

Kiitämme arvioijaa ja suosittelijaa kehittävästä ja asiantuntevista huomioista. Olemme käsitelleet jokaisen aiheen erikseen tässä dokumentissa. Muutosten rivinumerot, taulukko- tai kuvionumerot sekä osiot on merkitty **keltaisella korostusvärillä**. Päivitetyssä käsikirjoituksessa muutokset on korostettu **punaisella tekstivärillä**. Myös liitetiedostoihin on tehty päivityksiä, ja muutokset merkitty punaisella tekstivärillä.

Käsikirjoituksen arvioinnin ulkopuoliset päivitykset

Suosittelijan luvalla olemme tehneet seuraavat päivitykset ensimmäisessä arviointikierroksessa arvioituun esirekisteröityyn suunnitelmaan:

- Lisäsimme kyselylomakkeen kysymyksenasettelut, alakysymykset ja vastausvaihtoehdot taulukkona **osioon 2.4.1.** (**taulukko 3**, ks. Arvio 3, kohta 24). Järjestimme samalla osion tekstiä hieman uudelleen ja tiivistimme tekstiä jonkin verran **[210–216]**.
- Yhdenmukaistimme lukujen muotoilut nykyisen arvion arvioijan pyynnöstä (ks. Arvio 3, kohta 32). Yhdenmukaistaminen vaikutti myös tekstimuotoisiin liitteisiin.
- Lisäsimme huomion suunnitelmasta poikkeamisesta EFA-aineiston dimensionaalisuuden arvioinnissa nykyisen arvion arvioijan pyynnöstä (ks. Arvio 3, kohta 25).
- Korjasimme **osiossa 2.6.2.** kohdan dynaamisten raja-arvojen laskennan mahdottomuudesta, sillä nykyisen arvion arvioijan kommenttien myötä pystyimme sittenkin suorittamaan ensimmäisessä kierroksessa suunnitellun dynaamisen raja-arvojen simuloinnin (ks. Arvio 3, kohta 31, alakohta 2). Muutos palauttaa siis tältä osin **osion 2.6.2.** vastaamaan ensimmäisessä kierroksessa tehtyä suunnitelmaa.
- Teimme kielenhuoltoa ja tekstin tiivistämistä **osioissa 1, 1.1., 1.2., 1.3., 2.1., 2.2., 2.3., 2.4.1., 2.5., 2.6.1. sekä 2.6.2.** Muutokset on tehty muuttamatta lauseiden merkitystä tai sisältöä.
- Vaihdoimme mahdollisuuksien mukaan vakiintuneisiin lyhenteisiin tiivistääksemme tekstiä. Tämä koskee **osioita 2.6.1. ja 2.6.2.**

Näiden lisäksi olemme tehneet seuraavat arvioinnin ulkopuoliset päivitykset ja muutokset tekstiin:

- Tarkistimme uudelleen invarianssitestauksen ongelmat kohderyhmien vertailun osalta. Kohderyhmävertailussa oli aikaisemmin todettu, ettei vertailua voitu suorittaa vähäisen varianssin myötä. Epäilimme syyksi kolmatta, määrällisesti varsin pientä kohderyhmää (järjestöt, liitot ja koulut). Koetimme poistaa tämän kohderyhmän vertailusta, mikä mahdollisti invarianssitestauksen. Olemme siten lisänneet kohderyhmäinvarianssitestauksen mittausinvarianssin taulukkoon (**taulukko 8**) sekä lisänneet testausta käsittelevän lauseen **[415–416]**. Olemme myös korjanneet johtopäätösten tekstiä vastaamaan nyt onnistunutta invarianssitestausta **[460–463]**.
- Sovittaaksemme käsikirjoituksen pituuden toivotun yhteistyölehden merkkirajaan olemme siirtäneet useita osioita liitteiksi. Osiot valittiin niiden verrannollisesti vähäisemmän tärkeyden vuoksi – osioiden tiedot ovat tärkeitä ymmärtääkseen analyysien luotettavuuden, mutta eivät vaikuta laajemmin tulosten tulkintaan. Numeroimme myös liitteet uudelleen vastaamaan niiden esiintymisjärjestystä käsikirjoituksessa. Osiot olivat:
 - o Vanha osio 2.6. ”Tilastollinen voima ja alfataso” **[liite 2]**
 - o Vanha osio 2.7. ”Katoanalyysi” sekä vanhasta osiosta 4.1 ”Otoksen kuvailu ja katoanalyysi” kaikki katoanalyysiin liittyvät tulokset **[liite 3]**. Vanhat liitteet 2a-2d yhdistettiin tähän liitteeseen.
 - o Lisäsimme näistä muutoksista huomion **osioon 2.6.** ”Tilastolliset menetelmät” **[252–253]**.
 - o Vanha taulukko 6 ”Esiintyvyys- ja työuupumusmittareiden sijainti- ja hajontaluvut” **[liite 5]**
- Huomasimme virheen kontrollimuuttujien listauksessa **osiossa 2.5.** ”Kontrollimuuttujat”, jossa kohderyhmät oli virheellisesti vielä kirjattu neljään kategoriaan. Kategorioitten 3–4 yhdistäminen

tapahtui suunnitelman mukaisesti osana aineiston anonymisointia, joten päivitimme kirjausta vastaamaan tätä tietoa [244–250].

