

Oliver Saal ja työryhmä:

Olen saanut kolme arviota, jotka edustavat asiantuntemusta tutkimuksen aiheen ja menetelmien osalta. Kaikki arviot ovat positiivisia siinä määrin, että ne kannustavat lähettämään seuraavan version uudelleenarvioitavaksi. Arvioista käy ilmi useita tärkeitä seikkoja, jotka on oleellista ottaa huomioon seuraavassa versiossa. Tiivistän alle keskeisimmät asiat ja lisään joitain omia huomioita.

1. Mitattavaa käsitettä ja sen perusteita on selvennettävä. Kaikki arvioijat näkevät keskeiseksi haasteeksi sen, ettei mittauksen kohteella ole selkeää empiiristä, teoreettista tai muuta perustaa. Yksi looginen ratkaisu (Arvio 2) olisi niin ikään suorittaa erillinen EFA pilottiaineistolla ja sen jälkeen validoida syntyvä malli CFA:lla. Jos suunniteltua aineistonkeräystä on mahdollista lykätä, kevyemmän pilottiaineiston voisi kerätä erikseen. Jos suunniteltu keräys on teknisistä tai muista syistä pakko aloittaa keväällä/kesällä, toinen vaihtoehto olisi suorittaa keräys kahdessa osassa EFA-CFA järjestyksessä (tarvittaessa EFA:n täyttänyt henkilö voisi osallistua myös CFA-kyselyyn, kun aineistoja ei yhdistetä).
2. Jos kahteen keräykseen ei ole mahdollisuutta, on toki mahdollista kehittää mittari myös pelkästään CFA:n avulla. Tällöin käsitteen määrittely edellyttää huolellista teoreettista, kirjallisuuspohjaista ja aiempien mittareiden motivoimaa rakenteiden tarkastelua. Todennäköisesti luotettavin ratkaisu tällöin tulisi olemaan kavennettu määritelmä, joka ei pyrkisi mittaamaan maksimaalisen laajaa vaan selkeästi kohdennettua ilmiötä, jonka rajojen tulkinta olisi vähemmän kiistelty. Jos syystä tai toisesta on selkeä tarve luoda mittausperusta useille rinnakkaisille ilmiöille, on myös mahdollista luoda esim. kaksi linkittyntä mutta toisistaan riippumatonta mittaria. Tutkimuksen menetelmällinen osuus on suhteellisen vahva, mutta raportointiin liittyen jaan lähteen, joka "checklist" (lopussa) voi olla hyödyllistä soveltaa tässäkin tutkimuksessa: Jackson ym. (2009) <https://www3.nd.edu/~kyuan/courses/sem/readpapers/Jackson-reportingSEM09.pdf>
3. Kuten arvioijat korostavat, olisi tärkeää saada lisää tietoja katsauksesta, johon tutkimus nyt osittain perustuu (Arvio 1). Jos katsaus on valmis muttei julkaisu, olisi esim. mahdollista sijoittaa preprint-versio OSF-alustalle ja viitata katsauksen versioon niin, että lukijan on mahdollista tarkastaa esitettyjen väitteiden perusta. Myös muut relevantit liitteet, kuten mittareiden kieliversiot (Arvio 3), olisi tärkeää olla mukana seuraavassa käsikirjoituksessa (mikäli päätätte ottaa pidemmän EFA-CFA polun, lopullista mittaria ei tietenkään voi olla valmiina tässä vaiheessa).
4. Tutkimuskysymysten (TK) osalta lisään muutaman huomion, joita ei tarvitse seurata eksplisiittisesti, mutta toivon että ne auttavat suunnitelman parantamisessa. Koska kyse on mittarin rakentamisesta, faktuaalinen vastaus TK1:een on jo etukäteen "kyllä" – mitata toki *voi*, mutta mittauksen tarkkuus, valideetti jne. ovat avoimia kysymyksiä. Koska jo nykyisen version suunnitelmaan kuuluu erilaisten mallien testaaminen, tuntuisi että loogisin TK tässä työssä olisi jotain tämän suuntaista: "Minkälainen malli soveltuu parhaiten [käsitteen]

mittaamiseen kyselyn avulla?" Jos työ lähestyisi tutkimuskohdetta tällaisen TK:n näkökulmasta, hypoteeseja ei välttämättä tarvita lainkaan, sikäli kun kriteerit joita vasten validiteettia testataan on esitetty selkeästi (kuten monet niistä jo ovat).

