



Validiteetti ja reliabiliteetti

Leena Hiltunen
Graduryhmä 18.2.2009

JYVÄSKYLÄN YLIOPISTO

Määrittelyä

- Validiteetti ja reliabiliteetti määritellään usein kahdesta eri näkökulmasta
 - Onko mittaus- tai tutkimusmenetelmä validi ja/tai reliabeeli?
 - Ovatko tuloksista johdetut päätelmät valideja ja/tai reliabeleja?

Validiteetti tutkimusmenetelmän näkökulmasta

- Validiteetti (pätevyys, engl. validity) ilmaisee sen, miten hyvin tutkimuksessa käytetty mitta- tai tutkimusmenetelmä mittaa juuri sitä tutkittavan ilmiön ominaisuutta, mitä on tarkoituskin mitata;
 - mitaako tutkimus sitä, mitä sen avulla on tarkoitus selvittää¹

¹ Hirsjärvi ym. 2002, 213

JYVÄSKYLÄN YLIOPISTO

Validiteetti on hyvä silloin, kun tutkimuksen kohderyhmä ja kysymykset ovat oikeat. Validiteetin arviointi kohdistaa yleensä huomionsa kysymykseen, kuinka hyvin tutkimusote ja siinä käytetyt menetelmät vastaavat sitä ilmiötä, jota halutaan tutkia. Voidakseen olla validi, sovellettavan tutkimusotteen tulee tehdä oikeutta tutkittavan ilmiön olemukselle ja kysymyksenasettelulle. Validiteetin kannalta ei kuitenkaan ole keskeisintä ryhtyä pohtimaan sitä, millä - kuinka valideilla - mittareilla tuloksia saadaan, vaan aivan ensimmäiseksi miettiä sitä, **millainen tutkimuksen strategia on validi**.

Tutkimustyössä käytetty menetelmä ei itsessään johda tietoon, vaan menetelmä on valittava sen mukaan, millaista tietoa halutaan. Tätä seikkaa joutuu tutkija pohtimaan aivan ensimmäisenä tehdessään valintaa eri tutkimusmenetelmien välillä. Varsinkin teoreettisten käsitteiden validiteetti on viime kädessä arviointikysymys, jossa tutkijayhteisössä pitkäaikaisen keskustelun tuloksena muodostunut vallitseva mielipide on ratkaiseva: tutkijayhteisö on "sopinut", että tiettyä käsitettä voidaan mitata tietyllä mittarilla. Jos mittaustulokset osoittavat, että saatu tieto vastaa vallalla olevaa teoriaa tai pystyy sitä tarkentamaan ja parantamaan, silloin tulos on validi. Käytännössä ongelma on siinä, että mittausten ulkopuolelle jäävä, todellista tietoa ei yleensä ole käytettävissä.

Validiteetin totaalinen puuttuminen tekee tutkimuksesta arvottoman. Tällöin tutkitaan todellisuudessa aivan muuta asiaa kuin mitä kuvitellaan tai alun perin on ollut tarkoitus. Puutteellinen validiteetti merkitsee sitä, että empiiriset havainnot ja koko tutkimus itsessään kohdistuu enemmän tai vähemmän sivuun siitä, mikä oli tarkoituksena.