- Päivitimme kaikki liitetiedostot vastaamaan arviointimuutosten myötä syntyneitä muutoksia liitteiden tiedoissa.

Suosittelijan huomiot

1. ”Yleisesti ottaen kaikki ovat tyytyväisiä, mutta kaipaavat ennen kaikkea pohdinnan syventämistä. Olen samaa mieltä: etenkin tilanteessa, jossa tulokset ovat ristiriitaisia ja viittaavat menetelmällisiin ja/tai käsitteellisiin haasteisiin tutkimuskentällä, olisi tärkeää käydä läpi pohdinnassa erilaisia selityksiä ja tuottaa mahdollisesti uusia hypoteeseja siitä, miten tutkimuskenttä voisi oppia tuloksista ja tehdä tulevaisuudessa asioita toisin.”

Vastaus: Olemme syventäneet ja laajentaneet johtopäätöksiä laajasti (ks. arviokommentit alla).

2. ”On varmasti totta, että epäasiallisen ja väkivaltaisen kohtelun käsitteiden monitulkintaisuus on keskeinen haaste. Olen arvioijien kanssa samaa mieltä, että tätä olisi syytä purkaa sekä menetelmällisesti että teoreettisesti.”

Vastaus: Selvensimme monitulkinnallisuuden mahdollisuutta johtopäätöksissä, ja hyödynsimme suosittelijan ehdottamaa concept creep-käsitettä. Kiitämme suosittelijaa ja arvioijia erittäin tärkeästä huomiosta.

3. ”Menetelmien osalta asiaa voisi lähestyä esim. ilmiväliditeetin (face validity) kannalta haastatteleamalla erilaista kohtelua kokeneita ihmisiä ja tutkimalla miten sisältö vastaa näitä kokemuksia. Vaikka aiemmassa vaiheessa mukana oli ammattilaisten työryhmä, voivat ihmisten arkiset kokemukset ja tulkinnat poiketa. Toinen mahdollisuus olisi yhdistää potentiaalisen mittarin itemeihin avoimia kysymyksiä (”cognitive pretesting”), joiden avulla vastaajat kertovat minkälaisia kokemuksia he yhdistävät kysymyksiin.”

Vastaus: Kiitämme hyvistä kehitysehdotuksista. Lisäsimme maininnan näistä johtopäätösten loppuosuuteen [500–507].

4. ”Teorian kannalta voisi vuorostaan olla tärkeää käydä läpi esim. onko tutkimuskentällä tarve peruskäsitteiden ja niiden suhteiden päivittämiselle. Mieleen tulee etenkin Haslamin ”concept creep” – paljon käytetyillä käsitteillä on taipumus laajentua, jolloin niiden teoreettinen ja käytännöllinen teho heikkenee: Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. Psychological inquiry, 27(1), 1-17. DOI: 10.1080/1047840X.2016.1082418”

Vastaus: Kiitämme huomiosta ja viittauksesta. Olemme lisänneet kommentin asiasta johtopäätöksiin [494–499].

5. ”Etenkin kolmas arvioija esittää useita menetelmällisiä huomioita, jotka on hyvä ottaa vakavasti. Teknisenä huomiona: erilaisten ratkaisujen kokeileminen on tosiaan sallittua ja toivottua, mutta suosittelen hyödyntämään liitteitä niin, että artikkeli pysyy luettavana ja ylimääräisten testien yksityiskohtia siirretään OSF:lle. Arvioijan ehdottama jako suunniteltuihin ja suunnittelemissiin analyysein yleensä selkeyttää rakennetta huomattavasti. Silloin on helppoa pudottaa vähemmän oleelliset suunnittelemat analyysit OSF:lle ja viitata pelkällä linkillä, kun lukija ymmärtää että tieto on ylimääräistä.”

Vastaus: Olemme siirtäneet bifaktorimallin analyysin rekisteröimättömien analyysien osioon, lisänneet sen OSF-liitteenä ja käytännössä poistaneet sen merkityksen johtopäätöksissä (vrt. Arvio

3, kohta 27).

6. "Mainitsen vielä erikseen, että arvioijan esittämä ratkaisu H1 raportoinnista on mielestäni toimiva."
Vastaus: Olemme eri mieltä arvioijan kanssa H1-hypoteesin vastausmahdollisuudesta, mutta olemme ehdottaneet ratkaisua tähän erimielisyyteen (ks. Arvio 3, kohta 42).

