

Macht toxisches Führungsverhalten krank? Eine systematische Übersicht und Meta-Analyse zum Zusammenhang von destruktiver Führung und der psychischen Gesundheit Beschäftigter

Philip NEGТ

*Institut für interdisziplinäre Arbeitswissenschaft,
Wirtschaftswissenschaftliche Fakultät, Leibniz Universität Hannover,
Schloßwender Straße 7, D-30159 Hannover*

Kurzfassung: Die Beziehung zwischen Führungskraft und Beschäftigten gilt als wichtiger Einflussfaktor auf psychisches Wohlergehen am Arbeitsplatz. Mit Begriffen wie Toxic Leadership oder Abusive Supervision ist betriebliche Führung zuletzt häufiger in ihren dysfunktionalen Anteilen beschrieben worden. Der Beitrag stellt die vorläufigen Ergebnisse eines systematischen Reviews zum Zusammenhang von destruktiven Führungsformen und psychischer Beschäftigtengesundheit dar. Für 35 empirische Primärstudien ergab sich eine gepoolte Effektstärke von $r = 0.31$ (CI₉₅ 0.26 – 0.37). Aufgrund der überwiegend querschnittlichen Studiendesigns sowie einiger Limitationen in der Erfassung destruktiver Führung erfolgt eine methodenkritische Diskussion des gegenwärtigen Forschungsfeldes.

Schlüsselwörter: Destruktive Führung, dysfunktionale Führungsbeziehungen, psychische Belastungen am Arbeitsplatz

1. Einführung und theoretischer Hintergrund

Führung wird seit jeher als maßgebliche Einflussgröße auf die Zufriedenheit, Leistungsbereitschaft und Motivation der Beschäftigten sowie den wirtschaftlichen Erfolg von Unternehmen betrachtet. Die empirische Leadership- und Führungsforschung ist dabei sowohl durch einen Fokus auf die Person der Führungskraft als auch für Organisationen und Beschäftigte günstiges Führungsverhalten geprägt (Thoroughgood et al. 2018). Mit der Einführung von Ansätzen wie Toxic Leadership (Lipman-Blumen 2005), Destructive Leadership (Einarsen et al. 2007) oder Abusive Supervision (Tepper 2007) ist das Phänomen Führung in der jüngeren Vergangenheit vermehrt in seinen negativen, schädigenden und dysfunktionalen Anteilen untersucht worden. Trotz einiger Gemeinsamkeiten der Ansätze, erfolgen Definition und Operationalisierung bisher nicht vollständig einheitlich. Frühe epidemiologische Untersuchungen deuten auf eine hohe Prävalenz in der Arbeitswelt hin (Aasland et al. 2010).

Psychische Belastungen, Fehlbeanspruchungen sowie das Auftreten psychischer Störungen im Kontext der Erwerbsarbeit stellen Betroffene und Unternehmen gleichermaßen vor Herausforderungen. Für den deutschen Sprachraum zeigt sich eine seit den 1990er-Jahren kontinuierlich ansteigende Anzahl an Arbeitsunfähigkeitstagen aufgrund psychischer Störungen (Bär & Tophoven 2010). Die Ursachen können als multifaktoriell angenommen werden und umfassen neben einer Entstigmatisierung

psychischer Störungen sowie einer erhöhten diagnostischen Sensitivität im Versorgungssystem möglicherweise auch Merkmale der modernen Arbeitswelt, darunter beschleunigte Arbeitsprozesse oder die Zunahme von Dienstleistungstätigkeiten (Jacobi & Kessler-Scheil 2013). Unter Berücksichtigung der individuellen Beeinträchtigung und Funktionseinschränkung Betroffener sowie der potenziellen Krankheitskosten für Unternehmen kommt einer Untersuchung möglicher Prädiktoren für psychisches Belastungserleben am Arbeitsplatz hohe Aktualität und Relevanz zu. In diesem Zusammenhang sind Führung und Führungsbeziehungen als Einflussfaktoren erst in den vergangenen Jahren systematischer theoretisiert und einer empirischen Untersuchung zugänglich gemacht worden.