Esimerkkejä tutkimusmenetelmän validiteetista

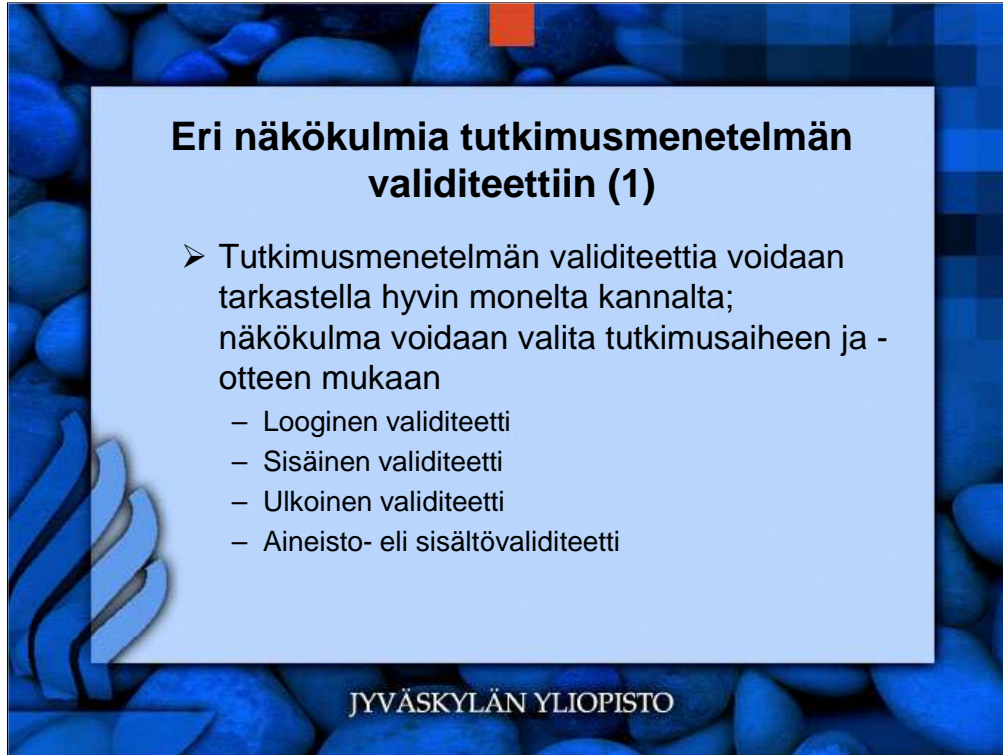
- Onko grafologia validi menetelmä persoonallisuuden analysoinnissa?
- Kertooko rajakokemus todella kuolemanjälkeisestä elämästä?

JYVÄSKYLÄN YLIOPISTO

Onko joku tietty mittari tai menetelmä hyvä juuri sen halutun ilmiön mittaamiseen tai tutkimiseen; mittaako se juuri haluttua ilmiötä vai jotain muuta?

Validiteettia tulee siis miettiä jo tutkimusmenetelmän valinnan yhteydessä.

Grafologia eli käsialatutkimus.



Eri näkökulmia tutkimusmenetelmän validiteettiin (1)

- Tutkimusmenetelmän validiteettia voidaan tarkastella hyvin monelta kannalta; näkökulma voidaan valita tutkimusaiheen ja -otteen mukaan
 - Looginen validiteetti
 - Sisäinen validiteetti
 - Ulkoinen validiteetti
 - Aineisto- eli sisältövaliditeetti

JYVÄSKYLÄN YLIOPISTO

1) **Looginen validiteetti** (koettu validiteetti, face validity) edellyttää, että tutkimusta tarkastellaan kokonaisuudessaan kriittisesti. Se tarkoittaa tutkijan omaa käsitystä suorituksen oikeellisuudesta. Näyttääkö tulos oikealta?

2) **Sisäisen validiteetin** tarkastelussa tehdään kysymys: Aiheutuvatko empiirisen tutkimuksen koetilanteessa saadut tulokset / muuttujien väliset erot niistä tekijöistä, joiden oletetaan niihin vaikuttavan? Laadullisen tutkimuksen ollessa kysymyksessä tarkastellaan mm. kysymystä siitä, onko tutkimusstrategia valittu tutkittavan kohteen olemuksen mukaisesti? Sisäisen validiteetin heikkouteen varsinkin kvantitatiivisissa tutkimuksissa saattaa vaikuttaa useita tekijöitä, joiden kontrollointiin on syytä kiinnittää huomiota:

•Aika: Eri mittauskertojen välillä on saattanut tapahtua erilaisia asioita, jotka vaikuttavat mittauskohteeseen (erilaisia tapahtumia, henkilöiden vanheneminen, kehittyminen ja oppiminen, fyysisen tilan muuttuminen, asenteiden muuttuminen, olosuhteiden/kontekstin muuttuminen jne.)

•Mittaus tapahtuma: Testaus on saattanut vaikuttaa kohteeseen? Mittareissa, arviointisijoissa tai havaintojen tekijöissä on voinut tapahtua muutoksia?

