

LC WHITE PAPER SERIES

LCP-BIAS-TEST

Lani Van Dusen, Ph.D.

Ausführung im Leadership Circle

LEADERSHIP
CIRCLE®

ZUSAMMENFASSUNG

Es gibt heute ein zunehmendes Bewusstsein für mögliche Verzerrungen bei der Beobachtung und Beurteilung von Menschen. Solche Verzerrungen lassen sich nicht vollständig eliminieren, da der Mensch bei seiner Urteilsbildung naturgemäß durch Stereotype und Wahrnehmungen beeinflusst wird, die auf seinen Erfahrungen beruhen. Allerdings können die Auswirkungen des persönlichen Bias auf das Ergebnis einer Beurteilung minimiert werden, wenn darauf geachtet wird, dass Messmodell und Leistungsergebnisse keine ausschließlich auf der Gruppenzugehörigkeit der beurteilten Personen basierenden signifikanten Unterschiede aufweisen.

Es gibt mehrere Quellen für Messverzerrungen und entsprechende Tests zur Analyse von Beurteilungsinstrumenten auf solche Fehler. Der Artikel „Bias in Psychological Assessment“ (Reynolds und Suzuki, 2012; <https://onlinelibrary.wiley.com/doi/full/10.1002/9781118133880.hop21004>) enthält einen Überblick über die Forschungsliteratur zu diesen Tests.

Wir haben anhand dreier bekannter und wissenschaftlich anerkannter Tests untersucht, ob das LCP für alle Führungspersonen die gleiche Relevanz aufweist, unabhängig von deren Geschlecht oder ethnischer Zugehörigkeit. (Geschlecht und ethnische Zugehörigkeit werden beim LCP von den untersuchten Personen angegeben. Diese Selbstauskunft wurde für den Test verwendet.) Im Folgenden wird das Testergebnis präsentiert.

LCP-BIAS-TEST

Ausführung im

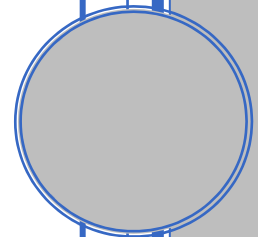
Leadership Circle

Autorin:

Lani Van Dusen, Ph.D.



März 2021



Einführung

Es gibt heute ein zunehmendes Bewusstsein für mögliche Verzerrungen bei der Beobachtung und Beurteilung von Menschen. Solche Verzerrungen lassen sich nicht vollständig eliminieren, da der Mensch bei seiner Urteilsbildung naturgemäß durch Stereotype und Wahrnehmungen beeinflusst wird, die auf seinen Erfahrungen beruhen. Allerdings können die Auswirkungen des persönlichen Bias auf das Ergebnis einer Beurteilung minimiert werden, wenn darauf geachtet wird, dass Messmodell und Leistungsergebnisse keine ausschließlich auf der Gruppenzugehörigkeit der beurteilten Personen basierenden signifikanten Unterschiede aufweisen.

Es gibt mehrere Quellen für Messverzerrungen und entsprechende Tests zur Analyse von Beurteilungsinstrumenten auf solche Fehler. Der Artikel „Bias in Psychological Assessment“ (Reynolds und Suzuki, 2012; <https://onlinelibrary.wiley.com/doi/full/10.1002/9781118133880.hop21004>) enthält einen Überblick über die Forschungsliteratur zu diesen Tests.

Wir haben anhand dreier bekannter und wissenschaftlich anerkannter Tests untersucht, ob das LCP für alle Führungspersonen die gleiche Relevanz aufweist, unabhängig von deren Geschlecht oder ethnischer Zugehörigkeit. (Geschlecht und ethnische Zugehörigkeit werden beim LCP von den untersuchten Personen angegeben. Diese Selbstauskunft wurde für den Test verwendet.) Im Folgenden wird das Testergebnis präsentiert.

Konstrukt-Bias

Messverzerrungen können entstehen, wenn der Inhalt (Items) eines Beurteilungsmodells auf Wissen oder Werte abzielt, die den Erfahrungen einer Gruppe entsprechen, nicht aber denen einer anderen (bzw. die von einer Gruppe als anstößig empfunden werden könnten). Ein Beispiel aus der Forschungsliteratur ist die Analyse standardisierter Tests der Mathematikkompetenz von Schülern. Basieren Textaufgaben (die „Items“) sämtlich auf Sportthemen, könnte ein Mädchen möglicherweise weniger gut abschneiden, weil sie sich in diesem Bereich eventuell weniger auskennt.

