dgp informationen

Fachexpertise
Forschung
Praxis

2021

62. Jahrgang Heft 71 Dezember 2021 ISSN 0948-0579

Geschlechtergerechtigkeit: Vision oder Illusion?

Fachexpertise

Geschlechtergerechtigkeit: Vision oder Illusion?
Seite 6

Ein Diagnoseinstrument für das Team
Seite 18

Sprachanalyse mit künstlicher Intelligenz zur Eignungsdiagnostik
Seite 34

Von Präsenz zu Remote: Erfahrungen mit und Forschungsbefunde zu Online-Assessment-Centern
Seite 46

Ein kritischer Kommentar zur DIN SPEC 91426
Seite 58

Forschung

Messung der Bedeutung von Demokratie – Methoden der Differenzierung
Seite 68

Vorstellung der dgp Testbatterie OV1
Seite 86

Vorstellung der dgp Testbatterie OG1
Seite 90

Praxis

Die Macht der Selbstwirksamkeit
Seite 98

Virtuelle Zusammenarbeit und Homeoffice – Veränderte Beanspruchung und Bewältigungsstrategien
Seite 104

Führen auf Distanz – (wie) gelingt das?
Seite 118

Live-Online-Seminare bei der dgp – ein Plädoyer für ein neues Format
Seite 126

Fachexpertise
Forschung
Praxis

dgp informationen

Liebe Leser*innen,

die aktuelle Zeit verlangt uns allen viel ab. Im Personalwesen fordert und ermöglicht die digitale Transformation, verstärkt durch die Covid19-Pandemie, unter anderem neue Wege der Personalauswahl- und Personalentwicklung. Der demographische Wandel führt weiterhin zu einem „war for talents“, da insbesondere die Nachwuchskräfte häufig fehlen. Bestehende Mitarbeiter*innen wünschen sich flexiblere Arbeitszeit- und ortmodelle, um beispielsweise mehr Zeit für Familie bzw. Privatleben zu haben. Dabei müssen die Prozesse im Personalmanagement fortwährend optimiert und akzeleriert werden.

Es macht uns stolz, dass wir Ihnen in den vorliegenden dgp-Informationen erfolgreiche Ansätze/Lösungen aus unserem Haus zeigen können, wie wir – in Zusammenarbeit mit unseren Kund*innen – den oben genannten Rahmenbedingungen begegnen, und zum Beispiel innerhalb kürzester Zeit zielführende Online-Formate geschaffen haben. Ganz nach dem Motto von Willy Brandt „Der beste Weg, die Zukunft vorherzusagen, ist, sie zu gestalten.“

Lesen Sie auch zu überaus wichtigen gesellschaftlichen Themen wie der Geschlechter(un)gerechtigkeit, der Messung der Bedeutung von Demokratie oder zur Entwicklung von Künstlicher Intelligenz im Kontext von Personalauswahl und natürlich zu fachpsychologischen Themen!

Wir wünschen Ihnen spannende Einblicke und viel Freude beim Lesen.

Ihre Ina Voigt
Leitende Psychologin der dgp e. V.

Fachexpertise

Geschlechtergerechtigkeit: Vision oder Illusion? S. 6

Ein Diagnoseinstrument für das Team S. 18

Sprachanalyse mit künstlicher Intelligenz zur Eignungsdiagnostik S. 34

Von Präsenz zu Remote: Erfahrungen mit und Forschungsbefunde zu Online-Assessment-Centern S. 46

Ein kritischer Kommentar zur DIN SPEC 91426 S. 58

Geschlechtergerechtigkeit: Vision oder Illusion?

Renée Schweisfurth & Ina Voigt

Sieht man sich die rechtlichen Rahmenbedingungen an, so dürfte es in Deutschland um das Thema Geschlechtergerechtigkeit gut bestellt sein:

Unsere Verfassungsnorm des Artikels 3, Absatz 2 sagt eindeutig „Frauen und Männer sind gleichberechtigt." Da die gelebte Praxis jedoch jahrzehntelang hinter diesem Grundsatz zurückblieb, wurde 1995 der folgende Satz ergänzt: „Der Staat fördert die tatsächliche Durchsetzung der Gleichberechtigung von Frauen und Männern und wirkt auf die Beseitigung bestehender Nachteile hin." 2001 kam das Bundesgleichstellungsgesetz hinzu, das auf eine gleichberechtigte Teilhabe von Frauen und Männern in der öffentlichen Verwaltung und die verbesserte Vereinbarkeit von Beruf und Familie zielt. Es schreibt fest, dass bei gleicher Qualifikation Frauen bevorzugt zu berücksichtigen sind, sofern nicht in der Person eines Mitbewerbers liegende Gründe überwiegen. Dieses Gesetz ist 2015 geschärft worden, u.a. dadurch, dass es Leitungskräfte besonders in die Pflicht nimmt, und die jeweilige Behörde einen Gleichstellungsplan als Instrument der Personalplanung und -entwicklung für vier Jahre mit konkreten Zielen erstellen muss.

Für Privatwirtschaft und öffentlichen Dienst gilt seit 2015 das „Gesetz für die gleichberechtigte Teilhabe von Frauen und Männern an Führungspositionen in der Privatwirtschaft und im öffentlichen Dienst“ (FüPoG), in dem für die Privatwirtschaft in Abhängigkeit vom Unternehmensstatus (Börsennotierung, Mitbestimmung) u.a. eine Quote für Aufsichtsratsposten festgeschrieben ist. Für den öffentlichen Dienst müssen nach dem FüPoG die obersten Bundesbehörden dem Statistischen Bundesamt jeweils zum 30. Juni Kontrolldaten (→ „Gleichstellungsindex“) liefern, die die Geschlechteranteile an Führungspositionen beschreiben. Im Juni 2021 haben der Bundestag und Bundesrat nun das FüPoG II beschlossen, das die Wirksamkeit des FüPoG weiter verbessern soll, indem u.a. ein Mindestfrauenanteil für Vorstände mit mehr als drei Mitgliedern verlangt wird. Zudem verbietet das Allgemeine Gleichbehandlungsgesetz (seit 2013) Diskriminierung und das Entgelttransparenzgesetz (seit 2017) Lohndiskriminierung.

Auch international werden Strategien zur Geschlechtergerechtigkeit formuliert. So betont die UN Geschlechtergerechtigkeit im Rahmen ihrer Nachhaltigkeitsstrategie, die OECD fordert Offenlegungspflichten und Frauenquoten, die EU setzt u.a. das Ziel einer mindestens 40-prozentigen Repräsentanz von Frauen auf allen Führungsebenen.

Frauen in Deutschland müssten zu Beginn ihrer Berufslaufbahn die besten Voraussetzungen mitbringen: Sie sind bei den höheren Schulabschlüssen (DESTATIS: Abitur 2019 39,1% der Mädchen versus 30,4% der Jungen) überrepräsentiert und stellen 2019 erstmals nach jahrzehntelanger Unterrepräsentanz auch mehr als die Hälfte der Studienabsolvent*innen – wenngleich unterschiedlich auf die möglichen Studienfächer verteilt. Ein genauerer Blick zeigt, dass Mädchen inzwischen im Durchschnitt bessere Noten haben und somit auch rein formal die besseren Voraussetzungen (siehe Abb. 1).

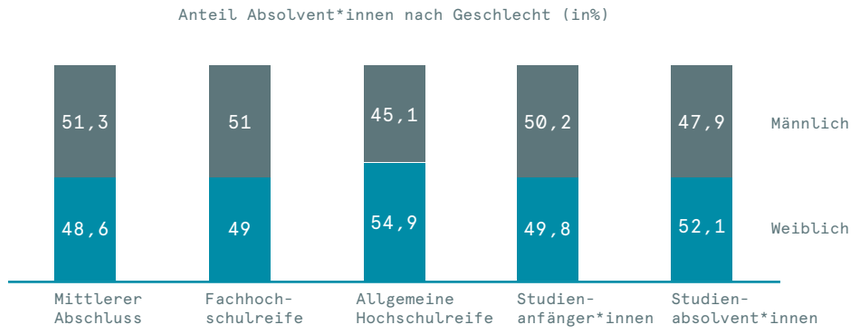


Abbildung 1: Frauen haben im Vergleich zu Männern rein formal die gleichen bzw. sogar bessere Voraussetzungen zum Berufsstart
Quellen: Statistisches Bundesamt, 2019; Ruß, Uwe (2021), Frauen in der Wissenschaft, Statistisches Monatsheft Baden-Württemberg, 1/2021

Warum dann dieser Artikel?

Die Praxis: Deutschland ist weit entfernt von Geschlechtergerechtigkeit

Nimmt man den Anteil von Frauen an Führungspositionen als einen Indikator für Geschlechtergerechtigkeit, so ist Deutschland im europäischen Vergleich (Eurostat) mit 29,4% in 2019 auf Platz 20 von 28. Der Frauenanteil an den Vorständen der 160 Börsenunternehmen beträgt in Deutschland lediglich 10,1% (AllBright September 2020). In einem Vergleich der AllBright Stiftung 2020 mit Unternehmen in Frankreich, Großbritannien, Polen, Schweden und den USA ist

Deutschland das einzige Land, in dem kein einziges Unternehmen einen Frauenanteil von 30% im Vorstand erreicht.

Der Anteil von Frauen in den Vorständen der jeweils 30 größten DAX-Konzerne beträgt im Jahr 2020 12,6%. In 11 der DAX-Konzerne wird nach wie vor keiner der Vorstandsposten von einer Frau besetzt. Von insgesamt 160 börsennotierten Unternehmen, die gesetzlich verpflichtet sind, Zielgrößen zur Steigerung des Frauenanteils an Vorstandssitzen zu veröffentlichen, haben 55 Unternehmen, die noch keine Frau im Vorstand haben, „Zielgröße Null“ (dies wird allerdings nach Inkrafttreten des FüPog II begründungspflichtig) gesetzt und zeigen damit, dass sie weiterhin nichts an diesem Umstand ändern möchten. Auch die Aufsichtsratsposten in deutschen Unternehmen sind nach wie vor anteilig nur mit 3,1 % Frauen besetzt.

Nach Angaben des Statistischen Bundesamtes gibt es in Deutschland 18.566 öffentliche Unternehmen. Wie die im Juli 2021 von der Zeppelin Universität vorgelegte „Fit-Public-Management-Studie“ zeigt, ist trotz Vorbildfunktion des öffentlichen Sektors auch hier die Repräsentanz von Frauen auf Top-Managementebenen defizitär (Städte 0% bis 50%, im Mittel 20,7%; Bundesländer 4,8% bis 35%; Bund 26,5%).

Das bedeutet, Frauen sind in Deutschland im Berufsleben nach wie vor benachteiligt, sie haben weniger Führungspositionen inne und liegen auch in Bezug auf das Gehalt hinter ihren männlichen Kollegen. Insgesamt zeigt sich eine starke Ungleichverteilung, was Machtpositionen in Deutschlands Unternehmen angeht. Auch werfen die pandemie-bedingten Krisenzeiten beim Thema Gleichstellung zurück: Während der Corona-Krise 2020 hat jedes fünfte weibliche Vorstandsmitglied den Vorstandsposten verlassen, was ein höherer Anteil ist als in den Jahren zuvor (siehe Abb. 2).

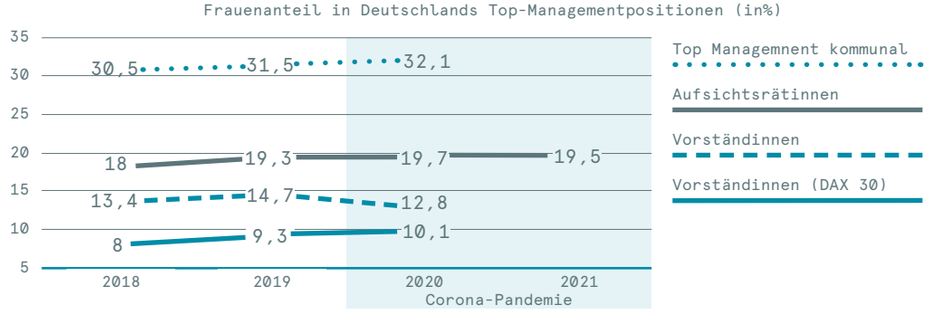


Abbildung 2: Der Frauenanteil in Deutschlands Top-Managementpositionen steigt über die Jahre nur marginal an bzw. sinkt sogar während der Corona-Krise leicht. (Quellen: Allbright, 2018,2019,2020; Papenfuß et al., 2021)

Es gibt Anzeichen dafür, dass uns die Pandemie diesbezüglich weiter zurückwirft. Bei Unsicherheit greifen Menschen bevorzugt wieder auf alte Rollenbilder zurück. Die Pandemie scheint u.a. eine Retraditionalisierung in den Geschlechterrollen zu bewirken. So sind es eher die Mütter, die sich um Homeschooling kümmern und die ihre Arbeitszeit zur Betreuung ihrer Kinder reduziert haben. Auch wenn hier noch keine abschließenden Zahlen vorliegen, steht zu erwarten, dass die Diskrepanzen beim Elterngeld mindestens bleiben (Mütter nahmen dies im Zeitraum 2010 bis 2018 im Mittel 13,5 und Väter 3,4 Monate in Anspruch; s.a. Allmendinger).

Ein weiterer Indikator für Geschlechtergerechtigkeit ist die Entgeltgerechtigkeit, die ja nicht nur für die aktuelle Lebenssituation von Frauen bedeutsam ist, sondern insbesondere Auswirkungen auf die (zu erwartende) Rente hat: Altersarmut in Deutschland ist weiblich! Das Durchschnittsbruttostundenentgelt von Frauen lag 2020 um insgesamt 18% (2017: 21%, 2007: 23%; DESTATIS) unter dem von Männern; der EU-Schnitt liegt hier 2019 bei einer Differenz von 14% (DESTATIS). Mütter in Deutschland verdienen langfristig 61% weniger als im Jahr vor der Geburt ihrer Kinder. Die Rentendifferenz beträgt etwa 53% (DIW 43/2017). Der Alterssicherungsbericht benennt bei den ab 65-Jährigen eine Differenz für 2020 von 41% (Frauen 833,- versus Männer 1.409,- Euro durchschnittliche Rente).

Die nach wie vor beschämende – und sich Pandemie bedingt voraussichtlich verschärfende – Diskrepanz in den beruflichen Perspektiven für Männer und Frauen hat vielfältige Ursachen:

Gender Care Gap

Frauen wenden im Durchschnitt täglich 52,4% mehr Zeit für unbezahlte Sorgearbeit auf als Männer (BMFSFJ, 2019). Dazu zählen Tätigkeiten der Haushaltsführung, die Pflege und Betreuung von Kindern und Erwachsenen und ehrenamtliches Engagement. So ist nach Angaben des Statistischen Bundesamts (Dezember 2020) die Anzahl Pflegebedürftiger in 2019 auf 4,1 Millionen angewachsen, etwa 80% wurden zu Hause versorgt. In Haushalten mit Kindern ist die Diskrepanz des täglichen Zeitaufwandes mit 83,8% am größten. Dieser Unterschied wird als Gender Care Gap bezeichnet. Wie die Erwerbs- und Sorgearbeit in Paarbeziehungen aufgeteilt ist, hat Einfluss auf die individuelle finanzielle und berufliche Perspektive. Wendet eine Frau mehr Zeit für unbezahlte Care-Arbeit auf, so bleibt ihr weniger Zeit für bezahlte Erwerbsarbeit. Dies zeigt sich auch an dem hohen Anteil der Mütter in Teilzeitbeschäftigung. Teilzeit ermöglicht es, einerseits Familie und Beruf besser miteinander vereinbaren zu können, hat aber andererseits auch eingeschränkte berufliche Entwicklungsmöglichkeiten, weniger Lohn und somit auch weniger Altersvorsorge zur Folge.

Mangelnde Vereinbarkeit von Beruf und Familie

Bestehende Arbeitsmodelle und strukturelle Gegebenheiten beeinflussen, wie Paare mit Kindern die Arbeit in ihrem Alltag aufteilen. Häufig sind die Rahmenbedingungen so, dass eine gleichberechtigte Teilhabe an Erwerbsarbeit nur möglich ist, wenn auf die private Sorgearbeit verzichtet wird. Beispielsweise wird es häufig als wichtig gesehen, dass eine Führungskraft ihre Position in Vollzeit wahrnimmt, wodurch teilzeitarbeitende Frauen tendenziell weniger in den Blick genommen werden, wenn es um den entsprechenden Aufstieg geht. Beruf und Familie lassen sich vor allem durch entsprechende Betreuungsangebote einfacher vereinbaren. Häufig kommen Eltern aufgrund langer Wartelisten nur verzögert an einen Betreuungsplatz, die Kosten für eine Betreuung sind hoch, wodurch die Familienfreundlichkeit der*s Arbeitgeber*in zunehmend an Bedeutung gewinnt. Auch hängt der Wunsch, Karriere zu machen und das Nachverfolgen von Karriereplänen davon ab, für wie realistisch eine Frau die Chance einschätzt, Kinder und Karriere zu vereinbaren.

Die zunehmende Digitalisierung und Flexibilisierung der Arbeitswelt bietet hinsichtlich der Vereinbarkeit von Beruf und Familie einerseits Chancen, andererseits zeigt sich in Studien bereits, dass Männer selbstbestimmtere Arbeitszeiten tendenziell nutzen, um länger zu arbeiten und mehr zu verdienen, wohingegen die Möglichkeit, Arbeitszeit und -ort stärker selbst bestimmen zu können nicht immer dazu führt, dass Frauen ihre Arbeitszeit auch aufstocken (Lott & Chung, 2016). Zwar lässt sich in manchen Fällen durch orts- und zeitflexible Arbeit die Vereinbarkeit von Beruf und Familie erleichtern, allerdings gibt es kaum Hinweise, dass sich dadurch die Verteilung der unbezahlten Sorgearbeit verändert (Carstensen & Demuth, 2020). Die Möglichkeit, immer mehr Situationen als Erwerbsarbeitszeit nutzbar zu machen, führt zu einer Entgrenzung bzw. Normalisierung eines Arbeitens immer und überall und so auch zu dem Versuch, Sorgearbeit und Erwerbsarbeit gleichzeitig zu erfüllen.

Die Möglichkeiten des orts- und zeitflexiblen Arbeitens können also zu zusätzlichen Belastungen führen. Es zeigt sich, dass eine solche Flexibilität nicht für alle Beschäftigten zur Verfügung steht. So wird von Beschäftigten mit geringerem Status und geringerer Autonomie häufig Anwesenheit vor Ort erwartet. Dies ist häufig der Fall in den nun als systemrelevant geltenden Berufen, in denen es eine hohe Anzahl an weiblichen Beschäftigten gibt (z.B. Pflegeberufe, Kassier*innen, Verwaltungen). Vereinbarkeit von Beruf und Familie bleibt trotz zunehmender Flexibilisierung durch Digitalisierung ein Problem, mit dem sich vermehrt Frauen auseinandersetzen (müssen) und ist oft mit der Erwartung verbunden, sich gut und ausreichend um die Kinder und den Haushalt zu kümmern, möglichst viel gleichzeitig zu schaffen und dabei auch noch im Beruf karriereorientiert zu handeln. Die Folge ist die Verunsichtbarung von Belastungen (Carstensen & Demuth, 2020).

Geschlechtsrollenstereotype und Sexismus

Häufig werden Männern und Frauen bestimmte Eigenschaften zugeschrieben. In Studien zeigt sich, dass sogenannte männliche Eigenschaften (z.B. durchsetzungsfähig, dominant) im Vergleich zu weiblichen Eigenschaften (z.B. liebevoll, sanft) stärker mit einer erfolgreichen Führungskraft in Verbindung gesetzt werden. Dies führt dazu, dass es Frauen einerseits schwerer haben, als Führungskraft bei Beförderungen in Betracht zu kommen und andererseits in Führungspositionen häufig mit einer Ambivalenz zurechtkommen müssen: Einerseits müssen sie – um als kompetente Führungskraft wahrgenommen zu werden – typisch „männliche“ Eigenschaften zeigen und gleichzeitig werden sie in ihrer Rolle als Frau gesehen, in der zu viele typisch „männliche“ Eigenschaften wiederum negativ bewertet werden (z.B. fordernd vs. „bossy“). Unabhängig vom Hintergrund einer Führungsposition sind Frauen im Beruf mehr Vorurteilen gegenüber ihrem Geschlecht ausgesetzt als Männer. Auch erleben Frauen mehr Sexismus als Männer (s.a. BMFSFJ-Pilotstudie „Sexismus im Alltag“).

Die Gläserne Decke

Der Begriff Gläserne Decke steht als Metapher für die unsichtbaren, kaum wahrzunehmenden Mechanismen, die Frauen trotz gleicher oder sogar besserer Voraussetzungen (wie beispielsweise der beruflichen Qualifikation) daran hindern oder es ihnen schwermachen, in höhere Managementpositionen zu kommen. Hierunter fallen traditionelle Geschlechterrollen, die dazu führen, dass Männer nach wie vor die „Familienernährer“ sind und Frauen sich um den Haushalt und den Nachwuchs kümmern. Spitzengremien sind somit nach wie vor männerdominiert. Diese „Monokultur“ kann auf Frauen wiederum abschreckend und dann als Selbstselektionsmechanismus wirken. Oftmals sind Auswahlprozesse bei der Besetzung von Spitzenpositionen wenig transparent oder standardisiert, wodurch Verzerrungen in der Wahrnehmung, wie beispielsweise Geschlechtsrollenstereotype oder Gender Biases, unbewusst eine Rolle spielen können.

In der Forschung zeigt sich, dass Männer bei gleicher Qualifikation bei der Besetzung einer neuen Stelle bevorzugt werden und das auch, wenn das Auswahlgremium zu einer höheren Anzahl aus Frauen besteht. Zudem wird die Qualität von männlichen Bewerbern tendenziell überschätzt (Eagly & Carli, 2007; Bagues & Esteve-Volart, 2010). Auch beim Netzwerken ist es für Frauen schwieriger, denn in sogenannten „Old Boy Networks“, in denen sich Mitglieder befinden, die bereits in Machtpositionen sind, haben Frauen schlechtere Chancen, überhaupt hereinzukommen und dort von anderen gezielt beruflich gefördert zu werden. Dies liegt darin begründet, dass solche stärker wirksamen männerdominierten Netzwerke auf informellen „Seilschaften“ basieren, die ohne Rahmen und nur auf Ähnlichkeit und Sympathie beruhen.

Dieses Prinzip der Selbstähnlichkeit wirkt auch bei der Nachwuchsförderung und bei Personalentscheidungen: (Ältere) Männer fördern und bevorzugen lieber jüngere Männer (männlicher „Karrierehabitus“) als junge Frauen. Einen Gegenpol dazu versuchen Unternehmen zu setzen, indem sie ihren Mitarbeiterinnen Frauennetzwerke oder gezielte Förderprogramme ermöglichen. Leider haben solche Programme auch aufgrund ihrer mangelnden Wirksamkeit das Image, Pink-Washing zu betreiben. Einige Programme wirken demnach nicht karrierefördernd, sondern vermitteln das Bild, dass Frauen in Entwicklungs- und Coachingprogrammen stecken, während die Männer die Jobs bekommen (Gutmann & Seisl, 2014).

Die aufgeführten Ursachen sind nur eine Auswahl der unterschiedlichen Wirkmechanismen. Sie zeigen, dass obwohl Mädchen im Vergleich zu Jungen im Schulsystem besser abschneiden, häufiger Abitur machen und somit rein formal die besseren Voraussetzungen haben, dennoch geringere Chancen haben, ihre Qualifikationen so einzusetzen, dass daraus auch die gleichen Karriere- und Einkommensmöglichkeiten entstehen.

Ziele und Maßnahmen

Nach dem AllBright-Bericht von September 2020 würde es bei der Fortschreibung des aktuellen Tempos noch 100 Jahre dauern, bis ein ausgeglichenes Verhältnis von Männern und Frauen in Vorständen erreicht ist. Auch andere Parameter stimmen wenig hoffnungsfroh.

Angesichts der nach wie vor starken Benachteiligung von Frauen und des zögerlichen Fortschritts ist es unausweichlich darüber zu sprechen, welche Strategien über bereits bestehende Maßnahmen hinaus die Gleichstellung von Frauen im Arbeitsleben antreiben können.

Hier gibt es nicht die eine Maßnahme, sondern es sind Maßnahmenbündel erforderlich. Die gesetzlichen Rahmenbedingungen sind gegeben, gleichwohl gibt es noch große Defizite in der Umsetzung. Es bleibt eine große gesellschaftspolitische Aufgabe, den bestehenden Klischees entgegenzuwirken und positive Rollenvorbilder zu schaffen. Insbesondere bleiben Kinderbetreuung und Pflege essenzielle Felder, in denen die Politik bessere Rahmenbedingungen für die Vereinbarkeit von Beruf und Familie schaffen muss.

Wir konzentrieren uns in diesem Artikel auf Handlungsfelder in der Arbeitswelt, insbesondere auf Strategien der Organisationsentwicklung und Personalentwicklung mit einem besonderen Fokus auf Führungsverhalten und Personalauswahl.

Organisationsentwicklung

Geschlechtergerechtigkeit ist auch ein unternehmenskulturelles Thema im Sinne grundlegender gemeinsamer Werte. Essenziell ist eine Verflechtung von Gleichstellungs- und Diversitätszielen in Leitbild, strategischen Zielen etc. Damit diese konkret handlungsleitend sind, arbeiten viele Organisationen mit Verhaltenskodizes wie etwa Führungsleitlinien oder Kommunikationsgrundsätzen, aus denen auch Regeln in Bezug auf respektvoll-diskriminierungsfreies Verhalten abgeleitet und formuliert werden. Organisationsentwicklerisch nicht zu unterschätzen ist beispielsweise auch eine klare Sprachregelung wie z.B. zum Genderstar o.ä., die durchaus bewusstseinsbildend sensibilisieren kann. ´

Ein schon vor der Pandemie zu beobachtender und nun beschleunigter Wandel von der Präsenz- zur Effizienzkultur sowie mobiles Arbeiten könnten in Bezug auf die Vereinbarkeit von Beruf und Privatleben positive Effekte haben. Viele angesichts der demografischen Situation und des damit verbundenen Fach- und Führungskräftemangels eingeleiteten Strategien in Bezug auf die Veränderung von (Leitungs-)Strukturen und Prozessen können mittelbar die Chancen von Frauen verbessern. Volkswirtschaftlich ist es nicht mehr zu verantworten, auf die gut ausgebildeten Frauen in dem jetzt gegebenen Ausmaß in der Arbeitswelt zu verzichten (ein nicht ganz unproblematisches Argument, da nicht wertegeleitet, sondern utilitaristisch).

Personalauswahl

Essenziell für eine treffsichere Personalauswahl ist das zu Grunde liegende Anforderungsprofil, das dann im eignungsdiagnostischen Prozess mit den Kompetenzen des*r Bewerber*in abgeglichen wird. Schon bei der Formulierung des Anforderungsprofils ist darauf zu achten, dass nicht Beschreibungen als Operationalisierung gewählt werden, die potenziell diskriminieren oder unbewusst Geschlechtsrollenstereotype bedienen könnten. Anforderungsprofile für Führungspositionen sollten noch stärker an den erforderlichen sozialkommunikativen Führungsfähigkeiten ausgerichtet werden, die als besonders erfolgskritisch gelten. Bei der Vorauswahl nach Bewerbungsunterlagen ist zu beachten, dass die Information über das Geschlecht die Bewertung von Bewerbungsunterlagen beeinflussen kann. Frauen werden als weniger kompetent eingestuft, dies umso mehr, je attraktiver sie sind (siehe auch Studie Uni München). Fehler in der ersten Vorauswahlstufe sind i.d.R. im weiteren Prozess nicht mehr zu heilen; anonymisierte oder automatisierte Vorauswahl kann diese Fehler reduzieren.

Nicht nur aus Gründen von Qualitätsstandards ist es zwingend, dass Mitglieder von (möglichst divers besetzten) Auswahlkommissionen sensibilisiert werden für Bewertungsfehler. Dies ist auch erforderlich, um unbewusst durch Geschlechtsrollenklischees geleitete selektive Wahrnehmungen und Bewertungen, die gleiches Verhalten in Abhängigkeit vom Geschlecht unterschiedlich werten, zu verdeutlichen.

Nach der Eignungsfeststellung folgt die Auswahlentscheidung: Hier gilt es, bei Gleicheignung quotenbezogen zu entscheiden und zu bedenken, dass neben dem grundlegenden Wert der Geschlechtergerechtigkeit Diversität i.d.R. die Problemlösekompetenz in Teams erhöht. Regelhaft sollte der gesamte eignungsdiagnostische Prozess kritisch im Hinblick auf Diskriminierungsfreiheit evaluiert werden.

(Weitere Ausführungen zu gendergerechter Personalauswahl finden sich im u.a. Artikel von Ristel und Jobmann in den dgp-Informationen 2017.)

Personalentwicklung

Ziel von Personalentwicklung ist es, ein optimales Verhältnis zwischen dem zu gewährleisten, was die Organisation braucht und dem, was die Mitarbeiter*innen an Potenzialen und Motiven mitbringen. Dies ist in einer sich schnell wandelnden Arbeitswelt ein nie abgeschlossener Prozess. Grundsätzlich gilt es natürlich, Ressourcen und Potenziale sorgfältig zu ermitteln (siehe auch die nicht genutzten Potenziale hoch qualifizierter Frauen) und eine aufgabenangemessene Personalzumessung zu leisten. Letztere ist auch wichtig, um in Teams, in denen viele Frauen mit Familienaufgaben befasst sind, Konflikten im Zusammenhang mit Aufgabenverteilung vorzubeugen.

Personalentwicklung bedient sich eines Pools an Instrumenten. Im Hinblick auf das Thema Geschlechtergerechtigkeit bieten sich hier unter anderem Schulungen an, die Mitarbeiter*innen und Führungskräfte sensibilisieren für unbewusste Stereotypen und zur diesbezüglichen Reflexion anregen. In Hinblick auf spezifische Förderung von Frauen empfehlen sich (wie schon von vielen praktiziert) gute Mentoring- und Coaching-Angebote an und z.B. auch die temporäre Übernahme von Führungsverantwortung als Mittel der Personalentwicklung (hier für Frauen) sowie natürlich gezielte Qualifizierungsangebote.

Eine weitere Strategie ist das lebensphasenorientierte Personalmanagement, das die individuelle Berufs- und Lebenssituation von Mitarbeiter*innen in den Fokus nimmt und bei der Ausgestaltung von Arbeit und im Führungsverhalten berücksichtigt: Unterschiedliche Personalentwicklungs- bzw. Führungsinstrumente müssen bei unterschiedlichen Personen unter Berücksichtigung ihrer unterschiedlichen Hintergründe/Lebenssituationen passgenau eingesetzt werden. Lebensphasenorientierung (eigentlich müsste es Lebens- und Berufsphasenorientierung heißen) bedeutet, persönliche Belange der Mitarbeiter*innen zu sehen, individuelle Lösungen zu finden und so auch als Arbeitgeber attraktiv zu bleiben. Dies kann auch eine Verbesserung der Möglichkeiten für Führung in Teilzeit bedeuten.

Die jährliche Veröffentlichung statistischer Datenblätter zu gleichstellungsrelevanten Kennzahlen und auch gezielte Aktionstage mit gleichstellungsrelevanten Themen verdeutlichen einen Selbstverpflichtungscharakter des Arbeitgebers, ein Commitment auf Gleichstellungsziele. Auch die Sichtbarmachung von Vorbildern weiblicher Führungskräfte, die zeigen, dass Vereinbarkeit von Beruf und Familie möglich ist, dürfte im Sinne positiver Rollenmodelle zielführend sein.

Führung

Führungskräfte müssen sich mit den gewünschten Werten der Organisation (und mit gesetzlichen Vorgaben) identifizieren und diese vertreten. Sie müssen also authentische Vorbilder für Gleichstellung/Geschlechtergerechtigkeit sein, und darüber hinaus überzeugen, dass Vielfalt den Unternehmenszielen und der Unternehmensidentität dient. Dies heißt auch, dass sie ihre eigenen Einstellungen in Bezug auf Geschlechtergerechtigkeit reflektieren und ihr Verhalten beispielgebend ausrichten. Dabei geht es nicht nur darum, die eigenen Werte transparent zu machen, sondern darum, wie diese (vor)gelebt werden. Werte als ethische Normen können z.B. auch über Verhaltenskodizes bzw. Führungsgrundsätze transportiert werden. Geschlechtergerechtigkeit ist hierbei die Norm, Gleichstellung die konkrete Führungsaufgabe und damit auch eine handwerkliche Frage in Bezug auf die Anwendung bestimmter Führungsinstrumente. Dies heißt einerseits, im Führungsalltag eine Kommunikationskultur zu gestalten, die jegliche Form von Diskriminierung unterbindet, Überzeugungsarbeit zu leisten in Bezug auf Geschlechtergerechtigkeit und aktiv einzuschreiten, wo mit sexistischen Bemerkungen und Herabwürdigungen wie z.B. „Quotenfrau“ hantiert wird. Die Me-Too-Debatte hat hier zu einer erhöhten Sensibilität geführt, was auch dem hier in Rede stehenden Thema zu Gute kommt.

Essenziell ist eine kritische Reflexion der eigenen Bereitschaft z.B. in Bezug auf flexiblere Arbeitszeitmodelle und insbesondere in Bezug auf das eigene Bewertungsverhalten. Nach wie vor wird das gleiche Verhalten, je nachdem, ob es von einem Mann oder einer Frau gezeigt wird, unterschiedlich bewertet. Dies kann sich z.B. negativ niederschlagen, wenn es um dienstliche Beurteilung geht.

Führungskräfte müssen insgesamt mit guten sozialkommunikativen Fähigkeiten ausgestattet sein und über einen ethischen Kompass verfügen. Sieht man sich den deutlichen Führungskulturwandel an (s.a. INQA), so wird deutlich, dass eine Kultur des „Das haben wir immer schon so gemacht“ bzw. des „Das haben wir noch nie so gemacht“ abgelöst werden muss durch eine Ermöglichungskultur, in der es darum geht, mit Mut, Kreativität und Phantasie neue Problemlösungen zu finden und Räume zu öffnen.

Ausblick

Betrachtet man sich die Gesamtsituation von Frauen in Deutschland, so scheint das Ansinnen verwirklichter Geschlechtergerechtigkeit illusionär. Dies ist jedoch nicht hinnehmbar! Die Vision (als angestrebter Idealzustand in der Zukunft) von Gleichstellung ist sowohl aus ethischen als auch aus rechtlichen Gründen nicht verhandelbar.

Offenkundig ist jedoch, dass hierzu sehr konkrete Ziele zu formulieren sind sowie klare Strategien und insbesondere auch eine Evaluierung dieser erfolgen muss. Neben den durch die Politik zu verantwortenden gesellschaftlichen Rahmenbedingungen bleibt es insbesondere Führungsaufgabe, aktiv Gleichstellung zu gestalten und hierzu Rechenschaft abzulegen. Die in der Vergangenheit immer wieder geäußerte Hoffnung, dass aufgrund der zunehmenden Qualifikation von Frauen das Problem sich quasi von selber löse, hat sich als trügerisch erwiesen. Abwartende Strategien sind nicht mehr angezeigt. Gerade nach der Pandemie, in der

viele Themenfelder einer Neuaufstellung bedürfen, bietet es sich an, sich auch zum Thema Gleichstellung aktiv neu zu verorten. Dies passt sehr gut in die schon länger angeschobene Entwicklung hin zu einer neuen Führungskultur.

Neben ethischen Argumenten für mehr Geschlechtergerechtigkeit gilt volkswirtschaftlich, dass Deutschland es sich nicht (mehr) leisten kann, auf die Potenziale der gut ausgebildeten Frauen zu verzichten. Insbesondere die demografische Situation, die sich u.a. in einem Mangel an Fach- und Führungskräften zeigt, aber auch die grundlegenden Veränderungen in der Arbeitswelt (s.a. digitale Transformation, COVID-getriebene Veränderungen etc.) erfordern einen Aufbruch. Die aktuelle Diskussion um Diversität in der Arbeitswelt bietet auch die Chance, Geschlechtergerechtigkeit als konkretes Ziel zu setzen und die Verantwortlichen in Politik sowie Führungskräfte und auch die Frauen selbst in die Pflicht zu nehmen. Der öffentliche Dienst hat hierbei eine Vorbildfunktion.

Literatur:

AllBright Stiftung gGmbH. (2020, 01.September). Deutscher Sonderweg. Frauenanteil in DAX-Vorständen sinkt in der Krise. Abgerufen von: https://www.allbright-stiftung.de/berichte

Allmendinger, J. (2021): Es geht nur gemeinsam! Wie wir endlich Geschlechtergerechtigkeit erreichen; Ullstein

Bagues, M. F., & Esteve-Volart, B. (2010). Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment. Review of Economic Studies, 77(4), 1301–1328. https://doi.org/10.1111/j.1467-937X.2009.00601.x

Bundesministerium für Familie, Senioren, Frauen und Jugend (2020). Auf dem Weg zur Entgeltgleichheit von Frauen und Männern. Daten, Ursachen, Maßnahmen.

Bundesministerium für Familie, Senioren, Frauen und Jugend (2020). Sexismus im Alltag. Wahrnehmungen und Haltungen der deutschen Bevölkerung. Pilotstudie.

Carstensen, T. & Demuth, U. (2020). Wandel der Geschlechterverhältnisse durch Digitalisierung. Working Paper Forschungsförderung, 201 (Dezember 2020). Hans-Böckler-Stiftung.

Deutsches Institut für Wirtschaftsforschung e.V. (2017). Gender Pay Gap. DIW Wochenbericht 43/2017. Abgerufen von: https://www.diw.de/documents/publikationen/73/diw_01.c.567545.de/17-43.pdf

Eagly, A. H., & Carli, L. L. (2007). Women and the labyrinth of leadership. Harvard Business Review, 85(9), 62–71, 146.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. Psychological Review, 109(3), 573–598. https://doi.org/10.1037/0033-295X.109.3.573

Eurostat (2021). Geschlechtspezifisches Verdienstgefälle, ohne Anpassungen, nach NACE Rev. 2 Tätigkeit - Methodik: Lohnstrukturerhebung. Abgerufen von: https://ec.europa.eu/eurostat/databrowser/view/earn_gr_gpgr2/default/table?lang=de

Fesefeldt, J. & Voigt, I.: Führungsethik. Philosophische Ansätze und führungspsychologische Konsequenzen. dgp-Informationen, 53. Jahrgang, Heft 62, September 2012 (s.a. www.dgp.de)

Groll, T. (2017, 6. März). Klüngeln für die Karriere. Süddeutsche-Zeitung. https://www.zeit.de/karriere/beruf/2017-02/netzwerken-karriere-vorteil-frauen/komplettansicht

INQA: Monitor Führungskultur im Wandel https://www.inqa.de/.../monitor-fuehrungskultur-im-wandel.html

Lott, Y. & Chung, H. (2016). Gender Discrepancies in the Outcomes of Schedule Control on Overtime Hours and Income in Germany. European Sociological Review, August 2016. doi: 10.1093/esr/jcw032

Ristel, N. & Jobmann, A.-L.: Gendergerechte Personalauswahl – rechtliche, wissenschaftliche und praktische Hinweise; in dgp-Informationen 2017, 58. Jahrgang, Heft 67.

Ruß, Uwe (2021). Frauen in der Wissenschaft. Statistisches Monatsheft Baden-Württemberg, 1/2021

Statistisches Bundesamt (Destatis). (2021, 9. März). Gender Pay Gap 2020: Frauen verdienten 18 % weniger als Männer. Abgerufen von: https://www.destatis.de/DE/Presse/Pressemitteilungen/2021/03/PD21_106_621.html;jsessionid=25DC020D30F71B8383442C69CA3764BB.live711

Statistisches Bundesamt (Destatis). (2019). Frauen in Führungspositionen in der EU. Abgerufen von: https://www.destatis.de/Europa/DE/Thema/Bevoelkerung-Arbeit-Soziales/Arbeitsmarkt/Qualitaet-der-Arbeit/_dimension-1/08_frauen-fuehrungspositionen.html

Voigt, I.: Lebensphasenorientiertes Personalmanagement; in dgp-Informationen 2015, 56. Jahrgang, Heft 65.

Voigt, I.: MeToo und Machtmissbrauch: Eine führungsethische Einordnung; in dgp-Informationen 2018, 59. Jahrgang, Heft 68.

Von Terzi, P. (2020). A man's world? – Wie sich Frauen in Männerberufen behaupten. In F. C. Brodbeck (Hrsg.), Evidenzbasierte Wirtschaftspsychologie, (36). Ludwig-Maximilians-Universität München. http://www.evidenzbasiertesmanagement.de.

Voyer, D., & Voyer, S. D. (2014). Gender Differences in Scholastic Achievement: A Meta-Analysis. Psychological Bulletin, 140(4), 1174-1204. doi: 10.1037/a0036620

Zeppelin Universität, Lehrstuhl für Public Management and Public Policy: „Fit-Public-Management-Studie", 2018 und 2021; www.zu.de

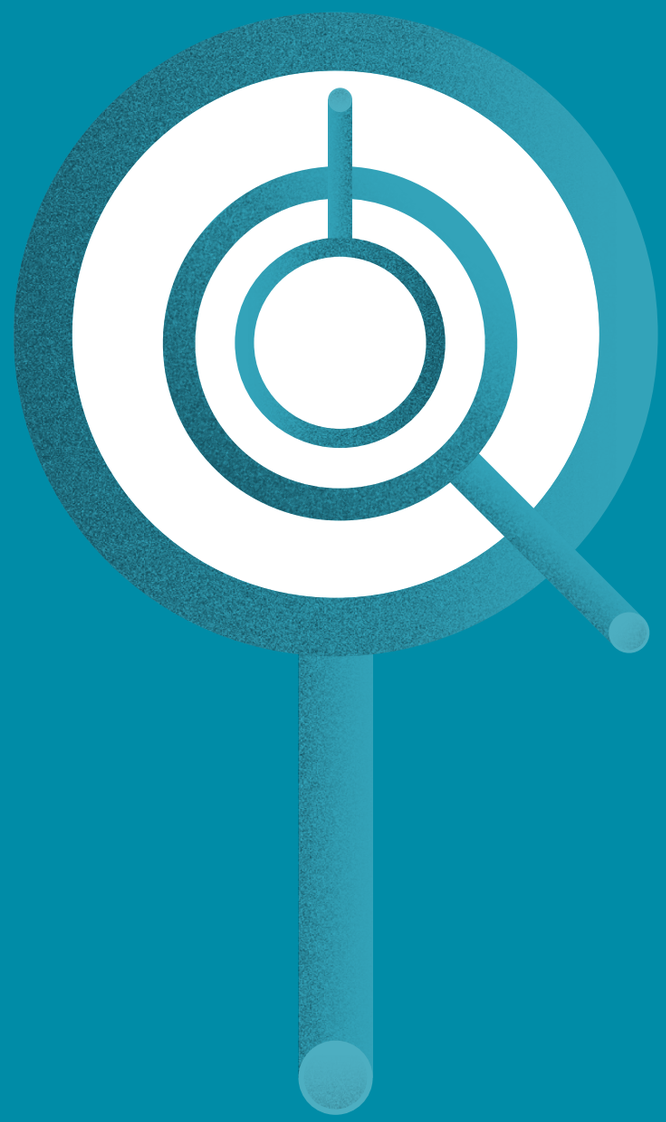
Kontakt:

Ina Voigt
Leitende Psychologin der dgp
Deutsche Gesellschaft für Personalwesen e.V.
voigt@dgp.de

Renée Schweisfurth,
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
schweisfurth@dgp.de

Ein Diagnoseinstrument für das Team

Isabel Lorenz & Sabrina Köpke



„Die Kommunikation könnte besser sein … wir wollen mal einen Teamtag dazu machen“ – so oder so ähnlich lautet oft die erste Anfrage einer Teamleitung an uns als beratende Psycholog*innen. Oft stehen dahinter Vorstellungen von oder auch bereits Erfahrungen mit Team-Building-Spaß-Aktivitäten oder allgemeinen Fortbildungen zu Teamthemen. Diese sind zeit- und kostenunaufwändig und beruhigen das Gewissen, weil man ja etwas fürs Team getan hat… Sie helfen aber zumeist weder dem Team dabei, sich wirklich zu hinterfragen und bestehenden Schwierigkeiten ernsthaft auf den Grund zu gehen, noch hinterlassen sie einen verbindlichen Nachhall, im Sinne eines systematisch gesteuerten und evaluierbaren Veränderungsprozesses.

Meist bleibt der Handlungsspielraum zum Erkennen und Behandeln von Teamschwierigkeiten bereits dadurch begrenzt, dass sich erst während des angesetzten Teamtags selbst andeutet, was das eigentliche Thema ist (wenn überhaupt…) und dann geht man wieder auseinander ohne nachhaltigen Effekt.

Wir wollen Teamentwicklungsbedarfe ernst(er) nehmen und wirksam angehen. Deshalb haben wir uns zum Ziel gesetzt, Teams die Möglichkeit zu einer systematischen grundständigen und individuellen Selbst-Diagnose zu geben.

Dazu haben wir einen umfassenden Fragebogen für Teams entwickelt, der verbreitete theoretische Modelle und unsere Praxiserfahrung in der Arbeit mit Teams integrativ aufgreift.

In diesem Artikel erläutern wir das theoretische Fundament, das die Bestimmungsstücke für unser Verständnis von Team-Sein, Team-Bedarfen, Team-Entwicklung und den abgeleiteten Anspruch an Teamentwicklungsmaßnahmen bietet, bei denen der Diagnose-Fragebogen gezielt und aussichtsreich zum Einsatz kommen soll. Der Fragebogen selbst wird dann im zweiten Teil des Artikels ausführlicher beschrieben.

Ein Team erkennen

Aufgrund der zahlreichen und unterschiedlichen Definitionen von Team und Gruppe möchten wir damit beginnen, eine Arbeitsdefinition aufzustellen, die unser Verständnis von Team und damit auch Teamentwicklung aufgreift und verdeutlicht.

Wir verstehen das Team als eine besondere Art der Gruppe, in der gruppendynamische Prozesse wirksam sind und bewusste wie unbewusste Vorgänge auftreten. Ein Team zeichnet sich durch Ergebnis- und Prozessabhängigkeit aus – es bestehen teilweise gemeinsame, teilweise geteilte Aufgaben, die ineinandergreifen müssen und mindestens ein übergeordnetes Ziel haben, an dem gearbeitet wird. Diese Abhängigkeiten zwischen den Teammitgliedern haben eine gegenseitige Beeinflussung von Verhalten und Handlungen zur Folge: Jede Interaktion zwischen Teammitglied und Umgebung führt zu stetigen Veränderungsprozessen (vgl. Ilgen, Hollenbeck, Jonson & Jundt, 2005).