Arvio 1

7. "Tutkimuksen aineisto käsitti 960 kyselyyn vastaajaa, joista miehiä oli noin 65 % ja kohderyhmästä noin 60 % oli pelastuslaitoksista ja ensihoidosta. Aineiston koko ja yksipuoleisuus vaikuttaa tuloksiin ja tätä tulee käsitellä pohdinnassa vielä tarkemmin."
Vastaus: Olemme lisänneet huomion tästä pohdintaan [474–480].
8. "Kuviossa 2 esitetään kohtelun ja työuupumuksen muuttujien välisiä korrelaatioita. Kuviossa muuttujien nimet ovat esitetty muodossa k12_5, k3_2 jne. Nämä lyhenteet olisi hyvä lisätä Taulukon 6 ensimmäiseen sarakkeeseen selkeyden vuoksi."
Vastaus: Olemme lisänneet muuttujien nimet uuteen taulukkoon 3. Lisäsimme myös työuupumuksen mittareiden koodit niitä käsittelevään osioon [234–237].
9. "Toinen osuus oli konfirmatorinen ja siinä tarkasteltiin ensimmäisen osatutkimuksen faktorirakenteen validiteettia ja invarianssia, n=680. [...] Ulkoisen validiteetin tarkastelut työuupumuksen mukaan osoittivat vaihtelevia tuloksia. Tätä tulisi tarkastella lisää pohdinnassa."
Vastaus: Olemme lisänneet pohdintaa työuupumuksen vaihtelevasta yhteydestä epäasialliseen ja väkivaltaiseen kohteluun [444–451].
10. "Mittaristo ei ollut täysin yhdenmukainen kaikissa konteksteissa (ryhmissä, suureissa ja malleissa). Tämä saattaa vaikuttaa johtopäätöksiin ryhmien välisistä eroista tai mittauksen luotettavuudesta. Tämä vaihtelevuus tulee huomioida vielä tarkemmin pohdinnassa."
Vastaus: Olemme lisänneet keskustelua mittariston osittaisesta epäinvarianttiudesta johtopäätöksiin [460–473].
11. "Pohdintaosuudessa tuodaan esiin mittariston toimivuuteen vaikuttavia tekijöitä kuten väkivallan määrittelyn ongelman. Pohdintaa mittarin soveltuvuuden osalta tulisi lisätä."
Vastaus: Olemme lisänneet kappaleen mittarin soveltumattomuudesta muille ammattiryhmille [477–480].
12. "Pohdinnan loppuosa on yleisellä tasolla ja siinä tulisi ottaa selkeämmin kantaa kehitetyn mittarin käytettävyyteen."
Vastaus: Olemme lisänneet kappaleen mittarin käytettävyydestä tulevaisuudessa [500–507].
13. "Käsikirjoitus tulee tarkistaa vielä aikamuotojen ja oikeinkirjoituksen osalta."
Vastaus: Olemme käyneet käsikirjoituksen aikamuodot ja kieliopin läpi (ks. käsikirjoituksen arvioinnin ulkopuoliset päivitykset).

Arvio 2

14. "Abstraktissa esitetään selkeä yhteenveto tutkimuksen sisällöstä. Harkitsisin kuitenkin ensimmäisen lauseen muokkausta. Johdannossa viitataan myös tutkimuksiin epäasiallisen ja väkivallan yleisyydestä, mutta abstraktissa viitataan vain uutisoinnin perusteella syntyneeseen mielikuvaan."

Vastaus: Olemme korjanneet abstraktin ensimmäistä lausetta koskemaan myös tutkimustuloksia [5].

15. "Lopputulosten tekstin oikeinkirjoitus ja tekstin sujuvuus olisi hyvä vielä tarkistaa."

Vastaus: Olemme käyneet käsikirjoituksen aikamuodot ja kieliopin läpi (ks. käsikirjoituksen arvioinnin ulkopuoliset päivitykset).

16. "Esimerkiksi lauseessa "Kohtelun tekoympäristö näytti vaikuttavat kohtelun esiintyvyyteen eri tavoin" on kirjoitusvirhe (vaikuttavat pitäisi olla vaikuttavan)."

Vastaus: Kiitämme huomiosta. Koska kyseinen sana ei sopinut artikkelin menetelmiin, vaihdoimme sanan toiseen (ks. Arvio 2, kohta 19).

17. "Lause "Ensimmäiseen, eksploraatiiviseen tutkimuskysymykseen voimme todeta" olisi sujuvampi esim. "Vastauksena ensimmäiseen, eksploraatiiviseen tutkimuskysymykseen voimme todeta"."

Vastaus: Olemme kirjoittaneet lopputulokset uudestaan, jolloin kyseinen lause on myös poistunut [426–507].

18. "Lopputuloksissa oli useita kohtia, joita olisi hyvä vielä selkeyttää. Tärkeimpänä tekstissä voisi pohtia enemmän sitä, että mitä tulokset tarkoittavat mittarin validiteetin kannalta nyt ja jatkotutkimuksissa."

Vastaus: Olemme lisänneet useita kappaleita liittyen mittarin validiteettiin, käytettävyyteen ja jatkokehitykseen johtopäätöksiin (ks. Arvio 1, kohta 10; Arvio 1, kohta 12).

19. "Ensimmäisessä kappaleessa [pohdinnassa] ei käsittääkseni pitäisi käyttää sanaa "vaikuttaa" (asetelma ei mahdollista kausaalisuhteiden arviointia)?"

Vastaus: Osana lopputulosten uudelleenkirjoitusta väitteet kausaalisuhteista on poistettu [426–507].

