

Exploration von Motion-Capture zur Replikation von Erwerbsarbeit für die betriebliche Kompetenzförderung

Philip Hermann SCHÄFER¹, Marvin GOPPOLD¹, Oliver BRUNNER¹,
Christopher BRANDL^{1,2}, Verena NITSCH^{1,2}

¹ *Institut für Arbeitswissenschaft, RWTH Aachen University,
Eilfschornsteinstraße 18, D-52062 Aachen*

² *Forschungsabteilung Produkt- und Prozessergonomie,
Fraunhofer-Institut für Kommunikation,
Informationsverarbeitung und Ergonomie FKIE,
Campus-Boulevard 55, D-52074 Aachen*

Kurzfassung: Der Trend der Digitalisierung gibt immer wieder neue Impulse für die berufliche Bildung. In dieser Machbarkeitsstudie wird anhand eines Arbeitsprozesses aus der Metallbearbeitung untersucht, welche Grenzen und Möglichkeiten eine gegenwärtige Motion Capture (MoCap) Lösung für digitale Anwendungen in der beruflichen Bildung aufweist. Dazu wurden die Bewegungen eines Arbeiters bei der Ablängung eines Werkstückes mittels MoCap aufgezeichnet und in einem Experiment qualitativ mit einem Video des Arbeitsprozesses verglichen. Die Ergebnisse deuten darauf hin, dass bei der reinen Betrachtung von Bewegungsdaten (im Vergleich zu Videodaten) ein für Lernprozesse relevanter Informationsverlust einhergeht. Anhand der Studienergebnisse werden Implikationen für den Einsatz von MoCap in der beruflichen Bildung diskutiert.

Schlüsselwörter: Berufliche Bildung, Digitalisierung, Motion Capture

1. Einleitung

Eine aktuelle Herausforderung der beruflichen Bildung liegt in der Gestaltung digitaler Bildungsprozesse (Windelband 2018) und der Eingliederung von digitalen Menschmodellen für die lernmethodische Nutzung (z. B. Rueckert et al. 2020; Björkenstam et al. 2020). Angefangen bei niederschwelligen betrieblichen Unterweisungskonzepten (z. B. REFA Bundesverband e.V. 2001; Schelten 2005) bis hin zur arbeitsprozessorientierten Ausbildung (Becker 2013) mit hohem Bildungsanspruch existieren eine Vielzahl von Maßnahmen, welche von digitalen Abbildungen der Realität profitieren könnten. Mittels Virtual Reality (VR) können Lernräume geschaffen werden, welche digitale Abbilder erlebbar machen, und die Möglichkeiten des physischen Gegenstücks übersteigen. Somit könnten Auszubildende einen digitalisierten Arbeitsprozess in VR aus unterschiedlichen Betrachtungswinkeln analysieren und Abschnitte nach ihren Bedürfnissen wiederholen oder in veränderter Geschwindigkeit abspielen. Im Vergleich zu video-basierten Lernprozessen, ermöglicht VR potenziell die Adressierung höherer kognitive Ebenen und Erzielung größerer Lerneffekte (Queiroz et al. 2022).

Um unverfälschte Lerninhalte zu generieren, darf die Aufnahmetechnik diesen

Arbeitsprozess nicht beeinflussen. Deshalb eignen sich insbesondere optisch markerlose Motion Capture (MoCap) Systeme. Hiermit geht jedoch in den meisten Fällen eine Informationsreduktion einher, da lediglich die Arbeitsperson aufgezeichnet wird und z. B. Bewegungsdaten von Arbeitsmaterial, Zuständen von Maschinen, aber auch Informationen über Fingerbewegungen verloren gehen. Die vorliegende Arbeit nutzt ein solches MoCap System (Captury Live), um einen Arbeitsprozess aus einem Reallabor aufzunehmen und anhand der erfassten Motion Tracking Daten der Arbeitsperson zu untersuchen, inwiefern ausgeführte Handlungen, Operationen und Grundbewegungen einer Arbeitsperson sowie deren Ausgangs-, Zwischen- und Zielzustände anhand der aufgezeichneten Bewegungsdaten subjektiv klassifiziert werden können. Dabei wird die Handlungsregulationstheorie (Jenewein 2010; Volpert 1983) als Referenztheorie in der beruflichen Bildung zugrunde gelegt. Ziel dieser Machbarkeitsstudie ist es, bereits vor der Implementierung eines digitalen Lernsystems einordnen zu können, welche Möglichkeiten und Limitationen sich auf Basis der aufgezeichneten Bewegungsdaten einer Arbeitsperson für eine digitale Lernumgebung ergeben.