Diese Veränderungsprozesse können eine große Herausforderung für die Teammitglieder darstellen, da sie eine ständige Anpassung und auch eigenverantwortliche Mitgestaltung erfordern (Kauffeld & Lehmann-Willenbrock, 2008). Neben den Abhängigkeiten und interaktionsbasierten Veränderungsprozessen entsteht im Team bestenfalls eine Zusammengehörigkeit, ein „Wir-Gefühl“ im Sinne einer persönlichen Identifikation der einzelnen Teammitglieder mit dem Team im Spannungsfeld zwischen „ich“ und „wir“ (Busch, & Oelsnitz, 2018). Stärker identifizierte Teammitglieder unterstützen sich hierbei gegenseitig und verbessern somit indirekt das Teamklima (Riketta & Van Dick, 2005). Das Ausmaß der Identifikation ist allerdings schnell veränderbar und damit abhängig von der Situation und dem Wandel des Kontexts.

Kurz gesagt, lässt sich ein Team daran erkennen, dass es sich um eine zusammengesetzte Gruppe handelt, deren Mitglieder zum einen auf einer strukturellen Ebene durch eine klare Aufgaben- und Zielorientierung miteinander verbunden sind, zum anderen auf einer Ebene (psycho-)sozialer Prozesse ein Wir-Gefühl entwickeln können, je nachdem wie stark und beständig die Identifikation der einzelnen Teammitglieder mit dem Team ausfällt. Die Teammitglieder sind also auf verschiedenen Ebenen voneinander abhängig und beeinflussen sich gegenseitig.

Beim Team-Sein handelt es sich also nicht um ein statisches, sondern um ein prozesshaft dynamisches Zusammenspiel: Das Team ist wandelbar und reagiert auf äußere und innere Veränderungen.

Die Bestimmung zentraler Teammerkmale, die wir hier nach Ist-Zustand einerseits und Prozessverläufen andererseits und darin verankerten sachorientieren (Aufgaben, Ziele) Strukturen und (psycho-) sozialen Prozessen vornehmen, orientiert sich am Ordnungsschema von Kauffeld (2001). Sie unterscheidet hinsichtlich der Teamdiagnose prozessanalytische und strukturanalytische Instrumente. Der von ihr entwickelte Fragebogen zur Arbeit im Team (F-A-T) ordnet Teammerkmale im heuristischen Modell der Kasseler Teampyramide an, wonach Zielorientierung → Aufgabenbewältigung → Zusammenhalt → Verantwortungsübernahme als Bestimmungsmerkmale von Teams in einer hierarchischen Ordnung stehen, die Prozesse auf der Sachebene also das Fundament für das Zusammenwachsen auf der sozialen Ebene bilden.

Integrativ gedacht, möchten wir das Team in seinem Entwicklungskontext betrachten, positive Entwicklungsfortschritte anregen und zielgerichtet beeinflussen. Nachfolgend werden wir abgeleitet von der hier skizzierten Team-Definition einige theoretische Konstrukte und Modelle aufgreifen, die sich mit Teamentwicklung und deren Steuerung befasst haben. Diese werden wir anschließend in die Vorstellung unseres eigenen Teamdiagnose-Fragebogens übernehmen, welcher im Rahmen von Teamentwicklungsmaßnahmen seine gezielte Anwendung finden soll.

Team braucht Entwicklung

Ein Team kann sich auf unterschiedlichen Entwicklungsstufen bewegen, wie es Tuckman in seinem viel zitierten Phasenmodell (1977) beschreibt. Im Modell werden fünf Stufen benannt, die jedoch nicht schrittweise nacheinander stattfinden müssen, sondern sich wiederholen, übersprungen oder ausgelassen werden können. Die Anfangsstufe nennt er *Forming*: Mitglieder der Gruppe lernen sich kennen und sind in dieser Phase auf sich selbst und die Gruppe konzentriert. Aufgaben spielen eher eine untergeordnete Rolle. Die nächste Phase ist das *Storming*: Hier verfestigen sich Positionen, und Rollen werden verteilt. Es kann zu Intragruppenkonflikten kommen, das Zusammengehörigkeitsgefühl und die Identifikation mit der Gruppe ist noch nicht entwickelt. In der Phase des *Norming* entwickelt sich nach der Aufarbeitung des Intragruppenkonflikts das „Wir-Gefühl“. Die Phase des *Performing* bezeichnet eine Gruppe, die arbeits- und entscheidungsfähig ist. Jetzt können die Fähigkeiten einzelner Mitglieder gezielt eingebracht werden, das Leistungspotenzial voll ausgeschöpft werden, da soziale Konflikte und Unsicherheiten nicht mehr den Großteil der Ressourcen binden. Die letzte Phase des *Adjourning* beschreibt eigentlich einen Übergang in wieder anfängliche Phasen, denn dabei geht es um die Teamauflösung, z.B. am Ende eines Projektes oder aber auch, wenn ein Team durch Fluktuationen neu aufgestellt wird.

Für die Entwicklung von Teams spielt, wie oben erwähnt, auch das Ausmaß der Identifikation der einzelnen Teammitglieder eine maßgebliche Rolle. Identifikation entwickelt sich vor allem aufgrund von wahrgenommener Ähnlichkeit, geteilter Werte und

Normen (siehe auch: Theorie der sozialen Identität, Tajfel, 1974). Die Identifikation mit dem Team hängt mit der Wahrnehmung eines guten Teamklimas und der Bereitschaft, sich für die Kolleg*innen einzusetzen und sich gegenseitig zu unterstützen, zusammen (Riketta & van Dick, 2005). Von einer ausgeprägten Identifikation kann dabei ausgegangen werden, wenn ein enges Zugehörigkeits- und Verbundenheitsgefühl entsteht. Identifikationsprozesse sind dabei abhängig von der jeweiligen Situation, der Umgebung und persönlichen Faktoren. Für eine starke, gemeinsame Identität sind u.a. Symbole und Rituale hilfreich (van Dick & Schuh, 2016).

Auch die Teamreflexion, nachfolgend beschrieben nach West (2012), nimmt eine fördernde und gestaltende Rolle in Entwicklungsprozessen von Teams ein. Die Teamreflexion hilft einem Team zum einen, sich an verändernde Umgebungsfaktoren anzupassen und zum anderen, Unterschiedlichkeiten im Team wertzuschätzen (West, 2012). Nach West (2012) beinhaltet sie folgende Aspekte:

- → Regelmäßige Überprüfung der Teamziele nach Relevanz, Angemessenheit und Grad der Erfüllung;
- → Sensibilität für äußere Veränderungen, die die Teamarbeit beeinflussen;
- → Wahrnehmung, Bewertung und Diskussion der Teamarbeit mit dem Fokus der Performanz-Steigerung;
- → Kreativität, Flexibilität und Anpassungsfähigkeit;
- → Toleranz für Unsicherheit;
- → Wertschätzung der Teammitglieder für unterschiedliche Perspektiven, Wissensstände, Fähigkeiten und Erfahrungen anderer Teammitglieder.

Für ein Team ist Entwicklung nicht nur notwendig, sondern sowieso unvermeidbar. Es befindet sich durch Interaktionen mit inneren und äußeren Faktoren in einem stetigen Veränderungsprozess (vgl. Ilgen, Hollenbeck, Jonson & Jundt, 2005). Diese unabwendbaren Entwicklungsprozesse können positive (z.B. Wertschätzung anderer Teammitglieder und deren Perspektiven) und auch negative Auswirkungen (z.B. Überforderung, anhaltende Konflikte) nach sich ziehen.

Um die Dynamik im Team in positive Entwicklungsbahnen zu leiten, ermöglichen bewusst und gezielt angesetzte Teamentwicklungsmaßnahmen eine Aufdeckung und klare Steuerung von Optimierungsfeldern, aber auch von auszuschöpfenden bestehenden Ressourcen im Team.

Teamentwicklung bewusst steuern

Oben Gesagtes verdichtend, geht es in unserem Verständnis von Maßnahmen zur Teamentwicklung also darum, sich bewusst und gezielt mit gruppendynamischen Erscheinungsformen, Kohärenz, Identifikationsausmaß sowie Abläufen und Strukturen innerhalb eines Teams zu beschäftigen, da dies die Bausteine sind, die sowohl den aktuellen Status als auch die Veränderungspotenziale eines Teams bestimmen.

Um diese Themenfelder entsprechend ihrer Bedeutung für die individuelle Teamentwicklung einkreisen und mit dem Gesamtteam behandeln zu können, ist eine sukzessive Annäherung notwendig. Eine Teamentwicklungsmaßnahme besteht somit im Idealfall aus mehreren Schritten:

1. Kontaktphase (Klärung der Rahmenbedingungen, Erwartungen, Zielsetzung, Vorgehensweise),
2. Diagnosephase (Datensammlung, Problemanalyse, Bewusstsein schaffen),
3. Planungsphase (individuelle Maßnahme auf Grundlage von Diagnosedaten),
4. Durchführungsphase (Selbstverpflichtung der Teilnehmer*innen, Transfer) und
5. Evaluation (Nacherfassung) (nach Kauffeld, 2001).

Diese Vorgehensweise ermöglicht es, die reflektierte Perspektive und Wahrnehmung der einzelnen Teammitglieder zu erheben und zu integrieren, Ergebnisse transparent darzustellen und gemeinsam zu bearbeiten, daraus Ansatzpunkte für gewünschte Veränderungen zu generieren und eine Selbstverpflichtung gegenüber konkreten Umsetzungsmaßnahmen sicherzustellen sowie den Entwicklungsfortschritt zu begleiten und ggf. nachzusteuern.

Es wird deutlich, dass es sich bei Teamentwicklung also nicht um eine Einzelmaßnahme handelt (Kauffeld & Lehmann-Willenbrock, 2008), sondern um einen sich kontinuierlich fortschreibenden dynamischen Prozess. Im Rahmen eines Teamentwicklungsprozesses kann dabei die Einleitung jeder Phase als Einzelmaßnahme begriffen werden, so dass es irreführend wäre, einen einzelnen Team-Workshop oder Teamtag mit einer Gesamtteamentwicklung gleichsetzen zu wollen.

Besonders der Diagnosephase wird eine unabdingbare, zentrale Rolle zugesprochen: Sie gilt als notwendige Voraussetzung, um überhaupt Maßnahmen ableiten zu können, die individuell zugeschnitten erfolgversprechend für die Weiterentwicklung des Teams wirken können (Kauffeld & Lehmann-Willenbrock, 2008). „Die Diagnose kann als Teil des Prozesses der Teamentwicklung selbst verstanden und hier sinnvoll realisiert werden“ (Kauffeld, 2001, S. 34). Um also überhaupt von einer Teammaßnahme sprechen zu können, die eine ENTWICKLUNGsmaßnahme sein soll, muss eine Diagnose als Basis, also als Dreh- und Angelpunkt, inbegriffen sein.

Neben den Inhalten, einer ausführlichen Diagnostik und individuellen, passenden Maßnahmen, spielen auch die Rahmenbedingungen einer Teamentwicklungsmaßnahme eine ausschlaggebende Rolle. Ein vertraulicher Kontext ist notwendig, um den Teammitgliedern eine offene Kommunikation zu ermöglichen, ein Verständnis für die Herausforderungen und Chancen im Team zu entwickeln und Selbstanteile zu reflektieren. Dabei nimmt das Konstrukt der Selbstentwicklungskompetenz bzw. Selbstmanagementkompetenz, welches u. a. von Graf (2019) beschrieben wurde, eine Schlüsselfunktion ein. Unter Selbstentwicklungskompetenz wird die Bereitschaft und Fähigkeit verstanden, das eigene Verhalten zu reflektieren, Stärken und Grenzen realistisch einzuschätzen sowie sich persönlich und fachlich zu motivieren und weiterzuentwickeln. Für Teamentwicklungsmaßnahmen ist es entscheidend, dass jedes Teammitglied grundsätzlich Bereitschaft zu Entwicklung und damit Veränderung sowie zur Übernahme von Selbstverantwortung zeigt. Ziel ist es, dass Teammitglieder im Rahmen der Teamentwicklungsmaßnahme ein Bewusstsein für das eigene Verhalten und mögliche Kritikpunkte entwickeln sowie persönliche Ziele, Werte und Wünsche reflektieren.

Im Kontext von Teamentwicklungsmaßnahmen existieren nicht nur viele ver-

schiedene Definitionen, sondern auch unterschiedliche Kategorisierungen und Bezeichnungen. Einige davon sollen an dieser Stelle aufgezählt und kurz beschrieben werden (vgl. Kauffeld, 2001):

→ Team-Designing: Strukturelle Interventionsebene mit Fokus auf Teammerkmalen wie Fähigkeiten, Größe, Ressourcen, Ziele, die so gestaltet werden, dass Teameffektivität optimiert wird

→ Team-Building: Teamentwicklung als gruppenbezogenes Interventionsbündel mit Fokus auf Teaminteraktionen und Prozessen (sozial und aufgabenbezogen) z.B. Rollen- und Zielklarheit, interpersonale / Konfliktlösungsziele, generelle Problemlösungsansätze

→ Team-Coaching: Umgrenzte Arbeit an beschriebenen Fragestellungen, die bereits definierte Probleme aufgreifen, oder etwas Spezielles aufarbeiten oder daran anschließen; kann auch eine Einzelmaßnahme sein, evtl. Folgemaßnahme bei vorhandenem Problembewusstsein.

Wenn wir hier von Teamentwicklungsmaßnahmen sprechen und Einsatzmöglichkeiten unseres Diagnose-Fragebogens thematisieren, beziehen wir uns hauptsächlich auf Team-Building im oben definierten Sinne.

Neben verschiedenen Arten von Interventionen, die beim Team ansetzen, unterscheiden sich natürlich auch die Anlässe, die beim Team zum Wunsch nach gezielten Steuerungsmaßnahmen führen.

Anlässe für Teamentwicklungsmaßnahmen

Häufig geht es, wenn Teamentwicklungsmaßnahmen ins Gespräch kommen, um Probleme bei sozialen und aufgabenbezogenen Prozessen oder interaktionsspezifische Schwierigkeiten (Rollen, Hemmnisse, zwischenmenschliche Beziehungen...). Zu den typischen Szenarien, die Anlass für Teamentwicklungsmaßnahmen bieten können, zählen u. a. die folgenden:

→ Es fehlen Regeln und Strukturen wie eine eindeutige Rollenverteilung oder klare gemeinsame Ziele.

→ Mangelnde Kommunikation führt zu Missverständnissen oder Konflikten, die die Zusammenarbeit erschweren, die Effektivität des Teams verringern und sich negativ auf die Zufriedenheit der Teammitglieder auswirken.

→ Konkrete unbearbeitete Konflikte spalten das Team, blockieren den offenen, ehrlichen Umgang, überschatten Produktivität.

→ Negative gruppendynamische Prozesse laufen ab, derer sich die Teammitglieder nicht bewusst sind und die sie folglich auch nicht beeinflussen können.

→ Es fehlen unterstützende Techniken wie Moderation oder Metaplan für eine effiziente Teamarbeit.

→ Innerbetriebliche Umstrukturierungen oder Fusionen führen zur Neubildung von Teams, die in kurzer Zeit gut zusammenarbeiten müssen.

→ Eine Neuzusammensetzung eines Teams bringt es mit sich, dass Rollen und Umgangsformen sich erst finden müssen und (Selbst-)Unsicherheit abgebaut werden muss.

→ Veränderungen diverser Art bringen Irritationen mit sich (Personalwechsel, Standortwechsel, Aufgabenspektrum, Homeoffice, Digitalisierung etc.) – aktuell insbesondere: Virtuelle Teams sind schnell entstanden, was Ängste und Unsicherheiten schürt und somit das Team aus der gewohnten Routine und Performanz herauswirft.

→ Eine Rückbesinnung (Auffrischung) auf bereits Erarbeitetes (z. B. Umgangsregeln) scheint angebracht, z.B. weil die letzte Entwicklungsmaßnahme schon länger her ist, nicht „gefruchtet“ hat, zu wenig nachgehalten wurde – hier muss auch ermittelt werden, warum das Erarbeitete wieder „untergegangen“ ist.

→ Der Umgang mit Diversität (z. B. Alter/Generationen, Lebensphasen, Nationalitäten, Geschlechter, Behinderung, kontrastreiche Charaktere, Unterschiede in der Leistungsfähigkeit) erzeugt bisweilen zu viele Reibungspunkte.

→ Unbekannte oder zumindest (noch) nicht bewusst gemachte Gründe für Schwierigkeiten im Team treten auf.

Diese Auflistung möglicher Anlässe ist weder vollständig noch trennscharf. Verschiedene Anlässe bedingen sich gegenseitig oder gehen ineinander über. Unabhängig vom Anlass ist es für eine Teamentwicklungsmaßnahme entscheidend, dass sie sich an den Anforderungen und Bedürfnissen des Teams orientiert und die Teammitglieder weder über- noch unterfordert (Kauffeld & Lehmann-Willenbrock, 2008).

Hilfreich zur Kategorisierung von Anlässen bezogen auf unterschiedliche Teambedürfnislagen ist die Unterscheidung von West (1994). Er beschreibt:

→ Team Start-ups: Hierbei geht es in der Anfangsphase eines Teams um eine präventive Intervention, in der eine Klärung von Zielen, Vorgehen, Rollen, Strategien erfolgt.

→ Regular formal reviews: Hierbei findet eine reflexive Betrachtung des Teamgeschehens aus der Metaebene im Rahmen von einem Workshop, z.B. nach Abschluss eines Projektes, statt.

→ Adressing known task-related problems: Hierbei geht es darum, ein umrissenes und bekanntes Problem punktuell zu behandeln.

→ Identifying what the problems are: Dabei geht es um die Diagnose von (z.B. aufgabenbezogenen) Problemen, bei denen die eigentlichen Ursachen unbekannt sind.

→ Social process interventions: Hierbei liegt der Schwerpunkt auf den interpersonalen Beziehungen, der sozialen Unterstützung und dem Teamklima. Ziel ist die Förderung eines positiven sozialen Klimas und des Wohlbefindens der Teammitglieder.

Nachdem wir nun sowohl Erkennungsmerkmale von Teams, Entwicklungsprozesse und Ansätze zu deren gezielter Steuerung sowie Anlässe für Teaminterventionen skizziert haben, möchten wir diese Basis nutzen, um Sinn, Zweck und Inhalt des von uns konstruierten Team-Diagnose-Instrumentes zu erläutern.

Der neue Team-Diagnose-Fragebogen

Vorweg muss gesagt werden, dass der Fragebogen sich zum Zeitpunkt des Verfassens dieses Artikels noch in der Entwicklung befindet, so dass nachfolgende Erläuterungen zu Inhalten und zur Umsetzung vorläufig zu verstehen und als Soll-Planungsschritte formuliert sind.

Ausgangspunkt und Anspruch

Ausgangspunkt für die Entwicklung eines umfassenden, strukturierten Team-Diagnose-Fragebogens war die Erkenntnis, dass Unterstützungsbedarfe von Teams sich sehr häufig nicht mit einem einmaligen Teamtag oder Team-Workshop abdecken lassen, auch wenn solche punktuellen „event-haften“ Angebote sehr verbreitet sind und auch dem laienhaften Verständnis von „Team-Building“ sehr nahe kommen. Solche Events mögen mehr Spaß bringen als eingängige, kritische Selbstreflexion und ernsthaftes Commitment gegenüber notwendigen Veränderungen, die anstrengend sind und Selbstverantwortung sowie ausdauerndes „am Ball bleiben“ erfordern.

Wenn ein Team jedoch ernsthaft daran interessiert ist, bestehende Schwierigkeiten und Blockaden aufzulösen und sich nachhaltig zu optimieren und voranzubringen, steht eine eingängige und sorgfältige Analyse des Teams in seiner Gesamtkomplexität unabdingbar auf dem Plan und erfordert Systematik und (psychologische) Expertise im Kontext Team.

Wir verorten grundständige Teamentwicklungsbedarfe auf diesem Level und verbinden dies mit dem Anspruch, Teams individuell zu begreifen, sich selbst begreifbar zu machen und von da aus angestoßene Weiterentwicklungsprozesse auch längerfristig zu begleiten. Deshalb haben wir nun ein Instrument entwickelt, das unser Vorgehen und die Qualität unserer Arbeit mit Teams auf ein fachtheoretisches Fundament stellt und einen Praxis-Standard für bestimmte Einsatzzwecke etablieren soll.

Bei dem Fragebogen handelt es sich um ein struktur- und prozessanalytisches Diagnoseinstrument, das von allen Teammitgliedern einzeln bearbeitet wird. Hauptziel ist es, Aufdeckungsarbeit zu leisten, indem das Team sich auf verschiedenen Dimensionen hinsichtlich unterschiedlicher Facetten des Team-Seins selber reflektiert und einschätzt. Angelehnt an die Diagnose als Dreh- und Angelpunkt von Teamentwicklungsmaßnahmen gehen wir soweit zu behaupten, dass die Diagnosearbeit, die mit dem Fragebogen geleistet wird, bereits den Hauptteil der Gesamtintervention ausmacht. Schon das Ausfüllen durch das Team ruft Erkenntnisgewinne hervor. Diese können dann in der Auswertung mit dem und durch das Team – welche in der Durchführung eines festgelegten Teamtags / Team-Workshops erfolgt – in die gezielte Ableitung von Veränderungsmaßnahmen münden.

Auch die anfängliche Kontaktphase wird mit dem Einsatz des Fragebogens initiiert, da das Team von Beginn an als Ganzes eingebunden, transparent informiert und in die Selbstverantwortung genommen wird. Zudem können der mit dem Fragebogen erhobene Ist-Stand und die Auswertungsergebnisse zu einem späteren Zeitpunkt der Evaluation abgeleiteter und umgesetzter Maßnahmen dienen, indem auf die Veränderung der bedeutsamen Parameter geschaut wird. Darüber hinaus dient der Einsatz des Fragebogens auch der Kontrolle des ursprünglichen Auftrages für eine Teamentwicklungsmaßnahme. Möglicherweise ergeben sich aus der Teamdiagnose ganz andere, sicher aber schärfer gezeichnete Optimierungsfelder als diejenigen, die vorab angenommen wurden oder nur vage formuliert werden konnten („Die Kommunikation könnte besser sein“).

Insgesamt erlaubt der Fragebogen also eine strukturierte, dennoch hochindividuelle Erfassung des Teams durch sich selbst und ermöglicht damit auch einen genauen Zuschnitt von Folgemaßnahmen auf die aufgedeckte „wahre“ Bedürfnislage. Selbstverständlich ist der Einsatz aber nur unter bestimmten Voraussetzungen effizient und maximal erfolgversprechend. Damit kommen wir zu einigen Rahmenbedingungen und Einschränkungen für den Einsatz.

Rahmenbedingungen für den Einsatz

Ein Einsatz des Fragebogens als umfassendes Diagnoseinstrument ist dann besonders sinnvoll, wenn es, angelehnt an die von West (1994) beschriebenen möglichen Anlässe, Schwierigkeiten im Team gibt, deren Wesen und Ursachen bisher nicht klar reflektiert und identifiziert sind (Identifying what the problems are) und um sich mit den interpersonalen Beziehungen zu beschäftigen, deren Charakteristika in der Zusammenschau der Selbst-Fremdbewertungen aller einzelnen Teammitglieder zum (sozialen) Teamgeschehen zum Vorschein treten (social process interventions).

Geht es hingegen eher um umschriebene aufgabenbezogene Optimierungsbedarfe und Zielfortschrittsüberprüfungen (regular formal reviews, addressing known task-related problems), eine Teamzusammenstellung nach Leistungseffizienzkriterien (Team-Designing) oder ein Coaching zu bereits klar umrissenen Problemstellungen, steht der umfassend aufdeckende, mehrdimensionale Charakter des Fragebogens nicht im rechten Verhältnis zur Zielsetzung. Möglicherweise können sich aber Anlässe für weitere umschriebene Maßnahmen durchaus aus der Auswertung des Fragebogens ergeben. Identifying what the problems are impliziert ja bereits, dass es nach dem Identifizieren mit dem Intervenieren weitergehen muss. Unsinnig wäre ein originärer Einsatz des Fragebogens auch, wenn es bei der angefragten Unterstützung des Teams gar nicht um eine grundständige Teamentwicklungsmaßnahme im hier definierten Sinne geht, sondern bloß um eine methodisch-fachliche Fortbildung, z. B. zur Zusammenarbeit in virtuellen Teams oder die Moderation eines einzelnen Arbeitsfortschritts, z. B. der Entwicklung von Kommunikationsregeln oder der Aufstellung eines Zuständigkeits-Organigramms.

Eine sorgfältige Auftragsklärung, aber auch Beratung des anfragenden Teams (oft zunächst der Teamleitung) hinsichtlich der tatsächlichen Bedürfnislage und entsprechend zweckmäßiger Interventionen (mit oder ohne Diagnose-Fragebogen) sind also unerlässlich. Neben dem was bedürfnisgerecht wäre, spielen natürlich auch finanzielle und zeitliche Begrenzungen in die Abwägung hinein, was getan werden sollte und was getan werden kann. Keinesfalls sollte ein Einsatz des Fragebogens „erzwungen“ werden oder nur in Teilen erfolgen, wenn definitiv klar ist, dass entsprechender Zeit- und Kostenaufwand nicht zu leisten sind. Er sollte auch dann nicht eingesetzt werden, wenn das Team mit der damit verbundenen Anstrengung und dem Verlassen der Komfortzone nicht bereit ist. Auch dies sollte bei der Auftragsklärung thematisiert werden – Änderungsbereitschaft, Vermögen zu Selbstkritik und Selbstverantwortung sowie ein grundständiges Maß an Offenheit und Ehrlichkeit im Umgang miteinander im Team müssen hier bereits eruiert werden. Auch muss Aufklärung darüber erfolgen, welche Aufwände vom Team erforderlich sind und welche Aufdeckung mit dem Fragebogen losgetreten (und dann nicht „zurück in die Tube gedrückt“) werden kann.

Diese Hinweise sind nicht als „Angstmache“ zu verstehen, vielmehr ist uns daran gelegen, dass Bedürfnisse, Motivation und Leistungsvermögen eines Teams ein bestmögliches Matching mit unserem Fragebogen eingehen und eine bestmögliche Hilfe zur Selbsthilfe angeboten werden kann.

An dieser Stelle ist nun allenfalls abstrakt der Mehrwert des Fragebogens bei bestimmten Anlässen abzulesen. Was fehlt, ist nun das Konkrete: Welche Inhalte hat der Fragebogen, und wie werden diese abgefragt? Wie soll er eingesetzt und ausgewertet werden?

Inhalt

Wie bereits angedeutet, haben wir uns bei der Entwicklung des Team-Diagnose-Fragebogens zum Ziel gesetzt, die oben beschriebenen etablierten Ansätze zur Bestimmung von Teammerkmalen, Teamentwicklung und deren Steuerung sowie Interventionsanlässen in ein Instrument zu integrieren, um eine möglichst umfassende Team-Diagnose vorzunehmen.

Entsprechend unserer Definition von Team und den hier eingeflossenen Ansätzen von Kauffeld, West und Tuckman, erfasst der Fragebogen:

→ Merkmale auf der Strukturebene, die die eher sachbezogene Selbstorganisation des Teams betreffen,

→ Merkmale auf der Ebene sozialer Prozesse, die sich auf die Verbundenheit untereinander und das persönliche Einlassen der Einzelnen auf das Team beziehen,

→ eine separate Ebene der Standortbestimmung, die vor allem umreißt, auf welcher Entwicklungsstufe (nach Tuckman) sich das Team aktuell verortet, welche Hauptstärken und Hauptoptimierungsfelder es bei sich erkennt und wie fähig es bereits ist, sich selber systematisch zu überprüfen, flexibel an Veränderungen anzupassen und mit Diversität konstruktiv umzugehen (angelehnt an die Teamreflexion nach West),

→ Abschließend wird auf der Ebene der Erwartungen an den Teamtag erfragt, was auf Grundlage der Fragebogenbearbeitung für positive Erwartungen als auch Befürchtungen bestehen und auch hinsichtlich der/des Psycholog*in der dgp, der/die die Maßnahme mit dem Team durchführt.

Tabelle 1 verdeutlicht die Struktur des Fragebogens nach den hier beschriebenen Ebenen, deren Dimensionen und Unterfacetten.

Erkennbar ist, dass der Ebene der sozialen Prozesse der größte Anteil am Fragebogen zukommt, da es sich hierbei um ein primär psychologisch-diagnostisches Instrument handelt, das Erleben und Verhalten im Sozialkontext Team abbildet. Wird hingehen hauptsächlich Struktur- und Organisationsoptimierung in den Fokus genommen und nach Zielerreichung und Leistungsfortschritt gefragt, sind betriebswirtschaftliche Instrumente womöglich aussagekräftiger. Davon abgesehen ist es gerade das zwischenmenschliche Geschehen, welches komplex und dynamisch verläuft und oft nicht regelhaft und rein sachlich reflektiert und Veränderungen zugeführt wird. Motive, Gewohnheiten, Ängste, Lernerfahrungen sitzen hier tief verankert und verbinden das persönliche eng mit dem professionellen Miteinander. Möchte man hier keinen Aufruhr erzeugen oder/und sich lieber nur an der Sachebene /den Arbeitsinhalten entlangarbeiten, kann man es bei einem Team-Spaß-Tag und Zahlenanalysen belassen.

Da, wie bereits verdeutlicht, der Fragebogen nur dann zum Einsatz kommen sollte, wenn die psychosoziale Komplexität eines individuellen Teamgeschehens gründlich durchleuchtet werden soll, spannt die Ebene der sozialen Prozesse mehrere Dimensionen auf. Zum Ersten geht es um die gefühlte Verbundenheit und Identifikation mit dem Team, zum Zweiten um eine geteilte soziale Werteorientierung, die auch im gelebten Umgang miteinander Ausdruck findet und zum Dritten um die Kommunikationskultur, insbesondere beim Umgang mit Konflikten und Kritik.

Angelehnt an die Annahme von Entwicklungsphasen (Tuckman) oder hierarchischen Stufen (Kasseler Teampyramide) der Teamentwicklung, sollte nicht übersehen werden, dass das soziale Zusammenwachsen und Sich-Reiben im Team auf dem Fundament der gemeinsamen Zielausrichtung, Zuständigkeitsverteilung und Wissens- und Informationsaustausch aufbaut. Mängel auf dieser Ebene wie Ineffizienz, Unklarheit sowie Intransparenz finden dann auch Ausdruck im Miteinander. Aus diesem Grund werden diese Merkmale bzw. deren Qualität auch im Fragebogen erfasst (Strukturebene).

Ebene	Dimensionen	Unterfacetten
Merkmalsebene: Struktur	Zielorientierung	Klarheit, Zielsetzung (SMART), Identifikation mit Zielen
	Aufgabenverteilung / Zuständigkeiten	Transparenz, Verbindlichkeit
	Wissens- und Informationsaustausch	Effizienz der Formate
Merkmalsebene: Soziale Prozesse	Verantwortungsübernahme	Bereitschaft, Fairness
	Identifikation	Wir-Gefühl
	Soziale Werteorientierung	Gegenseitige Unterstützung, Offenheit / Vertrauen, Loyalität, Wertebewusstsein
	Konfliktumgang	Verdeckte Konflikte, Offene Konflikte / Impulsivität, Bewältigungsverhalten / -strategien
	Kritikkultur	Kritisches Feedback geben und nehmen
	Kommunikationskultur	Kanäle (persönlich vs. Email), Gleichberechtigung & Wertschätzung
Ebene der Standort-bestimmung	Performanz / Reife	Entwicklungsphasen nach Tuckman
	Rollenklarheit	
	Besondere Stärken	
	Hauptoptimierungsfelder	
	Teamreflexion (nach West)	Systematische Selbstkontrolle, Anpassung an Veränderungen, Synergien durch Diversität
Ebene der Erwartungen an den Teamtag	Positiv	
	Negativ	
	Kernthema	
	Psycholog*in	

Tabelle 1: Struktur des Team-Diagnose-Fragebogens

Gehen wir weiter von der Facetten-Ebene zu den konkreten Items, sollen jeder Unterfacette vier Items zukommen – zwei davon sind Selbstaussagen, die auf einer 5er Skala einzuschätzen sind, zwei davon sind offene Fragen, die dazu dienen, die subjektiven Einschätzungen zu reflektieren und zu spezifizieren. Sie regen also dazu an, darüber nachzudenken, warum eine bestimmte Einschätzung spontan so vorgenommen wurde. Zudem beziehen sich die Fragen je zur Hälfte auf die Wahrnehmung des Teams als Ganzes und die Selbstwahrnehmung des einzelnen Teammitglieds in Bezug zum Team. Das nachfolgende Beispiel zur Unterfacette „Bereitschaft zur Verantwortungsübernahme“ verdeutlicht diese Struktur.

Item-Beispiele:

In unserem Team sind alle gleichermaßen bereit, Verantwortung zu übernehmen. Einschätzung (1= stimme gar nicht zu; 2 = stimme weitgehend nicht zu; 3 = teils, teils; 4 = stimme weitgehend zu; 5 = stimme absolut zu)

Meine eigene Bereitschaft zur Verantwortungsübernahme in unserem Team ist sehr hoch. Einschätzung (1= stimme gar nicht zu; 2 = stimme weitgehend nicht zu; 3 = teils, teils; 4 = stimme weitgehend zu; 5 = stimme absolut zu)

In welchen Situationen haben Sie den Eindruck, dass einzelne Teammitglieder vermeiden, Verantwortung zu übernehmen? Woran kann das liegen? (offen)

In welchen Situationen scheuen Sie sich selbst davor, Verantwortung im Team zu übernehmen und warum? (offen)

Neben dieser Fragebogenversion für das Gesamtteam soll es eine gesonderte Version für die Teamleitung geben, die dieselben Ebenen, Dimensionen und Facetten sowie auch Einschätzungs- und offene Fragen beinhaltet. Die Fragen sind aus dem Blickwinkel der Führungskraft auf ihr Team formuliert und können in Zusammenschau mit den Antworten des Teams interessante Selbst-Fremd- Ansichten im hierarchischen Gefüge vermitteln. Nachfolgend ein Beispiel für ein Item-Set aus dieser Version (Unterfacette: „Wertebewusstsein“):

Item-Beispiele für Teamleitung:

Ich verstehe mich als Vorbild für mein Team, wenn es darum geht, geteilte Werte auch bewusst zu leben. Einschätzung (1= stimme gar nicht zu; 2 = stimme weitgehend nicht zu; 3 = teils, teils; 4 = stimme weitgehend zu; 5 = stimme absolut zu)

Wann kann es vorkommen, dass Sie als Vorbild für Ihr Team versagen, wenn es darum geht, Werte auch vorzuleben, die Sie selbst predigen? (offen)

Einsatz

Der Fragebogen soll im Digital-Format über die Test-Plattform der dgp zur Verfügung stehen. Die Einladung an die einzelnen Teammitglieder, per Zugangslink am Fragebogen teilzunehmen, soll daneben eine Selbstvorstellung des/der Psycholog*in enthalten, der/die mit dem Team arbeiten wird sowie einige Informationen zum Zweck des Einsatzes im Rahmen der beauftragten Teamentwicklungsmaßnahme (ganz im Sinne der 1. Phase einer Teamentwicklungsmaßnahme: Kontaktaufnahme).

Da der Fragebogen relativ umfangreich ist (nach jetzigem Stand ca. 50 Einschätzungs-Items und ca. 50-60 offene Items enthalten wird; die Version für die Teamleitung ca. halb so viele), ist es sehr wichtig, dem Team dessen Bedeutsamkeit als eigentlichem Hauptteil der Intervention zu verdeutlichen und damit an die Vollständigkeit der Bearbeitung durch alle Teammitglieder zu appellieren. Hier wird wieder deutlich, warum es unerlässlich ist, den Auftrag und den möglichen Einsatz des Fragebogens vorab gründlich zu besprechen und sofern das Gespräch nur mit der Teamleitung erfolgt, dafür Sorge zu tragen, dass das Team von Anfang an einbezogen und nicht plötzlich vom Fragebogen „überrascht“ wird.

Damit die Auseinandersetzung mit dem Fragebogen neben der Arbeit Akzeptanz findet und machbar ist, sollte zum einen die Einladungsmail rechtzeitig, aber auch nicht zu früh versendet werden (wir schlagen vier Wochen vor dem angesetzten Teamtag vor), andererseits sollte es möglich sein, den Fragebogen auch mit Unterbrechungen auszufüllen (Zwischenspeicherung von Daten). Allerdings sollten zwei Wochen vor dem Teamtag alle Ergebnisse vorliegen, denn der/die Psycholog*in muss diese noch zusammenfassend aufbereiten und eine erste Auswertung vornehmen, um darauf die Moderation am Teamtag und die Behandlung von sich abzeichnenden „identified problems“ aufzubauen.

Auswertung

Der Fragebogen dient der individuellen Selbsteinschätzung eines Teams anhand vorgegebener Fragen. Während die Antworten auf die Einschätzungsfragen summarisch zusammengefasst, grafisch abgebildet und Durchschnittswerte errechnet werden können, müssen die freien Antworten auch nach weitgehend freiem Vorgehen inhaltlich verdichtet werden. Diese Ergebnisaufbereitung und erste Interpretation liegt bei dem/der durchführenden Psycholog*in.

In diesem Aufbereitungsschritt geht es darum, die Antworten aller Teammitglieder zwar anonymisiert zusammengefasst darzustellen, aber die Inhalte möglichst pro Item unverfälscht an das Team zurück zu geben. Dabei stellt sich die noch offene Frage, wieviel Komprimierung hier aus praktischen Gründen und in Hinblick auf die Verarbeitungskapazität der Teammitglieder angezeigt ist, vor allem wenn das Team recht groß ist und sehr ausführliche Antworten auf die offenen Items erfolgen. Die Handhabung muss diesbezüglich noch im Rahmen von Testläufen geklärt werden. Möglicherweise lassen sich einige offene Item-Antworten doch recht schematisch darstellen (z. B. wie die einzelnen Teammitglieder ihre eigene Rolle im Team und die Rollen der anderen beschreiben würden) oder Auszählungen hinsichtlich der Häufigkeit ähnlicher Antworten treffen.

In Bezug auf die Einschätzungs-Items wird eine nutzbare Erkenntnis die Streuung der Antworten im Team sein, wobei Extremwerte besonders spannend sind und wenn diese in Bezug auf ein Item auch noch an den verschiedenen Polen auftauchen. Beispielsweise könnten einige die Teamziele für absolut klar halten und einige dem gar nicht zustimmen. Dabei werden auch Vergleiche zwischen den Gesamtteameinschätzungen und den Selbst-im-Team-Einschätzungen aufschlussreich sein, insbesondere bei Abweichungen. Beispielsweise könnten die meisten im Team meinen, sie selbst identifizierten sich stark mit dem Team, das Gesamtteam jedoch als eine Einheit mit starkem Wir-Gefühl wird aber eher angezweifelt.

In Bezug auf die offenen Antworten sind häufige Übereinstimmungen, was sowohl Schwierigkeiten als auch Stärken angeht, aussagekräftig in Bezug auf die Veränderungsbedarfe und Entwicklungsreife des Teams. Aber auch Abweichungen und Ausreißer deuten auf Teamdynamiken hin, z. B. Untergruppenbildung und nicht geteilte Perspektiven.

Letztlich ist der Fragebogen als exploratives, aufdeckendes Instrument zu verstehen, mit dem eine umfassende Selbstauseinandersetzung des Teams angeregt wird.

Dem/der Psycholog*in werden anhand der Antworten (und der eigenen Vorerfahrungen mit Teams) Hypothesen darüber ermöglicht, was in diesem Team im Kern los ist. Diese sollen wiederum in die Arbeit mit dem Team bei dessen eigener Bewertung der Ergebnisse steuernd einfließen. Es kommt also darauf an, was das Team selbst aus den Ergebnissen machen kann und will, aber auch darauf, dass der/die Psycholog*in den Rahmen dafür schafft und dazu bewegt, dass vor allem deutlich hervorstechende Diskrepanzen und Änderungswünsche am Teamtag nicht unangesprochen und unbehandelt bleiben.

Der Fragebogen wird also immer individuell und frei in Zusammenarbeit zwischen Psycholog*in und Team ausgewertet, wobei der Prozess des Ausfüllens und Auswertens an sich den Hauptteil der Teamentwicklungsmaßnahme ausmacht. Selbstreflexion, Auseinandersetzung, Erkenntnis und Austausch sind angeregt worden. Was als Nächstes passieren soll (Veränderungsmaßnahmen), kann im Rahmen dieser einen Intervention zumeist nur noch angerissen werden – sofern es nur diesen einen Termin für einen Teamtag gibt.

Ausblick

Unsere Erfahrungen in der Arbeit mit Teams, eine theoretische Aufarbeitung und Integration von Ideen zur strukturierten und umfassenden Diagnose von Teams sind nun in die Entwicklung eines Fragebogens zur Team-Selbsteinschätzung geflossen. Ab jetzt wird es darum gehen, die bestehende Entwurfsversion des Fragebogens fachlich zu diskutieren und auf Mehrwert und Praxistauglichkeit hin zu testen. Funktionierende Teamarbeit gewinnt vor allem unter VUCA-Arbeitsbedingungen an Gewicht. Gleichzeitig vergrößern und vervielfältigen sich dabei Reibungsflächen. Wir halten deshalb die Möglichkeit, dgp-seitig eine passende, grundständige und systematische Selbst-Diagnose für Teams vorhalten und begleiten zu können, für äußerst bedarfs- und zukunftsorientiert.

Literatur:

Graf, A. (2019). Selbstmanagementkompetenz in Organisationen stärken. Springer Fachmedien.

Ilgen, D. R., Hollenbeck, J. R., Jonson, M. & Jundt, D. (2005). Teams in organizations: From Input-Process-Output models to IMOI models. Annual Review of Psychology, 56, 517-543.

Kauffeld, S. (2001). Teamdiagnose. Verlag für Angewandte Psychologie.

Kauffeld, S., & Lehmann-Willenbrock, N. (2008). Teamdiagnose und Teamentwicklung. In Erfolgreiche Gruppenarbeit (S. 29-41). Gabler.

Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. Journal of organizational Behavior, 13(2), 103-123.

Riketta, M., & Van Dick, R. (2005). Foci of attachment in organizations: A meta-analysis comparison of the strength and correlates of work-group versus organizational commitment and identification. Journal of Vocational Behavior, 67, 490-510.

Tajfel, H. (1974). Social identity and intergroup behaviour. Information (International Social Science Council), 13(2), 65-93.

van Dick, R., & Schuh, S. (2016). Führung von Gruppenprozessen: Identität und Identifikation bei den Mitarbeitern stiften. In Handbuch Mitarbeiterführung (S. 41-52). Springer.

West, M. A. (1994). Effective Teamwork. BPC Wheatons.

West, M. A. (2012). Effective teamwork: Practical lessons from organizational research. John Wiley & Sons.

Kontakt:

Isabel Lorenz
M.A. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
lorenz@dgp.de

Dr. Sabrina Köpke
M. Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
koepke@dgp.de

Sprachanalyse mit künstlicher Intelligenz zur Eignungsdiagnostik

Johannes Fesefeldt

Viel Lärm um nichts?
Im Jahr 2018 sorgte der deutsche, weltweit tätige Versicherungskonzern Talanx AG für Aufsehen.[1] Talanx berichtete davon, seine Führungskräfte mithilfe einer auf Künstlicher Intelligenz (KI) basierenden Sprachanalyse des Aachener Software-Anbieters Precire auszuwählen. Die Kandidat*innen erhielten dazu einen „Telefonanruf“ der KI und wurden gebeten, zum Beispiel von einem typischen Tagesablauf, von einem negativen beruflichen Erlebnis oder von persönlichen Gefühlen zu berichten. Aus dem so gewonnenen Audiomaterial versucht die KI, Rückschlüsse auf die Ausprägung von insgesamt elf zuvor definierten Persönlichkeitsmerkmalen zu ziehen. Diese wurden als besonders relevant für die Tätigkeit als Führungskraft eingeschätzt, z.B. „Zielorientierung“, „Dominanzstreben“ oder „Statusbewusstsein“.

[1] Vgl. den Online-Artikel von Bös und Marx (2019) in der Frankfurter Allgemeinen Zeitung oder den Online-Artikel von Tichy (2021) in der Personalwirtschaft (s. Literaturverzeichnis).

Aufregend erschien nicht nur die Möglichkeit, mithilfe des KI-Algorithmus Audio- und Textmaterial psychologisch zu analysieren, sondern auch, dass der Algorithmus Personalexpert*innen, vor allem Psycholog*innen, ersetzen könnte.

Talanx berichtete von einer etwa neunzigprozentigen Übereinstimmung der KI-Einschätzungen mit Expert*innenurteilen. Das traditionelle eignungsdiagnostische Vorgehen mithilfe eines Assessment-Centers schien deswegen überflüssig geworden zu sein (vgl. Fesefeldt, 2018).[2]

Im Jahr darauf, in 2019, erschien dann im Report Psychologie (die Zeitschrift des Bundesverbandes deutscher Psycholog*innen) die erste wissenschaftliche Rezension des Verfahrens. Das Gesamturteil fiel ernüchternd aus, da es große Kritikpunkte gab. So bemängelten die Autoren Schmidt-Atzert, Kühnemann und Zimmermann die lückenhafte Dokumentation der Stichprobe, an der das Verfahren normiert wurde, teilweise fehlende Norm- bzw. Vergleichswerte, das Fehlen einer echten Theorie hinter dem Verfahren, die Gefahr der Verfälschung sowie mangelnde Belege für seine Vorhersagekraft („prognostische Validität") in Bezug auf den späteren Berufserfolg.

Trotzdem machten die von Precire berichteten Daten durchaus Hoffnung, dass die Messung von Persönlichkeitsdimensionen per KI-Analyse möglich ist: Denn (wie auch schon in Studien zu anderen automatisierten Sprachanalysen) es zeigten sich in den Daten recht hohe Zusammenhänge zwischen den per KI ermittelten Faktoren und der Persönlichkeit der Kandidat*innen. Konkret wurden dabei die Skalen des sog. „Hexaco-Persönlichkeitsfragebogens" herangezogen, der sich auf das in der Persönlichkeitsforschung weit verbreitete Modell der „Big Five" (Costa & McCrae) stützt.[3] Die Logik dahinter ist simpel: Gelingt es, die Ausprägung der fünf grundlegengen Persönlichkeitsdimensionen mithilfe des Algorithmus einzuschätzen, können weitgehende Aussagen über die Persönlichkeitsstruktur eines Menschen getroffen werden. Und auch eine Studie von Schaumlöffel, Hübner, Thile & Stulle (2018) bestätigte die substanziellen Korrelationen zwischen den algorithmisch erfassten Faktoren und den Hexaco-Skalen (im Bereich von -0.66 bis maximal +0.49). Trotz der berechtigten wissenschaftlichen Kritik erscheint der Ansatz der Sprachanalyse per KI damit auf den ersten Blick durchaus sinnvoll. Auf den zweiten Blick zeigen sich jedoch eine Reihe von weiteren, grundsätzlichen Problemen. Diese werden im nächsten Abschnitt diskutiert.