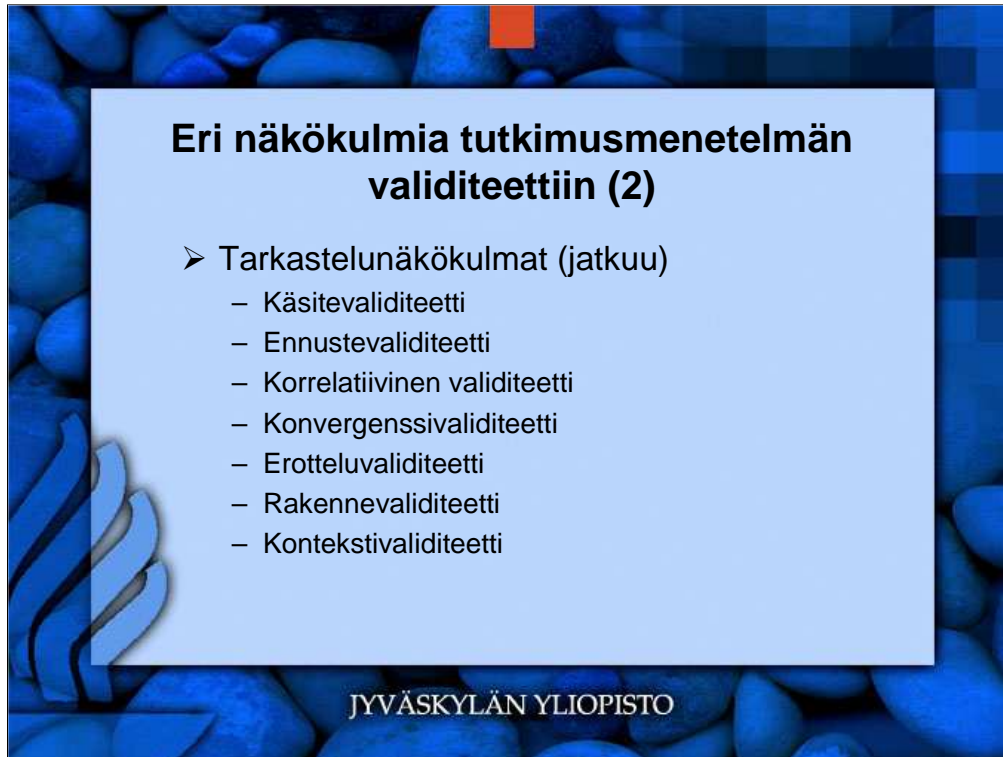
•Mittari itsessään: Onko mittari validi? Mittaako se sitä mitä sen tuleekin mitata? Onko mittari itsessään validiksi määritelty, esim. kalibroitu metrin mitta? Pohjautuuko mittarin validius muista vastaavista mittareista saatuihin empiirisiin kokemuksiin? Onko kyselylomake validi? Kaikissa edellä mainituissa tapauksissa validiteetti on aina suhteellinen, määritelty riittäväksi johonkin tiettyyn tarkoitukseen. Mahdolliset mittarin epätarkkuudet ja virheet kuuluu poistaa esim. kalibroimalla mittarit sekä selvittää esim. keskivirheen ja hajonnan määrä mitatuissa kohteissa.

•Vinoutumat? Onko valintakriteereissä vinoutumia? Silloin otokseen on tullut valituksi huomaamatta joidenkin tiettyjen kriteereiden mukaisia tapauksia.

•Kato? Tapahtuuko koehenkilöiden tai kohteiden häviämistä aineistosta? Esimerkiksi kyselytutkimuksissa ei ole saatu vastauksia kaikilta riittävässä määrin.

3) **Ulkoinen validiteetti** kiinnostaa silloin, kun pohditaan kvantitatiivisen tutkimuksen tulosten yleistettävyyttä: Missä populaatiossa, missä tilanteissa, missä asetelmissa saatu tulos voidaan yleistää? Ulkoisen validiteetin heikkouteen saattavat vaikuttaa monet tekijät, esimerkiksi tutkittavan kohteen valinta on vino perusjoukkoon nähden, jolloin tutkittavat edustavat vain jotakin osaa perusjoukosta. Esikoe tai -testaus: vaikuttaa ja muuttaa esitestattuja toisin kuin varsinaisia koehenkilöitä. Koemenetelmien yhteisvaikutus voi myös olla erilainen kuin yksittäisten koemenetelmien vaikutus ja vaikeasti erotettavissa.

