



Psychologische Rundschau

Herausgeber

Karl Christoph Klauer
Jens Bölte
Oliver Dickhäuser
Klaus Moser
Silja Vocks

Offizielles Organ der Deutschen Gesellschaft für Psychologie (DGPs)

Diskussionsforum

Positionspapier zur Lage der Pädagogischen
Psychologie in Forschung und Lehre

TBS-TK-Rezension

Leadership Style Assessment. Ein Situational Judgment Test zur Erfassung von Führungsstilen

Jan Schilling, Kommunale Hochschule für Verwaltung in Niedersachsen
 Sonja Lippke, Jacobs University Bremen

Allgemeine Informationen

Das „Leadership Style Assessment“ (LSA) soll unterschiedliche Führungsstile erfassen und situationsbasiert Rückmeldung über das selbstwahrgenommene (Version für Führungskräfte) bzw. fremdeingeschätzte Verhalten (Version für Mitarbeitende) geben. Es richtet sich an Verantwortliche für die Auswahl, Entwicklung und Beurteilung von Führungskräften. Angesichts der Vielfalt bestehender Theorien zum Thema „Führung und Führungsverhalten“ und darauf aufbauender Verfahren verweisen die Autoren insbesondere auf die wissenschaftlich-theoretische Fundierung in Form des „Full Range of Leadership Models“ (FRL; Bass & Riggio, 2006). Die Nutzung eines Situational Judgment Tests ermöglicht es, den Gedanken eines situativ angepassten Führungsverhaltens abzubilden. Der LSA besteht aus acht Situationsbeschreibungen, zu denen jeweils acht mögliche Reaktionsmöglichkeiten präsentiert werden. Die acht Reaktionsmöglichkeiten entsprechen den acht Dimensionen des FRL, sodass über die acht Situationen hinweg acht theoriebasierte Skalen gebildet werden können. Insgesamt sind die Verfahrenshinweise informativ, fachlich korrekt und umfassend sowie ausreichend sparsam.

Theoretische Grundlagen

Das empirisch gut gestützte FRL-Modell von Bass und Riggio (2006), das ein breites Spektrum von ineffektiver (laissez-faire) über eher effektive (transaktional) bis zu sehr effektiver Führung (transformational) abbildet, wird gut in den Erhebungsmaterialien repräsentiert. Alle Bereiche sind nachweisbar von Bedeutung für effektives Führungsverhalten, Teambildung und die Förderung der Mitarbeitenden. Die Autorinnen und Autoren verknüpfen die Führungsstile mit situativen Einflüssen, wobei Letztere auf Basis von Fokusgruppen-Diskussionen mit Führungskräften mittels der Critical Incident Technique entstanden sind. Die aktuelle Forschung zu destruktiver Führung (z. B. Schyns & Schilling, 2013; Padilla, Hogan & Kaiser, 2007) wird nicht mit einbezogen, was einerseits die Bandbreite möglicher Führungsverhaltensweisen im Test einschränkt, andererseits angesichts der intendierten

Selbstbewertung des Führungsverhaltens verständlich ist. Ebenso bleibt die empirische Evidenz des Test-Entwicklungsplanes unklar.

Objektivität

Es finden sich detaillierte Instruktionen zur Durchführung des als Papier- oder Computerversion nutzbaren Tests. Bei letzterer ist positiv zu bewerten, dass über generierte Codes eindeutige Zuordnungen von Mitarbeiter- und Führungskraftantworten ohne Anonymitätsgefährdung möglich sind. In der Papier-Version ist dagegen eine Beeinflussung des Antwortverhaltens nicht auszuschließen, da die Führungskraft die Fragebogen selbst an Mitarbeitende ausgeben kann. Die Auswertungsobjektivität ist durch standardisierte Verfahren sichergestellt, wie etwa einen automatisierten Bericht bzw. Instruktionen und Auswertungsschablonen. Gleichzeitig gilt für die Interpretation, die mit Anweisungen und Tabellen im Manual, Auswertungs- sowie Profildbogen standardisiert durchgeführt werden kann. Sie erfolgt im Vergleich zu einem theoretisch definierten Idealbereich bzw. zu kumulierten Häufigkeiten aus Normstichproben und wird durch Fallbeispiele und eine Broschüre für die eingeschätzte Führungskraft unterstützt.