5. Liittyen edelliseen, olisi hyvä listata validiteetin ehdottomat ja riittävät kriteerit yhteenvedona ja perustella, miksi juuri nämä on valittu. Käsikirjoitus tarjoaa jo nyt huolellisen suunnitelman ja on monelta osin vahva, mutta joitain tarkennuksia olisi hyvä tehdä. Kuten arvioijat mainitsevat, esim. nykyisen oletetun mallin rakenne ei tunnu täysin perustellulta aiemman tutkimuksen valossa. Jos on vaikkapa tiedossa, että naiset kohtaavat enemmän epäasiallista kohtelua kuin miehet, olisi loogista että omakohtainen_a ja omakohtainen_b ovat yhteydessä (koska vastaajien omakohtaisuus on vakio organisaation sisällä ja ulkona). Etenkin jos tutkimus päättyy tekemään pelkän CFA:n, olisi tärkeää käydä läpi tarkkaan kaikki keskeiset oletukset ja tekijät (sekä niiden perusteet). Tämä koskee myös oletuksia, jotka on nyt upotettu hypoteeseihin: onko empiirisiä, teoreettisia tai muita perusteita olettaa, että mittari ei toimisi joissain ryhmissä tai joillain yhteydenottomenetelmillä? Jos ei ole, en koe näiden testaamista tarpeelliseksi. Tämä ei tietenkään poissulje sitä mahdollisuutta, että työssä raportoidaan kokeilevasti vaikkapa liitteessä ryhmien mahdolliset erot.
6. Arvioijat ovat jo jättäneet paljon menetelmällistä palautetta, joten lisään vain pari seikkaa. En huomannut mallin kuvauksissa cross-loadings-rajoja; tämän huomion voi sivuuttaa jos ne menivät minulta ohi. Arvioijat pohtivat myös, onko mittarin validointi pelkästään työuupumuksen perusteella optimaalista ja se onkin relevantti kysymys. Olisiko mahdollista selvittää aiemmasta kirjallisuudesta, mitkä muut korrelaatiot voisivat olla oleellisia? Esim. Ilmarisen työkyvyn mittarista on olemassa napakka 1-kysymyksen versio, jonka lisääminen ei pidentäisi kyselyä juurikaan, mutta toisi helposti lisää tietoa (huom. en tiedä onko epäasiallisella kohtelulla osoitettua yhteyttä alhaiseen työkykyyn).
7. Monille arvioijille on epäselvää, mihin tutkimuksen efektikoko viittaa. Oletan, että tällä viitataan mittarin ja työuupumuksen väliseen korrelaatioon. Suunnitelmassa on käytetty Cohenin valmiita arvoja, mutta (kuten Arvioija 1 myös esittää), on epäselvää mikä suuruusluokka vastaisi "kultaista" korrelaatiota näiden kahden käsitteen väillä (ei liian suuri, ei liian pieni). Jos maailmalla on luotettavia epäasiallisen kohtelun mittareita, yksi mahdollisuus olisi poimia niiden aiemmat efektikoot ja määrittää tämän perusteella toimivan mittarin raja-arvot. Vaikka tämä ei tietenkään yksin vahvista mittarin toimivuutta, loisi se huomattavasti vahvemman perustan validiteetille (ja tekisi siitä ainakin osittain vertailukelpoisen kv-mittauksen kanssa).
8. Aineiston osalta olisi vielä hyvä varmistaa, ettei mukaan jää huolimattomia vastaajia. Yksi hyväksi todettu tekniikka näiden tunnistamiseksi on Mahalanobis-menetelmä, mutta toki muitakin on. Jätän pohdittavaksi, mikä sopii tähän tutkimukseen, mutta Curran (2016) on tämän suhteen hyvä lähde: <https://doi.org/10.1016/j.jesp.2015.07.006>

Yhteenvedona: suunnitelma on lupaava ja tutkimus on potentiaalisesti arvokas. On luonnollista, että tutkijoiden aika, resurssit ja muut käytännölliset asiat rajoittavat sitä,

miten perusteellisesti mittarin kehitys on mahdollista tehdä. Jätän pohdittavaksi, mikä on mahdollista. Oleellista on, että kehitetyn mittarin perusta on vankka: sen myötä tulevien mittarien kehittäjien/päivittäjien on mahdollista jatkaa näistä löydöksistä, eikä prosessia tarvitse aloittaa tulevaisuudessa kokonaan alusta. Kannustan tutkijoita siis keskittymään vankan empiirisen ja teoreettisen perustan rakentamiseen, mikä vie käsitteen/käsitteiden mittauksen kehitystä eteenpäin ja takaa selkeän kontribuution.

Pyydän lukemaan kaikki arviot huolellisesti ja vastaamaan niissä esitettyihin huomioihin yksitellen vastauskirjeessä. Pyydän myös liittämään seuraavaan versioon liitteeksi käsikirjoituksen kopion, jossa muutokset on merkitty värillä tai muilla tavoin (jos muutokset ovat suuria ja rakenne muuttuu kauttaaltaan, tämän voi sivuuttaa). Muistutan, että vastauskirje on julkinen ja se arkistoidaan osaksi RR-keskuksen arviointihistoriaa. Mikäli jokin arvioiden tai itseni esittämistä huomioista kaipaa tarkennusta, voitte ottaa minuun suoraan yhteyttä koska tahansa prosessin aikana.