2. Vorgehen

Tabelle 1: Ausschnitt des Arbeitsplans für den betrachteten Teil des Fertigungsablaufs

Arbeitsschritt	Vorgehensweise	Werkzeug/Arbeitsschutz	Material
Material bereitstellen	Profil aus Regal entnehmen		Boschprofil
Anzeichnen	Profil auflegen, Anreißermaß einstellen, anreißen	Reißnadel, Streichmaß	Boschprofil
Material ablängen	Profil einspannen, Absägen	Kappsäge, Gehörschutz, Schutzbrille	Boschprofil

In der Werkstattumgebung eines Reallabors wurde eine Arbeitsperson beim Ablängen eines Boschprofils nach einem vorgegebenen Arbeitsplan (Tabelle 1) von einem MoCap System aufgenommen. Dieser umfasste drei Arbeitsschritte: „Material bereitstellen“, „Anzeichnen“ und „Material ablängen“, welche sich aus Teilhandlungen mit Operationen und Grundbewegungen zusammensetzen (Tabelle 2).

Das eingesetzte MoCap System „Captury Live v152“ nutzte im Studienaufbau acht Kameras des Typs FLIR Blackfly S (1280x720@60fps). Es wurden sowohl Video- als auch Bewegungsdaten der Arbeitsperson erfasst. Veränderte Positionen von Werkzeugen, Material und weiteren Objekten (Tabelle 1) sind nur in den Videos (Abbildung 1, links) erkennbar. Verglichen werden die Kodierungsübereinstimmung zwischen einem Video (Abbildung 1, links), welches die Arbeitsperson inklusive des überlagerten Skelettmodells (V), und einem Video (Abbildung 1, Mitte), welches nur ein Skelettmodell (S) über einem statischen Hintergrund bei der Arbeitsaufgabe zeigt.

Tabelle 2: Kategorien zur Codierung der Videosequenz

Kategorie	Bezeichnung
Handlung	Materialbereitstellung und Anzeichnen
Operation	Anreißen des Profils
Grundbewegung (MTM-1)	Greifen der Reißnadel
Zustand Arbeitsobjekt	Profil [abgelängt / nicht abgelängt]
Zustand Arbeitsmittel	Säge [in Betrieb / nicht in Betrieb]
Zustand Arbeitsperson	Arbeitsperson [PSA tragend / nicht tragend]

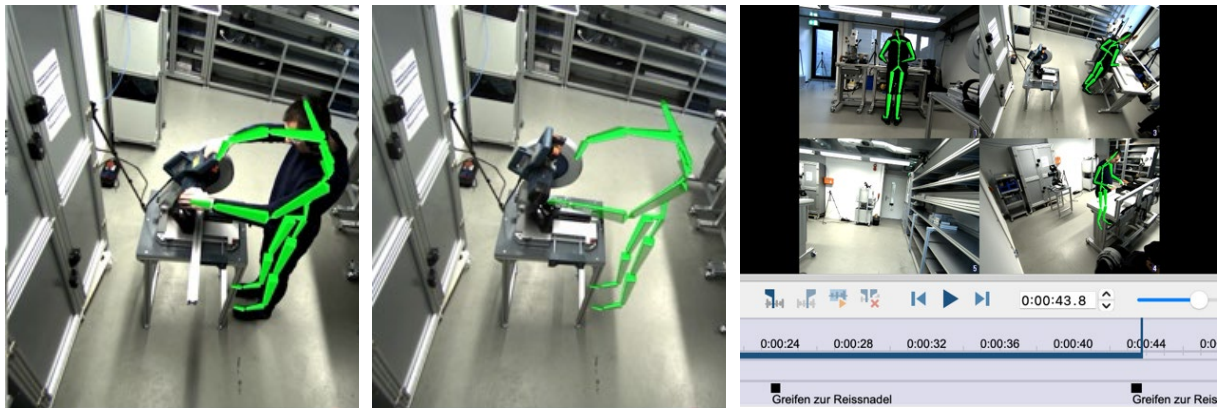


Abbildung 1: *Arbeitsperson inklusive überlagertem Skelettmodells (links); Skelettmodell (S) über statischen Hintergrund (Mitte); Ansicht des Videomaterials in MAXQDA (rechts)*