Im vorliegenden Beitrag werden die vorläufigen Ergebnisse eines systematischen Reviews zum Zusammenhang von destruktiven Führungsformen und der psychischen Gesundheit Beschäftigter dargestellt. Ergänzend zu einer Systematisierung von Studienmerkmalen und der Ermittlung aggregierter Effektstärken sollen auch mögliche methodische Verzerrungen innerhalb des Forschungsfeldes überprüft werden. Folgende Fragestellungen sind tiefergehend zu untersuchen: 1) Welche empirischen Indikatoren für destruktive Führung sowie psychisches Belastungserleben der Beschäftigten werden zugrunde gelegt? 2) Wie ist die Beschaffenheit (z. B. Studiendesign) und Qualität (z. B. Kontrolle konfundierender Effekte) der identifizierten Primärstudien zu bewerten, insbesondere hinsichtlich der Möglichkeit kausaler Inferenz? 3) Wie sind die Ergebnisbefunde einzuordnen, wenn Effektstärken über Einzelbefunde hinweg gepoolt und potenzielle Verzerrungen berücksichtigt werden?

2. Methodisches Vorgehen

Orientiert an den PRISMA-Richtlinien zur Anfertigung systematischer Reviews (Moher et al. 2009) wurde zunächst eine Liste an Suchbegriffen generiert, um anschließend in vier wissenschaftlichen Literaturdatenbanken (Scopus, Social Sciences Citation Index, PubMed, PsycInfo) nach empirischen Primärstudien zu suchen. Eine Übereinstimmung mit folgenden Einschlusskriterien wurde vorausgesetzt: 1) Bezug zu einem spezifischen Modell destruktiver Führung, für das empirische Erhebungsinventare vorliegen, 2) Ermittlung von psychischer Belastung über psychometrisch überprüfte Inventare, 3) Querschnittliches-, prospektives- oder experimentelles Studiendesign, 4) Untersuchte Proband*innen sind erwerbstätig, 5) Numerische Daten werden hinreichend ausführlich berichtet, sodass Effektstärken verwendet oder berechnet werden können, 5) Veröffentlichung in englischer oder deutscher Sprache.

Es erfolgte eine Codierung der Primärstudien in Bezug auf verschiedene Studienmerkmale (u.a. Ansatz destruktiver Führung, Form psychischer Belastung, Studiendesign, Region, berufliches Tätigkeitsfeld). Als Assoziations- und Effektstärkemaß wurde der Korrelationskoeffizient festgelegt. Um einen einheitlichen Effektstärkeindex über die eingeschlossenen Studien hinweg zu gewährleisten, wurde in einzelnen Fällen eine Transformation vorgenommen (Regressionskoeffizient- $\beta \rightarrow$ Korrelationskoeffizient) (Peterson & Brown 2005). Zur Berechnung gepoolter Effektstärken wurde ein Random-Effects Modell verwendet. Die Durchführung einer Subgruppenanalyse diente der Untersuchung möglicher Faktoren, welche mit Heterogenität der Befunde in Zusammenhang stehen könnten (z.B. Studiendesign). Neben der Analyse verzerrender Effekte auf Ebene der jeweiligen Primärstudien wurde auch

der Einfluss von Publication Bias (Duval & Tweedie 2000) auf die gepoolten Effektstärken ermittelt.

3. Ergebnisse

Abbildung 1 stellt die wesentlichen Ergebnisse des systematischen Reviews in einer zusammenfassenden Übersicht dar.

3.1 Deskriptive Merkmale der eingeschlossenen Primärstudien

Insgesamt konnten 35 Primärstudien identifiziert werden, die mit den definierten Einschlusskriterien übereinstimmten. Einzelne Veröffentlichungen führten Ergebnisse verschiedener Studien auf, sodass voneinander unabhängige Effektstärken berechnet werden konnten. 10 der Studien stammten aus Europa, 14 aus Nordamerika, einschließlich Kanada, und 11 aus dem asiatischen Raum. Der Großteil der Primärstudien ($k = 25$) nutzte den Ansatz der Abusive Supervision zur Operationalisierung destruktiver Führung. In 6 weiteren Primärstudien wurden andere Modelle destruktiver Führung (z. B. Petty Tyranny oder Destructive Leadership) zugrunde gelegt. Einige Untersuchungen verwendeten Inventare, die sich auf unterschiedliche Modelle bezogen ($k = 4$). Während 26 Primärstudien eine ausschließlich querschnittliche Datenerhebung durchführten, waren 6 weitere Studien prospektiv angelegt, jedoch häufig mit kurzen Erhebungsintervallen (2 Wochen bis 12 Monate) und ohne Kontrolle der abhängigen Variable zum initialen Zeitpunkt. Lediglich in 3 Primärstudien wurden experimentelle Vorgehensweisen ausgeführt, in der Regel durch Darstellung von Fallvignetten. Hinsichtlich des beruflichen Tätigkeitsfeldes basierten 12 der eingeschlossenen Studien auf Proband*innen-Stichproben aus umgrenzten Branchen (z. B. Versicherungsbranche, Krankenpflege, High-Tech Industrie), wohingegen die Mehrzahl der allgemeinen Erwerbsbevölkerung entsprach.

