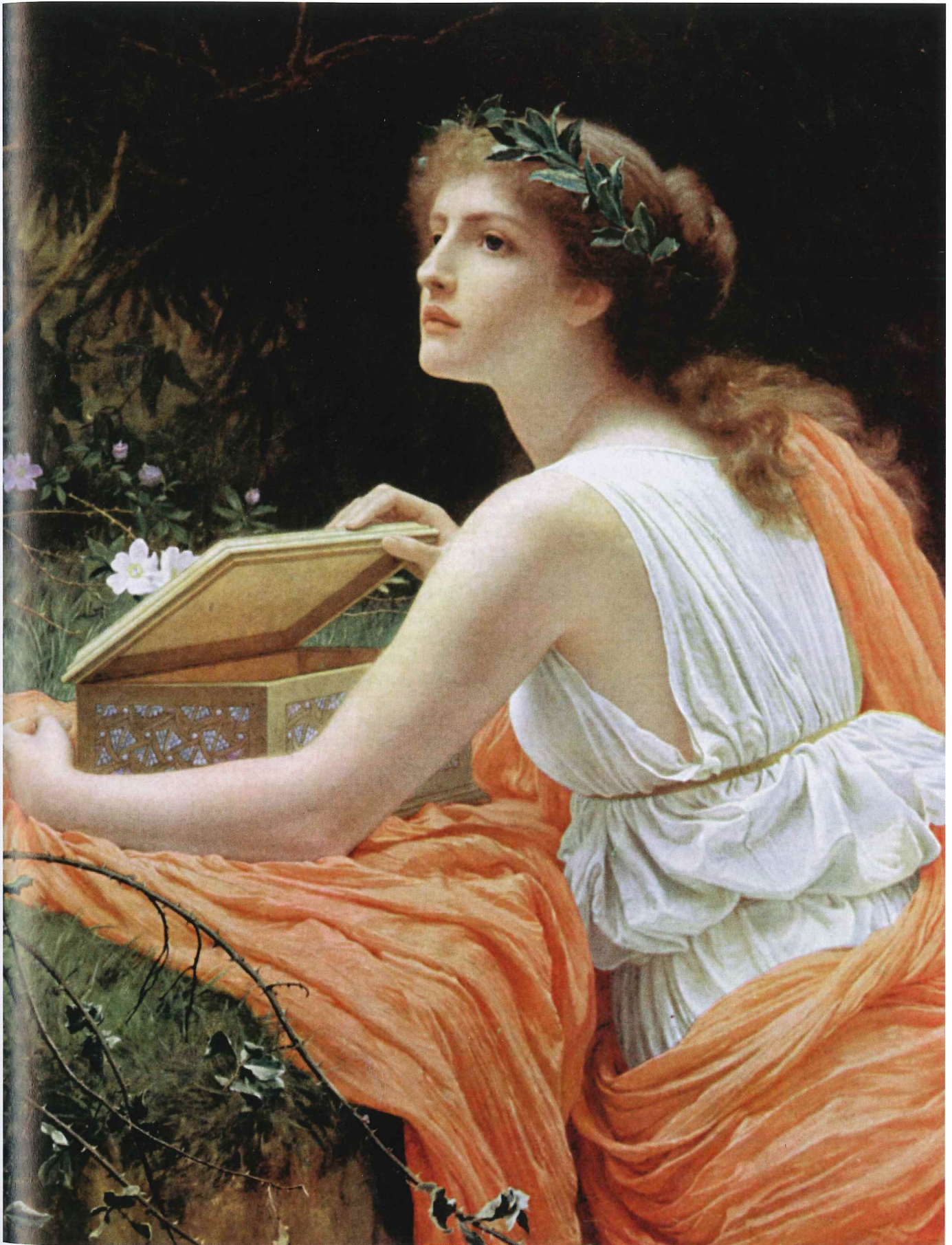


Kersting, M. (2019). Wovon man nicht reden kann, darüber muss man schweigen. Reflexionen zur Eignungsdiagnostik mit PRECIRE. Personalmagazin, 11/2019, S. 60-65.

