutaan yleisiin kokouskäytänteisiin eli valitaan puheenjohtaja ja sihteeri. Haastattelujen perusteella varmaankin tavallisin valinta puheenjohtajaksi on projektipäällikkö, mutta joissakin tapauksissa asiakkaan ollessa paikalla, hän ottaa puheenjohtajan roolin.

ICMM pitää sisällään ajatuksen siitä, että tarkastusten johtajat on viisasta kouluttaa ensin. Vaatimus muidenkin työntekijäryhmien koulutuksesta tulee mallin seuraavalla tasolla. Lähimpänä tarkastuksen johtajan roolia on varmasti-kin projektipäällikkö. Siksi näiden kohdeorganisaatioiden kohdalla tämän kri-terin voisi katsoa täyttyvän projektipäällikköiden kouluttamisella.

5.2.5 Testitapausten tarkastus

Testitapauksia pidetään selvästi tärkeinä dokumentteina kohdeorganisaatioissa, sillä niitä katselmoidaan yhtä usein kuin korkean tason suunnitelmia. Kolmessa organisaatiossa testisuunnitelmat testitapauksineen tarkastetaan jokaisessa projek-tissa. Muissakin organisaatioissa niitä katselmoidaan hiukan epäsäännöllii-semmin tai käyttäen epämuodollisempia käytänteitä.

5.2.6 Koodin tarkastus

Koodia ei kohdeorganisaatioissa kovinkaan usein tarkasteta. Vain puolessa or-ganisaatioista koodia tarkastetaan edes satunnaisesti. Asenteet koodin tarkas-tamiseen olivat haastatteluissa vaihtelevia. Joidenkin mielestä puutteet koodin tarkastamisessa olivat ongelma, jolle on yritettävä tehdä jotain. Toisaalta jotkut haastateltavat eivät pitäneet koodin tarkastusta lainkaan tärkeänä.

Ciolkowskin ym. (2003) tutkimuksessa 28 % vastaajista ilmoitti katsel-moivansa koodia säännöllisesti. On vaikea sanoa, kuinka haastateltavat tämän tutkimuksen kohdeorganisaatioissa olisivat vastanneet tässä samassa muodossa esitettyyn kysymykseen. Mahdollisesti kukaan heistä ei olisi sanonut koodia katselmoitavan säännöllisesti omassa organisaatiossa. Tämä on kuitenkin mie-lenkiintoinen ero verrattuna Ciolkowskin ym. tutkimukseen, jossa ero vaati-musten ja koodin katselmoinnissa ei ollut niin merkittävä. Tämän tutkimuksen kohdeorganisaatioissa taas vaatimukset katselmoidaan lähes aina, mutta koodia hyvin harvoin.

5.2.7 Määritelty tarkastusprosessi

Neljällä kohdeorganisaatiolla on määritelty tarkastusprosessi, joka kertoo, kuin-ka tarkastukset tulee organisoida. Näiden lisäksi vielä organisaatiolla O5 on yleisempää ohjeistusta, joka käytännössä määrää, että katselmointeja on järjes-tettävä. Se ei kuitenkaan samalla tavalla anna suoranaisia käytännön ohjeita kuin muiden organisaatioiden ohjeistus. Siis ainoastaan yhdellä organisaatiolla ei ollut minkäänlaista tarkastuksiin liittyvää prosessidokumentaatiota.

Haastatteluissa selvisi, että tarkastusprosessin määritelleissäkin organisaa-tioissa prosessi tunnettiin huonosti. Kaikki haastateltavat eivät olleet edes tie-toisia prosessimäärityksen olemassaolosta ja lisäksi useat muut tiesivät hyvin vähän sen sisällöstä. Tämä saattaa luonnollisesti olla yhteydessä koulutuksen puutteeseen. Toiseksi voidaan myös miettiä, onko tiedolla prosessimäärityksis-tä mitään merkitystä tavallisen työntekijän kannalta. Onko mahdollista, että on

totuttu toimimaan täsmälleen määritysten mukaisesti, vaikka ei tiedetä kirjallisesta prosessidokumentaatiosta mitään?

5.2.8 Koulutus kaikille

Kuten jo edellä todettiin, kohdeorganisaatioissa ei ole minkäänlaista katselmoiteihin liittyvää koulutusta. Myös Ciolkowski ym. (2003) totesivat kyselytutkimuksessaan samanlaisen ilmiön. Ainoastaan 12 % heidän tutkimuksensa vastaajista ilmoitti saaneensa koulutusta katselmoiteja varten.

Koulutusta voi epäilemättä suositella kaikille kohdeorganisaatioille. Työntekijöiden motivoiminen lukemaan toisten tuotoksia oli yksi suurimmista haasteista jokaisessa organisaatiossa. Lisäksi yleinen tietämys tarkastuksista ei ollut erityisen hyvä. Siis ainakin työntekijöiden motivoimisen ja yleisen tietämyksen lisäämisen tarkoituksessa koulutus näyttäisi olevan tarpeen kaikissa kohdeorganisaatioissa.

5.2.9 Mukautettu tukimateriaali

Yksikään tutkimuksen kohdeorganisaatioista ei käytä minkäänlaista tukimateriaalia virheiden etsimisessä. Kahdella organisaatiolla on olemassa jonkinlaisia tarkistuslistoja, mutta haastateltavien mukaan niitä ei kuitenkaan käytetä. Kaikki haastateltavat eivät edes tieneet niiden olemassaolosta mitään.

Tavallinen asenne tarkistuslistoja tai vastaavaa tukimateriaalia kohtaan oli haastatteluissa aika skeptinen. Harvalla oli kokemusta tällaisen materiaalin käytöstä ja useimmat arvelivat niistä olevan hyötyä korkeintaan aloittelijoille. Hyvästä tukimateriaalista on kuitenkin tutkimusten mukaan havaittu olevan hyötyä (Porter ja Votta 1998). Vaikuttaa, että tukimateriaalin kohdalla sen käytöstä täytyy organisaatiossa saada omakohtaisia positiivisia kokemuksia ennen kuin sellainen otetaan. Myös koulutuksella voidaan varmasti parantaa yleistä tietoisuutta tässäkin asiassa.

5.2.10 Tiedon keräys ja käyttö

Yksikään kohdeorganisaatioista ei kerää systemaattisesti tietoa tarkastuksista. Tarkastusraportti (jos sellainen tehdään) on yleensä tekstidokumentti, joka tavallisesti sisältää dokumentin tunnustietojen lisäksi valmistautumiseen käytetyn ajan, löytyneet virheet ja korjaavat toimenpiteet, jotka päätetään tehtäviksi. Siis monessa organisaatiossa tarkastuksista kirjoitetaan raportteihin monenlaisia hyödyllistä tietoa, mutta sitä ei kuitenkaan kerätä talteen. Kukaan ei vaikuta olevan kiinnostunut raporteista tarkastuksen jälkeen.

Kaksi organisaatioista luokittelee virheet eri vakavuusluokkiin tarkastusraporteissaan, mikä on käytännössä edellytys kunnollisten löydettyjen virheiden määrään perustuvien metriikoiden käyttöön. Sen ansiosta näissä organisaatioissa on olemassa validia historiatietoa ja niissä ei pitäisi olla kovin korkeaa kynnystä ryhtyä seuraamaan tarkastusten toteutumista metriikoiden avulla.

5.2.11 Organisaation politiikka

Neljässä kohdeorganisaatiossa oli dokumentoituna jonkinlainen organisaatiotasoinen politiikka katselmointien järjestämisestä. Tavallisesti se on kirjoitettu tarkastusprosessin kuvauksen yhteyteen ja siinä vaaditaan, että kaikki ohjelmistokehityksessä tuotettavat dokumentit on katselmoitava. Haastattelujen tuloksista voidaan nähdä, että tällainen politiikka ei toteudu käytännössä lähiainakaan. Joillakin organisaatioilla ei ole selvästi edes halua noudattaa tällaista politiikkaa, jonka ovat kuitenkin itse asettaneet. Yhdessäkään organisaatiossa ei käytännössä seurata aktiivisesti, toteutetaanko määriteltyä politiikkaa.

5.2.12 Määritellyt vastuutehtävät

Kaikissa organisaatiossa oli olemassa laatupäällikkö tai sitä vastaava rooli, joka on vastuussa laadusta ja ainakin jostain osasta prosessikehitystä. Käytännössä vastuu tarkastuksista on jätetty projektipäälliköille. Laatupäälliköt eivät käytännössä seuraa tarkastusten toteutusta, vaan ovat korkeintaan mukana käytänteitä kehitettäessä. Tarkastusten toteutukseen liittyvät vastuualueet ovat siis yleisesti melko heikosti määritelty, vaikka jonkinlainen työnjako sisäisesti onkin olemassa. Selkeän määrittelyn puuttuminen näkyy joidenkin laadunhallintaan liittyvien toimintojen, kuten tarkastusten seurannan, puutteena.

Käytännössä asiakas vaikuttaa olevan keskeisin osapuoli tarkastusten seurannassa. Jotkut asiakkaat vaativat hyvinkin tarkasti dokumenttien tarkastusta ja valvovat niiden toteutusta. Organisaatioissa oli selvästi nähtävissä että mitä vähemmän asiakas on kiinnostunut tarkastuksista, sitä harvemmin niitä toteutetaan. Vaikuttaa siis siltä, että tarkastuksia tehdään suurimmaksi osaksi asiakasta varten, ei niinkään oman toiminnan tehostamista ajatellen.

5.2.13 Allokoidut resurssit

Vain yhdessä kohdeorganisaatioista allokoidaan tavallisesti resurssit erikseen tarkastuksia varten. Sekin varaa resurssit hyvin yleisesti projektisuunnitelmassa, ei yksittäisten työntekijöiden työn suunnittelun tasolla. Tavallisesti organisaatioissa oletetaan tarkastusten sisältyvän suurempien tehtäväkokonaisuuksien sisään.

Tässä kohdassa vaikuttaa olevan kysymys enemmän organisaatiokulttuurista kuin muodollisesti allokoiduista resursseista. Joissakin organisaatioissa kaikki haastateltavat olisivat nähneet hyvänä resurssien varaamisen työntekijöiden henkilökohtaisiin työsuunnitelmiin. Muuten työntekijöillä saattaa olla tunne siitä, että pitäisi tehdä jotain "tuottavaa" tarkastuspalaveriin valmistautumisen sijasta. Toisaalta joissakin organisaatioissa ei ollut lainkaan tällaisia kokemuksia, eivätkä he tahdo suunnitella resursseja niin yksityiskohtaisesti.

5.3 Tutkimustulosten analysointia

Edellä esitetyistä tuloksista voidaan nähdä, että säännöllisestikään tarkastuksia toteuttavassa organisaatiossa ne eivät toteudu täydellisesti. Todellista tarkastusten tehokkuutta ei voitu tutkimuksessa mitata, koska organisaatiot eivät käytä siihen tarvittavia metriikoita. Kuitenkin aiempien tutkimusten valossa on ilmeistä, että kaikilla kohdeorganisaatioilla on vielä kehittämisen varaa heidän tarkastuskäytänteissään.

Organisaatio O5 (ks. taulukko 3) on esimerkki siitä, kuinka säännöllisestikin toteutetuissa tarkastuksissa voi olla merkittäviä puutteita. Tässä luvussa tarkasteltujen kriteerien mukaan käytänteet toimivat säännöllisesti, mutta toiminta keskittyy suurelta osin projektin tarkastuspisteissä toteutettaviin katselmoointeihin, joiden tarkoitus ei ole enää etsiä virheitä, vaan ainoastaan hyväksyä vaihetuotokset ennen projektin seuraavaan vaiheeseen siirtymistä. Varsinainen tarkastus, joka keskittyy virheiden etsimiseen, toteutetaan tavallisesti yhden tarkastajan toimesta. Haastateltavien mukaan tämän yhden tarkastajan työn laadussa on usein toivomisen varaa. Kaikki haastateltavat organisaatiossa näkivät nykyisten käytänteiden ongelmat ja kaipasivat lisää formaalimpia tarkastuksia.

Ciolkowskin ym. (2003) tutkimuksen ja tämän tutkimuksen tuloksilla oli voitakin mielenkiintoisia eroja. Tämän tutkimuksen kohdeorganisaatiot vaikuttavat katselmoivan vaatimuksia ja suunnitelmia säännöllisemmin kuin Ciolkowskin ym. kyselyyn vastanneet. Kuitenkin heidän kyselyynsä vastanneista suurempi osa ilmoitti katselmoivansa koodia säännöllisesti.

Eräs selitys tutkimustulosten eroihin saattaa olla asiakassuhteen vaikutus kohdeorganisaatioissa. Haastattelujen tuloksissa oli nähtävissä, että mitä enemmän asiakkaat ovat kiinnostuneita tarkastuksista, sitä kurinalaisemmin niitä järjestetään. Tarkastuksista yleisesti kiinnostuneetkaan asiakkaat eivät vaikuta olevan kovin kiinnostuneita koodin ja hyvin teknisten suunnitelmien tarkastuksesta. Ilmeisesti sen vuoksi myös säännöllisesti tarkastuksia toteuttavissa kohdeorganisaatioissa tarkastetaan näitä dokumentteja hyvin harvoin. Näistä tuloksista voidaan siis päätellä, että kohdeorganisaatiot eivät ole kovin sitoutuneita tarkastuksiin oman sisäisen toimintansa vuoksi, vaan tekevät niitä pääasiassa asiakasta varten. Kenties Ciolkowskin ym. tutkimuksessa harvemmat säännöllisesti katselmoiteja toteuttavat organisaatiot ovat omaksuneet niiden merkityksen omalle toiminnalle paremmin. Ciolkowski ym. eivät suoraan sitä raportoi, mutta heidän tulostensa perusteella voidaan pitää todennäköisenä, että suuri osa vaatimuksia säännöllisesti katselmoivista organisaatioista katselmoivat säännöllisesti myös koodia.

Koulutuksen puute ja organisaation sisäinen kommunikointi olivat selvästi yleisiä heikkouksia jokaisen kohdeorganisaation tarkastuksiin liittyvässä käytänteissä. Yleinen tietämys tarkastuksista vaikutti olevan melko rajallista. Työntekijät kenties tietävät, että katselmointi on ”hyvä juttu”, mutta heillä ei ole kuitenkaan kokonaisvaltaista ymmärrystä tarkastusten merkityksestä tuotantopro-

sessiin. Työntekijät vaikuttavat lisäksi olevan heikosti perillä määritellystä tarkastusprosessista tai organisaation politiikasta. Siis tarkastuskäytänteisiin liittyvälle koulutukselle vaikuttaisi olevan tarvetta kaikissa kohdeorganisaatioissa.

Eräs selvä heikkous kaikissa organisaatioissa oli metriikoiden ja tarkastuskäytänteiden seurannan puute. Työntekijät organisaation kaikilla tasoilla vaikuttavat olevan skeptisiä uusien menetelmien suhteen, vaikka niiden toimivuudesta olisi näyttöä jossain toisessa organisaatiossa. Lisäksi tarkastukset eivät vaikuta olevan ohjelmistosuunnittelijoille intuitiivinen tai mukava tapa toimia. Suurin haaste jokaisessa kohdeorganisaatiossa haastattelujen perusteella on saada motivoitua työntekijät lukemaan huolellisesti toistensa tuotoksia. Rationaalisia perusteluja työntekijöiden motivoimista ja toiminnan kehittämistä varten on vaikea saada muuten kuin mittaamalla toimintaa. Tämän hetkisessä tilanteessa kohdeorganisaatiot eivät voi muuta kuin arvata, kuinka mahdollinen tarkastuskäytänteiden kehittäminen vaikuttaa heidän tuotantoprosessiinsa.

Tässä luvussa esitettyä organisaatioiden tarkastuskäytänteistä esitettyä vertailua ei pidä ottaa liian vakavasti. Yleisesti tämän kaltaisen kypsyystasomallin käyttö johtaa tiettyihin ongelmiin. On esimerkiksi hyvin tunnettua, että työntekijän yksilöllinen suorituskyky on suurin tarkastusten tehokkuuteen vaikuttava tekijä (Sauer ym. 2000). On mahdollista, että alemmalla kypsyystasolla olevalla organisaatiolla on erityisen hyvin tai korkealla kypsyystasolla olevalla organisaatiolla heikosti motivoituneet ja osaavat työntekijät. Korkea kypsyystaso ei siis välttämättä takaa käytänteiden tehokkuutta.

Toinen huomioon otettava asia organisaatioiden vertailussa on siinä käytetty ICM-malli. Käytetty malli vasta alustava versio, eikä sitä ole laajasti hyväksytty tai tässä kehitysvaiheessa edes arvioitu empiirisesti. Kuitenkin mallin rooli tässä tutkimuksessa on olla ainoastaan viitekehys organisaatioiden välisessä vertailussa. Käytännössä se toimi hyvin tuossa roolissa. Siis yksittäisten mallin sisältämien käytänteiden toteutumista organisaatioissa voidaan hyvin vertailla keskenään. Sen sijaan mallin tuottamaan tasoluokituksen on hyvä asennoitua kriittisemmin.