Veli-Matti Karhulahti

Arkistohuomio: tutkijat ottivat yhteyttä RR-keskukseen tammikuussa 2024, mikä mahdollisti arvioijien aikaisen etsimisen ja ajan varaamisen niin, että kolme arviota saapuivat heti 2 viikon sisällä käsikirjoituksen saapumisesta.

Arvio 1: 21.3.2024

Kiitos mahdollisuudesta arvioida alkamassa oleva tutkimus epäasiallisen kohtelun mittarin validoinnista pelastus- ja ensihoitoaloille. Alla muutamia huomioitani:

- Olemme selvittäneet äskettäisessä tutkimushankkeessa epäasiallisen kohtelun ja asiakasväkivallan riskin ennustamisen mahdollisuuksia sekä toteuttaneet ko. konfliktitilanteiden vähentämiseen tähdänneen kehittämishankkeen kunta-alalla. Hankkeesta on useita julkaisuja, joista saattaisi olla kirjoittajille apua johdantoon ja validiteettitutkimuksen suunnitteluun:

[Työnantajan vaikutusmahdollisuudet epäasialliseen kohteluun ja asiakasväkivallan uhkaan \(julkari.fi\)](#)

[Prediction of violence or threat of violence among employees in social work, healthcare and education: the Finnish Public Sector cohort study - PubMed \(nih.gov\)](#)

[Association of workplace violence and bullying with later suicide risk: a multicohort study and meta-analysis of published data - PubMed \(nih.gov\)](#)

[Prediction of bullying at work: A data-driven analysis of the Finnish public sector cohort study - PubMed \(nih.gov\)](#)

[Effectiveness of a workshop-based intervention to reduce bullying and violence at work: A 2-year quasi-experimental intervention study - PubMed \(nih.gov\)](#)

- Johdannossa keskeistä on kertoa, onko jo olemassa validoitua mittaria. Tässä kohtaa kirjoittajat viittaavat julkaisemattomaan kirjallisuuskatsaukseen, mikä saattaa tässä kohtaa prosessia olla perusteltua, jos on oletettavissa, että julkaisu tapahtuu ennen kuin tämä käsikirjoitus on lopullisesti valmis. Mikäli niin ei tapahdu, kirjallisuuskatsaukseen viittaaminen ei tule onnistumaan ja johdanto menee pitkälti uusiksi sillä kirjallisuuskatsauksen tulokset (+keskeiset viitteet) pitää lisätä tähän artikkeliin.
- Pidän intuitiivisesti ja jossain määrin myös juridisesti (eri lakipohja: rikoslaki vs. työturvallisuuslaki) heikkona ratkaisuna sitä, että epäasiallisen kohtelun kattokäsitteen alle on sisällytetty kaikki: kiusaaminen, syrjintä, häirintä ja väkivalta. Ratkaisussa (taulukko 1) viitataan Jäntin ja Nordqvistin artikkeliin vuodelta 2022, mutta se artikkeli keskittyi vain työpaikkaväkivaltaan ja siinä _väkivaltaan_ on sisällytetty seksuaalinen häirintä. Eli kehottaisin kirjoittajia vielä kerran pohtimaan, mikä olisi oikea kattokäsite työpaikkakiusaamiselle, syrjinnälle, häirinnälle ja väkivallalle. Olen itse käyttänyt jossain yhteyksissä käsitettä työpaikan konfliktitilanteet, mutta tämä ei ole ollut erityisen teoreettinen, tai teoriaan perustuva ratkaisu, joten en väitä, että se olisi se oikea. Tärkeää on kuitenkin erotella työyhteisön sisältä ja toisaalta ulkopuolelta (asiakkaat, potilaat yms.) tulevat tilanteet sekä työsuojelulailliset ja rikoslakitilanteet. Tärkeää on myös erotella työnantajan syrjivä toiminta/käytännöt esimerkiksi kollegoiden tekemästä syrjinnästä. Jotenkin se, että esimerkiksi puukonheiluttelua tai nyrkillä lyömistä kuvaa "epäasiallisena kohteluna", tuntuu tilanteen aliarvioimiselta.

Minusta olisi siis tärkeää, että väkivalta on väkivaltaa eikä siitä puhuta epäasiallisena kohteluna.