Die Konstrukte des LCP basieren auf einem universellen Führungsmodell und der Inhalt ist an alltäglichen Geschäftserfahrungen und Interaktionen zwischen Erwachsenen orientiert. Da alle Bewerter/innen und Führungspersonen in einem geschäftlichen Umfeld agieren und entsprechende Erfahrungen unabhängig von ihrer Gruppenzugehörigkeit teilen, ist anzunehmen, dass der LCP-Inhalt für alle Führungspersonen relevant ist. Zudem sind in den meisten Organisationen (insbesondere in größeren Unternehmen) Männer und Frauen sowie Führungspersonen verschiedener Ethnien im Geschäftsumfeld und in Arbeitsgruppen vertreten, sodass es unwahrscheinlich ist, dass bei arbeitsbezogenen Verhaltensweisen und Einstellungen die Erfahrungen einer Gruppe gegenüber denen einer anderen vorwiegen.

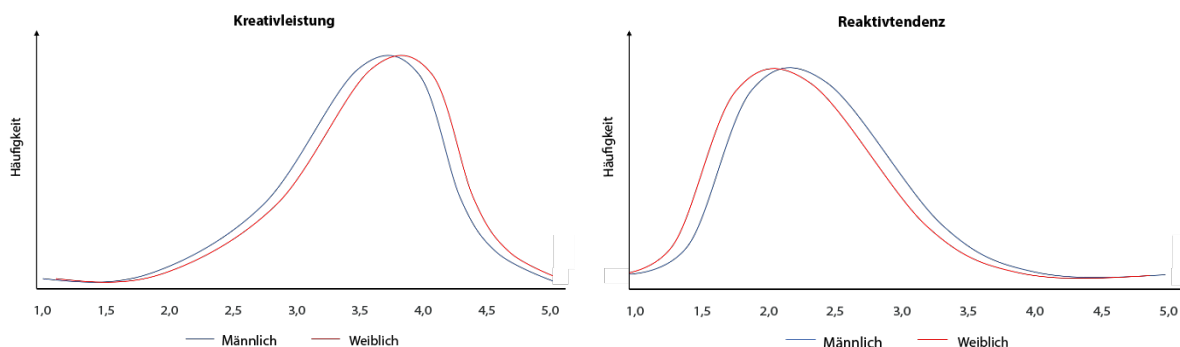
Messfehler

Eine weitere Ursache für Bias kann die nicht faire Messung aller Gruppen sein. Das bedeutet nicht, dass bei der Beurteilung keine Unterschiede bei den durchschnittlichen Leistungen der verschiedenen Gruppen festgestellt werden dürfen. Diese kann es durchaus geben und eine Messung sollte ausreichend empfindlich sein, um sie zu erkennen. Eine im Jahr 2009 von Okozi, Smith, Harvey und Sherman unter dem Titel „Leadership Styles of Ethnic Minority Leaders“ durchgeführte Studie ergab beispielsweise Unterschiede bei den Führungsfähigkeiten und -praktiken von Führungskräften verschiedener Ethnien (siehe Website der American Psychological Association auf <https://www.apa.org/pi/oema/resources/communique/2009/08/ethnic-leaders>).

Eine Beurteilungsverzerrung entsteht, wenn ein systematischer Fehler bei der Schätzung eines Wertes vorliegt. Hierfür müsste das LCP eines oder mehrere der Dimensionsergebnisse einer spezifischen Gruppe systematisch über- oder unterschätzen. Die mittlere (durchschnittliche) Leistung kann in jeder Dimension unterschiedlich sein, doch die Verteilung um die Mittelwerte sollte ähnlich ausfallen.