5.4 Yhteenveto

Tässä luvussa on käsitelty tarkastuskäytänteiden toteutumista käytännössä tutkimuksen kuudessa kohdeorganisaatiossa. Tuloksia verrattiin joissakin kohdissa aiempaan kyselytutkimukseen (Ciolkowski ym. 2003), joka selvitti katselmointien toteuttamista ohjelmistoyrityksissä. Tämän tutkimuksen kohdeorganisaatiot toteuttavat tarkastuksia säännöllisemmin kuin Ciolkowskin ym. tutkimuksessa. Poikkeuksena oli koodin katselmointi, jota puolestaan toteutettiin säännöllisemmin Ciolkowskin ym. tutkimuksen mukaan.

Keskeisin johtopäätös tässä luvussa esitetyistä tuloksista oli, että tarkastuksia säännöllisestikin toteuttavien organisaatioiden käytänteissä oli merkittäviä puutteita, jotka oletettavasti syövät toiminnan tehokkuutta. Tehokkuusvai-

kutusta ei tosin voitu mitata, sillä organisaatioissa ei ole käytössä siihen soveltuvia metriikoita.

Haastattelujen tuloksissa oli havaittavissa seuraavia mielenkiintoisia trendejä:

- Asiakassuhteella vaikutti olevan selvä vaikutus tarkastuskäytänteiden kurinalaisuuteen. Tarkastuksia tehdään suurelta osin asiakasta varten, ei niinkään organisaation omien tarpeiden vuoksi.
- Kohdeorganisaatioissa ei järjestetä minkäänlaista tarkastuksiin liittyvää koulutusta. Yleinen tietämys tarkastuksista, määritellyistä käytänteistä ja politiikoista vaikutti melko vähäiseltä.
- Tarkastuksista ei kerätä tietoa eikä tarkastusten toteutumista seurata. Tästä syystä käytössä ei ole myöskään metriikoita, joiden avulla tarkastusten hyödyllisyyttä organisaatiossa voisi arvioida.

Kaikki edellä mainitut löydökset viittaavat johtopäätökseen, että kohdeorganisaatioissa ei ehkä ole täysin ymmärretty tarkastusten mahdollisia hyötyjä oman toiminnan kannalta. Useimmat organisaatioista ovat ylittäneet kehityksessä jo ensimmäisen kynnyksen ja ovat saaneet tarkastukset säännölliseen käyttöön ainakin joidenkin tuotosten osalta. Seuraava askel olisi kehittää käytänteiden tehokkuutta. Tässä luvussa organisaatioiden vertailuun käytetty ICMM-malli voi toimia kehittämisen apuvälineenä, mutta se ei itsessään riitä takaamaan käytänteiden toimivuutta. Säännöllisestikin toimivat käytänteet voivat pitää sisällään ongelmia, jotka tekevät käytänteistä tehottomia. Tällaisiin tarkastuskäytänteiden ongelmiin palataan luvussa 6.

Tässä luvussa esitetyistä tuloksista nousee mielenkiintoisia kysymyksiä tulevaa tutkimusta varten. Tarkastuskäytänteiden tyypillisiä puutteita ja ongelmia tulisi kyetä ymmärtämään erilaisissa organisaatioissa. Tämän tutkimuksen kokemusten perusteella saattaa olla mahdollista luoda kysely, jolla voidaan saavuttaa suurempi otos erityyppisiä organisaatioita. On oletettavaa, että suuremmasta otoksesta voidaan todeta erityyppisillä organisaatioilla olevan erilaisia ongelmia. Toiseksi tulevaisuudessa tulisi kyetä ymmärtämään paremmin käytänteiden kehittämisen todellisia vaikutuksia organisaatioissa. Sitä varten täytyisi tutkia useampia sellaisia kohdeorganisaatioita, joilla on käytössään sopivia tarkastusten toimintaa seuraavia metriikoita.

6 TARKASTUSTEN ONGELMAT

On luonnollista, että tarkastusten toteuttamiseen inhimillisenä toimintana liittyy monenlaisia ongelmia. Aiemman tutkimuksen perusteella suurin osa yrityksistä ei edes toteuta tarkastuksia säännöllisesti (Johnson 1998 ja Ciolkowski ym. 2003). Sitä kuinka hyvin tarkastuksia toteutetaan, ei aiemmin ole juuri tutkittu. Kuitenkin on hyvin todennäköistä, että tarkastuksia säännöllisestikin toteuttavien yritysten toiminnassa on kehittämisen varaa. Ainakin tämän tutkimuksen empiirinen osa antaa vahvasti tukea sille oletukselle.

Saattaa olla, että yritys on huolellisesti määritellyt tarkastusprosessin ja noudattaa sitä säännöllisesti, mutta silti toiminnassa saattaa olla ongelmia, jotka syövät merkittävästi tarkastusten tehokkuutta. Huonoimmassa tapauksessa tarkastukset saatetaan hylätä liian kalliina ja tehottomana menetelmänä ennen kuin ne saadaan toimimaan ilman merkittäviä ongelmia. Kun mietitään tarkastusten kehittämistä ohjelmistoyrityksessä, ei siis riitä ainoastaan prosessitasolla tapahtuva toiminnan kehittäminen. Sen lisäksi on analysoitava tarkastusten toteutumista käytännössä ja ymmärrettävä, minkälaisilla asioilla on oleellinen vaikutus tarkastusten tehokkuuteen.

Kirjallisuudessa tarkastuksiin liittyviä ongelmia ei ole käsitelty aiemmin systemaattisesti, eikä siten ole olemassa kovin monia tutkimustuloksia, joista voitaisiin päätellä erilaisten ongelmien konkreettinen vaikutus tarkastusten tehokkuuteen. Tässä luvussa yritetään kuitenkin päästä alkuun ongelmien tutkimisessa hahmottamalla yleisesti tarkastusten toteuttamiseen liittyvää ongelma- kenttää kirjallisuuden ja empiirisen tutkimuksen avulla.

Seuraava kohta 6.1 käsittelee katselmointien toteuttamiseen liittyviä ongelmia kirjallisuuden pohjalta. Se pyrkii kasaamaan yhteen aiemmassa tutkimuksessa mainittuja ongelmatilanteita ja samalla pohtimaan aiemman tutkimuksen valossa ongelmien mahdollisia vaikutuksia tarkastusten tehokkuuteen. Kohta 6.2 esittelee tuloksia empiirisestä tutkimuksesta, jonka tarkoituksena oli kartoittaa tarkastuksiin liittyvien ongelmien esiintymistä käytännössä. Kohdassa 6.3 esitetään yhteenveto tärkeimmistä tarkastusten ongelmia käsittelevistä johtopäätöksistä.

6.1 Tarkastusten ongelmat kirjallisuudessa

Tarkastusten toteutumiseen liittyviä ongelmia ei ole aiemmin systemaattisesti tutkittu. Kuitenkin tarkastuksia eri näkökulmista tarkastelevista tutkimusraporteista löytyy jonkinlaisia viitteitä käytännössä havaittuihin ongelmatilanteisiin. Seuraaviin alakohtiin on listattu joukko tällaisia kirjallisuudessa mainittuja ongelmia ja mahdollisia ratkaisuja niihin.

Tässä kohdassa (6.1) esitettävää listaa käytettiin tutkimuksen empiirisessä osassa kartoittamaan sitä, millaisia ongelmia organisaatioissa käytännössä koetaan. Tässä ei ole siis tarkoitus väittää esitettyjä ongelmatilanteita todellisiksi, vaan ne toimivat tutkimuksessa hypoteeseina, joiden merkittävyyttä käytännössä arvioitiin tutkimukseen osallistuneissa kohdeorganisaatioissa. Lista on myös liitetty perusteluineen pari oletettua ongelmaa, joihin kirjallisuudesta ei löytynyt viitteitä.

Seuraavissa alakohdissa esitellään lyhyesti jokainen mahdollinen ongelmatilanne tehokkuuden näkökulmasta ja samalla esitetään mahdollinen ratkaisu ongelmaan. Tämä lista on alun perin julkaistu artikkelissa Kollanus (2005b). Listan asiat eivät ole missään tietyssä järjestyksessä, eivätkä kaikki välttämättä ole todellisia ongelmia käytännössä. Tämä lista edustaa ainoastaan hypoteettisia ongelmatilanteita ja niiden mahdollisia ratkaisuja. Tässä mainittujen asioiden merkittävyyttä käytännössä tarkastellaan kohdassa 6.2.

6.1.1 Palaverin aikataulutus voi aiheuttaa viivettä

Yhteisen ajan löytäminen palaveria varten saattaa usein olla ongelmallista. Yleensä tarkastusten aikataulu suunnitellaan myöhemmin kuin useimpien muiden tehtävien resursointi. Votta (1993) tutkimuksen mukaan jopa 20 % vaatimusmäärittelyyn ja suunnitelmien tuottamiseen kuluvasta ajasta saattaa mennä tarkastuspalaverin odotteluun. Hänen tutkimuksessaan tyypillinen odotusaika suuremmissa projekteissa oli 20 %. Chatzigeorgiou ja Antoniadis (2003) havaitsivat omassa tutkimuksessaan, että heikosta projektin resurssien suunnittelusta johtuen katselmointien ja kehitystyön vaatimat työkuormat osuivat usein samalle viikolle. Tämä saattaa olla selitys myös Votta (1993) tutkimuksen tuloksissa.

Tutkimuksessaan Votta (1993) kyseenalaistaa palaverin merkityksen tarkastusten tehokkuudelle. Gilbin ja Grahamin (1993) mukaan optimaalinen suhde tarkastuksessa olisi, että 80 % virheistä löydetään valmistautumisen aikana ja loput 20 % palaverissa. Votta (1993) tutkimuksessa kuitenkin palaverissa löytyi ainoastaan 4 % kaikista virheistä. Sen perusteella hän esittää, että tavallinen tarkastuspalaveri ei välttämättä ole tehokas tapa toimia. Hän ei kuitenkaan suosittelen kokonaan palaverista luopumista, vaan esittää, että palaveriin osallistuisi tuotoksen tekijän lisäksi ainoastaan yksi tarkastaja.

Myöhemmin useimmat tutkijat (esim. Harjuma ym. 2001; Johnson ja Tjahjono 1998; Perpich ym 1997) virtuaalista tarkastusta vastaukseksi aikatauluongelmaan. Virtuaalisella tarkastuksella tässä tarkoitetaan työkalutuettua

tarkastusprosessia, johon ihmiset voivat antaa oman työpanoksensa ajasta ja paikasta riippumatta. Joissakin tehdyissä tutkimuksissa on todettu virtuaalinen tarkastus yhtä kyvykkääksi löytämään virheitä ja samalla kustannuksiltaan pienemmäksi kuin perinteinen palaveripohjainen tarkastus (Johnson ja Tjahjono 1998; Perpich ym. 1997). Myös Stein ym. (1997) ovat samansuuntaisia tuloksia virtuaalisesta tarkastuksesta. Kuitenkaan he eivät suosittelle palaverista luopumista, vaan pitävät virtuaalista tarkastusta ennemmin täydentävänä menetelmänä, jota voidaan käyttää silloin kuin palaveria ei voida järjestää. Heidän tutkimuksessaan tietyn tyyppisiä virheitä löytyi heikosti ilman fyysistä palaveria.

Mahdollinen ratkaisu: Katselmointien ja sen vaatimien resurssien suunnittelun tulisi olla osa koko projektin suunnittelua, jotta tarkastusten vaatimat resurssit ovat käytettävissä oikeaan aikaan. Optimaalisessa tilanteessa tarkastusten aika-tila ja resursointi tulisi tehdä osana projektisuunnitelmaa yhdessä muiden tehtävien suunnittelun kanssa. Aiempien tutkimusten valossa toinen ratkaisu tähän ongelmaan saattaa olla jokin vaihtoehtoinen menetelmä perinteisen tarkastuspalaverin järjestämisen sijaan.

6.1.2 Palaveri kuluttaa resursseja, mutta siellä ei löydetä paljon virheitä

Jo edellisessä kohdassa viitattiin tähän ongelmaan. Votta (1993) tutkimuksessa tarkastuspalaverissa löytyi merkittävän vähän uusia virheitä niiden lisäksi, jotka oli löydetty jo valmistautumisvaiheessa. Myöhemmin Porterin ja Votta (1998) tekemä tutkimus on tukenut tätä havaintoa. Molemmat tutkimukset kuitenkin havaitsivat palaverin suodattavan tehokkaasti vääriä löydöksiä. Vertailuryhmä, joka ei järjestänyt palaveria lainkaan, oli löydettyjen todellisten virheiden määrässä suunnilleen yhtä tehokas, mutta se raportoi huomattavasti enemmän vääriä löydöksiä. Toiseksi, pelkkä tehokkuus virheiden määrässä mitattuna on rajoittunut näkökulma palaverin merkitykseen. Esimerkiksi Perpich ym. (1997) ja Johnson (1998) näkevät palaverin tärkeänä oppimisen paikkana.

Palaverin tehottomuudelle voi olla myös muita selittäviä tekijöitä, kuten huono valmistautuminen. Jos osallistujat ovat huonosti valmistautuneita, löydetään vain vähän merkittäviä virheitä (Gilb ja Graham 1993). Edellä mainituissa tutkimuksissa (Votta 1993; Porter ja Votta 1998) data kerättiin ympäristössä, jossa tarkastukset toteutettiin poikkeuksellisen hyvin. Johtopäätöksissä he ehdottavat, että perinteinen palaverikäytäntö tarkastuksissa on tehoton ja tulisi korvata jollain vaihtoehtoisella menetelmällä. Palaverilla on kuitenkin nähty olevan oma merkityksensä ainakin menetelmän käyttöönottovaiheessa. Esimerkiksi Laitenberger ja DeBaud (2000) esittävät, että organisaation olisi hyvä aloittaa perinteisellä palaveripohjaisella tarkastuksella ja miettiä vaihtoehtoisia toteutustapoja myöhemmässä vaiheessa, kun tarkastukset on saatu toimimaan säännöllisesti.

Mahdollinen ratkaisu: Ensimmäiseksi pitäisi tunnistaa syy tähän ongelmaan. Jos tarkastuksissa on jokin muu ilmeinen ongelma, kuten huono valmistautuminen, se on käsiteltävä ensin. Jos mitään tällaisia muita selittäviä tekijöitä pa-

laverin tehottomuudelle ei ole, voidaan ajatella jotain vaihtoehtoista toteutustapaa palaverin järjestämiselle. Tällaista päätöstä tehdessä on hyvä kuitenkin miettiä palaverin merkitys toiminnalle laajemmasta näkökulmasta. Vaikka palaveri ei kannattaisikaan virheiden löytämisen kannalta, se saattaa olla tärkeä väline oppimisessa ja tiedon jakamisessa. Joissakin organisaatioissa palaveri voi olla myös tärkeä elementti kurinalaisuuden ylläpitämisessä. Jokin toinen tapa koota valmistautumisen tulokset ei välttämättä luo samanlaista sosiaalista painetta tehdä valmistautumistyötä kuin fyysinen palaveri.

6.1.3 Huono tukimateriaali

Riittämätön tai huono tarkastuksen tukimateriaali voi vaikuttaa tarkastuksen tehokkuuteen. Materiaali voi olla huonosti suunniteltua tai se saattaa puuttua kokonaan. Tukimateriaalilla tarkoitetaan tässä yhteydessä tarkistuslistoja tai muuta vastaavaa materiaalia, jonka avulla ohjataan tarkastajia virheiden etsinnässä. Käytännössä tyypillinen tapa etsiä virheitä perustuu tarkistuslistoihin tai ad-hoc menettelyyn ilman mitään tukimateriaalia (Laitenberger ja DeBaud 2000). Laitenberger ym. (1999) kritisoivat tarkistuslistoja siitä, että ne ovat usein liian yleisiä. Tällaiset tarkistuslistat kertovat ainoastaan sen, millaisia virheitä dokumentissa voi olla, eikä opasta kuinka niitä voi löytää. Porter ja Votta (1998) totesivat, että hyvä tukimateriaali voi parantaa merkittävästi jopa kokeneiden ammattilaisten kykyä löytää virheitä.