- Terveysvaikutusten redusointi työuupumukseen tuntuu kapealta lähestymistavalta. Ensinnäkin, esimerkiksi väkivallalla lienee myös ihan somaattiseen terveyteen liittyviä vaikutuksia. Toiseksi, myös mielenterveysvaikutukset voivat olla työuupumusta laajempia. Tämä valinta vaatii enemmän perusteluja tai uudelleen pohdintaa.
- Tutkimuskysymyksistä ei käy selväksi, miten työuupumus liittyy mittarivalidointiin
- Alaikäisten kohderyhmä vaatii mielestäni lisäpohdintaa. Miksi alle 18-vuotiaat ovat mukana? Miksi on tarpeen tutkia mittarivaliditeettia alaikäisillä opiskelijoilla, joilla ei voi varsinaisesti olla paljoakaan alan työkokemusta?
- Onko tutkimus läpikäynyt eettisen arvioinnin? Tämä on ensiarvoisen tärkeää tutkimuksen julkaisun kannalta.
- Tilastollisen voiman arvioinnissa puhutaan efektikoosta, mutta mikä tässä on se odotettavissa oleva yhteys, eli efekti? Tutkitaanko siis, onko "epäasiallisen kohtelun" mittari yhteydessä työuupumukseen? Viittaa tässä aiempiin kohtiin arviossani. Tuntuu hieman köykäiseltä, että ulkoinen validiteetti todetaan riittäväksi, jos havaitaan mittarin korreloivan työuupumukseen. Tarkalleen ottaenhan siitä voitaisiin päätellä ainoastaan, että mittarinne mittaa jossain määrin samaa asiaa kuin työuupumusmittari. Kirjoittajat toteavat kyllä, että erotteluväliditeettia mitataan vastaavasti _liian suurilla_ korrelaatioilla. Useinhan mittarivaliditeettia tutkitaan suhteessa "gold standard" mittariin. Eli jotta voisitte tutkia, mittaako epäasiallisen kohtelun mittarinne riittävästi tutkittavan ilmiön ominaisuutta, sitä tulisi verrata johonkin mahdollisimman objektiiviseen tietoon/mittariin ko. ilmiöstä, eli epäasiallisesta kohtelusta. Tässä siis ajatellaan, että mikäli mittari mittaa riittävän suuressa, mutta ei liian suuressa määrin työuupumusta, sen voidaan tulkita mittaavan validisti epäasiallista kohtelua. Näen, että tässä on tutkimuksen suurin kompastuskivi. Kehotan miettimään vielä uudelleen ja laajemmin sitä, mihin kehittävää mittaria verrataan.

Yhteenvedon totean, että näen tutkimuksen edessä kolme merkittävää ongelmaa tai vähintään asiaa, jotka tulee selkiyttää: 1) eettinen käsittely; 2) epäasiallinen kohtelu kattokäsitteenä monenlaisille tilanteille ja jopa fyysiselle väkivallalle; 3) mittarivaliditeettitarkastelun rakentaminen työuupumusmittariin vertailun varaan. Lopuksi toivotan tutkijoille onnea tutkimuksen jatkokehittelyyn ja toivon, että arviostani on teille hyötyä.

Ystävällisin terveisin

Jenni Ervasti

johtava tutkija, FT, dos., Työterveyslaitos

Arvio 2: 25.3.2024

Epäasiallisen kohtelun mittarin validointitutkimuksen suunnitelmassa esitetty tutkimuskysymys on merkityksellinen, ja suunniteltu tutkimus pystyy vastaamaan siihen osittain. Kuitenkin suunnitelman perusteella on epäselvää, mitä tutkimus itse asiassa pyrkii valdimoimaan, mikä validoitavan mallin merkitys on tutkimukselle, ja miten laajasti kyseistä mallia voitaisiin soveltaa sekä suomalaisessa että eri kulttuuriympäristöjen tutkimusasetelmissa.

Kirjoittajat kuvaavat soveltavansa epäasiallisen kohtelun määrittelyyn Tolhurstin ja kollegoiden (1999) kehittämää työpaikkaväkivallan kuutta ulottuvuutta, ja täydentävänsä näitä ulottuvuuksia Alspachin (2008) kuvaamalla vihamielisyyden ulottuvuudella sekä yhteisvertaisuuslain syrjinnän määritelmällä. Tolhurstin ja kollegoiden määrittelyä he muokkaavat yhdistämällä väkivaltaan sekä toisaalta häirintään viittaavan epäasiallisen seksuaalisen kohtelun yhdeksi kokonaisuudeksi. Lukija olettaa kirjoittajien päätyvän näin seitsenulotteiseen epäasiallisen kohtelun malliin. Kuitenkin Taulukon 1 alla kirjoittajat toteavat käyttävänsä seitsenulotteista mittaristoa, koska yhdistävät "ulkoisen ja sisäisen työpaikkaväkivallan yhdeksi käsitteeksi". Tämä kuvaus jättää lukijan epäselvyyden tilaan sen suhteen, miten seitsemään ulottuvuuteen todellisuudessa päädyttiin. Kattavammat viittaukset aiempaan kirjallisuuteen sekä kuvaukset siitä, minkälaisiin teoreettisiin ja empiirisiin malleihin – sekä minkälaiseen dataan – Tolhurstin ja kollegoiden sekä Alspachin tutkimuksen perustuvat auttaisivat arvioimaan ehdotetun seitsenulottuvuudensis­en mallin perusteltavuutta ja riittävyyttä.