20. "Tekstissä [pohdinnassa] voisi avata tarkemmin, että mitä tässä tarkoitetaan kohtelun tekoympäristöllä. Tämä selviää tietysti aiemmasta tekstistä, mutta lukijaa voisi muistuttaa."

Vastaus: Olemme lisänneet sivulauseen muistuttamaan siitä, mitä tarkoitamme tekoympäristöllä [429–430].

21. "Tekstissä todetaan, että "Pitkittäistutkimus voisi tukea kausaliteetin selvittämisessä, mutta mahdollisia välillisiä tekijöitä tulisi silloin selvittää tarkemmin." Pitkittäistutkimus ei kuitenkaan itsessään mahdollista kausaliteetin selvittämistä (ja toisaalta poikittaistutkimuksesta voidaan joissain tapauksissa päätellä kausaalisuhteita). Ehkä tässä voisi viitata yleisemmin kausaalipäätelyn mahdollistaviin asetelmiin? Tai vaihtoehtoisesti antaa tarkemman ehdotuksen asetelmasta."

Vastaus: Kiitämme tärkeästä huomiosta. Olemme muokanneet lausetta vastaamaan tarkoitustaan paremmin [486–487].

22. "Sovittamamme mallit eivät sopineet aikaisempaan tutkimukseen pohjautuvan teorian mukaisesti, jolloin käsitteistöä olisi syytä tarkistaa uudelleen." Mitä tämä tarkoittaa? Voisi kirjoittaa auki."

Vastaus: Olemme lisänneet selvennystä tästä lauseesta sen yhteyteen johtopäätöksissä [488–499].

23. "Väkivalta voi tällä hetkellä olla niin monitulkinnallinen käsite mitattavaksi, että vastaajat tulkitsevat käsitettä hyvin vaihtelevasti." Tämä on varmasti totta ja tärkeä huomio, mutta millä perusteella vastaajat tulkitsevat käsitettä vaihtelevasti?"

Vastaus: Olemme lisänneet perustelun tälle ajatukselle käyttäen concept creep -käsitettä [494–499].

Arvio 3

24. "Käsikirjoituksen liitteeksi tulisi laittaa kyselylomake vähintään faktorianalyysiin kuuluneiden kysymysten ja kysymyksenasetteluiden osalta. [...] Viite tällaiseen liitteeseen sopisi esimerkiksi luvun 2.4.1 päätteeksi eli kohtaan, jossa kuvataan mittarin kehitysvaiheita. Toinen ja suosittelemani vaihtoehto on raportoida kysymykset ja kysymyksenasettelut suoraan käsikirjoituksessa, osana EFA:n raportointia."

Vastaus: Kiitämme huomiosta. Saimme suositteijalta luvan lisätä tiedot taulukkona osion 2.4.1. [taulukko 3].

25. "Ehdotan, että esirekisteröinnissä määritellyn ja käsikirjoituksessa raportoidun Hullin menetelmän ja CFI, CAF sekä RMSEA indeksien lisäksi kirjoittajat raportoisivat eksploratiivisesti myös EFA-tutkimuksessa useammin käytettyjen Minimum Average Partial (MAP) sekä erityisesti Parallel Analysis (PA) -testien tulokset, mitä tulee faktoreiden lukumäärän määrittämiseen kaikkien muuttujien mallissa sekä omakohtaisesti koetun kohtelun malleissa. [...] Ehdottaisin itse ainakin parallel analysis -testin käyttöä Hullin menetelmän rinnalla, sillä satunnaisdatasta johdetut vertailut estävät ylisovituksen ja tuovat faktoreiden lukumäärän määrittämiseen tilastollisen kontrollin, mikä tekee tästä testistä eksploratiivisemmän ja vähemmän riippuvaisen teoreettisista ennakkoletuksista."

Vastaus: Kiitämme tärkeästä huomiosta. Olemme lisänneet Parallel Analysis-testin tulokset (käyttäen 95 persentiiliin rajaa) sekä suorittaneet Velicerin Minimum Average Partial -testin [346–349]. Lisäsimme myös huomion osioon 3. suunnitelmasta poikkeamisesta näiden osalta (ks. arvio 3, kohdat 26 ja 38). Näiden pohjalta päädyimme myös asettamaan uudet eksploratiiviset faktorimallit [ks. 350–353]: kaikkien muuttujien malleja ei voitu analysoida, sillä determinantti ja ongelmalliset muuttujat eivät puoltaneet niiden analyysia, ja faktorimääräarviot olivat suuresti ristiriitaisia. Omakohtaisesti koettujen muuttujien mallien osalta datasopivuus puolsi analyysia, mutta faktorimääräarviot puolsivat joko kahta tai kolmea faktoria. Olemme muuttaneet eksploratiivista analyysiosuutta vastaamaan näitä huomioita [346–380].