4) **Aineisto- eli sisältövaliditeetti** - joka on eri asia kuin sisäinen validiteetti - tarkoittaa tutkimusaineistoon liittyvää validiteettia ja kuvastaa sitä, kuinka hyvin aineiston analysointimenetelmä vastaa tutkimusaineistoa. Kun reliabiliteetti kuvastaa sitä, kuinka hyvin menetelmä toimii aineiston analysoinnissa, sisältövaliditeetti kuvastaa sitä, kuinka hyvin koottu aineisto vastaa ulkopuolisia kriteereitä. Sisältövaliditeetin arviointi kuuluu varsinkin laadullisen tutkimuksen moniin erilaisiin sisällönanalyysin menetelmiin. Tutkimusprosessin on oltava arvioitavissa ja arvioijan on kyettävä seuraamaan tutkijan päättelyä. Hänen on kyettävä näkemään, etteivät tulokset perustu pelkästään tutkijan henkilökohtaiseen intuitioon. Siksi tutkijan tulee mahdollisimman selvästi kuvata aineistonsa, tekemänsä tulkinnat sekä ratkaisut- ja tulokset.



5) **Käsitevaliditeetti** (teoriavaliditeetti) kertoo siitä, onko teoreettinen lähestymistapa valittu oikein? Onko käytetty oikeita käsitteitä? Onko ilmiön luonnetta tulkittu oikein? Tässä suhteessa kysymys liittyy tutkimusfilosofisiin kysymyksiin ja tutkittavan ilmiön ontologisiin taustaoletuksiin. Erityisen tärkeä tämä asia on silloin, jos on asetettu tutkimushypoteeseja, joihin on teorian pohjalta laadittu erilaisia mittareita. Esimerkkinä ovat sellaiset testit, joita käytetään taitosuoritusten mittareina. Elleivät ne täytä käsitevaliditeetin vaatimuksia, saattaa taitosuorituksen rakenteen tulkinta muodostua vääräksi.

6) **Ennustevaliditeetti** tarkoittaa sitä, pystytäänkö tutkimusmenetelmän tai mittareiden avulla saamaan aikaan tuloksia, joilla on ennustearvoa? Tavallisesti tuloksia voidaan verrata käytännössä havaittuihin tosiasioihin. Jos ne eivät täsmää tutkimustulosten kanssa, saattaa vika olla paitsi mittarissa itsessään, myös tutkimusmenetelmässä tai -tilanteessa tai -ajoituksessa tms. Tavallinen esimerkki tämälantapaisista mittareista on oppilaitosten pääsykoetehtävät, joiden validiteetti voidaan todeta vasta myöhemmin: Ennustavatko ao. mittarit menestymistä opinnoissa vai myöhemmässä elämässä vai onko niillä mitään ennustearvoa? Toinen esimerkki on työhönottohaastattelu tai yleensä työhönottotapahtuman kriteerit. Niihin harvemmin kohdistetaan varsinaista tutkimusta, mutta aihetta ehkä olisi, nimenomaan ennustevaliditeetin kohottamiseksi.

7) **Korrelatiivinen validiteetti** tarkoittaa sitä, että tutkimuksen tuloksilla on korkea korrelaatio jossakin toisessa tutkimuksessa saatujen tulosten kanssa.

8) **Konvergenssivaliditeetti** tarkoittaa, osoittavatko mittaustulokset kaikki samaan suuntaan? Saataisiinko joillakin muilla mittareilla toisenlainen tuloskuva?

9) **Erotteluvaliditeetti** tarkoittaa, erottelevatko eri mittareilla saadut tulokset?

10) **Rakennevaliditeetti** tarkoittaa, kattavatko mittarit tutkittavan kohteen riittävästi?

11) **Kontekstivaliditeetti** tarkoittaa, liittyykö mittari ja tutkimusaineisto tutkittavaan ilmiöön; onko jokaisella tutkimusaineiston osalla jokin yhteys tutkimustulokseen?

Validiteetti tulosten näkökulmasta

- Validiteetillä tarkoitetaan tutkimuksen tai väitteen pätevyyttä eli oikeuttavatko käytetty aineisto, tutkimusmenetelmät ja saadut tulokset esitetyt väitteet
- Validiteetilla tarkoitetaan niiden päätelmien sopivuutta, mielekkyyttä ja käyttökelpoisuutta, joita mittaustuloksista tehdään²

² Nummenmaa ym. 1997, 203

JYVÄSKYLÄN YLIOPISTO

Tulosten näkökulmasta tarkasteltuna validiteetin katsotaan koskevan mittaustuloksista tehtäviä päätelmiä eikä mittaria. Toiseksi määritelmässä on huomioitu myös tulosten käyttökelpoisuus. Validiteetissa on kyse siitä, onko tutkimus pätevä; onko se perusteellisesti tehty, ovatko saadut tulokset ja tehdyt päätelmät "oikeita"; kuinka hyvin tutkijan konstruktiot vastaavat tutkittavien tuottamia ja kuinka hyvin hän tuottaa nämä konstruktiot ymmärrettäväksi myös muille. . Tutkimuksessa voi ilmetä virheitä esimerkiksi siinä, että tutkija näkee suhteita tai periaatteita virheellisesti, ei näe niitä tai kysyy väärää kysymyksiä. (ks. esim. Kirk & Miller 1986, 29-30.)