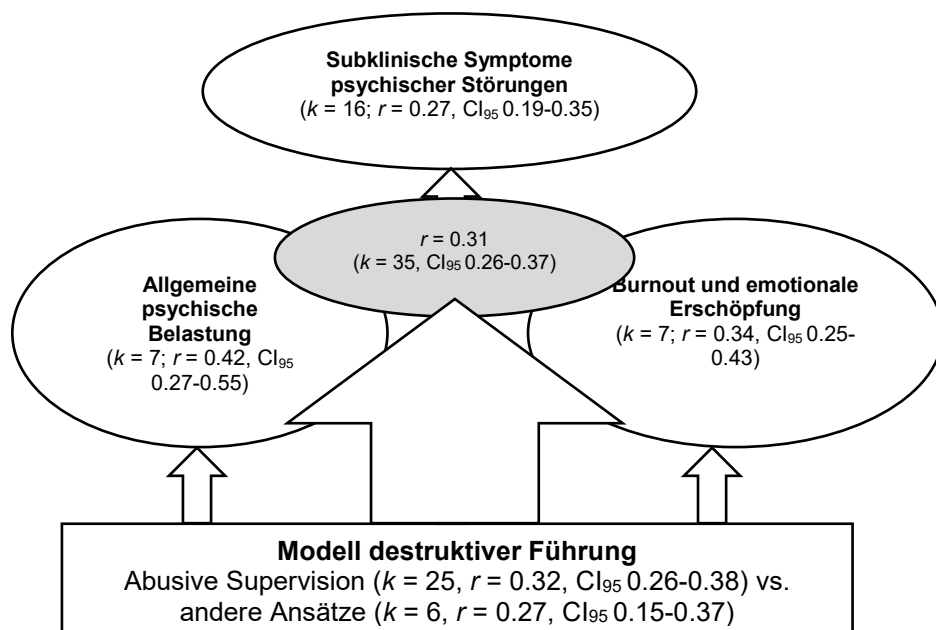


Abbildung 1: Zusammenfassende Ergebnisübersicht. Neben der Anzahl jeweils einbezogener Primärstudien werden gepoolte Effektstärken und deren Konfidenzintervalle (95%)

Bezogen auf den Gesamtstudienpool und die zugrunde gelegten empirischen Indikatoren für psychisches Belastungserleben ließen sich drei unterschiedliche Cluster bilden: 1) Subklinische Symptome psychischer Störungen, darunter Depression, Angst, paranoides Erleben, welche mittels klinisch gebräuchlicher Skalen erfasst wurden. 2) Burn-out und emotionale Erschöpfung, jeweils ermittelt über psychometrisch überprüfte Inventare, wie das *Maslach Burnout Inventory*. 3) Allgemeine psychische Belastung und Indikatoren für erhöhtes Stresserleben (z. B. Schlafstörungen), überwiegend erfasst über validierte Selbstbefragungsinstrumente.