Wissenschaftliche Kritik und Schwächen der KI-Sprachanalyse

Wie der Professor für Arbeitspsychologie Kanning (2018) kritisch anmerkt, fehlt bislang vor allem eine unabhängige Validierung der Ergebnisse des geschilderten Verfahrens. Es ist also unklar, ob sich die KI-Sprachanalyse auch an einem Datensatz bewähren würde, der nicht vom Anbieter der KI selbst stammt. Insofern sind die berichteten Korrelationen mit Vorsicht zu genießen. Wie Kanning in seinem Vortrag „Künstliche Intelligenz in der Eignungsdiagnostik – ein kritischer Blick" auf dem dgp-Kongress „Personalarbeit ohne Personal?!" (2019) ausführte, gibt es zudem eine Reihe von fundamentalen Problemen.

[2] 2019 erhielt Precire in Europa und der Schweiz das Patent für "Sprachanalyse".

[3] Die fünf übergeordneten Dimensionen des Big-5-Modells sind: Offenheit für Erfahrungen, Gewissenhaftigkeit, Extraversion, Verträglichkeit und Neurotizismus. Für eine nähere Erklärung des Big-Five-Modells, seiner Dimensionen und Facetten siehe z.B. Fesefeldt (2017).

Hierzu ist zunächst daran zu erinnern, dass das Ziel der Eignungsdiagnostik nicht in der Messung von Persönlichkeitsmerkmalen besteht. Vielmehr stellt die „Eignung" ein allgemeines Kriterium dar, das die zukünftig erwartbare berufliche Leistung und den beruflichen Erfolg insgesamt betrifft. Die Persönlichkeit ist hierbei nur ein Einflussfaktor unter vielen (z.B. spielen offensichtlich auch Aspekte wie Motivation oder soziale Kompetenz wichtige Rollen für den Berufserfolg).

Aus dieser Perspektive ist zu sagen: Selbst, wenn die zuvor berichteten Zusammenhänge zwischen den Ergebnissen der Sprachanalyse und den Big Five allgemein zutreffen sollten, wäre auf dieser Basis trotzdem nur ein kleiner Bruchteil der beruflichen Eignung vorhersagbar. Nimmt man z.B. – ganz konkret und optimistisch – an, dass sich über eine Sprachanalyse 50% der Datenvarianz im Zielkriterium Persönlichkeit vorhersagen lassen, führt dies in Bezug auf das Kriterium beruflicher Erfolg trotzdem nur zu einer mageren Vorhersage von 5%. Denn: Selbst die Zusammenhänge besonders hervorstechender Dimensionen der Persönlichkeit (wie z.B. Gewissenhaftigkeit und Selbstwirksamkeit) belaufen sich allen Studienerkenntnissen zufolge auf höchstens ca. 10% (vgl. hierzu etwa Obermann, 2013). Dies bedeutet, dass die angenommenen 50% Vorhersage der Sprachanalyse durch ein maximal 10% großes „Nadelöhr" von 10% Relevanz der Persönlichkeit für das Zielkriterium führen. Am Ende bleibt damit nur sehr wenig an Vorhersagekraft übrig!

Obwohl auch ein vorhersageschwaches Instrument diagnostisch nützlich sein kann, sobald es nur irgendeinen Mehrwert an Vorhersagekraft bietet (dies ist bei Persönlichkeitstests durchaus der Fall – vgl. wiederum Obermann, 2013), bereiten weitere Punkte Sorge: Wie wir aus unserem alltäglichen Sprachgebrauch wissen, liegt es geradezu in der Natur von Sprache, diese flexibel an verschiedene Situationen und Adressat*innen anzupassen (vgl. etwa den klassischen Aufsatz von Goffman (1959) hierzu). Studien zeigen, dass unser Sprachstil stark von externen Ereignissen und unserer persönlichen Lebenslage beeinflusst wird (vgl. Pennebaker et al., 2003).

Dies führt zur Frage, wie die KI-Sprachanalyse mit den natürlichen Variationen von Sprache, aber auch mit künstlichen Variationen, d.h. mit gezielter Selbstdarstellung (sog. „sozial erwünschte" Antworten) umgeht.

Zwar stellen beide Aspekte eine methodische Herausforderung für jede Form von Auswahlverfahren dar, aber sie lassen sich vor allem im Rahmen eines Assessment- Centers (AC) besser kontrollieren. Beispielsweise kann eine als nicht authentisch empfundene Antwort im Interview gezielt hinterfragt werden, indem nach Belegen aus der Erfahrung der Kandidat*innen gefragt wird. Außerdem ergeben sich über die verschiedenen Stationen eines AC hinweg Informationen, die zueinander in Beziehung gesetzt werden können, um die Stabilität und Typizität von Kompetenzen zu belegen bzw. zu widerlegen. Beispielsweise kann die von einer Kandidatin im Interview behauptete „hohe Konfliktkompetenz" in einem konfliktär angelegten Rollenspiel direkt überprüft werden. Derartige Möglichkeiten einer „Kreuzvalidierung" stehen einer „mono-methodisch" arbeitenden (d.h. sich nur auf Sprachanalyse verlassenden) KI jedoch nicht zur Verfügung. Daher ist ihre Anfälligkeit für übermäßig positive Selbstdarstellung als hoch einzuschätzen. Die relevanten Sprachparameter können von dem/der Befragten ja einfach angepasst werden.

Eine Strategie der Selbstdarstellung könnte z.B. darin bestehen, häufig positiv klingende Worte wie „Zukunft“, „motiviert“, „Erfolg“ (mit „positiver emotionaler Valenz“) zu verwenden und Verneinungen (z.B. „nicht“, „kein/e“) sowie Füllwörter (z.B. „äh“ und „öhm“) zu vermeiden. Darüber hinaus könnte angestrebt werden, schnell zu sprechen und häufiger Pronomen der ersten Person singular zu nutzen („ich“, „mein“, „mir“, „mich“).[4] Durch diese kleinen Anpassungen des Stils würde das Gesprochene wesentlich tatkräftiger, optimistischer und ehrlicher wirken (vgl. Pennebaker et al., 2003). Rechtlich bedenklich an solchen Strategien ist, dass unehrliche Kandidat*innen sich mit ihnen leicht Vorteile verschaffen können. Der oder die Dumme ist dann der oder die Ehrliche. Das mit der Methode der „Selbstauskunft“ typischerweise verbundene Problem der sozialen Erwünschtheit verschärft sich im Kontext der KI-basierten Sprachanalyse also nochmals deutlich. [5]

Eine geringe Aussagekraft in Bezug auf die berufliche Eignung gepaart mit einer hohen Verfälschbarkeit stellt an sich eigentlich bereits den Knockout für ein eignungsdiagnostisches Instrument dar. Ein weiterer Nachteil lässt den Einsatz von KI-Sprachanalyse (zumindest in ihrer derzeitigen Form) jedoch noch unattraktiver erscheinen: Grundsätzlich bedeutet beinahe jedes Instrument zur Prüfung der beruflichen Eignung einen Eingriff in die Persönlichkeitssphäre eines Menschen. Allerdings wird dieser Eingriff durch das Erkenntnisinteresse des Arbeitgebers an der Eignung bis zu einem gewissen Grad gerechtfertigt. Der Eingriff liegt im Rahmen des Ansatzes der Persönlichkeitsdiagnostik ganz offensichtlich in besonderer Weise vor. Anbieter*innen von Tests stellen daher sog. „berufsbezogene Persönlichkeitsverfahren“ zur Verfügung, welche Fragen ausschließlich zu den vermeintlich beruflich relevanten Aspekten stellen. Trotzdem lässt sich naturgemäß auch bei dieser Herangehensweise ein starker Bezug der Fragen zu allgemeinen Merkmalen der Persönlichkeit erkennen. Daraus ergibt sich: Ein KI-Verfahren zur Sprachanalyse wie das geschilderte, das nur wenig Bezug zur beruflichen Eignung aufweist, anfällig für Verfälschungen ist und zu dem Preis eines starken Eingriffes in die Persönlichkeitssphäre der Kandidat*innen kommt, ist äußerst schwer zu rechtfertigen.

Ein weiterer gravierender Kritikpunkt betrifft den Vorwurf möglicher Diskriminierung: Durch die „Trainingsdaten“ (der Datensatz, mithilfe dessen das KI-System den Auswertungsalgorithmus erlernt) werden implizit Vorurteile eingespeist. Denn für die Optimierung des Algorithmus ist allein entscheidend, inwiefern die in den Daten steckenden Zielkriterien erfüllt werden. Die Zielkriterien können aber prinzipiell nicht völlig eigenständig von der KI aufgestellt werden.

Vielmehr müssen menschliche Expert*innen den Trainingsdatensatz „labeln“: Dies bedeutet, dass diese die Daten hinsichtlich des Erfüllungsgrades der Zielkriterien beurteilen müssen, z.B. danach, ob das Profil eines/einer Kandidat*in erfolgversprechend erscheint oder danach, ob ein/e Kandidat*in ein Assessment-Center bestanden hat. Für jeden „Datenpunkt“ im Trainingsdatensatz wird damit entschieden, wie stark sein Zusammenhang mit den Zielkriterien ist. Bestehende „Biases“, z.B. gegen einen bestimmten Ausbildungshintergrund, gegen ein bestimmtes Geschlecht, gegen einen bestimmten Dialekt oder gegen andere Merkmale werden dadurch in das System integriert. Besser gesagt: Sie werden auf versteckte Art und Weise Teil der Zielstellung des Systems.

Für Furore hat beispielweise die systematische Diskriminierung von Frauen durch eine KI-basierte Auswahl von Kandidat*innen anhand der Bewerbungsunterlagen bei Amazon gesorgt (s. z.B. Beck et al., 2019). Eine Diskriminierung in Bezug auf das Geschlecht kann im Rahmen der KI-Sprachanalyse z.B. drohen, weil deutliche Hinweise bestehen, dass Frauen und Männer tendenziell unterschiedliche Sprachstile pflegen (vgl. Pennebaker et. al., 2003).

In Summe muss leider festgestellt werden, dass die KI-basierte Sprachanalyse gleich mehrere zentrale Qualitätsstandards der Personalauswahl verletzt.

Einem Vergleich mit Anforderungen, wie sie z.B. in der Norm für Personalauswahl DIN 33430 formuliert sind, hält sie nicht stand (siehe z.B. den „DIN-Screen“ von Kersting, 2008). Da trotzdem ein gewisses Potenzial der Technologie zu erkennen ist und es sich um ein spannendes Gebiet handelt, soll im nächsten Abschnitt ein Blick darauf geworfen werden, worauf die Sprachverarbeitung generell basiert und welchen Herausforderungen sie dabei gegenübersteht.

Allgemeine Methoden und Herausforderungen der Sprachanalyse

Eine wichtige Unterscheidung ist die zwischen der Spracherkennung (engl. Automatic Speech Recognition, kurz ASR) und der Sprachverarbeitung (engl. Natural Language Processing, kurz NLP). Die traditionelle Spracherkennung beruht auf statistischen Analysen (entweder des Audio- oder des Textmaterials) und setzt u.a. ein akustisches Modell, ein Sprachmodell sowie ein „Wörterbuch“ voraus, mit dem die Signale verglichen und dann in computerlesbaren Text transkribiert werden. Die Herausforderung der Spracherkennung (ASR) liegt im Umgang mit unterschiedlichen Stimmen, akustischen Umgebungen, Störgeräuschen, aber auch dem Kontext. Linguistische Merkmale wie die Prosodie (z.B. ansteigende Betonung bei einer Frage), Füllwörter (wie „ähs“ und „öhms“) und andere sind in der Regel ebenfalls zu berücksichtigen (vgl. Liu & Whitaker, 2019).

Hingegen besteht bei der Verarbeitung der Sprache (NLP) das Ziel darin, den Sinn (des Gesprochenen oder Geschriebenen) zu erschließen. Ein NLP-System soll außerdem oftmals in der Lage sein, sinnvoll zu reagieren, d.h. mit dem Menschen in Kommunikation zu treten. Ein in der NLP häufig verwendetes auf den Sprachwissenschaftler Ferdinand de Saussure zurückgehendes Modell (das sog. „synchrone Sprachmodell“) unterscheidet folgende Kategorien bzw. Sprachmerkmale: „Morphologie“ (Form und Struktur der Wörter), „Lexikon“ (Zerlegung eines Textes in Worte), „Syntax“ (Regeln für Wörter im Satz), „Semantik“ (Bedeutung der Wörter), „Diskurs“ (Bedeutung zwischen Sätzen), „Pragmatik“ (Absicht hinter dem Satz), „Akustik“ (Repräsentation der Geräusche), „Phonetik“ (Erkennen von Sprachelementen), „Phoneme“ (Verarbeitung zu Sprache) und „Prosodie“ (Betonung, Tonhöhe, Rhythmus u.a.) – s. Liu & Whitaker (2019).

[4] Es ist ein recht hoher Zusammenhang der Sprechgeschwindigkeit mit Maßen der Intelligenz belegt (s. etwa Becker, 2016). Hinsichtlich der Anzahl von Personalpronomen der 1. Person Singular ist ein Zusammenhang mit der wahrgenommenen Authentizität belegt (je mehr, desto authentischer). Andererseits korreliert ein hoher Gebrauch der ersten Person mit der Big-Five-Dimension Neurotizismus, deren hohe Ausprägung auf mangelnde emotionale Stabilität eines Menschen hinweisen kann (vgl. Pennebaker et. al., 2003).

[5] Da die Kontrolle der sozialen Erwünschtheit, z.B. durch sog. „Lügenitems“ in einem Fragenbogen, bereits bei den herkömmlichen Persönlichkeitsverfahren problematisch ist, besteht auf den ersten Blick wenig Hoffnung, dass eine KI dies könnte. Wie die „Arbeitsgruppe IT-Sicherheit, Privacy, Recht und Ethik“ der 2017 vom Bundesministerium für Bildung und Forschung gegründeten Plattform „Lernende Systeme“ berichtet, stoßen Versuche, die sog. „Fairness“ der Behandlung verschiedener Gruppen ebenfalls zum Ziel der KI zu machen, leider (noch) an konzeptuelle und technische Grenzen (vgl. Beck et al. 2019).

Es drängt sich hier die Frage auf, inwiefern ein die natürliche Sprache verarbeitendes System Sprache wirklich verstehen kann. Der britische Mathematiker Alan Turing schlug mit seinem bekannten „Turing-Test“ (1950) vor, dies davon abhängig zu machen, wie oft es eine Maschine schafft, einem Menschen vorzugaukeln, ebenfalls ein Mensch zu sein. Konkret schlug Turing vor, dass eine Maschine dann als „intelligent“ betrachtet werden sollte, wenn bei einem „Gespräch“ zwischen ihr und einem Menschen eine Minimalchance der erfolgreichen Täuschung von ca. 30% besteht (in mind. 30% der Fälle würde die Maschine nicht entdeckt). Als eine frühe Umsetzung des Turing-Tests und ein gutes Beispiel hierfür kann das 1966 von Joseph Weizenbaum entwickelte Programm Eliza gesehen werden. In einer einfachen Version basiert Eliza auf einem Wörterbuch, welches nach Oberbegriffen und Synonymen strukturiert ist. Außerdem enthält es eine Sammlung von Phrasen zu verschiedenen Themengebieten. Betrachten Sie die folgende „Konversation“ zwischen Eliza und einem Patienten des Psychotherapeuten Carl Rogers (frei übersetzt nach Jurafsky und Martin, 2008, S.7):

Patient:
In gewisser Weise bist du wie mein Vater.

Eliza:
Welche Ähnlichkeit siehst du?

Patient:
Du bist nicht besonders aggressiv, aber ich glaube du willst nicht, dass ich das bemerke.

Eliza:
Warum glaubst du, dass ich nicht aggressiv bin?

Patient:
Du diskutierst nicht mit mir.

Eliza:
Warum glaubst du, dass ich nicht mit dir diskutiere?

Patient:
Du hast Angst vor mir.

Eliza:
Befriedigt es dich, zu glauben, dass ich Angst vor dir habe?

Eine Gegenposition zu Turing formuliert das Gedankenexperiment des „Chinesischen Zimmers“ des Philosophen John Searle (1980): Searle stellte sich vor, was passieren würde, wenn eine nur des Englischen mächtige Person in einem Zimmer isoliert würde. In das imaginierte Zimmer werden Kärtchen mit chinesischen Zeichen gereicht, welche die Person im Zimmerinneren mithilfe eines Regelbuches zu komplexen Zeichenfolgen kombiniert. Die so zusammengestellten Kärtchenfolgen werden dann nach draußen gereicht. Dort werden sie von Chinesisch-Sprecher*innen gelesen. Die zentrale Einsicht des Experimentes ist, dass nicht davon gesprochen werden sollte, dass die Person im Zimmer Chinesisch „versteht“, obwohl er bzw. sie ein Programm zur Verarbeitung von chinesischen Sätzen ausführt, das zu sinnvollen Sätzen führt. Eine weitere Pointe ist, dass die oft bemühte Analogie zwischen dem Computer und unserem Gehirn falsch sein könnte. Es könnte sich um eine völlig falsche Metapher handeln, die uns bezüglich des Verständnisses der Bedeutung der natürlichen Sprache, wie sie „in“ unserem Gehirn generiert und interpretiert wird, in die Irre führt.

Interessanterweise basiert die weit überwiegende Zahl der Verfahren der natürlichen Sprachverarbeitung jedoch inzwischen gerade auf Algorithmen, die vom menschlichen Gehirn inspiriert sind: Seit ca. 2006 hat sich der Ansatz des sog. „Deep Learning“ in „künstlichen neuronalen Netzwerken“ (kürzer als „neuronale Netze“ bezeichnet) als das vorherrschende Paradigma der Sprachverarbeitung etabliert (s. z.B. Liu & Whitaker, 2019). Die Deep-Learning-Technologie steckt auch hinter vielen bekannten Anwendungen wie dem Google Assistant, Apples Siri oder Amazons Alexa. Die Besonderheit des Deep Learning liegt darin, dass die sprachlichen Analysemerkmale in Trainingszyklen selbständig erlernt werden, d.h. sie werden dem Netz nicht per Modell vorgegeben. Entscheidend für das Training des Netzes ist die Einspeisung einer möglichst großen und reichhaltigen Datenbasis (an Audio- oder Textmaterial) sowie die Vorgabe von Zielkriterien.[6] Letztere sind in der Regel menschliche Bewertungen der Daten danach, ob z.B. Wörter einer bestimmten Kategorie oder grammatisch korrekte Sätze vorliegen (sog. „überwachtes Lernen“, insofern die Kriterien in den Beispieldatensatz eingegeben werden). Anhand dieser Vorgaben ist das Netz in der Lage, durch Selbstkorrekturen seines Algorithmus die Leistung immer weiter zu verbessern. In den unzähligen Schichten des Netzwerkes (in der Regel mehr als Tausende) entstehen somit automatisch immer bessere Repräsentationen sprachlicher Merkmale. Dabei gilt, dass die unteren Schichten des Netzes primitivere Merkmale repräsentieren, die oberen komplexere (vgl. Fesefeldt, 2018): Auf unterster Ebene wird also im Prinzip nur das Spektrogramm des eingehenden Audiosignals erfasst (bzw. das Lichtmuster eines optischen Signals). Danach folgen höhere Schichten, die gleichsam zur Erkennung von Buchstaben oder Phonemen dienen. Auf oberster Ebene werden schließlich gar vollständige Sätze repräsentiert. Auch der im ersten Abschnitt beschriebene Algorithmus von Precire setzt auf das tiefe Lernen in einem solchen künstlichen neuronalen Netzwerk.

Einige klassische Methoden der Sprachverarbeitung, die auch in neuronalen Netzen angewandt werden, sollen hier zudem kurz vorgestellt werden: Bei der Analyse eines Satzes ist es allgemein nötig, diesen in seine natürlichen Wortelemente zu zerlegen. Bei dieser sog. „Tokenization“ wird z.B. der Satz „Dies__ist__ein__Beispiel__für__die__

[6] Entsprechend große Audio- oder Textsammlungen, sog. „Korpora“ sind heutzutage z.T. kostenlos im Internet zugänglich, sodass die Netze an ihnen trainiert werden können.

[7] Für eine Übersicht inklusive weiterer wichtiger NLP-Methoden siehe Liu & Whitaker (2019), S. 90 ff.

Zerlegung_ _.“ entsprechend der „_“ zerlegt. Um ein einzelnes Wort als bestimmtes zu erkennen, wird vom sog. „Stemming“ Gebrauch gemacht. Hierbei wird der Wortstamm durch Erkennen von Prä- und Suffixen freigelegt, z.B. lässt sich das Wort „Zerlegung“ so in den Stamm „Zerleg-“ und das Suffix „-ung“ zerlegen. Unter „Lemmatization“ versteht man eine Technik, mit der Worte um ein sog. „Lemma“ gruppiert werden, das die Kernbedeutung der Gruppe wiedergibt. Beispielsweise können die Worte „zerlegt“, „Zerlegung“ und „Zerlegerin“ um „zerlegen“ gruppiert werden. Bei „Part-of-Speech-Tagging“ werden syntaktisch unterschiedliche Worttypen erkannt, sodass z.B. im Satz „Die Fledermaus flog davon.“ das Nomen „Fledermaus“, der Artikel „die“, das Verb „flog“ und das Adverb „davon“ erkannt und zugeordnet werden. „Named Entity Recognition“: Hier werden abstrakte Entitäten erkannt, z.B. Organisationen oder Orte: „Die Deutsche Gesellschaft für Personalwesen [Organisation] ist ein eingetragener Verein [Organisation] mit Hauptsitz in Hannover [Ort].“ Beim „Chunking“ werden Worte zu inhaltlichen Einheiten gebündelt, sodass z.B. der hier genannte Beispielsatz auf „dgp“, „ist“ und „e.V.“ reduziert wird. [7]

Trotz immer raffinierterer Methoden sieht sich die natürliche Sprachverarbeitung einer gewaltigen Aufgabe gegenüber. Denn unsere Sprache ist nicht einfach nur eine formal zerlegbare Struktur, sondern sie spiegelt die Gesamtheit unseres Menschseins wider.

Das berühmte Diktum des Philosophen Ludwig Wittgenstein „Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.“ macht diese enge Beziehung zwischen Welt, Geist und Sprache deutlich.[8] Sprache kann zu verschiedensten Zwecken und in verschiedenster Weise verwendet werden und ist insofern nicht einfach isolierbar. Wie etwa der KI-Forscher Jerry Kaplan (2017) ausführt, kommt es beim richtigen Sprachgebrauch häufig auf einen realen, physikalischen Bezugsrahmen an. Beispielsweise ist es nötig, die Umwelt zu kennen, wenn eine Person eine andere darum bittet, sich auf diesen und nicht auf jenen Stuhl zu setzen. Der Mathematiker und Philosoph Gottlob Frege hat diesen Aspekt bereits im 19. Jahrhundert klar herausgearbeitet, indem er die Unterscheidung zwischen Sinn und Bedeutung eines Ausdruckes vorschlug (Frege, 1892): In seinem bekannten Beispiel „Der Morgenstern ist der Abendstern“ ist zwar die „Bedeutung“ (für Frege der Bezugsgegenstand) der Ausdrücke „Morgenstern“ und „Abendstern“ dieselbe, nämlich der Planet Venus, jedoch ist ihr Sinn in Bezug auf die Erscheinung am Himmel verschieden.[9] Diese sinnhafte Mehrdeutigkeit der Sprache wird z.B. auch an dem Satz „I saw her duck“ deutlich (s. Jurafsky und Martin, 2008, S. 4). Ein weiteres Beispiel verdeutlicht zudem die zentrale Rolle von Pronomen für das Verständnis eines Satzes: „John ging zum Laden, um etwas Brot zu kaufen. Als er es hatte, fuhr er heim.“ (übersetzt aus Pennebaker, 2003, S. 570).

Viele weitere Herausforderungen kommen hinzu. Beispielweise zeichnet sich die gesprochene Sprache durch Sprechpausen, halbfertige Sätze und spontane Korrekturen des Gesprochenen aus. Eine weitere grundlegende Bedeutungsdimension der Sprache, die ebenfalls auf Frege zurückgeht, ist die Unterscheidung von Kraft und Gehalt: Wir können einen Satz mit unterschiedlicher Kraft sprechen, um z.B. zu befehlen oder eine Frage zu stellen. So haben die Sätze „Die Tür ist offen.“, „Die Tür ist offen(?)“ und „Die Tür ist offen(!)“ nach Frege denselben Gehalt, aber verschiedene Kraft (eine noch weiter ausgefeilte Theorie dessen, wie wir mit Sprache handeln, stellt die sog. „Sprechakttheorie“ der Sprachphilosophen Austin und Searle dar).[10] Wenn wir zudem noch an Sprachformen wie Ironie, Sarkasmus, Metaphern oder Poesie denken, werden der immense Reichtum und die Vielschichtigkeit der Sprache uns deutlich vor Augen geführt.

Fazit

In diesem Artikel wurden am Beispiel eines Auswahlverfahrens bei einer großen Kapitalgesellschaft Vor- und Nachteile von KI-basierten Sprachanalyseverfahren diskutiert. Als Ergebnis zeigte sich, dass die KI-Sprachanalyse mit grundsätzlichen Problemen zu kämpfen hat, die ihren Einsatz in der Eignungsdiagnostik nicht empfehlenswert machen. Insgesamt lässt sich festhalten, dass der persönlichkeitsbezogene Ansatz in der Sprachanalyse an der allgemein geringen Relevanz von Persönlichkeitsdimensionen für die berufliche Eignung krankt. Auch wenn hier sicherlich noch näher zu differenzieren ist (z.B. variiert die Relevanz bestimmter Dimensionen je nach Berufsbild), legt die Abwägung zwischen dem relativ großen Eingriff in die persönliche Sphäre von Kandidat*innen und dem generell geringen prognostischen Mehrwert der Methode nahe, von ihr Abstand zu nehmen. Weitere Risiken ergeben sich durch die leichte Verfälschbarkeit von sprachanalytisch basierten Auswahlmethoden, mögliche Verzerrungen der Ergebnisse durch die natürliche Variation von Sprache und nicht zuletzt durch die mögliche Diskriminierung bestimmter Personengruppen. Trotz des unbestritten großen Potenzials der Deep-Learning-Technik in künstlichen neuronalen Netzwerken stellen diese Punkte ein ethisches und rechtliches Risiko dar, das wohl den wenigsten vertretbar erscheint.

Wie der letzte Abschnitt dieses Beitrages zeigen sollte, liegen viele Probleme der KI-Sprachanalyse in grundsätzlichen Herausforderungen der natürlichen Verarbeitung von Sprache (NLP) begründet. Geschildert wurden beispielsweise die Kontextabhängigkeit von Aussagen in Bezug auf innersprachliche und umweltbezogene Zusammenhänge, die zahlreichen Gebrauchsweisen von Sprache (z.B. metaphorisch vs. beschreibend), die Mehrdeutigkeit von Wörtern und Sätzen sowie Eigenheiten der gesprochenen Sprache wie z.B. Gedankenpausen, Betonungen oder Korrekturen des Gesagten.

Es bleibt also spannend zu sehen, ob durch die Verbesserung der linguistischen Modellannahmen und immer mächtigere KI-Systeme, vor allem im Deep-Learning-Bereich, diese Probleme nach und nach behoben werden können, sodass die KI-Sprachanalyse irgendwann eignungsdiagnostisch ernst zu nehmende Ergebnisse produziert.

[8] Dieser bekannte Satz findet sich unter Ordnungsnummer 5.6 im Tractatus (Wittgenstein, 1922).

[9] Da die Venus sich auf ihrer Bahn schneller bewegt als die Erde, überholt sie unseren Planeten. Bei der Annäherung an die Erde steht sie östlich der Sonne und erscheint als „Abendstern“ am Himmel. Hat sie die Erde überholt und entfernt sich wieder von ihr, steht die Venus westlich der Sonne und erscheint als „Morgenstern“ am Himmel (Quelle: https://www.weltderphysik.de/thema/hinter-den-dingen/abendstern-und-morgenstern/).

[10] Darstellung der Fregeschen „Kraft-Gehalt-Unterscheidung“ nach Stoutland (2004, S. 25).

Literatur:

Beck, S., Grunwald, A., Jacob, K. & Matzner, T. (2019). Künstliche Intelligenz und Diskriminierung. Herausforderungen und Lösungsansätze – Whitepaper der Arbeitsgruppe IT-Sicherheit, Privacy, Recht der Plattform Lernende Systeme. Online unter: file:///C:/Users/FESEFE~1/AppData/Local/Temp/AG3_Whitepaper_250619.pdf

Bös, N. & Marx, U. (27.11.2019). Wenn Computer Bewerber aussuchen. Online-Artikel der Frankfurter Allgemeinen Zeitung. Online unter: https://www.faz.net/aktuell/karriere-hochschule/buero-co/vorstellungsgespraech-mit-ki-wenn-computer-bewerber-aussuchen-16507199.html#void

Fesefeldt, J. (2017): Persönlichkeitstests – Was ist Persönlichkeit? dgp-Informationen 2017. Online unter https://www.dgp.de/files/dgp/downloads/dgp%20Informationen/dgp-Informationen_2017.pdf.

Fesefeldt, J. (2018). Künstliche Intelligenz im Personalmanagement. In: dgp-Informationen 2018, S. 6-37. Online unter: https://www.dgp.de/files/dgp/downloads/dgp%20Informationen/dgp-Informationen-2018_web.pdf [aufgerufen: 30.08.2021]

Frege, Gottlob (1892). Über Sinn und Bedeutung. In: Patzig, G. (Hrsg., 2008). Funktion, Begriff, Bedeutung. Fünf logische Studien. Göttingen: Vandenhoeck & Ruprecht.

Goffman, E. (1959). The Presentation of Self in Everyday Life. New York: Doubleday.

Jurafsky, D. & Martin, J. H. (2008). Speech and Language Processing. New Jersey: Pearson - Prentice Hall.

Obermann, C. (2013). Assessment Center - Entwicklung, Durchführung, Trends mit aktuellen AC-Übungen. Wiesbaden: Springer-Gabler.

Pennebaker, J., Mehl, M. & Niederhoffer, K. (2003). Psychological Aspects of Natural Language Use: Our words, Our Selves. Annual Revue of Psychology, Heft 54, S. 547-577.

Kanning, U. P. (2017). Computertechnologie in der Personalauswahl – Zwischen Hoffen und Bangen. Skeptiker, Heft 4, S. 168-173.

Kanning, U. (2019). Künstliche Intelligenz in der Eignungsdiagnostik – ein kritischer Blick [Vortrag auf dem dgp-Kongress „Personalarbeit ohne Personal?!“ in Berlin 2019].

Kaplan, Jeremy (2017). Künstliche Intelligenz – eine Einführung. Köln: Mitp-Verlag.

Kersting, M. (2008). Qualität in der Diagnostik und Personalauswahl – der DIN-Ansatz. Göttingen: Hogrefe.

Redaktion der Zeitschrift Wirtschaftspsychologie aktuell (13.08.2019). Big Brother Award für Precire. Online unter: https://www.wirtschaftspsychologie-aktuell.de/magazin/bigbrotheraward-fuer-precire/14/ [aufgerufen: 31.08.2021].

Searle, J. (1980). Minds, Brains and Programs. In: Behavioral and Brain Sciences, Heft 3, S. 417–457.

Stoutland, F. (2004). Philosophie und Bedeutung 1 – Studienbrief der FernUniversität zu Hagen (Kurs 03395) [unveröffentlicht, neueste Auflage aus 2021].

Schaumlöffel, L., Hübner, R., Thile, S. und Stulle, K. P. (2018), Du bist, was du sprichst – Validierung der Sprachanalysetechnologie PRECIRE® anhand des HEXACO®-Persönlichkeitsmodells (Kap.3), in: Psychologische Diagnostik durch Sprachanalyse, Stulle, K.P. (Hrg.), Wiesbaden: Springer Gabler.

Tichy, N. (02.03.2021), Precire: KI-Pionier ist pleite, Personalwirtschaft, online unter: https://www.personalwirtschaft.de/hr-organisation/hr-software/artikel/der-ki-pionier-precire-steht-vor-dem-aus.html [aufgerufen: 31.08.2021]

Turing, A. M. (1950). Computing Machinery and Intelligence. In: Mind, Heft 59, S. 433–460.

Wittgenstein, L. (1922). Tractatus Logico-Philosophicus (Erstausgabe). London: Kegan Paul.

Kontakt:

Johannes Fesefeldt
Diplom-Psychologe
Deutsche Gesellschaft für Personalwesen e.V.
fesefeldt@dgp.de

Von Präsenz zu Remote: Erfahrungen mit und Forschungsbefunde zu Online-Assessment-Centern

Julia Wagenschieber & Renée Schweisfurth

Die Corona-Pandemie prägt die Arbeitswelt derzeit auf beispiellose Weise. Viele Unternehmen sahen sich im Zuge der Kontaktbeschränkungen und Corona-Verordnungen Veränderungen und Herausforderungen gegenüber. Arbeitsstrukturen und Prozesse mussten teils umstrukturiert oder neugestaltet werden, um Mitarbeiter*innen durch mobiles Arbeiten vor Ansteckungen zu schützen. Auch wir standen gemeinsam mit einigen unserer Kund*innen im Kontext der Kontaktbeschränkungen zur Bekämpfung der Pandemie vor neuen Herausforderungen und der Frage: Wie können wir Auswahlverfahren virtuell gestalten?

Nun haben wir als dgp mit unseren Kund*innen in den Jahren 2020 und 2021, seit Beginn der Pandemie, 37 Online-Assessment-Center (Verfahren mit mehr als einer Station und mindestens einem simulationsorientierten Element, Stand Sommer 2021) konzipiert und gemeinsam remote durchgeführt. Dabei haben wir insgesamt 750 Bewerber*innen eignungsdiagnostisch beobachtet und beurteilt. Über erste Eindrücke, Besonderheiten, Vorteile und Herausforderungen bei der Gestaltung von Online-Auswahlverfahren bzw. Online-Assessment-Centern haben Jaeger und Kuhn bereits in den dgp-Informationen 2020 informiert und einige ganz praktische Hinweise zur Gestaltung und Durchführung gegeben.

Knapp eineinhalb Jahre nach Ausbruch der Pandemie ist es in unseren Augen Zeit für eine Bestandsaufnahme und Evaluation unserer Online-Assessment-Center und der Erfahrungen, die wir in diesem Zusammenhang gesammelt haben.

Darüber hinaus möchten wir im vorliegenden Artikel auf neue Befunde und Trends aus der Wissenschaft zu jüngsten Entwicklungen und Befunden hinsichtlich Online-ACs eingehen und aufzeigen, welche Forschungsfragen in diesem Kontext untersucht wurden bzw. welche weiterhin offen sind. Zudem möchten wir auch den Anwender*innen eine Stimme geben – also unseren Auftraggeber*innen, mit denen wir in den letzten eineinhalb Jahren Online-ACs gemeinsam umgesetzt haben.

Aus diesem Grund haben wir mit einigen unserer Auftraggeber*innen im Bereich Online-AC eine qualitative Befragung zu Erfahrungen und Eindrücken der gemeinsam durchgeführten Online-ACs durchgeführt. An unserer Befragung teilgenommen haben sowohl Vertreter*innen von Bundesverwaltungen als auch Landeshauptstadt-Verwaltungen. Die Sichtweisen und Erfahrungen unserer Auftraggeber*innen haben wir in diesem Artikel eingeflochten und um wissenschaftliche Befunde aus der Forschung ergänzt. Auch unsere eigenen praktischen Erfahrungen als Eignungsdiagnostiker*innen bei der Konzeption und Moderation von Online-ACs werden aufgegriffen.

Eignungsdiagnostische Standards – auch in Remote-Assessment-Centern

Mitarbeiter*innen, die die Anforderungsprofile der vakanten Positionen erfüllen und demnach gut geeignet sind – das ist der Schlüssel zum Erfolg für unsere Kund*innen bei der Rekrutierung. Um zu einer möglichst passgenauen und treffsicheren Auswahlentscheidung zu gelangen, sind bestimmte eignungsdiagnostische Standards unabdingbar. Diese Qualitätsstandards, welche durch die DIN 33430 definiert sind, gilt es auch bei der Konzeption und Durchführung von Online-Auswahlverfahren zu beachten. Sowohl die DIN 33430 als auch die Standards, die das Forum Assessment-Center e.V. erarbeitet hat und die Eignungsdiagnostiker*innen schon seit vielen Jahren als Orientierung dienen, gelten nach wie vor, unabhängig davon, ob ein AC in Präsenz oder online durchgeführt wird.

Sicherlich gibt es einige Standards, die auf die Besonderheiten eines Remote-Assessment-Centers zugeschnitten werden müssen und bei der Konzeption und Umsetzung besonders zu beachten sind. So kann es beispielsweise sinnvoll sein, in Rollenspielen die Bewertungskriterien anzupassen. Weitere praktische Hinweise dazu lassen sich bei Jaeger und Kuhn (2020) finden. Zum einen sind einige Verhaltensweisen online einfach nicht so gut zu beobachten (z.B. Gestik), zum anderen fallen Verhaltensweisen stärker auf, wenn ein Rollenspiel online durchgeführt wird. Das Unterbrechen des Gegenübers ist präsenter, wenn die Gesprächssimulation remote stattfindet, im Vergleich zu einem Gespräch im Präsenz, weil insgesamt weniger Verhaltensweisen zu beobachten sind.

Grundsätzlich gelten für uns als Psycholog*innen aus eignungsdiagnostischer Sicht also die gleichen Qualitätsstandards für Präsenz- wie auch für Online-ACs. Daher werden unsere bisherigen Standards nicht außer Kraft gesetzt, sondern lediglich für Online-Verfahren angepasst und neu diskutiert.

Auch aus diesem Grund wurde wohl die DIN SPEC 91426 (Qualitätsanforderungen für videogestützte Methoden der Personalauswahl) als Ergänzung der DIN 33430 erarbeitet und Ende 2020 veröffentlicht. Nicht erst seit dem wachsenden Einsatz von videogestützten Methoden zur Personalauswahl, sondern bereits seit sich Eignungsdiagnostiker*innen mit der Güte von Auswahlverfahren befassen, ist eines der wichtigsten Gütekriterien die Validität.

Die Validität ist ein Indikator dafür, ob wir mit einer Auswahlmethode das messen, was wir auch zu messen beabsichtigen. Greifen wir bei der Personalauswahl auf die Methode von ACs zurück, ist die Absicht, dass wir mit dem AC vorhersagen möchten, wie erfolgreich ein/e Bewerber*in auf der zu besetzenden Position sein wird. Darum stellt auch für Anwender*innen die Validität eines der wichtigsten Kriterien zur Auswahl von Instrumenten der Personalauswahl dar. Gleichzeitig sehen Anwender*innen heute die Praktikabilität und die Akzeptanz seitens der Bewerber*innen der eingesetzten Verfahren als wichtige Kriterien (Armoneit, Schuler & Hell, 2020). Den Autoren zufolge bestimmen diese drei Kriterien die Entscheidung über die Nutzung eines bestimmten Personalauswahl-Instruments für Anwender*innen nachhaltig. Die Validität ist nach den Ergebnissen der Trendanalyse von Armoneit et al. (2020) der wichtigste Indikator. Danach spielen Praktikabilitäts- und dann Akzeptanzüberlegungen eine wesentliche Rolle für Anwender*innen.

Aus diesem Grund betrachten wir den Status quo aus den aktuellen wissenschaftlichen Befunden und unseren eigenen Erfahrungen sowie denen unserer Auftraggeber*innen vor dem Hintergrund dieser drei Kriterien.

Validität: Wie gut können wir die soziale Kompetenz virtuell überhaupt beobachten?

Insbesondere im Bezug auf die Validität ist uns im Laufe der letzten 1,5 Jahre häufiger der Einwand begegnet, dass soziale Kompetenzen in Online-Verfahren schlechter zu beurteilen sind. Oftmals wird befürchtet, bestimmte Verhaltensweisen über Videoformate schlechter bis gar nicht erfassen zu können. Nach einer Theorie von Potosky („Framework of Media Attributes“, 2008), wird medienvermittelte Kommunikation durch unterschiedliche Merkmale beeinflusst. Unter anderem die soziale Bandbreite.

So ist die Kommunikation einfacher, wenn während des Austauschs neben verbalen Kommunikationspfaden auch nonverbale Pfade (z.B. Gesten, Mimik, ...) verwendet werden. Das Fehlen von nonverbalen Verhaltensweisen führt nach Potosky demnach

zu Einschränkungen in zwischenmenschlichen Interaktionen, weil nonverbale Signale nicht als zusätzliche Informationsquelle während eines Gesprächs zur Verfügung stehen. Bei videobasierten Methoden zur Personalauswahl fehlen nonverbale Verhaltensweisen zwar nicht zur Gänze, sind jedoch eingeschränkter beobachtbar.

Diese Theorie spiegelt sich auch in den Erfahrungen eines/einer unserer Auftraggeber*innen, dem Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, wider: Dadurch, dass über die Kamera nur ein Ausschnitt der Bewerbenden zu sehen ist, gehen Informationen verloren, die in einem Präsenzverfahren zu Verfügung stehen. Es geht um Informationen, die aus dem Verhalten von Bewerber*innen gewonnen werden, z.B. wenn sie den Raum betreten – sowie Informationen der Kleidung, der informellen Begrüßung usw. Joachim Turre, Referent der Personalgewinnung des BMZ, spricht von sogenannten fehlenden „social cues“ (dt. „soziale Hinweisreize“), die unsere Wahrnehmung in der Realität beeinflussen, und die uns mehr Sicherheit für unser eigenes Urteilsvermögen geben. Durch das Fehlen dieser Informationen käme den verbalen Äußerungen ein höheres Gewicht zu, so Herr Turre.

Die sogenannten „social cues“ in einer Beobachtungssituation während des ACs (z.B. in einem Rollenspiel) können auch aus eignungsdiagnostischer Sicht wertvolle Informationen für Kommissionsmitglieder zur Beurteilung liefern. Sie bergen jedoch auch das Risiko, dass die Urteile einer Auswahlkommission durch ebendiese Informationen unbewusst beeinflusst und verzerrt werden. Dieses Risiko fällt daher nach aktueller Auffassung bei digitalen Interviewformaten geringer aus. Die Digitalisierung könnte also in diesem Fall helfen, Verzerrungen zu verringern (Suen et al., 2019; zitiert nach Woods et al., 2020). Eine Meta-Analyse von Blacksmith, Willford und Behrend (2016) zeigt, dass die Bewertungen von Bewerber*innen in onlinebasierten Auswahlverfahren insgesamt schlechter ausfallen als in Präsenzverfahren. Ein Forschungsbefund dazu ist, dass Bewertende in Online-Verfahren strenger zu bewerten scheinen (Melchers, Petrig, Basch, & Sauer, 2021). Dies könnte damit zusammenhängen, dass „social cues“ fehlen und die Abwesenheit dieser zu strengeren Beurteilungen verleitet oder dass nur das ausgesprochene Wort für die Beurteilung berücksichtigt wird. In letzterem Fall könnte es sein, dass „social cues“ von den verbalen Aussagen ablenken oder diese aufwerten. Unter Validitätsgesichtspunkten ist dies insofern relevant, als dass für Online-ACs dann (je nach Anforderung) Aufgaben oder Fragen anders gestaltet werden, damit das gleiche Konstrukt wie in einem Präsenzauswahlverfahren beobachtbar gemacht wird. Auch das BMZ denkt darüber nach, für zukünftige Online-Auswahlverfahren bisher verwendete Interviewfragen zu prüfen, zu schärfen und zu optimieren, um sie verständlicher zu gestalten und die anforderungsrelevanten Aspekte im Antwortverhalten noch besser beobachtbar machen zu können.

Eine kurze qualitative Befragung von drei weiteren Auftraggeber*innen, die im vergangenen Jahr mit der dgp Online-ACs durchgeführt haben, ergab u.a. dass die Befragten den Informationsverlust, die erschwerte Interaktion der Bewerber*innen untereinander (z.B. im Rahmen von Gruppendiskussionen) und die Tatsache, dass sie die Sozialkompetenz als Auswahlgremium schwieriger zu beurteilen vermochten, als wesentliche Unterschiede zwischen Online- und Präsenzverfahren ansehen. Die Sozialkompetenz könne, so berichtete ein Mitarbeiter aus dem Personalbereich der Landeshauptstadt Hannover, die seit vielen Jahren Präsenzauswahlverfahren mit der dgp gemeinsam durchführt, schlechter durch situative Fragen aufgefangen werden.

Unser Auftraggeber führt hier an, er habe häufig Bewerber*innen im Online-Verfahren erlebt, die im durchschnittlichen Bereich abschnitten, d.h. weniger Extremfälle, und die situativen Interviewfragen „erscheinen anfällig für sozial erwünschtes Antwortverhalten.“ Auch der Blickkontakt sei im Online-Setting schwerer zu deuten. Auch wenn die Landeshauptstadt Hannover grundsätzlich von Online-ACs überzeugt ist, wird deutlich, dass eine Übersetzung von Präsenzverfahren zu virtuellen Verfahren nicht eins zu eins funktioniert. Um möglichst wenig Informationsverlust zu erleiden und die relevanten sozialen Kriterien auch online beobachten und bewerten zu können, scheinen Aufgabenstellung wie Bewertungskriterien ein Update zu benötigen. Das sollte jedoch – wie oben bereits angedeutet – in keinem Fall die Standards außer Acht lassen, die für Assessment-Center bereits gelten.

Bei einer neueren Studie zum Vergleich von videogestützten vs. Face-to-Face-Interviews gaben Teilnehmer*innen an, dass sie denken, weniger Impression Management betreiben zu können (Basch, Melchers, Kurz, Krieger & Miller, 2020). Impression Management ist die bewusste oder auch unbewusste Steuerung des Eindrucks, den Personen in einem Auswahlverfahren auf andere machen, also der Versuch sich gegenüber anderen beispielsweise positiv darzustellen. Ein Grund für die schlechteren Beurteilungen in Online-Verfahren kann also auch darin liegen, dass Bewerber*innen weniger Impression Management betreiben (können), da Impression Management nicht nur auf verbalen, sondern auch auf nonverbalen Kommunikationswegen abläuft, und diese im Online-AC zum Teil nicht zur Verfügung stehen. Somit sind die Beurteilungen der Kommission weniger anfällig für Verzerrungstendenzen und könnten dadurch sogar objektiver ausfallen.