Tulosten kannalta validiteetti on traditionaalisesti jaettu kolmeen alaryhmään: *sisältövaliditeettiin*, *kriteerivaliditeettiin* ja *käsitevaliditeettiin*.

1) **Sisältövaliditeetti** ilmaisee sen, onko testiin valitut tehtävät edustava otos mittaukseen liittyvästä tehtäväuniversiumista. Tämä validiteetin laji on tärkeä laadittaessa koulukokeita, sillä kokeen tulee vastata annettua koealuetta ja opetussuunnitelmaa. Koska kaikkea ei voida kysyä, on edustavan otoksen valitseminen kaikista mahdollisista koetehtävistä välttämätön ehto onnistuneelle kokeelle.

2) **Kriteerivaliditeetti** jaetaan usein kahteen ryhmään, *ennustevaliditeettiin* ja *rinnakkaisvaliditeettiin*. Esimerkkinä ennustevaliditeetista on oppilaitosten valintakokeet, sillä valintakokeiden on kyettävä ennustamaan tulevaa opintomenestystä. Rinnakkaisvaliditeettia käytetään usein tieteellisessä tutkimuksessa, jolloin rakennettua uutta tehtäväsarjaa verrataan samaa asiaa mittaavaan tunnettuun ja standardoituun testiin. Kriteerivaliditeetin arvoksi saadaan korrelaation lukuarvo.

3) **Käsitevaliditeetti** liittyy saadun pistemäärän merkityksen tarkasteluun, minkä ominaisuuden, piirteen tai taidon määrää pisteet ilmaisevat. Tässä tarkastelussa on erittäin merkittävä asema mittaukseen liittyvällä teoreettisella viitekehyksellä, konstruktiolla, jonka vuoksi käsitevaliditeettia kutsutaan myös **konstruktio- tai rakennevaliditeetiksi**.

Edellisten validiteettilajien lisäksi puhutaan usein myös **face-** tai **ilmivaliditeetista**, jolloin arvioidaan intuitiivisesti testin uskottavuutta. Validiteetin lajeista käsitevaliditeetti on nuorin. Käsitevaliditeettia käytetään usein silloin, kun sisältövaliditeetin ja kriteerivaliditeetin avulla ei riittävän tarkasti voida osoittaa testin mittaavan sitä, mitä sen haluttiin mittaavan.

Modernin validiteetin teoreettisen yleistyksen on esittänyt Messick (1989a, 1989b, 1994), jonka mukaan validiteetti sisältää myös mittauksen tulkinnasta seuraavat eettiset ja yhteiskunnalliset seuraamukset. Hän esittää käsitevaliditeetin sisältävän kaikki validiteetin lajit, sillä käsitevaliditeettia arvioidaan myös sisältö- ja kriteerivaliditeetin kautta.

Esimerkkejä tulosten validiteetista

- Kysymys koulun ilmanvaihdosta ei vielä riitä mittaamaan sitä, kuinka hyvin opettaja viihtyy työssään
- Ei pidä haastatella ainoastaan koulun nykyisiä oppilaita, kun halutaan saada selville koulun imago

JYVÄSKYLÄN YLIOPISTO

Mitä tuloksista oikeasti voidaan päätellä? Onko aineisto riittävä, jotta siitä voidaan vetää toivottuja johtopäätöksiä?

