

Oliver Saal ja työryhmä:

Kiitos huolellisista muutoksista käsikirjoitukseen. Kaksi arvioijaa palasi jakamaan näkemyksensä toiselle kierrokselle, ja toinen heistä on tyytyväinen nykyiseen versioon.

Toisella arvioijalla oli vielä kriittinen huomio koskien mittarin ulottuvuuksia, ja tämä huomio tulisi harkita lisämuutoksena ennen tutkimuksen aloittamista. Koska asia on mittarinkehityksen kannalta keskeinen, konsultoin vielä erikseen tilastotieteen ulkoista ammattilaista, joka oli arvioijan kanssa samaa mieltä siitä, että yksi kysymys ei ole riittävä itsenäisen ulottuvuuden luotettavaan validointiin tässä kontekstissa. Kannustan työryhmäänni vielä harkitsemaan arvioijan esittämää kahta ratkaisuvaihtoehtoa ja tarvittaessa ottamaan minuun yhteyttä toimivan ratkaisun löytämiseksi.

Lisäksi pohdin optimaalista otosta EFA:lle (nyt 20 %). Vaikka tarkkaa määrää on vaikea arvioida, monet tutkimukset kallistuvat 50/50 suuntaan koska hyvä EFA parantaa myös CFA:n onnistumisen mahdollisuuksia. EFA:n nostaminen 40%- tai edes 30%-luokkaan voisi vahvistaa suunnitelmaa yhä, mikäli koette sen mahdolliseksi.

Kyriazos, T. A. (2018). Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, 9(08), 2207.

Vaikka emme saaneet ensimmäiseltä arvioijalta uusia kommentteja, haluaisin kuitenkin vielä palata hänen aiemmin esittämiin huomioihin ja niitä koskeviin ratkaisuihin. Olette päättäneet pitäytyä epäasiallisen kohtelun käsitteen laajassa tulkinnassa, ja sitä on perusteltu yhä kattavammin. Viitatus kirjallisuuden valossa tuntuu kuitenkin yhä selvemmin siltä, että "väkivallan" käsite ja sen tutkimus erottuu vahvasti omana sektorinaan. Olisiko mahdollista ratkaista tämä kahtiajako niin, että suunnitelma koskisi "Epäasiallisen ja väkivaltaisen kohtelun mittarin" validaatiota? Tällöin moniulotteisuus olisi läpinäkyvä eikä mittari olisi niin altis käsiterakenteelliselle kritiikille.

Tiedän, että kyselyn pituuden kontrolloimiseksi on tärkeää, ettei se ole liian pitkä. Pohdin kuitenkin, että validaatio voisi hyötyä toisen ulkoisen muuttujan testaamisesta työuupumuksen rinnalla. Esim. "työn tarmokkuutta" (3 kysymystä) ([Hakanen 2009](#)) voisi mitata kokeilevasti ilman, että korrelaatiota asetetaan mittarin toimivuuden edellytykseksi (koska siitä ei ole aiempaa evidenssiä). Teorian kannalta olisi kuitenkin loogista, että kohtelu vaikuttaa myös tarmokkuuteen. Jos epäasiallisen ja väkivaltaisen kohtelun mittari korreloi sekä uupumuksen että tarmokkuuden kanssa, tieto voisi auttaa ymmärtämään uuden mittarin toimintaa yhä paremmin. Ehdotuksen voi sivuuttaa, jos se koetaan liian raskaaksi toteuttaa.

Yhteenvetona: suunnitelman on parantunut huomattavasti ja lähestyy valmista. Odotan mielenkiinnolla seuraavan (ehkä lopullisen) version tarkastamista.

Veli-Matti Karhulahti

Arvio 1 [Ei päivämäärää]

[Ei uutta arviota]

Arvio 2: 16.5.2024

Kirjoittajat ovat ottaneet arviointipalautteet kattavasti huomioon ja muokanneet suunnitelmaa merkittävillä tavoilla. Nykyinen suunnitelma vastaa esittämäni kysymyksiin ja kritiikkeihin hyvin, ja voin näiltä osin puoltaa tutkimussuunnitelman hyväksymistä. Haluan kuitenkin nostaa esille edelleen yhden kriittisen kysymyksen, joihin nähdäkseni kirjoittajien täytyy vielä vastata.

Kysymyksen liittyä epäasiallisen kohtelun mittarin dimensionaalisuuden tarkasteluun sekä siihen, minkälainen instrumentti kuvatulla tutkimuksella voidaan periaatteessakaan validoida. Kirjoittajat (rivit 198-214) kuvaavat mittaavansa epäasiallista kohtelua neliosaisella mittarilla. Kutakin kuvatuista osista (omakohtaiset kokemukset organisaation sisällä, havaitut kokemukset organisaation sisällä, omakohtaiset kokemukset organisaation ulkopuolella, havaitut kokemukset organisaation ulkopuolella) mitataan 7 muuttujan kysymyssarjalla. Suunnitelman mukaan nämä seitsemän muuttujaa muotoillaan siten, että kukin muuttuja mittaa yhtä oletetuista epäasiallisen kohtelun ulottuvuuksista (Taulukko 1).

Näin kuvattu 7 muuttujan instrumentti, joka toistetaan tai toisinnetaan (versioidaan) erikseen kuhunkin yllä mainittuun neljään osioon ei voi perustellusti mitata epäasiallista kohtelua kuin yksi- tai kaksifaktorisena konstruktiona. Näin on siksi, että yhden faktorin tai ulottuvuuden mittaamiseen vaaditaan vähintään kolme muuttujaa/kysymyssarja eli kyselyosio.