Neue wissenschaftliche Untersuchungen belegen zudem, dass technologie-basierte Einstellungsinterviews (telefonisch oder per Videokonferenz) im Vergleich zu Face-to-Face-Interviews mit schlechteren Bewertungen der Interviewten einhergehen und dies auf eine tatsächlich schwächere Leistung der Teilnehmer*innen im AC zurückzuführen ist und nicht auf verändertes Beurteilungsverhalten (Melchers et al., 2021). Interessanterweise lassen sich diese Unterschiede in der Bewertung durch unabhängige Bewerter*innen vor allem für die Antworten auf standardisierte Interviewfragen finden. Ob es sich dabei um ein telefonisches oder videobasiertes Interview handelte, machte keinen Unterschied mehr; in Interviews mit beiden Medien fielen die Bewertungen schlechter aus, als wenn das Interview im Rahmen eines persönlichen Gesprächs geführt wurde. Für Aufgabenstellungen, die den Bewerber*innen eine Vorbereitungszeit gewährten (in der Studie handelte es sich um eine maximal 3-minütige Selbstvorstellung, auf die sich die Bewerber*innen 4 Minuten lang vorbereiten durften), entstanden keine größeren Unterschiede in der Bewertung der Leistung.

Auf Basis dieser Ergebnisse können wir schließen, dass es für das Abschneiden der Bewerber*innen sehr wohl einen Unterschied macht, in welchem Format Auswahlverfahren durchgeführt werden.

Die Autoren der Untersuchung versuchten bereits eine mögliche Ursache auszuschließen: Sie erhoben das Stress- bzw.

Testangst-Niveau der Bewerber*innen in der Situation sowohl durch physiologische Maße, als auch durch Selbsteinschätzung, um ggf. Unterschiede zu erfassen, die auf-grund der Modalität des Interviews (Face-to-Face, videobasiert, telefonisch) zustande kommen könnten. Sie fanden keine signifikanten Unterschiede. Weitere Erklärungsansätze hierfür könnten darin liegen, dass Interviewte Präsenzverfahren vor virtuellen Verfahren bevorzugen und ihre eigene Leistung in Präsenzverfahren besser einschätzen und somit im Sinne des psychologischen Effekts der sich selbsterfüllenden Prophezeiung auch tatsächlich in virtuellen Verfahren schlechtere Leistungen zeigen (Vgl. Chapman, Uggerslev & Webster, 2003; Sears, Zhang, Wiesner, Hackett & Yuan, 2013).

Nach Betrachtung aktueller Forschungsergebnisse zum Thema Remote-Verfahren bleibt also immer noch die Frage offen: Wie kommt es zu Unterschieden in der Leistungsbewertung? Liegt es an den tatsächlichen Leistungen der Bewerber*innen oder daran, dass diese aufgrund fehlender „social cues“ weniger Leistungen erbringen können? Oder liegt es daran, dass Bewerter*innen schlechter bewerten, möglicherweise weil ihnen Informationen fehlen und sie dadurch „strenger“ sind? Auch ist fraglich, ob sich diese Bewertungsunterschiede auch bei der Bewertung von simulationsorientierten Verfahren (z.B. Rollenspielen, Fallstudien etc.) belegen lassen. Interessant wären ähnliche Studien wie die von Melchers et al. (2021), die nicht nur die Bewertungsunterschiede in Interviewformaten, sondern auch anderen Assessment-Center-Stationen untersuchen.

Wir dürfen gespannt sein, welche Erklärungen die wissenschaftliche Forschung in Zukunft wider- und belegen wird. Schon heute sollten alle Beteiligten an einem Online-AC- oder -Interviewformat sich dieser Unterschiede jedoch bewusst sein. Leistungen von Bewerber*innen aus Präsenz-Auswahlverfahren und Online-Auswahlverfahren sollten wir nach derzeitigem Kenntnisstand jedoch nicht miteinander vergleichen, da wir nicht genau wissen, welche psychologischen Prozesse die Unterschiede in den Bewertungen hervorrufen.

Praktikabilität: Das Online-AC als Ressourcen-Sparschwein

Wir können festhalten, dass Online-Assessment-Center Präsenz-ACs hinsichtlich der Praktikabilität oft in den Schatten stellen: So steht häufig im Vordergrund für unsere Auftraggeber*innen, dass sowohl für Bewerber*innen wie auch für Mitglieder der Auswahlkommission zum Teil enormer Reiseaufwand entfällt. Insbesondere für Auftraggeber*innen, die überregional Bewerber*innen rekrutieren oder sogar regelmäßig Bewerber*innen im Auswahlverfahren kennen lernen wollen, die sich gerade im Ausland befinden, stellt das Online-AC eine Erleichterung dar. Diese Rückmeldung gab es vereinzelt auch am Ende des Online-ACs von Bewerber*innen im Bundesministerium für wirtschaftliche Entwicklung und Zusammenarbeit (BMZ). Das BMZ sieht den organisationalen Wert der Nachhaltigkeit, für den es neben vielen weiteren steht, durch die Online-ACs mehr gewahrt als durch Präsenz-ACs. Für das BMZ interessante Bewerber*innen befinden sich nicht selten im Ausland und können so flexibel aus der ganzen Welt „anreisen“, ohne finanzielle und zeitliche Ressourcen aufzuwenden. Als Ministerium mit zwei Dienstsitzen, die gleichermaßen an der Auswahl neuen Personals beteiligt sind, werden diese Ressourcen auch innerhalb der Belegschaft eingespart.

Die befragten Auftraggeber*innen von Online-ACs der dgp gaben insbesondere an, dass die organisatorischen Rahmenbedingungen von Online-ACs gut gestaltbar sind und gut gestaltet wurden. Hier wurde beispielsweise die Flexibilität im Umgang mit technischen Störungen genannt. Ein Mitarbeiter aus dem Personalbereich der Landeshauptstadt Hannover gab auch an, dass sich positiv bemerkbar gemacht habe, dass sich die Bewerber*innen, anders als in Präsenzverfahren, im Online-AC weniger zu den Inhalten der jeweiligen Stationen austauschen. So hat das Ausbleiben des informellen Austauschs unter den Bewerber*innen also auch einen positiven Effekt, da das Verhalten der Bewerber*innen nicht verfälscht wird.

Im Rahmen unserer Umfrage gaben die Auftraggeber*innen u.a. auch an, wie Bewerber*innen das Verfahren wahrgenommen haben bzw. welche Rückmeldungen die Auftraggeber*innen von den Bewerber*innen bekommen haben. Auch hier wurden Aspekte genannt, die für eine höhere Praktikabilität von Online-ACs gegenüber Präsenz-ACs sprechen. Die „Vermeidung der Anreise“ sowie eine gute Betreuung während des Verfahrens wurden hier u.a. zurückgespiegelt.

Auch Armoneit et al. (2020) belegen in ihrer Trendanalyse, dass aufwendige und kostenintensive Verfahren, wie es Assessment-Center beispielsweise sind, insbesondere zur Personalauswahl in Führungspositionen herangezogen werden. Die dahinterstehende Annahme ist, dass für Anwender*innen die Relevanz und der Einfluss der zu besetzenden Position oft die aufwendigen Mittel heiligt. Auch für Traineepositionen werden Assessment-Center häufiger angewendet. Hier verpflichtet man sich ebenfalls zu einer Beschäftigung für einen mittelfristigen Zeitraum; zudem sind Trainees häufig auch Potenzialträger*innen in Organisationen, mit denen man längerfristig plant und sich daher zugesteht, aufwendigere Auswahlverfahren durchzuführen.

Akzeptanz: Wie Bewerber*innen Remote-Verfahren wahrnehmen

Unseren Auftraggeber*innen ist es wichtig, dass der Auswahlprozess von Bewerber*innen als fair und aussagekräftig wahrgenommen wird. Auch durch die Veränderung des Bewerber*innenmarkts und Konkurrenz um Bewerber*innen gewinnt der Faktor Akzeptanz bzw. Attraktivität der/des Arbeitgeber*in zunehmend an Bedeutung. Bereits der Bewerbungsprozess und das Auswahlverfahren können beeinflussen, wie der/die Arbeitgeber*in von Bewerbenden wahrgenommen wird.

Die Forschung zeigt, dass Bewerber*innen Face-To-Face-Interviews den Videointerviews vorziehen und befürchten, in Videointerviews ihre Leistung weniger zeigen zu können (Proost, Germeys & Vanderstukken, 2021). Allerdings werden videobasierte Interviews generell durch Bewerber*innen fairer wahrgenommen, wenn der Einsatz der Interviews zuvor beispielsweise durch höhere Standardisierung oder eine höhere Flexibilität erklärt wird (Basch & Melchers, 2019). Auch haben strukturierte Interviews, die aus eignungsdiagnostischer Sicht den unstrukturierten Interviews vorzuziehen sind, über ein Videoformat eine höhere Akzeptanz, als wenn diese face to face durchgeführt werden.

Aus der Trendanalyse von Armoneit et al. (2020) geht hervor, dass zusätzlich zu strukturierten Einstellungsinterviews und der Analyse von Bewerbungsunterlagen „neuerdings die Arbeitsprobe / Fallstudie am ehesten bei Bewerbenden akzeptiert ist“ (S. 12). Arbeitsproben und Fallstudien sind

den simulationsorientierten Aufgabentypen zuzuordnen und daher nicht selten Bestandteil von Assessment-Centern.

Unter Akzeptanzgesichtspunkten ist es daher ratsam, auch in Online-ACs simulationsorientierte Verfahren in Form von Arbeitsproben oder Fallstudien aufzunehmen.

Dies können z.B. Rollenspiele wie Mitarbeiter*innen- oder Verhandlungsgespräche sein.

Melchers et al. (2021) zeigen auf, dass sich die Wahrnehmung von technologiebasierten Auswahlmethoden im Vergleich zu Präsenzauswahlmethoden im Laufe eines Auswahlprozesse durchaus verändert: In ihrer Untersuchung schätzen Bewerber*innen, die ein video- oder telefonbasiertes Einstellungsinterview durchführten, diese Methoden vor dem Interview wesentlich schlechter ein als Bewerber*innen, die ein Präsenz-Interview durchführten. Dieser Unterschied hob sich jedoch auf, nachdem die Interviews durchgeführt wurden. Es sind also eher die Vorbehalte, welche die Wahrnehmung der Bewerber*innen negativ beeinflussen. Diese werden ausgeräumt, sobald sie an einem technologiebasierten Interview tatsächlich teilgenommen haben. Gerade letzteres ist ein wichtiger Befund in Bezug auf die Auswirkungen der Wahrnehmungen von Bewerber*innen von Auswahlverfahren: Eine positive Wahrnehmung geht in der Regel auch mit einem gesteigerten Attraktivitätsempfinden (Hausknecht et al, 2004; zitiert nach Melchers et al., 2021) einher – ebenso wie mit einer höheren Wahrscheinlichkeit, sich wieder zu bewerben (Gilliland et al, 2001; zitiert nach Melchers et al., 2021), sowie der Weiterempfehlung des/der Arbeitgeber*in (Harold et al, 2016; zitiert nach Melchers et al., 2021).

Daher ist es lohnenswert für Organisationen, diese Parameter bei der Gestaltung von Auswahlverfahren zu berücksichtigen. Die Ergebnisse von Melchers et al. (2021) legen nahe, dass auch technologiebasierte Auswahlmethoden zu einer positiven Wahrnehmung bei Bewerber*innen führen können, und Vorbehalte sich auflösen, die vor dem Verfahren bestehen. Jedoch ist es auch möglich, zwischen Ankündigung des technologiebasierten Auswahlverfahrens und der Durchführung potenziell Geeignete zu verlieren. In der Praxis lässt sich hier eventuell durch Aufklärung der Bewerber*innen, Transparenz beim Vorgehen und einer entsprechenden Vorbereitung auf das Verfahren (z.B. durch Technik-Checks oder Videotutorials) Abhilfe schaffen, um Vorbehalten bereits zu Beginn entgegen zu wirken.

Ein weiterer Aspekt, der sich auf die Akzeptanz von Online-Assessment-Center auswirkt, ist die Realitätsnähe, die Gespräche oder Präsentationen über ein Videokonferenzsystem aufweisen. Mittlerweile gehören diese Kommunikationsformate zum beruflichen Arbeitsalltag. So beschreiben Paschen und Brandt (in Stulle, 2020), dass es mittlerweile als natürlicher erlebt werden könnte, die typischen Assessment-Center-Stationen in einem Online-Format durchzuführen, als dafür ein aufwendiges, durchstrukturiertes und dennoch oft als „künstlich hergestellt“ erlebtes Präsenzauswahlverfahren durchzuführen. Diese Unterschiede in der Wahrnehmung aller Beteiligten sollten zukünftig jedoch empirisch untersucht werden, um diese bisher theoretisch hergeleitete These zu überprüfen.

Interessant ist auch, dass Bewerber*innen unseren Auftraggeber*innen im Rahmen der Rückmeldung zu den Online-ACs spiegelten, sie seien aufgrund der Tatsache, dass das Verfahren online stattfand, „weniger nervös und angespannt gewesen“. Dies könnte in der Folge dazu führen, dass Bewerber*innen ihre Leistungsfähigkeit in Online-ACs besser unter Beweis stellen können, weil z.B. Prüfungsangst ihre Leistungen weniger beeinflussen würde. Auch das subjektive Empfinden, die eigene Leistungsfähigkeit im Auswahlverfahren darstellen zu können, könnte sich dadurch steigern und würde evtl. zu einer höheren Akzeptanz des ACs als Auswahlinstrument führen. Die These, ob „psychologische Sicherheit“, also die Tatsache, dass Bewerber*innen sich wohler und sicherer fühlen, wenn sie in einer bekannten Umgebung an einem AC teilnehmen, tatsächlich die Ursache für mehr Akzeptanz bei Bewerber*innen ist, verlangt jedoch noch nach wissenschaftlicher Fundierung. Gegebenenfalls könnte dieses Sicherheitsempfinden auch mit der Gewöhnung an virtuelles (Zusammen-) Arbeiten in Verbindung stehen, und stärker werden, je mehr wir uns als Gesellschaft an ein Remote-Arbeitsumfeld gewöhnen.

Einsatz von Online-ACs auch nach Rückkehr in die Normalität?

Aus den Rückmeldungen unserer Befragung können wir das Fazit ziehen, dass alle Organisationsvertreter*innen, die an der Umfrage teilgenommen haben, ein Online-AC in der durchgeführten Form noch einmal durchführen würden. Eine/r der Befragten gab an, dass generell ein Präsenz-AC vorzuziehen wäre, Online-ACs aber eine gute Alternative für „pandemische Bedingungen“ oder für die Abwesenheit von Bewerber*innen oder anderen am Verfahren Beteiligten wäre. Die drei weiteren befragten Organisationsvertreter*innen gaben an, dass für sie Online-ACs eine Alternative sind, die sie auch ohne Kontaktbeschränkungen etc. weiter in Erwägung ziehen. Zwei der befragten Organisationen denken bereits konkret darüber nach, wie sich diese Vorgehensweise in bestehende Auswahlprozesse und -richtlinien integrieren lässt. Wir können also feststellen, dass auch im Bereich der Assessment-Center die COVID19-Pandemie zu einem Umdenken und einem Digitalisierungsschub beigetragen hat. Jetzt, nachdem wir gezwungen wurden, konkret über Möglichkeiten nachzudenken und diese auch umzusetzen, bevor Vorgehensweisen und Prozesse bis zum Ende durchdacht wurden, fällt es vielen leichter, das Potenzial in der Methode Online-AC zu erkennen. Vieles ist ins Rollen gebracht worden, viel haben wir ausprobiert, bei vielem Erfolg gehabt und manches mussten wir auch wieder verwerfen. Weiterhin werden wir uns mit unseren Auftraggeber*innen Gedanken machen, wie wir Online-ACs optimal gestalten – für Bewerber*innen und Arbeitgeber*innen, für eine effiziente Durchführung sowie für eine qualitativ hochwertige Eignungsdiagnostik. Wir dürfen gespannt sein, welche Erkenntnisse wir in der Zukunft hierzu gewinnen werden.

Literatur:

Armoneit, C., Schuler, H., & Hell, B. (2020). Nutzung, Validität, Praktikabili-tät und Akzeptanz psychologischer Personalauswahl-verfahren in Deutschland 1985, 1993, 2007, 2020. Zeitschrift für Arbeits- und Organi-sationspsychologie (2020),64(2), 1–16. Doi: https://doi.org/10.1026/0932-4089/a000311

Basch, J.M. & Melchers, K.G. (2019). Fair & Flexible?! Explanations Can Improve Applicant Reactions Toward Asynchronous Video Interviews. Personnel Assesment and Decisions, 5 (3).

Basch, J.M., Melchers, K. G., Kegelmann, J., & Lieb, L. (2018). Smile fort he Camera! The role of social presence and impression management in perceptions of technology-mediated interviews. Journal of Managerial Psychology, 35 (4), 285-299.

Basch, J.M., Melchers, K. G., Kurz, A., Krieger, M., & Miller, L. (2020). It Takes More Than a Good Camera: Which Factors Contribute to Differ-ences Between face-to-Face Interviews and Videoconference Interviews Regarding Performance Ratings and Interviewee Perceptions?Journal of Business and Psychology. Doi: https://doi.org/10.1007/s10869-020-09714-3

Blacksmith, N., Willford, J. C., & Behrend, T. S. (2016). Technology in the employment interview: A meta-analysis and future research agenda. Pers.Assess. Decis. 2, 12–20. doi: 10.25035/pad.2016.002

Chapman, D. S., Uggerslev, K. L., & Webster, J. (2003). Applicant reac-tions to face-to-face and technology-mediated interviews: A field investga-tion. J.Appl.Psychol. 88, 944–953.doi: 10.1037/0021-9010.88.5.944

Melchers, K. G., Petrig, A., Basch, J. M., & Sauer, J. (2021). A comparison of conventional and technology-mediated selection interviews with regard to interviewees' performance, perceptions, strain, and anxiety. Frontiers in Psychology, 11, 3851. Doi: 10.3389/fpsyg.2020.603632

Potosky, D. (2008). A conceptual framework for the role of the administra-tion medium in the personnel assessment process. Academy of Manage-ment Review, 33, 629–648.

Proost, K., Germeys, F, & Vanderstukken, A. (2021). Applicants' pre-test reaction towards video interviews: the role of expected chances to de-monstrate potential and to use nonverbal cues. European Journal of Work and Organizational Psychology, 30 (2), 265-273.

Sears, G. J., Zhang, H., Wiesner, W. H., Hackett, R. D., & Yuan, Y.(2013).

A comparative assessment of videoconference and face-to-face employ-ment interviews. Manag. Decis. 51, 1733–1752.doi: 10.1108/MD-09-2012-0642

Stulle, K. P. (2020). Digitalisierung der Management-Diagnostik. Springer.

Woods, A.S., Ahmed, S., Nikolaou, I., Costa, A.C., & Anderson, N.R. (2020). Personnel selection in the digital age: a review of validity and appli-cant reactions, and future research challenges. Europe-an journal of work and organizational psychology, 29 (1), 64-77.

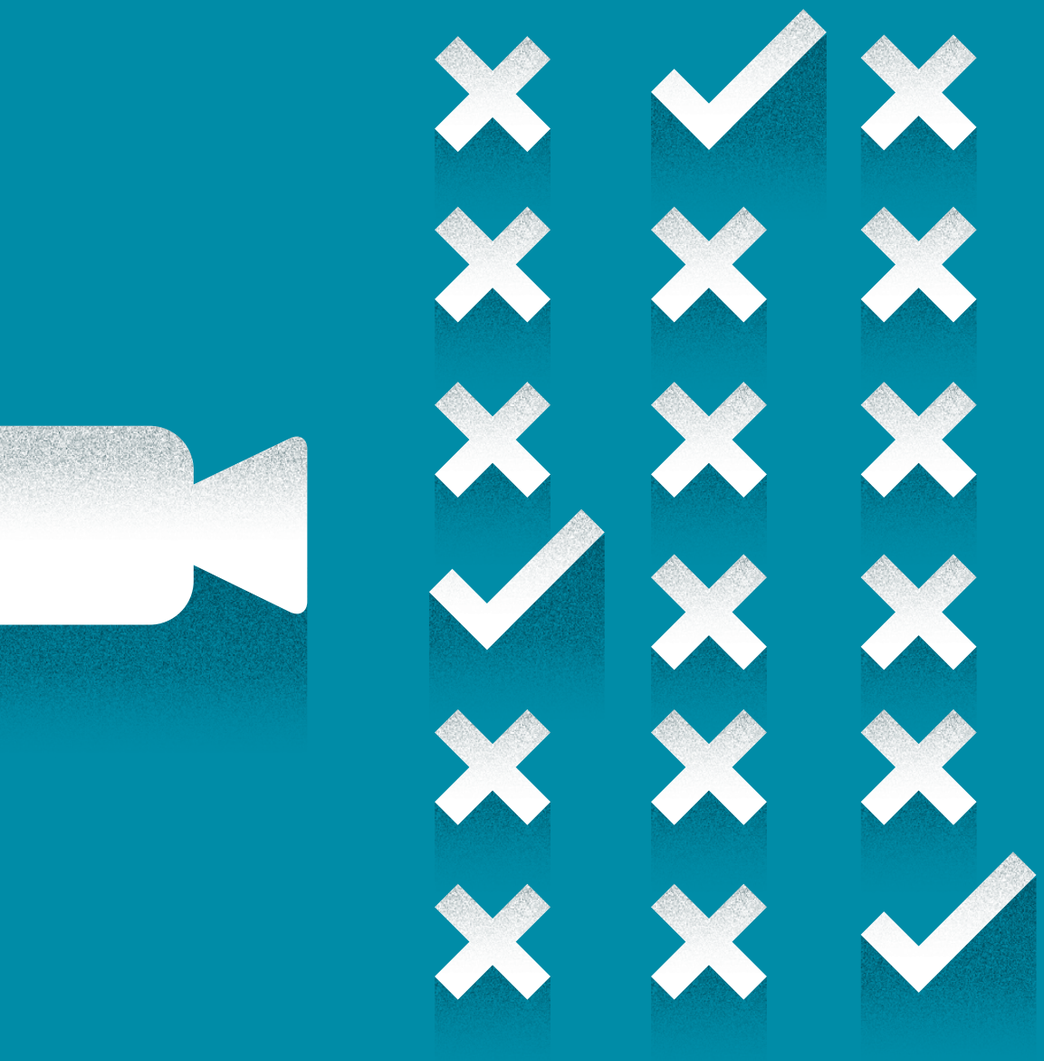
Kontakt:

Julia Wagenschieber
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
wagenschieber@dgp.de

Renée Schweisfurth
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
schweisfurth@dgp.de

Ein kritischer Kommentar zur DIN SPEC 91426

Julia Wagenschieber & Christian Reiß



Videogestützte Methoden der Personalauswahl haben pandemiebedingt einen Aufschwung erlebt. Um in Corona-Zeiten Auswahlverfahren durchführen zu können und/oder dies möglichst kontaktlos zu tun, mussten viele Organisationen schnell und flexibel auf Personalauswahlverfahren ausweichen, die „remote“ und mit Hilfe von Videokonferenzsystemen durchzuführen sind. Während dieser Umstellung von Präsenzauswahlformaten auf videogestützte Auswahlverfahren wurde schnell deutlich, dass die Handhabung beider Durchführungsmodi (Präsenz/remote) Unterschiede aufweist. Es erscheint daher sinnvoll, eine fachliche Empfehlung zu formulieren, wie videogestützte Auswahlverfahren bzw. durch Videokonferenzsysteme gestützte Auswahlinstrumente (im Folgenden: Online-Auswahlinstrumente) so eingesetzt werden, dass man mit ihnen treffsicher geeignete Personen aussuchen kann.

Aus diesem Grund ist die Intention, eine Hilfsrichtlinie zur Durchführung von Online-Auswahlverfahren zu entwickeln, sehr zu begrüßen.

Die Ausgestaltung dieser Hilfestellung (DIN SPEC 91426) sollte weiter vorangetrieben werden, besonders hinsichtlich der Punkte, die die etablierte DIN33430 unnötigerweise aufweichen. In diese Diskussion möchten wir gerne mit diesem Artikel, den folgenden Gedanken und Anmerkungen einsteigen.

Probleme von videogestützten Auswahlverfahren

Die größten Probleme, die wir bei videogestützten Auswahlverfahren sehen, sind:

- **Mangelnde Durchführungsobjektivität**
  Die Bedingungen, unter denen die Bewerber*innen an einem Online-Auswahlverfahren teilnehmen, sind bei videogestützten Verfahren nicht gleich und können nicht vollständig kontrolliert werden. Wir können nicht mit Sicherheit ausschließen, dass Bewerber*innen Hilfsmittel, z.B. auf dem Bildschirm, verwenden oder dass dritte Personen im Raum sind, während Bewerber*innen sich vorbereiten etc. Insbesondere ist damit die Unabhängigkeit von „irrelevanten Einflüssen" wie z.B. situativen Besonderheiten nicht gewährleistet (vgl. DIN33430, Juli 2016, 2.14; S.8). Dadurch sind die damit verbundenen Gütekriterien der Reliabilität und der Validität eingeschränkt, die zur Beurteilung der Qualität von Auswahlinstrumenten (z.B. „Übungen" im Assessment-Center) herangezogen werden.

- **Erleichterte Inhaltsweitergabe**
  Darüber hinaus ist bei diesen Online-Auswahlverfahren der Schutz der Inhalte eher gefährdet als bei Präsenz-Auswahlverfahren. Die „Hürde", Inhalte weiterzuleiten, ist geringer, wenn man, wie bei Online-Auswahlverfahren, allein und „unbeobachtet" in einem Raum sitzt, als bei Präsenz-Auswahlverfahren. Für Bewerber*innen bestehen bei videogestützten Verfahren, im Vergleich zu einem Auswahlverfahren, bei dem eine Verfahrensleitung anwesend ist, daher einfache und unkontrollierte Möglichkeiten, die Aufgaben und Fragen zu dokumentieren und weiterzugeben. Screenshots und Aufnahmen von Fragen, Aufgaben und auch dem gesamten Verfahren können ohne großes Risiko angefertigt werden, wodurch Inhalte des Auswahlverfahrens deutlich einfacher und präziser weitergegeben werden können als in Präsenzverfahren. Auch dies wirkt sich unmittelbar auf die Objektivität und die Validität des Auswahlverfahrens aus.

- **Technische Probleme**
  Trotz einer immer größer werdenden Expertise im Umgang mit Videokonferenz-Tools kommt es sowohl auf Seiten der Bewerber*innen als auch der Auswahlgremien zu technischen Störungen. Diese können mit mangelnder Vertrautheit von Individuen mit dem Tool zusammenhängen, mit schlechter Internetverbindung an (Homeoffice-)Arbeitsplätzen usw. Auch die unterschiedliche Qualität von Bild und Ton durch unterschiedliches technisches Equipment ist als technische Herausforderung zu bewerten.

- **Systematische Unterschiede in der Leistung der Bewerber*innen**
  Erste Studien zeigen, dass die Bewertungen der Bewerber*innen in Online-Auswahlverfahren insgesamt schlechter ausfallen können als in Präsenzauswahlverfahren (Basch et al., 2020). Inwiefern dieser Unterschied bestehen bleibt, jetzt wo wir uns als Gesellschaft mehr und mehr daran gewöhnen, remote zu arbeiten und auch Auswahlverfahren remote durchzuführen, bleibt abzuwarten. Zu wenig Konkretes wissen wir über die Ursachen für diese Bewertungsunterschiede. Ist die Leistung der Bewerber*innen schlechter oder sind die Beurteiler*innen kritischer, und welchen Einfluss haben dabei technische Gegebenheiten? Wir haben bisher erste Anhaltspunkte. Wesentlich ist jedoch, dass das Format des Auswahlverfahrens nach bisherigem Kenntnisstand das Abschneiden der Bewerber*innen beeinflusst. Damit beeinflusst die Wahl des Formats (Präsenz oder remote) auch die Validität des Auswahlverfahrens. Implikationen, die sich daraus ergeben, sind beispielsweise das Vermeiden von Mischformen, die Nicht-Vergleichbarkeit von Ergebnissen zwischen Präsenz- und Remote-Auswahlverfahren und die Anpassung von Aufgabenstellungen, Antwort- und Verhaltensankern, sodass die abzubildenden Anforderungen in Präsenz wie remote gemessen werden können.

Natürlich gibt es weitere Schwierigkeiten bzw. offene Fragen, die sich ergeben, wenn man eine Qualitätsnorm für Remote-Auswahlverfahren festlegen möchte. Nicht alle sind hier aufgeführt.

Dieser Artikel konzentriert sich vorrangig auf Faktoren, die die Hauptgütekriterien von Auswahlverfahren betreffen und letzten Endes für eine angemessene Vorhersagekraft der Auswahlverfahren entscheidend sind.

Auch andere Aspekte sind relevant, die der aktuelle Entwurf der DIN SPEC91426 zum Teil bereits aufnimmt.

Inhalte der DIN SPEC 91426

Wie groß ist der Grad der Hilfestellung, den uns die DIN SPEC 91426 bei den beschriebenen Problemen bietet, und in welchen anderen Bereichen unterstützt sie ihre Nutzer*innen?

Problem der Durchführungsobjektivität: In der DIN SPEC 91426 wird nicht deutlich, dass aufgrund der Einschränkungen hinsichtlich der Durchführungsobjektivität von online-basierten Auswahlinstrumenten Qualitätseinschränkungen entstehen können, die bei Durchführung des gleichen Auswahlverfahrens in Präsenz nicht entstehen würden. Konkret kann bspw. nicht kontrolliert werden, welche Hilfsmittel die Bewerber*innen während eines Interviews nutzen. Ob sie etwas vom Bildschirm ablesen oder etwas auswendig gelernt haben, lässt sich in wenigen Fällen eindeutig sagen.

Vor diesem Hintergrund, sollte aus der DIN SPEC 91426 hervorgehen, dass, solange es keine Kontrollmöglichkeiten für die vorgenannten Aspekte gibt, bei der Anwendung von videogestützten Auswahlverfahren „ohne die persönliche Anwesenheit eines Verfahrensleiters" mit „Einschränkungen der Durchführungsobjektivität" und erhöhter Manipulierbarkeit der Ergebnisse gerechnet werden muss (vgl. DIN33430, Juli 2016, 6.2; S. 19). Wünschenswert wären hier Hinweise, wie diese Problematik minimiert werden kann.

Problem der Inhaltsweitergabe: Aus diesem Problem ergibt sich die Implikation für die/den verantwortliche*n Eignungsdiagnostiker*in, bspw. Inhalte von Auswahlverfahren (wie z.B. Interviewfragen) regelmäßiger auszutauschen, zu aktualisieren und / oder zu schützen, damit Bewerber*innen, die Zugang zu diesen Informationen (bspw. durch Internetrecherche) haben, keinen

systematischen Vorteil gegenüber anderen Bewerber*innen gewinnen können. Auch hier sollte ein entsprechender Hinweis in der DIN SPEC 91426 aufgenommen werden. Alternativ wäre ein Hinweis hilfreich, der dafür sensibilisiert, mit Aufgabenstellungen zu arbeiten, deren Kenntnis vorab den Bewerber*innen keinen eindeutigen Vorteil verschafft.

Probleme mit der technischen Umsetzung: In den Ausführungen hierzu wird der Fakt vernachlässigt, dass die Teilnahme über unterschiedliche Endgeräte zu einer unterschiedlichen Bewertung der Leistung durch die Kommission führen kann (Bewertungs-Bias, der in der Technik begründet liegt). Möglich wäre hier z.B. darauf hinzuweisen, dass Bewerber*innen eines Online-Auswahlverfahrens darum gebeten werden, möglichst einheitlich mit einer Sorte „Endgerät" teilzunehmen, auf optimale Kamera- und Toneinstellungen hingewiesen werden und auch das Auswahlgremium aus Vergleichbarkeitsgründen um die Nutzung einer Sorte „Endgerät" gebeten wird. Konkrete Hinweise, welche technischen Rahmenbedingungen unbedingt gegeben sein müssen (z.B. zuverlässige Internetverbindung), könnten hilfreich für alle Beteiligten sein.

Problem der systematischen Unterschiede in der Leistung der Bewerber*innen (Präsenz vs. virtuell): Anwender*innen der DIN SPEC 91426 sollten darüber aufgeklärt werden, dass sich die Wahrnehmung von bzw. das Verhalten der Bewerber*innen in Online-Auswahlverfahren von denen im Präsenz-Verfahren unterscheidet. Ggf. sind aktuelle Erkenntnisse aus Wissenschaft und Praxis in der Zukunft einzuflechten. Mindestens sollte darauf eingegangen werden, dass man ein Auswahlverfahren nicht sowohl mit Bewerber*innen, die online teilnehmen, und anderen, die in Präsenz teilnehmen, gestalten sollte. So wiese man zumindest darauf hin, dass Vorsicht geboten ist, und andere Wahrnehmungsmechanismen zugrunde liegen, die gegebenenfalls zu ungleichen Voraussetzungen für Bewerber*innen führen können.

Anwendungsbereich der DIN SPEC 91426

Neben den auszubauenden Hilfestellungen ist der Anwendungsbereich der DIN SPEC 91426 teilweise unnötig eingeschränkt. Der Anwendungsbereich beschränkt sich auf Interviewformate und vernachlässigt dabei andere, eher simulationsorientierte, videobasierte Auswahlinstrumente wie z.B. Rollenspiele, Gruppendiskussionen, Präsentationsaufgaben – Verfahren, die in der Regel Remote-Verfahren genannt werden. Auch mit der mittlerweile gängigen Videokonferenz-Technik können DIN 33430-konforme Verfahren zur Verhaltensbeobachtung und -beurteilung durchgeführt werden, die in der DIN SPEC 91426 nicht erwähnt werden. Das erweckt den Eindruck, dass diese Instrumente virtuell nicht umzusetzen wären. Dass stattdessen das zeitversetzte Videointerview in die Kategorie der Verhaltensbeobachtung im Sinne der DIN 33430 eingeordnet wurde, erscheint kaum nachvollziehbar. Sowohl in einem zeitversetzten Videointerview wie auch in einem Echtzeit-Videointerview können den Bewerber*innen Fragen gestellt werden, die Rückschlüsse auf das Verhalten der interviewten Person erlauben. I.d.R. sind dies situative oder biographische Fragestellungen. Es ist demnach nicht klar, auf welchem Unterscheidungsmerkmal die Zuordnung zur „Verhaltensbeobachtung" bzw. zur „direkten mündlichen Befragung" erfolgte. Die Einordnung der DIN 33430 sieht hier Instrumente wie „Rollenspiele, Gruppendiskussionen, Präsentationsübungen, Arbeitsproben" vor (DIN 33430:2016-07; 5.1; S. 12).

Eignungsdiagnostische Einschränkungen

Eine inhaltliche Empfehlung der DIN SPEC 91426 beurteilen wir kritisch: Künstliche Intelligenz (KI), deren Algorithmen vom Anforderungsprofil unabhängig sind, sollen laut der DIN SPEC 91426 eingesetzt werden können. Dies ist aus eignungsdiagnostischer Sicht problematisch und legitimiert eignungsdiagnostisch weniger aussagekräftige Ansätze. Die DIN 33430 positioniert sich zurecht mehrfach für einen Anforderungsbezug von Auswahlverfahren. Für die Auswertung von Verfahren wird in der DIN 33430 (S. 19) klar formuliert: „Es dürfen nur Informationen zu anforderungsrelevanten Eignungsmerkmalen ausgewertet werden." In der DIN SPEC 91426 wird dieser eindeutigen Vorgabe ohne Angabe von Gründen widersprochen. Zweifelhaft ist, ob dieses Vorgehen den „Richtlinien für den verantwortungsvollen Einsatz von Künstlicher Intelligenz und weiteren digitalen Technologien in der Personalarbeit" des Ethikrats HR Tech entspricht (insbesondere Richtlinie 2, „Fundierte Lösungen"), die sich ebenfalls an der DIN 33430 orientieren.

Daneben ergeben sich noch folgende inhaltliche Kritikpunkte: Teilweise erscheinen Empfehlungen hier nicht sinnvoll, wie z.B. die Möglichkeit, das Mehrfachprinzip außer Kraft zu setzen, oder einen Konsenszwang bei größeren Wahrnehmungsunterschieden durchzusetzen

Wichtig ist anzumerken, dass sich die Technik den Voraussetzungen für treffsichere Eignungsdiagnostik anpassen muss und nicht die Eignungsdiagnostik so ausgelegt wird, dass sie gut zu technischen Lösungen von Personalauswahlverfahren verschiedener Anbieter*innen passt.

Allgemeine Hinweise zur Benutzer*innenfreundlichkeit

Der Benutzer*innenfreundlichkeit wird ein großer Raum in der DIN SPEC 91426 eingeräumt. Sicher, die empirische Überprüfung der Benutzer*innenfreundlichkeit wäre gewinnbringend, insbesondere mit Blick auf die Akzeptanz des Auswahlverfahrens seitens der Bewerber*innen. Zum Teil erscheinen die dort gestellten Vorgaben jedoch willkürlich. Warum wird explizit ein Vorgehen in der Überprüfung der Benutzer*innenfreundlichkeit vorgegeben? Muss die Benutzer*innenfreundlichkeit wirklich mittels einer DIN derart festgeschrieben werden oder sind auch andere Methoden und Vorgehensweisen zulässig? Festgeschriebenes Vorgehen der DIN: „Die Benutzerfreundlichkeit der VMP für Anwender*innen und die vorgesehenen Kandidat*innengruppen muss durch den Anbieter vor Markteinführung empirisch geprüft werden. Nach Markteinführung wird die Benutzerfreundlichkeit laufend anhand von Nutzerdaten bewertet und ggf. verbessert. Die Studien müssen dokumentiert und auf Nachfrage den (auch potenziellen) Anwender*innen zur Verfügung gestellt werden."

Insgesamt werden die Anforderungen an die Benutzer*innenfreundlichkeit sehr spezifisch beschrieben, und sie erscheinen für Anwender*innen und Anbieter*innen derartiger Verfahren eventuell eher als Hürde, als dass hieraus ein konkreter Nutzen gezogen werden kann. Insbesondere vor dem Hintergrund, dass einige wesentliche eignungsdiagnostische Schwierigkeiten gar nicht behandelt werden, wird sehr detailliert beschrieben, wie in der Benutzer*innenfreundlichkeit vorzugehen ist. Dabei entsteht wenig Spielraum in der Definition und Sicherstellung der Benutzer*innenfreundlichkeit. An anderer Stelle ergibt

sich aber an einem wesentlicheren Punkt ein viel größerer Spielraum, und zwar bei der Validität der Auswahlverfahren – einem der vorrangigsten Gütekriterien. Aus unserer Sicht legt die DIN SPEC 91426 hier eine nicht angemessene Gewichtung von unterschiedlichen Themen für Online-Auswahlinstrumente fest.

Fazit

Mit der Veröffentlichung der DIN SPEC 91426 zu videogestützten Personalauswahlverfahren könnte ein großer Schritt im Sinne der Qualitätssicherung solcher Auswahlverfahren getan werden. Das Dokument kann in einigen Fragestellungen Orientierung geben, dringend ist dabei aber gleichzeitige Orientierung an der DIN 33430 sicher zu stellen, wenn man von einer qualitätssichernden Norm sprechen möchte.

In der vorliegenden Form reglementiert die DIN SPEC 91426 einige Aspekte von Online-Auswahlverfahren noch unnötig, diskutiert wichtige Probleme noch nicht und fällt in Teilen fachlich hinter der DIN 33430 zurück. Es bedarf noch einer Überarbeitung, damit Anwender*innen, Bewerber*innen und Ratsuchende die DIN SPEC 91426 als ein für sich stehendes, qualitätssicherndes Dokument nutzen können.

Literatur:

Basch, J. M., Melchers, K. G., Kurz, A., Krieger, M., & Miller, L. (2021). It takes more than a good camera: Which factors contribute to differences between face-to-face interviews and videoconference interviews regarding performance ratings and interviewee perceptions?. Journal of business and psychology, 36(5), 921-940.

DIN (2016). DIN 33430: Anforderungen an berufsbezogene Eignungsdiagnostik. Berlin: Beuth.

DIN (2020). DIN SPEC 91426:2020-12: Qualitätsanforderungen für video-gestützte Methoden zur Personalauswal (VMP). Berlin: Beuth.

Richtlinien_Download_deutsch_final.pdf (ethikbeirathrtech.de)

Kontakt:

Julia Wagenschieber
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
wagenschieber@dgp.de

Dr. Christian Reiß
Diplom-Psychologe und Leiter
der dgp-Geschäftsstelle Düsseldorf
Deutsche Gesellschaft für Personalwesen e.V.
reiss@dgp.de

Forschung

Messung der Bedeutung von Demokratie – Methoden der Differenzierung S. 68

Vorstellung der dgp Testbatterie OV1 S. 86

Vorstellung der dgp Testbatterie OG1 S. 90

Messung der Bedeutung von Demokratie – Methoden der Differenzierung[1]

Ulrich Stadelmaier & Norma Osterberg-Kaufmann

Zusammenfassung:
In der aktuellen Forschung zu Einstellungen zur Demokratie mehren sich die Zweifel an der Gültigkeit der These, dass Demokratie weltweit überwältigende Unterstützung findet. Zahlreiche empirische Studien zeigen eine divergierende Bedeutung von Demokratie in Ländern jenseits des Westens – im Widerspruch zu bisherigen Annahmen eines universell homogenen Verständnisses von Demokratie. Die Entwicklung von Konzepten des „D-Worts" jenseits eurozentrischer Hegemonie wurde von einer kritischen Überprüfung methodischer Ansätze begleitet. Vor dem Hintergrund dieser methodischen Herausforderungen werden in der Literatur die Forderungen nach differenzierteren oder mixed-method-Ansätzen diskutiert. In diesem Artikel wird ein kombinierter Ansatz aus Repertory-Grid und semantischem Differenzial vorgeschlagen, um die innovative methodische Dynamik der Untersuchung der Bedeutung von Demokratie zu bereichern. Der Artikel gibt eine Einführung in die Repertory-Grid-Methode und veranschaulicht anhand einer Fallstudie, wie die singapurische Mittelschicht Demokratie sieht. Diese Repertory-Grid-Face-to-Face-Interviews dienen als Ausgangspunkt für die Erstellung von validen Polaritätsprofilen für die semantische Differenzialmethode – eine Methode, die wie das Repertory-Grid zur Messung der konnotativen oder affektiven Bedeutung von Objekten eingesetzt wird, jedoch in einem quantitativen Design, so dass Repräsentativität erreicht werden kann. Durch diese Vorgehensweise wird der konstruktivistische Ansatz des Repertory-Grid teilweise mit positivistischer Umfrageforschung kombiniert – und damit induktive mit deduktiver Forschung.

[1] Beim vorliegenden Beitrag handelt es sich um eine gekürzte Überführung des folgenden Artikels ins Deutsche: Osterberg-Kaufmann, N. & Stadelmaier, U. (2020). Measuring meanings of democracy—Methods of differentiation. Zeitschrift für Vergleichende Politikwissenschaft. https://doi.org/10.1007/s12286-020-00461-6

1. Einleitung

In der aktuellen Forschung zu Einstellungen zur Demokratie werden zunehmend Zweifel an der Gültigkeit der These einer weltweit überwältigenden Unterstützung der Demokratie geäußert (Dalton et al. 2008; Diamond 2010; Welzel 2014; Lu und Shi 2015; Cho 2015; Welzel und Kirsch 2017). Zahlreiche empirische Studien (Schubert 2012; Cho 2015) zeigen eine divergierende Bedeutung von Demokratie in Ländern jenseits des Westens, was im Widerspruch zu bisherigen Annahmen eines universellen, homogenen Demokratieverständnisses steht. Aus diesem Grund soll auch hier explizit zwischen Demokratieverständnis und Demokratiebedeutung[2] unterschieden werden. Während Demokratieverständnis individuell unterschiedliche Repräsentationen eines immer noch identischen Gegenstandes meint und damit eine einheitliche Messung möglich ist, lässt das Konzept der Demokratiebedeutung begriffliche Mehrdeutigkeit zu. Begriffliche Mehrdeutigkeit wird für die Wissenschaft im Allgemeinen immer wichtiger, um rein lineare Ansätze hinter sich zu lassen und der Volatilität, Unsicherheit, Komplexität und Mehrdeutigkeit unserer heutigen globalisierten Welt gerecht zu werden. Insbesondere in der Organisations- und Managementforschung haben solche Ansätze bereits fruchtbare alternative Erklärungen hervorgebracht (Jarzabkowski et al. 2017; Miron-Spektor 2018; Waldman 2019).

Die Abkehr von der eurozentristischen Hegemonie (Yildiz 2012) der Konzepte des „D-Wortes" (Bratton 2010) wurde von einer kritischen Überprüfung der methodischen Ansätze begleitet. Angesichts dieser methodologischen Herausforderungen werden in der Literatur die Anforderungen an differenziertere Ansätze (Yildiz 2012) bzw. Mixed-Methods-Ansätze (Pickel 2009) diskutiert.

Dieser Artikel wird einen kombinierten Ansatz aus Repertory-Grid und semantischem Differenzial vorschlagen, um die innovative methodische Dynamik der Erforschung der Bedeutung von Demokratie als Ergänzung zu deduktiven durch induktive Methoden zu bereichern (Schubert 2012; 2016; Cho 2015). Die spezifischen Vorteile der Repertory-Grid-Methode liegen darin, dass die Ausgangsdaten qualitativ sind, indem die Sprache, Wörter und Assoziationen der Befragten verwendet werden, während die Analyse statistische Methoden verwendet, die einen Vergleich der Konzepte und Bedeutungen ermöglichen. Und da die Ergebnisse direkt auf dem individuellen Wertekontext der Befragten beruhen, werden Probleme der Äquivalenz, der sozialen Erwünschtheit und der Lippenbekenntnisse reduziert. Der Einsatz des Repertory-Grid ermöglicht daher einen völlig neuen Ansatz in der transnationalen und kulturübergreifenden Forschung, neue Einblicke in den internationalen Vergleich und ein differenzierteres Verständnis dessen, was beispielsweise Demokratie für die Menschen bedeutet. Da einer groß angelegten Face-to-Face-Umsetzung des Repertory-Grid klare Grenzen gesetzt sind, da die methodisch gut strukturierten Interviews Zeit benötigen und der damit verbundene Aufwand im Vergleich zu standardisierten Befragungen relativ hoch ist, sind repräsentative Stichproben und damit verallgemeinerbare Ergebnisse kaum möglich. Eine methodische Lösung wäre eine Online-Befragung im Anschluss an die persönlichen Interviews (Kaulartzs und Heckmann 2014). Auf Basis von Repertory-Grid-Face-to-Face-Interviews können valide Polaritätsprofile für semantische Differenziale erstellt

[2] Zur Differenzierung zwischen Demokratieverständnis (understanding of democracy) und Demokratiebedeutung (meanings of democracy) siehe auch Osterberg-Kaufmann et al. 2020 Referenz: Osterberg-Kaufmann, N., Stark, T. & Mohamad-Klotzbach, C. Challenges in conceptualizing and measuring meanings and understandings of democracy. Z Vgl Polit Wiss 14, 299–320 (2020). https://doi.org/10.1007/s12286-020-00470-5

werden – ein Verfahren, wie das Repertory-Grid, zur Messung der konnotativen oder affektiven Bedeutung von Objekten, jedoch in einem quantitativen Design, sodass Repräsentativität erreicht werden kann (Osgood 2009; Rosenberg und Navarro 2018).