Mahdollinen ratkaisu: Laitenberger ym. (2000) esittävät perspektiivipohjaista tarkastusta ratkaisuksi tähän ongelmaan. Menetelmä perustuu skenaarioihin, jotka ohjeistavat virheiden löytämisessä tavallisia tarkistuslistoja yksityiskohdaisemmin. Toinen ratkaisu voi löytyä työntekijöiden tukimateriaalin käyttöön keskittyvästä kouluttamisesta. Rifkin ja Daimel (1994) totesivat, että tarkastajien dokumentin lukutaitoon keskittyvä lyhytkin koulutus paransi huomattavasti virheiden löytymistä tarkastuksissa. Heidän koulutuksessaan jokainen osallistuja teki oman tarkistuslistan ja opetteli käyttämään sitä tehokkaasti. Yhteenvetona voidaan todeta, että hyvän tukimateriaalin käyttö voi parantaa merkittävästi tarkastusten tehokkuutta.

6.1.4 Osallistujat eivät ymmärrä tarkastusprosessia

Saattaa olla, että osallistujat eivät ymmärrä omaa rooliaan tarkastusprosessissa. Dokumentin tekijä saattaa olla tietämätön siitä, mitkä dokumentit tulee tarkastaa tai milloin ja miten tarkastus tulee toteuttaa. Wiegerson (1998) mukaan yleinen katselmointeihin liittyvän ymmärryksen puute voi johtaa epä johdonmukaisuuteen katselmointien toteutuksessa. Tällöin katselmointien tavoite, tiimin koostumus, katselmointimenetelmät, raportointi ja palaverikäytänteet saattavat vaihdella katselmoinnista toiseen. Wiegerson mukaan saattaa myös olla epäselvää, kuka johtaa palaveria ja palaveri saattaa ajautua käsittelemään toissijaisia asioita, kuten ratkaisemaan ongelmia tai arvostelemaan ohjelmointityyliä. Yhteenvetona hän toteaa, että tämänkaltaiset ongelmat vaikuttavat katselmointien

tehokkuuteen ja pahimmassa tapauksessa saattavat johtaa siihen, että katselmoiteja ei järjestetä ollenkaan.

Mahdollinen ratkaisu: Wiegertsin (1998) mukaan koulutus on paras keino varmistaa kaikkien osallistujien yhteinen ymmärrys katselmoitiprosessista. Myös Fagan (1986) korostaa artikkelissaan koulutuksen tärkeyttä. Hän jopa varoittaa ottamasta tarkastusmenetelmää käyttöön organisaatiossa ilman asianmukaista koulutusta. Faganin mukaan jotkut koulutuksen väliin jättäneet yritykset eivät ole onnistuneet kovin hyvin tarkastusten toteuttamisessa.

6.1.5 Kritiikki kohdistetaan tekijään

Wiegertsin (1998) mukaan ensimmäiset yritykset järjestää katselmoiteja saattavat joskus johtaa dokumentin tekijän kykyjen tai tyylin arvosteluun. Jos dokumentin tekijä kokee tullessa hyökkäyksen kohteeksi, saattaa hän jatkossa olla haluton tuomaan omia tuotoksiaan katselmoitavaksi. Pahimmassa tapauksessa hän saattaa odottaa tilaisuutta kostaa arvostelijoille silloin, kun hänen tuotoksiaan on katselmoitavana.

Mahdollinen ratkaisu: Wiegerts (1998) painottaa tässä tarkastuksen johtajan roolia, sillä hänen tulisi hallita tilanne palaverissa ja estää tällainen käyttäytyminen. Siis tähän kykenevät ja koulutetut tarkastusten johtajat voivat olla ratkaisu tällaisiin ongelmiin.

6.1.6 Huono tarkastusten suunnittelu

Wiegertsin (1998) mukaan monissa projekteissa katselmoinnit nähdään projektin tarkistuspisteinä (milestone), joiden kohdalla projekti siirtyy seuraavaan vaiheeseen. Jos katselmointi järjestetään juuri tarkistuspisteessä, ei aikataulussa ole enää aikaa virheiden korjaukselle. Se puolestaan saattaa johtaa projektin aikataulun venymiseen. Gilb (2000) suosittelee järjestämään tarkastukset ennen projektin tarkistuspisteitä, mutta sen lisäksi myös useampaan kertaan prosessin aikana.

Mahdollinen ratkaisu: Luonnollisesti ratkaisu tähän ongelmaan on suunnitella ja järjestää tarkastukset hyvissä ajoin. Se ei kuitenkaan välttämättä ole aivan yksinkertaista. Sekä projektipäällikön, että ylemmän tason johdon tulee ymmärtää tarkastusten merkitys prosessissa. Näiden ryhmien koulutus saattaa siis olla ratkaisu tähän ongelmaan.

6.1.7 Palaveri ajautuu epäoleellisiin asioihin

Jo Fagan (1976) painotti palaverissa keskittymistä virheisiin ja varoitti ylimääräisestä keskustelusta tarkastuspalaverissa. Myös Gilb ja Graham (1993) ovat tästä samaa mieltä. Jos palaverissa sallitaan ongelmien ratkaisua ja muuta keskustelua, se voi helposti viedä koko ajan. Tästä ollaan kirjallisuudessa laajasti

yhtä mieltä, mutta kuitenkin esimerkiksi Johnson (1998) on kritisoinut keskustelun kieltämistä. Hänen mukaansa oppiminen on palaverin merkittävin hyöty ja ongelmanratkaisu ryhmässä on tärkeää oppimisen kannalta. Johnson esittää samassa artikkelissa, että virheiden kokoamiseen tarkastuksessa käytettäisiin jonkinlaista asynkronista työkalutuettua menetelmää fyysisen palaverin sijaan. Joka tapauksessa vaikuttaa vaikealta sijoittaa tehokasta virheiden kirjaamista ja keskustelua samaan palaveriin. On hyvä, jos tuotoksen tekijä saa tiimiltä apua ongelmien ratkaisussa, mutta se ei voi tapahtua tarkastuspalaverissa, jonka tulisi keksittyä löytyneiden virheiden kirjaamiseen.

Mahdollinen ratkaisu: Tässäkin kohdassa koulutus saattaa olla ratkaisu. Tarkastuksen johtajat ovat keskeisessä roolissa ja heidän tulisi olla koulutettuja. Myös muiden osallistujien kouluttaminen saattaa auttaa asiaa. Jos palaverin tarkoitus ja sitä koskevat ohjeet ovat selviä kaikille, on pienempi riski ajautua palaverissa epäoleellisiin asioihin.

6.1.8 Huono valmistautuminen

Gilbin ja Grahamin (1993) valmistautuminen on tarkastusprosessin kriittisin osa. Heidän mukaansa tarkastus ilman valmistautumista saattaa löytää vain 10 % virheistä, jotka voidaan löytää hyvällä valmistautumisella. Huonoon valmistautumiseen voi organisaatiossa olla monta syytä. Johtotasolla voi olla kiusaus säästää tarkastusten vaatimista resursseista tai työntekijät voivat kokea itsensä liian kiireisiksi valmistautuakseen kunnolla.

Mahdollinen ratkaisu: Wiegiers (1998) esittää, että jokaisen osallistujan valmistautumiseen käyttämä aika tulisi kirjata muistiin. Se ei kuitenkaan välttämättä takaa hyvää valmistautumista, jos ei kukaan konkreettisesti valvo kerätty tietoa. Jollakin henkilöllä organisaatiossa tulisi olla vastuu tarkastuskäytänteiden toteutumisesta ja työntekijöiden motivoinnista. Organisaatiolla olisi myös hyvä olla määriteltynä suositukset valmistautumisajat erilaisille tarkastuksille.

6.1.9 Tarkastuksiin osallistuu vääriä ihmisiä

Wiegiers (1998) toteaa, että katselmointeihin saattaa joskus osallistua ihmisiä, joilla ei ole virheiden löytämiseen tarvittavaa kykyä tai tietämystä. Kuitenkin osallistujien henkilökohtainen osaaminen vaikuttaa olevan suurin yksittäinen tekijä tarkastuksen tehokkuudessa (Sauer ym. 2000).

Mahdollinen ratkaisu: Tämä saattaa olla hankala hallinnollinen ongelma, sillä parhaimman osaamisen omaavat työntekijät eivät ole aina käytettävissä. Kuitenkin kouluttamalla voidaan nostettua yleistä osaamistasoa. Rifkin ja Daimel (1994) saivat hyviä tuloksia koulutuksesta, jossa keskityttiin virheiden löytämiseen. Koulutuksen jälkeen tarkastuksissa löydettyjen virheiden määrä nousi huomattavasti paremmalle tasolle kuin kontrolliryhmissä, joiden koulutus käsitteli tarkastuksia prosessitasolla.

6.1.10 Roolit saattavat puuttua

Jossain määrin erilaisten roolien määrittäminen eri tarkastajille on kuulunut tarkastusmenetelmään aivan alusta lähtien (Fagan 1976). Porterin ym. (1995) mukaan kuitenkin käytännössä osallistujat tekevät työnsä täysin samalla tavalla silloin, kun käytetään ad-hoc lähestymistapaa tai tarkistuslistoja virheiden etsimisessä. Tällöin osallistujat tekevät paljon päällekkäistä työtä ja se sillä saattaa olla vaikutusta tarkastusten tehokkuuteen.

Mahdollinen ratkaisu: Tämä on luonnollisesti organisointikysymys. Porter ym. (1995) esittävät, että tarkastajilla tulisi olla erilaisten prosessitason roolien lisäksi myös erilaisia virheiden etsintään liittyviä tehtäviä. He esittävät tutkimuksensa ratkaisuksi skenaariopohjaista menetelmää, jossa tietynlaisten skenaarioiden avulla ohjataan tarkastajia heidän tehtävissään. Skenaarioiden avulla hoidetaan tehtävänjako niin, että osallistujat etsivät dokumentista erityyppisiä virheitä. Porter ym. (1995) totesivat tutkimuksessaan, että skenaariopohjaisella tekniikalla löydettiin enemmän virheitä kuin ad-hoc lähestymistavalla tai tarkistuslistojen avulla.

6.1.11 Huonolaatuinen dokumentti vie turhaan aikaa

Kelly ja Shepard (2002) viittaavat artikkelissaan tähän ongelmaan. He tutkivat erityisesti koodin katselmointia ja totesivat, että tarkastusten kannalta selkeästi kirjoitettu ja hyvin dokumentoitu koodi on erittäin tärkeää. Huonot dokumentit haaskaavat tarkastajien aikaa. Gilb ja Graham (1993) suosittelevat, että liian huonolaatuista dokumenttia ei tulisi viedä lainkaan tarkastettavaksi, vaan sen laatua tulee ensin parantaa kelvolliselle tasolle.

Mahdollinen ratkaisu: Hyvät organisaatiotasoiset ohjeet ja standardit auttavat dokumentin laatijaa tekemään laadukasta työtä. Myös hyvät tarkistuslistat voivat auttaa jo dokumenttia tehdessä. Gilb ja Graham (1993) kehottavat laatimaan selkeät esiehdot tarkastukselle. Nämä ehdot toimivat ohjeena jo dokumentin tekijälle. Tarkastusta ennen tarkastuksen johtaja tarkistaa esiehtojen täyttymisen ennen kuin käynnistää tarkastusprosessin.

6.1.12 Huono asenne

Dokumentin tekijä saattaa olla haluton tuomaan tuotoksiaan tarkastukseen, ellei ole pakko. Tämä on luonnollisesti asennekysymys. Monissa yrityksissä tarkastukset vaaditaan prosessissa järjestettäväksi. Kuitenkin, jos tarkastus koetaan ainoastaan prosessin vaatimaksi inhottavaksi tehtäväksi, eivät osallistujat ole motivoituneita tekemään työtä kunnolla. Se puolestaan vaikuttaa varmasti tarkastuksen tehokkuuteen.

Mahdollinen ratkaisu: Tässäkin kohdassa koulutus saattaa auttaa. Osallistujien tulisi selkeästi ymmärtää tarkastusten merkitys heidän oman työnsä kannalta.

Parhaassa tapauksessa työntekijät näkevät tarkastukset tilaisuutena, jossa tasa-puolisesti autetaan toisia ja saadaan apua omaan työhön. Myös alemman johdon tulisi ymmärtää tämä, sillä heillä on tärkeä rooli tällaisen ilmapiirin rakentamisessa organisaatiossa.

6.1.13 Liian paljon materiaalia

Gilb ja Graham (1993) varoittavat tarkastamasta liikaa materiaalia yhdessä tarkastuksessa, sillä heidän mukaansa silloin tarkastuksen teho laskee. Heidän mukaansa optimaalinen aika on usein noin sivu materiaalia tunnissa. He väittävät, että todella merkittävät virheet voidaan löytää vain, jos tarkastukseen käytetään riittävästi aikaa.

Mahdollinen ratkaisu: Materiaalin määrässä on usein kysymys asenteista, joita voi olla vaikea muuttaa. Vie paljon aikaa ja vaivaa rakentaa organisaatioon uutta kulttuuria, joka osaa arvostaa tarkastuksia organisaation kaikilla tasoilla. Aina ei ole välttämättä tarpeen tarkastaa huolellisesti dokumentaation jokaista palasta. Esimerkiksi Gilb ja Graham (1993) ehdottavat, että laajoista dokumenteista voisi tarkastaa otoksia, joiden perusteella arvioidaan dokumentin laatu.

6.1.14 Tarkastuksiin ei ole varattu resursseja

Usein tarkastukset on määriteltävä kuuluviksi laajempiin tehtäväkokonaisuuksiin, kuten esimerkiksi vaatimusmäärittelyyn. Silloin tarkastuksiin ei ole tavallisesti varattu erikseen resursseja. Jos tarkastuksiin ei ole erikseen varattua aikaa, voivat työntekijät kokea se ylimääräiseksi työksi, johon ei ole aikaa. Se puolestaan johtaa huonoon valmistautumiseen ja tehottomiin tarkastuksiin.

Mahdollinen ratkaisu: Resurssien varaaminen on hallinnollinen asia, joka on eräs näkyvä signaali johdon sitoutumisesta. Tässä kysymyksessä projektipäälliköt ovat avainasemassa ja heidän kouluttamisensa voi auttaa asiaa. Organisaatiossa olisi myös hyvä olla projektisuunnitteluvaiheessa sellaiset käytänteet ja standardit, jotka ottavat tarkastukset huomioon.

6.2 Kokemukset yrityksissä

Tämä kohta käsittelee kohdeorganisaatioista kerättyjä katselmointien ongelmiin liittyviä kokemuksia. Organisaatioissa tehdyissä haastatteluissa käytettiin yleisesti termiä katselmointi, koska se on käsitteenä tarkastusta tutumpi. Käytännössä organisaatioiden katselmointikäytännöt useimmissa tapauksissa muistuttivat hyvin paljon tarkastusprosessia. Katselmointi määriteltiin haastattelun aluksi tarkoittamaan virheiden etsintään keskittyvää toimintaa, johon sisältyi oletus ihmisten valmistautumisesta. Siis esimerkiksi spontaanisti tapahtuva koodin läpikäyminen ei tämän rajauksen mukaan ole katselmointia. Koska

haastatteluissa käytettiin käsitettä katselmointi, käytetään sitä myös tässä kohdassa haastattelujen tuloksista raportoidessa.

Tässä raportoidut kokemukset kerättiin haastatteluilla kuudessa eri ohjelmistojen tuottavassa organisaatiossa. Aluksi haastateltaville esitettiin avoin kysymys siitä, mitkä ovat heidän mielestään keskeisimmät tarkastuksiin liittyvät ongelmat organisaatiossa sillä hetkellä. Sen jälkeen heidän kanssaan käytiin läpi edellisessä kohdassa (6.1) esitetty lista mahdollisista tarkastuksiin liittyvistä ongelmista. Haastateltavilta kysyttiin, kuinka tutulta mainittu ongelma tuntuu omassa työssä. Tuttuutta pyydettiin kuvaamaan numeerisesti asteikolla 1-5 (5 on erittäin tuttu). Empiirisen tutkimuksen toteutus on paremmin esitelty luvussa 4.

Haastatteluissa selvisi, että esitettyihin kysymyksiin oli vaikeaa saada yksiselitteisiä vastauksia. Jokainen tulkitsee kysymyksessä esitettyä mahdollista ongelmaa omalla tavallaan oman työnsä kontekstissa. Siksi jopa samassa yrityksessä eri tehtävissä työskentelevät ihmiset vastasivat samassa kohdassa hiukan eri asioihin. Siitä syystä tässä kohdassa raportoiduissa kokemuksissa ei painoteta haastateltavien vastausten numeerisia arvoja. Sen sijaan tässä keskitytään keräämään yhteen kokemuksia eri näkökulmista esitettyihin ongelmiin.