Kirjoittavat perustelevat suunniteltua tutkimusta seuraavilla asiantiloilla: aikaisempien tutkimusten välillä ei ole syntynyt konsensusta epäasiallisen kohtelun määrittelylle tai mittaamiselle, aiempia mittareita on kehitetty muissa kulttuuriympäristöissä kuin Suomessa, sekä mittaristojen kohderyhmätasoinen validointi on ollut puutteellista. Rekisteröidyn tutkimusraportin suunnitelmaosiossa tulisi avata nykyistä enemmän aikaisempien instrumenttien soveltumattomuutta suomalaiseen ympäristöön. Missä kulttuurisissa konteksteissa nämä aiemmat mittarit ovat kehitetty ja miksi voidaan perustellusti olettaa, että ne soveltuisivat huonosti suomalaisiin käyttöyhteyksiin? Entä miten suunnitelmassa kuvattu validaatio ylittäisi nämä aiemman tutkimuksen haasteet vai tulisiko nyt suunniteltu tutkimus tulkita sekin kontekstuaaliseksi mittariksi, joka kehitetään ja validoidaan yksinomaan suomalaiseen kontekstiin soveltuvaksi instrumentiksi?

Tutkijat esittävät ensimmäisenä tutkimuskysymyksenään: "Voiko epäasiallista kohtelua mitata määritellyn seitsemän ulottuvuuden avulla, tekijä- ja tekokontekstista riippumatta?" Tämä sinänsä selkeä tutkimuskysymys hämmentää lukijaa kahdesta syystä. Ensinnäkin pidän erikoisena asetelmaa, jossa kirjoittajat pyrkivät tutkimaan seitsenulottuvuudensis­en mittarin kontekstikohtaisuuden ylittävää validiteettia. Tämä on erikoista siksi, että he perustelivat edellä suomen­kielisen mittariston kehittämistä toteamalla aikaisempien mittareiden kehitetyn Suomesta poikkeaviin kulttuuriympäristöihin. Tämä luulisi indikoivan, että epäasiallisen kohtelun mittari on

lähtökohtaisesti jotakin, jonka validiteettia tulee tutkia suhteessa kuhunkin kulttuuriympäristöön sekä siinä vallitseviin tekijä- ja tekokonteksteihin.

Toinen lukijaa hämmentävä seikka on ensimmäistä paljon merkittävämpi, mitä tulee Tutkimuskysymyksen 1 merkityksiin suunnitellun tutkimuksen kannalta. Tutkimuskysymys 1 ja suunnitellun tutkimuksen kuvaus tutkimuskysymyksiä edeltävissä tekstiosioissa alleviivaavat epäasiallisen kohtelun koostuvan seitsemästä ulottuvuudesta. Ensinnäkin, koska tutkijat itse kuvaavat ettei ilmiön ymmärryksestä saati sen mittaamisesta ole olemassa konsensusta, esitetty seitsenulottuvuudellinen ilmiön rakenne on sekin väistämättä hypoteettinen. Näin ollen on avoin empiirinen tutkimuskysymys, rakentuuko epäasiallisen kohtelun mittari todella seitsemästä ulottuvuudesta – ja jos rakentuu, vastaavatko tunnistetut seitsemän ulottuvuutta kirjoittajien esittämää teorialähtöistä kuvausta. Suunnitelman merkittävin puute on, että tämä seitsemän ulottuvuuden hypoteettinen rakenne otetaan annettuna eikä siihen kohdennu suunnitelman mukaan empiiristä tarkastelua. Sen sijaan mittarin oletetaan olevan yksiulotteinen vaikka tämä oletama eroaa suoraan kirjoittajien omista argumenteista ja esitetystä 7-ulotteisesta rakenteesta.

Tutkimuskysymyksen jälkeen esitetyt hypoteesit mainitsevat neljän faktorin mallin, mutta lukijalle on hyvin epäselvää, mistä nämä neljä faktoria nousevat, ja mikä on niiden teoria- tai aineistolähtöinen perustelu. Asiaan saadaan tarkennusta vasta mittareita kuvaavassa alaluvussa, jossa 4-faktorisuus yhdistetään metodologiseen ratkaisuun mitata epäasiallista kohtelua "neliosaisella mittarilla, jonka jokaisessa osassa on yksi mittari per epäasiallisen kohtelun ulottuvuus", tarkemmin: "omakohtaiset kokemukset organisaation sisällä", "havaitut kokemukset organisaation sisällä", "omakohtaiset kokemukset organisaation ulkopuolella" ja "havaitut kokemukset organisaation ulkopuolella". Kirjoittajat esittelevät tämän neliosaisen mittarirakenteen muodostavan 4-faktorisen mallin, koska *a priori* olettavat kunkin osan mittaavan yksiulottuvuuksista epäasiallista kohtelua seitsemän indikaattorin mittarilla.