Näiden lisäksi eksploratiivinen osuus tunnisti yhden lisäselvitettävän mallin (11 muuttujaa, 2 faktoria, pohjautuen ensimmäiseen uuteen EFA-malliin) (ks. arvio 3, kohta 31), jota ei kuitenkaan valittu jatkomalliksi. Muutimme siksi eksploratiivisten analyysien vertailutaulukon (taulukko 7) vastaamaan kolmen mallin vertailua: omatoimisesti koettu kohtelu, kaksi faktoria (Malli 1); omatoimisesti koettu kohtelu, kolme faktoria (Malli 2); sekä omatoimisesti koettu kohtelu, vain 11 muuttujaa, kaksi faktoria (Malli 3).

26. "Koskien edellä kuvattua ehdotusta täydentävien testien raportoinnista, ja myös koskien konfirmatorisen faktorianalyysin lukua, kirjoittajien tulisi tekstissään selvästi erotella, mitkä testit ja analyysit ovat suoraan esirekisteröinnin suunnitelmaan perustuvia ja mitkä ovat laajennuksia ja täydennyksiä esirekisteröidyssä suunnitelmassa kuvattuun malliin nähden. [...] Tämä palaute koskee erityisesti konfirmatorisen faktorianalyysin lukua sekä toisaalta myös Lopputulokset-lukua."

Vastaus: Olemme erottelleet rekisteröidyt ja rekisteröimättömät suunnitelmat selkeästi käyttäen uutta osiota 3 [301–317]. Lopputuloksiin liittyen, ks. Arvio 3, kohta 27.

27. "Lopputulokset-luku tulisi siten [viitaten Arvio 3, kohta 26] jäsentää myös tavalla, joka selvästi kuvaa lukijalle, 1) mitkä analyysit perustuivat esirekisteröityyn suunnitelmaan, ja mitkä ovat näihin tuloksiin liittyvät päätelmät sekä 2) mitkä analyysit ovat laajennuksia suunnitelmasta, ja minkälaisia

päätelmiä näihin liittyen voidaan tehdä.”

Vastaus: Koska päätimme vähentää selkeästi rekisteröimättömän, eksploratiivisen bifaktorimallin osuutta tekstissä (vrt. Arvio 3, kohta 26), olemme poistaneet lopputuloksista keskustelun liittyen bifaktorimalliin. Pysymme rekisteröityjen analyysien keskustelussa, sillä ne ovat luotettavampia ja käsikirjoituksen tarkoituksen mukaisia.

28. ”Käsikirjoituksesta puuttuu EFA-tutkimuksen motivointi ja rationale. Tarkoitin, että käsikirjoituksessa olisi syytä kuvata ne syyt, miksi analyysi edellyttää eksploratiivista analyysia, ja miten valitut indeksit/testit tukevat juuri analyysin eksploratiivisuutta sekä näiden analyysien tavoitetta. Implisiittisesti esirekisteröidystä suunnitelmasta tämä käy ilmi, mutta kaipa analyysin tulokset esittävän EFA-luvun alkuun menetelmän käyttöä tukevaa tiivistelmää, joka selkeästi perustelee myös valitut menetelmät, mitä tulee faktoreiden lukumäärän tunnistamiseen.”
- Vastaus:** Lisäsimme tiivistelmän ja motivoinnin EFA-menetelmän valinnasta ensimmäisen osatutkimuksen tulososioon [337–340].
29. ”Taulukon 8 luentaa vaikeuttaa se, että näiden indeksien ja testien viitearvoja EFA-kontekstissa ei ole kerrottu; kirjoittajat toteavat Mallin 2B sopivuusarvojen olevan muita malleja paremmat, mutta käsikirjoituksessa tulisi raportoida myös, miten 2B-mallin ja muiden mallien arvot vertautuvat kirjallisuudessa määriteltyihin tyydyttäviin/hyviin indeksi- ja testiarvoihin EFAn viitekehelyssä.”
- Vastaus:** Lisäsimme yhden kappaleen sopivuusindeksien viitearvoista ja mallien paremmuudesta taulukon jälkeen [361–364].

30. ”Jos kirjoittajat raportoivat eksploratiivisen faktorianalyysin mukaiset standardoidut lataukset Kuvion 3 mukaisesti, heidän tulisi myös keskustella tässä yhteydessä siitä, minkälaisen lähtökohdan instrumentti/item-kohtaiset lataukset antavat mittarin validoinnille ja konfirmatoriselle faktorianalyysille. Pelkästään latausten perusteella Kuviossa 3 kuvatun kahden faktorin mallin instrumenteista olisi perusteet poistaa molemmat ristiinlatautuneet muuttujat sekä ”Fyysinen väkivalta, sisäinen”, ”Omaisuus-väkivalta, sisäinen” ja mahdollisesti myös ”Omaisuus-väkivalta, ulkoinen” -muuttujat ennen konfirmatoriseen vaiheeseen siirtymistä – jos tähän on olemassa teorialähtöiset perusteet.”
- Vastaus:** Olemme korvanneet mainitun kuvion taulukolla, josta esiintyy jokaisen muuttujan faktorikohtaiset, rotatoidut lataukset sekä kommunaliteettiarvot [taulukko 7] (ks. myös muutokset analyysissa, Arvio 3, kohta 31).