Ensin jokaiselle haastateltavalle siinä esitettiin avoin kysymys siitä, mitkä ovat heidän mielestään keskeisimmät katselmointien ongelmat omassa organisaatiossa. Vastaukset luonnollisesti vaihtelivat organisaatiosta ja vastaajasta riippuen, mutta seuraavat kolme asiaa olivat tavallisimpia vastauksia:

- Kiire, erityisesti osaavilla ihmisillä
- Heikko valmistautuminen
- Kaivataan formaalimpaa toimintaa

Näistä vastauksista kiire mainittiin lähes jokaisessa organisaatiossa. Erityisenä ongelmana koettiin parhaiden osaajien aikapula. Kiire johtaa organisaatioissa monenlaisiin ongelmiin. Tavallisin ongelma vaikuttaa olevan seuraavana listalla mainittu valmistautuminen, joka tosin voi johtua muustakin kuin kiireestä. Kiireessä ihmiset eivät ennätä valmistautumaan palaveriin ja silloin dokumenteista löytyy vain vähän virheitä. Joissakin organisaatioissa kiire johtaa usein siihen, että katselmointi jätetään kokonaan tekemättä.

Monessa organisaatiossa tuli esille se, että katselmointeihin kaivataan formaalimpaa toimintaa. Käytännössä tämä tarkoittaa sitä, että monet kaipasivat enemmän katselmointeja ja kurinalaisuutta niiden toteutukseen. Yllättäen tällaisia vastauksia tuli työntekijätasolta yhtä paljon kuin laatupäälliköitä. Tavallinen käsitys laatupäälliköilläkin oli, että työntekijät olisivat hiukan vastahakoisia nykyistä formaalimpaan toimintaan. Haastateltavia pyydettiin arvioimaan millaista muutosvastarintaa organisaatiossa on odotettavissa, jos toimintaan kehitetään formaalimpaan suuntaan. Työntekijät arvioivat useassa yrityksessä odotettavissa olevan vastarinnan pienemmäksi kuin laatupäälliköt.

Seuraaviin alakohdissa esitellään lyhyesti organisaatioiden kokemukset edellisessä kohdassa (6.1) esiteltyihin mahdollisiin ongelmatilanteisiin.

6.2.1 Palaverin aikataulutus voi aiheuttaa viivettä

Pisteiden keskiarvoa laskettaessa tämä kohta sai kolmanneksi korkeimmat pisteet. Käytännössä se tarkoitti, että aikataulun viivästymistä katselmointien vuoksi tapahtuu silloin tällöin. Kuitenkaan myöhästyminen ei ole niin suurta, että tätä olisi pidetty organisaatioissa kovin merkittävänä ongelmana.

Tämän kohdan tulkinnessa oli jonkin verran eroja eri haastateltavien välillä. Alkuperäinen ajatus ongelmasta Vottan (1993) mukaan oli, että ihmisten aikataulujen järjestäminen yhteistä palaveria varten aiheuttaa viivettä. Osa ongelmista kohdeorganisaatioissa johtui kyllä tästä, mutta ainakin yhtä suuri ongelma organisaatioissa oli materiaalin saaminen ajoissa katselmoitavaksi. Osassa organisaatioista on tavallista jättää katselmointi väliin ennemmin kuin myöhästyttää sen vuoksi aikataulua.

6.2.2 Palaveri kuluttaa resursseja, mutta siellä ei löydetä paljon virheitä

Tämä osoittautui hankalaksi asiaksi lähestyä. Suurin osa haastateltavista ei osannut sanoa tähän juuri mitään. Jotkut haastateltavat arvioivat, ettei palaverissa löydy kovin paljon virheitä, mutta näkivät palaverit muuten hyödyllisinä. Näiden kokemusten perusteella voisi suositella, ettei palaverin järjestämisen tarpeellisuutta asetettaisi kyseenalaiseksi ainakaan ennen kuin määrämuotoisten katselmointien järjestämisestä on riittävän vahva kokemus. Tätä tukevat ainakin Laitenberger ja DeBaud (2000) joiden mukaan organisaation kannattaisi aloittaa tarkastusten järjestäminen perinteisellä tavalla ja miettiä vaihtoehtoisia toteutustapoja myöhemmin. Vottan (1993) tutkimuksessa, johon tämä ongelma viittaa, kohteena oli tarkastuskäytänteissä kehittynyt organisaatio, joka saattaa kyetä pohtimaan palaverin merkitystä perusteellisesti.

Tässä kohdassa on huomattava, että myös löydettyjen virheiden määrään voi vaikuttaa useampi tekijä. Votta (1993) olettaa, että prosessi toimii muuten moitteettomasti ja vikaa on ainoastaan perinteisessä tarkastusmenetelmässä, joka ei hänen mukaansa ole tehokas. Tässä tutkimuksessa palaveriin valmistautuminen todettiin merkittävimmäksi ongelmaksi kohdeorganisaatioiden katselmoinneissa. Jo huono valmistautuminen selittää sen, jos näissä organisaatioissa palaverissa löytyy vähän virheitä.

6.2.3 Huono tukimateriaali

Myös tämä oli eräs vaikeista kohdista haastatteluissa. Vaikeus johtui siitä, että yhdelläkään organisaatioista ei ollut minkäänlaisia tarkistuslistoja tai muuta vastaavaa materiaalia käytössä. Yhdellä organisaatiolla oli jonkinlaisia tarkistuslistoja, mutta ainakaan haastateltavat eivät olleet kuulleet niitä käytettävän käytännössä. Yhtä lukuun ottamatta kohdeorganisaatioissa oli olemassa määritetty katselmointiprosessi, mutta myös tietoisuus siitä oli aika puutteellista. Monet haastateltavat tiesivät, että sellainen ohje on olemassa, mutta eivät osanneet sanoa, mitä se pitää sisällään.

Tarkistuslistojen tai muun virheiden löytämisen tueksi tarkoitettua materiaalin käyttö oli siis tuntemattomia haastateltaville. Mielenpitoet tällaisen materiaalin tarpeellisuudesta vaihtelivat laidasta laitaan. Osa oli sitä mieltä, että kyllä ihmiset osaavat toimia ilman tarkistuslistojakin. Toiset taas pitivät tällaista materiaalia hyödyllisenä. Tämän kohdan vakavampi kommentointi olisi siis selvästi vaatinut haastateltavilta enemmän pohjatietoa ja kokemusta.

6.2.4 Osallistujat eivät ymmärrä tarkastusprosessia

Prosessin ymmärtämiseen liittyvät ongelmat jakaantuivat tässä oikeastaan kahteen eri kategoriaan: katselmointien toteutukseen ja niiden organisointiin tuotantoprosessissa. Osallistujat saattavat olla siis tietämättömiä siitä, mitä dokumentteja täytyy katselmoida, kuinka ne tulee toimittaa katselmoitavaksi. Toisaalta voidaan olla tietämättömiä siitä, miten katselmointi organisoidaan, mitä sen aikana tulee tehdä, kuinka palaverissa tulee käyttäytyä jne.

Yleisesti tämä ei ollut mitenkään merkittävä ongelma. Kohdeorganisaatioissa koettiin, että osallistujat tietävät melko hyvin, kuinka tulee toimia katselmoineissa. Joissakin organisaatioissa koettiin jossain määrin ongelmaksi ihmisten tietämättömyys katselmoineista yleensä ja niiden katselmointien organisoinnista tuotantoprosessissa. Työntekijä ei esimerkiksi välttämättä ymmärrä selvästi katselmointien merkitystä tai tiedä, mitä hänen tuotoksistaan tulee katselmoida.

6.2.5 Kritiikki kohdistetaan tekijään

Tämä kohdan haastateltavat kokivat kaikkein vähiten ongelmalliseksi. Useimmat sanoivat, etteivät ole koskaan havainneet mitään ongelmaa tässä. Tämä oli kirjallisuuden perusteella hiukan yllättävää, sillä ainakin kaikissa aiheita käsittelevissä oppikirjoissa muistetaan varoittaa painavasti tuotoksen tekijän kritisoimisesta. Kulttuuritekijöillä saattaa olla tässä kohdassa oma osansa, mutta joka tapauksessa missään näistä kohdeorganisaatioista tällaisia ongelmia ei ollut koettu.

6.2.6 Huono tarkastusten suunnittelu

Haastateltavat eivät pitäneet suunnittelua kovinkaan merkittävänä ongelmana. Tavallisesti katselmoinnit suunnitellaan kyllä järjestettäväksi ajoissa, mutta ongelma on ennemminkin katselmoitavan materiaalin valmistuminen ajoissa. Katselmoinnin järjestäminen saattaa siis venyä katselmoitavan tuotoksen valmistumista odotellessa. Ongelmia voi tuottaa se, että tällaisessa tilanteessa katselmoinnit saatetaan järjestää liian nopealla aikataululla, jolloin osallistujilla ei ole riittävästi aikaa valmistautua. Joissakin organisaatioissa katselmointi saattaa jäädä aikataulun kiristyessä kokonaan tekemättä.

Tarkastusten suunnittelun ongelmiin voidaan myös yhdistää työntekijöiden kiire varsinkin jonkin projektin tarkistuspisteen lähestyessä. Esimerkiksi Chatzigeorgiou ja Antoniadis (2003) havaitsivat eräässä yrityksessä, että työ-

määrän huiput projektissa yleensä ja katselmoinneissa osuivat samoille viikoille. Tätä ei suoranaisesti kysytty haastatteluissa, mutta merkkejä tällaisesta oli havaittavissa. Jotkut haastateltavat kommentoivat erityisesti kokeneempien työntekijöiden aikapulan aiheuttavan ongelmia.

6.2.7 Palaveri ajautuu epäoleellisiin asioihin

Haastateltavat arvioivat tämän keskiarvoissa mitattuna toiseksi tavallisimmaksi ongelmaksi. Tätä ei kuitenkaan pidetty kovin suurena ongelmana yhdessäkään organisaatioissa. Vaikka katselmointien järjestäminen ei jossakin organisaatioissa olisikaan kovin säännöllistä, ovat työntekijät kuitenkin tottuneet yleisiin kokouskäytänteisiin. Siis tavallisesti asia saadaan kuitenkin käsiteltyä, vaikka keskustelu välillä hairahtuisikin sivupoluille.

Asenne sivuraiteille ajautuvaa keskustelua kohtaan vaihteli selvästi haastateltavan mukaan. Toiset pitivät keskustelua mahdollisista ongelman ratkaisuista hyödyllisenä, eivätkä haluaisi missään nimessä karsia sellaista keskustelua pois katselmointipalaverista. Varsinaisesti haitalliseksi tällaista keskustelua ei tavallisesti koettu, mutta jotkut haastateltavat pitivät sitä tehottomana toimintana, johon koko ryhmältä kuluu turhaan aikaa.

6.2.8 Huono valmistautuminen

Haastateltavat pitivät valmistautumista selvästi tavallisimpana ja merkittävimpänä ongelmakohtana katselmoinneissa. Syyt heikkoon valmistautumiseen voivat organisaatioissa olla monenlaiset. Eräs suurimmista syistä on varmasti ihmisten heikko motivaatio lukea toisten tuotoksia. Organisaatioissa eletään jatkuvasti tiukkojen aikataulujen kanssa ja ”omat työt” menevät usein toisten töiden lukemisen edelle. Tällaiseen kokemukseenkin vaikuttavat tietysti monet asiat kuten organisaatiokulttuuri, tietämys katselmointien merkityksestä ja se, onko katselmoiteja varten erikseen varattu resurssit työsuunnitelmissa.

Haastattelujen pohjalta vaikuttaa siltä, että toiminnan konkreettiselle valvonnalle on organisaatioissa tarvetta. Sellaisissakin organisaatioissa, joissa toteutetaan säännöllisesti määrämuotoisia tarkastuksia, oli ongelmia valmistautumisen kanssa. Parissa organisaatioissa kerättiin tietoa valmistautumiseen käytetystä ajasta, mutta niissäkään tiedolla ei tehty mitään. Työntekijät voivat siis käytännössä jättää jatkuvasti valmistautumisen vähälle ilman minkäänlaista kontrollia. Vaikuttaa siis siltä, että työntekijät eivät oma-aloitteisesti valmistaudu katselmoiteihin riittävän hyvin. Tähän tarvitaan katselmointien todellisen toteutumisen valvontaa esimerkiksi laatupäällikön toimesta.

6.2.9 Tarkastuksiin osallistuu väriä ihmisiä

Suurin osa haastateltavissa ei nähnyt tässä kohdassa minkäänlaista ongelmaa omassa organisaatioissa. Harvat tähän kohtaan saadut kommentit viittasivat teknisen osaamisen puutteeseen. Joissakin tapauksissa nimenomaan teknisiä dokumentteja katselmoitaessa, jolloin kaikilla osallistujilla ei välttämättä ole

riittävän syvällistä teknistä tietämystä. Erityisesti sellaisessa tilanteessa, kun ollaan tekemisissä uuden tekniikan kanssa, voi olla vaikeaa löytää teknisesti riittävän tietämyksen osaavia katselmoijia. Hankalimmassa tapauksessa tuotoksen tekijä on ainoa, joka hallitsee käytetyn tekniikan.

Projektiliiketoiminnassa katselmoiteja tehdään usein myös asiakkaiden kanssa. Haastattelujen perusteella projektitoiminnassa asiakkaan tekninen osaaminen saattaa olla vaihtelevaa. Joskus puutteellinen tekninen tietämys asiakkaan teknisessä tietämyksessä saattaa aiheuttaa hankaluuksia. Kuitenkin yleisesti ottaen väärin ihmisten osallistuminen katselmoiteihin koettiin hyvin pieneksi ongelmaksi.

6.2.10 Roolit saattavat puuttua

Roolien puuttumista ei useimmissa tapauksissa koettu ongelmaksi. Tämä oli haastatteluissa eräs vaikeasti lähestyttävissä olevia asioita, sillä tässä tarkoitettua selkeää työnjakoa virheiden etsinnässä ei ole organisaatioissa aiemmin ajateltu. Useat haastateltavat totesivat, että tällainen työnjako saattaisi olla hyvä ajatus, mutta eivät osanneet kommentoida asiaa sen tarkemmin.

Tällaista asiaa ei ehkä yleensä kukaan mielletä ongelmaksi toiminnassa, vaan ehkä potentiaaliseksi kehittämiskohteeksi. Tämän asian pohtiminen organisaatioissa vaatisi selvästi enemmän tietämystä katselmoineista. Lisäksi organisaatioilla olisi ehkä hyvä olla nykyistä vakiintuneemmat katselmoitikäytänteet ennen kuin tällainen katselmoitien kehittäminen on ajankohtaista.

6.2.11 Huonolaatuinen dokumentti vie turhaan aikaa

Tämä oli eräs vähiten ongelmaksi koetuista kohdista organisaatioissa. Huonolaatuisia dokumentteja tuodaan haastateltavien mukaan useimmissa organisaatioissa katselmoitavaksi hyvin harvoin. Jotkut haastateltavat kommentoivat, että dokumentit saisivat olla paremmin viimeistelyjä, mutta eivät pitäneet sitä mitenkään suurena ongelmana.

Tähän vaikuttaa selvästi organisaation toimintakulttuuri. Joissakin organisaatioissa katselmointi saatetaan jättää tekemättä, jos tuotos ei näytä valmistuvan ajoissa. Joissakin organisaatioissa tätä taas ei koettu ongelmaksi, koska materiaalin ollessa hiukan keskeneräinen, voidaan pitää yhteinen työstöpalaveri ja katselmoida tuotos myöhemmin. Se taas ei onnistu millään sellaisessa toiminnassa, jonka kannalta on kriittistä, että katselmointi saadaan järjestettyä juuri tietynä päivänä.

6.2.12 Huono asenne

Tämä oli myös eräs vähiten ongelmana pidetyistä kohdista. Lähes poikkeuksetta haastateltavat kommentoivat, että työntekijät päinvastoin haluavat palautetta omasta työstään. Monessa organisaatiossa erityisesti työntekijätason haastateltava (esim. ohjelmistosuunnittelija) kaipasi lisää ja muodollisempia katselmoiteja.

Lähes kaikki haastateltavat olivat yhtä mieltä siitä, että asenteissa hankalampi ongelma on saada työntekijät lukemaan toisten tuotoksia. Kuten jo useampaan kertaan aiemmin on todettu, valmistautumista pidettiin organisaatioissa katselmointien merkittävimpänä ongelmana. Useimmissa organisaatioissa suurin syy puutteelliseen valmistautumiseen löytyy työntekijöiden asenteissa. Syy ei välttämättä ole ainoastaan katselmoijien asenteessa, vaan luonnollisesti asenteet kaikilla organisaation tasoilla vaikuttavat suoriutumiseen. Jos esimerkiksi projektipäällikkö ei arvosta katselmoiteja, ei niihin todennäköisesti varata riittävästi resursseja.

6.2.13 Liian paljon materiaalia

Tätä ei koettu kovin merkittäväksi ongelmaksi joitakin yksittäisiä haastateltavia lukuun ottamatta. Organisaatioista ei tässä tutkimuksessa varsinaisesti selvitetty tavallista katselmoitavan materiaalin. Kuitenkin haastatteluissa käytyjen keskustelujen perusteella voidaan arvioida, että tavallinen määrä katselmoitavaa materiaalia oli jokaisessa organisaatiossa todella paljon kirjallisuuden esittämiä suosituksia suurempi. Lisäksi haastateltavien arvioima tavallinen valmistautumiseen käytetty aika oli melko pieni. Tyypillinen arvio oli noin puoli tuntia.