Sen lisäksi, että epäasiallisen kohtelun mittarin dimensionaalisuus (onko kyseessä 7-ulotteinen, yksiulotteinen tai jokin muu faktorirakenne) on avoin empiirinen tutkimuskysymys, on myös erikoista tarkastella mallia 4-faktorin rakenteena perustuen metodologiseen ratkaisuun kysyä epäasiallisesta kohtelusta keskenään eri tavoilla. Arvioni mukaan tällöin olisi kyse neljästä yksifaktorisesta mallista (sillä oletuksella että yksiulotteinen malli todella voitaisiin konfirmoida), joiden välisiä suhteita ja riippuvuuksia tutkimus pyrkii tarkastelemaan. Sen sijaan epäselväksi jää, miksi nämä neljä oletettavasti yksiulottuvuuksista mallia yhdessä muodostaisivat teoreettisesti perustellun kokonaisuuden eli nelifaktorisen mallin. Lisäksi tämän mallin kohdalla herää kysymys siitä, onko perusteltua yrittää tuoda yhteen malliin omakohtaisen kokemuksen arvioinnin näkökulmia sekä toimintaympäristön havainnointiin perustuvia näkökulmia. Tämä oman toiminnan sekä toimintaympäristön havainnoinnin itsearviointin yhdistelmä pitäisi vähintään problematisoida ja perustella, mielellään yhdistämällä nämä perustelut sekä teoriaan että aiempiin empiirisiin tutkimuksiin, joissa vastaavaa lähestymistapaa on sovellettu menestyksekkäästi.

Tutkimussuunnitelman erilaisten validointilähestymistapojen arviointi on hyvin vaikeaa ja osin mahdotonta juuri siksi, että koko asetelma on rakennettu olettamalle, että itse epäasiallisen kohtelun mittari on yksiulotteinen. Tutkimusasetelmaa ei tästä syystä ole kuvattu riittävällä tarkkuudella, eikä toteutusta siten voida arvioida suunnitelman nykytilassa täsmällisesti. Jotta yksiulotteisen epäasiallisen kohtelun mittarin kehittäminen ja myöhemmin validointi olisi mahdollista, tulisi tutkia ensin ilmiön oletettua seitsemää ulottuvuutta.

Ehdotan kirjoittajia harkitsemaan asetelmaa uudestaan siten, että ensinnäkin seitsenulotteisen mallin teoreettinen perustelu avattaisiin nykyistä kattavammin kuvaamalla, miten aiempi tutkimus on teorioiden sekä empirian näkökulmasta päätynyt dimensioihin, joista kirjoittajat johtavat seitsemän ulottuvuuden mallinsa. Lisäksi tämä johtaminen pitää perustella huolellisesti ja esittää hypoteettisena rakenteena sillä, kuten kirjoittaja toteavat, aiempi tutkimus ei ole päätynyt yhteisymmärrykseen ilmiön määrittelystä tai mittaamisesta.

Käytännössä tämä tarkoittaa, että tutkijoiden tulisi ensinnä kehittää epäasiallisen kohtelun mittari siten, että "kysymyspooli" kattaisi kaikki seitsemän oletettua ulottuvuutta tavalla, joka vastaa faktorianalyysin hyviä käytänteitä. Näin ollen kullekin oletetulle seitsemälle dimensiolla tulisi kehittää teoria- ja aineistolähtöisesti sekä fokusryhmätyöskentelyllä vähintään 3-5 kysymystä/indikaattoria, ja että kyselyvastaajat vastaisivat näin muodostettuihin kysymyksiin kaikkiin keskenään samalla asteikolla. Mittarille tulisi sitten kerätä kattava määrä kyselyvastaajia, mieluiten sekä suomen- että ruotsinkielisiä. Aineiston ei kuitenkaan tarvitsisi olla sidoksissa työnkuviin vaan aineisto voitaisiin kerätä laajalla rajauksella koko populaation ikä- ja sukupuolirakennetta vastaavalla tavalla. Näin siksi, että valtakunnallisesti edustava aineisto toisaalta vastaisi joka tapauksessa riittävällä tarkkuudella tarkasteltavana olevaa kulttuurista kontekstia, ja toisaalta edustava aineisto antaisi hyvän lähtökohdan kyselymittarin aineistolähtöiselle lyhentämiselle. Mittarille tulisi mieluiten tehdä sekä eksploratiivinen faktorianalyysi että erillinen konfirmatorinen faktorianalyysi, ja nämä tulisi tehdä keskenään eri aineistoille. Eksploratiivisen faktorianalyysin perusteella tutkijat saisivat jo käsityksen siitä, tukeeko empiria seitsenulotteisen mallin teoriaa vai syntyykö näin toinen tulkinta epäasiallisen kohtelun dimensionaalisuudesta. Vasta kattavalla kyselyinstrumentilla tehdyn eksploratiivisen faktorianalyysin jälkeen mittarin lyhentäminen on perusteltua, ja toisaalta konfirmatorinen faktorianalyysi sitten voi validoida tämän rakenteen.