Der entsprechende Test an LCP-Daten ergab eine große Ähnlichkeit bei den Verteilungskurven aller Gruppen. Als Beispiel können die Verteilungskurven der kreativen Leistung und der reaktiven Tendenz männlicher und weiblicher Führungskräfte dienen. Hier besteht ein Unterschied beim Durchschnitt: Der Scheitelpunkt der Kurve liegt bei weiblichen Führungskräften geringfügig weiter rechts im Bereich Kreativ und geringfügig weiter links im Bereich Reaktiv, was auf eine höhere durchschnittliche Leistung weiblicher Führungskräfte hindeutet (in Übereinstimmung mit weiteren von uns durchgeführten Studien).

Die Gesamtform der Verteilung ist jedoch ähnlich, und die Unterschiede zwischen Kreativleistung und Reaktivtendenz sind bei den Geschlechtern gleich. (Es gibt viele weibliche Führungskräfte, die unter dem Durchschnitt der männlichen Führungskräfte liegen, und viele männliche Führungskräfte, die über dem Durchschnitt der weiblichen Führungskräfte liegen, was darauf hindeutet, dass das Instrument keine durchgängig höhere Leistung bei den Frauen misst.)



Differenzielle prädiktive Validität

Das Ziel jeder standardisierten Messung ist eine präzise Ergebnisvorhersage. Messverzerrungen führen zu einer unterschiedlichen Genauigkeit der Vorhersagen für verschiedene Gruppen. Für den entsprechenden Test des LCP wurde die Entwicklungsleistung in jeder Dimension mit der allgemeinen Führungseffektivität (Wert auf der Skala Führungseffektivität) männlicher und weiblicher Führungskräfte unter Berücksichtigung der angegebenen ethnischen Zugehörigkeit korreliert.

Auf der Grundlage des LCP-Modells wäre anzunehmen, dass sich die Effektivität einer Führungsperson verbessert, wenn sie ihre kreativen Fähigkeiten entwickelt (positive Korrelation), und ebenso, wenn sie ihre reaktiven Tendenzen eliminiert oder verringert (negative Korrelation). Die Tabelle unten belegt eine starke und sehr ähnliche Validität dieser Vorhersagen für alle Gruppen.

Das Ergebnis deutet darauf hin, dass die mit dem LCP gemessene Entwicklungsleistung der Führungspersonen unabhängig von deren Gruppenzugehörigkeit einen erheblichen Anteil an der Varianz ihrer Gesamteffektivität hat. Anders gesagt: Das LCP kann für Vorhersagen über die Effektivität von Führungspersonen ohne gruppenbezogene Verzerrung verwendet werden.

Insgesamt deuten unsere Tests darauf hin, dass die LCP-Beurteilung kein systematisches Bias in Bezug auf Geschlecht oder ethnische Zugehörigkeit aufweist. Sie kann als wirksames Instrument zur Gewinnung von Erkenntnissen über die Entwicklung von Führungsqualitäten für alle Gruppen von Führungspersonen dienen.

Angeg. Gruppen-zugehörigkeit	Kreativ Korrelation mit Führungseffektivität	Reaktiv Korrelation mit Führungseffektivität
Männlich	$r = 0,93$ ($p < 0,001$)	$r = -0,71$ ($p < 0,001$)
Weiblich	$r = 0,94$ ($p < 0,001$)	$r = -0,76$ ($p < 0,001$)
Indigene Bevölkerung (US)	$r = 0,91$ ($p < 0,001$)	$r = -0,65$ ($p < 0,001$)
Asiatisch/Pazifische Inseln	$r = 0,94$ ($p < 0,001$)	$r = -0,79$ ($p < 0,001$)
Schwarz/Afroamerikanisch	$r = 0,94$ ($p < 0,001$)	$r = -0,73$ ($p < 0,001$)
Hispanisch/Latein-amerikanisch	$r = 0,93$ ($p < 0,001$)	$r = -0,70$ ($p < 0,001$)
Weiß	$r = 0,94$ ($p < 0,001$)	$r = -0,74$ ($p < 0,001$)

Über WIRE

WIRE ist eine unabhängige 1978 gegründete Agentur, die Organisationen zur Verbesserung oder Beurteilung menschlicher Leistung berät und Programme bzw. Produkte, die auf die menschliche Leistung abzielen, evaluiert. WIRE ist Anbieter technischer Unterstützung und vielfältiger Dienstleistungen, von Sekundäranalysen bis hin zu umfassenden nationalen und internationalen Studien. Weitere Informationen über WIRE siehe www.worldwideinstitute.biz.