Tämä vaikuttaa siis olevan selvästi sellainen kysymys, joka vaatisi enemmän tietoa ja kokemusta katselmoineista. Katselmoiteihin satsataan resursseja hyvin varovasti ja yksikään kohdeorganisaatioista ei ole yrittänyt tunnistaa konkreettisesti niiden vaikutusta tuotantoprosessiin. Jos organisaatiot yrittäisivät mitata katselmointien vaikutusta, saattaisivat ne ainakin kirjallisuuden raportoimien tutkimustulosten perusteella päätyä käyttämään enemmän aikaa katselmoiteihin.

6.2.14 Tarkastuksiin ei ole varattu resursseja

Yhdessäkään kohdeorganisaatiossa tarkastuksiin ei ole selkeästi varattu resursseja työntekijöiden työsuunnitelmiin. Kuitenkin mielipiteet tästä kohdasta vaihtelivat melko paljon. Kahdessa organisaatiossa haastateltavat pitivät tätä eräänä merkittävimmistä ongelmista. Toisaalta yhdessä yrityksessä tämä oli selvästi vähiten ongelmaksi koettujen asioiden joukossa.

Tässä on hyvä huomata, että resurssien varaaminen yksistään ei ole ratkaisu tähän liittyviin ongelmiin. Jos organisaatiokulttuuri ei pidä tarkastuksia tärkeänä, voidaan ne tehdä edelleen huonosti, vaikka niihin muodollisesti olisi varattu resurssit olemassa. Siis jos tarkastuksia ei arvosteta, löytyy valmistautumisen sijasta aina jotain tärkeämpää tekemistä.

6.3 Yhteenveto tarkastusten ongelmista

Tässä luvussa on käsitelty katselmointien toteuttamiseen liittyviä ongelmia sekä kirjallisuuden että kohdeorganisaatioissa tehdyn haastattelututkimuksen perusteella. Tarkoituksena on ollut kuvata ongelmakenttää ja tunnistaa, mitkä ongelmat ovat todellisia tämän päivän ohjelmistotuotannossa.

Haastateltavia pyydettiin arvioimaan, kuinka tavallisia ja merkittäviä ovat tämän luvun kohdassa 6.1 esitetyt mahdolliset ongelmat heidän omassa organisaatiossaan. Kolme haastateltavien merkittävimmäksi koettua ongelmaa olivat:

- Huono valmistautuminen
- Palaveri ajautuu epäoleellisiin asioihin
- Palaverin aikataulutuksen voi aiheuttaa viivästystä

Mainituista tavallisimmiksi kuvatuista ongelmista valmistautuminen kuvattiin selvästi merkittävimmäksi. Palaverin ajautuminen epäoleellisiin asioihin ja palaverin aikataulutuksen aiheuttama viivästys olivat asioita, joita tapahtuu silloin tällöin, mutta niitä ei koettu kuitenkaan kovin merkittävästi toimintaa vaikeuttaviksi asioiksi.

Mielenkiintoista tuloksissa oli se, että ainakin näissä kohdeorganisaatioissa ongelmat olivat samansuuntaisia. Organisaatioiden välillä oli kuitenkin huomattavia eroja katselmointien toteutuksen systemaattisuudessa. Tämä tukee etukäteen tehtyä hypoteesia siitä, että myös käytänteissään kehittyneessä ja tarkastuksia säännöllisesti toteuttavassa organisaatioissa voi olla joitakin merkittävästi tehokkuuteen vaikuttavia ongelmia.

Haastatteluissa kävi selväksi, että koko organisaation tilaa on vaikeaa kuvata kolmen haastattelun tuottamien vastausten perusteella. Yksilöiden väliset erot monissa kysymyksissä olivat suuria. Eri työntekijät tulkitsivat jokainen kysymyksiä oman tietonsa ja ennakoasenteensa pohjalta. Saattoi olla, että haastateltavien joukossa oli joku, jonka mielestä kaikki toimii lähes täysin ongelmattomaksi. Kuitenkin toinen haastateltava samasta organisaatiosta saattoi nähdä merkittäviä ongelmia usealla osa-alueella. Myös erot eri projektiryhmien toiminnassa samassa organisaatioissa saattavat selittää vastausten eroja, mutta varmasti suurin osa eroista johtuu tietoon ja ennakoasenteeseen perustuvista tulkintaeroista.

Jotkut haastatteluissa esitetyt mahdolliset ongelmakohdat olivat selvästi sellaisia, jotka olisivat vaatineet haastateltavilta enemmän tietoa ja kokemusta. Esimerkiksi roolijaon merkitystä virheiden etsinnässä ei kukaan haastateltavista ollut ajatellut aiemmin ja siksi he eivät kokeneet ongelmana. Sama asia koski tukimateriaalin käyttöä virheiden etsinnässä. Yksikään organisaatioista ei käyttä katselmoinneissa tällaista materiaalia, mutta haastateltavat eivät myöskään kokeneet sitä ongelmana. Kokemus esimerkiksi hyvin laadittujen tarkistuslistojen käytöstä saattaisi muuttaa vastauksia.

Monilla katselmointien toimintaan liittyvillä ongelmilla on riippuvuus-suhde toisiinsa. Vastausten tulosten tulkintaa vaikeuttaa erityisesti sellainen kohta, jossa tietty ongelma estää toisen esiintymisen tai jonkin ongelman realisoitumiselle on olemassa jonkinlainen esiehto. Esimerkkinä voisi olla katselmointipalaverin järjestelyn aiheuttama viivästys, jota parissa kohdeorganisaatioissa ei haastateltavien mukaan tapahdu kovin usein. Syy oli se, että näissä organisaatioissa tavallisempi ratkaisu aikataulun kiristymiseen on jättää katselmointi pitämättä kuin viivästyttää sen järjestämistä.

Eräs haastattelututkimuksen taustalla oleva ajatus oli, että tässä luvussa esitetyn kaltaista ongelmalistaa voisi käyttää katselmointikäytänteiden kehittämisen apuvälineenä. Muutaman haastattelun avulla organisaatioissa voitaisiin kartoittaa todelliset katselmointien ongelmakohdat, jotta voidaan keskittää kehittämistoimenpiteet oikeisiin asioihin ja parantaa siten toimintaa mahdollisimman kustannustehokkaasti. Johtopäätös haastattelujen jälkeen oli, että ongelmalista voi toimia apuvälineenä, mutta haastateltavien antamat vastaukset ovat itsessään liian rajallinen tietolähde kehittämiskohteiden valintaa varten.

Yhteenvetona haastattelujen tuloksista voidaan sanoa, että niiden määrällisiin tuloksiin ei kannata kiinnittää liian suurta huomiota. Edellä mainittujen tulosten tulkintaan liittyvien ongelmallisuuksien kysymyslistan avulla ei voida tunnistaa kaikenlaisia ongelmia. Esimerkiksi organisaation asenteen mittaaminen on hankala tehtävä, koska se ei välttämättä näy ulkonaisina tunnusmerkeinä käytänteissä. Haastattelutulosten lisäksi tarvitaan siis runsaasti katselmointien erityisasiantuntemusta ja organisaation nykytilan tulkintaa, jotta saadaan perustellusti valittua tärkeimmät kehittämiskohteet katselmointikäytänteissä.

7 JOHTOPÄÄTÖKSET

Tässä tutkimuksessa on käsitelty tarkastusten toteuttamiseen liittyviä puutteita ja ongelmia aiemman kirjallisuuden sekä kuudessa kohdeorganisaatiossa toteutetun haastattelututkimuksen avulla. Työn teoreettinen pohja rakentuu aiemman kirjallisuuden varaan, jota on esitelty kattavasti luvussa 2. Eräs aiemman kirjallisuuden analysoinnin pohjalta syntynyt johtopäätös oli, että alalla on tehty melko vähän tutkimusta todellisessa yritys-elämän kontekstissa. Tämä tutkimus on omalta osaltaan pyrkinyt toimimaan juuri sillä alueella.

Tutkimuskysymystä on lähestytty kahdesta eri näkökulmasta. Ensimmäiseksi tutkimusta varten luotiin luvussa 3 esitelty tarkastuskäytänteiden kypsyttä arvioiva ICM-malli, jonka avulla tarkasteltiin olemassa olevia käytänteitä kohdeorganisaatiossa (luku 5). Toinen näkökulma tutkimusongelmaa oli tarkastusten toteuttamiseen liittyvät ongelmat, joita tarkasteltiin sekä teoriassa että käytännössä kohdeorganisaatioissa (luku 6).

Sekä aiempien tutkimusten (Johnson 1998 ja Ciolkowski ym. 2003) että kohdeorganisaatioissa toteutettujen haastattelujen perusteella voidaan sanoa, että tarkastusten systemaattiseen käyttöön on käytännön työelämässä vielä matkaa. Jokaisessa kohdeorganisaatioissa vaatimuksia katselmoidaan säännöllisesti vähintäänkin jollain tarkastusta epämuodollisemmalla menetelmällä, joten niiden kohdalla katselmointien merkitys ymmärretään hyvin. Sen sijaan yksikään kohdeorganisaatio ei katselmoi koodia säännöllisesti. Tarkastusten säännöllisyydessä on organisaatioiden välillä selviä eroja.

Joillekin kohdeorganisaatioille on vielä selvä haaste saada tarkastukset säännölliseen käyttöön. Toisissa organisaatioissa ainakin vaatimusten ja suunnitelmien tarkastaminen toimii jo säännönmukaisesti. Selvä johtopäätös kuitenkin oli, että tarkastuksia säännöllisestikin toteuttavien organisaatioiden käytänteissä on merkittävästi tarkastuksen tehokkuuteen vaikuttavia puutteita ja ongelmia. Esimerkiksi valmistautumisessa koettiin olevan merkittäviä ongelmia kaikissa organisaatioissa riippumatta siitä, kuinka säännöllisesti tarkastuksia toteutetaan.

Puutteista käytänteissä on nähtävissä selviä yhteyksiä havaittuihin ongelmiin. Keskeisin ongelmatekijä organisaatioissa vaikutti liittyvän työntekijöi-

den asenteisiin ja motivaatioon. Tavallisimpia käytännön ongelmia olivat huono valmistautuminen ja palaverin ajautuminen epäoleellisiin asioihin. Olemassa olevia käytänteitä tarkasteltaessa havaittiin, että yhdessäkään kohdeorganisaatioissa ei järjestetä koulutusta, eikä tarkastusten toteutumista seurata vastuuhenkilön toimesta. Näillä puutteilla voidaan todennäköisesti selittää ainakin osa havaituista ongelmista.

Eräs mielenkiintoinen havainto oli se, että asiakassuhteella vaikuttaa kohdeorganisaatioissa olevan selvä yhteys tarkastusten kurinalaisuuteen. Mitä enemmän asiakas on kiinnostunut tarkastuksista, sitä säännöllisemmin niitä toteutetaan. Kaikkein heikoimmin tarkastuksia toteuttavat sellaiset organisaatiot, joissa asiakas ei ole niistä kiinnostunut tai toiminta on luonteeltaan sellaista, että asiakasta ei ole. Tämä suuntaus kuvaa sitä, että motivaatio tarkastusten järjestämiseen organisaation oman sisäisen toiminnan tehostamiseksi ei ole kovin korkea.

Tutkimuksen taustalla oleva näkökulma on tarkastusten kehittäminen ohjelmistotuotannossa. Tässä vaiheessa tutkimuksessa on kartoitettu tarkastuskäytänteitä ja niihin liittyviä ongelmakohtia kirjallisuuden ja kohdeorganisaatioiden avulla. Tämän tutkimuksen perusteella voidaan selvästi nähdä tarve soveltavalle tutkimukselle, joka tähtää tarkastuskäytänteiden kehittämisen ymmärtämiseen. Seuraava askel on tarkastella näiden tutkimustulosten pohjalta sitä, kuinka tarkastuskäytänteitä saadaan kehitettyä organisaatioissa.

Tämän tutkimuksen yhteydessä kehitetty ICMM-malli voi toimia tarkastusten kehittämisen tukivälineenä. Sen avulla voidaan tunnistaa nykyisten käytänteiden keskeiset puutteet. Malli on tosin tässä vaiheessa vasta kirjallisuuden pohjalta kehitetty alustava versio. Sen rakennetta ja käyttökelpoisuutta käytännössä täytyy vielä arvioida.

Toinen kehittämiseen suuntautunut näkökulma tässä tutkimuksessa on olemassa olevien ongelmien analysointi. Eräs haastattelujen pohjalta syntynyt keskeinen johtopäätös oli, että säännöllisestikin toteutetut tarkastukset voivat sisältää merkittävästi tarkastusten tehokkuutta rajoittavia käytännön ongelmia. Ei siis riitä, että kehitetään käytänteitä ulkonaisesti, vaan ne täytyy myös saada toimimaan tehokkaasti. Artikkelissa Kollanus (2005b) on lyhyesti esitetty alustavia ajatuksia siitä, kuinka luvussa 6 esiteltyä listaa tarkastuksiin liittyvistä mahdollisista ongelmista voi käyttää kehittämisen tukena. Tällainen ongelmakeskeinen lähestymistapa vaatii kuitenkin vielä lisää tutkimusta ja jatkokehitystä.

Tässä tutkimuksessa otoksena oli pieni joukko ohjelmistoja tuottavia organisaatioita. Sen pohjalta ei voida tehdä vahvoja johtopäätöksiä tarkastuksiin liittyvästä havaittujen puutteiden ja ongelmien yhteydestä organisaation ominaisuuksiin. Tutkimuksen pohjalta voidaan kuitenkin esittää hypoteesi siitä, että erilaisissa organisaatioissa tarkastusten toteuttamiseen liittyy erilaisia ongelmia. Ongelmiin vaikuttavia tekijöitä saattavat olla esimerkiksi tarkastuskäytänteiden kypsyys, organisaation toiminnan luonne ja organisaation koko. Erilaisten kontekstien ymmärtäminen tarkastuksia käsittelevässä tutkimuksessa on yleisestikin haaste tulevaisuudessa. Tarkastuksiin liittyvien ilmiöiden ei voida ajatella toimivan samalla tavalla kaikenlaisissa organisaatioissa. Erilaisten kon-

tekstien tutkiminen vaatisi kuitenkin huomattavan laajan empirisen aineiston erilaisista organisaatioista.

SUMMARY

Software inspection is an already 30 years old method, which has been acknowledged as an important part of software engineering. Inspections have been reported to increase productivity and quality as well as to decrease costs and time to market. However, inspections are still not rigorously implemented in industry. In addition, we still know quite little about inspections in the real context.

The main research question in this thesis is to identify the most essential weaknesses and problems in inspection practices. *Weakness* in this study refers to a missing or improperly implemented area in inspection process. *Problem* means a practical issue, which could occur in practicing inspections. In this thesis both of these aspects are studied in theory and practice.

The first part of the thesis includes a systematic literature survey, which characterizes the previous inspection related research and its trends over the last 15 years (1990-2004). The survey covers 110 scientific articles from the most relevant journals and proceedings of International Conference on Software Engineering. The number of publications has increasingly grown during the studied time period. There were more inspection related articles published during the last five years than during the whole decade before that. According to the analysis, the best studied research area deals with different reading techniques and other effectiveness factors in inspection process. Also estimation of the defect content after an inspection has been an active research area, especially during the last few years. One of the major conclusions from the survey is that much more research is needed to understand, how inspections work in their real context.

The empirical part of the work includes case studies with 18 interviews in six organizations. These organizations are from five software companies with different sizes. Each case included three interviews representing different roles in the organization. Typical roles were quality manager, project manager and software developer. The two main goals in the interviews were 1) to assess the state of the current inspection practices and 2) to identify inspection related problems.

The thesis introduces Inspection Capability Maturity Model (ICMM), which applies the idea of capability maturity levels presented in the well known CMM, but it focuses only on inspection process. The purposes of ICMM are to support inspection process maturity assessment and inspection process improvement. It was constructed based on the existing literature, because there was no existing model suitable for this purpose. So, the model is still a preliminary version, and it has to be evaluated and improved in the future work. However, it served well in this study as a framework in analyzing existing inspection practices in the case organizations. ICMM was used to compare the case organizations and to identify the strengths and the weaknesses in their inspection practices.

Some previous studies have found that inspections are still used in an irregular manner in industry. The empirical part of this study confirms this finding, but it also studies the topic in more detail. For example, requirement documents are more or less formally reviewed in a relatively regular basis in the case organizations. On the other hand, general knowledge about inspection benefits, organizational processes and related organizational policies were relatively low. Moreover, the organizations don't provide any kind of training related to inspections. So, lack of proper training was probably the most remarkable weakness in the current practices.