31. ”En myöskään täysin ymmärrä, miten tähän malliin päädyttiin, sillä mallissa raportoidaan kolme muuttujaa, joiden faktorilataukset ovat alle 0.4, vaikka kirjoittajat ovat määritelleet: ”Sisällytämme muuttujan faktoriin, jos sen faktorilataus on yli .40”.”
- Vastaus:** Kiitämme erittäin tärkeästä huomiosta. Suorittaessamme analyysimme, teimme virheen ja jätimme noudattamatta ennalta rekisteröityä sisäänottokriteeristöä (lataus yli ,40 ja erotus suurimman ja toiseksi suurimman latauksen välillä yli ,30). Olemme korjanneet tämän virheen käsikirjoituksessa alkaen osiosta 4.2. Virheen korjaus on muuttanut useita käsikirjoituksen osioita merkittävästi:
1. Oikealla, ennalta rekisteröidyllä analyysilla päädyimme testaamaan kolmea eksploratiivista mallia (Malli 1: 14 muuttujaa, 2 faktoria; Malli 2: 14 muuttujaa, 3 faktoria; Malli 3: 11 muuttujaa, 2 faktoria), joiden pohjalta valitsimme Malli 1:een pohjautuvan 11 muuttujan kaksiulotteisen mallin CFA-analyysiin.
 2. Mallin muutos mahdollisti dynaamisten raja-arvojen käytön, joten vaihdoimme käyttämään näitä raja-arvoja. Teimme sen myötä muutoksen ensimmäisessä arvioissa arvioituun käsikirjoituksen

osioon 2.6.2 [292–295].

3. Mallin muutos vaikutti lievästi modifikaatioindeksien tarkasteluun. Päädyimme kuitenkin vielä epäilemään bifaktorirakennetta, mutta kevensimme bifaktorirakenteen osuutta käsikirjoituksessa (sillä tarkastelua ei ollut rekisteröity) ja siirsimme tarkastelun osaksi rekisteröimättömien analyysien osiota (ks. Arvio 3, kohta 26). Kevennyksen myötä luovuimme bifaktorimallin invarianssitestauksesta.

4. Korjasimme taulukot 6, 7 ja 8 vastaamaan uuden mallin tuloksia.

5. Korjasimme kuvion 3 arvot uuden mallin tuloksilla. Lisäsimme samalla myös virhevarianssien merkitsevyydet kuvioon.

32. ”Faktorilatausten osalta kehotan myös yhtenäistämään kirjoitusasua, kirjoittajat käyttävät välillä .XX-mallia ja välillä 0,xx-mallia, sama koskee myös indeksien raja-arvojen raportointia.”

Vastaus: Kiitämme huomiosta. Olemme vaihtaneet kaikki desimaaliluvut käyttämään suomalaista muotoa pilkulla edeltävällä nollalla (0,XX-muoto). Muutokset on tehty kautta käsikirjoituksen ja merkitty punaisella värillä. Teimme myös vastaavat muutokset kuvioihin.

33. ”En itse pidä myöskään EFAn tulosten esittämistä Kuvion 3 -muotoisena parhaana ratkaisuna. [...] Pyydän siis kirjoittajia raportoimaan täyden faktorilataustaulukon 2B-mallille, jonka kanssa he etenevät CFA-vaiheeseen.”

Vastaus: Olemme vaihtaneet Kuvion 3 taulukoksi (taulukko 7), josta esiintyy standardisoidut faktorilataukset ja instrumenttien kommunaliteetit (ks. arvio 2, kohta 34).

34. ”Täyden faktorilataustaulukon [viitaten Arvio 2, kohta 34] tulisi sisältää myös uniqueness- tai communality-arvo jokaiselle instrumentille/itemille.”

Vastaus: Lisäsimme kommunaliteettiarvot uuteen faktorilataustaulukkoon (ks. arvio 2, kohta 33) [382–383].

35. ”CFAn osalta (Kuvio 4) kirjoittajien tulisi raportoida instrumentti/item-tasolla faktorilatausten lisäksi myös residuaali/virhevarianssi sekä sisällyttää näihin liittyvä keskustelu osaksi analyysia.”

Vastaus: Olemme lisänneet täysin standardisoidut virhevarianssit kuvioon 3. Olemme myös lisänneet vertailun virheistä ja faktorilatauksista [397–400] sekä keskustelleet asiasta enemmän johtopäätöksissä [452–459].

36. ”Toiseksi yleisten sopivuusmittojen χ^2 , RMSEA, CFI, SRMR lisäksi faktoreille tulisi laskea CR (composite reliability / komposiittireliabiliteetti) sekä AVE (Average Variance Explained / keskimääräinen selitetty varianssi) -arvot, ja suhteuttaa jälkimmäinen faktoreiden välisen korrelaation neliöön (Fornellin ja Larckerin kriteeri).”

Vastaus: Lisäsimme CR- ja AVE-arvot sekä tarkastimme Fornellin ja Larckerin kriteerin täyttymisen [401–404].

37. ”Lisäksi mallissa (Kuvio 4) on useita heikosti latautuneita instrumentteja/itemeitä.”