In beiden Videosequenzen wird die Arbeitsperson bzw. das Skelettmodell aus vier Perspektiven gezeigt (Abbildung 1, rechts). Das Skelettmodell ist in beiden Videosequenzen identisch und besitzt keine Informationen zu Fingerbewegungen. In (S) sind keine Bewegungen der Arbeitsmittel und der Arbeitsobjekte erkennbar, da es sich um einen statischen Hintergrund handelt.

3. Versuchsdurchführung

Um zu erfahren, inwieweit Handlungen, Operationen und Grundbewegungen als auch deren Ausgangs-, Zwischen- und Zielzustände anhand der aufgenommenen Bewegungsdaten in beiden Videos subjektiv klassifiziert werden, wurde das Probandenkollektiv (n=16, 11m, 5w) randomisiert in zwei Gruppen von jeweils acht Personen unterteilt. Als Methode wurde eine qualitative Inhaltsanalyse (vgl. Mayring 2015) basierend auf den Aufnahmen des Arbeitsprozesses von den Probanden ausgeführt. Die Teilnehmenden bekamen entweder die Videoaufzeichnung (Gruppe V) der Arbeitsperson inklusive überlagertem Skelettmodell; oder (Gruppe S) die Bewegungen des Skelettmodells vor einem statischen Hintergrund vorgelegt. Die Stichprobe besitzt geringe fachliche Erfahrung im Gegenstandsbereich Metallbearbeitung (ohne Ausbildungs- und Berufserfahrung) und in der Arbeitsanalyse (keine bis gering) sowie MTM-1 Nutzung (keine bis gering). Sie wurden gebeten, das vorgelegte Videomaterial mit der Kodierungssoftware MAXQDA v20 (Abbildung 1, rechts) zu sichten und zu kodieren. Hierbei kodierten sie Start sowie Endzeitpunkte erkannter Handlungen, Operationen und Grundbewegungen der Arbeitsperson und Zustände von Arbeitsobjekt, Arbeitsmittel und Arbeitsperson (Tabelle 2) in der Videosequenz. Die Kategorien wurden nicht weiter detailliert, als dies in Lerninhalten von Ausbildungen üblich ist. Neben dem Videomaterial standen den Teilnehmenden die technische Zeichnung des zu fertigenden Werkstückes, der Arbeitsplan zur Fertigung, Fotos der Werkstattarbeitsplätze, der Bewegungszyklus von MTM-1 Grundbewegungen sowie ein Ankerbeispiel der Handlungsregulationstheorie für die Kodierung zur Verfügung. Nach der Begrüßung und der standardisierten Erhebung demographischer Daten und Vorerfahrungen wurde das variierte Videomaterial vom Probandenkollektiv kodiert. Die Versuchsleitung stand für technische Rückfragen zur Softwarehandhabung zur Verfügung.

4. Ergebnisse

Die Ergebnisse zeigen in Form von Heatmaps (Abbildung 2) die Häufigkeitsverteilung der codierten Segmente über die Zeit für alle Teilnehmenden in beiden Gruppen. Eine dunklere Einfärbung signalisiert eine stärkere Übereinstimmung innerhalb einer Gruppe. Insgesamt zeigt sich, dass sich die eingeschätzten Dauern der codierten Abschnitte sowohl innerhalb der Gruppen als auch zwischen den Gruppen Vollinformation (V) und Skelettmodell (S) unterscheiden.

Abbildung 2 zeigt für „Rohmaterial bereitstellen und Anzeichnen“ insgesamt eine starke Streuung in beiden Gruppen. Dabei fällt auf, dass bei vollen Informationen zum Arbeitsprozess die Varianz der Streuung in Gruppe V größer als in Gruppe S ist. Beim „Anreißen des Profils“ ist ebenfalls eine starke Streuung zu verzeichnen, wobei Gruppe V einen deutlich abweichenden Bereich für diese Operation kodiert. Fast alle der Teilnehmenden, denen ausschließlich Skelettdaten zur Verfügung standen, beschreiben die Dauer des Anreißens über einen deutlich längeren Zeitraum.