There were also some other serious weaknesses in the current inspection practices in the case organizations. The motivation for discipline in inspections often appeared to come from the customer. Typically, even in the best case only the higher level documents, which are interest of the customer, were properly inspected. In the worst case, inspections may not be conducted at all, if there is no customer available. The organizations do very little to monitor and control inspection practices. The responsibility is usually left to the project managers and no one is interested in inspection data. These findings can be summarized to conclude that the case organizations haven't really understood the benefits they could gain from using inspections rigorously.

Practical problems in inspection practices were another part of this study. The first stage in this part was to collect a list of the problems mentioned in the literature. The interviews included both open question and structured part. The participants were first asked to freely mention the most serious problems in their inspections. Then the literature based list of proposed problems was presented to them. They were asked to estimate how usual the problems are in their own organizations.

The most common inspection related problem proposed in the case organizations was poor preparation. Other experienced problems were straying to irrelevant issues in the meetings and problems with scheduling. The most critical issue mentioned in an open question was tight schedule. Especially the most skilled people are very busy and they are wanted to participate in many different activities. An interesting finding was that the problems were mostly the same in all the case organizations, and they were not dependent on the current inspection process maturity. The final conclusion was that even regularly used inspection practices may include problems that cause significant loss of inspection effectiveness.

LÄHTEET

- de Almeida, J.R., Jr., Camargo, J.B., Jr., Basseto, B.A. & Paz, S.M. 2003. Best practices in code inspection for safety-critical software. *IEEE Software* 20(3), 56-63.
- Anderson, P., Reps, T. & Teitelbaum, T. 2003. Design and implementation of a fine-grained software inspection tool. *IEEE Transactions on Software Engineering* 29(8), 721-733.
- d'Astous, P. & Robillard P.N. 2000. Characterizing implicit information during peer review meetings. In *Proceedings of the 22nd International Conference on Software Engineering*, Limerick, Ireland, June 4-11. New York: ACM Press, 460 - 466.
- Aurum, A., Petersson, H. & Wohlin, C. 2002. State-of-the-art: software inspections after 25 years. *Software Testing, Verification and Reliability*, 12(3), 133-154.
- Barnard, J. & Price, A. 1994. Managing code inspection information. *IEEE Software* 11(2), 59-69.
- Basili, V.R., Green, S., Laitenberger, O., Lanubile, F., Shull, F., Soerumgaard, S. & Zelkowitz M. 1996. The empirical investigation of perspective-based reading. *Empirical Software Engineering* 1(2), 133-164.
- Basili, V.R. & Selby, R. 1987 Comparing the effectiveness of software testing strategies. *IEEE Transactions on Software Engineering* 13(12) 1278-1296.
- Bianchi, A., Lanubile, F. & Visaggio, G. 2001. A controlled experiment to assess the effectiveness of inspection meetings. In F. Titsworth (Eds.) *Proceedings of the International Symposium on Software Metrics*, London, England, April 4-6. Los Alamitos: IEEE Computer Society Press, 42-50.
- Biffi, S., Freimut, B. & Laitenberger, O. 2001. Investigating the cost-effectiveness of reinspections in software development. In *Proceedings of the 23rd International Conference on Software Engineering*, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 155 - 164.
- Biffi, S. & Grossmann, W. 2001. Evaluating the accuracy of defect estimation models based on inspection data from two inspection cycles. In *Proceedings of the 23rd International Conference on Software Engineering*, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 145 - 154.
- Bisant, D.B. & Lyle, J.R. 1989. A two-person inspection method to improve programming productivity. *IEEE Transactions on Software Engineering* 15(10), 1294-1304.
- Boehm, B.W. 1981. *Software engineering economics*. Englewood Cliffs, NJ: Prentice-Hall.
- Bourgeois, K.V. 1996. Process insights from a large-scale software inspections data analysis. *CrossTalk*, Oct. 1996.
- Briand, L., El Emam, K., Laitenberger, O. & Fussbroich, T. 1998. Using simulation to build inspection efficiency benchmarks for development projects. In *Proceedings of the 20th International Conference on Software Engineer-*

- ing, Kyoto, Japan, April 19-25. Washington: IEEE Computer Society, 340 – 349.
- Burnstein, I. Homeyen, A., Suwanassart, T., Saxena, G. & Grom, R. 1999. A testing maturity model for software test process assessment and improvement. *Software Quality Professional* 1(4), 8-21.
- Bush, M. 1990. Improving software quality: the use of formal inspections at the JPL. In *Proceedings of the 14th International Conference on Software Engineering*, Nice, France, March 26-30. Los Alamitos: IEEE Computer Society Press, 196-199.
- Card, D.N. 1990. Software quality engineering. *Information and Software Technology* 32(1), 3-10.
- Chaar, J.K., Halliday, M.J., Bhandari, I.S. & Chillarege, R. 1993. In-process evaluation for software inspection and test. *IEEE Transactions on Software Engineering* 19(11), 1055-1070.
- Chatzigeorgiou, A. & Antoniadis, G. 2003. Efficient management of inspections in software development projects. *Information and Software Technology* 45(10), 671-680.
- Chernak, Y. 1996. A statistical approach to the inspection checklist formal synthesis and improvement. *IEEE Transactions on Software Engineering* 22(12), 866-874.
- Chillarege, R., Bhandari, I.S., Chaar, J.K., Halliday, M.J., Moebus, D.S., Ray, B.K. & Wong, M.-Y. 1992. Orthogonal defect classification - a concept for in-process measurements. *IEEE Transactions on Software Engineering* 18(11), 943-956.
- Christenson, D.A, Huang, S.T. & Lamperez, A.J. 1990. Statistical quality control applied to code inspections. *IEEE Journal of Selected Areas of Communication* 8(2), 196-200.
- Ciolkowski, M., Laitenberger, O. & Biffel, S. 2003. Software reviews, the state of the practice. *IEEE Software* 20(6), 46-51.
- Doolan, E. 1992. Experience with Fagan's inspection method. *Software - Practice and Experience* 22(2), 173-182.
- Dunsmore, A. 2000. Survey of object-oriented defect detection approaches and experience in industry. Technical Report—EFoCS-37-2000, July 2000.
- Dunsmore, A., Roper, M. & Wood, M. 2000. Object-oriented inspection in the face of delocalisation. In *Proceedings of the 22nd International Conference on Software Engineering*, Limerick, Ireland, June 4-11. New York: ACM Press, 467 - 476.
- Dunsmore A., Roper M. & Wood M. 2001. Systematic object-oriented inspection technique. In *Proceedings of the 23rd International Conference On Software Engineering*, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 123-144.
- Dunsmore A., Roper M. & Wood M. 2002. Further investigations into the development and evaluation of reading techniques for object-oriented code inspection. In *Proceedings of the 24th International Conference on Software Engineering*, Orlando, Florida, May 19-25. New York: ACM Press, 47 - 57.

- Dunsmore, A., Roper, M. & Wood, M. 2003. The development and evaluation of three diverse techniques for object-oriented code inspection. *IEEE Transactions on Software Engineering* 29(8), 677-686.
- Ebert, C., Parro, C.H., Suttels, R. & Kolarczyk, H. 2001. Improving validation activities in a global software development. In *Proceedings of the 23rd International Conference on Software Engineering*, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 545 – 554.
- Eick, S.G., Loader, C.R., Long, M.D., Votta, L.G. & Wiel, S.V. 1992. Estimating software fault content before coding. In *Proceedings of the 14th International Conference on Software Engineering*, Melbourne, Australia, May 11-15. New York: ACM Press, 59-65.
- El Emam K., Laitenberger O. & Harbich T. 2000. The application of subjective estimates of effectiveness to controlling software inspections. *Journal of Systems and Software* 54(2), 119-136.
- Ericson, T., Subotec, A. & Ursing, S. 1997. TIM – A test improvement model, *Software Testing, Verification and Reliability* 7(4), 229-246.
- Erlikh, L. 2000. Leveraging legacy system dollars for E-business. *IT Professional* 2(3), 17-23.
- Fagan, M.E. 1976. Design and code inspection to reduce errors in program development. *IBM Systems Journal* 15(3), 182-211.
- Fagan, M.E. 1986. Advances in software inspections. *IEEE Transactions on Software Engineering* 12(7), 744-751.
- Freedman, D. & Weinberg, G. 1990. *Handbook of Walkthroughs, Inspections, and Technical Reviews*. New York: Dorset House.
- Fusaro, P., Lanubile, F. & Visaggio, G. 1997. A replicated experiment to assess requirements inspection techniques. *Empirical Software Engineering* 2(1), 39-57.
- Gilb, T. 2000. Planning to get the most out of inspections. *Software Quality Professional* 2(2), 46-57.
- Gilb, T. & Graham, D. 1993. *Software Inspection*. Wokingham, England: Addison-Wesley.
- Grady, R. & Van Slack, T. 1994. Key lessons in achieving widespread inspection use. *IEEE Software* 11(4), 46-57.
- IEEE, 1988. *IEEE Standard for Software Reviews and Audits*, IEEE Std 1028-1988.
- Harjumaa, L., Hedberg, H. & Tervonen, I. 2001. A path to virtual software inspection. In D. Young (Eds.) *Proceedings of Asia-Pacific Conference on Quality Software (APAQS 2001)*, Hong Kong, December 10-11. IEEE Computer Society, 283-287.
- Harjumaa, L, Tervonen, I. & Vuorio, P. 2004. Using software inspection as a catalyst for SPI in a small company. In *Proceedings of 5th International Conference of Product Focused Software Process Improvement*, Kausai Science City, Japan, April 5-8. Springer LNCS 3009, 62-75.
- Hedberg, H. 2004. Introducing the next generation of software inspection tools. In D. Young (Eds.) *Proceedings of 5th International Conference of Product*

- Focused Software Process Improvement, Kausai Science City, Japan, April 5-8. Springer LNCS 3009, 234-247.
- Jacob, A.L. & Pillai, S.K. 2003. Statistical process control to improve coding and code review. *IEEE Software* 20(3), 50-55.
- Jalote, P. & Haragopal, M. 1998. Overcoming the NAH syndrome for inspection deployment. In *Proceedings of the 20th International Conference on Software Engineering*, Kyoto, Japan, April 19-25. Washington: IEEE Computer Society, 371 – 378.
- Johnson, P.M. 1994. An instrumented approach to improving software quality through formal technical review. In *Proceedings of the 16th International Conference on Software Engineering*, Sorrento, Italy, May 16-21. Los Alamitos: IEEE Computer Society Press, 113-122.
- Johnson, P.M. 1998. Reengineering inspection. *Communications of the ACM* 41(2), 49-52.
- Johnson, P.M. & Tjahjono, D. 1998. Does every inspection really need a meeting? *Empirical Software Engineering* 3(1), 9-35.
- Kelly, D. & Shepard, T. 2002. Qualitative observations from software code inspection experiments. In D. Stewart & H. Johnson (Eds.) *Proceedings of the 2002 Conference of the Centre for Advanced Studies on Collaborative research*, Toronto, Canada, September 30 - October 3. IBM Press.
- Kelly, D. & Shepard, T. 2004. Task-directed software inspection. *Journal of Systems and Software* 73(2), 361-368.
- Kitchenham, B.A., Pfleeger, S.L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., El Emam, K. & Rosenberg, J. 2002. Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering* 28(8), 721-734.
- Knight, J.C. & Myers, E.A. 1993. An improved inspection technique. *Communications of the ACM* 36(11), 51-61.
- Kollanus S. 2005a. ICMM – inspection capability maturity model. In M. Hamza (Eds.) *Proceedings of the IASTED International Conference on Software Engineering*, Innsbruck, Austria, February 15-17. ACTA Press, 372-377.
- Kollanus S. 2005b. Issues in software inspection practices. In F. Bomarius & S. Komi-Sirviö (Eds.) *Proceedings of the PROFES 2005 Conference*, Oulu, June 13-18. Springer LNCS 3547, 429-442.
- Kusumoto, S., Kikuno, T., Matsumoto, K. & Torii, K. 1996. Experimental evaluation of time allocation procedure for technical reviews. *Journal of Systems and Software* 35(2), 119-126.
- Kusumoto, S., Chimura, A., Kikuno, T., Matsumoto, K. & Mohri, Y. 1998. A promising approach to two-person software review in educational environment. *Journal of Systems and Software* 40(2), 115-123.
- Laitenberger, O., Atkinson, C., Schlich, M. & El Emam, K. 2000. An experimental comparison of reading techniques for defect detection in UML design documents. *Journal of Systems and Software* 53(2), 183-204.
- Laitenberger, O., Beil, T. & Schwinn, T. 2002. An industrial case study to examine a non-traditional inspection implementation for requirements specifications. *Empirical Software Engineering* 7(4), 345 - 374.

- Laitenberger, O. & DeBaud, J.-M. 1997. Perspective-based reading of code documents at Robert Bosch GmbH. *Information and Software Technology* 39(11), 781-791.
- Laitenberger, O. & DeBaud, J.-M. 2000. An encompassing life cycle centric survey of software inspection. *Journal of Systems and Software* 50(1), 5-31.
- Land, L., Sauer, C. & Jeffery, R. 2000. The use of procedural roles in code inspections: an experimental study. *Empirical Software Engineering*, 5(1), 11-34.
- Martin, J. & Tsai, W.T. 1990. N-Fold inspection: a requirements analysis technique. *Communications of the ACM* 33(2), 225-232.
- Macdonald, F. & Miller, J. 1997. A software inspection process definition language and prototype support tool. *Software Testing, Verification and Reliability* 7(2), 99-128.
- Macdonald F. & Miller J. 1998. A comparison of tool-based and paper-based software inspection. *Empirical Software Engineering* 3(3), 233 - 253.
- Macdonald, F. & Miller, J. 1999. ASSIST-a tool to support software inspection. *Information and Software Technology* 41(15), 1045-1057.
- Macdonald, F., Miller, J., Brooks, A., Roper, M. & Wood, M. 1996. Applying inspection to object-oriented code. *Software Testing, Verification and Reliability* 6(2), 61-82.
- Madachy, R. 1996. System dynamics modeling of an inspection-based process. In *Proceedings of the 18th International Conference on Software Engineering*, Berlin, Germany, March 25-29. Washington: IEEE Computer Society, 376 - 386.
- McConnell, S. 2000. The best influences on software engineering. *IEEE Software* 17(1), 10-17.
- Mashayekhi, V., Drake, J.M., Tsai, W.-T. & Riedl, J. 1993. Distributed, collaborative software inspection. *IEEE Software* 10(5), 66-75.
- Miller J., Wood M. & Roper M. 1998. Further Experiences with Scenarios and Checklists. *Empirical Software Engineering* 3(1), 37-64.
- Müller, M. 2004. Are reviews an alternative to pair programming? *Empirical Software Engineering* 9(4), 335 - 351.
- Neumann, P. 2005. Risks to the public. *ACM SIGSOFT Software Engineering Notes* 30(3), 22-37.
- Padberg, F. 2002. Empirical interval estimates for the defect content after an inspection. In *Proceedings of the 24th International Conference on Software Engineering*, Orlando, Florida, May 19-25. New York: ACM Press, 58 - 68.
- Parnas, D.L. & Lawford, M. 2003. The role of inspection in software quality assurance. *IEEE Transactions on Software Engineering* 29(8), 674-676.
- Parnas, D.L. & Weis D.M. 1987. Active design reviews: Principles and practices. *Journal of Systems and Software* 7(4), 259-265.
- Perpich, J.M., Perry, D.E., Porter, A.A., Votta, L.G. & Wade, M.W. 1997. Anywhere, anytime code inspections: using the web to remove inspection bottlenecks in large-scale software development. In *Proceedings of the 19th International Conference on Software Engineering*, Boston, Massachusetts, May 17-23. New York: ACM Press, 14 - 21.