Der Artikel will einen Beitrag zur Diskussion über den angemessenen methodischen Ansatz leisten, um die Bedeutung von abstrakten Konzepten wie Demokratie jenseits westlicher Normativität zu erforschen.

Im Kontext der Frage nach den methodischen Herausforderungen der Demokratieforschung will der Artikel einen Beitrag zum induktiven/bottom-up Diskussionsstrang leisten. Daher wird der Artikel die Vorteile von Mixed-Methods-Konzepten zur Messung der Bedeutung von Demokratie unter der Travelling-Prämisse (Schubert und Weiß 2016) reflektieren. Im nächsten Kapitel wird der Artikel das Repertory-Grid als eine solche gemischte Methode vorstellen und anhand der Ergebnisse einer Fallstudie illustrieren, wie die singapurische Mittelschicht Demokratie sieht. Im letzten Schritt wird der Artikel die semantische Differenzialmethode als eine Möglichkeit vorstellen, die Ergebnisse der Repertory-Grid-Interviews als induktiven/bottom-up-Ansatz in ein deduktives/top-down-Design zu integrieren, um eine Studie mit großer Stichprobe und damit Repräsentativität zu ermöglichen. Durch diese Vorgehensweise wird der konstruktivistische Ansatz des Repertory-Grid in gewisser Weise mit positivistischer Umfrageforschung kombiniert.

2. Methodisch differenzierte Ansätze und Mixed-Methods-Designs

Die gängige empirische Forschung zum Demokratieverständnis ist in der Lage, das normative Verständnis von Demokratie (deduktiv/top-down) aus westlicher Perspektive abzubilden, nicht aber die heterogene Bedeutung von Demokratie weltweit zu erfassen. Die Politische Theorie fordert eine Methode mit einem innovativeren und differenzierteren Umgang mit der Frage nach der weltweiten Bedeutung von Demokratie (Yildiz 2012). Diskutiert als Travelling-Prämisse (Schubert und Weiß 2016) argumentiert Yildiz auf der Grundlage der Idee, dass politische Ideen und Begriffe über Kulturen hinweg reisen. Die Travelling-Prämisse besagt, dass sich beispielsweise Ideen und Techniken des Regierens über Zeit und Raum verändern. Sie werden angepasst und können dabei auch widersprüchliche Elemente enthalten. Die Vergleichende Politikwissenschaft sollte keine Mühe scheuen, diese adaptierten Bedeutungen zu verstehen (Yildiz 2012, S. 216). Stattdessen folgt die Vergleichende Politikwissenschaft der Einschätzung von Sartori (1997), der davon ausgeht, dass es den einen und wahren Kern der Demokratie geben soll, den es nur zu identifizieren gilt. Und wenn diese wahren Werte und Ideale der Demokratie identifiziert sind, neigt der von Sartori (1997) inspirierte Forscher dazu, nicht-westliche Bedeutungen anhand vorgefertigter (westlicher) Kriterien zu analysieren. Ein solcher Ansatz führt, der Argumentation von Yildiz folgend, zu einer Version von Modernisierung, die eurozentrisch und normativ überhöht ist (Dallmayr 2010).

Unter Berücksichtigung dieser Überlegungen gibt es erste Ansätze, die Methoden zu differenzieren und zu erneuern, um die Bedeutung von Demokratie weltweit zu verstehen. Indem sie die zehn Demokratie-Items des World Value Survey Wave 2005-2008[3] zusammenfasst, versucht Schubert (2012) beispielsweise, Gemeinsamkeiten und Unterschiede hinter diesen Items zu identifizieren. Dabei identifizierte sie einen Kern von Demokratie, bestehend aus Wahlen, Gleichheit, Bürgerrechten und Referenden, der westlichen Bedeutungen entspricht. Gleichzeitig beobachtete sie Unterschiede zwischen den Ländern und Unstimmigkeiten über den Demokratiebegriff bei nicht-westlichen Ländern und sogar ganz eigene Definitionen von Demokratie (Schubert 2012, S. 198). Diese Befunde deuten auf eine heterogene Bedeutung von Demokratie mit einem demokratischen Kern nach westlichem Verständnis hin, wie auch Chu und Huang (2010) feststellten. Was bei diesem Ansatz neben dem gemeinsamen Kern nicht berücksichtigt wird, sind die unterschiedlichen Auffassungen der Menschen von Demokratie. Mittels multidimensionaler Skalierung verglich Schubert (2012) daher die Relationen der zehn Demokratie-Items in den Vorstellungsräumen der Befragten. Durch Annäherungen und Abstände der Items können Gemeinsamkeiten und Unähnlichkeiten, in diesem Fall typisch demokratische und undemokratische Elemente der Demokratiedefinition, dargestellt werden. Diese Herangehensweise führt zu dem Ergebnis, dass es eine idealtypische, eine realtypische[4] und eine atypische Definition[5] von Demokratie gibt, die z.B. religiöse Führung in das gemeinsame Kernverständnis von Demokratie einschließt[6]. In den atypischen und realtypischen Definitionen finden sich ausschließlich nicht-westliche Länder (Schubert 2012, S. 204ff.)[7].

Die multidimensionale Skalierung ist ein großer Fortschritt, um die Unterschiede in der Sichtweise der Menschen auf die Demokratie zu systematisieren, ist aber immer noch an Nachteile der Umfrageforschung gebunden, wie oben diskutiert. Nächste Schritte sollten Ansätze sein, die Vorteile und Potenziale der quantitativen Umfrageforschung und qualitativer Methoden kombinieren, um die reine Messung normativ definierter Elemente der Demokratie zu überwinden und die individuellen Bewertungskriterien der Befragten selbst zu verstehen. Mögliche Forschungsdesigns könnten Mehrebenen- oder Mixed-Methods-Verfahren sein. Während Mixed-Methods verschiedene methodische Ansätze parallel innerhalb einer Studie impliziert, integriert Multi-Level diese innerhalb der jeweiligen Forschungsphase.

3. Repertory-Grid als alternativer Ansatz

Trotz aller Innovationen innerhalb des politikwissenschaftlichen Methodenspektrums bei der Erhebung von Einstellungen, für die die multidimensionale Skalierung als herausragendes Beispiel steht, stößt diese Perspektive dennoch an ihre Grenzen, weshalb wir in diesem Artikel darüber hinausgehen und über den disziplinären Horizont hinausschauen.

Die Messung der Bedeutung von Demokratie ist an sich ein psychologisches Messpro-

[4] Die vier klassischen Elemente der idealen Definition von Demokratie werden mit sozioökonomischen Elementen und Sicherheit kombi-niert. Der Gegensatz zu den atypischen Elementen bleibt bestehen.
[5] Bestehend aus Wahlen, Bürgerrechten, Gleichberechtigung und Volksabstimmungen, entsprechend der westlichen Bedeutung von Demokratie - gegen die atypischen Elemente (Militär und Religion).
[6] Die Trennung von demokratischen und nicht-demokratischen Elementen ist aufgelöst.
[7] Atypische Definition: Indonesien, Ägypten, Ghana, Jordanien, Irak; real-typische Definition: Rumänien, Marokko, Mexiko, Südafrika, Bulgarien, Thailand, China, Sambia, Malaysia, Trinidad und Tobago, Peru, Russland, Japan, Südkorea, Moldawien, Zypern und Serbien.

blem, weil der Forscher erforschen will, wie Individuen darüber denken und fühlen und wie diese individuellen Wahrnehmungen zu einer verallgemeinerbaren Skala aggregiert werden können, die auf ganze Populationen angewendet werden kann.

Die Psychometrie beschäftigt sich mit ähnlichen Fragestellungen in der Psychologie. Als interdisziplinärer Ansatz erscheint daher die Anwendung psychometrischer Methoden zur Erhebung differenzierterer Bedeutungen von Demokratie vielversprechend. Im Folgenden wird daher der psychometrische Messansatz kurz vorgestellt.

Auf der Basis der jeweiligen Fachtheorie muss ein sogenanntes Itemuniversum abgeleitet werden. Das Itemuniversum definiert, wie der interessierende Gegenstand mit welchen Aussagen und entsprechenden Teilnehmerantworten gemessen werden kann. Nachdem dieses theoretische Item-Universum definiert wurde, wird es mit Inhalten gefüllt. Dies geschieht in der Weise, dass die definierte Aussagenstruktur mit verschiedenen bewertbaren Erfahrungs- und Verhaltensbeispielen gefüllt wird. Dieser Vorgang wird als Itemgenerierung bezeichnet. Im Zuge der Itemgenerierung produziert man in der Regel mindestens dreimal so viele Items wie man später im endgültigen Messinstrument verwenden möchte. Der Grund dafür ist die sogenannte Skalen- und Itemanalyse, bei der sich meist die wenigsten Items als valide und reliabel erweisen (Moosbrugger und Kelava 2008; Rost 2004). Bei diesen theoriegeleiteten Schritten werden die (zukünftigen) Probanden nicht einbezogen. Dadurch besteht die Gefahr, dass entscheidende Aspekte der Wahrnehmung des Messobjekts verloren gehen, was im Hinblick auf die begrenzte Erklärungskraft quantitativer Umfrageforschung auch für die quantitativen Ansätze in der Politikforschung zu gelten scheint (siehe Kapitel Einleitung). Ausschließlich theoretische Top-Down-Ansätze zur quantitativen Messung des Demokratieverständnisses lassen bei der Verallgemeinerung westlicher Konzepte auf andere Kulturen offenbar konzeptionelle Tiefe vermissen.

Sowohl in der Politikforschung als auch in der Psychometrie kann diese Einschränkung durch die Befragung von Personen der Zielpopulation entlang der Itementwicklungsschritte verringert werden.

Daraus ergibt sich zwar eine Fülle von Information, aber die Interpretation, Strukturierung und Priorisierung ist immer noch ein rein qualitatives Urteil des Testentwicklers. Eine ausschließlich qualitative Bottom-up-Messung mag also konzeptionell tiefer sein als eine theoriebasierte quantitative Messung, letztere ist aber viel stärker von unkontrollierten Faktoren beeinflusst, vergleichbar mit den im Einführungskapitel besprochenen qualitativen Ansätzen.

Das Grundproblem für eine umfassende Beurteilung von etwas, das man messen möchte, besteht in der Regel darin, dass der Forscher den Befragten entweder sagt, was sie zu beantworten haben (quantitativer Top-down-Ansatz) (Jankowicz 2004). Oder, wenn qualitative Erweiterungen hinzukommen, besteht die Gefahr, dass das Messobjekt aus einer stark subjektiven Perspektive des Forschers betrachtet wird (qualitativer Bottom-up-Ansatz).

Das Repertory-Grid stellt eine vielversprechende methodische Alternative zum üblichen Vorgehen dar (Kelly 1980; Walker und Winter 2007), insbesondere wenn Forscher eine breitere Bedeutung eines Konzepts erforschen wollen, also auf der Suche nach dessen Differenzierung sind. Das Repertory-Grid bildet die Brücke zwischen qualitativen und quantitativen Methoden und kann die Grundlage für einen Mixed-Methods/Multi-Level-Ansatz sein, der qualitative und quantitative Instrumente innerhalb einer einzigen Studie kombiniert und die Daten für eine anschließende groß angelegte Befragung nutzt. In diesem Sinne können Repertory-Grids nach Jankowicz (2004) als eine strukturierte Interviewmethodik, kombiniert mit einem multidimensionalen Ratingskalenansatz, betrachtet werden.

Repertory-Grids bestehen aus vier Aspekten (Jankowicz 2004; Shcheglova 2010): Der erste Aspekt ist das Thema. Das Thema repräsentiert die Erfahrungen der (zukünftigen) Befragten innerhalb des Interessensgebietes des Forschers. Im vorliegenden Fall ist das Thema Demokratie. Der zweite Aspekt des Repertory-Grid ist die Menge der Elemente. Die Elemente sind Beispiele für das Thema (z.B. Institutionen, Personen o.ä., die repräsentativ für Demokratie sind). Der dritte Aspekt eines Repertory-Grid sind die sogenannten Konstrukte. Dieser Begriff stammt aus der humanistischen Psychologie, in der davon ausgegangen wird, dass jeder Mensch seine eigene Welt durch Erfahrung konstruiert. In seinen Konstrukten spiegelt sich seine Konstruktion der Wirklichkeit wider. Die Konstrukte der Befragten in der Fallstudie in Singapur stellen die individuellen Konstruktionen von Demokratie dar und stehen daher im Mittelpunkt des Interesses. Einer der theoretischen Pioniere der Psychologie der persönlichen Konstrukte, George A. Kelly, nahm an, dass ein Konstrukt immer bipolar ist (Kelly 1980). In gewissem Sinne kann „links“ nur dadurch verstanden werden, dass der Gegenpol „rechts“ ist. Der vierte Aspekt eines Repertory-Grid sind die quantitativen Bewertungen der Elemente durch die Befragten auf deren Konstrukten.

Ein Repertory-Grid-Interview wird üblicherweise in den folgenden Schritten durchgeführt. Den Befragten werden drei Elemente präsentiert, und sie werden gefragt, welche zwei davon sich ähneln und sich gleichzeitig vom dritten unterscheiden. Die Befragten nennen dann die Art und Weise, in der sich die beiden Elemente ähneln. Dies ist der erste Konstruktpol. Der gegenüberliegende Konstruktpol wird ermittelt, indem die Befragten gebeten werden zu beantworten, in welcher Weise sich das dritte Element von den beiden ähnlichen unterscheidet. Nachdem die beiden Konstruktpole ermittelt wurden, führt der Befragte ein Rating aller Elemente auf dem Konstrukt durch. Dieser Vorgang wird mehrmals mit verschiedenen Element-Triaden wiederholt. Als Faustregel gilt, dass ca. 10 Konstrukte pro Proband*in ermittelt werden. Einerseits erhalten Forscher*innen dadurch ein äußerst profundes Bild der individuellen Bewertungsmuster. Andererseits entsteht durch die jeweiligen Bewertungen eine Datenstruktur, die mit multivariaten statistischen Methoden analysiert werden kann. Tabelle 1 auf der folgenden Seite zeigt eine Datenstruktur, die sich aus dem Interview eines einzelnen Teilnehmers ergeben hat.

Konstrukte	Bürgerrechte	gleiche Rechte	Wahlen	legitimste Herrschaftsform	Demokratie	geeignetste Regierungsform	Sicherheit	Wohlstand	Ihr eigenes Land	Einkommensgleichheit	Umverteilung	Arbeitslosengeld	Gehorsam	Militär	religiöse Führer	
frei	55	70	85	95	85	95	25	65	50	45	55	70	15	0	0	stimmlos
ansprechbar	65	65	60	65	65	80	25	65	75	70	70	65	15	10	10	aufzwingend
gleich	75	75	55	70	65	75	55	50	60	80	70	70	10	10	10	hierarchisch
Fairness	75	60	70	90	95	90	50	70	90	80	60	70	40	5	5	Vertrauen
Überleben	75	60	65	70	65	80	65	70	70	75	65	65	50	40	5	Unterdrückung
erfolgreich	20	25	40	55	55	75	60	60	65	70	75	75	5	0	5	dominierend
	Bewertungen															

Tabelle 1: Datenstruktur eines Repertorium-Gitters (Repertory-Grid) eines einzelnen Teilnehmers. Quelle: Eigene Daten.

Bei einem multiplen Repertory-Grid, wie es in der vorliegenden Studie verwendet wurde, werden alle Teilnehmer*innen zu den gleichen Elementen befragt, aber die Teilnehmer*innen tragen ihre individuellen Konstrukte zur Datenstruktur bei. Die Gemeinsamkeiten hinter den einzelnen Bewertungsmustern können dann mit den üblichen Methoden der Hauptkomponentenanalyse sowie der Clusteranalyse ausgewertet werden.

All diesen inhaltlichen Vorteilen steht der erhebliche praktische Nachteil gegenüber, dass Repertory-Grid-Interviews für den Forscher extrem aufwendig sind und bei groß angelegten, multinationalen Erhebungen einen unverhältnismäßig hohen Zeit- und Ressourcenaufwand erfordern. Daher haben sich die Autor*innen entschieden, eine Repertory-Grid-Analyse als Grundlage für ein semantisches Differenzial zu verwenden, das auf einer methodischen Lösung basiert, die auf den Vorteilen von Repertory-Grid-Analysen aufbaut, aber große Stichproben ermöglicht.

4. Datenanalyse und Ergebnisse: Wie die singapurische Mittelschicht die Demokratie sieht

Die Analyse von Repertory-Grid-Daten und die Übertragung dieser Daten in ein semantisches Differenzial wird anhand eines Auszugs aus einer Studie über die Sicht der singapurischen Mittelschicht auf die Demokratie illustriert. Als erster Schritt eines Mehrebenenansatzes wurde Singapur als prototypischer Ausreißer im Kontext der Diskussion um die Bedeutung von Demokratie identifiziert. Die Auswahl dieses Falles basiert auf einer Regression von Lu und Shi (2015), in der sie für nicht-westliche Länder zeigen konnten, dass in einigen Fällen, insbesondere in China, Vietnam, Thailand und Singapur, die Bewertung des eigenen Landes als Demokratie von derjenigen von Freedom House abweicht. Aktuellere Umfragedaten unterstützen die Analysen von Lu und Shi (2015). Abbildung 1 veranschaulicht die aktuellen Ergebnisse. Die wahrgenommene demokratische Qualität wird durch den Prozentsatz der Befragten operationalisiert, die positive Antworten auf die Frage „Wie zufrieden oder unzufrieden sind Sie mit der Art und Weise, wie die Demokratie in [Land] funktioniert“ geben. Dem wird der Freedom House Score als objektive Bewertung der Qualität der Demokratie gegenübergestellt.

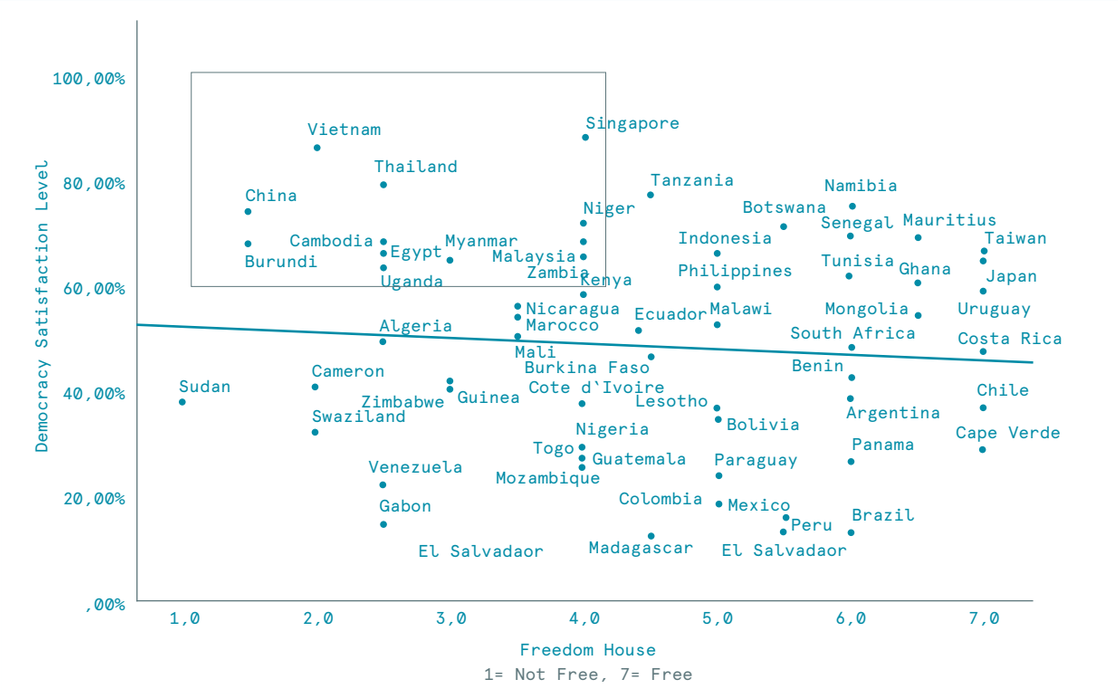


Abbildung 1: Wahrgenommene demokratische Qualität und Freedom-House-Score
Quelle: Eigene Berechnungen in Osterberg-Kaufmann und Stadelmaier (2020) mit Daten aus Freedom in the World 2017, African (Round 6 in 2016), Asian (Wave 4 von 2014 bis 2016)[8], und Latino (Survey 2017) Barometers. Zufriedenheitsgrad operationalisiert durch Fragen wie „Wie zufrieden oder unzufrieden sind Sie mit der Art und Weise, wie die Demokratie in [Land] funktioniert“. Die Antworten „Sehr zufrieden“ und „Eher zufrieden“ (im Gegensatz zu „Nicht sehr zufrieden“ und „Überhaupt nicht zufrieden“) wurden als positive Antworten kategorisiert, die eine Zufriedenheit mit der Demokratie anzeigen.

[8] Die in diesem Artikel analysierten Daten wurden im Rahmen des Asian Barometer Project (2013-2016) erhoben, das von den Professoren Fu Hu und Yun-han Chu mitgeleitet wurde und vom taiwanesischen Bildungsministerium, der Academia Sinica und der National Taiwan University maßgeblich finanziell unterstützt wurde. Das Asian Barometer Project Office (www.asianbarometer.org) ist allein für die Datenverteilung verantwortlich. Die Autor*innen bedanken sich für die Unterstützung bei der Bereitstellung von Daten durch die oben genannten Institute und Personen. Die hierin zum Ausdruck gebrachten Ansichten sind die der Autor*innen.

Definiert man Ausreißer als Länder mit einem Freedom-House-Score von 4 („Partly Free“) und darunter in Kombination mit einer Demokratiezufriedenheit von 60 Prozent und mehr, so erfüllen Sambia, Malaysia, Niger, Singapur, Myanmar, Uganda, Ägypten, Kambodscha, Thailand, Vietnam, Burundi und China diese Bedingungen.[9] Singapur ist in dieser Gruppe von Ausreißern von besonderem Interesse, weil seine Voraussetzungen für eine Demokratisierung nach westlichliberaler Demokratietradition im Vergleich zu den anderen Fällen äußerst vorteilhaft gewesen wären. Aber mit Freedom House (2017) kann das politische System Singapurs nur als teilweise frei (4) charakterisiert werden. Singapur ist nicht demokratisierend (Ortmann 2011), sondern der Ursprung der sogenannten asiatischen Wertedebatte, in welcher der ehemalige Premierminister die Besonderheiten der asiatischen Kultur und deren Unvereinbarkeit mit der Demokratie betonte. Diese Debatte fand auch ihren Weg in die Wissenschaft (Emmerson 1995, Sen 1999). Wie die singapurische Mittelschicht die Demokratie sieht, wird daher im Fokus der Repertory-Grid-Analyse stehen, die in diesem Kapitel skizziert wird. Die Mittelschicht ist in dieser Debatte von besonderem Interesse, da es eine große Diskussion über die asiatische Mittelschicht und die Präferenz für Demokratie gibt (Koo 1991; Chen 2013; Glassman 1991; Goodman 1999; Chen 2002; Xiao 2003; He Li 2003; O‘Brian und Stern 2008; Zhang 2010; Wu et al. 2017).

4. 1 Vorbereitende Schritte für die Repertory-Grid-Technik: die Auswahl der geeigneten Elemente

Bei der Repertory-Grid-Methodik müssen die Elemente adäquate Repräsentationen des Themas sein (Dinge, Personen, Situationen, z.B. (Jankowicz 2004)). Wie in der aktuellen Studie zur Bedeutung von Demokratie besteht das entwickelte Set von Elementen, als erster Schritt und Grundlage des Repertory-Grid-Interviews, aus 15 theoretischen Repräsentationen der Demokratieforschung, nämlich „Demokratie“, „Wahlen“, „Bürgerrechte“, „Gleichberechtigung“, „geeignetste Regierungsform“, „legitimste Herrschaftsform“, „Einkommensgleichheit“, „Umverteilung“, „Arbeitslosengeld“, „Militär“, „religiöse Führer“, „Gehorsam“, „Sicherheit“, „Wohlstand“ sowie das Referenzelement „eigenes Land“. Die empirische Grundlage der Elementauswahl orientiert sich an den Items des World Value Survey (Welle 6) zum Demokratieverständnis; dazu kommen das normative Element „legitimste Herrschaftsform“, das realistische Element „geeignetste Regierungsform“, „Ihr eigenes Land“, um das politische System Singapurs im Sinne der Demokratie zu bewerten sowie „Sicherheit“ und „Wohlstand“, um zwei Elemente der Output-Legitimität einzubeziehen, die eine Option zur Legitimierung nicht-demokratischer Alternativen sein könnten. „Wahlen“ (70,9%)[10] , „Gleichberechtigung“ (66,8%) und „Bürgerrechte“ (57,9%) stehen für den liberalen Kern der Demokratie, während „Militär“ (20,2%) und „religiöse Führer“ (17,4%) nicht-demokratische Elemente darstellen. Soziale oder sozialpopulistische Aspekte der Demokratie werden durch „Arbeitslosengeld“ (52,1%), „Umverteilung“ (42,1%) und „Einkommensgleichheit“ (36,5%) abgedeckt. Und „Gehorsam“ (37%) deckt ein eher populistisch-autoritäres Element der Bedeutung von Demokratie ab (Schubert 2016, S. 292 ff.).

4.2 Repertory-Grid-Analyse durch räumliche Anordnung von Elementen und Konstrukten

Das oben beschriebene Verfahren der Repertory-Grid-Befragung führt zu einem Datensatz, der der Ausgangspunkt für quantitative Analysen mittels Cluster- und Hauptkomponentenanalyse ist. Die Hauptkomponentenanalyse eines multiplen Repertory-Grid liefert zwei grundlegende Informationen darüber, wie eine Gruppe von Personen über ein bestimmtes Thema denkt und fühlt. Die allererste Information ist die räumliche Anordnung der einzelnen Elemente. Je weiter die Elemente voneinander entfernt sind, desto unterschiedlicher werden sie wahrgenommen. Je näher sie beieinander liegen, desto ähnlicher werden sie wahrgenommen. In der vorliegenden Studie wurde der räumliche Abstand der Elemente durch ihre Faktorwerte für die identifizierten Hauptkomponenten bestimmt. Die zweite Information eines multiplen Repertory-Grid ist die räumliche Anordnung der Konstrukte im Vergleich zu den räumlichen Anordnungen der Elemente. Diese wurde in der vorliegenden Studie durch die jeweiligen Faktorladungen eines Konstrukts auf jeder Hauptkomponente bestimmt. Je näher ein Konstrukt an einem Element liegt (ermittelt durch die Faktorwerte der Elemente im Vergleich zu den Faktorladungen der Konstrukte), desto charakteristischer ist dieses Konstrukt für dieses Element.

Für die statistischen Analysen wurden die Programmumgebung R (R Core Team 2018) und die Software congrid® (permit-to GmbH 2019) verwendet. Als Ergebnis der Interviews haben die 17 Teilnehmer*innen insgesamt 190 Konstrukte geliefert. Daraus wurden drei Hauptkomponenten extrahiert, die in Summe 87 Prozent der Varianz erklären. Die congrid®-Software liefert eine grafische Übersicht über die beschriebene räumliche Struktur des multiplen Repertory-Grid, wie in Abbildung 2 auf der nächsten Seite[11]. Die zugrunde liegenden euklidischen Distanzen sind in Tabelle 2 dokumentiert.

[9] Die in diesem Artikel analysierten Daten wurden im Rahmen des Asian Barometer Project (2013-2016) erhoben, das von den Professoren Fu Hu und Yun-han Chu mitgeleitet wurde und vom taiwanesischen Bildungsministerium, der Academia Sinica und der National Taiwan University maßgeblich finanziell unterstützt wurde. Das Asian Barometer Project Office (www.asianbarometer.org) ist allein für die Datenverteilung verantwortlich. Die Autor*innen bedanken sich für die Unterstützung bei der Bereitstellung von Daten durch die oben genannten Institute und Personen. Die hierin zum Ausdruck gebrachten Ansichten sind die der Autor*innen.

[10] Die Prozentzahlen verdeutlichen die Zustimmung der World Value Survey Welle 6-Befragten, dass diese Elemente die Demokratie charakterisieren (Schubert 2016, S. 293 f.).

[11] Die räumliche Anordnung von Elementen und Konstrukten ist am aussagekräftigsten, wenn sie durch interaktive Software visualisiert wird.

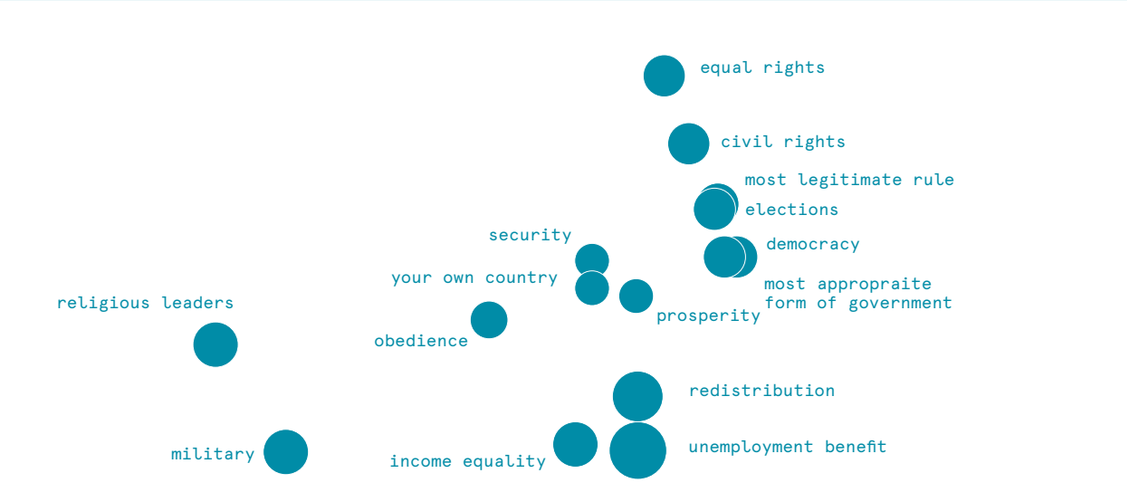


Abbildung 2: Räumliche Anordnung der Elemente
Quelle: Osterberg-Kaufmann und Stadelmaier (2020, S. 412).

	civil rights	democracy	elections	equal rights	income equality	military	most appropriate from of government	most legitimate rule	obedience	prosperity	redistribution	religious leaders	security	unemployment benefit	your own country
civil rights		427	408	273	572	773	382	345	612	495	507	789	495	528	490
democracy			364	472	644	765	395	406	621	487	576	775	526	579	479
elections				428	622	786	403	372	609	474	567	801	529	604	488
equal rights					566	770	426	370	615	514	517	768	517	534	505
income equality						634	565	560	562	596	406	676	556	361	532
military							786	784	541	688	644	355	657	629	633
most appropriate from of government								332	598	441	508	798	467	526	426
most legitimate rule									582	488	476	784	469	509	455
obedience										562	550	573	500	579	445
prosperity											553	705	463	569	414
redistribution												675	515	311	471
religious leaders													695	668	635
security														522	427
unemployment benefit															513
your own country															

Tabelle 2: Euklidische Distanzen zwischen Elementen.
Quelle: Osterberg-Kaufmann und Stadelmaier (2020, S. 413).

Die Daten zeigen, dass es eine Kernbedeutung von Demokratie gibt, die aus Wahlen als integralem Bestandteil des Demokratiekonzepts besteht. Sowohl Demokratie als auch Wahlen werden als die angemessenste Regierungsform (Demokratie) und die legitimste Herrschaftsform (Wahlen) angesehen. Das politische System Singapurs wird mit zwei Ausgangselementen identifiziert, Wohlstand und Sicherheit, die beide nicht als demokratisch wahrgenommen werden. Die beiden sozialdemokratischen Elemente der Umverteilung und der Arbeitslosenunterstützung gehören nicht zur Bedeutung der Demokratie. Definitiv nicht zum Demokratieverständnis der Befragten und mit der größten Distanz zur Demokratie gehören das Militär und die religiösen Führer als undemokratische Elemente (eine atypische Bedeutung von Demokratie kann daher ausgeschlossen werden) sowie Gehorsam und Einkommensgleichheit als soziale oder autoritär-populistische Elemente.

Weniger offensichtlich ist die Frage, ob Bürgerrechte und Gleichberechtigung zu diesem Kern der Demokratie gehören oder nicht. Eine mögliche Interpretation ist, dass es einen inneren Kern der Demokratie mit Wahlen gibt und einen erweiterten Kern mit Bürgerrechten, Gleichberechtigung und dem eigenen Land (Singapur) und einer Gruppe von Elementen, die aus Wohlstand und Sicherheit bestehen. Diese Interpretation würde die Annahme eines liberalen Kerns der Demokratie (Wahlen, Bürgerrechte, Gleichberechtigung) begünstigen. Die zweite Interpretation könnte sein, dass es eine Gruppe von Elementen gibt, die aus Bürgerrechten und Gleichberechtigung besteht und eine andere Gruppe, die aus dem eigenen Land (Singapur), Wohlstand und Sicherheit besteht. Diese Interpretation würde das Argument unterstützen, dass es nur eine minimalistische Bedeutung von Demokratie gibt, die aus Wahlen besteht, die nicht notwendigerweise Bürgerrechte und Gleichberechtigung beinhaltet. Typischerweise bedeutet die Feststellung, dass ein Element in der Mitte eines Repertory-Grid liegt, dass die Befragten diesem Element gegenüber indifferent sind[12] und in der qualitativen Analyse des Repertory-Grid fokussiert werden sollten.

Die Clusteranalyse aller Elemente identifizierte fünf Cluster erster Ordnung, die Ähnlichkeiten in der Art und Weise aufzeigen, wie die Teilnehmer*innen auf den Konstrukten bewertet haben. Die Analyse der fünf Cluster (I Demokratie und Wahlen, II Bürgerrechte und Gleichberechtigung, III geeignetste Regierungsform und legitimste Herrschaftsform, IV Arbeitslosengeld und Umverteilung, V Militär und religiöse Führer) anhand von Konstruktwolken ermöglicht ein besseres Verständnis dafür, wie die Befragten die Elemente und ihre gemeinsamen Merkmale beschrieben haben. Konstruktwolken sind eine congrid®-Analysefunktion, die Distanzen zwischen einzelnen Konstrukten und Gruppen von Elementen untersucht. So wird das Demokratie-Wahl-Cluster (I) als *fürsorglich*, *fair*, *fortschrittlich* und frei charakterisiert[13]. Neben der Charakterisierung dieses Clusters als *frei* und *fair* finden sich auch Beschreibungen, die näher an kommunitaristischen und gemeinschaftsorientierten Vorstellungen von Staat sind als an liberalen Vorstellungen von Demokratie. Der Bürger- und Gleichberechtigungs-Cluster (II) wird als *erfolgreich*, *gewählt*, *fair* und *frei* beschrieben. Das bedeutet, dass beide Cluster als *frei* und *fair* wahrgenommen werden, wobei für die Bürger- und Gleichberechtigung am

[12] Diese Indifferenzen müssen im Zusammenhang mit dem beobachteten Antwortverhalten während der Interviews gesehen werden. Die Befragten fanden relativ leicht gegensätzliche Konstrukte, aber bei der Bewertung der Elemente innerhalb dieser Konstruktskalen tendierten sie zu eher moderaten Bewertungen der Elemente.

[13] Kursivschrift steht für Originalformulierungen in den Interviews. Nicht-kursive Schrift steht für grammatikalische und, im Hinblick auf das semantische Differenzial, auch methodische Modifikationen der Autoren.

ehesten die Beschreibung *erfolgreich* und für die *Demokratie Wahlen* als *fürsorglich* zutrifft. Mit der Beschreibung von Cluster II als *frei* und *fair* und der gleichzeitigen Wahrnehmung als *erfolgreich* und *gewählt* scheint es also einerseits eine liberale Bedeutung von Demokratie zu geben und andererseits eine pragmatische, durch Kommunitarismus legitimierte Sichtweise von Good Governance.

Das Cluster angemessenste Regierungsform und legitimste Herrschaftsform (III) wird als *ermächtigend*, *fürsorglich* und *gemeinsam* wahrgenommen. Das Cluster Demokratie und Wahlen (I) und das Cluster angemessenste Regierungsform und legitimste Herrschaftsform (III) werden mit sehr ähnlichen Eigenschaften wahrgenommen (Abbildung 2). Was sie vor allem gemeinsam haben, wird als *Fürsorge* beschrieben. Auch wenn Freiheit und Fairness neben pragmatischen Erwägungen einen demokratierelevanten Wert zu haben scheinen, dominieren letztere bei der Wahrnehmung der angemessensten und legitimsten Herrschaft als *ermächtigend*, *fürsorglich* und *gemeinsam*. Dies führt zu einer Bestätigung der bisherigen Interpretation einer eher minimalistischen Bedeutung von Demokratie in der singapurischen Mittelschicht, die weitgehend durch Wahlen definiert ist. Auf der Basis einer größeren Datenbasis wäre es interessant zu untersuchen, ob das Attribut *fürsorglich* im Kontext der konfuzianischen Staatstradition im Sinne einer Vater-Sohn-Beziehung gesehen werden könnte und damit auf die asiatische Wertedebatte verweist, oder ob es eher aus der singapurischen Ideologie des Pragmatismus stammt.

Cluster IV, Umverteilung und Arbeitslosenunterstützung werden überwiegend als *ausgleichend*, *sicher*, *anpassungsfähig* und *konzessionsbereit* wahrgenommen. Und Cluster V Militär und religiöse Führer werden am häufigsten als *einschränkend* und *untertänig* beschrieben. Und Singapur schließlich wird hauptsächlich als *erfolgreich*, *gebildet*, zustimmend, *gerecht*, stabil, *essentiell*, pragmatisch, *patriarchalisch* und *vertraglich* wahrgenommen. Singapur teilt die Attribute Erfolg mit dem Element Wohlstand und gebildet, gerecht und zustimmend mit dem Element Sicherheit.

Diese qualitativen Interpretationen der quantitativen Ergebnisse der Clusteranalyse unterstützen die Interpretation einer minimalen Bedeutung von Demokratie, die Bürgerrechte und Gleichberechtigung einschließen kann, aber nicht unbedingt muss. Singapur scheint zur Gruppe der Sicherheit und des Wohlstandes zu gehören.

4.3 Repertory-Grid-Analyse durch Konstrukt-Lademuster auf Hauptkomponenten

Die zweite der beiden grundlegenden Informationen, die ein multiples Repertory-Grid über die Gedanken und Gefühle einer Gruppe von Personen zu einem bestimmten Thema liefert, sind die gemeinsamen Muster hinter den einzelnen Konstruktstrukturen. Aus diesem Grund kann die Datenstruktur des multiplen Repertory-Grid nach Elementen als Zeilen und Konstrukten als Spalten geordnet werden. In einem ersten Schritt wird diese Datenstruktur mittels Hauptkomponentenanalyse analysiert, was zu der in Tabelle 3 dargestellten Faktorenstruktur führt (reduziert auf die fünf Konstrukte mit den höchsten Ladungen auf jeder Komponente).

Pol konstruieren	Ladung auf PC 1	Ladung auf PC 2	Ladung auf PC 3	Gegenpol
Autorität	.63			Freiheit
autoritär	.57			gewählt
autoritativ	.45			Fairness
effizient	-.40			Behörde
Offenheit	-.42			Unanfechtbarkeit
Gleichmacherei		.95		Hierarchie
idealistisch		.79		physisch
gewählt		.71		autoritär
nicht-politisch		.69		politisch
umverteilend		.68		wohlhabend
exklusiv			.97	gemeinsam
sicher			.92	gut dran
wohlhabend			.85	Umverteilung
Einkommens-ungleichheit			.85	Fairness
gebildet			.77	unmoralisch

Tabelle 3: Konstruktladungen und Hauptkomponenten.
Quelle: Osterberg-Kaufmann und Stadelmaier (2020), Konstrukte mit Ladungen kleiner .40 werden nicht gezeigt.

Diese Hauptkomponenten können im Großen und Ganzen als die Struktur des politischen Systemmodells von Easton (1965) interpretiert werden, das aus den Inputs durch Forderungen und Unterstützung durch die Bevölkerung, dem politischen System mit seinen jeweiligen Institutionen und Prozessen und den Outputs in Form von Entscheidungen und Handlungen besteht. Die Hauptkomponente PC 1 spiegelt die Input-Dimension der Legitimität wider, die Aspekte der Freiheit und Ungebundenheit unterstreicht. PC 2 entspricht der Throughput- oder Regierungsdimension mit ihren Institutionen und Verfahren. Und PC 3 ist der Output, der sich auf den sozialen Nutzen und andere Ergebnisse konzentriert.

Dennoch war es nicht das Ziel dieser ersten Studie, die endgültigen Bedeutungsdimensionen von Demokratie in Singapur zu bestimmen. Der vorliegende Datensatz ist auf die singapurische Mittelschicht beschränkt und kann aufgrund seiner Kleinheit nur erste Anhaltspunkte liefern. Diese ersten Ideen müssen in einer größeren Studie getestet werden. Die identifizierten Hauptkomponenten des Repertory-Grid werden daher die Grundlage für die Begriffspaare eines späteren semantischen Differenzials sein, das Studien mit großen Stichproben ermöglicht.

5. Semantisches Differenzial

Die semantische Differenzialmethode geht auf Charles Osgood zurück (Osgood 1952). Die Methode wurde in der Psychologie ent-

wickelt, um die Assoziationen von Menschen mit bestimmten Begriffen, Fakten oder Plänen zu ermitteln.

Es werden keine direkten Fragen wie „Was denken Sie über …?“ verwendet. Stattdessen werden die Personen indirekt befragt, indem ihnen die Möglichkeit gegeben wird anzugeben, wie stark sie z.B. den Begriff „Demokratie“ (oder einen anderen Gegenstand) mit bestimmten Merkmalen assoziieren. Sie erhalten eine Reihe von Merkmalspaaren wie „gewählt – ernannt“ oder „offen – geschlossen“ und können angeben, ob sie „Demokratie“ mit „gewählt“ oder „ernannt“ assoziieren und wie stark sie dies tun (Hofstätter 1986). Ähnlich wie das Repertory-Grid wertet die semantische Differenzialanalyse Bewertungsmuster aus, die auf gegensätzlichen Begriffspaaren basieren. Die im semantischen Differenzial häufig verwendeten Begriffspaare stammen in der Regel aus der akademischen Expertise des Forschers oder aus vordefinierten Sets von semantischen Listen (Osgood et al.1957; Rosenberg und Navarro 2018) und sind somit top-down generiert.

Im vorliegenden Fall haben die Autor*innen die identifizierten Konstrukte, die sich in der Repertory-Grid-Analyse als die stärksten erwiesen haben, als konzeptuelle Paare verwendet. Starke Konstrukte werden dadurch identifiziert, dass sie die höchsten Ladungen auf den Hauptkomponenten des multiplen Repertory-Grid aufweisen. Die konzeptuellen Paare des skizzierten Forschungsdesigns sind dabei bottom-up.

Die Autor*innen entschieden sich aus ökonomischer Sicht für die Auswahl von jeweils 10 Konstrukten pro Hauptkomponente, was zu 30 semantischen Differenzialitems führt. Durch die Kombination von statistischen Informationen mit aktuellen Theorien der Politikforschung wurden 30 Items ausgewählt. Tabelle 4 zeigt ausgewählte Beispiele der Konstrukte, die aus Gründen der sprachlichen Klarheit leicht umformuliert wurden. Die Teilnehmer*innen der semantischen Differenzial-Befragung werden gebeten, ihre Vorstellungen von Demokratie (die als Elemente in der Repertory-Grid-Analyse verwendet wurden) zwischen den beiden Polen der Konstrukte zu bewerten.

Konstrukt-Pol	-2	-1	0	1	2	Konstrukt-Gegenpol
dynamisch						geordnet
ansprechbar						aufzwingend
glücklich						unterdrückt
egalitär						hierarchisch
idealistisch						pragmatisch
Theokratie						säkular
exklusiv						gemeinsam
sicher						wohlhabend
unsicher						patriarchalisch

Tabelle 4: Begriffspaare eines möglichen semantischen Differenzials.
Quelle: Osterberg-Kaufmann und Stadelmaier (2020), Begriffspaare für ein semantisches Differenzial zur Bedeutung von Demokratie in Singapur als Ergebnis der multiplen Repertory-Grid-Analysen.

Durch dieses Studiendesign können Daten aus großen Stichproben gewonnen werden, die dennoch die relevantesten kollektiven Wahrnehmungen von Demokratie fokussieren. Daher ist die Integration von Repertory-Grid und semantischen Differenzialmethoden in der Lage, eine effiziente Kombination von qualitativ tiefen Einsichten zu liefern, die durch übliche groß angelegte Datenanalysemethoden, die auf große Stichproben angewandt werden.

6. Fazit

Ausgangspunkt für diesen Artikel waren die Zweifel an der These der überwältigenden Unterstützung für die Demokratie weltweit, der ein großer Teil der politikwissenschaftlichen Literatur über einen langen Zeitraum gefolgt ist. Stattdessen folgen die Autor*innen dieses Artikels dem Argument, dass diese weltweite überwältigende Unterstützung für die Demokratie künstlich ist, erstens weil Demokratie kein universelles homogenes Konzept ist und insbesondere in nicht-westlichen Ländern anders verstanden wird. Zweitens besteht daher ein Bedarf an alternativen methodischen Ansätzen, die Möglichkeiten bieten, die Unterstützung und vor allem die Bedeutung von Demokratie jenseits der Frage nach der Unterstützung normativer (westlicher) Konnotationen von Demokratie (top-down) zu messen, indem deduktive Ansätze durch induktive Ansätze ergänzt werden. In diesem Sinne schlägt der Artikel einen Methodenmix und einen Mehrebenenansatz aus Repertory-Grid und semantischem Differenzial als methodische Alternativen vor. Die Kombination dieser beiden Methoden ermöglicht es, die semantische Tiefe des Demokratiebegriffs zu erfassen und umfangreiche Daten über den jeweiligen kulturellen Kontext zu sammeln. Zusammengenommen reduziert dieser kombinierte Ansatz die Einschränkungen durch kulturelle und sprachliche Äquivalenz sowie die Phänomene des Lippenbekenntnisses und der sozialen Erwünschtheit und ermöglicht gleichzeitig Repräsentativität und Generalisierung.