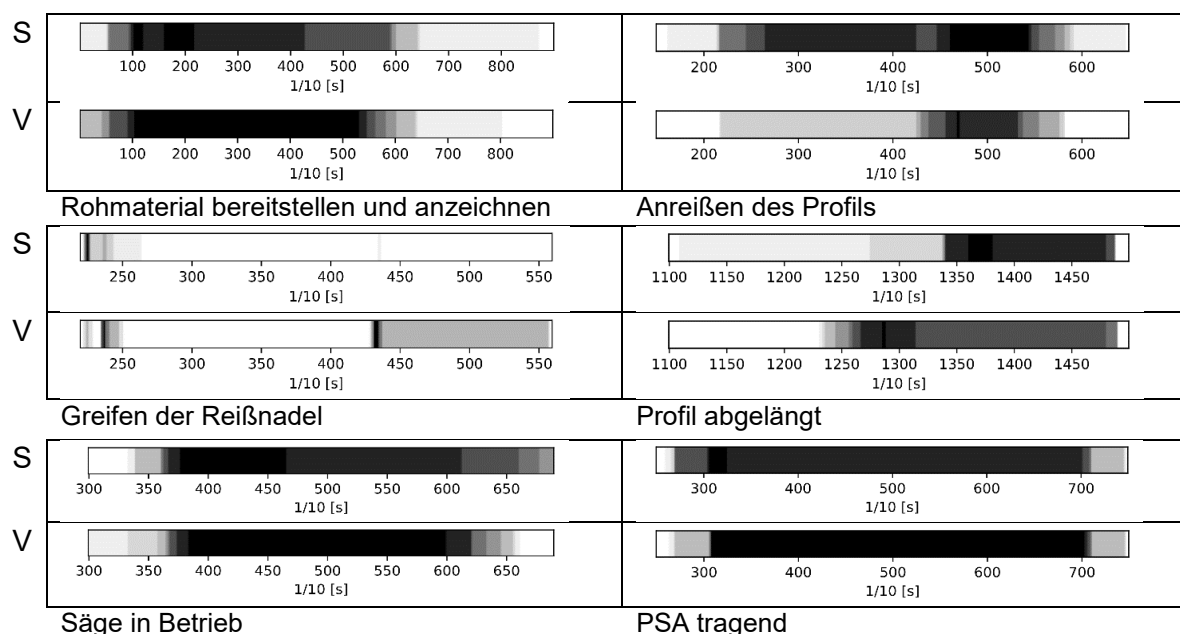


Abbildung 2: Häufigkeitsverteilung in den Gruppen Vollinformation (V) und Skelettmodell (S). Farbverlauf von weiß nach schwarz indiziert steigende Datendichte.

Insgesamt sind die Abweichungen bei der Grundbewegung „Greifen der Reißnadel“ innerhalb beider Gruppen deutlich geringer als bei Operation und Handlung. Allerdings ist zu beachten, dass bei Grundbewegungen nach MTM-1 Standard auch kleinste Abweichungen eine erhebliche Auswirkung besitzen. In Gruppe S wurde ein zweites Greifen fast gar nicht erkannt, wohingegen in Gruppe V der zweite Greifvorgang mit unterschiedlicher Dauer kodiert wurde. Bei der Zustandsbeschreibung „Profil abgelängt“ liegen offensichtlich unterschiedliche Interpretationen des beeinflussenden Trennprozesses vor.

Bei „Säge in Betrieb“ als sicherheitskritischer Zustand weisen beide Gruppen deutliche Abweichungen auf. Bei der Zustandsangabe zum „Tragen von PSA“ (Persönliche Schutzausrüstung) zeigt sich in beiden Gruppen eine homogene Verteilung und eine Definitionsdifferenz auf. In beiden Gruppen wird sowohl nur die vollständige PSA als auch bereits ein Teil der PSA tragend kodiert.