- Perry, D.E., Porter, A., Wade, M.W., Votta, L.G. & Perpich, J. 2002. Reducing inspection interval in large-scale software development. *IEEE Transactions on Software Engineering* 28(7), 695-705.
- Petersson, H., Thelin, T., Runeson, P. & Wohlin, C. 2004. Capture-recapture in software inspections after 10 years research – theory, evaluation and application. *Journal of Systems and Software* 72(2), 249-264.
- Porter, A.A. & Johnson, P.M. 1997. Assessing software review meetings: results of a comparative analysis of two experimental studies. *IEEE Transactions on Software Engineering* 23(3), 129-145.
- Porter, A., Siy, H., Mockus, A. & Votta, L. 1998. Understanding the sources of variation in software inspections. *ACM Transactions on Software Engineering and Methodology* 7(1), 41-79.
- Porter, A.A., Siy, H.P., Toman, C.A. & Votta, L.G. 1997. An experiment to assess the cost-benefits of code inspections in large scale software development. *IEEE Transactions on Software Engineering* 23(6), 329-346.
- Porter, A. & Votta, L.G. 1994. An experiment to assess different defect detection methods for software requirements inspections. In *Proceedings of the 16th International Conference on Software Engineering, Sorrento, Italy, May 16-21*. Los Alamitos: IEEE Computer Society Press, 103-112.
- Porter, A. & Votta, L. 1998. Comparing detection methods for software requirements inspections: a replication using professional subjects. *Empirical Software Engineering*, 3(4), 355-379.
- Porter, A.A., Votta, L.G., Jr. & Basili, V.R. 1995. Comparing detection methods for software requirements inspections: a replicated experiment. *IEEE Transactions on Software Engineering* 21(6), 563-575.
- Rifkin, S. & Deimel, L. 1994. Applying program comprehension techniques to improve software inspections. In *Proceedings of the 19th Annual NASA Software Engineering Workshop, Greenbelt, Maryland*, 115-126.
- Roper M., Wood M. & Miller J. 1997. An empirical evaluation of defect detection techniques. *Information and Software Technology* 39(11), 763-775.
- Russell, G.W. 1991. Experience with inspection in ultra large-scale developments. *IEEE Software* 8(1), 25-31.
- Sabaliauskaite G., Kusumoto S. & Inoue K. 2004. Assessing defect detection performance of interacting teams in object-oriented design inspection. *Information and Software Technology* 46(13), 875-886.
- Sabaliauskaite, G., Matsukawa, F., Kusumoto S. & Inoue, K. 2003. Further investigations of reading techniques for object-oriented design inspection. *Information and Software Technology* 45(9), 571-585.
- Sandahl, K., Blomkvist, O., Karlsson, K., Krysander, C., Lindvall, M. & Ohlson, N. 1998. An extended replication of an experiment for assessing methods for software requirements inspections. *Empirical Software Engineering* 3(4), 327 - 354.
- Sauer, C., Jeffery, D.R., Land, L. & Yetton, P. 2000. The effectiveness of software development technical reviews: a behaviorally motivated program of research. *IEEE Transactions on Software Engineering* 26(1), 1-14.

- Seaman, C.B. & Basili, V.R. 1998. Communication and organization: an empirical study of discussion in inspection meetings. *IEEE Transactions on Software Engineering* 24(7), 559-572.
- SEI (Software Engineering Institute) 2002. Capability Maturity Model Integration, Version 1.1. (URL: <http://www.sei.cmu.edu/cmm/>)
- Shull, F., Lanubile, F. & Basili, V.R. 2000. Investigating reading techniques for object-oriented framework learning. *IEEE Transactions on Software Engineering* 26(11), 1101-1118.
- So, S., Cha, S., Shimeall, T. & Kwon, Y. 2002. An empirical evaluation of six methods to detect faults in software. *Software Testing, Verification and Reliability* 12(3), 155-171.
- Sommerville, I. 2001. *Software Engineering* (6. edition). Harlow: Addison-Wesley.
- Stein, M., Riedl, J., Harner, S. & Mashayekhi, V. 1997. A case study of distributed, asynchronous software inspection. In *Proceedings of the 19th International Conference on Software Engineering*, Boston, Massachusetts, May 17-23. New York: ACM Press, 107 - 117.
- Tervonen, I. 1996. Support for quality-based design and inspection. *IEEE Software* 13(1), 44-54.
- Tervonen, I, Iisakka, J. & Harjumaa, L. 2001. Looking for inspection improvements through the base practices. In M. Lawford & D.L. Parnas (Eds.) *Proceedings of the 1st Workshop on Inspection in Software Engineering*, Paris, France, July 23. McMaster University, Canada, 145-152.
- Thelin, T. 2004. Team-based fault content estimation in the software inspection process. In *Proceedings of the 26th International Conference on Software Engineering*, Edinburgh, Scotland, May 23-28. Washington: IEEE Computer Society, 263- 272.
- Thelin, T., Petersson, H., Runeson, P. & Wohlin, C. 2004a. Applying sampling to improve software inspections. *Journal of Systems and Software* 73(2), 257-269.
- Thelin, T., Runeson, P. & Regnell, B. 2001. Usage-based reading — an experiment to guide reviewers with use cases. *Journal of Information and Software Technology* 43(15), 925-938.
- Thelin, T., Runeson, P. & Wohlin, C. 2003. An experimental comparison of usage-based and checklist-based reading. *IEEE Transactions on Software Engineering* 29(8), 687-704.
- Thelin, T., Runeson, P., Wohlin, C., Olsson, T. & Andersson, C. 2004b. Evaluation of usage-based reading — conclusions after three experiments. *Empirical Software Engineering* 9(1-2), 77 – 110.
- Traore, I. & Aredo, D.B. 2004. Enhancing structured review with model-based verification. *IEEE Transactions on Software Engineering* 30(11), 736- 753.
- Tyran, C.K. & George, J.F. 2002. Improving software inspections with group process support. *Communications of the ACM* 45(9), 87 - 92.
- UKSMA (United Kingdom Software Metrics Association) 2002. *Quality Standards Defect Measurement Manual Release 1.a.* (URL: <http://www.ukσμα.co.uk/>)

- Votta, L. 1993. Does every inspection need a meeting? *ACM Software Engineering Notes* 18(5), 107-114.
- Weinberg, G.M. 1992. *Quality Software Management, Volume 1: Systems thinking*, New York : Dorset House, cop.
- Wiegers, K. 1998. Seven deadly sins of software reviews. *Software Development* 6(3).
- Wiegers, K. 2002. Seven truths about peer reviews. *Cutter IT Journal* 15(7).
- Xu, J. 2003. Making software timing properties easier to inspect and verify. *IEEE Software* 20(4), 34-41.
- Yin, Z., Dunsmore, A. & Miller, J. 2004. Self-assessment of performance in software inspection processes. *Information and Software Technology* 46(3), 185-194.

LIITE 1 – TÄYDELLINEN LISTA KIRJALLISUUSANALYYSIIN SISÄLTÄNEISTÄ ARTIKKELEISTA

Tässä liitteessä on listattu kaikki artikkelit, jotka on valikoitu mukaan kirjallisuusanalyysiin. Julkaisut on järjestetty niistä löytyneiden katselmointeja käsittelevien artikkeleiden määrän mukaan. Jokaisen julkaisun kohdalla artikkelit on järjestetty aikajärjestykseen uusimmasta vanhimpaan.

International Conference on Software Engineering (vv. 1990-2004)

1. Thelin, T. 2004. Team-based fault content estimation in the software inspection process. Proceedings of the 26th International Conference on Software Engineering, Edinburgh, Scotland, May 23-28. Washington: IEEE Computer Society, 263- 272.
2. Gantner, T. & Barth, T. 2003. Experiences on defining and evaluating an adapted review process. Proceedings of the 25th International Conference on Software Engineering, Portland, Oregon. Washington: IEEE Computer Society, 506-511.
3. Dunsmore, A., Roper, M. & Wood, M. 2002. Further investigations into the development and evaluation of reading techniques for object-oriented code inspection. Proceedings of the 24th International Conference on Software Engineering, Orlando, Florida, May 19-25. New York: ACM Press, 47 – 57.
4. Padberg, F. 2002. Empirical interval estimates for the defect content after an inspection. Proceedings of the 24th International Conference on Software Engineering, Orlando, Florida, May 19-25. New York: ACM Press, 58 – 68.
5. Dunsmore, A., Roper, M. & Wood, M. 2001. Systematic object-oriented inspection — an empirical study. Proceedings of the 23rd International Conference on Software Engineering, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 135 – 144.
6. Biffl, S. & Grossmann, W. 2001. Evaluating the accuracy of defect estimation models based on inspection data from two inspection cycles. Proceedings of the 23rd International Conference on Software Engineering, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 145 – 154.
7. Biffl, S., Freimut, B. & Laitenberger, O. 2001. Investigating the cost-effectiveness of reinspections in software development. Proceedings of the 23rd International Conference on Software Engineering, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 155 – 164.
8. Ebert, C., Parro, C.H., Suttels, R. & Kolarczyk, H. 2001. Improving validation activities in a global software development. Proceedings of the 23rd International Conference on Software Engineering, Toronto, Canada, May 12-19. Washington: IEEE Computer Society, 545 – 554.
9. d’Astous, P. & Robillard, P. 2000. Characterizing implicit information during peer review meetings. Proceedings of the 22nd International Conference on Software Engineering, Limerick, Ireland, June 4-11. New York: ACM Press, 460 – 466.
10. Dunsmore, A., Roper, M. & Wood, M. 2000. Object-oriented inspection in the face of delocalization. Proceedings of the 22nd International Conference on Software Engineering, Limerick, Ireland, June 4-11. New York: ACM Press, 467 – 476.
11. Briand, L., El Emam, K., Laitenberger, O. & Fussbroich, T. 1998. Using simulation to build inspection efficiency benchmarks for development projects. Proceedings of the 20th International Conference on Software Engineering, Kyoto, Japan, April 19-25. Washington: IEEE Computer Society, 340 – 349.
12. Jalote, P. & Haragopal, M. 1998. Overcoming the NAH syndrome for inspection deployment. Proceedings of the 20th International Conference on Software Engineering, Kyoto, Japan, April 19-25. Washington: IEEE Computer Society, 371 – 378.

13. Perpich, J.M. & Perry, D.E., Porter, A.A., Votta, L.G. & Wade, M.W. 1997. Anywhere, any-time code inspections: using the Web to remove inspection bottlenecks in large-scale software development. Proceedings of the 19th International Conference on Software Engineering, Boston, Massachusetts, May 17-23. New York: ACM Press, 14 – 21.
14. Seaman, C. & Basili, V.R. 1997. An empirical study of communication in code inspections. Proceedings of the 19th International Conference on Software Engineering, Boston, Massachusetts, May 17-23. New York: ACM Press, 96 – 106.
15. Stein, M., Riedl, J., Harner, S. & Mashayekhi, V. 1997. A case study of distributed, asynchronous software inspection. Proceedings of the 19th International Conference on Software Engineering, Boston, Massachusetts, May 17-23. New York: ACM Press, 107 – 117.
16. Johnson, P.M. & Tjahjono, D. 1997. Assessing software review meetings: a controlled experimental study using CSRS. Proceedings of the 19th International Conference on Software Engineering, Boston, Massachusetts, May 17-23. New York: ACM Press, 118 – 127.
17. Porter, A.A., Siy, H.P. & Votta, L.G. 1997. Understanding the effects of developer activities on inspection interval. Proceedings of the 19th International Conference on Software Engineering, Boston, Massachusetts, May 17-23. New York: ACM Press, 128 – 138.
18. Madachy, R. 1996. System dynamics modeling of an inspection-based process. Proceedings of the 18th International Conference on Software Engineering, Berlin, Germany, March 25-29. Washington: IEEE Computer Society, 376 – 386.
19. Porter, A.A. & Votta, L.G. 1994. An experiment to assess different defect detection methods for software requirements inspections. Proceedings of the 16th International Conference on Software Engineering, Sorrento, Italy, May 16-21. Los Alamitos: IEEE Computer Society Press, 103-112.
20. Johnson, P.M. 1994. An instrumented approach to improving software quality through formal technical review. Proceedings of the 16th International Conference on Software Engineering, Sorrento, Italy, May 16-21. Los Alamitos: IEEE Computer Society Press, 113-122.
21. Eick, S.G., Loader, C.R., Long, M.D., Votta, L.G. & Wiel, S.V. 1992. Estimating software fault content before coding. Proceedings of the 14th International Conference on Software Engineering, Melbourne, Australia, May 11-15. New York: ACM Press, 59-65.
22. Bush, M. 1990. Improving software quality: the use of formal inspections at the JPL. Proceedings of the 14th International Conference on Software Engineering, Nice, France, March 26-30. Los Alamitos: IEEE Computer Society Press, 196-199.

IEEE Transactions on Software Engineering (vv. 1990-2004)

1. Traore, I. & Aredo, D.B. 2004. Enhancing structured review with model-based verification. *IEEE Transactions on Software Engineering* 30(11), 736- 753.
2. Miller, J. & Zhichao Y. 2004. A cognitive-based mechanism for constructing software inspection teams. *IEEE Transactions on Software Engineering* 30(11), 811- 825.
3. Padberg, F., Ragg, T. & Schoknecht, R. 2004. Using machine learning for estimating the defect content after an inspection *IEEE Transactions on Software Engineering* 30(1), 17 - 28.
4. Dunsmore, A., Roper, M. & Wood, M. 2003. The development and evaluation of three diverse techniques for object-oriented code inspection. *IEEE Transactions on Software Engineering* 29(8), 677-686.
5. Thelin, T., Runeson, P. & Wohlin, C. 2003. An experimental comparison of usage-based and checklist-based reading. *IEEE Transactions on Software Engineering* 29(8), 687-704.
6. Anderson, P., Reps, T. & Teitelbaum, T. 2003. Design and implementation of a fine-grained software inspection tool. *IEEE Transactions on Software Engineering* 29(8), 721-733.
7. Biffl, S. & Halling, M. 2003. Investigating the defect detection effectiveness and cost benefit of nominal inspection teams. *IEEE Transactions on Software Engineering* 29(5), 385-397.
8. Perry, D.E., Porter, A., Wade, M.W., Votta, L.G. & Perpich, J. 2002. Reducing inspection interval in large-scale software development. *IEEE Transactions on Software Engineering* 28(7), 695- 705.

9. El Emam, K. & Laitenberger, O. 2001. Evaluating capture-recapture models with two inspectors. *IEEE Transactions on Software Engineering* 27(9), 851-864.
10. Laitenberger, O., El Emam, K. & Harbich, T.G. 2001. An internally replicated quasi-experimental comparison of checklist and perspective based reading of code documents. *IEEE Transactions on Software Engineering* 27(5), 387-421.
11. Sauer, C., Jeffery, D.R., Land, L. & Yetton, P. 2000. The effectiveness of software development technical reviews: a behaviorally motivated program of research. *IEEE Transactions on Software Engineering* 26(1), 1-14.
12. Seaman, C.B. & Basili, V.R. 1998. Communication and organization: an empirical study of discussion in inspection meetings. *IEEE Transactions on Software Engineering* 24(7), 559-572.
13. Ebrahimi, N.B. 1997. On the statistical analysis of the number of errors remaining in a software design document after inspection. *IEEE Transactions on Software Engineering* 23(8), 529-532.
14. Porter, A.A., Siy, H.P., Toman, C.A. & Votta, L.G. 1997. An experiment to assess the cost-benefits of code inspections in large scale software development. *IEEE Transactions on Software Engineering* 23(6), 329-346.
15. Porter, A.A. & Johnson, P.M. 1997. Assessing software review meetings: results of a comparative analysis of two experimental studies. *IEEE Transactions on Software Engineering* 23(3), 129-145.
16. Chernak, Y. 1996. A statistical approach to the inspection checklist formal synthesis and improvement. *IEEE Transactions on Software Engineering* 22(12), 866-874.
17. Porter, A.A., Votta, L.G., Jr. & Basili, V.R. 1995. Comparing detection methods for software requirements inspections: a replicated experiment. *IEEE Transactions on Software Engineering* 21(6), 563-575.
18. Chaar, J.K., Halliday, M.J., Bhandari, I.S. & Chillarege, R. 1993. In-process evaluation for software inspection and test. *IEEE Transactions on Software Engineering* 19(11), 1055-1070.

Journal of Systems and Software (vv. 1990-2004)

1. Briand, L., Freimut, B. & Vollei, F. 2004. Using multiple adaptive regression splines to support decision making in code inspections. *Journal of Systems and Software* 73(2), 205-217.
2. Thelin, T., Petersson, H., Runeson, P. & Wohlin, C. 2004. Applying sampling to improve software inspections. *Journal of Systems and Software* 73(2), 257-269.
3. Kelly, D. & Shepard, T. 2004. Task-directed software inspection. *Journal of Systems and Software* 73(2), 361-368.
4. Petersson, H., Thelin, T., Runeson, P. & Wohlin, C. 2004. Capture-recapture in software inspections after 10 years research—theory, evaluation and application. *Journal of Systems and Software* 72(2), 249-264.
5. Biffi, S. 2003. Evaluating defect estimation models with major defects. *Journal of Systems and Software* 65(1), 13-29.
6. Miller, J. 2002. On the independence of software inspectors. *Journal of Systems and Software* 60(1), 5-10.
7. El Emam, K., Laitenberger, O. & Harbich, T. 2000. The application of subjective estimates of effectiveness to controlling software inspections. *Journal of Systems and Software* 54(2), 119-136.
8. Laitenberger, O., Atkinson, C., Schlich, M. & El Emam, K. 2000. An experimental comparison of reading techniques for defect detection in UML design documents. *Journal of Systems and Software* 53(2), 183-204.
9. Höst, M. & Johansson, C. 2000. Evaluation of code review methods through interviews and experimentation. *Journal of Systems and Software* 52(2-3), 113-120.
10. Dunsmore, A., Roper, M. & Wood, M. 2000. The role of comprehension in software inspection. *Journal of Systems and Software* 52(2-3), 121-129.