Vastaus: Osana bifaktorimallin merkityksen vähentämistä käsikirjoituksessa sen rekisteröimättömyyden takia (ks. Arvio 3, kohta 26 ja 31) olemme poistaneet tämän kuvion.

38. ”CFA-raportoinnin osalta luku olisi syytä selkeästi jakaa edellä mainitsemallani tavalla eli esirekisteröinnissä esitetyn suunnitelman mukaisesti ja sitä täydentäviin eksploratiivisiin kokeiluihin, joihin tulkintani mukaan osittain siirrytään käsikirjoituksen rivillä 423, vaihtoehtoisten mallien vertailun takia. Tällaisen jaon voisi tehdä hyvin jo otsikkotasolla, tämä auttaisi lukijaa.”

Vastaus: Erottelimme rekisteröimättömät analyysit omaksi osioksi (ks. Arvio 3, kohta 26).

Poistimme myös lopputuloksista rekisteröimättömien analyysien päätelmät (ks. Arvio 3, kohta 27).

39. "Vaihtoehtoisten mallien luenta CFA:n yhteydessä on periaatteellisella tasolla myös hieman horjuvaa, koska CFAn tulisi teorialähtöisenä mallina olla riippumaton eksploratiivisesta otteesta ja perustua empiria- ja teorialähtöisen yhden mallin testaamisen. Toki jos kahteen kilpailevaan malliin on teoria- ja empirialähtöiset perusteet, tämä tulisi selvästi ilmaista ja perustella jo ennen CFA-analyysiin ryhtymistä."

Vastaus: Olemme siirtäneet bifaktorimallin arvioinnin pienempään rooliin ja selkeästi merkinneet, että kyseessä on rekisteröimätön analyysi. Haluamme kuitenkin huomauttaa, että myös CFA-mallinnukseen voi, nimestään huolimatta, kuulua vaihtoehtoisten mallien arviointia sekä eksploratiivista analyysia (esim. Kline, 2023, ss. 33, 37, 231).

40. "Lisäksi CFA-mallin hyväksymiskriteeristämme: "Hyväksyimme mallin, jos vähintään kaksi sopivuusindeksiä neljästä olivat raja-arvojen sisällä" tulee todeta, että julkaisuilla on käytössä mallien hyväksyntäkriteerien suhteen erilaisia malleja. Toisissa julkaisuissa mallit ovat hyväksyttäviä vain, jos kaikki sopivuusindeksit pysyvät raja-arvojen sisällä. Tähän liittyen olisi hyvä lisätä lähteitä, joissa todetaan 2/4 täyttyvien kriteerien oleva riittävää mallin hyväksynnän kannalta."

Vastaus: Kiitämme tärkeästä huomiosta. Osana vertaisarviointiin valmistautumista olemme selvittäneet laajemmin raja-arvojen käyttöä. Selvityksessä meille kävi muun muassa ilmi, ettei normalisoitua khii neliö –suhdetta suositella käytettävän ollenkaan (Kline, 2023, s. 161), joten olemme jättäneet tämän pois arvioinnista kokonaan. Samaan aikaan saimme mallimuutoksen myötä (ks. Arvio 3, kohta 31) dynaamiset raja-arvot toteutettua, joten olemme vaihtaneet raja-arvot näihin. Päätimme samalla poistaa valikoivan raja-arvosääntöämme (jossa 3/4 arvoista tulisi täyttyä), sillä juuri tälle rajalle ei ole tieteellistä perustaa (osio 2.6.2.). Raja-arvojen binaarikäyttö on kuitenkin haasteellista ja osittain ongelmallista. On totta, että useat julkaistut mallit vaativat hyväksynnälle kaikkien sopivuusindeksien pysymisen raja-arvojen sisällä. Tämä ei kuitenkaan välttämättä ole kestävä menetelmä, sillä globaalit sopivuusindeksit 1) mittaavat osittain eri asioita (esim. SRMR mittaa korrelaatiomatriisin residuaalin keskipitkyyttä, ja RMSEA mittaa khii neliö –suureen suhteellista laskua mallissa verrattuna keskiarvomalliin), 2) ovat herkkiä eri virhespesifikaatioille (esim. otoskoko, mallin virhespesifikaation laajuus, vapausasteet, tai virheiden hajautuminen), ja 3) mittaavat vain kokonaissopivuutta, eivät muuttujien paikallista sopivuutta (Kline, 2023, ss. 165–170). Suosituksena onkin keskittyä enemmän paikalliseen sopivuuteen, eli muuttujien virheisiin, kuin globaaliin sopivuuteen. Olemme lisänneet pyynnöstä virheiden analyysin sekä tulokset- että johtopäätökset-osioon (ks. Arvio 3, kohta 35), jolloin käsittelemme mahdollista virhesopivuutta myös paikallisella tasolla.

41. "Riviltä 423 alkaen esitelty Malli 2, joka kuvataan graafisesti rivillä 446, on mielestäni hämmentävä. Jos tässä eksploratiivisessa mallissa kaikki itemit katsotaan latautuviksi Yleinen-faktorille että Organisaatio-sisäinen/Organisaatio-ulkoinen-faktoreille, eikö tällöin ole kyseessä II-tason faktorianalyysi? [...] Kuitenkin huomautan, että Mallissa 2 on varsin matalia latauksia, jotka eivät tue tätä eksploratiivista mallia."

