

LC WHITE PAPER SERIES

TESTING AV LCP FOR BIAS

Lani Van Dusen, Ph.D.

Utført for Leadership Circle

LEADERSHIP
CIRCLE®

SAMMENDRAG

I dagens kultur er det en økende sensitivitet for potensiell bias når det gjelder hvordan mennesker observeres og evalueres. Det er umulig å fjerne all personlig bias fra individers evalueringer – det er i sakens natur at mennesker gir bedømmelser basert på stereotyper og har oppfatninger basert på egne erfaringer. Påvirkningen av personlig bias på vurderingsresultater kan imidlertid minimeres ved å sikre at målemodellen og relaterte prestasjonsresultater ikke er vesentlig forskjellig, basert utelukkende på gruppetilhørighet til personene som evalueres.

Det finnes flere kilder til målebias og tilhørende tester som kan utføres for å fastslå om et vurderingsverktøy viser slik bias. En god gjennomgang av forskningslitteraturen angående disse testene er gitt i en artikkel av Reynolds og Suzuki (2012) med tittelen Bias in Psychological Assessment (<https://onlinelibrary.wiley.com/doi/full/10.1002/9781118133880.hop21004>).

Vi valgte de tre testene som er mest kjente og aksepterte i forskningsmiljøet for å fastslå i hvilken grad LCP er like relevant for alle ledere, uavhengig av kjønn eller etnisk tilhørighet. (Vær oppmerksom på at kjønn og etnisitet ble selvrapportert på LCP, og at disse selvidentifiseringene ble brukt for å teste relevansen av LCP for ulike grupper.) Resten av denne avhandlingen gir resultatene av disse testene.

TESTING AV LCP FOR BIAS

Utført for:

Leadership Circle

av:

Lani Van Dusen, Ph.D.



Mars 2021

Introduksjon

I dagens kultur er det en økende sensitivitet for potensiell bias når det gjelder hvordan mennesker observeres og evalueres. Det er umulig å fjerne all personlig bias fra individers evalueringer – det er i sakens natur at mennesker gir bedømmelser basert på stereotyper og har oppfatninger basert på egne erfaringer. Påvirkningen av personlig bias på vurderingsresultater kan imidlertid minimeres ved å sikre at målemodellen og relaterte prestasjonsresultater ikke er vesentlig forskjellig, basert utelukkende på gruppetilhørighet til personene som evalueres.

Det finnes flere kilder til målebias og tilhørende tester som kan utføres for å fastslå om et vurderingsverktøy viser slik bias. En god gjennomgang av forskningslitteraturen angående disse testene er gitt i en artikkel av Reynolds og Suzuki (2012) med tittelen *Bias in Psychological Assessment* (<https://onlinelibrary.wiley.com/doi/full/10.1002/9781118133880.hop21004>).

Vi valgte de tre testene som er mest kjente og aksepterte i forskningsmiljøet for å fastslå i hvilken grad LCP er like relevant for alle ledere, uavhengig av kjønn eller etnisk tilhørighet. (Vær oppmerksom på at kjønn og etnisitet ble selvrapportert på LCP, og at disse selvidentifiseringene ble brukt for å teste relevansen av LCP for ulike grupper.) Resten av denne avhandlingen gir resultatene av disse testene.

Konstruktbias

En kilde til målebias dannes når innholdet (elementene) i vurderingsverktøyet fokuserer på kunnskap eller verdier som samsvarer med erfaringene til én gruppe, men ikke en annen (eller som kan ses som støtende av én gruppe). Et eksempel på dette i forskningslitteraturen kommer fra analyser utført på standardiserte tester av matematiske ferdigheter gitt til skolebarn. Hvis alle de matematiske tekstoppgavene (elementene i vurderingen) er basert på sportstemaer, kan det hende at en jente ikke gjør det så bra fordi hun kan muligens ha mindre erfaring på dette området.

Konstruktene til LCP er basert på en universell modell for ledelse, og innholdet er orientert rundt dagligdagse forretningsopplevelser og samspill blant voksne. Gitt at alle evaluatorene og ledere er en del av et forretningsmiljø og deler disse erfaringene uavhengig av deres gruppetilhørighet, er det sannsynlig at LCP-innholdet vil være relevant for alle ledere. I tillegg, gitt at de fleste organisasjoner (spesielt større selskaper) har forretningsmiljøer og arbeidsgrupper som består av både menn og kvinner, samt ledere fra forskjellige etniske bakgrunner, vil det være usannsynlig at arbeidsrelaterte atferder og holdninger vil favorisere erfaringene til én gruppe fremfor en annen.