Reliabiliteetti tutkimusmenetelmän näkökulmasta

- Reliabiliteetti (luotettavuus, engl. reliability) ilmaisee sen, miten luotettavasti ja toistettavasti käytetty mittaus- tai tutkimusmenetelmä mittaa haluttua ilmiötä
- Reliabiliteetin yhteydessä ilmoitetaan usein mittavirhe ja reliabiliteettia voidaan arvioida esim. toistomittauksilla
- Reliabiliteetissa erotetaan kaksi osatekijää: **stabiliteetti ja konsistenssi**

JYVÄSKYLÄN YLIOPISTO

Mittauksen reliabiliteetti tarkoittaa mittaustuloksen toistettavuutta, ei-sattumanvaraisuutta. Oletetaan, että mitataan kalibroiduin mittarein kutakuinkin samanlaisissa olosuhteissa työhuoneen ilman kosteutta ja lämpötilaa esimerkiksi kvasikokeellista tutkimusmenetelmää varten ja todetaan, että tulos on jokaisen päivän osalta lähestulkoon sama. Mittaus on siis ollut reliabeeli. Jos taas mittauksen tulos olisi jokaisen päivän osalta erilainen poiketen ensimmäisestä mittauksesta milloin yhteen milloin toiseen suuntaan, mittaus olisi likimain satunnaista.

Stabiliteetissa on kysymys mittarin tai menetelmän pysyvyydestä ajassa. Epästabiilissa mittarissa tai menetelmässä näkyvät olosuhteiden ja vastaajan mielialan ynnä muiden satunnaisvirheiden vaikutukset helposti. Mittarin tai menetelmän pysyvyyttä voidaan tarkastella vertaamalla useampia ajallisesti peräkkäisiä mittauksia. Tällöin aikavälin pituus tulisi osata optimoida: Sen pitää olla tarpeeksi pitkä, jotta vastaaja ei muista vastauksiaan, mutta toisaalta niin lyhyt, ettei todellisia muutoksia asioissa ole ehtinyt tapahtua. Monissa tapauksissa tämä reliabiliteetin mittaustapa ei ole toteuttamiskelpoinen, sillä huono reliabiliteettikerroin voidaan usein helpommin selittää ajassa tapahtuneilla todellisilla muutoksilla kuin epästabiililla mittarilla (Wright 1979, s. 47).

Mittarin konsistenssilla eli yhtenäisyydellä tarkoitetaan sitä, että kun useista väittämistä koostuva mittari jaetaan kahteen joukkoon väittämiä, kumpikin väittämäjoukko mittaa samaa asiaa. Tällöin molempien väittämäjoukkojen kokonaispistemäärien välinen korrelaatiokerroin saa suuren arvon. Koska ei ole mitään ulkoista kriteeriä, jolla testattaisiin mittarin reliabeeliutta, on tyydyttävä edellä kuvatulla tavalla "sisäisiin" kriteereihin eli samaan tutkimusjoukkoon ja mittariin itseensä. (Procter 1998, s. 128). Tämän toteamiseksi yleisesti käytetään mm. Cronbachin alfakerrointa, joka perustuu väittämien välisiin korrelaatioihin. On kuitenkin todettava, että on mahdollista luoda väittämäpatteristo, joka sisältää täysin eri asioita mittaavia, mutta keskenään voimakkaasti korreloivia muuttujia. Toisaalta saman ilmiön osaluaita mittaavat muuttujat eivät aina välttämättä korreloi keskenään ja kuitenkin niitä on tarpeen tarkastella yhdessä.

Samalla reliabiliteetin käsitteellä on määritelty siis kaksi varsin erilaista mittarin ominaisuutta. Stabiili mittari ei välttämättä ole konsistentti eikä konsistentti mittari välttämättä stabiili. Käytännössä reliabiliteetti liitetään pääasiassa mittarin konsistenssiin. Vaikka mittari olisi sekä konsistentti että stabiili, se ei riitä. Tutkimuksen mittari voi nimittäin mitata väärääkin asiaa hyvin johdonmukaisesti. Mittarin on oltava myös validi.

Nykyisten tilasto-ohjelmien aikana reliabiliteettikysymys saattaa kaventua mittausvirheen arvioinniksi, joka on tilasto-ohjelmiin valmiiksi sisäänrakennettu.