3.2 Aggregation von Effektstärkemaßen und meta-analytische Ergebnisbefunde

Für den Gesamtstudienpool ergab sich eine Effektstärke von $r = 0.31$ ($CI_{95} 0.26 - 0.37$) für den Zusammenhang zwischen erlebter destruktiver Führung und den verschiedenen Indikatoren für psychische Belastung. Die numerisch höchsten aggregierten Effektgrößen konnten für den Pool an Primärstudien ermittelt werden, in denen ausschließlich allgemeine psychische Belastung als Outcome betrachtet wurde ($k = 7$, $r = 0.42$, $CI_{95} 0.27 - 0.55$). Die gemittelten Effektstärken für die Studienpools, in denen entweder subklinische Symptome psychischer Störungen ($k = 16$, $r = 0.27$, $CI_{95} 0.19 - 0.35$) oder Burn-out und emotionale Erschöpfung ($k = 7$, $r = 0.34$, $CI_{95} 0.25 - 0.43$) erfasst wurden, fielen in der deskriptiven Analyse kleiner aus. Um Abhängigkeiten von Effektstärken zu reduzieren, wurden in die Aggregation nach dem Indikator psychischer Belastung nur Studien einbezogen, welche sich auf einen Outcome beschränkten (z. B. Burn-out und emotionale Erschöpfung). Diejenigen Primärstudien, welche sich lediglich auf den Ansatz der Abusive Supervision bezogen ($k = 25$, $r = 0.32$, $CI_{95} 0.26-0.38$), erreichten ähnlich große Effektstärken, wie Primärstudien, in denen andere Modelle zur Operationalisierung destruktiver Führung herangezogen wurden, z. B. Petty Tyranny oder Toxic Leadership ($k = 6$, $r = 0.27$, $CI_{95} 0.15-0.37$). Auch hier wurden im Sinne einer konservativen Schätzung nur Primärstudien einbezogen, die entweder Inventare zur Erfassung von Abusive Supervision oder aber alternativer Modelle (z. B. Petty Tyranny) verwendeten.

3.3 Subgruppenanalyse und Verzerrungspotenzial

Die Ergebnisse der Subgruppenanalyse lieferten keine substanziellen Hinweise dafür, dass die untersuchten kategorialen Variablen mit den Effektstärken assoziiert sind. Für das spezifische Modell destruktiver Führung (Abusive Supervision vs. andere Ansätze; $Q_b(1) = 0.54$, $p = 0.46$), den Indikator für psychisches Belastungserleben (subklinische Symptomausprägung vs. Burn-out vs. allgemeine psychische Belastung; $Q_b(2) = 4.39$, $p = 0.11$), das Studiendesign (querschnittlich vs. prospektiv vs. experimentell; $Q_b(2) = 0.77$, $p = 0.68$) oder den kulturellen Arbeitskontext (Nordamerika vs. Europa vs. Asien; $Q_b(2) = 2.65$, $p = 0.26$) unterschieden sich die in den Subgruppen gepoolten Effektstärken nicht signifikant voneinander. Aufgrund der Tatsache, dass ein Großteil der eingeschlossenen Primärstudien einem querschnittlichen Studiendesign mit nur einem Indexzeitpunkt der Erhebung folgt ($k = 26$), sich darüber hinaus die Operationalisierung destruktiver Führung häufig auf das subjektive Erleben einzelner Beschäftigter beschränkt, ist das Verzerrungspotenzial insgesamt als höher einzuschätzen. Den Ergebnissen einer Trim-and-Fill Analyse (vgl. Abbildung 2) zufolge ergaben sich zudem Hinweise für Publikationsverzerrungen.

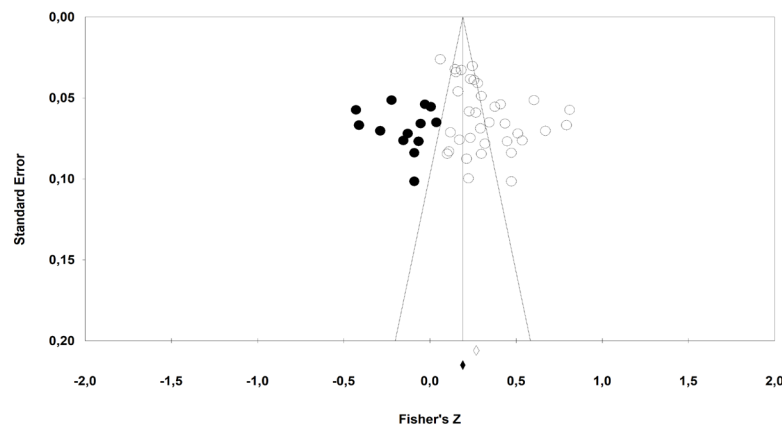


Abbildung 2: Dargestellt wird ein Funnel Plot, in dem die Effektstärken (Fisher's Z transformierte Korrelationen), die Präzision der jeweiligen Primärstudie (SE) sowie gepoolte Effektgrößen abgebildet sind. Ausgefüllte Punkte repräsentieren geschätzte Effektstärken aus potenziell unveröffentlichten Studien.