Um die methodischen Möglichkeiten der beiden hier vorgestellten Differenzierungsmethoden zu veranschaulichen, werden in diesem Artikel einige Ergebnisse von Repertory-Grid-Interviews vorgestellt, die in Singapur durchgeführt wurden, um die Bedeutung von Demokratie für die singapurische Mittelschicht zu ermitteln. Die wichtigsten Ergebnisse zu diesem Punkt sind zunächst eine Bestätigung der Debatten über Demokratie jenseits des Westens (Schubert 2016; 2012), da auch in der Mittelschicht Singapurs eine Bedeutung von Demokratie zu dominieren scheint, die Kernelemente der westlichen liberalen Bedeutung von Demokratie teilt, sich aber auf eine Minimaldefinition beschränkt, die nur Wahlen einschließt. Freiheits- und Bürgerrechte scheinen nur im weiteren Sinne Teil des Demokratiebegriffs zu sein. Zweitens kann festgestellt werden, dass es keine atypische Bedeutung von Demokratie in der singapurischen Mittelschicht gibt, da Militär und religiöse Führer definitiv nicht als Teil des Demokratiekonzepts angesehen werden. Auch sozialdemokratische Aspekte entsprechen nicht den Demokratievorstellungen der singapurischen Mittelschicht.

Allerdings ist, wie im Artikel diskutiert, die Interpretation der Daten jenseits der beiden Extreme von Demokratie und antidemokratischen Elementen nicht explizit möglich, da die Befragten der Demokratie generell eher indifferent gegenüberstehen, wie

bereits in früheren Forschungen diskutiert (Schubert 2012). Mit Hilfe der räumlichen Anordnung der Elemente konnte eine Bedeutung von Demokratie abgebildet werden, die Ähnlichkeiten und Unterschiede zur westlichen, liberalen Bedeutung aufweist.

Die eigentliche Bedeutung, die hinter den Gemeinsamkeiten und Unterschieden liegt, erschließt sich jedoch erst durch die qualitativen Daten des Interviews. Diese Daten konnten gewonnen werden, weil das Repertory-Grid mit den eigenen Worten des Befragten arbeitet und die Befragten ihre eigene Bewertung innerhalb ihrer eigenen Bewertungsskalen vornehmen. Diese subjektive Bewertung eliminiert die Effekte der sprachlichen und kulturellen Äquivalenz innerhalb des Interviews und ermöglicht es sogar, die unterschiedlichen kulturellen und sprachlichen Bedeutungen der Wörter und Bewertungssysteme während der Analyse zu verstehen.

Innerhalb der Methodendiskussion um alternative methodische Ansätze zur Untersuchung der weltweiten Bedeutung von Demokratie stellt das Repertory-Grid einen induktiven Bottom-up-Ansatz dar. Um diesen induktiven Charakter der Repertory-Grid-Methode in zukünftigen Forschungsprojekten zu stärken, könnte die Auswahl der Elemente zu Beginn des Interviewprozesses vielleicht auch bottom-up erfolgen, anstatt die Elemente top-down aus der bestehenden Forschung abzuleiten. Eine Möglichkeit, die in diese Richtung weist, wären z.B. Fokusgruppeninterviews, die auf diese Weise gemeinsam mit den Interviewpartnern Elemente der Demokratie entwickeln (Jankowicz 2004). Eine Herausforderung bleibt die Vergleichbarkeit der Interviews zwischen verschiedenen Interviewgruppen und Ländern.

Um Repräsentativität und damit verallgemeinerbare Ergebnisse zu erreichen, plädieren die Autor*innen für einen mehrstufigen Ansatz. Durch die Analyse der Hauptkomponenten der Datenstruktur des Repertory-Grid können die Ergebnisse auf die Methode des semantischen Differenzials übertragen werden, die in einem nächsten Schritt die Durchführung einer großen Anzahl von Interviews ermöglicht, da es sich um einen rein quantitativen Ansatz handelt. Der Vorteil dieses Mehrebenen-Designs ist, dass die Konstruktpole des semantischen Differenzials auf den eigenen Konstrukten des Repertory-Grid-Interviews basieren und somit bottom-up statt top-down definiert werden. In diesem nächsten Schritt wird es möglich sein, erstens zu analysieren, wie die Singapurer*innen Demokratie sehen sowie zweitens die Vorstellungen von Demokratie zwischen verschiedenen Gruppen in der singapurischen Gesellschaft zu vergleichen.

Literatur:

Almond G. A. und S. Verba. 1963. The Civic Culture. Political Attitudes and Democracy in Five Nations. Princeton: Princeton University Press.

Arzheimer K. 2002. Politikverdrossenheit. Bedeutung, Verwendung und empirische Relevanz eines politikwissenschaftlichen Begriffes. Wiesbaden: Westdeutscher Verlag.

Bertin, J. 1973. Sémiologie graphique. Flammarion, Paris. http://www.jstor.org/stable/pdf/25131373.pdf

Braizat F. 2010. Was Araber denken. In Journal of Democracy 21 (4): 131-138.

Bratton M. 2010. Die Verankerung des „D-Wortes" in Afrika. In Journal of Democracy 21 (4): 106-113.

Chen, A. 2002. Kapitalistische Entwicklung, unternehmerische Klasse und Demokratisierung in China. In: Political Science Quarterly 117(3): 401-422.

Chen, J. 2013. A Middle Class without Democracy: Economic Growth and the Prospects for Democratization in China. New York: Oxford University Press.

Cho Y. 2015. How Well are Global Citizenries Informed about Democracy? Zur Erfassung der Breite und Verteilung ihres demokratischen Wissens und seiner Quellen. In: Political Studies 63: 240-258.

Chu Y. und M. Huang 2010 Solving an Asian Puzzle. In Journal of Democracy 21 (4): 114-130.

Cohen, J. 1992. A power primer. Psychological Bulletin 112 (1): 155-159.

congrid ®. 2019. permitto GmbH, Wiesbaden, Deutschland.

Dahl R. 1989. Democracy and its Critics. New Haven: Yale University Press.

Dallmayr F. (Eds.). 2010 Comparative Political Theory: An introduction. New York: Palgrave Macmillan.

Dalton R. J., D. C. Shin und W. Jou. 2008. How People Understand Democracy. In How People View Democracy, Eds. L. Diamond and M. F. Plattner, 1-15. Baltimore: The Johns Hopkins University Press.

Diamond L. 2008. The Spirit of Democracy. The Struggle to Build Free Societies Throughough the World. Ney York: Times Books.

Diamond L. 2010. Introduction. In Journal of Democracy 21 (4): 102-105.

Easton D. 1975. A Re-assesment of the Concept of Political Support. In: British Journal of Political Science 5 (4): 435-457.

Emmerson D. K. 1995. Singapur und die „Asian Values"-Debatte. In: Journal of Democracy 6 (4): 95-105.

Ferrín M. und H. Kriesi. 2016. How Europeans View and Evaluate Democracy. Oxford: Oxford University Press.

Fuchs D. 1987. Trends politischer Unterstützung der Bundesrepublik. In. Politische Kultur in Deutschland. Bilanz und Perspektiven der Forschung, Eds. D. Berg-Schlosser und J. Schissler, 357-377. Opladen: Westdeutscher Verlag.

Fuchs D. 1989. Die Unterstützung des politischen Systems der Bundesrepublik Deutschland. Opladen: Westdeutscher Verlag.

Glassman, R. 1991. China im Übergang: Communism, capitalism, and democracy. New York: Prager.

Goodman, D. G. 1999. The new middle class. In The paradox of China's post Mao reforms. (Eds.) Merle Goldman and Roderick MacFarquhar, 241-261. Cambridge: Harvard University Press.

He, L. 2003. Middle Class: Friends or Foes to Beijing's New Leadership. In Journal of Chinese Political Science 8(1&2): 88-100.

Hofstätter, P. R. 1986. Gruppendynamik: Kritik der Massenpsychologie. Reinbek bei Hamburg: Rowohlt Taschenbuch.

Inglehart R. 2003. How Solid is Mass Support for Democracy - and How Can We Measure It? Politikwissenschaft und Politik 36 (1): 51-57.

Jankowicz, D. 2004. The easy guide to repertory grids. Chichester, West Sussex, England; Hoboken, N.J: Wiley.

Jarzabkowski, P., W. K. Smith, A. Langley und M.W. Lewis (Eds.). 2017. The Oxford handbook of organizational paradox. Oxford: Oxford University Press.

Kaulartz M. und M. Heckmann. 2014. Comparing interactive online and face-to-face repertory grid interviews in terms of construct quality and practicability. Vortrag gehalten auf der 12th Biennial Conference of the European Personal Construct Association (EPCA), Brno, Tschechische Republik.

Kausikan B. 1998. The ‚Asian Values' Debate: A View from Singapore. In. L. Diamond und M. F. Plattner. Eds. Democracy in East Asia. 17-27. Baltimore & London: Johns Hopkins University Press.

Kelly, G. A. 1980. Theory of Personality: the psychology of personal constructs (1. Aufl. im Norton Taschenbuch 1963). New York: Norton.

King G. und J. Wand 2007. Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes. Political Analysis 15: 46-66.

King G., C. J. L. Murray, J.A. Salomon und A. Tandon 2004. Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. American Political Science Review 98 (1): 191-207.

Klingemann H. D. 1999. Mapping Political Support in the 1990s: A Global Analysis, In. Eds. P. Norris. Critical Citizens: Global Support for Democratic Government. 31-56. Oxford: Oxford University Press.

Koo, H. 1991. Middle Classes, Democratization, and Class Formation: The Case of South Korea. Theorie und Gesellschaft 20(4): 485-509.

Lauth H. J., G. Pickel und S. Pickel. 2009. Methoden der vergleichenden Politikwissenschaft. Wiesbaden: VS Verlag für Sozialwissenschaften.

Lijphart A. 1971. Vergleichende Politik und die vergleichende Methode. The American Political Science Review 65 (3): 682-693.

Lipset 1959. Some Social Requisites of Democracy: Wirtschaftliche Entwicklung und politische Legitimität. The American Political Science Review 53 (1): 69-105.

Lu J. und T. Shi 2015. Der Kampf der Ideen und Diskurse vor dem demokratischen Übergang: Unterschiedliche Demokratiekonzeptionen im autoritären China. In: International Political Science Review 36 (1): 20-41.

Maehler D. und U. Schmidt-Denter 2013. Migrationsforschung in Deutschland. Leitfaden und Messinstrumente zur Erfassung psychologischer Konstrukte. Wiesbaden: Springer VS.

Mahoney J. 2007. Qualitative Methodologie und vergleichende Politik. Comparative Political Studies 40 (2): 122-144.

Miron-Spektor, E., A. Ingram, J. Keller, W.K. Smith und M.W. Lewis. 2018. Microfoundations of Organizational Paradox: The Problem Is How We Think about the Problem. Academy of Management Journal 61(1): 26-45.

Moore B. 1966. Social Origins of Dictatorship and Democracy. Lord and Peasant in the Making of the Modern World. Boston: Beacon Press.

Moosbrugger, H., & Kelava, A. (Eds.). 2008. Testtheorie und Fragebogenkonstruktion: mit 43 Tabellen. Berlin: Springer.

Norris P. 1999. The Growth of Critical Citizens?, In P. Norris. Eds. Critical Citizens: Global Support for Democratic Government. 1-29. Oxford: Oxford University Press.

O'Brien, K. J. und R. E. Stern. 2008. Introduction: Studying Contention in Contemporary China. In. Popular Protest in China. Eds. K. J. O'Brien. 11-25. Cambridge: Harvard University Press.

Ortmann, S. 2011. Singapur: Authoritarian but Newly Competitive. In Journal of Democracy 22(4): 153-164.

Osgood, C. E. 1952. The nature and measurement of meaning. Psychological Bulletin, 49 (3): 197-237.

Osgood, C. E. 2009. Semantic Differential Technique in the Comparative Study of Cultures. American Anthropologist 66 (3): 171-200.

Osgood, C. E., G. J. Suci und P. H. Tannenbaum. 1957. The Measurment of Meaning. Urbana und Chicago: University of Illinois Press.

Pickel G. 2006. Die Einschätzung der Demokratie im internationalen Vergleich. In G. Pickel und S. Pickel Eds. Demokratisierung im internationalen Vergleich. Neue Erkenntnisse und Perspektiven. 267-300. Wiesbaden: VS Verlag für Sozialwissenschaften.

Pickel G. 2009. Der Einbezug des Individuums in die Länderanalyse - Umfrageforschung und vergleichende Politikwissenschaft. In S. Pickel, G. Pickel und H.-J. Lauth. Eds. Methoden der vergleichenden Politikwissenschaft. Neue Entwicklungen und Anwendungen. 297-315. Wiesbaden: VS Verlag für Sozialwissenschaften.

Pickel G. und D. Walz 1997. Politikverdrossenheit in Ost- und Westdeutschland. Dimensionen und Ausprägungen. In: Politische Vierteljahresschrift 38: 27-49.

Pickel G. und S. Pickel 2009. Qualitative Interviews als Verfahren des Ländervergleichs. In S. Pickel, G. Pickel und H.-J. Lauth. Eds. Methoden der vergleichenden Politikwissenschaft. Neue Entwicklungen und Anwendungen. 441-464. Wiesbaden: VS Verlag für Sozialwissenschaften.

R Core Team. 2018. Eine Sprache und Umgebung für statistisches Rechnen. R Foundation for Statistical Computing, Wien, Österreich. URL https://www.R-project.org/.

Rosenberg, B. und M. Navarro, M. 2018. Semantic Differential Scaling. In Frey, B.B. (Eds.), The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. 1504–1507. Thousand Oaks: SAGE Publications, Inc.

Rost, J. 2004. Lehrbuch Testtheorie-Testkonstruktion. Bern: Huber.

Sartori G. 1997. Demokratietheorie. Darmstadt: Wissenschaftliche Buchgesellschaft.

Schubert S. 2012. Die globale Konfiguration politischer Kulturen. Eine theoretische und empirische Analyse. Wiesbaden: VS Verlag für Sozialwissenschaften.

Schubert S. 2016. Inwiefern universell? Zum Demokratiebegriff in der vergleichenden Demokratieforschung. In S. de la Rosa, S. Schubert und H. Zapf Eds. Transkulturelle Politische Theorie. Eine Einführung. 285- 303. Wiesbaden: Springer Fachmedien

Schubert S. und A. Weiß. 2016. Demokratiekonzeptionen jenseits des Westens: Ein theoretischer Rahmen für ein neues Forschungsfeld. In S. Schubert und A. Weiß Eds. „Demokratie" jenseits des Westens. Theorien, Diskurse, Einstellungen. 6-34. Baden-Baden: Nomos.

Sen A. K. 1999. Democracy as a Universal Value. In Journal of Democracy 10 (3): 3-17.

Shcheglova, M. 2010. Eine integrierte Methode zur Bewertung der Verbrauchermotivation in schwierigen Marktnischen: Ein Fallbeispiel für das Premium-Autosegment in Russland. Abgerufen von https://depositonce.tu-berlin.de/handle/11303/2685

Shi T. und J. Lu 2010. Der Schatten des Konfuzianismus. In: Journal of Democracy 21 (4): 123-130.

Stegmüller D. 2011. Äpfel und Orangen? Das Problem der Äquivalenz in der vergleichenden Forschung. Politische Analyse 19: 471-487.

van de Vijver F. J. R. 2003. Bias und Äquivalenz: Cross-Cultural Perspectives. In J. A. Harkness, F. J. R. van de Vijver und P. Ph. Mohler Eds. Cross-Cultural Survey Methods. 143-155. Hoboken: John Wiley & Sons.

van Deth J. W. 2013. Äquivalenz in der vergleichenden Politikforschung. In Comparative Politics: The Problem of Equivalence, Eds. J. W. van Deth. 1-19. Colchester: ECPR Press.

Waldman, D. A., L.L. Putnam, E. Miron-Spektor und D. Siegel, D. 2019. Die Rolle der Paradoxie-Theorie in der Entscheidungsfindung und Managementforschung. Organizational Behavior and Human Decision Processes 155: 1-6.

Walker, B. M., & Winter, D. A. 2007. Die Ausarbeitung der Personal Construct Psychology. Annual Review of Psychology 58 (1): 453-477. https://doi.org/10.1146/annurev.psych.58.110405.085535

Wand J. 2012. Credible Comparisons Using Interpersonally Incomparable Data: Non-parametric scales with anchoring vignettes. American Journal of Politics, Early View (Online-Version eines Datensatzes, der vor der Aufnahme in eine Ausgabe veröffentlicht wurde) 25 Jun 2012.

Welzel C. 2014. Freedom Rising. Human Empowerment and the Quest for Emancipation. Cambridge: Cambridge University Press.

Welzel C. und H. Kirsch. 2017. Democracy Misunderstood: Authoritarian Notion of Democracy around the Globe. In World Values Research 9(1): 1-29.

Welzel C. und R. Inglehart 2008. The Role of Ordinary People in Democratization. In How People View Democracy. In L. Diamond and M. F. Plattner. Eds. 16-30. Baltimore: The Johns Hopkins University Press.

Weßels B. 2009. Spielarten des Euroskeptizismus. In F. Decker und M. Höreth. Eds. Die Verfassung Europas. 50-68. Wiesbaden: VS Verlag für Sozialwissenschaften.

Westle B. 1989. Politische Legitimität - Theorien, Konzepte, empirische Befunde. Baden-Baden: Nomos.

Wu, W.-C., Y.-T. Chang und H.-H. Pan 2017: Does China's middle class prefer (liberal) democracy? In Democratization 24(2): 347-366.

Xiao, G. 2003. The rise of the technocrats. In: Journal of Democracy 14: 59-65.

Yldiz T. 2012. Übersetzungspraktiken und die widersprüchliche Logik des Politischen. Eine praxistheoretische Grundlegung der Transformationsforschung. In H. Zapf Eds. Nichtwestliches politisches Denken. 213-228. Wiesbaden: Springer Fachmedien.

Zhang, Ch. 2010. Inflation Uncertainty and Monetary Policy in China. In China und Weltwirtschaft 18(3): 40-55.

Kontakt:

Dr. Ulrich Stadelmaier
Diplom-Psychologe
Leiter der dgp-Geschäftsstelle Leipzig
Deutsche Gesellschaft für Personalwesen e.V.
stadelmaier@dgp.de

Dr. Norma Osterberg-Kaufmann
M.A. Politikwissenschaften
Institut für Sozialwissenschaften der Humboldt-Universität zu Berlin
norma.osterberg-kaufmann@hu-berlin.de

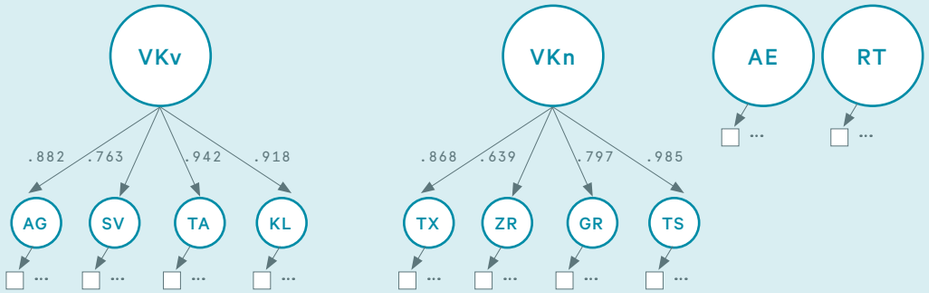


Abbildung 1. Faktorenmodell der Testbatterie OV1

Abbildung 1 zeigt die angenommene Faktorstruktur des OV1.

Die Korrelationen zwischen den latenten Variablen VKv (verbale Verarbeitungskapazität), VKn (numerische Verarbeitungskapazität), AE (Arbeitseffizienz) und RT (Rechtschreibung) wurden im Modell ebenfalls geschätzt. Es zeigen sich durchgehend hohe Korrelationen zwischen den latenten Variablen (siehe Tabelle 1). Wie theoretisch erwartet, ist die Korrelation zwischen verbaler und numerischer Verarbeitungskapazität am höchsten (r = .827), während die Korrelationen von Verarbeitungskapazität zu Arbeitseffizienz und Rechtschreibung geringer ausfallen (r =.530 bis r = .762). Hohe Korrelationen von über r =.50 zwischen verschiedenen kognitiven Fähigkeiten sind auf Grundlage des sogenannten „positive manifold“ zu erwarten und sprechen für die konvergente Validität des Verfahrens (Brody, 2005).

Weitere Gütekriterien

Ökonomie

Mit einer Gesamtzeit von 120 bis 150 Minuten ist der OV1 eine zeitökonomische Testbatterie, um die numerische und verbale Verarbeitungskapazität, die Arbeitseffizienz und die Rechtschreibung zu erfassen. Durch die Online-Durchführung am eigenen Endgerät entsteht kein zusätzlicher Materialbedarf und keine Kosten für Anreise, Testaufsicht o.ä.

Fairness

Die Geschlechterfairness der Testbatterie wurde auf Itemebene (Differenzial Item Functioning) mit Hilfe der logistischen Regression überprüft (Swaminathan and Rogers, 1990). Bei der Methode werden Männer und Frauen in Eignungskategorien eingeteilt. Für jede Eignungskategorie wird für jedes Item die Lösungswahrscheinlichkeiten der Items für Männer und Frauen betrachtet. Bei gleicher Eignung sollten Männer und Frauen sich nicht in ihrer Wahrscheinlichkeit, das Item korrekt zu lösen, unterscheiden. Die Methode eignet sich, um sowohl uniformes als auch non-uniformes Differenzial Item Functioning zu überprüfen.

Für alle Items des OV1 zeigen sich nach Jodoin und Gierls (2001) unerhebliche Effektgrößen in der logistischen Regression. Das bedeutet, dass für kein Item eine Benachteiligung auf Grund der Zugehörigkeit zu einer der Geschlechtergruppen besteht.

	VKv	VKn	AE	RT
VKv	1			
VKn	.827	1		
AE	.664	.762	1	
RT	.737	.690	.530	1

Tabelle 1. Korrelationen der Faktoren des OV1

Testsicherheit

Der Zugang zum OV1 erfolgt über eine ETAN, die eine einmalige Testteilnahme in einem festgelegten Zeitraum erlaubt. Die Aufgabengruppen sind auch zeitlich begrenzt, um Verfälschungen zu minimieren. Personenkontrollen und Überwachungen der Testsituation sind mittels zusätzlicher Software durch die einstellende Behörde möglich. Im Vergleich zu Vor-Ort-Tests ist die Testsicherheit herabgesetzt.

Verfälschungen nach unten sind wie bei allen Leistungstests möglich, aber in Auswahlsituationen nicht zu erwarten. Um die Verfälschbarkeit nach oben durch Trainingseffekte zu vermeiden, sind Personen nach Bearbeitung der Testbatterie für sechs Monate von der erneuten Teilnahme am Test ausgeschlossen.

Fazit

Die Testbatterie OV1 erlaubt die reliable, faire und ökonomische Erfassung kognitiver Fähigkeiten im Rahmen der Personalauswahl für den mittleren und gehobenen öffentlichen Dienst. Instruktion, Auswertung und Interpretation sind durch eine hohe Standardisierung objektiv. Die Normstichprobe der Testbatterie entspricht der Einsatzzielgruppe und ist aktuell. Die Reliabilität der Testbatterie ist sowohl auf Skalenebene als auch auf Testebene sehr hoch, sodass Gesamtwert und Skalenwerte interpretiert werden können. Evidenz für faktorielle Validität des Tests liegt vor und spricht für ein gutes Auswahlinstrument.

Die Qualität der Testbatterie wird in zukünftigen Untersuchungen weiter gesichert: Es erfolgt in Kürze eine Überprüfung der Normen an einer größeren Stichprobe. Ebenfalls werden die Fairnessanalysen mit mehr Daten wiederholt. Eine Bewährungskontrolle wird durchgeführt, sobald hierfür Daten vorliegen.

Literatur:

Brody, N. (2005). To g or Not to g--That Is the Question. In O. Wilhelm & R. W. Engle (Eds.), Handbook of understanding and measuring intelligence (pp. 489–502). Sage Publications, Inc. https://doi.org/10.4135/9781452233529.n26

Hülsheger, U. R., Maier, G. W., & Stumpp, T. (2007). Validity of general mental ability for the prediction of job performance and training success in Germany: A meta-analysis. International Journal of Selection and Assessment, 15(1), 3-18. https://doi.org/10.1111/j.1468-2389.2007.00363.x

Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). BIS-TEST Berliner Intelligenzstruktur-Test – Form 4. Göttingen: Hogrefe.

Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I error and power rates using an effect sizemeasure with logistic regression procedure for DIF detection. Applied Measurement in Education, 14, 329-349. doi:10.1207/S15324818AME1404_2

Kersting, M. (2009). Die DGP Testverfahren - Ein kurzer Rückblick und eine aktuelle Studie zur Konstrukt und Kriteriumsvalidität des BIS-r-DGP Tests. dgp Informationen, S. 22-37.

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. Journal of Educational Measurement, 27(4), 361–370. https://doi.org/10.1111/j.1745-3984.1990.tb00754.x

Kontakt:

Dr. Anna-Lena Jobmann
Diplom-Psychologin
Deutsche Gesellschaft für Personalwesen e.V.
jobmann@dgp.de

Amelie Kleinmanns
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
kleinmanns@dgp.de

Vorstellung der dgp Testbatterie OG1

Amelie Kleinmanns & Anna-Lena Jobmann

Allgemeine Informationen

Der OG1 ist eine Online-Testbatterie zur Beurteilung der kognitiven Eignung für gewerbliche, medizinische und soziale Ausbildungsberufe sowie den mittleren feuerwehrtechnischen Dienst, die seit September 2021 eingesetzt wird. Basierend auf einem Intelligenzstrukturmodell erfasst der Test verbale, numerische und figurale Verarbeitungskapazität mit insgesamt acht Untertests sowie figurale und numerische Bearbeitungsgeschwindigkeit mit vier Untertests. Des Weiteren werden das Praktische Verständnis und die Deutschkenntnisse (Grammatik & Rechtschreibung) überprüft. Für die Testbatterie können sowohl ein Gesamtwert als auch Skalenwerte für numerische, verbale und figurale Verarbeitungskapazität sowie numerische und figurale Bearbeitungsgeschwindigkeit, Praktisches Verständnis, Grammatik und Rechtschreibung ausgegeben werden. Eine Interpretation des Gesamtwertes und der Skalenwerte für numerische, verbale und figurale Verarbeitungskapazität ist möglich. Der OG1 wird online über die dgp-Testplattform durchgeführt und dauert ca. 120 bis 150 Minuten. Die einzelnen Aufgabengruppen sind zeitlich begrenzt.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion

Der OG1 basiert auf dem Berliner-Intelligenz-Struktur Modell (BIS-Modell) und erfasst die operativen Fähigkeiten Verarbeitungskapazität und Bearbeitungsgeschwindigkeit in verbalen (sprachgebundenen), numerischen (zahlengebundenen) und figuralen (figural-bildhaften) Inhaltsbereichen (Jäger, Süß, & Beauducel, 1997). Die Verarbeitungskapazität erfasst die Verarbeitung komplexer Informationen zur Lösung von Aufgaben unter Zuhilfenahme von formallogischem Denken und korrekter Beurteilung von Informationen bezogen auf Sprache, Zahlen und bildhaftes Material. Die Aufgaben zur Bearbeitungsgeschwindigkeit erfassen das Arbeitstempo und die Konzentrationskraft beim Lösen von einfach strukturierten Aufgaben mit geringer Schwierigkeit. Die beiden im OG1 abgebildeten Inhaltsbereiche stellen für die Eignungsbeurteilung in den Zielberufen die wichtigsten kognitiven Fähigkeiten dar.

Um die Anforderungen für die Zielberufe noch passender abzufragen, werden in der Testbatterie zwei weitere Kenntnisbereiche abgefragt. Die Kenntnisse praktisch-technischer Gesetzmäßigkeiten und Zusammenhänge in verschiedenen anwendungsnahen Situationen werden im Untertest Praktisches Verständnis erfasst. Zudem werden die Deutschkenntnisse in Bezug auf Rechtschreibung und Grammatik erfasst.

Objektivität

In Bezug auf die Durchführungsobjektivität wird durch eine standardisierte Online-Testinstruktion sowie Tipps zur Gestaltung der Testsituation am eigenen Endgerät (z.B. ungestörte Umgebung, sichere Internetverbindung) eine möglichst standardisierte Situation hergestellt.

Die Auswertung erfolgt automatisiert und computergestützt mit entsprechend hoher Auswertungsobjektivität. Als Interpretationshilfe werden Bewerber*innen Empfehlungsgrade zugeordnet. Diese Empfehlungsgrade mit den Werten 1-5 umfassen einen a priori festgelegten Bereich an Standardwerten. Diese werden mit Hilfe einer Tabelle auf den Profilkarten (der Ergebnisdarstellung) visualisiert. Eine Eingruppierung in die Gruppen: Leistung entspricht „…den Anforderungen“, „…weitgehend den Anforderungen“, „nur teilweise den Anforderungen“, „weitgehend nicht den Anforderungen“ und „nicht den Anforderungen“ bietet eine weitere Hilfe bei der Interpretation der Testergebnisse.

Normierung

Für den OG1 liegt eine Gesamtnorm in Form von Standardwerten mit einem Mittelwert von M = 100 und einer Standardabweichung von SD = 10 vor. Die Normstichprobe umfasst 730 Bewerber*innen für gewerbliche, medizinische und soziale Ausbildungsberufe im öffentlichen Dienst in Deutschland sowie den mittleren feuerwehrtechnischen Dienst. 75,6% der Personen sind männlich, 23,8% weiblich und 0,5% divers. Die Daten zur Normierung wurden im Jahr 2021 erhoben.

Zuverlässigkeit

Die Berechnung der internen Konsistenzen mithilfe von Cronbach’s Alpha ergeben für Verarbeitungskapazität numerisch α = .84, verbal α =.80 und figural α =.91. Für die Skalen Praktisches Verständnis, Deutsch

Rechtschreibung und Deutsch Grammatik liegen die Cronbach's Alpha-Werte zwischen α = .59 und α = .70. Die interne Konsistenz der gesamten Testbatterie liegt bei α =.92. Die Split-Half-Reliabilitäten liegen für die Skalen numerisch rr = .84, verbal rr = .78 und figural rr = .93. Die gesamte Batterie hat eine Split-Half-Reliabilität von rr = .92.

Gültigkeit

Konstruktvalidität

Aufgrund der rationalen Konstruktion der Testbatterie kann die theoretische Struktur im Rahmen von konfirmatorischen Faktorenanalysen überprüft werden. Aus den Subtests wurden dazu je drei bis fünf zufällige Parcel gebildet, die auf den latenten Variablen für die Subtests laden. Diese latenten Variablen der Subtests laden wiederum auf den latenten Variablen numerische, figurale und verbale Verarbeitungskapazität. Die Verarbeitungskapazität wird als latente Variable höherer Ordnung mit den drei latenten Variablen für figurale, numerische und verbale Verarbeitungskapazität als Indikatoren abgebildet.

Die Bearbeitungsgeschwindigkeit wird als eindimensionales Modell mit drei manifesten Indikatoren eingebunden. Rechtschreibung und Grammatik werden als zwei latente Variablen mit zufälligen Parceln als manifeste Indikatoren gebildet. Ihnen übergeordnet ist die latente Variable Deutschfähigkeiten, auf welche beide latenten Variablen laden.

Praktisches Verständnis wird als eindimensionale latente Variable mit manifesten Indikatoren (zufällige Parcel) in dem Modell geschätzt. Zwischen den genannten latenten Variablen werden die Zusammenhänge geschätzt (s. Tabelle 1).

Es zeigt sich eine gute Modellpassung (ML, X2 =1862.92, df =1205, p = .000, CFI = .938, RMSEA = .027, 90% CI [.025; .030]). Die Faktorladungen der latenten Variablen für die Subtests auf den latenten Variablen der numerischen, verbalen und figuralen Verarbeitungskapazität sind durchgehend sehr hoch. Auch die Ladungen von numerischer, verbaler und figuraler Verarbeitungskapazität auf die allgemeine Verarbeitungskapazität sind sehr hoch (siehe Abbildung 1). Die Ladungen von Deutsch Rechtschreibung und Deutsch Grammatik auf den übergeordneten latenten Variablen liegen bei r = .84 und r = .96. Die Korrelationen zwischen den latenten Variablen Verarbeitungskapazität, Bearbeitungsgeschwindigkeit und Deutschfähigkeiten sind mit Werten zwischen r = .58 und r = .87 wie erwartet sehr hoch. Die Korrelationen der Deutschfähigkeiten und Bearbeitungsgeschwindigkeit zu Praktischem Verständnis fallen moderater aus.

Insgesamt zeigt sich die theoretisch erwartete Struktur empirisch bestätigt, sodass die Skalen- und Gesamtwerte des OG1 gut interpretierbar sind.

	PV	D	VK	BG
PV	1			
D	.42	1		
VK	.71	.87	1	
BG	.49	.58	.79	1

Tabelle 1. Korrelationen der Faktoren des OG1

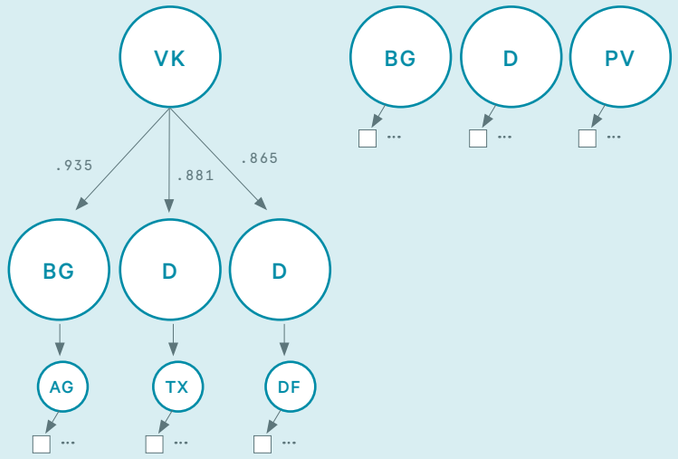


Abbildung 1. Faktorenmodell der Testbatterie OG1

Kriteriumsvalidität

Für die äquivalenten Papierversionen des OG1 wurden bisher mehrere Bewährungskontrollen zur Untersuchung der prognostischen Validität durchgeführt (Rieke, Bendler & Knorz, 2007, Riestel & Haarhaus, 2014, Stadelmaier, Thiem, & Kleinmanns, 2018). Es zeigten sich Zusammenhänge zwischen den dgp-Empfehlungsgraden und Kriterien für Ausbildungserfolg von .28-.43. Diese Korrelationen deuten auf eine gute prädiktive Validität der Papierversion hin. Durch die konzeptionelle und inhaltliche Nähe der Online- und Papierversion kann vom ähnlichen prognostischen Validitäten für den OG1 ausgegangen werden (Kersting, 2009).

Weitere Gütekriterien

Ökonomie

Mit einer Gesamtzeit von 120 bis 150 Minuten ist der OG1 eine zeitökonomische Testbatterie, um die numerische, verbale und figurale Verarbeitungskapazität, die figurale und numerische Bearbeitungsgeschwindigkeit, das Praktische Verständnis und die Deutschkenntnisse zu erfassen. Durch die Online-Durchführung am eigenen Endgerät entstehen kein zusätzlicher Materialbedarf und keine Kosten für Anreise, Testaufsicht o.ä.

Fairness

Die Geschlechterfairness der Testbatterie kann auf Itemebene (Differenzial Item Functioning) im Rahmen der Logistischen Regression überprüft werden (Swaminathan and Rogers, 1990). Bei der Methode werden Männer und Frauen in Eignungskategorien eingeteilt. Für jede Eignungskategorie wird für jedes Item die Lösungswahrscheinlichkeit der Items für Männer und Frauen betrachtet. Bei gleicher Eignung sollten Männer und Frauen sich nicht in ihrer Wahrscheinlichkeit unterscheiden, das Item korrekt zu lösen. Für den OG1 sowie für die äquivalenten Papierversionen des Tests ergeben sich Herausforderungen dadurch, dass deutlich mehr Männer den Test bearbeiten. Diese Ungleichheit in den Stichprobengrößen kann zu erheblichen Problemen der statistischen Methode führen. Für den OG1 liegt zum Zeitpunkt der Analysen noch

keine hinreichend große Stichprobengröße für Frauen vor. Die entsprechenden Analysen der sehr ähnlichen Papierversion G3 zeigen für alle Testteile niedrige Effektstärken (Zumbo & Thomas, 1997, Jodoin & Gierl, 2001). Die Analysen zur Geschlechterfairness werden bei geeigneter Stichprobengröße wiederholt.

Testsicherheit

Der Zugang zum OG1 erfolgt über eine ETAN, die eine einmalige Testteilnahme in einem festgelegten Zeitraum erlaubt. Die Aufgabengruppen sind auch zeitlich begrenzt, um Verfälschungen zu minimieren. Personenkontrollen und Überwachungen der Testsituation sind durch zusätzliche Software durch die einstellende Behörde möglich. Im Vergleich zu Vor-Ort-Tests ist die Testsicherheit herabgesetzt.

Verfälschungen nach unten sind wie bei allen Leistungstests möglich, aber in Auswahlsituationen nicht zu erwarten. Um die Verfälschbarkeit nach oben durch Trainingseffekte zu vermeiden, sind Personen nach Bearbeitung der Testbatterie für sechs Monate von der erneuten Teilnahme am Test ausgeschlossen.

Fazit

Die Testbatterie OG1 erlaubt die faire, ökonomische und unverfälschte Erfassung kognitiver Fähigkeiten und Kenntnisse für gewerbliche, medizinische und soziale Ausbildungsberufe sowie den mittleren feuerwehrtechnischen Dienst. Die Objektivität in Durchführung, Auswertung und Interpretation ist durch entsprechende Maßnahmen optimiert. Die vorliegenden Normen sind aktuell und wurden auf Basis einer hinreichend großen Stichprobe von Bewerber*innen aus den relevanten Berufen berechnet. Die Reliabilität der Testbatterie kann als gut bewertet werden, sodass eine Interpretation des Gesamtergebnisses empfohlen werden kann. Die Evidenz für prädiktive und faktorielle Validität des Tests liegt vor und spricht für ein gutes Auswahlinstrument.

Die Qualität der Testbatterie wird in zukünftigen Untersuchungen weiter gesichert, insbesondere mit Fokus auf die Geschlechterfairness, einer Überprüfung der Normen sowie einer Untersuchung der prognostischen Validität des Testverfahrens.

Literatur:

Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). BIS-TEST Berliner Intelligenzstruktur-Test – Form 4. Göttingen: Hogrefe.

Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I error and power rates using an effect sizemeasure with logistic regression procedure for DIF detection. Applied Measurement in Education, 14, 329-349. doi:10.1207/S15324818AME1404_2

Kersting, M. (2009). Die DGP Testverfahren - Ein kurzer Rückblick und eine aktuelle Studie zur Konstrukt- und Kriteriumsvalidität des BIS-r-DGP Tests. dgp Informationen, 51, 22-37.

Rieke, J., Brendler, K., & Knorz, C. (2007). Auswahl von Bewerbern/innen in den gewerblichen Berufen: Eine Bewährungskontrolle für den Test „Gewerbliche Ausbildung“ (GA). dgp Information, 59, 62-73.

Ristel, N., & Haarhaus, B. (2014). Bewährungskontrolle der Testbatterie G2. dgp Informationen, 64, 74-79.

Stadelmaier, U., Thiem, H., & Kleinmanns, A. (2018). Eignungsdiagnostik für die Feuerwehr: Bewährungskontrolle des dgp-Tests G2 bei der Berufsfeuerwehr München. dgp Informationen, 68, 96-107.

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. Journal of Educational Measurement, 27(4), 361–370. https://doi.org/10.1111/j.1745-3984.1990.tb00754.x

Thielepape, M. (2003). Ergebnisse einer Bewährungskontrolle der DGP-Eignungstests für eine große Krankenpflegeschule in Niedersachsen. dgp Information, 47, 19-23.

Zumbo, B. D., & Thomas, D. R. (1997). A measure of effect size for a model-based approach for studying DIF. Prince George, Canada: University of Northern British Columbia, Edgeworth Laboratory for Quantitative Behavioral Science.

Kontakt:

Dr. Anna-Lena Jobmann
Diplom-Psychologin
Deutsche Gesellschaft für Personalwesen e.V.
jobmann@dgp.de

Amelie Kleinmanns
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
kleinmanns@dgp.de

Praxis

Die Macht der Selbstwirksamkeit S. 98

Virtuelle Zusammenarbeit und Home-office – Veränderte Beanspruchung und Bewältigungsstrategien S. 104

Führen auf Distanz – (wie) gelingt das? S. 118

Live-Online-Seminare bei der dgp – ein Plädoyer für ein neues Format S. 126

Die Macht der Selbstwirksamkeit

Luciana von Römer

Im Rahmen eines Auswahlprozesses ist es üblich, dass Bewerber*innen den Beobachter*innen und Ausschussmitgliedern ihre Kompetenzen offen darlegen. Im Nachhinein berichten Bewerber*innen, oftmals sehr aufgeregt und nicht besonders zufrieden mit ihrer Leistung gewesen zu sein. Welches psychologische Phänomen liegt dieser Selbsteinschätzung zugrunde, und inwiefern steuert dieses auch die Erfolgschancen der Bewerber*innen?

Die Selbsteinschätzung ist ein komplexes Konstrukt, dem mehrere psychologische Prozesse zugrunde liegen. Im folgenden Artikel beschränken wir uns auf zwei für die Arbeitspsychologie relevante Aspekte der Selbsteinschätzung: die Selbstwirksamkeit (SW) und die Selbstwirksamkeitserwartung (SWE) von Bewerber*innen und Mitarbeiter*innen.

Die Konzepte der SW sowie der SWE wurden maßgeblich von den Arbeiten des kanadischen Psychologen Albert Bandura in den 70er-Jahren geprägt. In seiner weltbekannten Publikation „*Self-efficacy: towards a unifying Theory of Behavioral Change*“ (1977) erläutert A. Bandura die Vorhersagekraft einzelner psychologischer Faktoren auf die SWE. Bandura durchleuchtet die Wirkmechanismen der subjektiven Wahrnehmung und inwiefern diese den Glauben daran, ein Ziel zu erreichen, steuert.

Im Folgende soll zunächst definiert werden, was unter SW und SWE – im psychologischen Sinne – zu verstehen ist (1). Anschließend wird auf die zentrale Rolle der SW und der SWE in der Motivationspsychologie eingegangen (2). Zuletzt soll die Bedeutung der SW und der SWE im Rahmen der Personalauswahl (3) und der Personalentwicklung (4) hervorgebracht werden.

1. Was ist Selbstwirksamkeit und was ist Selbstwirksamkeitserwartung?

Unter SW versteht man die Überzeugung, etwas erreichen zu können. Die damit verbundene Einstellung und Erwartungshaltung ist die SWE. Alexander D. Stajkovic und Fred Luthans (1998) forschten mit A. Bandura an der „Self-efficacy Theory“ und definierten SW als:

„Der Glaube einer Person, dass sie fähig ist, eine bestimmte Aufgabe in einem bestimmten Kontext erfolgreich zu erledigen.“

Die SW beschreibt also einen Zustand, der einem das Gefühl vermittelt, bestimmte Fähigkeiten zu haben. Es ist die Annahme, man könne gezielt Einfluss auf sich und seine Umwelt nehmen. Somit ist die SW maßgeblich für das Verhalten und daher die Motivation von Mitarbeiter*innen mitverantwortlich.

In der Motivationspsychologie spielen die SW sowie die SWE schon seit Langem eine zentrale Rolle. Zahlreiche Studien belegen, dass Menschen mit einer hohen SW und SWE sich realistische und anspruchsvolle Ziele setzen. Bei erschwerten Umständen weisen sie eine höhere Ausdauer auf und halten an ihren Vorsätzen länger fest. Auch bei Rückschlägen intensivieren Personen mit hoher SW und SWE ihre Anstrengungen, um ihre Vorsätze zu erfüllen. Zudem weisen Menschen mit einer hohen SW eine höhere Resilienz auf: Wo andere aufgeben, sehen sie die Möglichkeit, etwas zu verändern (Stajkovic und Luthans, 1998).

Die SW steuert das Verhalten von Mitarbeiter*innen wesentlich mit. Daher ist es durchaus günstig, diese zu erkennen und sie zu fördern. Hierbei gilt es, die situative SW von der allgemeinen SW zu unterscheiden. Die situative SW bezieht sich auf eine konkrete Handlung, z. B. bei einem Vorstellungsgespräch oder der Bearbeitung eines konkreten Projektes. Die allgemeine SW ist die Überzeugung, das Leben als Ganzes und die damit verbundenen Herausforderungen grundsätzlich gut bewältigen zu können. In diesem Artikel wird auf beide Aspekte eingegangen.

2. SW, SWE und Motivation am Arbeitsplatz

Die Motivation am Arbeitsplatz ist sowohl für Mitarbeiter*innen als auch für Führungskräfte von wachsender Bedeutung. Die allgemeine sowie die situative SW spielen hierbei eine wichtige Rolle. Eine ausgeprägte SW wirkt sich positiv auf die Erhöhung der Arbeitsleistung (Shantz et al. 2013) und der Innovation (Hakanen et al. 2008) von Mitarbeiter*innen aus. Studien zufolge kommt es ebenfalls zu einer Reduzierung von Fehlverhalten am Arbeitsplatz (Shantz et al. 2013) sowie zu geringeren Fehlzeiten bei Mitarbeiter*innen (Shaufeli et al. 2009).

Die Auswirkung einer starken SW ist ebenfalls auf der Organisationsebene spürbar. So ist ein positiver Zusammenhang zwischen Profitabilität, Fluktuation sowie Kundenzufriedenheit, Kundenbindung und der Motivation von Mitarbeiter*innen zu verzeichnen (Harter et al. 2002) (Siehe Ab.1).

Abb 1. Auswirkung der Mitarbeiter*innen-motivation (Becker, 2019)

Die SW und SWE sind somit Kernbausteine der Motivation, sie beruhen auf dem Glauben, eine intrinsische Gestaltungsmacht zu besitzen. Arbeitnehmer*innen mit einer hohen SWE machen nicht extrinsische Umstände wie Zufall oder Glück für ihre Erfolge und/oder Misserfolge verantwortlich. Sie rechnen sich diese Leistungen selber an. In der Psychologie spricht man hierbei auch von einer ausgeprägten „Kontrollüberzeugung“.