Suunnitelmassa kuvattu seitsemän indikaattorin ja yhden faktorin epäasiallisen kohtelun malli on kirjoittajien omaa argumentaatiota seuraten toisen asteen faktorianalyysi, jossa kutakin oletettua seitsemää ulottuvuutta edustaa "single-item" -tarkastelu. Tällaisessa asetelmassa on suuri riski sille, että mallin kutakin oletettua seitsemää faktoria edustavat kysymykset sisältävät suuren virhemarginaalin, mikä osaltaan laskee faktorilatauksia ja tekee validoinnista teknisestikin todella vaikeaa tai mahdotonta. Toisin sanoen, nyt esitetty suunnitelma hyppää usean tutkimusvaiheen yli esittäessään

olettamuksen siitä, että epäasiallisen kohtelun mittari olisi yksiulotteinen, ja että sitä voisi perustellusti mitata seitsemällä yksittäisellä indikaattorilla, joista kukin edustaisi yhtä olettua epäasiallisen kohtelun ulottuvuutta. Lisäksi jää avoimeksi kysymykseksi, miten nämä seitsemän yksittäistä "single-item" -kysymystä muotoiltaisiin sellaisiksi, että ne varmuudella mittaisivat juuri argumentoituja ulottuvuuksia. Suunnitelmassa tulisi kuvata tämä prosessi ja siihen liittyvä asiantuntijatyö sekä testaaminen kattavasti.

Edellä mainitut seikat eivät tietenkään tarkoita sitä, etteikö epäasiallista kohtelua voisi viime kädessä perustellusti mitata yksiulotteisella mittarilla, joka mahdollisesti rakentuisi tismalleen seitsemästä indikaattorista. Kyseessä on kuitenkin avoin empiirinen ja osin teoreettinenkin kysymys, jonka selvittämiseksi tarvitaan laajaa eksploratiivista (EFA) ja konfirmatorista faktorianalyysia (CFA). Lisäksi vasta näin tehtäessä voidaan myös datalähtöisesti muotoilla ja valita mittarin kysymyksille ne spesifit sanamuodot ja kuvaukset, jotka parhaalla tavalla kuvaavat kutakin seitsemää ulottuvuutta – sillä oletuksella, että EFA ja CFA todellakin vahvistaisivat juuri tällaisen rakenteen.

Nyt esitetty suunnitelma lähtee siitä, että edellä mainitut validaatiotutkimuksen vaiheet olisivat jo tehty, ja esittelee siten analyysit, jotka osittain vastaavat toisen asteen faktorianalyysin periaatteita sekä toisaalta rakenneyhtälömallinnuksen periaatteita (esim. Alaluku 2.3.1). Kuitenkin validaatiotutkimuksen rakenneyhtälömallinnusta edeltää aina mittamallin konfirmointi, ja tässä tapauksessa mittamallin konfirmointi edellyttää yllä mainittuja EFA- ja CFA-aineistonkeruita sekä asianmukaisia analyyseja. Vasta mittamallin ja sen kunkin dimension validointi mittamallissa mahdollistaa tutkijoiden kuvaaman tutkimusastelman, jossa selvitetään epäasiallisen kohtelun henkilöön kohdistuvia *terveysvaikutuksia*. Lisäksi toisen asteen konfirmatoriset mittamallit tulisi perustella teoria- ja aineistolähtöisesti sen sijaan, että testattaisiin mallien soveltuvuutta dataan ja valittaisiin sitten parhaiten sopiva vaihtoehto (ks. Alaluku 2.3.1) Luvun 2.5. Tilastollinen voima ja alfataso kohdalla lukijalle on epäselvää, mihin mallin efektikoolla viitataan? Minkä muuttujien tai faktoreiden efekteistä mihin vastemuuttujaan tai latenttiin tässä on kyse? Invarianssitestaamisen kohdalla herää kysymys siitä, minkälaiset ryhmäkoot mahdollistavat kyseisen analyysin kohderyhmien, ammattiryhmien, kieliversioiden ja yhteydenottomenetelmien välillä.

Myös aineiston hankintaan ja käsittelyyn liittyy joukko avoimia ja kriittisiä kysymyksiä: Miten kyselyaineiston keruussa ja aineiston käsittelyssä huolehditaan anonymiteetistä? Kyselyssä on mukana terveysdataa, ja siitä syystä johtuen tämä seikka on erityisen tärkeä. Entä mikä on tutkimuksen suunnitelma eettisten tutkimuslupien suhteen? Suosittelen, että tutkijat laativat näistä asioista yksityiskohtaisen suunnitelman ja liittävät sen osaksi rekisteröidyn tutkimusraportin ensimmäistä arviointivaihetta. Olisi myös perusteltua laatia hankkeen tietosuojaseloste sekä perustella siinä kaikki Tutkimuseettisen neuvottelukunnan edellyttämät tietosuojaan liittyvät seikat, mukaan lukien aineiston säilytys-, käsittely- sekä arkistointisuunnitelma sekä erittely henkilötietojen keruusta ja käsittelytavoista.