Demnach könnten 13 Effektstärken eine Überrepräsentation positiver Ergebnisbefunde darstellen. Eine annäherungsweise Adjustierung der ursprünglich gepoolten Effektstärke ($r = 0.31$, $CI_{95} 0.26 - 0.37$) ist mit einer Verringerung des mittleren Zusammenhangs auf $r = 0.20$ ($CI_{95} 0.13 - 0.26$) verbunden.

4. Diskussion

Die dargestellten Ergebnisse deuten zusammengefasst auf einen kleinen bis mittleren korrelativen Zusammenhang zwischen Indikatoren für das Erleben destruktiver Führung und psychischer Belastung der Beschäftigten hin. Das gegenwärtige Forschungsfeld ist jedoch durch Einschränkungen charakterisiert, welche die Ableitung robuster Schlussfolgerungen erschweren. Dies betrifft insbesondere eine häufig nur unzureichende Berücksichtigung möglicher methodischer Artefakte. Mit Begriffen wie Common Method Bias oder Single Source Bias (Podsakoff et al. 2003; Podsakoff et al. 2012) sind in der Vergangenheit methodische Verzerrungen in den Sozial- und Verhaltenswissenschaften beschrieben worden. Der überwiegende Anteil bisheriger Primärstudien ist so aufgebaut, dass sowohl destruktive Führung als beeinflussende Variable als auch psychische Belastung als angenommene Wirkung aus derselben methodischen Quelle stammen, d.h. der subjektiven Evaluation der Beschäftigten. Hieraus ergeben sich einige Limitationen. So findet sich inzwischen eine breite empirische Studienbasis, die darauf hinweist, dass negative Affektivität als Disposition sowohl mit psychischer Belastung als auch negativen Interpretationsmustern in Bezug auf den Erwerbsarbeitskontext zusammenhängt (Falco et al. 2013; Stanton & Watson 2014). Darüber hinaus ist unter dem Begriff des Hostile Interpretation Bias in vergangenen Untersuchungen wiederholt gezeigt worden, dass psychisch belastete Personen mehrdeutige soziale Interaktionen häufiger als feindselig interpretieren (Smith et al. 2016), ohne dass die Interaktion selbst Ursache der Belastung sein muss. Die fehlende Kontrolle für derartig konfundierende Variablen erschwert die Interpretation der Befunde. Sofern es sich um Querschnittstudien mit nur einem Indexzeitpunkt der Erhebung handelt, gilt dies umso mehr.

Um dem Einwand einer methodisch artifiziellen Forschungslogik künftig begegnen zu können, erscheint eine Weiterentwicklung des Forschungsfeldes von Bedeutung.

Neben der Berücksichtigung einer stärker systemischen Perspektive auf das Phänomen destruktive Führung und einer damit verbundenen Auflösung des Leader-Zentrismus (Thoroughgood et al. 2018) könnten insbesondere methodische Triangulationen sowie Mixed-Method Designs eine gewinnbringende Ergänzung darstellen. Innerhalb der derzeit vorherrschenden Studienlandschaft fehlen zudem aussagekräftige prospektive Kohortenstudien, in denen die Erfahrung destruktiven Führungsverhaltens bei psychisch möglichst unbelasteten Individuen als Prädiktor fungiert, um psychische Belastungen zu einem späteren Zeitpunkt vorherzusagen. Trotz des hohen Aufwands solcher Studiendesigns sind sie als „Goldstandard“ zur Untersuchung vermuteter Wirkrichtungen anzusehen. Vor dem Hintergrund der aufgeführten methodischen Einschränkungen können die im vorliegenden Review ermittelten gepoolten Effektstärken als eher gering bewertet werden.

Interpersonelles Erleben und Verhalten am Arbeitsplatz ist zweifellos wichtiger Einflussfaktor auf psychisches Wohlergehen. Führung und Führungsbeziehungen kommt in diesem Zusammenhang hervorgehobene Bedeutung zu. Angesichts der hohen Relevanz des Themas sowohl für Wissenschaft als auch Praxisfelder sollten Limitationen im derzeitigen Forschungsfeld in künftigen Untersuchungen eine stärkere Berücksichtigung erfahren. Dies stellt eine wichtige Voraussetzung zur Ableitung zielgerichteter Präventionsmaßnahmen in Organisationen und Unternehmen dar.