Normierung

Die Normen für das Verfahren stammen aus einer Stichprobe von Führungskräften ($N = 242$), die zu ihrem eigenen Verhalten befragt wurden, und einer Stichprobe von Mitarbeitenden ($N = 513$), die zu ihrer Führungskraft Auskunft gaben. Beide Stichproben aus dem deutschen Sprachraum wurden mittels Online-Version des Fragebogens befragt. Problematisch ist, dass relativ wenige weibliche Führungskräfte ($N = 56$) und kaum weitere Diversitätsfaktoren berücksichtigt wurden. Ebenso sind die Stichproben für eine Normierung eher klein, was spezifischere (zum Beispiel branchenbezogene) Vergleichswerte nicht ermöglicht. Auch gibt es keine Hinweise, inwieweit unterschiedliche Hierarchieebenen vertreten sind. Bei der heterogenen zusammengesetzten Stichprobe der Mitarbeitenden wird von den Autoren darauf hingewiesen, dass diese

in Teilen abhängig von der Führungskräftestichprobe ist (da gemeinsam erhoben). Auch wird festgestellt, dass die Branchen Industrie und Handel deutlich unterrepräsentiert sind.

Zuverlässigkeit

Interne Konsistenzen der Dimensionen werden getrennt nach Version (LSA-F, Version für Führungskräfte: Selbsteinschätzung, und LSA-M, Version für Mitarbeitende: Fremdeinschätzung) berichtet. Die Koeffizienten des LSA-M sind mit einer Ausnahme als (sehr) gut einzuschätzen, während sich die Werte des LSA-F überwiegend im akzeptablen Bereich befinden (aber auch knapp darunter: .69 für „Individuelle Wertschätzung“ und „Charismatisches Verhalten“). Das Fehlen von Retest-Reliabilitäten wird durch die Veränderbarkeit des Führungsverhaltens erklärt. Interrater-Reliabilitäten für den LSA-M werden mittels rwg(j), ICC (1) und ICC (2) abgebildet, um die Übereinstimmung mehrerer Personen in Bezug auf eine Führungskraft abzubilden, die aber überwiegend moderat bis niedrig ausfällt. Die Autoren weisen darauf hin, dass hier die geringe Anzahl von Mitarbeitenden pro Führungskraft zu beachten ist. Die begrenzte Übereinstimmung erscheint mit Blick auf die LMX-Forschung (z. B. Schyns & Day, 2010), nach der Führungskräfte unterschiedliche Beziehungsqualitäten mit Mitarbeitenden haben, nachvollziehbar.

Gültigkeit

Mittels Pilotstudie und Expertenratings wurden Relevanz/Prototypikalität der Situationen und Repräsentativität der Antworten für die Dimensionen eingeschätzt (Inhaltsvalidität). Die Konstrukt- und Kriteriumsvalidität wurden nur für den LSA-M überprüft, so dass die Güte des LSA-F als unbekannt gelten muss. Zur Konstruktvalidität wurde eine konfirmatorische Faktorenanalyse herangezogen, die einen besseren Fit des Acht-Faktoren-Modells im Vergleich zum Generalfaktor-Modell bestätigte (hier fehlt leider der Vergleich zu einem dreidimensionalen Modell). Die Korrelationen mit dem Multifactor Leadership Questionnaire (MLQ) fallen mittel bis hoch aus. Allerdings liegen für vier der acht Skalen höhere Zusammenhänge mit einer anderen als der erwarteten MLQ-Skala vor. Die Kriteriumsvalidität wurde anhand der Aspekte Vertrauen und Loyalität geprüft, wobei die Auswahl der Variablen nicht genauer erläutert wird. Der LSA-M leistete hier allgemein eine etwas bessere Varianzaufklärung als der MLQ.

Weitere Gütekriterien

Das Verfahren ist ökonomisch, da die mittlere Bearbeitungszeit bei 20 bis 30 Minuten liegt. Es ist zu befürchten, dass ein gemeinsames Ausfüllen mehrerer Befragter oder

die Weitergabe des Fragebogens durch die Führungskraft Verfälschungstendenzen fördern können. Sinnvoll ist der Hinweis der Autoren, dass Ergebnisse von Mitarbeitenden nur dann ausgewertet werden sollen, wenn mindestens drei Einschätzungen vorliegen. Als problematisch ist aufgefallen, dass einige Aussagen sogenannte „double-barreled questions“ sind: Will die bzw. der Ausfüllende nur einem Teil zustimmen, ist nicht klar, wie die Antwort adäquat gegeben werden sollte. Auch teilweise auftretende doppelte Verneinungen sowie die Ähnlichkeit von Aussagen könnten dazu führen, dass Antwortende verwirrt werden und möglicherweise eine Antwort geben, die nicht ihrer tatsächlichen Erfahrung entspricht.