5. Diskussion und Fazit

Die Klassifizierung beruflicher Handlungen zeigt Unterschiede in der Interpretation der Start- und Endzeitpunkte in beiden Gruppen. Die Reduktion von Informationen durch MoCap führt zu einer geringeren Varianz für einen logischen Start- und Endzeitpunkt. Dies kann bei der Gestaltung von Lernprozessen berücksichtigt werden, sodass MoCap z. B. zur Bildung mentaler theoretischer Modelle auf Handlungsebene geeignet sein könnte, da eine auf die wesentlichen Elemente reduzierte Informationslage entsteht. Ein höherer Erfahrungswert der Versuchsteilnehmenden oder die Vorgabe eindeutiger Definitionen der beruflichen Handlungen könnten den Interpretationsspielraum ebenfalls reduzieren.

Auch bei der Klassifizierung der Operationen konnten Start- und Endzeitpunkte nicht eindeutig identifiziert werden. Hier sind allerdings bei Vollinformation die Unterschiede geringer als beim Skelettmodell, was sehr wahrscheinlich mit fehlenden Informationen zu Bewegungen in Gruppe S zusammenhängt, da wichtige Informationen, wie z. B. welches Werkzeug gegriffen wird, nicht unterschieden werden kann. Selbst ein höherer Erfahrungswert wird vermutlich diesen Informationsmangel nicht vollständig kompensieren, weshalb Operationen entweder in der Realität oder in VR mit einem zum Lernprozess passend detaillierten Menschmodell betrachtet werden sollten (z. B. inklusive Handmodell).

Beim genauen Vergleich der Grundbewegungen zwischen den Gruppen wird ersichtlich, dass die Dauer der ersten Greifbewegung gruppenübergreifend homogen erkannt wird, obwohl im Skelettmodell Informationen über Fingerpositionen sowie über Positionen der Werkzeuge fehlen. Lediglich Start- und Endzeitpunkt variieren, weil die fehlenden Informationen auf Grundlage des MTM-1 Bewegungszyklus erschlossen werden können. Eine zweite Greifbewegung wurde im Skelettmodell jedoch nicht erkannt, da, ähnlich wie bei der Beurteilung der Operationen, manuelle Handbewegungen nicht eindeutig unterschieden werden können und keine Information zum Zustand der Reißnadel vorhanden ist. Für Grundbewegungen eignet sich evtl. ein Ansatz mit Methoden der künstlichen Intelligenz, um aus den MoCap Informationen genauere Aussagen treffen zu können.

Sehr positiv fällt die Zustandsangabe zu „PSA tragend“ in beiden Faktorstufen auf, sodass sich aus den für die Zustandsänderung der Arbeitsperson notwendigen Bewegungen eine gute Anwendungsmöglichkeit in der beruflichen Bildung schließen lässt. Deutliche körperliche Bewegungen der Arme und Hände bei der hierfür notwendigen Operation werden trotz Informationsunterschied vergleichbar identifiziert, sodass sich für MoCap mit einem Skelettmodell bei solchen Charakteristika ein erster Einsatz in Lernprozessen anbieten könnte.

Ein zeitlicher Versatz aufgrund fehlender Informationen des Arbeitsobjekts ist beim Zustand „Profil abgelängt“ im Skelettmodell vorhanden. Der Zustand „Säge in Betrieb“ wird bei fehlenden Informationen im Skelettmodell risikoavers und wie bei der Handlung weniger unterschiedlich interpretiert. Trotz einer einfachen Zustandsangabe haben die vollständigen Videoinformationen keine Eindeutigkeit ergeben. Hieraus ergibt sich für die berufliche Bildung eine Limitation, da objektive, dynamische Informationen zu Arbeitsobjekten und Arbeitsmitteln aus den Skelettdaten nicht ableitbar sind. Sobald diese aus anderen Quellen bereitstehen, könnten extreme Potenziale für VR Umgebungen erschlossen werden, wobei die Verknüpfung der Daten ein Problemfeld ist (Wolf et al. 2020). Auf Basis der erhobenen Daten lässt sich eine bessere Einordnung anzunehmender Effektstärken für zukünftige konfirmatorische Studien mit

MoCap Daten in VR für die berufliche Bildung vornehmen. Für Studien mit stark kontrollierten Bedingungen wird empfohlen, deutlich präzisere Merkmale für Handlungen und Operationen zu definieren als dies in typischen betrieblichen Arbeitsplänen als Lerninhalt von Ausbildungen der Fall ist.