5. Literaturverzeichnis

- Aasland MS, Skogstad A, Notelaers G, Birkeland N, Einarsen S (2010) The prevalence of destructive leadership behaviour. *British Journal of Management* 21:438–452.
- Bär T, Tophoven C (2010) Psychische Erkrankungen im Fokus der Krankenkassen – eine Analyse der Gesundheitsreporte seit 2000. *Gesundheits- und Sozialpolitik* 64(2): 38–48.
- Duval S, Tweedie R (2000) A non-parametric “trim and fill” method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association* 95(449): 89–98.
- Einarsen S, Aasland MS, Skogstad A (2007) Destructive leadership behavior: A definition and conceptual model. *The Leadership Quarterly* 18: 207–216.
- Falco A, Girardi D, Marcuzzo G, De Carlo A, Bartolucci GB (2013) Work stress and negative affectivity: A multi-method study. *Occupational Medicine* 63: 341–347.
- Jacobi F, Kessler-Scheil S (2013) Epidemiologie psychischer Störungen: Häufigkeit und Krankheitslast in Deutschland. *Psychotherapeut* 58: 191–206.
- Lipman-Blumen J (2005) Toxic leadership: When grand illusions masquerade as noble visions. *Leader to Leader* 36: 29–36.
- Moher D, Liberati A, Tezlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Open Medicine* 3(2): 123–130.
- Peterson PL, Brown SP (2005) On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology* 90(1): 175–181.
- Podsakoff PM, MacKenzie, SB, Lee JY, Podsakoff NP (2003) Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88(5): 879–903.
- Podsakoff PM, MacKenzie, SB, Podsakoff NP (2012) Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology* 63:539–569.
- Smith HL, Summers, BJ, Dillon KH, Macatee, RJ, Cogle, JR (2016) Hostile interpretation bias in depression. *Journal of Affective Disorders* 203: 9–13.
- Stanton K, Watson D (2014) Positive and negative affective dysfunction in psychopathology. *Social and Personality Psychology Compass* 8/9: 555–567.
- Tepper BJ (2007) Abusive supervision in work organizations: Review, synthesis, and research agenda. *Journal of Management* 33(3): 261–289.
- Thoroughgood CN, Sawyer KB, Padilla A, Lunsford L (2018) Destructive leadership: A critique of leader-centric perspectives and toward a more holistic definition. *Journal of Business Ethics* 151:627–649.



Gesellschaft für Arbeitswissenschaft e.V.

Nachhaltig Arbeiten und Lernen

**Analyse und Gestaltung lernförderlicher
und nachhaltiger Arbeitssysteme
und Arbeits- und Lernprozesse**

69. Kongress der
Gesellschaft für Arbeitswissenschaft e.V.

Gottfried Wilhelm Leibniz Universität Hannover

01. – 03. März 2023

GfA-Press

Bericht zum 69. Arbeitswissenschaftlichen Kongress vom 01. – 03. März 2023

**Fakultät Maschinenbau, Institut für Berufswissenschaften der Metalltechnik (IBM) und
Institut für Fabrikanlagen und Logistik (IFA), Leibniz Universität Hannover**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Sankt Augustin: GfA-Press, 2023
ISBN 978-3-936804-32-4

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle (s. u.) erhältlich.

Alle Rechte vorbehalten.

© GfA-Press, Sankt Augustin

Schriftleitung: Prof. Dr. Rolf Ellegast

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet:

- den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen,
- den Kongressband oder Teile daraus in Print- und/oder Nonprint-Medien (Webseiten, Blog, Social Media) zu verbreiten.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

Geschäftsstelle der GfA

Simone John, Tel.: +49 (0)30 1300-13003

Alte Heerstraße 111, D-53757 Sankt Augustin

info@gesellschaft-fuer-arbeitswissenschaft.de · www.gesellschaft-fuer-arbeitswissenschaft.de

Screen design und Umsetzung

© 2023 fröse multimedia, Frank Fröse

office@internetkundenservice.de · www.internetkundenservice.de