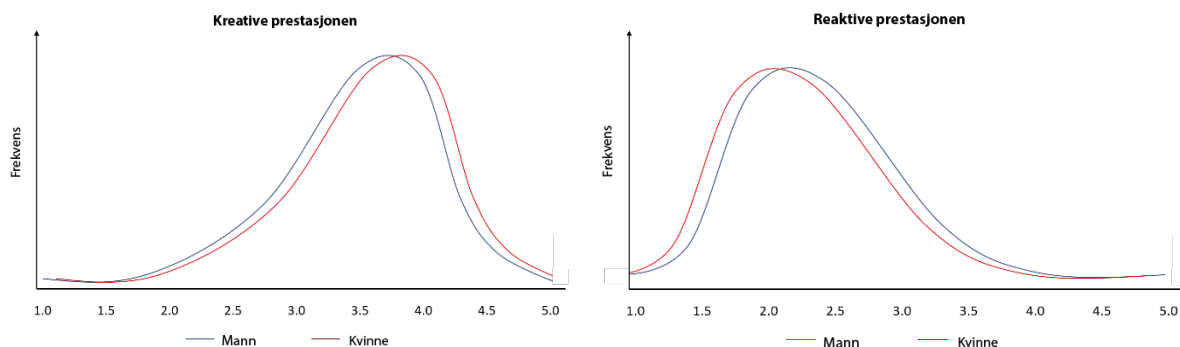
Målefeil

En annen viktig kilde til bias blir introdusert når en vurdering ikke måler alle grupper rettferdig. Det er spesielt viktig å forstå at dette ikke betyr at vurderingen ikke vil finne forskjeller i gjennomsnittlig prestasjon for forskjellige grupper. Det kan faktisk være reelle forskjeller mellom grupper, og vurderingen bør være sensitiv nok til å fange opp disse forskjellene. En forskningsstudie utført av Okozi, Smith, Harvey og Sherman (2009) med tittelen *Leadership Styles of Ethnic Minority Leaders*, som ble publisert i *Communique* av American Psychological Association (<https://www.apa.org/pi/oema/resources/communique/2009/08/ethnic-leaders>) demonstrerte for eksempel at det er forskjeller i lederegenskapene og tilnærmingene som brukes av ledere av forskjellige etnisiteter.

Det som gjør en vurdering partisk, er om det finnes «systematisk feil ved estimering av en verdi.» Det betyr at LCP systematisk vil måtte overvurdere eller undervurdere én eller flere av dimensjonsskårene for en bestemt gruppe. Den gjennomsnittlige prestasjonen kan være forskjellig for en hvilken som helst dimensjon, men fordelingen rundt gjennomsnittene bør være lik.

Da vi testet dette ved hjelp av LCP-dataene, fant vi ut at fordelingskurvene for alle grupper var ganske like. Et eksempel på dette er fordelingskurven for den kreative og reaktive prestasjonen til mannlige og kvinnelige ledere. Det finnes en gjennomsnittlig forskjell – det betyr at toppen av kurven er litt til høyre på Kreativ for kvinnelige ledere og litt til venstre for Reaktiv, noe som indikerer høyere gjennomsnittlig prestasjon for kvinnelige ledere (i samsvar med andre av våre forskningsstudier).

Den generelle formen på fordelingen er imidlertid lik, og forskjellene som ble observert mellom Kreativ og Reaktiv er konsekvente på tvers av kjønn. (Vær oppmerksom på at det finnes mange kvinnelige ledere som skårer under den gjennomsnittlige mannlige lederen, og det finnes mange mannlige ledere som skårer over den gjennomsnittlige kvinnelige lederen, noe som tyder på at instrumentet ikke konsekvent måler kvinnelig prestasjon høyere.)



Differensiell prediktiv validitet

Målet med enhver standardisert måling er å kunne forutsi utfall nøyaktig. Hvis det finnes målebias, vil nøyaktigheten (eller styrken) av prediksjonene være forskjellig for forskjellige grupper. For å teste dette for LCP korrelerte vi utviklingsprestasjon på alle dimensjoner med generell effekt av ledelse (skår på skalaen Effekt av ledelse) for mannlige og kvinnelige ledere, samt ledere som identifiserte seg med forskjellige etnisiteter.

Basert på LCP-modellen ville vi forutsi at etter hvert som en leder utvikler mer kreative ferdigheter, bør effektiviteten hennes/hans bli bedre (positiv korrelasjon). Og etter hvert som en leder eliminerer eller reduserer sine reaktive tendenser, bør effektiviteten hennes/hans bli bedre (negativ korrelasjon). Som anvist i den medfølgende tabellen, er validiteten av disse prediksjonene sterk og veldig lik for alle grupper.

Dette funnet tyder på at uavhengig av lederens selvidentifiserte bakgrunn, utgjør deres utviklingsprestasjon som målt av LCP, en betydelig del av variansen i deres generelle effektivitet. Sagt på en annen måte, LCP kan brukes til å forutsi lederes effektivitet uten bias mot noen gruppe.

Totalt sett tyder testene våre på at LCP-vurderingen ikke viser noen systemisk bias for kjønn eller etnisitet. Det kan være et effektivt verktøy for å gi innsikt i ledelsesutvikling for alle grupper av ledere.

Gruppe Identifiserer som:	Kreativ Korrelasjon med LE	Reaktiv Korrelasjon med LE
Mann	$r = 0,93, (p < 0,001)$	$r = -0,71, (p < 0,001)$
Kvinne	$r = 0,94, (p < 0,001)$	$r = -0,76, (p < 0,001)$
Amerikansk urinnvåner	$r = 0,91, (p < 0,001)$	$r = -0,65, (p < 0,001)$
Asiatisk stillehavsboer	$r = 0,94, (p < 0,001)$	$r = -0,79, (p < 0,001)$
Svart / afrikansk- amerikansk	$r = 0,94, (p < 0,001)$	$r = -0,73, (p < 0,001)$
Spansk / latinamerikansk	$r = 0,93, (p < 0,001)$	$r = -0,70, (p < 0,001)$
Hvit	$r = 0,94, (p < 0,001)$	$r = -0,74, (p < 0,001)$

Om WIRE

WIRE er et uavhengig oppdragsbyrå som ble etablert i 1978 for å tilby konsulenttjenester til organisasjoner som ønsker å forbedre eller vurdere menneskelig prestasjon og for å evaluere programmer eller produkter som er målrettet mot menneskelig prestasjon. Som sådan tilbyr vi teknisk støtte og tjenester i full skala som strekker seg fra sekundæranalyseforskning til omfattende nasjonale og internasjonale studier. Ytterligere informasjon om WIRE finner du på nettstedet vårt: www.worldwideinstitute.biz.