Tutkimusmenetelmän reliabiliteetin tarkistus

- Yhdenmukaisuus
 - miten eri indikaattorit mittaavat samaa asiaa
- Tarkkuus
 - toistuvan ilmiön havainnointitarkkuus
- Objektiivisuus vs. subjektiivisuus
 - kuinka muut ymmärtävät tekijän tarkoituksen ja viitekehysten
- Jatkuvuus
 - havainnon jatkuva samankaltaisuus

JYVÄSKYLÄN YLIOPISTO

Reliabiliteetin tarkistus voidaan jakaa neljään eri näkökulmaan:

- 1) **Kongruenssi eli yhdenmukaisuus**, jonka avulla tarkistetaan, miten eri indikaattorit mittaavat samaa asiaa. Mitä useampaa indikaattoria on mahdollista käyttää, sitä vakuuttavammaksi tieto tulee.
- 2) **Instrumentin tarkkuus**, jolla mitataan toistuvan ilmiön havainnointitarkkuus, ts. samanlaisen useasti toistuvan ilmiön rekisteröinnin yhtäläisyysaste. Se voidaan tarkistaa esimerkiksi havainnointitutkimuksessa toistuvien ilmiöiden osalta niin, että ilmiötä pyritään havainnoimaan mahdollisimman moneen kertaan ja saatuja tuloksia verrataan keskenään. Haastattelututkimuksissa voidaan samaa asiaa kysyä eri muodoissa samassa haastattelutilanteessa. Uusinta-haastattelussa voidaan myös kerrata samoja asioita ja näin saatuja tuloksia verrata keskenään.
- 3) **Instrumentin objektiivisuus vs. subjektiivisuus**, jolla tarkistetaan, miten pitkälle muut ymmärtävät havainnoinnin tekijän tarkoituksen. Se voidaan tarkistaa käyttämällä esimerkiksi useampaa kuin yhtä havainnoijaa samassa tilanteessa. Tämä kuitenkin edellyttää, että eri havainnoitsijoiden havainnointimahdollisuudet ovat suurin piirtein samanlaiset ja että tutkijoiden henkilökohtainen viitekehys tutkimuksen suorittamisessa on sama. Itse asiassa instrumentin subjektiivisuus on seikka, joka tällä tavoin tavallisesti tulee parhaiten esille. Eri havainnoitsijoiden tai haastattelijoiden saamissa tiedoissa ilmeneviä eroja voidaan tarkastella vertaamalla heidän ominaisuuksiaan keskenään (esim. ikä ja sukupuoli).
- 4) **Ilmiön jatkuvuus**, jolla ilmaistaan jonkin havainnon jatkuva samankaltaisuus. Ilmiön voidaan katsoa olevan jatkuva, jos se on todettu samankaltaisena eri aikoina. Jos ilmiö kuvataan jollakin tavalla tyyppillisenä, tutkijan on pystyttävä osoittamaan, että se ei ole ainutkertainen.

Alhainen reliabiliteetti voi kertoa esimerkiksi virheellisestä mittavälineestä tai sen väärästä tulkinnasta. Satunnaisvirheitä voi syntyä mittauksiin monista syistä. Esimerkiksi kysely- ja haastattelututkimuksissa vastaaja voi muistaa jonkin asian väärin, ymmärtää kysymyksen toisin kuin tutkija on ajatellut, hän tai haastattelija saattaa merkitä vastauksen väärin, tallennettaessa vastauksia tietokoneen muistiin saattaa tapahtua virheitä jne. Vaikka tavoitteeksi onkin asetettava mahdollisimman reliabeli mittaustapa, puutteellinen reliabiliteetti ei välttämättä ole tutkimukselle tuhoisa. On tärkeää pyrkiä arvioimaan käytettyjen mittarien reliabiliteettiä, jolloin myös sen vaikutukset ovat arvioitavissa.

Reliabiliteetin laskemista ja arviointia varten onkin kehitetty erilaisia menettelytapoja:

- 1) **Puolitusmenetelmän** (split-half -menetelmän) avulla lasketaan jonkin mittarin kahden eri puoliskon välinen korrelaatio. Voidaan ottaa ensin käsittelyyn vain puolet aineistosta, laatia sen pohjalta alustava tulkinta ja tarkastaa sitten jäljelle jääneen aineiston puolikkaan avulla voidaanko tulkinnot yleistää koko aineistoon.
- 2) **Rinnakaistestimenetelmä** tarkoittaa jollekin mittarille laadittua vastaavaa ominaisuutta mittaavaa toista mittaria. Hyvä reliabiliteetti osoittaa näiden mittareiden korkeaa korrelaatiota.
- 3) **Retestausten menetelmä** eli toistomittaus tarkoittaa samaa mittausta kahdesti.
- 4) **Arvioitsijareliabiliteetti** tarkoittaa eri arvioitsijoiden välisten arviointien välistä kerrointa, esimerkiksi yksimielisyyss prosenttia tms. Jos niiden keskinäinen kerroin on korkea ($r =$ esim. .80 tai enemmän) voidaan reliabiliteetin sanoa olevan hyvä.