Für die berufliche Bildung sollten allerdings die gesamten Lösungsmöglichkeiten bei der Bearbeitung von Lernhandlungen berücksichtigt werden, um die selbstständige Lösungsgestaltung durch die Lernenden nicht einzuschränken. Dies stellt im Kontext von häufig begrenzten Entwicklungsressourcen für VR-Anwendungen einen prinzipiellen Zielkonflikt zur geforderten Arbeitsprozessdefinition für genaue Auswertungen dar, weil offene Lösungsräume sehr viel Implementierungsaufwand verursachen. Auch in der beruflichen Bildung ist der Entwicklungsaufwand stark abhängig von der definierten Zielgruppe und wächst mit deren Bildungs- und Erfahrungsstand. Besonderen Mehrwert kann der Einsatz von MoCap in Anwendungen bieten, die von einer schnellen Feedbackgabe profitieren.

6. Literatur

- Becker M (2013): Arbeitsprozessorientierte Didaktik. In: *bwp@ Berufs- und Wirtschaftspädagogik* (24).
 Björkenstam S, Mardberg P, Roller M, Carlson JS (2020): Digital Human Motion Planning of Operation Sequences Using Optimal Control of Hybrid Systems. In: Lars Hanson, Dan Högberg und Erik Brolin (Hg.): DHM2020. Proceedings of the 6th International Digital Human Modeling, 31-08.20-02.09.20. 11 Bände: IOS Press BV, S. 115–120.
 Jenewein K (2010): Handlungsorientiertes Lernen in der Berufsbildung. In: Bundesarbeitsgemeinschaften Elektrotechnik-Informatik und Metalltechnik (Hg.): lernen & lehren – Heft 98/ 25. Jahrgang/ 2010, S. 53–55.
 Mayring P (2015): Qualitative Inhaltsanalyse. Grundlagen und Techniken. 12. Aufl. Weinheim: Beltz (Beltz Pädagogik).
 Queiroz ACM, Fauville G, Herrera F, Da Leme MIS, Bailenson JN (2022): Do students learn better with immersive virtual reality videos than conventional videos? A comparison of media effects with middle school girls. In: *Technology, Mind, and Behavior* 3 (3). DOI: 10.1037/tmb0000082.
 REFA Bundesverband e. V. (Hg.) (2001): Arbeitsunterweisung und Qualifizierung. Lehrunterlage zu Modul 5100456. Darmstadt.
 Rueckert A, Niemann M, Kam E (2020): Human-Centric Assembly Cell & Line Validation. In: Hanson L, Högberg D, Brolin E (Hg.): DHM2020. Proceedings of the 6th International Digital Human Modeling, 31-08.20-02.09.20. 11 Bände: IOS Press BV, S. 279–287.
 Schelten A (2005): Grundlagen der Arbeitspädagogik. 4. Aufl. Stuttgart: Steiner.
 Volpert W (1983): Handlungsstrukturanalyse als Beitrag zur Qualifikationsforschung. 2. Aufl. Köln: Pahl-Rugenstein ((Sport, Arbeit, Gesellschaft, 5)).
 Windelband L (2018): Was bedeutet ‚prozessbezogen ausbilden‘ in der beruflichen Bildung im Zeitalter der Digitalisierung? Hg. v. AG BFN-Forum am 4./5. Juni 2018 in Nürnberg. Nürnberg.
 Wolf A, Miehl J, Wartzack S (2020): Challenges in interaction modelling with digital human models - A systematic literature review of interaction modelling approaches. In: *Ergonomics* 63 (11), S. 1442–1458. DOI: 10.1080/00140139.2020.1786606.

Danksagung: Der Beitrag entstammt dem Projekt „Arbeitswissenschaftliches Kompetenzzentrum für Erwerbsarbeit in der Industrie 4.0“. Es wird mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) innerhalb des Fachprogramms „Zukunft der Arbeit: Regionale Kompetenzzentren der Arbeitsforschung“ gefördert und vom Projektträger Karlsruhe unter dem FKZ 02L19C400 betreut.



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