Die persönliche Motivation vs. Kontrolle am Arbeitsplatz

In der heutigen von der Pandemie geprägten Arbeitswelt, die zahlreiche Mitarbeiter*innen ins Homeoffice gezwungen hat und noch immer zwingt, hat die persönliche Kontrolle von Mitarbeiter*innen stark abgenommen. Dieses Phänomen kennzeichnet nicht nur die Corona-Zeit, schon zuvor ging der Trend in diese Richtung. Denn in den Industrienationen steigt die Anzahl der Mitarbeiter*innen pro Führungskraft stetig (Guadelupe et al. 2013). Um diesem Phänomen entgegenzuwirken, ist es äußerst wichtig, die Selbstkontrolle der einzelnen Mitarbeiter*innen zu erhöhen.

Mitarbeiter*innen mit einer ausgeprägten SW und SWE verfügen über eine hohe Kontrollüberzeugung. Unter Kontrollüberzeugung versteht man das Bewusstsein, ein Ereignis auf irgendeine Weise herbeiführen zu können. Darüber hinaus sind sie der Ansicht, dieses Ereignis mit ihren Kompetenzen bewirken zu können (= SWE). Diese Kompetenzen gilt es gerade in Zeiten von Homeoffice und Dezentralisierung der Kontrolle zu stärken, um Arbeitnehmer*innen zu eigenständigem Handeln zu befähigen und zu motivieren.

Doch wie erkennt man SW und SWE, und wie kann man diese fördern?

3. Selbstwirksamkeit bei der Personalauswahl erkennen

Gerade bei der Personalauswahl liegt es im Interesse von Organisationen, Mitarbeiter*innen mit hoher SW zu erkennen. Wie im Vorfeld bereits erläutert, hängen eine hohe SW und SWE mit erstrebenswerten Verhaltensweisen wie Ausdauer, Zielverfolgung und Erhöhung von Anstrengung zusammen (Becker, 2019).

Positive Verhaltensweisen bei hoher SW wären z. B.:

- → Das Setzen von realistischen und anspruchsvollen Zielen.
- → Das hartnäckige Verfolgen von Zielen und die damit einhergehende Ausdauer.
- → Das Erhöhen von Anstrengungen, gerade wenn Erfolge nicht erzielt wurden, auch bei negativem Feedback von Führungskräften und Kolleg*innen.
- → Menschen mit einer hohe SWE zeigen eine hohe Resilienz gegenüber negativen Situationen auf. Soziale Konflikte, Stress und Misserfolge machen ihnen nicht so viel aus. Wo andere Mitarbeiter*innen aufgeben oder gar Burn-out erfahren, kämpfen sie weiter, um ihre Ziele zu erreichen (Bakker et al. 2005).

Diese Verhaltensweisen, denen eine positive SWE zugrunde liegt, beeinflussen die Leistung von Mitarbeiter*innen positiv (Bandura 1997).

Ausprägungen einer niedrigen SW und SWE wären z. B.:

- → Das Wählen von zu leichten oder zu schweren Zielen, also eine unpassende Selbsteinschätzung.

→ Das vorschnelle Aufgeben bei geringen Herausforderungen.

→ Die Reduzierung von Anstrengungen, wenn Ergebnisse nicht gut genug sind.

→ Das schnelle Aufgeben bei negativen Ergebnissen.

Selbstwirksamkeit ist allerdings nicht mit Selbstüberschätzung zu verwechseln. Diese hat eine negative Ausprägung auf die Performanz von Mitarbeiter*innen. Mitarbeiter*innen, die sich selbst überschätzen, wählen zu schwere und/oder zu viele Aufgaben auf einmal.

Wichtig für die Personalauswahl ist: SW hängt mit Persönlichkeitseigenschaften zusammen und kann mittelbar bei der Personalauswahl genutzt werden.

Studien zufolge weisen Bewerber*-innen mit einem hohen Grad an SW einen hohen Grad an Intelligenz, Gewissenhaftigkeit, Extraversion und emotionaler Stabilität auf

(Judge et al. 2007). Personen mit hohen Werten auf diesen Persönlichkeitsdimensionen, die sich auch in bestimmten Schlüsselkompetenzen widerspiegeln, haben meist eine höhere SWE als Menschen mit niedrigen Werten auf diesen Dimensionen (Becker, 2019). Daher gilt es, diese Dimensionen z. B. in Assessment-Centern und bei Einstellungsverfahren zu berücksichtigen, zu beobachten und zu bewerten. Es ist vorteilhaft, bei der Personalauswahl insbesondere Bewerber*innen auszuwählen, die eine solide Basis an SW und SWE aufweisen. Denn die SW hängt maßgeblich mit der Selbstentwicklungskompetenz von Bewerber*innen (Köpke, 2018) und Arbeitnehmer*innen zusammen. Nachfolgend wird auf die Entwicklung von Selbstwirksamkeit bei Beschäftigten eingegangen.

4. Selbstwirksamkeit entwickeln

Um SW bei Mitarbeiter*innen zu fördern, hat Albert Bandura (1997) im Rahmen seiner Studien vier Hypothesen untersucht:

1. Die Erfahrung von Erfolgserlebnissen: Erfolgserlebnisse sowie ein reicher Erfahrungsschatz nehmen den Menschen die Ängste vor neuen Herausforderungen. Wenn diese auch noch zur Routine werden, steigern sie das Gefühl der Beherrschung bei Mitarbeiter*innen. Das Wiederholen von Erfolgserlebnissen steigert die Selbstständigkeit und das Selbstvertrauen zur erfolgreichen Ausführung einer Aufgabe. Hierbei ist es wichtig, dass Erfahrungen positiv empfunden werden, daher sollten regelmäßige Erfolgserlebnisse gefördert, benannt und zelebriert werden.

2. Die Erweiterung von Kompetenzen: Wer etwas gut kann, wird es auch praktizieren. Daher ist es wichtig, die Fähigkeiten der Mitarbeiter*innen zu entwickeln und langfristig auszubauen. Hierdurch erlangen sie das Bewusstsein, kompetent zu sein.

3. Die Rolle der Vorbilder: Gleichgestellte bei Erfolgserlebnissen zu beobachten, steigert das eigene Gefühl, eine Aufgabe oder Herausforderung meistern zu können. Frei nach dem Motto „wenn die/der das kann, kann ich das doch erst recht!“. Daher ist es gerade als Führungskraft wichtig, eine Vorbildfunktion einzunehmen und auch Mitarbeiter*innen dazu zu bewegen, füreinander als Vorbilder zu funktionieren.

4. Die positiven Erwartungen des Umfeldes als Ansporn: Für die positive Einschätzung der eigenen Leistungsfähigkeit ist die Erwartungshaltung des sozialen Umfelds von zentraler Bedeutung. Das gilt umso mehr, wenn die Erwartungen in Form von Ermutigungen geäußert werden. Dieses Phänomen wird in der Psychologie auch Rosenthal-Effekt genannt oder die sich selbst erfüllende Prophezeiung. Studien zufolge zeigt eine positive Erwartungshaltung eine kontinuierliche positive Entwicklung des Verhaltens der Person. Niedrige Erwartungshaltungen bewirken das Gegenteil.

Es zeigt sich, dass SW sowie SWE eine entscheidende Rolle bei der Motivation und dem positiven Verhalten von Mitarbeiter*innen spielen. Allein der Glaube daran, ein Ziel zu erreichen, unabhängig von den eigenen Kompetenzen, wirken sich positiv auf das Verhalten aus. Darum gilt es, diesen „Placebo-Effekt“ gerade für die Mitarbeiter*innenmotivation zu nutzen.

Fazit

Ziel dieses Artikels war es, SW als messbare Ordnungsdimension in Auswahlverfahren zu erkennen sowie die SW und die damit verbundene SWE als wesentliche Triebfeder der Arbeitsmotivation darzustellen. Wichtig für die Personalauswahl ist der deutliche Zusammenhang zwischen Merkmalen einer Person und dem Grad der SW. Es gilt daher, insbesondere Mitarbeiter*innen auszuwählen, die hohe Werte bei der allgemeinen Intelligenz und den Persönlichkeitsdimensionen bzw. Kompetenzfacetten der Gewissenhaftigkeit, der Extraversion und der emotionalen Stabilität aufweisen.

Bei der Personalentwicklung ist es wichtig, die SW und SWE bei Mitarbeiter*innen zu fördern. Diese hängen nicht zwangsläufig mit einer höheren Kompetenz der Mitarbeiter*innen zusammen, wirken sich jedoch positiv auf die Motivation der Einzelnen aus. Dafür sind Erfolgserlebnisse, die Entwicklung und der Aufbau von relevanten Kompetenzen sowie die Präsenz von Vorbildern und die Bestärkung des Glaubens an das eigene Können von zentraler Bedeutung.

Literatur:

Guadelupe, M., Li., H., & Wulf, J. (2013). Who lives in the C-suite? Organizational structure and the division of labor in top management, Management Science, 60(4), 824-844

Stajkovic, A. D., & Luthans (1998). Self-efficacy and work-related performance: a meta-analysis. Psychological Bulletin, 124(2), 240-261

Bandura, A. (1977). Self-efficacy: the exercise of control. New Yorck: Freeman.

Bakker, A. B., Demerouti, E. and Euwema, M. C. (2005). Job ressources buffert he impact of job demands on burnout. Journal of Occupational Health Psychology, 10(2), 170-180.

Becker, F. (2019). Mitarbeiter wirksam motivieren. Berlin: Springer.

Hakanen, J. J., Perhoniemi, R. und Toppinen-Tanner, S. (2008). Positive gain spirals at work: from job ressources to work engagement, personal initiative and work-unit innovativeness. Journal of Vocational Behaviour, 73 (1), 78-91

Judge, T. A., Jackson, C. L., shaw, J. C., scott, B.A. und Rich B. L. (2007). Ist he effect of self-efficacy on job/task performance an epiphenomenon? Journal of Applied Psychology, 92, 107-127.

Köpke, S. (2018). Selbstentwicklungskompetenz - Eine Meta-Schlüsselkompetenz. dgp Informationen 2018.

Shantz, A., Alfes, K., Truss, C., und Soane, E. (2013). The role of employee engagement in the relationship between job design and task performance, citizenship and deviant behaviours. The international journal of human Ressource Management, 24 (13), 2608 -2627

Schaufeli, W. B., Taris, T. W. und Van Rhenen, W. (2008). Workaholism, burnout, and workengagement: three of a kind or three different kinds of employee well-beeing? Applied Psychology, 57(2), 173-203

Kontakt:

Luciana von Römer
M. Sc. Clinical Psychology
Deutsche Gesellschaft für Personalwesen e.V.
roemer@dgp.de

Virtuelle Zusammenarbeit und Homeoffice – Veränderte Beanspruchung und Bewältigungsstrategien

Kaschenbach, Riechmann & Schubert

„Sehen wir uns später noch, um über unser Projekt zu sprechen?“ Was vor der Corona-Pandemie noch relativ leicht mit „Ja“ oder „Nein“ zu beantworten gewesen wäre, lässt sich heute meist nur noch mit einem „Ja, aber…“ beantworten. Ja, aber ist denn unser Besprechungsraum groß genug? Ja, aber können wir denn ausreichend Abstand einhalten? Ja, aber haben wir auch genug Masken vor Ort? Ja, aber hast du mir die Einwahldaten schon geschickt? Ja, aber ich muss mich vorher noch schnell testen lassen. Ja, aber ich lasse die Kamera aus, ich bin im Homeoffice.

Unsere Arbeitswelt verändert sich. Was früher analog geschah, passiert heute digital. Treffen mit Kolleg*innen oder Kund*innen sind teilweise nur einen Mausklick entfernt. Nicht erst seit der Corona-Pandemie, denn Formen der Arbeit verändern sich stetig – Hierarchien werden flacher, neue Teamstrukturen und agiles Arbeiten ermöglichen neue Formen der Zusammenarbeit. Virtuelle Zusammenarbeit, wie im einleitenden Beispiel suggeriert, gibt es auch nicht erst seit der Corona-Pandemie, sondern sie ist unter anderem bei räumlich getrennten Teams schon immer eine sinnvolle Arbeitsform gewesen (vgl. Lorenz & Grünhagen-Scheele, 2019) und in vielen Organisationen bereits Gang und Gäbe. Dass wir mit unserer Kollegin, mit der wir ein Büro teilen, plötzlich ausschließlich per Videokonferenz kommunizieren, ist jedoch eine Veränderung, die sich nicht durch den ständigen Wandel der Arbeitswelt ergeben hat. Vielmehr haben die Corona-Pandemie und die daraus resultierenden Maßnahmen zum Gesundheitsschutz und der Kontaktbeschränkungen es notwendig gemacht, unsere Zusammenarbeit mit Kolleg*innen, Vorgesetzten oder Kund*innen zu verändern. Kontakte und nicht notwendige Treffen sollen vermieden werden, der „kurze“ Dienstweg ist plötzlich nicht mehr so kurz, und wir sehen unsere Kolleg*innen häufig nur noch per Video.

Aus diesen Entwicklungen, mit denen viele Arbeitnehmer*innen und Organisationen konfrontiert werden, resultieren veränderte Anforderungen, denen sich jede*r von uns gegenüber sieht: Anforderungen an Zusammenarbeit, an Arbeits- und Selbstorganisation sowie Zeitmanagement oder auch an den persönlichen Umgang mit der Pandemie- und Arbeitssituation, in der sich alle gegenwärtig befinden. Der Artikel thematisiert daher, welchen Herausforderungen sich Arbeitnehmer*innen und Organisationen aufgrund der virtuellen und mobilen Arbeitsweise stellen müssen, zeigt Bewältigungsstrategien im Umgang mit

diesen Herausforderungen auf und vertieft dabei insbesondere die Themen „Erfolgsfaktoren für das Arbeiten im Homeoffice“ sowie „Gesundheitsorientierte Führung im Homeoffice“. Hierzu werden erste Studienergebnisse berücksichtigt, die sich mit den oben beschriebenen Themen beschäftigen.

1. Mobiles Arbeiten während der Corona-Pandemie

Als es im März 2020, in der ersten Welle der Corona-Pandemie, hieß #stayhome, stand für viele Arbeitnehmer*innen und Organisationen eine große Umstellung an: Arbeitsplätze wurden nach Hause verlegt (wenn möglich), Kund*innentermine fielen aus oder fanden virtuell stand, Dienstreisen wurden abgesagt, berufliche und private Kontakte wurden reduziert bzw. vermieden. Organisationen sahen sich der Herausforderung gegenüber, plötzlich digitale Infrastruktur zu schaffen, um ihre Mitarbeiter*innen ins sogenannte Homeoffice zu schicken. In Organisationen, in denen mobile Arbeitsformen kaum denkbar waren, wurden neue Systeme etabliert, Endgeräte beschafft und Arbeitsprozesse an digitale und mobile Arbeitsformen angepasst, was viele Organisationen vor große technische und organisatorische Herausforderungen stellte.

Der jährlich herausgegebene DAK-Gesundheitsreport (Marschall, Hildebrandt, Kleinlercher & Nolting, 2020) hat in einer Sonderauswertung für das Jahr 2020 den Arbeitsstress vor dem Hintergrund der Digitalisierung der Arbeit unter Berücksichtigung der Corona-Pandemie untersucht. Neben gesundheitlichen Auswirkungen der Digitalisierung der Arbeit wurde auch der Umgang mit der Digitalisierung und der Ausweitung von Homeoffice untersucht. Auf die Aussage „In der Corona-Krise hat mein Arbeitgeber die Möglichkeiten, über digitale Arbeitsmethoden im Homeoffice zu arbeiten, sprunghaft ausgeweitet“ antworteten branchenübergreifend 57% aller Befragten mit „trifft genau/eher zu“ (Marschall et al., 2020). Innerhalb der Branchen zeigen sich jedoch große Unterschiede, die nur wenig überraschen. In Branchen, in denen viel Kontakt mit Kund*innen, Klient*innen oder Patient*innen herrscht, konnten digitale Arbeitsmethoden für das Arbeiten im Homeoffice kaum ausgeweitet werden. Im Sozial- und Gesundheitswesen stimmten lediglich 35% bzw. 29% der Befragten der oben genannten Aussage zu. In der öffentlichen Verwaltung gaben im Gegensatz dazu fast drei Viertel der Befragten an (72%), dass ihr Arbeitgeber die Möglichkeiten zur Arbeit im Homeoffice sprunghaft ausgeweitet habe (Marschall et al., 2020), ebenso im Bank- und Versicherungssektor (80%), die als „Spitzenreiter“ aus dieser DAK-Erhebung hinsichtlich der Ausweitung der digitalen Arbeitsmethoden hervorgehen. Es ist anzunehmen, dass diese Branchenunterschiede, wie bereits berichtet, unter anderem in der Art der Tätigkeit und der Möglichkeit sowie Sinnhaftigkeit, die Arbeit aus dem Homeoffice zu gestalten, begründet liegen.

Dabei ist das Homeoffice nur eine Form des mobilen Arbeitens, wozu unter anderem auch Dienst- und Geschäftsreisen oder Termine bei Kund*innen vor Ort zählen (vgl. DGB Index, 2020). Vor Ausbruch der Corona-Pandemie arbeiteten nach einer Studie der Hans-Böckler-Stiftung (2021) lediglich 4% der Befragten ausschließlich oder überwiegend von zu Hause aus. Wie die Studie, die Personen in den einzelnen Wellen der Corona-Pandemie befragte, zeigt, hat sich diese Zahl im ersten bundesweiten „Lockdown“ fast vervierfacht, flachte über den Herbst hinweg wieder ab und war im Januar 2021, als die „dritte Welle“ der Pandemie begann, wieder auf einem ähnlich hohen Niveau wie im April 2020 (vgl. Abbildung 1). Dass Organisationen und Mitarbeiter*innen auch nach der Corona-Pandemie weiterhin erwarten, im Homeoffice zu arbeiten, überrascht dabei nur wenig. Laut der Befragung der Hans-Böckler-Stiftung (2021) aus dem Juni 2020 erwarten 71% der Befragten, auch zukünftig häufiger im Homeoffice zu arbeiten als vor der Pandemie. Wo vorher Homeoffice kaum denkbar war, sah man sich nun, spätestens seit der Homeoffice-Verordnung des Bundes im März 2021, damit konfrontiert, einen Versuch zu wagen. Vielerorts haben sich neben Befürchtungen zum „von zu Hause aus arbeiten“ auch viele Vorteile, neben dem der Pandemiebekämpfung, gezeigt. Man hat gesehen, dass „es funktioniert“ – Bedenken von Arbeitgeber*innen- und Mitarbeiter*innen-Seite haben sich teilweise zerschlagen.

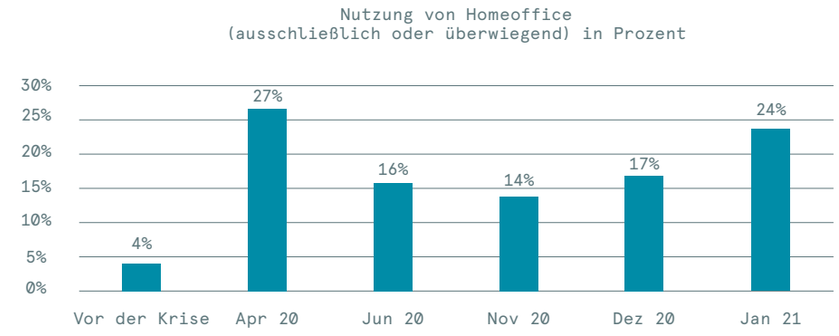


Abbildung 1. Arbeiten im Homeoffice während der Corona-Pandemie (nach einer Studie der Hans-Böckler-Stiftung).

2. Entgrenzung von Arbeits- und Privatleben

Das Arbeiten von zu Hause existiert nicht erst seit Corona, sondern wurde auch vorher schon mehr oder weniger regelmäßig von Arbeitnehmer*innen wahrgenommen. Dass das Arbeiten von zu Hause die Entgrenzung von Arbeits- und Privatleben auslöst, also ein „Verschwimmen“ der Grenzen der Arbeitszeit und der Freizeit, verwundert nicht. Diese Entgrenzung kann dabei positive wie negative Folgen haben. Der DGB-Index aus 2020 befragte hierzu Personen, die auch vor der Corona-Pandemie bereits von zu Hause arbeiteten.

Von diesen im Homeoffice tätigen Personen können 46% sehr häufig oder oft in ihrer arbeitsfreien Zeit nicht richtig abschalten. 34% der im Homeoffice arbeitenden Befragten gaben an, dass sie sehr häufig oder oft Schwierigkeiten haben, private Interessen mit ihrer Arbeit zeitlich zu vereinbaren.

Wie kann das sein – wenn doch gerade durch das Homeoffice z.B. Arbeitswegzeiten eingespart werden und demnach mehr Zeit zu Hause verbracht wird? Dass Mitar-

beiter*innen gerade aus Vereinbarkeitsgründen zwischen Beruf und Familie bzw. Privatem von zu Hause aus arbeiten, zeigen auch die Zahlen: 30% (38% bei Personen mit minderjährigen Kindern im Haushalt) der Befragten gaben an, gerade aus diesem Grund im Homeoffice zu arbeiten, um eine bessere Vereinbarkeit von Arbeit und Privatem gewährleisten zu können, 26% möchten Fahrtzeiten vermeiden.

In der Studie der Hans-Böckler-Stiftung (2021), die bei den Befragten auch Arbeitnehmer*innen miteinschließt, die erst seit Ausbruch der Corona-Pandemie im Homeoffice arbeiten, stellen sich die Zahlen etwas anders dar: Im Juni 2020 gaben 77% der Befragten an, dass das Homeoffice die Vereinbarkeit von Beruf und Familie „voll und ganz“ oder „eher ja“ erleichtern würde. Auf der anderen Seite sagen 53%, dass sie für Arbeitgeber*in, Kolleg*innen oder Kund*innen länger erreichbar sind, als vor der Krise. Gerade in Zeiten der Hochphase der Corona-Pandemie, als Arbeitnehmer*innen ins Homeoffice geschickt wurden sowie Kitas und Schulen geschlossen wurden, wurde aus dem Homeoffice plötzlich das „family office“, bei dem möglicherweise Eltern sowie Kinder von zu Hause aus arbeiten und lernen mussten. Dies könnte ein Erklärungsansatz sein, weshalb die Vereinbarkeit von Beruf und Privatem verschwimmt und eine Trennung nicht mehr gelingt.

Der dritte Gleichstellungsbericht der Bundesregierung (2020), der kurz vor Ausbruch der Corona-Pandemie veröffentlich wurde, zeigt, dass Arbeitnehmer*innen, die von zu Hause arbeiten, mehr Zeit in informelle Sorgearbeit (also z.B. Aufgaben im Haushalt, der Kindererziehung usw.) investieren, als wenn sie nicht von zu Hause arbeiten (Samtleben, Lott & Müller, 2020). Es ist anzunehmen, dass durch die Notwendigkeit, die eigenen Kinder zusätzlich während der Arbeitszeit betreuen oder zumindest beaufsichtigen zu müssen, der Anteil an Sorgearbeit, der sowieso durch das Arbeiten im Homeoffice zunimmt, noch mehr Zeitinvestment in Sorgearbeit notwendig macht. Dieser Effekt führt möglicherweise dazu, dass die Mittagspause schnell zur Hausaufgabenpause wird, zwischen zwei Meetings schnell private Termine im Kinderzimmer stattfinden oder auch der Haushalt nebenbei gepflegt wird, was automatisch zum Verschwimmen der Grenzen von Arbeit und Privatleben führt. Dass diese Entgrenzung nicht unbedingt ausschließlich negative Folgen haben muss, lässt sich leicht schlussfolgern: Durch die mögliche selbstständige und freie Arbeitszeitgestaltung gelingt es vielen Personen leichter, privaten Verpflichtungen nachzukommen und dennoch die Arbeitszeit nicht zu vernachlässigen. Abends nachdem man die Kinder ins Bett gebracht hat, oder noch einen Film geschaut hat, eine Mail zu schreiben, macht nicht jeder/m Arbeitnehmer*in etwas aus – ganz im Gegenteil: Die neu gewonnene Freiheit kann auch zur besseren Vereinbarkeit von Beruf- und Privatleben beitragen.

Es gilt daher, dass die reine Tatsache, von zu Hause aus arbeiten zu können, allein scheinbar nicht garantiert, dass das Arbeits- und Privatleben besser organisiert werden kann.

Vielmehr spielen der Umgang und das Verhalten im Homeoffice sowie die eigene Einstellung zur möglichen Entgrenzung zwischen Arbeits- und Privatleben sowie die häuslichen bzw. familiären Umstände eine tragende Rolle, worauf in der ersten Vertiefung „Erfolgsfaktoren für das Arbeiten im Homeoffice“ weiter eingegangen wird.

Neben der persönlichen Einstellung und dem individuellen Umgang mit der Entgrenzung von Arbeits- und Privatleben existieren weitere Faktoren, die erfolgreiches Arbeiten aus dem Homeoffice fördern können. Das Fraunhofer Institut (Hofmann, Piele & Piele, 2021) kommt in einer kürzlich veröffentlichten Studie zum Thema „Entgrenzungseffekte durch mobiles Arbeiten – und deren Vermeidung“ zu dem Schluss, dass eben nicht nur den Beschäftigten selbst, sondern auch den Führungskräften eine entscheidende Rolle zukommt, wenn es darum geht, Handlungsbedarfe der Arbeit im Homeoffice zu erkennen und angemessene Lösungen im Umgang mit der Entgrenzung zu finden. 52% der befragten Arbeitnehmer*innen der Fraunhofer-Studie sehen die direkte Führungskraft in der Hauptverantwortung, 48% die Unternehmensleitung und 44% sehen die Beschäftigten selbst in der Position, den Umgang mit der Entgrenzung von Arbeits- und Privatleben hauptverantwortlich zu gestalten (Hofmann et al., 2021). Da die Führungskraft und die Unternehmensleitung in ihren Positionen am meisten Gestaltungsspielraum haben, wenn es darum geht, Arbeitsbedingungen zu verändern oder Aufgabenverteilungen zu gestalten, überrascht es nicht, dass diese in der Hauptverantwortung gesehen werden.

Dass die Führungskräfte neben dieser erhöhten Organisation und Delegation von Aufgaben und Verantwortlichkeiten selbst vor einer großen Herausforderung stehen, nämlich dem virtuellen Führen bzw. dem Führen auf Distanz, thematisiert die zweite Vertiefung dieses Artikels, die sich mit dem Thema „Gesundheitsorientierte Führung im Homeoffice“ beschäftigt.

3. Erfolgsfaktoren für das Arbeiten im Homeoffice

Die Auswirkungen von Homeoffice scheinen auf den ersten Blick widersprüchlich zu sein. Wie kann es sein, dass Mitarbeiter*innen im Homeoffice zwar zufriedener sind, jedoch gleichzeitig häufiger unter Belastungssymptomen leiden? Wer früher im Büro zum Feierabend die Tür hinter sich zuzog, der zog damit in gewisser Weise auch eine Grenze zwischen dem Berufs- und dem Privatleben. Diese Grenze ist im Homeoffice nicht mehr so eindeutig. Die veränderten Arbeitsbedingungen im Homeoffice erfordern daher auch eine Verhaltensanpassung.

Ständige Erreichbarkeit, Arbeiten während der Freizeit, Multitasking – ob sich die Nutzung von Homeoffice gesundheits- und leistungsförderlich auswirkt oder nicht, hängt unter anderem von der eigenen „Homeoffice-Tauglichkeit“ ab.

Die positive Nachricht ist, dass die „Homeoffice-Tauglichkeit“ durch die Nutzung individueller Handlungsstrategien trainiert werden kann, sodass förderliche Auswirkungen von Homeoffice begünstigt werden. Die Auswahl der Handlungsstrategien sollten dabei unter anderem auf die individuellen Bedürfnisse der Personen abgestimmt sein. Dabei spielt das persönliche Abgrenzungs- respektive Integrationsbedürfnis eine wesentliche Rolle.

In diesem Zusammenhang veranschaulichen Ashforth, Kreiner und Fugate (2000) mit der Boundary-Theorie die Art und Weise, wie Individuen Grenzen setzen, aufrechterhalten oder verändern, um die Welt um sie herum zu vereinfachen und zu klassifizieren. Die Grenzen lassen sich auf

einem Kontinuum von schwach (durchlässig für andere Domänen) bis hin zu stark ausgeprägt (undurchlässig für andere Domänen) verorten (Nippert-Eng, 1996). Die „Boundary-Work-Tactics" beschreiben dabei Verhaltensstrategien, anhand derer Grenzen zwischen dem Berufs- und dem Privatleben gezogen und gehalten werden. Entscheidend ist, dass die Auswahl der Verhaltensstrategien zum einen von den situativen Faktoren abhängen und zum anderen auch von dem individuellen „Boundary"-Persönlichkeitstypen.

Die verschiedenen „Boundary"-Persönlichkeitstypen werden auf einem Kontinuum vom stark segmentierenden bis hin zum stark integrierenden Typen unterteilt.

Während stark segmentierende Persönlichkeitstypen großen Wert darauf legen, die Lebensdomänen Arbeit und Privat möglichst getrennt zu halten, legen stark integrierende Persönlichkeitstypen großen Wert darauf, die beiden Lebensdomänen optimal und gewinnbringend ineinander zu verschachteln (Gisin, Schulze & Degenhardt, 2016). Beispielhaft betrachten Segmentierer*innen das Wochenende als private Erholungsphase und ziehen damit eine klare Grenze, in dem sie sich von beruflichen Reizen abschirmen. Integrierer*innen hingegen lassen berufliche Reize am Wochenende zu und bleiben bezüglich der Einteilung von Freizeit oder Arbeitszeit flexibel.

Abbildung 2. Die „Boundary"-Persönlichkeitstypen auf einem Kontinuum der Segmentierung / Integration der Lebensdomänen Privat- und Berufsleben.

Zwischen diesen beiden gegensätzlichen Persönlichkeitstypen befindet sich der sogenannte Mischtyp („mixed type"), welcher sich situativ unterschiedlich, in Abhängigkeit der subjektiv zugesprochenen Relevanz einer Aufgabe, verhält. Anders als bei den zuvor genannten Persönlichkeitstypen, sind Mischtypen die statt seine Bedürfnisse hinsichtlich der Entgrenzung oder Abgrenzung beider Lebensdomänen weitestgehend unbewusst. Der Mischtyp nutzt „Boundary"-Taktiken daher weniger stringent und zeigt die größten Schwierigkeiten im Umgang mit Homeoffice.

Nach Kreiner, Hollensbe und Sheep (2009) lassen sich „Boundary"-Taktiken in die Kategorien verhaltensbasiert („behavioral), zeitlich („temporal"), physisch („physical") und kommunikativ („communicative") unterteilen. In der nachfolgenden Tabelle werden zu den jeweiligen Kategorien exemplarisch Taktiken genannt.

Kategorie	Taktik	Beispiel
Verhaltensbasiert	Personen von außen mit einspannen	Den Partner oder die Partnerin bei der Lösungsfindung einer beruflichen Aufgabe um Rat fragen
	Einsatz von Technologien, um Grenzen auf-/abzubauen	„Nach 18:00 Uhr schalte ich mein Arbeitshandy auf stumm."
	Priorisierung von arbeitsbezogenen und/oder privaten Anforderungen	Erstellung einer To-Do-Liste, in welcher die Aufgaben nach Wichtigkeit, Dringlichkeit und persönlicher Relevanz aufgelistet werden
Zeitlich	Bestimmung von Pufferzeiten für spontane berufsbezogene und/oder private Tätigkeiten	„In diesem Zeitraum reagiere ich nur auf Anfragen, wenn diese wirklich nicht warten können."
	Pausen einplanen, in denen man ausschließlich berufsbezogene oder private Dinge ausübt	„In diesem Zeitraum befasse ich mich ausschließlich mit der Aufgabe X, alles andere muss warten."
Physisch	Physische Distanz schaffen oder abbauen	„Mein Arbeitshandy bleibt auf meinem Schreibtisch."
Kommunikativ	Erwartungsmanagement im Voraus	„Das Wochenende ist mir heilig, in diesem Zeitraum stehe ich auf keinen Fall zur Verfügung."
	Konfrontation bei Grenzüberschreitung	„Sie erwischen mich in meiner freien Zeit, gerne können wir morgen, wenn ich wieder im Büro bin, einen Termin ausmachen."

Tabelle 1. „Boundary"-Taktiken nach den Kategorien: verhaltensbasiert, zeitlich, physisch und kommunikativ

Jeder „Boundary"-Typ hat seine spezifischen Schwierigkeiten im Umgang mit Homeoffice. So fällt es Integrierer*innen häufiger schwer, den Überblick über die geleistete Arbeit zu behalten, weshalb diese stärker gefordert sind, sich aktiv Freiräume zu schaffen und diese auch zu nutzen, um einer drohenden Arbeitsüberlastung und den damit einhergehenden Belastungssymptomen entgegen zu wirken.

Segmentierer*innen hingegen sollten lernen, sich hin und wieder Zeiträume zu schaffen, in denen sie flexibel sowohl auf private als auch auf berufliche Anforderungen reagieren. Die hierdurch erworbene Flexibilität lässt sich besser mit den speziellen Bedingungen im Homeoffice vereinbaren. Gleichermaßen können den Segmentierer*innen Rituale helfen, dem Bedürfnis nach Abgrenzung der beiden Lebensdomänen zu entgegnen. Beispielhaft hierfür wäre das Tragen formeller respektive arbeitsangemessener Kleidung im Homeoffice. Mit diesem Ritual wird zum einen der Arbeitsbeginn bzw. das Arbeitsende eingeläutet und zum anderen der Rollenwechsel in berufsbezogene oder eben private Rollen verdeutlicht.

Die Herausforderung des Mischtypen liegt hingegen darin, die eigenen Bedürfnisse besser zu verstehen und zu erkennen, um so bei der Auswahl der „Boundary"-Taktiken bedürfnisorientierter und stringenter vorzugehen. Diese Erkenntnisse können unter anderem durch Maßnahmen der Selbstbeobachtung gewonnen werden.

Neben den Merkmalen der Mitarbeiter*innen spielen äußere Faktoren eine wesentliche Rolle. So nutzen Homeoffice-Kompetenzen sowie die digitale Reife nur wenig, wenn die notwendige technische Ausstattung im Homeoffice unzureichend ist. Aus einer Untersuchung des Fraunhofer Instituts (Hofmann, Piele & Piele, 2021) geht hervor, dass die wahrgenommene Leistungserbringung den stärksten Einfluss auf die Zufriedenheit, Motivation und Leistung der Mitarbeiter*innen im Homeoffice hat. Die wahrgenommene Leistungserbringung hat demnach einen stärkeren Einfluss auf die Arbeitszufriedenheit der Homeoffice-Tätigen als die wahrgenommene Nähe zu den Arbeitskolleg*innen. Eine unzureichende technische Ausstattung wirkt dabei der empfundenen und sicherlich auch der tatsächlichen Leistungserbringung entgegen.

Daneben sind die individuellen Arbeitsbedingungen zu Hause ausschlaggebend. So konnte in der genannten Befragung festgestellt werden, dass die Arbeitszufriedenheit von Homeoffice-Tätigen geringer ist, wenn sich im selben Haushalt ein betreuungsbedürftiges Kind aufhält. Grund hierfür ist die zum Teil problematische Vereinbarkeit von Familie und Beruf. Auch wenn technische und arbeitsorganisatorische Faktoren einen stärkeren Einfluss auf die Leistung im Homeoffice haben als die menschlichen Faktoren, so spielen diese Faktoren eben auch eine wesentliche Rolle. Ein regelmäßiger Kontakt zu den Kolleg*innen, zum Beispiel durch (virtuelle) Teammeetings sowie der regelmäßige Austausch zur verantwortlichen Führungskraft tragen entscheidend dazu bei, dass die Zusammenarbeit auf Distanz gelingt.

Um die zuvor genannten Faktoren zu gewährleisten, muss eine passende Unternehmenskultur vorhanden sein. So geht Homeoffice in Teilen mit der Arbeit zu „unüblichen" Zeiten einher, weshalb eine Vertrauenskultur und damit auch ein auf Vertrauen und Flexibilität beruhendes Arbeitszeitmodell unausweichlich ist.

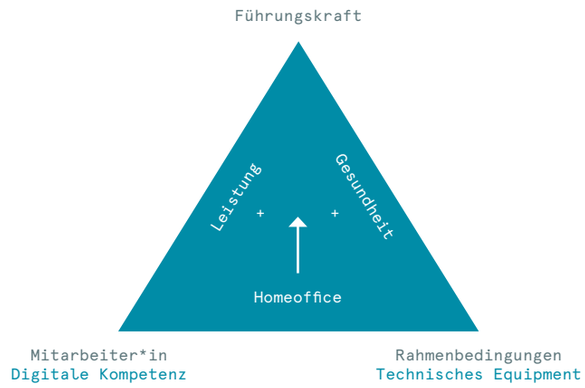


Abbildung 3. Darstellung der internen und externen Erfolgsfaktoren für das Arbeiten im Homeoffice.

Bisher noch nicht thematisiert wurde die Rolle der Führungskraft. Dabei haben die Führungskräfte in diesem Kontext eine entscheidende Schlüsselfunktion für ihre Mitarbeiter*innen und bilden praktisch das Bindeglied zwischen den externen und den personenbezogenen Erfolgsfaktoren im Homeoffice. Daher soll im nachfolgenden Abschnitt verstärkt auf die Rolle der Führungskräfte in Bezug auf die Zusammenarbeit auf Distanz eingegangen werden.

4. Gesundheitsorientierte Führung im Homeoffice

Wie bislang deutlich wurde, führen die veränderte Zusammenarbeit und die veränderten Arbeitsbedingungen in Zeiten von Homeoffice und in virtuellen Teams zu Veränderungen im Anforderungs-/Belastungsspektrum und den Ressourcen der Mitarbeiter*innen.

Die digitale Zusammenarbeit beeinflusst aber auch die Einflussmöglichkeiten von Führung und stellt Führungskräfte vor neue Herausforderungen im Vergleich zur Zusammenarbeit vor Ort.

Unter Berücksichtigung dieser beiden Aspekte gerät insbesondere die gesundheitsorientierte Führung in den Fokus: Wie können Vorgesetzte im Sinne der Fürsorgepflicht Einfluss auf die Gesundheit der Beschäftigten im Homeoffice nehmen?

Definiert man Gesundheit als das „...Ergebnis der Auseinandersetzung mit Belastungen und Anforderungen... auf psychosozialer und physischer Ebene" (Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie, 1997), so kommt Führung hier eine besondere Bedeutung zu. Führung wirkt auf Mitarbeiter*innen sowohl als Stressor als auch als Ressource bei der Bewältigung der täglichen Arbeitsaufgaben. Viele Studien belegen sogar einen entscheidenden Einfluss der Führungskräfte auf Arbeitsfähigkeit bzw. Gesundheit der Beschäftigten. Lassen sich diese Ergebnisse auf Führung in digitalen Kontexten übertragen?

In den vorliegenden Studien zu Führung und Zusammenarbeit konnten vier Einflussfaktoren eruiert werden, die gesundheitsorientierte Führung in virtuellen Teams unterstützen können (Franke und Felfe, 2011):

(1) Die direkte Interaktion wird bisweilen als definitorisches Merkmal von Führung beschrieben. Insofern kann eine Führungskraft durch ihr Verhalten und ihre Kommunikation die Gesundheit der Mitarbeiter*innen direkt beeinflussen.

(2) Darüber hinaus gestalten Führungskräfte Arbeitsbedingungen und Arbeitsumfeld. Damit nehmen sie indirekt Einfluss auf die Gesundheit und das Wohlbefinden der Beschäftigten.

(3) Wie gesundheitsbezogen ist die Selbstführung? Voraussetzung für eine gesundheitsorientierte Mitarbeiter*innenführung ist ein achtsamer und wertschätzender Umgang mit der eigenen Gesundheit.

(4) Zuletzt fungiert eine Führungskraft als Vorbild für ihre Mitarbeiter*innen. Dieser Faktor ist lange sträflich unterschätzt worden. Durch ihr eigenes Gesundheitsverhalten nehmen Führungskräfte erheblichen Einfluss auf das Verhalten der Mitarbeiter*innen.

Somit ergeben sich für die gesundheitsorientierte Führung vier Felder (Felfe, 2013):

Faktor	Ressource	Stressor
Selbstführung	•Auseinandersetzung mit dem eigenen Stresserleben und persönlichen Belastungsfaktoren •Sensibilisiert für eigene Grenzen •Selbstfürsorge als Grundlage für die Fürsorge der Mitarbeiter*innen	•Belastungssituation der Mitarbeiter*innen wird nicht angemessen wahrgenommen •Gesundheit der Mitarbeiter*innen wird nicht gefördert •Weitergabe von Stress und Druck
Direkter Einfluss (Kontakt)	•Lob und Anerkennung •Anbieten von Unterstützung •Wertschätzung und Respekt •Vertrauen	•Distanzierter, sachlicher Kontakt •Misstrauen und Kontrolle •Gleichgültigkeit •Bevormundung
Indirekter Einfluss (Arbeitsgestaltung)	•Kooperation im Team •Handlungs- und Entscheidungsspielräume •Terminpuffer •Partizipation und Beteiligung •Klare Prioritäten •Entsprechende Arbeitsmittel bereitstellen	•Ungünstige Aufgabenverteilung •Unklare Prioritäten •Enge organisatorische und zeitliche Vorgaben •Über- oder Unterforderung
Vorbildfunktion	•Authentizität und Glaubwürdigkeit des gesundheitsförderlichen Verhaltens •Multiplikator für das Betriebliche Gesundheitsmanagement	•Führungskraft lebt nicht vor, was sie fordert

Bislang existieren nur wenige Studien zu gesunder Führung in virtuellen Teams. Kann Gesundheit und Wohlbefinden auch über Distanz und virtuell vermittelt erreicht werden?

Die Anpassung des Führungsverhaltens an die daraus resultierenden Aufgaben in digitaler Zusammenarbeit ist von entscheidender Bedeutung. Im Folgenden sollen hierzu einige Anregungen gegeben werden.

Auch im Kontext der digitalen Zusammenarbeit kommt der Führungskraft eine Vorbildfunktion zu, z.B. bei der Einhaltung der Arbeits- und Pausenzeiten. Aufgrund des geringen direkten Kontaktes ist die Wahrnehmung dieses Vorbildverhaltens durch die Mitarbeiter*innen deutlich erschwert. Um dies sicherzustellen, könnte die Führungskraft z.B. ihr Informations- und Kommunikationsverhalten reflektieren und sich der Wirkung einer am Wochenende oder am späten Abend versendeten E-Mail bewusstwerden. Auch symbolische Handlungen und Rituale können die gesundheitsorientierte Selbstführung unterstreichen. Was spricht dagegen, dass sich die Führungskraft auch mediengestützt in den Feierabend verabschiedet oder die Mitarbeiter*innen im Homeoffice morgens bei Dienstantritt persönlich begrüßt?

Der eingeschränkte direkte Kontakt mit den Beschäftigten stellt eine weitere Herausforderung in der virtuellen Zusammenarbeit dar. Führungskräfte haben aufgrund der Distanz nur wenig Einblick in die (Arbeits-) Situation der Mitarbeiter*innen im Homeoffice. Wie ist es um die Arbeitsumgebung bestellt? Wie hoch ist die Gesamt-Arbeitsbelastung des Mitarbeiters / der Mitarbeiterin vor Ort? Gibt es Probleme, private Anforderungen mit der Arbeit zeitlich zu verbinden? Sind neue Aufgaben mit hoher Priorität hinzugekommen? Regelmäßige Mitarbeiter*innengespräche und Feedbacks seitens der Führungskraft bieten den Beschäftigten die Möglichkeit, über Ängste und Sorgen zu sprechen und ihr Beanspruchungserleben bei der Heimarbeit zu thematisieren. Führungskräfte können ihre Mitarbeiter*innen unterstützen, erforderliche Qualifikationen zu erwerben, um die Anforderungen zu bewältigen, z.B. durch online-basierte Seminare und Trainings zum Stress- und Selbstmanagement.

Die geringere Möglichkeit zum sozialen Austausch der Kolleg*innen untereinander kann ebenfalls als belastend empfunden werden (Risiko der sozialen Isolation). Daher ist es für die Zusammenarbeit wichtig, unterstützende soziale Beziehungen aufrechtzuerhalten und sich z.B. über Videokonferenzen auch informell, nicht aufgabenbezogen auszutauschen. Führungskräfte können solche Gesprächsformate anregen und fördern, indem sie zeitliche Ressourcen zur Verfügung stellen.

Dem Aspekt der Arbeitsgestaltung (indirekter Einfluss) kommt in digitalen Kontexten eine besondere Bedeutung zu. In erster Linie bezieht sich dieser Faktor auf die Bereitstellung adäquater digitaler Informations- und Kommunikationstechnologien (IKT). Welche IKT letztendlich eingeführt wird, ist immer in Abhängigkeit von den Anforderungen der Arbeitsaufgabe und den Kompetenzen der Beschäftigten zu sehen. Die Nutzung von IKT zur Zusammenarbeit und Kommunikation kann allerdings auch leichter zu Unsicherheiten und Missverständnissen führen. Der Wahl des Kommunikationsmittels kommt somit entscheidende Bedeutung zu. So empfiehlt sich der Einsatz eines audiovisuellen Mediums in unklaren oder komplexen Situationen.

Die Ausführungen im ersten Teil des Artikels machen deutlich, dass neben den möglichen positiven Effekten der Arbeit im Homeoffice auch Risiken bzw. Herausforderungen existieren. Die größere zeitliche und räumliche Flexibilität in Verbindung mit IKT, die ständige Erreichbarkeit und die ständige Möglichkeit zu arbeiten kann zu längeren Arbeitszeiten führen und zu der wahrgenommenen Anforderung, auch außerhalb der Arbeitszeit E-Mails beantworten zu müssen oder erreichbar zu sein. Gesundheitsorientierte Führungskräfte können hier unterstützen, indem sie die Bedeutung von Erholungs- und Regenerationszeiten thematisieren und auch konkrete Maßnahmen einbeziehen, wie z.B. festgelegte Zeiten der Nicht-Erreichbarkeit, abgestimmte Antwortzeiten für den E-Mailverkehr oder Regeln für die Priorisierung (weitere Aspekte zum Thema „Entgrenzung“ wurden von Autor Kaschenbach im zweiten Teil des Artikels aufgezeigt).