11. Thelin, T. & Runeso P. 2000. Robust estimations of fault content with capture–recapture and detection profile estimators. *Journal of Systems and Software* 52(2-3), 139-148.
12. Miller, J. & Macdonald, F. 2000. An empirical incremental approach to tool evaluation and improvement. *Journal of Systems and Software* 51(1), 19-35.
13. Laitenberger, O. & DeBaud, J.-M- 2000. An encompassing life cycle centric survey of software inspection. *Journal of Systems and Software* 50(1), 5-31.
14. Näslund, T. & Löwgren, J. 1999. Usability inspection in contract-based systems development – A contextual assessment. *Journal of Systems and Software* 45(3), 233-240.
15. Kusumoto, S., Chimura, A., Kikuno, T., Matsumoto, K. & Mohri, Y. 1998. A promising approach to two-person software review in educational environment. *Journal of Systems and Software* 40(2), 115-123.
16. Kusumoto, S., Kikuno, T., Matsumoto, K. & Torii, K. 1996. Experimental evaluation of time allocation procedure for technical reviews. *Journal of Systems and Software* 35(2), 119-126.
17. Kelly, J., Sherif, J. & Hops, J. 1992. An analysis of defect densities found during software inspections. *Journal of Systems and Software* 17(2), 111-117.

Empirical Software Engineering (vv. 1997-2004)

1. Müller, M. 2004. Are reviews an alternative to pair programming? *Empirical Software Engineering* 9(4), 335 - 351.
2. Thelin, T., Runeson, P. & Wohlin, C. 2004. Evaluation of usage-based reading—conclusions after three experiments. *Empirical Software Engineering* 9(1-2), 77 - 110.
3. Laitenberger, O., Beil, T. & Schwinn, T. 2002. An industrial case study to examine a non-traditional inspection implementation for requirements specifications. *Empirical Software Engineering* 7(4), 345 - 374.
4. Biffel, S. & Gutjahr, W. 2002. Using a reliability growth model to control software inspection. *Empirical Software Engineering* 7(3), 257 - 284.
5. d'Astous, P., Robillard, P., Détienne, F. & Visser, W. 2001. Quantitative measurements of the influence of participant roles during peer review meetings. *Empirical Software Engineering* 6(2), 143-159.
6. Regnell, B., Runeson, P. & Thelin, T. 2000. Are the perspectives really different? – Further experimentation on scenario-based reading of requirements. *Empirical Software Engineering* 5(4), 331 - 356.
7. Land, L., Sauer, C. & Jeffery, R. 2000. The use of procedural roles in code inspections: an experimental study. *Empirical Software Engineering* 5(1), 11-34.
8. Zhang, Z., Basili, V.R. & Shneiderman, B. 1999. Perspective-based usability inspection: an empirical validation of efficacy. *Empirical Software Engineering* 4(1), 43-69.
9. Runeson, P. & Wohlin, C. 1998. An experimental evaluation of an experience-based capture-recapture method in software code inspections. *Empirical Software Engineering* 3(4), 381 - 406.
10. Porter, A. & Votta, L. 1998. Comparing detection methods for software requirements inspections: a replication using professional subjects. *Empirical Software Engineering* 3(4), 355 - 379.
11. Sandahl, K., Blomkvist, O., Karlsson, J., Krysanter, C., Lindvall, M. & Ohlsson, N. 1998. An extended replication of an experiment for assessing methods for software requirements inspections. *Empirical Software Engineering* 3(4), 327 - 354.
12. Macdonald, F. & Miller, J. 1998. A comparison of tool-based and paper-based software inspection. *Empirical Software Engineering* 3(3), 233 - 253.
13. Miller, J., Wood, M. & Roper, M. 1998. Further experiences with scenarios and checklists. *Empirical Software Engineering* 3(1), 37-64.
14. Johnson, P.M. & Tjahjono, D. 1998. Does every inspection really need a meeting? *Empirical Software Engineering* 3(1), 9-35.
15. Fusaro, P., Lanubile, F. & Visaggio, G. 1997. A replicated experiment to assess requirements inspection techniques. *Empirical Software Engineering* 2(1), 39-57.

16. Basili, V.R., Green, S., Lanubile, F., Shull, F., Sorumgard, S. & Zelkowitz, M. 1996. The empirical investigation of perspective-based reading. *Empirical Software Engineering* 1(2), 133-166.

IEEE Software (vv. 1990-2004)

1. Ciolkowski, M., Laitenberger, O. & Biffl, S. 2003. Software reviews, the state of the practice. *IEEE Software* 20(6), 46-51.
2. Jacob, A.L. & Pillai, S.K. 2003. Statistical process control to improve coding and code review. *IEEE Software* 20(3), 50-55.
3. De Almeida, J.R., Jr., Camargo, J.B., Jr., Basseto, B.A. & Paz, S.M. 2003. Best practices in code inspection for safety-critical software. *IEEE Software* 20(3), 56-63.
4. Kazman, R. & Bass, L. 2002. Making architecture reviews work in the real world *IEEE Software* 19(1), 67-73.
5. Van Genuchten, M., van Dijk, C., Scholten, H. & Vogel, D. 2001. Using group support systems for software inspections. *IEEE Software* 18(3), 60-65.
6. Biffl, S. 2000. Using inspection data for defect estimation. *IEEE Software* 17(6), 36-43.
7. Porter, A. & Votta, L. 1997. What makes inspections work? *IEEE Software* 14(6), 99-102.
8. Tervonen, I. 1996. Support for quality-based design and inspection. *IEEE Software* 13(1), 44-54.
9. R.B. Grady & T. Van Slack 1994. Key lessons in achieving widespread inspection use. *IEEE Software* 11(4), 46-57.
10. Barnard, J. & Price, A. 1994. Managing code inspection information. *IEEE Software* 11(2), 59-69.
11. Weller, E.F. 1993. Lessons from three years of inspection data. *IEEE Software* 10(5), 38-45.
12. Mashayekhi, V., Drake, J.M., Tsai, W.-T. & Riedl, J. 1993. Distributed, collaborative software inspection. *IEEE Software* 10(5), 66-75.
13. Bias, R. 1991. Interface-walkthroughs: efficient collaborative testing. *IEEE Software* 8(5), 94-95.
14. Russell, G.W. 1991. Experience with inspection in ultra large-scale developments. *IEEE Software* 8(1), 25-31.

Information and Software Technology (vv. 1990-2004)

1. Sabaliauskaite, G., Kusumoto, S. & Inoue, K. 2004. Assessing defect detection performance of interacting teams in object-oriented design inspection. *Information and Software Technology* 46(11), 875-886.
2. Yin, Z., Dunsmore, A. & Miller, J. 2004. Self-assessment of performance in software inspection processes. *Information and Software Technology* 46, 185-194.
3. Chatzigeorgiou, A. & Antoniadis, G. 2003. Efficient management of inspections in software development projects. *Information and Software Technology* 45(10), 671-680.
4. Sabaliauskaite, G., Matsukawa, F., Kusumoto, S. & Inoue, K. 2003. Further investigations of reading techniques for object-oriented design inspection. *Information and Software Technology* 45(9), 571-585.
5. Thelin, T. & Runeson, P. 2002. Confidence intervals for capture-recapture estimations in software inspections. *Information and Software Technology* 44(12), 683-702.
6. d'Astous, P. & Robillard, P. 2002. Empirical study of exchange patterns during software peer review meetings. *Information and Software Technology* 44(11), 639-648.
7. Thelin, T., Runeson, P. & Regnell, B. 2001. Usage-based reading — an experiment to guide reviewers with use cases. *Journal of Information and Software Technology* 43(15), 925-938.
8. Macdonald, F. & Miller, J. 1999. ASSIST—a tool to support software inspection. *Information and Software Technology* 41(15), 1045-1057.
9. Roper, M., Wood, M. & Miller, J. 1997. An empirical evaluation of defect detection techniques. *Information and Software Technology* 39(11), 763-775.

10. Laitenberger, O. & DeBaud, J.-M. 1997. Perspective-based reading of code documents at Robert Bosch GmbH. *Information and Software Technology* 39(11), 781-791.
11. Raz, T. & Yaung, A. 1997. Factors affecting design inspection effectiveness in software development. *Information and Software Technology* 39(4), 297-305.

Journal of Software Testing, Verification and Reliability (vv. 1996-2005)

1. Kelly, D. & Shepard, T. 2004. Eight maxims for software inspectors. *Software Testing, Verification and Reliability* 14(4), 243-256.
2. Aurum, A., Petersson, H. & Wohlin, C. 2002. State-of-the-art: software inspections after 25 years. *Software Testing, Verification and Reliability* 12(3), 133-154.
3. So, S., Cha, S., Shimeall, T. & Kwon, Y. 2002. An empirical evaluation of six methods to detect faults in software. *Software Testing, Verification and Reliability* 12(3), 155-171.
4. Miller, J. 1999. Estimating the number of remaining defects after inspection. *Software Testing, Verification and Reliability* 9(3), 167-189.
5. Macdonald, F. & Miller, J. 1997. A software inspection process definition language and prototype support tool. *Software Testing, Verification and Reliability* 7(2), 99-128.
6. Macdonald, F., Miller, J., Brooks, A., Roper, M. & Wood, M. 1996. Applying inspection to object-oriented code. *Software Testing, Verification and Reliability* 6(2), 61-82.

Communications of the ACM (vv. 1990-2004)

1. Tyran, C. & George, J. 2002. Improving software inspections with group process support. *Communications of the ACM* 45(9), 87 - 92.
2. Johnson, P.M. 1998. Reengineering inspection. *Communications of the ACM* 41(2), 49 - 52.
3. Knight, J. & Myers, A. 1993. An improved inspection technique. *Communications of the ACM* 36(11), 51 - 61.
4. Martin, J. & Tsai, W. 1990. N-Fold inspection: a requirements analysis technique. *Communications of the ACM* 33(2), 225 - 232.

ACM Transactions on Software Engineering and Methodology (vv. 1992-2004)

1. Porter, A., Siy, H., Mockus, A. & Votta, L. 1998. Understanding the sources of variation in software inspections. *ACM Transactions on Software Engineering and Methodology* 7(1), 41-79.
2. Schneider, M., Martin, J. & Tsai, W. 1992. An experimental study of fault detection in user requirements documents. *ACM Transactions on Software Engineering and Methodology* 1(2), 188 - 204.

LIITE 2 – HAASTATTELURUNKO

Haastattelun alku

Yleiset tiedot haastattelusta ja haastateltavasta

- Aika
- Paikka
- Yritys
- Haastateltavan nimi
- Työtehtävät yrityksessä
- Kuinka pitkä työkokemus yrityksessä ja alalla yhteensä

Yleiset tiedot yrityksestä

- Yrityksen koko
- Yksikön koko
- Tavallinen projektin koko
- Onko laatujärjestelmän (ISO9001, CMM, SPICE) auditointia tehty?
- Tehdäänkö yksikössä pääasiassa tuotekehitystä vai asiakasprojekteja?

Johdantokysymykset

- Määrittele katselmointi (tarkoitus varmistaa, että puhutaan samasta asiasta)
- Anna arvosana asteikolla 1-5 (5 on paras) siitä, kuinka hyvin katselmoinnit toimivat omassa organisaatiossa.

Haastattelun 1. vaihe - käytänteet

1. Kuinka säännöllisesti seuraavia dokumentteja katselmoidaan? Vastaa ainoastaan oman tietämyksesi pohjalta. Suorituksen täytyy täyttää ICM-tason 2 kriteerit eli katselmointien täytyy sisältää valmistautuminen ja raportointi. Selitys: 1 = ei koskaan, 2 = joskus, 3 = joka toinen kerta, 4 = usein, 5 = aina

| | | | | | |
|--------------------------------|---|---|---|---|---|
| Vaatimusmäärittely | 1 | 2 | 3 | 4 | 5 |
| Arkkitehtuurit | 1 | 2 | 3 | 4 | 5 |
| Muut suunnitelmat | 1 | 2 | 3 | 4 | 5 |
| Pienet komponenttisuunnitelmat | 1 | 2 | 3 | 4 | 5 |
| Koodi (keskeiset pätkät) | 1 | 2 | 3 | 4 | 5 |
| Koodi (kaikki osat) | 1 | 2 | 3 | 4 | 5 |
| Testisuunnitelmat | 1 | 2 | 3 | 4 | 5 |
| Projektisuunnitelmat | 1 | 2 | 3 | 4 | 5 |

Toteutetaanko säännöllisesti joissakin edellisissä kohdissa vapaamuotoisempaa katselmointia?

2. Mitä edellisen kohdan dokumentteja katselmoidaan asiakkaan kanssa ja kuinka usein?
3. Onko katselmointiprosessi määritelty? Mitä määrittely pitää sisällään?
4. Toteutetaanko katselmoinnit aina samalla tavalla? Siis noudatetaanko prosessia?
5. Kuinka katselmointien suunnittelu toteutetaan?
 - a) Aikataulukus (kuka hoitaa ja milloin). Sisältyvätkö katselmoinnit projektisuunnitelmaan?
 - b) Kuinka hoidetaan osallistujien valinta? Ovatko osallistujat ja roolit päätetty etukäteen vai päätetäänkö osallistujat juuri ennen katselmointia?
6. Kuinka paljon valmistautumiseen käytetään aikaa? Antaako prosessimäärittely suositusten käytettävistä ajasta? Valvotaanko valmistautumista? Mitä tehdään, jos joku ei ole valmistautunut?
7. Miten palaveri toteutetaan? Kuka johtaa? Miten etenee? Järjestetäänkö palaveria vai organisoidaanko katselmointi esimerkiksi sähköpostilla?
8. Mitä katselmoinnista raportoidaan? Miten tieto käsitellään (esim. tekstidokumentti tai erillinen sovellus)?
9. Kuinka korjauksien seuranta toteutetaan? Ohjeistetaanko prosessissa vai päätetäänkö tapauskohtaisesti? Järjestetäänkö usein uusi katselmointikierrös?
10. Järjestetäänkö katselmointikoulutusta? Mitä koulutus pitää sisällään?
11. Millaista katselmointimateriaalia on olemassa (tarkistuslistat, skenaariot ym.)? Onko materiaalia kehitetty kokemuksen perusteella? Onko materiaalia mukautettu eri rooleille?
12. Kuinka katselmointidataa kerätään? Seurataanko katselmointien todellista toteutumista datan avulla? Kuka seuraa ja miten? Käykö joku joskus läpi dataa kehittämismielessä?
13. Onko katselmointeihin selkeästi varatut resurssit työsuunnitelmassa?
14. Onko olemassa organisaation yhteinen politiikka katselmointien toteuttamisesta? Siis määritelläänkö jossain esimerkiksi, mitä tuotoksia projektissa pitäisi katselmoida?
15. Onko organisaatiossa määritelty selkeästi vastuuhenkilö, joka on vastuussa katselmointien toteutumisesta, valvonnasta ja kehittämisestä?

Haastattelun toinen vaihe – katselmointien ongelmat

- Mitä ongelmia katselmointien toimivuudessa on? Miksi ne eivät toimi ihanteellisella tavalla? Missä asioissa katselmointien tulisi toimia paremmin?

Kuinka tutuilta seuraavat ongelmat vaikuttavat omassa organisaatiossasi? Vastaa oman kokemuksesi pohjalta asteikolla 1-5 (1 = ei tuttu, 5 = erittäin tuttu). Mikä vastaa sinun kokemustasi parhaiten?

1. Aikataulu saattaa viivästyä katselmointia odotellessa.
2. Katselmointikokouksissa löydetään vain vähän uusia virheitä, vaikka ne vievät paljon aikaa.
3. Tukimateriaali (tarkistuslistat ym.) on huonoa.
4. Osallistuja ei ymmärrä katselmointiprosessia ja omaa rooliaan siinä.
5. Kritisoidaan tekijää eikä tuotosta
6. Katselmoinnit on huonosti suunniteltu.
7. Palaverissa keskitytään epäolennaisiin asioihin, kuten ratkaisemaan ongelmia.
8. Osallistajat ovat huonosti valmistautuneita.
9. Katselmointeihin osallistuu väärinä ihmisiä.
10. Roolien määrittäminen puuttuu.
11. Huonolaatuinen tuotos vie turhaan katselmoijien aikaa.
12. Työntekijöillä on huono asenne.
13. Katselmoinnissa on liikaa materiaalia kerralla.
14. Katselmointeihin ei ole selkeästi varattu resursseja yksilöllisiin työsuunnitelmiin.