Vastaus: Kuten aikaisemmin mainittu, olemme vähentäneet bifaktorimallin merkitystä käsikirjoituksessa ja selkeästi osoittaneet sen rekisteröimättömäksi analyysiksi (ks. Arvio 3, kohta 39). Malli on kuitenkin bifaktorimalli, sillä bifaktorimallissa on kaksi rinnakkaista faktoritasoa: spesifit tai sisältöfaktorit (joiden itemit eivät risteä, ja joiden faktoreiden välillä ei esiinny korrelaatiota), sekä yksi yleinen faktori (jolle kaikki itemit lataavat). Tuolloin kaikki faktorit ovat eksogeenisiä, ja kaikki itemit saavat kaksi endogeenistä "syytä": yksi sisältöfaktori ja yksi

yleisfaktori. Toisen tason faktorianalyysissä sisältöfaktorit lataavat ylemmän tason faktorille, eli sisältöfaktorit ovat endogeenisiä ja ylemmän tason faktori eksogeeninen.

42. "Tutkimustulokset-luvussa raportoitu H1-testaus, joka perustuu riveillä 434-442 korrelaatioihin kehitteillä olevan mittarin ja työuupumuksen välillä on sikäli ongelmallinen, että tällainen rakenneyhtälömalli on mahdollista toteuttaa vain validoiduilla instrumenteilla, eikä kehitteillä olevan mittarin validaatiota voi pitää tutkimuksen perusteella pitävänä. [...] Minusta tässä yhteydessä pitäisi vähintään todeta tämä seikka ja samalla todeta, että suunnitelman mukaista H1-testausta ei tästä syystä voida toteuttaa. [...] Yksi vaihtoehto olisi keskustella sen sijaan yleisellä tasolla rivien 348-349 raportoidun korrelaatiotaulun tuloksista pohtien yksittäisten muuttujien korrelaatioita työuupumuksen vastaavien muuttujien kanssa, ja todeta, ettei H1 tarkastelu ollut suunnitellulla tavalla mahdollista. Tietystä mielessä tämä huomio koskee myös muita konfirmatoriselle faktorianalyysille asetettuja hypoteeseja. Yksittäisten itemien kohdalla invarianssitestaukset toki ovat silti mahdollisia."

Vastaus: Kiitämme arvioijaa huomiosta. Emme kuitenkaan yhdy tulkintaan, että tässä kontekstissa "kehitteillä olevan mittarin validaatiota [ei] voi pitää tutkimuksen perusteella pitävänä". Kriteeri- tai konseptuaalisesti liittyvien rakenteiden yhteyden käyttö validiteettitarkastelussa on vakiintunut menetelmä, jonka takia olimme valinneet tämän menettelyn työuupumuksen osalta. Kuten toteamme käsikirjoituksessa **osiossa 1.2.**, työuupumuksen yhteys epäasialliseen ja väkivaltaiseen kohteluun on näytetty aikaisemmassa tutkimuksessa, mikä motivoi työuupumuksen käyttöä validaatioasetelmassa. Vaikutusvaltainen *The Standards for Educational and Psychological Testing* -standardi ei myöskään tunnista, etteikö kyseistä validaatiotarkastelua voisi tehdä samanaikaisesti kerätyllä datalla, riippumatta siitä, onko analysoitava testi aikaisemmin validoitu vai ei (ks. standardit 1.16-1.24, American Educational Research Association ym., 2014, ss. 27–30). Toki työuupumuksen konstruktin yhteys epäasiallisen ja väkivaltaisen kohtelun konstrukteihin ei voi olla ainoa validiteetin todiste, mistä syystä keskustelemme johtopäätöksistä myös sisäisen rakenteen tulkinnasta sekä teoreettisista todisteista ja mahdollisista selityksistä faktorituloksille. Olemme joka tapauksessa lisänneet selventävän huomion johtopäätöksiin, ettei yksittäinen SEM-malli anna täydellisiä todisteita kyselytyökalumme validiteetista **[441–445]**.

43. "Tutkimustulokset-luku on hyvä, mutta mahdollisen jatkotutkimuksen osalta hieman suppea. Jos tutkijat toteuttaisivat ja raportoisivat täyden EFA-aulun, he voisivat tässä luvussa keskustella myös eri muuttujien latauksista sekä ristiinlatauksista tarkemmin, sekä pohtia minkälaisia implikaatioita tällä voisi olla mittarin jatkokehitykselle tai keskeisten käsitteiden kuten väkivallan tulkinnoilla."

Vastaus: Olemme lisänneet osioita, joissa käsittelemme valitun EFA-mallin ristilatauksia sekä latausten myötä tehtyjä merkitystulkintoja **[430–436]**.

Viitteet

American Educational Research Association, American Psychological Association, & National Council on

Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Kline, R. B. (2023). *Principles and Practice of Structural Equation Modeling* (5. p.). Guilford Press.