Tutkimuksen reliabiliteetti (1)

- Reliabiliteetilla tarkoitetaan tutkimustulosten ja väitteiden luotettavuutta: johtuuko tutkimustulos vain sattumasta vai kyetäänkö tulokset riippumattomasti toistamaan? Toistaminen voi tapahtua joko arvioitsijoiden välillä (inter-rater reliability) tai tutkimuskertojen välillä (test-retest reliability).
 - luotettava tutkimus on tehty siten, että se alkuperäisen tutkimuksen edellytyksin toistettuna antaa saman tuloksen

JYVÄSKYLÄN YLIOPISTO

Tutkimuksen reliabiliteetti tarkoittaa mittaustulosten toistettavuutta; tutkimustulokset eivät ole sattumanvaraisia tuloksia. Tutkimuksen reliabiliteetti on hyvä, kun tulokset eivät ole sattuman aiheuttamia. Jos tutkimus uusittaisiin, pitäisi samoissa olosuhteissa saada samat tulokset. Kysymysten tulee olla yksiselitteisiä ja ymmärrettäviä ja haastattelut tulee tehdä huolellisesti.

Reliabiliteetin arvioinnin ja tarkistamisen kannalta on tärkeää, että tutkimusaineisto muokataan sellaiseen muotoon, että se on tutkimuksen kommentoijien saatavilla ja tarkastettavissa. Periaatteessa tutkimusmenetelmä voi olla reliabiliteetti eli luotettava vaikka tutkimus ei olisikaan validi eli pätevä. Tällöin tutkimuksessa käytettyjen menetelmien avulla voidaan päätyä sinänsä kiinnostaviin tuloksiin, mutta tulokset eivät vastaa sitä, mihin tutkimuksella tähdättiin.

Yhteenvetoa

- Reliabiliteetin ja validiteetin suhteelle pätee: mitä alhaisempi reliabiliteetti, sitä alhaisempi validiteetti
- Kääntäen väite ei kuitenkaan pidä paikkansa, sillä mittaus voi antaa hyvinkin reliaabeleja tuloksia, mutta aivan eri asiasta kuin mitä on tarkoitus mitata

¹Polit & Hungler 1997

JYVÄSKYLÄN YLIOPISTO

Mittauksessa yksittäisen tilastoyksikön kohdalla saatu tieto voi olla oikea tai virheellinen. Mittareiden antamien tietojen ja muuttujien todellisten arvojen erot eli ns. **mittausvirheet** voidaan jakaa satunnaisvirheisiin, jotka alentavat mittauksen **reliabiliteettia** ja systemaattisiin, **validiteettia** alentaviin virheisiin. Virheen **satunnaisuus** tarkoittaa sitä, että virheiden keskiarvo koko aineistossa on nolla ja virheet eivät liity mihinkään systemaattisesti vaikuttavaan tekijään. Virhe ei liity itse mittariin, vaan on siihen nähden ulkoista, esim. lyöntivirhe atk-tallennuksessa. **Systemaattinen** virhe sisältyy itse mittariin, jolloin vääristymä koko aineistossa on tietyn suuntainen, esim. epäselvä kysymys.

Reliabiliteetin ja validiteetin suhteelle pätee: mitä alhaisempi reliabiliteetti, sitä alhaisempi validiteetti. Kääntäen väite ei kuitenkaan pidä paikkansa, sillä mittaus voi antaa hyvinkin reliaabeleja tuloksia, mutta aivan eri asiasta kuin mitä on tarkoitus mitata. Esim. haastattelussa vastaaja puhuu systemaattisesti totuudenvastaisesti.



Kun teette tutkimusta, tehkää se kerralla kunnolla. Miettikää huolella

- miten tutkimusongelmat tai –tehtävät kannattaa muotoilla, jotta pystyt saamaan niihin vastaukset?
- mitä tutkimusmenetelmää kannattaa käyttää, jotta saat selville sen mitä halusitkin ja että menetelmä on luotettava?
- miten aineisto kannattaa analysoida, jotta tulokset olisi luotettavia, yleistettäviä ja toistettavissa?
- miten perustelet tulokset ja millaisia johtopäätöksiä tuloksista teet?
- miten raportoit kaiken tämän siten, että joku muu voi toistaa tutkimuksesi niin halutessaan ja että raportointiin ei jää aukkoja?