Arvio 3: 3.4.2024

Arvioitavan tutkimuksen tarkoituksena on kehittää epäasiallisen kohtelun suomen- ja ruotsinkielinen mittaristo pelastusalan ja ensihoidon eri organisaatioiden kohderyhmille osana Pelastusopiston toteuttamaa tutkimushanketta "Pelastustoimen nollatoleranssi syrjinnälle ja häirinnälle – faktaa vai fiktiota?" Epäasiallista kohtelua on mitattu monella erilaisella mittarilla sen vuoksi että epäasiallinen kohtelu käsittää laajan kirjon eri asioita mm. syrjintä, häirintä, väkivalta, työpaikkakiusaaminen tms. Mittarin valintaan vaikuttaa myös se, mitataanko potilaiden vai henkilöstön eri tasojen suunnalta ilmenevää epäasiallista kohtelua. Tässä isona haasteena on saada nämä asiat sovitettua yhteen mittariin. Käsikirjoitus on selkeästi kirjoitettu, mutta useampi kohtaa kaipaava selvennystä:

1. Johdannon alku (sivu 1) nojautuu vahvasti populaariuutisiin. Aiheen ajankohtaisuuteen toki sopii perustella tutkimusta sitä kautta, mutta johdannon teoriaosan tieteellisyyttä tulisi syventää ja selvittää tarkemmin mihin teoreettisiin lähtökohtiin tutkimus perustuu.
2. Tutkimuksen tarkoitusta tulee tarkentaa ottaen huomioon epäasiallisen kohtelun moninaisuuden ja sen ilmenemisen potilaiden ja henkilöstön osalta. Tätä tulee selventää ja harkita aiheen rajaamista.
3. Johdannon lopussa s. 2, 1 kpl mainitusta tutkimushankkeesta tulee mainita mm. rahoittaja. Tämän tutkimuksen kysymys näyttää pohjautuvan kirjallisuuskatsaukseen Saal ym. joka on lähdeluettelon mukaan julkaisematon käsikirjoitus, lähetetäänkö se tieteelliseen kanavaan arviointiin? Samassa kohdassa mainitaan että tutkimushanke pohjautuu asiantuntijaryhmien työhön, tätä tulisi avata tarkemmin.
4. Kappaleessa 1.1 kuvataan epäasiallisen kohtelun määrittelyä ja mittaamista. Tässä tulee avata tarkemmin Tolhurstin ym. 1999 mittarin kuutta ulottuvuutta.
5. Taulukossa 1 esitetyssä epäasiallisen kohtelun ulottuvuuksissa ensimmäinen kohta 'verbaalinen väkivalta' vaikuttaa olevan päällekkäinen mm. kohtien 'uhkailu', 'vihamielisyys' ja 'syrjintä' kanssa, tätä tulee vielä muokata niin etteivät ulottuvuudet ole päällekkäisiä. Tässä ei tuoda esiin somekanavien välityksellä tuotettua epäasiallista kohtelua, joka tulee huomioida.
6. Kohdassa 1.2 kuvattu epäasiallisen kohtelun vaikutukset eivät näy tutkimuksen tarkoituksessa tai 1.3 kohdan tutkimuskysymyksissä ja tämä jää siten tässä irralliseksi. Epäasiallisen kohtelun vaikutukset tulee joko sisällyttää käsikirjoituksen tarkoitukseen tai jättää pois käsikirjoituksesta.
7. Tutkimuksen kohderyhmän kuvaus vaikuttaa perusteelliselta, tosin haasteena on se miten saada tutkimukseen mukaan sopimuspalokuntien ja palokuntayhdistysten suuri joukko henkilöitä, jotka eivät ole päätoimisia. Tämän lisäksi heidän toimenkuvansa eroo päätoimisista pelastusalan työntekijöistä ja siten on syytä olettaa etteivät he altistu epäasialliselle kohtelulle samassa määrin kuin päätoimiset pelastustoimen työntekijät.
8. Yhteydenottomenetelmät aineiston keruussa keskittyvät johdon/toimistojen kautta välitettyihin viesteihin. Kohdassa 2.2 tulee kuvata tarkemmin osallistujien anonymiteetti ja tutkimuseettiset lähtökohdat aiheen sensitiivisyyden huomioon ottaen jotta kyselyssä saadaan aidosti esiin henkilöiden omia kokemuksia aiheesta.

9. Kohdassa 2.3.1 kuvataan epäasiallisen kohtelun mittaria, tähän tulee liittää mukaan tässä tutkimuksessa käytettävät suomen- ja ruotsinkieliset mittarit (erillisenä liitteenä).
10. Epäasiallisen kohtelun mittari keskittyy lähinnä sen kokemuksen/havaintojen määriin, mutta toinen asia on, miten häiritseväksi epäasiallinen kohtelu koetaan ja miten vakavaa se on. Onko tätä mahdollista ottaa huomioon tässä tutkimuksessa?
11. Epäasiallisen kohtelun mittarin ruotsinkielistä käännöstä ja sen validointia suomenkieliseen verrattuna tulee kuvata, tämä puuttuu nyt kokonaan.