Kooperationserfordernisse im Team lassen sich z.B. über die Gestaltung der Arbeitsaufgaben erzeugen. Die Schaffung von Interdependenzen (z.B. Ziel-, Aufgaben- und Ergebnisabhängigkeiten) und Feedbackprozessen bei Erreichen von (Teil-)Zielen fördern die Identifikation als Team und das Vertrauen zueinander. Die Führungskraft nimmt dabei den Beschäftigten gegenüber keine kontrollierende Rolle ein, sondern sorgt im Team für Kooperation und die Koordination der jeweiligen Ziele. Sie sollte den Austausch darüber fördern, wie man sich im Team gegenseitig unterstützen kann und welche Rahmenbedingungen als sinnvoll für die gemeinsame gesunde Zielerreichung erachtet werden.

Arbeit im Homeoffice schafft zusätzliche Autonomie, Handlungsspielräume und Zeitsouveränität.

Freiräume, die sich hieraus ergeben, führen zu einer erhöhten Motivation und haben nachweisliche positive gesundheitliche Konsequenzen. Führungskräfte sind angehalten, Flexibilität und Freiheitsgrade durch Regeln und Vorschriften nicht zu stark einzuschränken, sondern die Selbstbestimmung der Mitarbeiter*innen durch Auflösen der zentralen Führungsverantwortung zu erhöhen.

Insgesamt lässt sich festhalten, dass bei digitaler Zusammenarbeit viele zentrale Führungsaufgaben auch weiterhin Gültigkeit besitzen. Die zunehmend digital vermittelte Einflussnahme stellt Führungskräfte vor neue Herausforderungen und birgt die Notwendigkeit von Anpassungsleistungen in den Führungspraktiken. Um eine gesundheitsorientierte Führung in digitaler Zusammenarbeit zu gewährleisten, sollten Führungskräfte zum Thema „Führung auf Distanz“ geschult werden, ein Training, welches z.B. auch die dgp als Online-Seminar anbietet.

.

Literatur:

Ashforth, B., Kreiner, G. & Fugate, M. (2000). All in a day's work: boundaries and micro role transitions. Academy of Manager Revue 25(3):472–491.

Bundesministerium für Forschung und Technik (BMFT). (1997). Definition der Gesund-heit. http://de.wikipedia.org/wiki/Gesundheit.

DGB-Index Gute Arbeit (2020). Mehr als Homeoffice – Mobile Arbeit in Deutschland. Institut DGB-Index Gute Arbeit, Berlin.

Efimov, I., Harth, V., Mache, S. (2020). Gesundheitsförderung in virtueller Teamarbeit durch gesundheitsorientierte Führung. In: Prävention und Gesundheitsförderung. Springer

Erwerbstätigenpersonenbefragung der Hans-Böckler-Stiftung (2020).

Franke, F., Felfe, J. (2011). Diagnose gesundheitsförderlicher Führung – Das Instrument „Health oriented Leadership“. In: Badura, B., Ducki, A., Schröder, H., Klose, J., Macco, K., (Hrsg). Fehlzeitenreport 2011: Führung und Gesundheit: Zahlen Daten, Analysen aus allen Branchen der Wirtschaft. Berlin, Heidelberg: Springer.

Gisin L., Schulze H., Degenhardt B. (2016) Boundary Management as a Crucial Success Factor for Flexible-Mobile Work, Demonstrated in the Case of Home Office. In: Deml, B., Stock, P., Bruder, R. & Schlick, C.M. (eds) Advances in Ergo-nomic Design of Systems, Products and Processes. Springer Vieweg, Berlin, Heidelberg.

Hofmann, J., Piele, A. & Piele, C. (2021). Arbeiten in der Corona-Pandemie. Entgrenzungseffekte durch mobiles Arbeiten – und deren Vermeidung. Studie des Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO.

Kreiner G., Hollensbe E., Sheep, M. (2009) Balancing borders and bridges: negotiating the work-home interface via boundary work tactics. Academy of Management Journal 52(4):704–730.

Lorenz, I. & Grünhagen-Scheele, A. (2019). Beziehungsgestaltung in der digitalen Arbeitswelt. In: dgp Informationen 2019.

Marschall, J., Hildebrandt, S., Kleinlercher, K. & Nolting, H. (2020). Gesundheitsreport 2020. Stress in der modernen Arbeitswelt. In: A. Storm. Beiträge zur Gesundheitsökonomie und Versorgungsforschung (33). DAK-Gesundheit, Hamburg.

Nippert-Eng, E. (1996) Home and work: negotiating boundaries through everyday life. University of Chicago Press, Chicago.

Samtleben, C., Lott, Y. & Müller, K. (2020). Auswirkungen der Ort-Zeit-Flexibilisierung von Erwerbsarbeit auf informelle Sorgearbeit im Zuge der Digitalisierung. In: Expertise für den Dritten Gleichstellungsbericht der Bundesregierung.

Staar, H., Gurt, J., Janneck, M. (2019). Gesunde Führung in vernetzter (Zusammen-)Arbeit – Herausforderungen und Chancen. In: Badura, B., Ducki, A., Schröder, H., Klose, J., Meyer, M., editors. Fehlzeitenreport 2019: Digitalisierung – gesundes Arbeiten ermöglichen. Berlin, Heidelberg: Springer.

Kontakt:

Lucas Maximilian Kaschenbach
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
kaschenbach@dgp.de

Stefan Riechmann
Dipl.-Psychologe
Deutsche Gesellschaft für Personalwesen e.V.
riechmann@dgp.de

Kira Schubert
M.Sc. Wirtschaftspsychologie
Deutsche Gesellschaft für Personalwesen e.V.
schubert@dgp.de

Führen auf Distanz – (wie) gelingt das?

Michael Jaeger & Maximilian Dicker

Mit Ausbruch der Corona-Pandemie haben mobile Arbeitsformen – insbesondere das Arbeiten im Homeoffice, rasant an Verbreitung gewonnen und werden mit hoher Wahrscheinlichkeit auch dauerhaft zu festen Bestandteilen der Arbeitswelt werden. An Führungskräfte wie Mitarbeiter*innen stellt sich damit die Anforderung, die Zusammenarbeit über räumliche Distanz erfolgreich zu bewältigen. Zwar helfen hierbei digitale Formate wie z.B. Videokonferenztools. Trotzdem wirken sich die Distanz und die über Medien vermittelte Kommunikation auf das Miteinander zwischen Führungskräften und Mitarbeiter*innen wie auch auf die Zusammenarbeit im Team selbst aus. Welche Risikofaktoren bestehen dabei, und wie gelingt Führung auch unter Bedingungen der räumlichen Entfernung? Und wie können Organisationen ihre Führungskräfte und Mitarbeiter*innen dabei unterstützen?

Merkmale und Herausforderungen der räumlichen Distanz

Die Anforderung, Mitarbeiter*innen über Distanz zu führen oder in räumlich verteilten Teams zusammenzuarbeiten, ist zunächst nichts Neues und gehört in der Privatwirtschaft, im Zuge international vernetzter Produktion und Logistik sowie globaler Absatzmärkte, zu den selbstverständlichen Arbeitsweisen. Auch in den öffentlichen Verwaltungen waren Arbeitsformen wie z.B. Telearbeit bereits vor der Corona-Pandemie vielerorts etabliert. Die Rolle der Pandemie ist vor diesem Hintergrund, wie auch mit Blick auf andere gesellschaftliche Teilbereiche, daher in erster Linie als die eines Treibers und Verstärkers bereits vorhandener Trends und Entwicklungen zu sehen. Mobiles und dezentrales Arbeiten hat demgemäß inzwischen auch in Branchen und Organisationen Raum gegriffen, in denen zuvor noch eine Skepsis gegenüber entsprechenden Arbeitsformen bestand.

Grundsätzlich kann sich Führung über Distanz auf unterschiedliche Formen der Zusammenarbeit beziehen. Remdisch und Utsch (2006) definieren Führen über Distanz vor diesem Hintergrund als Führungsverantwortung und -tätigkeit unter der Bedingung räumlicher Entfernung, sozialer und/oder kultureller Diversität, mediengestützter Kommunikation oder projekt- und netzwerkförmiger Organisationsformen. Im vorliegenden Beitrag greifen wir aus den hier in Betracht kommenden Konstellationen die Führung vom Mitarbeiter*innen in und aus dem Homeoffice auf.

Mit der Etablierung der Arbeit im Homeoffice verbinden sich zunächst einmal viele Chancen.

Arbeitnehmerseitig zählen hierzu beispielsweise eine bessere Vereinbarkeit von Beruf und Privatleben und die Einsparung von Wegezeiten, womit auch in ökologischer Hinsicht – Stichwort: Reduzierung des CO2-Fußabdrucks – Vorteile erschlossen werden können. Arbeitgeberseitig wird im Angebot solcher Arbeitsformen vielfach die Möglichkeit gesehen, die eigene Attraktivität für Fachkräfte zu steigern – in Zeiten des Fachkräftemangels ein wichtiges Argument. Gerade auch für die Nachwuchskräftegewinnung im öffentlichen Sektor erhält dieser Faktor mehr Bedeutung: Laut Nachwuchsbarometer Öffentlicher Dienst 2019 ist jedem/jeder vierten Studienteilnehmer*in mobiles Arbeiten (Homeoffice) sehr wichtig, während nur jede*r siebte diesen Punkt in der Verwaltung als gut umgesetzt erachtet. Zudem ist speziell für die vielfach mit eklatanter Raumnot kämpfenden öffentlichen Verwaltungen auch die Möglichkeit interessant, Einspareffekte im Bereich räumlicher Ressourcen erzielen zu können. Viele Dienstvereinbarungen zum Arbeiten vom heimischen Arbeitsplatz aus sehen daher bei der Gewährung dieser Arbeitsform eine Mindestzahl von Wochentagen vor, die dann im Homeoffice zu erbringen ist.

Auf der anderen Seite bringt die Zusammenarbeit im und aus dem Homeoffice heraus auch Risiken mit sich. Boos, Hardwig und Riethmüller (2017) arbeiten hier insbesondere die folgenden heraus:

- Die Intensität der sozialen Interaktion wird durch die Distanzsituation deutlich eingeschränkt, dies gilt sowohl auf Ebene der Zusammenarbeit zwischen Führungskraft und Mitarbeiter*innen wie auch im Team insgesamt. Vor allem das Wegfallen informeller Formate wie etwa dem Plausch in der Teeküche oder auf dem Flur kann sich auf Ebene von Bindung und Zusammenhalt im Team erschwerend auswirken.
- Eine unmittelbare Kommunikation „face to face“, wie sie in der Präsenzumgebung selbstverständlich ist, kann in bzw. aus dem Homeoffice heraus nicht stattfinden. Kommunikativer Austausch mit Vorgesetzten und Kolleg*innen bedarf hier stets des Einsatzes von Medien, wie bspw. in Form von Telefonkonferenzen und virtuellen Formaten. Im Vergleich zur Face-to-Face-Kommunikation bringt dies aber immer einen Informationsverlust mit sich.

So ist bspw. für die Empfänger*innen einer E-Mail nicht unbedingt ersichtlich, wie eine bestimmte Mitteilung gemeint ist oder in welchem situativen Kontext sie verfasst wurde – Missverständnisse und Fehlinterpretationen können die Folge sein. Auch für den oder die Absender*in einer E-Mail ist nicht unmittelbar sichtbar, wie eine bestimmte Botschaft beim Gegenüber ankommt, zumal auch die Möglichkeit der unmittelbaren Rückkopplung – z.B. durch Verständnisfragen – erschwert ist.

Demgegenüber beinhalten videogestützte Formen der Zusammenarbeit zwar nicht die eben beschriebenen Risiken der zeitversetzten und kontextbereinigten Kommunikation. Auch hier gehen aber wesentliche Anteile insbesondere auf Ebene der nonverbalen Kommunikation – Mimik, Gestik, Haltung – verloren, die für den Aufbau zwischenmenschlicher Beziehungen ausschlaggebende Bedeutung haben. Zudem wird die videobasierte Zusammenarbeit von den Teilnehmenden, gerade bei größeren Teilnehmer*innenzahlen und über längere Zeiträume hinweg, oft als anstrengend empfunden und kann zu Lasten der Aufmerksamkeitsfokussierung gehen.

Speziell im Zusammenhang mit der Homeoffice-Situation ist auf ein weiteres Risiko hinzuweisen: Viele im Homeoffice tätige Personen berichten, dass es ihnen in der Heimarbeitssituation schwerfällt, die berufliche und private Sphäre klar voneinander abzugrenzen. In der Folge kann es zu erlebten Beeinträchtigungen wie z.B. Erschöpfung, Konzentrationsproblemen oder Schlafstörungen kommen (vgl. z.B. AOK/WIdO 2019).

Was folgt daraus für Führungskräfte?

Mit Blick auf die Konsequenzen für das Führungshandeln ist zunächst eine Implikation unmittelbar offensichtlich:

Führung durch kontinuierliche Verhaltensbeobachtung von Mitarbeiter*innen am Arbeitsplatz und damit einhergehende Kontrolle kann, wenn überhaupt, nur noch partiell praktiziert werden.

An die Stelle der „Kontrolle über Anwesenheit“ muss demzufolge die Vereinbarung und Definition klarer Ziele und Aufgaben stehen, deren Erreichung dann Gegenstand der Überprüfung und Kontrolle durch die Führungskraft ist. Dies ist, im Sinne zeitgemäßer und partizipativer Mitarbeiterführung, zwar nicht Neues – neu hingegen ist, dass anwesenheitsbasierte Formen der Mitarbeiter*innenführung als alternative Führungsansätze bei räumlicher Distanz nur noch bedingt zur Verfügung stehen.

Vier weitere Aspekte erweisen sich als wesentliche Erfolgsfaktoren bei der Führung und Zusammenarbeit über Distanz:

- Die Abkehr von dem Ansatz, Führung über Beobachtung und (Anwesenheits-) Kontrolle zu praktizieren, erweist sich für viele Führungskräfte auch deswegen als Herausforderung, weil hiermit ein – zu-

mindest erlebter – Kontrollverlust verbunden ist. Vielen Führungskräften fällt es schwer, anstatt dessen Leistungsziele und Aufgaben zu definieren, die dann Gegenstand einer (späteren) Überprüfung der Zielerreichung sind. Damit wird deutlich, dass das Vertrauen zwischen Führungskraft und Mitarbeiter*in, und hier vor allem auch das Vertrauen der Führungskraft in die Mitarbeiter*innen, eine Schlüsseldimension beim Führen über Distanz ist: Erfolgreiche Führung über Distanz setzt bei der Führungskraft ein Grundvertrauen voraus, dass Mitarbeiter*innen motiviert sind, ihren Arbeitsaufgaben nachzugehen – auch wenn sie nicht wie in der Präsenzkultur jederzeit beobachtbar sind. Und für die Gestaltung der Beziehung zwischen Führungskraft und Mitarbeiter*innen erwächst daraus die Frage: Wie kann ich als Führungskraft das Vertrauen meiner Mitarbeiter*innen konkret fördern?

- Das Gegenstück zu dem „Vertrauensvorschuss“ muss die Etablierung eines klaren Rahmens der Zusammenarbeit sein, in dem neben Zielen und Aufgaben auch die Art und Weise der Zusammenarbeit mit im Homeoffice befindlichen Mitarbeiter*innen sowie die Überprüfung von Zielerreichung und Aufgabenerledigung als selbstverständliche Elemente festgelegt werden.

- Der in der Distanzsituation deutlich höhere Anteil medienvermittelter Kommunikation mit den damit einhergehenden Risiken macht es drittens erforderlich, den Einsatz von Medien in der Mitarbeiter*innenführung sorgsam abzuwägen und auch gezielt zu planen. Besprechungen in Präsenz werden in der Distanzsituation zu einer kostbaren Ressource, die führungsseitig auch gezielt für solche Anlässe eingesetzt werden sollte, die Kontext und unmittelbare Rückkopplungsmöglichkeit erfordern. Beispiele hierfür sind Mitarbeiter*innengespräche und Feedbacksituationen. Anlässe, für die das weniger notwendig ist und die z.B. eher der Informationsübermittlung dienen (z.B. Bekanntgabe von Terminen, Hinweise zu Prozessabläufen) eignen sich dann besser z.B. für die Übermittlung per E-Mail. Grundsätzlich gilt: Je komplexer die Kommunikationsaufgabe, desto differenzierter sollte das eingesetzte Medium sein (vgl. Abbildung).

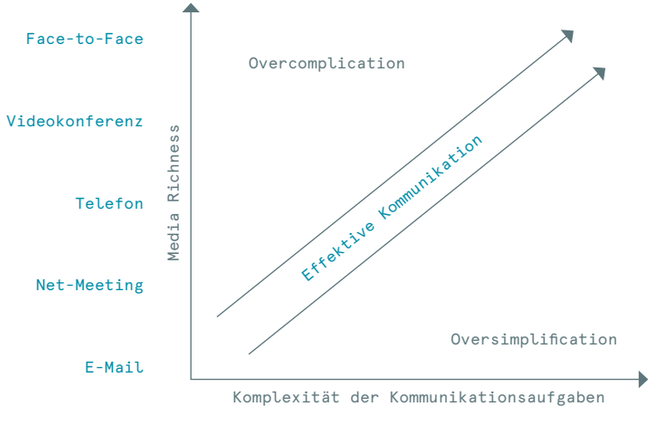


Abbildung: Media Richness Modell, nach Hülsbusch, Utsch, Remdisch & Groß, 2006

- Der vierte Aspekt bezieht sich auf die Ebene des Teams. Aufgrund der Schwächung des sozialen Zusammenhalts durch räumliche Distanz wird die Förderung des Teamzusammenhalts mehr denn je zu einer genuinen Führungsaufgabe. Auf der sachlichen Ebene kann die Erarbeitung und Definition gemeinsamer Teamziele hierzu beitragen. Aber auch die Beziehungsebene selbst bedarf einer Aufmerksamkeit und aktiven Gestaltung. Mit Blick auf die klassischen Phasen einer Teamentwicklung ist hier das Forming besonders wichtig, d.h. die Phase, in der ein Team sich zusammenfindet und sich Strukturen der Kooperation aufbauen. Die fortlaufende Stimmungs- und Vertrauenspflege, z.B. durch regelmäßige Zeiten für gemeinsame Reflexionsphasen, und die gezielte Förderung informeller Austauschformate können hierzu einen wichtigen Beitrag leisten. Das gleiche gilt für die Absprache von Regeln, nach denen die Kommunikation untereinander erfolgen soll.

Als generelles Fazit lässt sich also ziehen:

Vertrauen Sie Ihren Mitarbeiter*innen , beteiligen Sie diese aktiv an der Definition von Teamzielen und Arbeitsaufgaben und schaffen Sie einen klaren Rahmen der Zusammenarbeit, in dem auch die Überprüfung des Erreichens von Arbeitszielen und -aufgaben ihren Platz hat.

Entwickeln Sie Antennen für die Artikulation von Bedürfnissen durch Mitarbeiter*innen, gerade auch in medienvermittelten Kommunikationsformen, um einen Blick dafür zu bekommen, ob die häusliche bzw. mobile Arbeitssituation für den oder die Mitarbeiter*in gut funktioniert. Damit haben Sie die entscheidenden Voraussetzungen dafür im Blick, dass die Zusammenarbeit über Distanz gut gelingt und die Potenziale dieser Arbeitsform sowohl im Sinne der Mitarbeiter*innen als auch der Organisation erschlossen werden können.

Wie können Organisationen ihre Führungskräfte und Mitarbeiter*innen beim Überwinden der räumlichen Distanz unterstützen?

Abschließend wechseln wir noch einmal die Betrachtungsebene und gehen der Frage nach, wie Organisationen ihre Führungskräfte und Mitarbeiter*innen bei der Zusammenarbeit über räumliche Distanz unterstützen können. Hierbei greifen wir die Ergebnisse von Klausurtagungen und Workshops zu diesem Thema auf, die wir als dgp in unterschiedlichen Kontexten begleiten und bei denen seitens der Führungskräfte häufig Fragen wie etwa die folgenden auftauchen: Wie sieht eine zeitgemäße Führungskultur aus, die auch für die Zusammenarbeit über räumliche Distanz Orientierung gibt und Standards setzt? Wie sorgt man für genügend Kontakt und Zusammenhalt in der Belegschaft und in den einzelnen Teams? Und ist Homeoffice möglicherweise unfair gegenüber bestimmten Beschäftigten(gruppen)?

Organisationen können ihre Führungskräfte bei genau diesen Fragen unterstützen, indem sie klare Rahmenbedingungen fürs Homeoffice und damit auch Sicherheit schaffen, idealerweise im gemeinsamen Austausch. Den institutionellen Rahmen hierfür bildet dann der Abschluss einer Dienstvereinbarung, die bspw. Regelungen dazu enthält, an welche Voraussetzungen – aufgaben- wie auch kompetenzseitiger Natur – die Arbeit vom heimischen Arbeitsplatz aus gebunden ist, wie diese überprüft

werden und nach welchen Standards die Zusammenarbeit zwischen Führungskräften und Mitarbeiter*innen organisiert wird. Dies bietet Führungskräften und Beschäftigten einen sicheren Rahmen, in dem Homeoffice stattfinden kann, entlastet jedoch freilich nicht von der Betrachtung der individuellen Fallentscheidung.

Zusätzlich zu einem verbindlichen Regelwerk zum Umgang mit Homeoffice sollten Organisationen laut Hermann, Hünecke und Rohrberg (2013) auch gezielt einen Rahmen für virtuelle Teamarbeit schaffen. Hierzu gehört neben der Beförderung eines Commitments bei den Führungskräften für entsprechende Formen der Zusammenarbeit (z.B. mit gutem Beispiel voranzugehen, neuen Medien offen gegenüberstehen) insbesondere die Berücksichtigung des veränderten Ressourcenbedarfs von räumlich getrennten Teams.

Es müssen eine tragfähige IT-Infrastruktur aufgebaut und Prozesse entwickelt werden, welche die Zusammenarbeit über räumliche Distanz hinweg unterstützen. Dies geht über den Einkauf von Software für Videokonferenzen und Webcams hinaus, da auch zentrale Abläufe auf den Prüfstand gestellt werden müssen (z.B. Wie findet Bürger*innenkontakt statt? Wie kommen postalische Schreiben und Akten zu den Mitarbeiter*innen ins Homeoffice? Wie kommen Bescheide und andere Drucksachen aus dem Homeoffice zu den Empfänger*innen?).

Schließlich ist, neben der Schaffung klarer institutioneller Rahmenbedingungen, die Beförderung des Kompetenzaufbaus bei Führungskräften und Mitarbeiter*innen ein weiteres zentrales Gestaltungsfeld für Organisationen.

Dies bezieht sich nicht nur im engeren Sinne auf die Frage, wie und mit welchen Instrumenten Führung und Kommunikation über die räumliche Distanz gelingt, sondern darüber hinaus auf die Entwicklung der Kompetenzen, die benötigt werden, um z.B. mit digitalen Medien umzugehen und sich auf das hohe Tempo einzustellen, in dem sich auch in den öffentlichen Verwaltungen die – ganz unabhängig von der Corona-Pandemie – Arbeitswelt verändert.

Abschließend lässt sich festhalten: Organisationen sollten sich mit Blick auf Themen wie Führung und Zusammenarbeit über räumliche Distanz und Gestaltung der digitalisierten Arbeitswelt nicht von den (technologischen) Entwicklungen treiben lassen. Vielmehr sollten sie das Heft des Handelns in der Hand behalten und selbst Strategien und Leitsätze entwickeln, wie sie Führung und Zusammenarbeit in der sich verändernden Arbeitswelt gestalten wollen. Mehr denn je werden Organisations- und Personalentwicklung damit zu zentralen Gestaltungsfeldern, auch im Bereich öffentlicher Organisationen.

Literatur:

Boos, M., Hardwig, Th., & Riethmüller, M. (2017): Führung und Zusammenarbeit in verteilten Teams. Göttingen: Hogrefe.

Herrmann, D., Hünecke, K., & Rohrberg, A. (2012). Führung auf Distanz. Wiesbaden: Springer Gabler.

Hülsbusch, W., Utsch, A., Remdisch, S. & Groß, M. (2006). Führen auf Distanz. Vom E-Leading zum Blended Leading, Wissensmanagement, 01/2006, 40-42.

Next:Public. (2019). Nachwuchsbarometer Öffentlicher Dienst 2019. Internet: https://www.nachwuchsbarometer-oeffentlicher-dienst.de/wp-content/uploads/2019/06/Inhaltsverzeichnis_verlinkt_Nachwuchsbarometer_Oeffentlicher_Dienst_2019.pdf

Remdisch, S. & Utsch, A. (2006). Herausforderungen der Führung auf Distanz. Organisationsentwicklung: Zeitschrift für Unternehmensentwicklung und Change Management, 25 (3), 32–43

WIdO (Wissenschaftliches Institut der AOK). (2019). Fehlzeitenreport 2019 (Pressemappe). Internet: https://www.wido.de/news-events/aktuelles/2019/fehlzeiten-report-2019/.

Kontakt:

Dr. Michael Jaeger
Diplom-Psychologe und Leiter der
dgp-Geschäftsstelle Hannover
Deutsche Gesellschaft für Personalwesen e.V.
jaeger@dgp.de

Maximilian Dicker
M.Sc. Psychologie
Deutsche Gesellschaft für Personalwesen e.V.
dicker@dgp.de

Live-Online-Seminare bei der dgp – ein Plädoyer für ein neues Format

Beate Scholz

„Genauso interaktiv wie im „Live-Seminar!“

„Auch online ist eine abwechslungsreiche Methodik möglich!“

„Für mich das erste Online-Seminar: es war lebhaft, lustig und aufschlussreich!“

Online-Trainingsformate haben sich in zahlreichen Unternehmen längst etabliert, im öffentlichen Dienst waren sie dagegen bis 2020 eher selten nachgefragt. Nachdem aber – spätestens mit Eintreffen der zweiten Welle der Corona-Pandemie – Präsenztrainings aufgrund der Kontaktbeschränkungen für die nächsten Monate nicht oder nur sehr eingeschränkt stattfinden konnten, hat auch hier ein Wandel eingesetzt. Nach eher abwartender Haltung bei vielen Kund*innen, ob,

1. dgp-Seminare mit derselben Qualität auch online durchgeführt werden könnten,
2. die Ziele, Inhalte und Methoden eins zu eins ins digitale Medium übersetzbar seien und
3. die Umstellung/Entwicklung eines neuen Formats für die (kurze) Corona-Zeit überhaupt sinnvoll sei, ...

und auch Zurückhaltung bei den Trainer*innen, die sich in Reaktionen wie

- „Grundsätzlich eine gute Idee, aber dieses spezielle Thema funktioniert nur als Präsenz-Schulung."
- „Wie soll man denn die vielen Inhalte und Übungen virtuell meistern?"
- „Die Teilnehmer*innen bleiben nicht am Ball."
- „Und mit welchem Tool wollen wir arbeiten?"

äußerten, häuften sich Anfragen wie: „Wir hatten doch das Seminar xyz vereinbart. Können wir das auch online stattfinden lassen?"

So war es für die dgp nur logisch, in Zeiten größtmöglicher Kontaktreduktion und zunehmender Digitalisierung zusätzlich zu digitalen Auswahlformaten auch das Portfolio der Personalentwicklungsmaßnahmen um Online-Formate zu erweitern und die ersten Präsenzseminare in Live-Online-Trainings zu transformieren.

Die dgp bietet seit 2020 zusätzlich zu Präsenz-Seminaren Online-Seminare, Webinare und auch Hybridveranstaltungen an. Nach ersten Erfahrungen, z.B. bei der Vereinbarung von Online-Veranstaltungen, wurde deutlich, dass die Absprache von Online-Formaten in der Fortbildung einer genauen Begriffsbestimmung bedarf. Wann handelt es sich z.B. um ein Webinar, ein Online-Seminar oder um ein Hybridseminar? Dies ist wichtig, um weder bei den Kund*innen noch dgp-seitig falsche Erwartungen zu wecken. Den Präsenzveranstaltungen sind Online-Seminare am ähnlichsten. Zur Verdeutlichung folgen hier kurze Begriffsbestimmungen:

Online-Seminare

Ein Online-Seminar ist eine interaktive Lehrveranstaltung, in der ein*e Trainer*in und eine überschaubare Gruppe Teilnehmer*innen im virtuellen Raum aufeinandertreffen. Trainer*in und Teilnehmer*innen verfügen über Kamera und Mikrofon. Hier werden sowohl Wissen vermittelt als auch Kompetenzen bei den Teilnehmer*innen aufgebaut. Die Teilnehmer*innen gestalten den Lernprozess aktiv mit, indem sie Fragen stellen, miteinander diskutieren und interagieren sowie sich ausprobieren können. Den Präsenzveranstaltungen sind Online-Seminare am ähnlichsten.

Webinare

Ein Webinar ist eine Präsentation im virtuellen Raum. Die Teilnehmer*innen benötigen nicht unbedingt Kamera und Mikrofon. Im Chat können Fragen gestellt werden. Webinare vermitteln Wissen, können informieren und überzeugen, aber bauen in der Regel keine Kompetenzen auf.

Hybridseminare

Ein Hybridseminar ist eine Kombination aus Präsenz- und Onlineseminar, indem beispielsweise einem Präsenzseminar eine Follow-Up-Veranstaltung im Online-Format folgt. Durch die Kombination bieten Hybridseminare ein effizientes Gesamtkonzept, da die Vorteile beider Veranstaltungsarten genutzt werden.

Ablauf von Online-Seminaren

Um ein klares Bild zu erhalten, was ein Live-Online-Seminar an Möglichkeiten tatsächlich bietet, wie wirksam es ist, wie viel Spaß die Interaktion im virtuellen Seminarraum machen kann und welche Vorteile das für die Teilnehmer*innen, die Organisation und natürlich auch den/die Trainer*in bringt, erfolgt an dieser Stelle die Beschreibung eines möglichen Ablaufs von der Klärung technischer Voraussetzungen bis hin zur Evaluation. Die inhaltliche Klärung der Veranstaltung (sprich: Seminarthema, Zielgruppe...) bleibt an dieser Stelle außen vor.

Klärung technischer Voraussetzungen

Um einen technisch reibungslosen Ablauf zu gewährleisten ist es bei der Auftragsklärung sehr wichtig, genaue Absprachen über technische Voraussetzungen (stabiles Internet, Videoplattform [Browser/Datenschutz], sind Kamera, Mikro/Lautsprecher vorhanden) zu treffen (vergleichbar der „Raumausstattung" in Präsenzveranstaltungen).

Die Erfahrung hat gezeigt, dass ein optimaler Ablauf inhaltlicher und technischer Natur ganz erheblich von der Gruppengröße abhängig ist. Je nach Thema bieten sich Gruppenstärken von 6-10, in Ausnahmen auch bis zu 12 Teilnehmer*innen an. Bei dieser Personenzahl passen alle gut auf den Bildschirm, und das Internet ist (meist) noch stabil.

Legt man die oben genannten Begriffsbestimmungen zu Grunde, ist klar, dass für interaktive Online-Seminare sowohl auf Trainer*innen- als auch auf Teilnehmer*innen-Seite eine Kamera sowie Mikrofon und Lautsprecher zwingend erforderlich sind. Um als Trainer*in nicht nur mit Initialen oder Standbildern der Teilnehmer*innen zu kommunizieren – wie anfangs zeitweise geschehen – ist dieser Hinweis bei der Auftragsabsprache sehr wichtig.

Ebenso wichtig ist auch die Klärung, mit welcher Videoplattform gearbeitet wird und wer sie zur Verfügung stellt, sprich ist die dgp „Host" bzw. „Gastgeber*in" und lädt infolgedessen zum Seminar ein oder übernimmt dies der/die Auftraggeber*in? Bei der Vielzahl der Videokonferenztools steht immer wieder die Frage im Raum: Welches ist das beste? Wie immer ist die Antwort abhängig von der jeweiligen Nutzungsabsicht. Die dgp hat bislang gute Erfahrungen mit Zoom und Webex gemacht und bietet ihrerseits zusätzlich auch noch MS Teams an. Andere Videokonferenztools erfordern z.T. erhebliche Einarbeitung, bieten keine oder nur wenige interaktive/kollaborative Elemente und/oder bereiten browserabhängig oft Übertragungsprobleme. Auf dem Markt ist erfreulicherweise eine stetige Produktweiterentwicklung und Anpassung an Erfordernisse des Weiterbildungsalltags zu beobachten.

„Vielfältiger als gedacht im Hinblick auf digitale Fortbildung!"

Länge/Dauer der Online-Seminare

Die Praxis der letzten Monate hat gezeigt, dass interaktive Online-Seminare mit einer Gesamtdauer von 4-6,5 Stunden mit kurzen Pausen ca. alle 75 Minuten und einer längeren Mittagspause ohne merklichen Konzentrationsverlust gut durchführbar sind.

Einladung „Onboarding“ der Teilnehmer*innen

Zusammen mit dem Link zum Online-Seminar, sollten den Teilnehmer*innen alle notwendigen Informationen gemailt werden (Zeiten/Pausen und Information, was im „Notfall“ zu tun ist), um allen Teilnehmer*innen die Teilnahme so leicht wie möglich zu machen. Auch wenn viele schon an Teamsitzungen, Schul-Elternabenden oder Jahreshauptversammlungen über Zoom oder Skype teilgenommen haben oder sogar ihren Yoga-Kurs auf einer Videoplattform besuchen, zeigt die Erfahrung, dass die Barriere/Hemmung, beruflich an Online-Seminaren teilzunehmen, bei einigen noch groß ist – z. B. aus Angst, es technisch nicht zu schaffen. Diese Barrieren gilt es, von Anfang an abzubauen.

„Selten so eine Wohlfühlatmosphäre gehabt, trotz digital!“

Die Durchführung

Bei den ersten Online-Seminaren nutzten viele Teilnehmer*innen das Angebot, sich bereits 30-20 Minuten vor der Veranstaltung zwecks technischen Onboardings einzufinden. Mittlerweile erscheinen die meisten Teilnehmer*innen „just in time“.

Wie im Präsenzseminar beginnt das Abenteuer „Online-Seminar“ bereits, wenn die Teilnehmer*innen den Raum betreten, aber anders als in der Präsenzveranstaltung kommt es neben dem Begrüßungs-Smalltalk ggf. noch zu einem kurzen individuellen Technik-Check: Sind alle sicht- und hörbar?

Gleich zu Beginn werden den Teilnehmer*innen Hinweise zum zeitlichen Ablauf gegeben (besonders wichtig für Teilnehmer*innen, die parallel z.B. Kinder im Homeschooling betreuen mussten) und die grundlegenden Funktionen des Videokonferenztools erklärt, um eventuell bestehende Bedienungs-Unsicherheiten auszuräumen. Dies kommt praktisch einem Rundgang durch den Seminarraum gleich. An dieser Stelle werden die Funktionen und die Regeln z.B. zum An-/Abschalten von Kamera und Mikro, der Umgang mit den Reaktionszeichen, dem Chat und der Möglichkeit, den eigenen Bildschirm zu teilen, erklärt.

Didaktisch erfordert die Transformation der Präsenzveranstaltungen in Live-Online-Seminare in mehrerlei Hinsicht ein Umdenken und Umlernen.

1. Es ist bekannt, dass im virtuellen Seminarraum bereits Monologe von mehr als vier Minuten dazu führen, dass Teilnehmer*innen abschweifen, Mails checken, andere Programme öffnen, aufs Smartphone schauen etc. Während sich in Präsenzseminaren das Gros der Teilnehmer*innen an bestimmte „Benimm-Regeln“ halten und die Trainingsatmosphäre sowie ein gewisser Gruppendruck dazu beitragen, sich auch bei einsetzender Langeweile zu beteiligen, liegt beim Live-Online-Training die Hemmschwelle für Multi-Tasking deutlich niedriger. Wer nicht eingebunden ist und sich unbeobachtet fühlt, macht schnell etwas anderes.

2. Es besteht die Gefahr, dass keine Stimmung aufkommt. In Präsenz-Seminaren lernt man sowohl Trainer*in als auch Teilnehmer*innen mit ihrer vollständigen Körpersprache kennen, es kommt im Seminar und in den Pausen auch zu informellen Kontakten, es wird zwischendurch gescherzt und gelacht, und für etwas Small Talk findet man immer Zeit. Die Gefahr besteht, dass dies in Online-Seminaren anders ist, dass auf Grund der größeren Anonymität der Beziehungsaufbau erschwert ist und kein Funke überspringt.

Diesen Herausforderungen begegnet die dgp, indem sie einige ihrer Trainer*innen zu zertifizierten Online-Trainer*innen qualifiziert und dadurch mit entsprechendem Methodenrepertoire ausgestattet hat, so dass das Onboarding, die Netzwerkbildung – sprich der Beziehungsaufbau – durch den Einsatz entsprechender Methoden auch online bislang sehr gut gelungen ist.

Durch Kleingruppenarbeit in separaten Breakout-Rooms, kurze zwischengeschaltete Aktivierungen sowie gut angeleiteten Austausch kann sich auch im virtuellen Raum ein Wir-Gefühl entwickeln. Auch der Gefahr der Ermüdung und des Abschweifens durch zu lange Phasen der Inaktivität kann durch Gruppen- oder Partner*innenarbeit in Breakout-Rooms begegnet werden. Zur Aktivierung bietet sich außerdem die Nutzung der Kommentierfunktion zum gemeinsamen Arbeiten auf Folien oder dem Whiteboard an sowie der Einsatz von Umfragetools. Diskussionen sind natürlich – wie im Präsenzseminar – ebenfalls möglich. Auch für individuelles Feedback bieten sich mehrere Funktionen an. Alle angesprochenen Möglichkeiten sorgen dafür, die Teilnehmer*innen „bei der Stange“ und aktiv zu halten. Da alle Teilnehmer*innen sich per Video zuschalten, fällt ein zu offensichtliches Abschweifen auch auf, so dass der oben angesprochene soziale Druck im virtuellen Raum noch zusätzlich wirkt. Als einziges Manko bleibt, dass Pausengespräche zu zweit, ohne dass der Rest lauscht, kaum möglich sind, aber bei Bedarf z.B. mittels Chatfunktion oder durch Öffnen von „Pausenräumen“ realisiert werden können.

Viele praktische Erfahrungen resultieren aus interaktiven Online-Seminaren für Ausbildungsbeauftragte. Speziell in der praktischen Ausbildung der Anwärter*innen und Auszubildenden gibt es einen großen Bedarf an neuen Ausbilder*innen. Da es in einigen Verwaltungen zum Standard gehört, als Ausbildungskräfte vorab an einer Qualifizierungsreihe teilgenommen zu haben, war die Durchführung dieser Qualifizierungsreihe auch trotz Corona-Rahmenbedingungen unbedingt erforderlich.

Mittlerweile wird die gesamte Modulreihe für Praxisausbilder*innen vollständig sowohl als Präsenz- als auch als Online-Seminar angeboten. Es zeigte sich, dass diese

Veranstaltungen zusätzlich eine Unterstützung in Richtung Digitalisierung der Ausbildung darstellten. Viele Teilnehmer*innen erlebten einen Zuwachs an Digitalkompetenz, da sie zusätzlich zu den Modulinhalten der Ausbildungsreihe noch die Vielfalt der Möglichkeiten der Plattformen kennen- und damit auch praktisch umzugehen lernten. So konnten viele der eingesetzten Techniken im Anschluss auch im Ausbildungsalltag eingesetzt werden.

Dadurch, dass Ausbildung von einem auf den anderen Tag, also ohne Vorbereitung und Absprache oftmals „von Homeoffice zu Homeoffice“ stattfand, bestand bei diesen Seminargruppen darüber hinaus ein großer Bedarf an Austausch konkret z.B. zum Thema „Ausbilden von Homeoffice zu Homeoffice“ (wie halte ich Kontakt, welche Methoden bieten sich an, wie bewerte ich...?).

Natürlich gehen auch die Erfahrungen aus anderen Live-Online-Trainings wie z.B. Selbst- und Zeitmanagement (im Homeoffice), Prüfungsvorbereitung, Remote-Führung, Interkultureller Kompetenz oder „Service und Kund*innenorientierung“ in diesen Bericht ein.

„Zeitumfang war gut, die Themen interessant, ein sehr angenehmes Arbeiten!“

Vorteile von Live-Online-Seminaren und Webinaren

- Veranstaltungen können (relativ) spontan verabredet werden, da es weder einer Raumbuchung noch eines Caterings bedarf.
- Durch die Ortsunabhängigkeit von Online-Seminaren und Webinaren können sich Teilnehmer*innen von überall dazu schalten. Auch arbeitgeberübergreifende Veranstaltungen sind leicht möglich.
- Da die Teilnahme ortsunabhängig ist, können auch Homeoffice-Worker, Teilzeitkräfte, Menschen mit privaten Verpflichtungen teilnehmen, für die Dienstreisen nicht möglich sind (Stichwort: Homeschooling).
- Reisezeiten fallen weg, ggf. werden dadurch andere Termine noch realisierbar.
- Es entstehen keine Reisekosten, der CO2-Ausstoß verringert sich. Dies stellt einen großen Pluspunkt unter Umweltschutzgesichtspunkten dar.
- Ein weiteres Plus für den Klimaschutz bilden die Teilnehmer*innen-Unterlagen, die gleich im Anschluss inklusive der Arbeitsergebnisse als pdf-Datei zur Verfügung gestellt werden, wodurch eine Menge Papier eingespart wird.
- Durch die Nutzung digitaler Medien wird vorgelebt, was es heißt, lebenslang zu lernen und mit seinen Kompetenzen aktuell zu bleiben. Trainer*innen und Teilnehmer*innen bleiben auch nach der Pandemie digital „am Ball“.

Fazit

Dieser Praxisbericht hat einen Einblick in die wertvollen Lernerfahrungen im Online-Trainingsbereich der dgp innerhalb der letzten Monate gegeben. Unser Fazit nach vielen Monaten Praxiserfahrung lautet eindeutig, dass bei guter Live-Online-Didaktik (fast) alles möglich ist. So finden interaktive Online-Seminare weiterhin von Mensch zu Mensch mit direktem Austausch zwischen Trainer*in und Teilnehmer*in statt, aber eben nicht mehr mit allen Beteiligten am gleichen Ort, sondern im virtuellen Seminarraum. Die allermeisten Trainingsinhalte können problemlos als Live-Online-Trainings im virtuellen Seminarraum durchgeführt und alle gängigen Trainings-Methoden realisiert werden.

Die aufgeführten Vorteile von Online-Seminaren zeigen, dass diese Trainingsform auch in Zukunft, also in Nach-Corona-Zeiten zumindest einen Teil des Weiterbildungsportfolios ausmachen sollte. Bestimmte Themen könnten als Online-Seminare oder als Hybridveranstaltungen durchgeführt werden. Hybridveranstaltungen vereinen die Vorteile von Präsenz- und Online-Seminaren. Sie bieten die Möglichkeit, die Digitalisierungswelle in der Trainingsbranche „mitzureiten“ und nicht von ihr überrollt zu werden. Zusätzlich bieten sie den persönlichen Kontakt mit „echten“ Pausengesprächen bei duftendem Kaffee. Es gibt viele Möglichkeiten. Bieten Sie diese an!

Kontakt:

Beate Scholz
Dipl.-Psychologin
Deutsche Gesellschaft für Personalwesen e.V.
scholz@dgp.de

Die Deutsche Gesellschaft für Personalwesen e.V. (dgp) bietet ein breites Repertoire an Seminar- und Beratungsangeboten im Kontext von Personalarbeit, um unterschiedlichste Ausgangslagen von Organisationen zu berücksichtigen und lösungsorientiert zu bedienen. Alle Angebote sind zeitlich flexibel und werden individuell auf die jeweiligen Bedürfnisse der Kund*innen zugeschnitten. Erstklassig qualifizierte Trainer*innen vermitteln wissenschaftlich fundierte und hoch praxisbezogene Inhalte.

Wir begleiten Sie gern bei der Bearbeitung Ihrer Herausforderung und freuen uns auf den persönlichen Kontakt mit Ihnen, sei es per Telefon, Mail oder einem Treffen.

Einen Überblick über unser gesamtes Angebot finden Sie hier:

Passend zu den Themen der vorliegenden dgp-Informationen möchten wir Ihre Aufmerksamkeit besonders auf folgende, ausgewählte Seminarangebote lenken:

Digitale Arbeits- und Lebenswelten

Führung auf Distanz – Führung im digitalen Wandel

Wie kommt der Teamgeist in den virtuellen Raum?

Gestalten von Veränderungsprozessen

Diversity Management: Steuern von Verschiedenheit

Gleichstellung als Führungsaufgabe

Gesund leben bei hoher Arbeitsbelastung

dgp

Deutsche Gesellschaft
für Personalwesen e.V.

Die Deutsche Gesellschaft für Personalwesen e.V. (dgp) ist die älteste deutsche Personalberatung und arbeitet seit über sieben Jahrzehnten erfolgreich auf dem Gebiet der Personalauswahl sowie der Personal- und Organisationsentwicklung.

Wir sind Vordenker*innen und Vorbild für verantwortungsvolle Personalarbeit in Deutschland.

Unsere Kund*innen:
Zu den Kund*innen der dgp zählen beispielsweise Bundesministerien und Landesministerien sowie deren nachgeordnete Behörden, Kommunen, Hochschulen, Kulturorganisationen, Stiftungen und Verbände sowie Unternehmen zahlreicher Branchen, von mittelständischen Betrieben bis hin zu international agierenden Organisationen.

Unsere Regionalstruktur gründet auf fünf Geschäftsstellen in Berlin, Düsseldorf, Hannover, Leipzig und Stuttgart. Damit setzen wir auf kurze Wege und können gemäß den örtlichen Bedürfnissen handeln.

Unsere Leistungen:

Organisationsberatung, die unsere Kund*innen bei der Erreichung ihrer Ziele unterstützt

Zukunftsorientierte Personalauswahl, sowohl von Nachwuchs- als auch von Fach- und Führungskräften

Individuelle, beratende Begleitung (Coaching), insbesondere von Führungskräften

Eigene Forschung zur Überprüfung und Aktualisierung vorhandener fachpsychologischer Methodik

Erarbeitung von Problemlösungen für die Arbeitswelt von morgen

Sprechen Sie uns bitte an, wir sind für Sie da!

**Deutsche Gesellschaft
für Personalwesen e.V.**

Berlin
Kantstraße 153
10623 Berlin
T 030 / 3983718-0
F 030 / 3983718-29
E berlin@dgp.de

Düsseldorf
Hohenzollernstraße 11-13
40211 Düsseldorf
T 0211 / 688508-0
F 0211 / 688508-29
E duesseldorf@dgp.de

Hannover
Stammestraße 40 D
30459 Hannover
T 0511 / 94393-0
F 0511 / 94393-43 / 44
E hannover@dgp.de

Leipzig
Grassistraße 12
04107 Leipzig
T 0341 / 9782-223
F 0341 / 9782-225
E leipzig@dgp.de

Stuttgart
Kernerstraße 43
70182 Stuttgart
T 0711 / 9457-6701
F 0711 / 9457-2765
E stuttgart@dgp.de

www.dgp.de