Abschlussbewertung/Empfehlung

Das Verfahren bietet eine theoretisch fundierte Erfassung unterschiedlicher Führungsstile, wobei die Einbettung in unterschiedliche Situationen als eigentliche Innovation zu werten ist. Insbesondere der Vergleich der Selbst- und Fremdeinschätzungen kann dabei für die Entwicklung von Führungskräften als sehr wertvoll eingeschätzt werden, gerade weil das Verfahren theoretisch hergeleitete Idealbereiche als Vergleichsbasis anbietet. Während die Objektivität des Verfahrens durch standardisierte Vorgaben und Verfahrensweisen weitgehend sichergestellt ist, besteht noch deutlicher Forschungsbedarf in Bezug auf die Reliabilität und insbesondere Validität des Verfahrens, gerade für die Version bezüglich der Selbsteinschätzung des LSA. Auch sollten die Normstichproben deutlich vergrößert werden, um spezifischere Vergleiche mit Teilstichproben zu ermöglichen.

Überlegenswert wäre es auch, ob neben den anzugebenden Informationen zur Person auf dem Deckblatt oder im Fragebogen nicht zusätzlich Aspekte zu erlebtem Zeitdruck und anderen Anforderungen und wahrgenommener Ressourcenausstattung (z. B. finanzielle Mittel) angegeben werden sollten. Solche Rahmenbedingungen der Führungsinteraktion würden es ermöglichen, die Antworten in Bezug auf das Führungsverhalten besser einordnen und bei Führungskräfte- bzw. Organisationsentwicklung effizienter intervenieren zu können.

Generell werden das Instrument und die Erläuterungen sowie die Online-Ausführung als gut und praktikabel eingeschätzt, gleichermaßen in Forschung und Praxis. Es handelt sich um einen vielversprechenden Ansatz, der aber in starkem Maße weiterer empirischer Fundierung bedarf.

Leadership Style Assessment (LSA; 1. Auflage)	Die TBS-TK-Anforderungen sind erfüllt			
	voll	weit- gehend	teil- weise	nicht
Allgemeine Informationen, Beschreibung und dia- gnostische Zielsetzung	X			
Objektivität		X		
Zuverlässigkeit		X		
Validität			X	

Diese Testrezension wurde im Auftrag des Testkuratoriums der Förderung deutscher Psychologinnenvereinigungen (DGPs und BDP) gemäß den TBS-TK-Richtlinien (Testkuratorium, 2009, 2010) erstellt.

Literatur

- Testkuratorium (2009). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Förderung Deutscher Psychologinnenvereinigungen. Revidierte Fassung vom 09. September 2009. *Report Psychologie* 34, 470–478.
- Testkuratorium (2010). TBS-TK. Testbeurteilungssystem des Testkuratoriums der Förderung Deutscher Psychologinnenvereinigungen. Revidierte Fassung vom 09. September 2009. *Psychologische Rundschau* 61, 52–56.
- Bass, B. M. & Riggio, R. E. (2006). *Transformational Leadership* (2nd edition). Mahwah: Lawrence Erlbaum Associates.

- Padilla, A., Hogan, R. & Kaiser, R. B. (2007). The toxic triangle: Destructive leaders, susceptible followers, and conducive environments. *The Leadership Quarterly*, 18, 176–194.
- Schyns, B. & Day, D. (2010). Critique and review of leader-member exchange theory: Issues of agreement, consensus, and excellence. *European Journal of Work and Organizational Psychology*, 19, 1–29.
- Schyns, B. & Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly*, 24, 138–158.

Testinformationen

Peus, C., Braun, S. & Frey D. (2015). Leadership Style Assessment. Ein Situational Judgment Test zur Erfassung von Führungsstilen (LSA, 1. Auflage).

Bezugsquelle: Testzentrale Göttingen, Herbert-Quandt-Straße 4, 37081 Göttingen. Test komplett 294 €. 5 Broschüren „Hinweise für Teilnehmende“ 68 €, 15 Fragebogen 47 €, 50 Auswertungsbogen für Führungskräfte 16 €, 50 Auswertungsbogen für Mitarbeitende 32 €, 50 Profilbogen 16 €.

Bitte zitieren Sie diesen Artikel wie folgt: Schilling, J. & Lippke, S. (2019). TBS-TK Rezension: „Leadership Style Assessment. Ein Situational Judgment Test zur Erfassung von Führungsstilen (LSA; 1. Auflage)“. *Psychologische Rundschau*, 70, 168–170.

<https://doi.org/10.1026/0033-3042/a000442>