Toisin sanoen suunnitelmassa esitelty lähestymistapa ei voi validoida kuvatulla mittarinkehitysprosessilla seitsenulotteista tai mitään yli kolmiulotteista epäasiallisen kohtelun instrumenttia. Jotta kirjoittajat voisivat perustellusti tutkia, koostuuko epäasiallisen kohtelun kyselyinstrumentti teorian mukaisista seitsemästä ulottuvuudesta (Taulukko 1) kunkin neljän osion tulisi sisältää vähintään 21 (3 itemiä per 7 oletettua ulottuvuutta) muuttujaa (eli kyselyssä olisi yhteensä vähintään $21 \times 4 = 84$ muuttujaa).

Näen vähintään kaksi erilaista linjaa vastata yllä olevaan kriittiseen huomioon. Vaihtoehto 1: Koska epäasiallisen kohtelun määritelmästä ei ole olemassa konsensusta (rivi 26), tulisi kirjoittajien ensin tutkia, voidaanko 7-ulotteinen epäasiallisen kohtelun konstruktio validoida. Tämän asian selvittämiseksi tulisi kerätä erillinen aineisto, jossa jokaista oletettua 7 ulottuvuutta mittaamaan kehitettäisiin vähintään 3 kysymystä eli kehitettäisiin minimissään 21 kysymyksen mittari. Tälle tulisi tehdä eksploratiivinen faktorianalyysi, jolla selvitettäisiin, koostuuko näin muodostettu mittari todella seitsemästä ulottuvuudesta tai jonkinlaisesta muusta dimensionaalisuudesta. Tässä vaihtoehdossa etuna olisi, että epäasiallisen kohtelun dimensionaalisuus saisi empiiristä validaatiota. Toinen etu olisi, että tällaisen analyysin myötä kysymysten muotoilua voitaisiin tarkentaa sekä mittaria lyhentää aineistolähtöisesti esimerkiksi seitsemän kysymyksen mittaiseksi (tai $7 \times 4 = 28$ kysymyksen kokonaisuudeksi kuten

nykyisessä suunnitelmassa kuvataan). Tämä vaihtoehto olisi riskittävämpi ja tuottaisi todennäköisemmin parhaimman lopputuloksen.

Vaihtoehto 2 on edetä nyt esitellyn suunnitelman mukaisesti ja kehittää suoraan seitsemän kysymyksen mittariversio sekä toteuttaa suunnitelmassa esitellyt työvaiheet. Tässä muodossa tutkimus ei tuottaisi kuitenkaan lisätietoa epäasiallisen kohtelun dimensionaalisuudesta yli 1- tai 2-faktorisen rakenteen, ja siten tutkimus ei voisi validoida teoriaosiossa esitettyä 7-dimension rakennetta tai mitään muutakaan moniulotteista faktorirakennetta. Lisäksi näen, että suoraan 7 kysymyksen mittarin kehittämiseen liittyy merkittävä riski: teorian mukaiset dimensiot (esim. uhkailu, syrjintä, vihamielisyys...) voivat hyvinkin mitata paitsi epäasiallista kohtelua myös muita latenteja rakenteita, ja tämä seikka voi hankaloittaa 7 kysymyksen mittarin validointia. Lisäksi tutkijoiden esittämä malli, jossa kutakin teorian mukaista ulottuvuutta mittaa vain yksi kysymys tekee instrumentista haavoittuvaisen ja mittarinkehityksestä erittäin vaikeaa: tutkijat voivat joutua tiputtamaan instrumentista yhtä tai useampaa hypoteettista dimensiota vastaavan/vastaavat kysymykset mitä pienimmänkin kysymysmuotoiluun päätyvän merkitysvivahteen tai sanavalinnan takia. Jos tutkijat päättävät edetä vaihtoehdon 2 mukaisesti, tulisi heidän perustella tämä valinta suunnitelmassa ja argumentoida, miksi yksi kysymys tässä tapauksessa luotettavasti mittaa esimerkiksi fyysistä väkivaltaa, uhkailua, vihamielisyyttä ja muita neljää oletettua epäasiallisen kohtelun ilmentymismuotoa.

Vaihtoehdossa 2 tutkijat joutuvat ottamaan suuremman riskin sen suhteen, että he joutuvat viime kädessä pudottamaan validoitavasta mittarista yhden tai useamman kysymyksen (ja siten yhteyden oletettuun epäasiallisen kohtelun dimensioon, esimerkiksi vaikkapa uhkailuun tai syrjintään), jotta konfirmatorinen faktorianalyysi mahdollistuu ja malli voidaan validoida. Toisena puutteena tässä vaihtoehdossa on, että näin toteutettu prosessi ei tuota kattavaa lisätietoa epäasiallisen kohtelun latenteista dimensioista – asia, joka voisi toteutuessaan olla askel konsensuksen löytymiseen ilmiön määrittelyssä (vrt. rivi 26 käsikirjoituksessa).

Vaihtoehdossa 1 hankaluutena on, että tämä käytännössä edellyttää erillistä aineistonkeruuta sekä instrumentin kehittämistä laajempaan muotoon, mikä toisaalta on eksploratiivisen faktorianalyysin kannalta ensiarvoisen tärkeä vaihe. Vaihtoehto 1 siis lisäisi tutkimukseen työvaiheita, mutta toisaalta arvioin myös, että näin toteutetun tutkimuksen impakti tieteessä ja yhteiskunnassa olisi vaihtoehtoa 2 merkittävämpi.

Muilta osin suunnitelma on mielestäni edistynyt merkittävästi ja voin puoltaa suunnitelmassa esitetyn tutkimuksen toteutusta. Koen kuitenkin, että tutkijoiden tulee vastata yllä esittämäni kritiikkiin ja perustelemalla valintansa suunnitelmassa.

- Jukka Vahlo

Arvio 3: 19.5.2024

[Arvioija on tyytyväinen muutoksiin, ei uusia kommentteja.]