„Wovon man nicht reden kann, darüber muss man schweigen.“

Reflexionen zur Eignungsdiagnostik mit Precire –
von Prof. Dr. Martin Kersting

Die Verlockung ist groß: Endlich gibt es mit Precire ein digitales Produkt für die Eignungsdiagnostik, das HR in die Zukunft führt. Doch die Zweifel an der Tauglichkeit des Produkts sind groß, in wichtigen Aspekten fehlen wissenschaftliche Grundlagen.



● Pandora wurde von den Göttern mit vielfältigen Begabungen, Schönheit und einem großen Sprachtalent beschenkt, um Epimetheus zum Öffnen ihrer verhängnisvollen Büchse zu verführen. Prometheus warnte seinen Bruder vor dem Geschenk der Götter, Epimetheus aber war fasziniert. Ähnlich gespalten wie Epimetheus und Prometheus reagiert die HR-Welt auf Precire Jobfit (im Folgenden: Precire), ein Tool, das mittels komplexer Algorithmen aus 15 Minuten gesprochener Sprache Aussagen zu Fähigkeiten und Eigenschaften des Sprechers ableitet.

Faszination „Moderne Eignungsdiagnostik“

Die Meinungsdivergenzen lassen sich größtenteils darauf zurückführen, dass verschiedene Bewertungsaspekte für die Qualitätseinschätzung herangezogen werden. Precire punktet damit, dass die Datenerhebung unkompliziert und flexibel erfolgt. Die Teilnehmer können zu einer beliebigen Zeit von einem beliebigen Ort aus mit einem Chatbot telefonieren. Das Ganze geht schnell und dennoch werden in der kurzen Zeit große Datenmengen erfasst und computergestützt analysiert. Die resultierenden Persönlichkeitsgutachten sind sehr weitgehend und umfassend. Die Argumentation, dass die Sprache „zum Charakter des Menschen“ gehört (Francis Bacon), ist grundsätzlich ebenso überzeugend wie der Hinweis auf die hohe Bedeutung der Kommunikation für das Berufsleben. Das Verfahren soll gegen Verfälschungen immun sein.

Für Precire spricht vor allem, dass die Vorgehensweise zeitgemäß erscheint. Während in allen Geschäftsbereichen die Automatisierung voranschreitet, arbeitet die Personal diagnostik häufig noch nach Art eines mittelalterlichen Handwerksbetriebs mit Vorstellungsgesprächen und Beobachterkonferenzen. Mit Precire scheint das Versprechen „HR goes digital“ endlich auch für die Diagnostik eingelöst zu werden. Und noch ein zweiter Trend spielt Precire in die Karten: Der Trend zur Diversity beziehungsweise die Angst, dass man vom Management oder den Kunden für eine mangelhafte Diversity zur Rechenschaft gezogen wird. Ist die Homogenität der Belegschaft nicht letztlich den (unbewussten) Vorurteilen der HR-Mitarbeiter geschuldet? Wie gut, wenn man hier die Hände in Unschuld waschen kann, da man die Eignungsbeurteilung der vermeintlich objektiven Maschine überlassen hat. In der Begeisterung für Precire drückt sich ein diffuser Wille zum Vorstoß von HR in die Moderne aus.

So dick hinsichtlich des Anspruchs aufgetragen wird, so dünn ist die Faktenlage. Die Vision erweist sich weitgehend als Illusion. Im Folgenden lenke ich die Aufmerksamkeit kurz auf das Fachurteil von unabhängigen Gutachtern über Precire und gehe dann auf einige Besonderheiten des Verfahrens ein.

Unabhängige Experten bewerten Precire

Für die Bewertung der Qualität eines psychodiagnostischen Verfahrens sind zahlreiche Aspekte zu berücksichtigen. Die verbreitete Vorgehensweise, „sich selbst ein Urteil zu bilden“, indem man den zu beurteilenden Test selbst durchführt und „prüft“, ob das resultierende Gutachten „passt“, klingt gut, ist es aber nicht. Die Überzeugung, dass das resultierende Gutachten die eigene Seelenlage zutreffend beschreibt, lässt sich leicht hervorrufen. Dies funktioniert beispielsweise über Aussagen, die universell auf alle Menschen zutreffen (sogenannte „Barnum-Aussagen“, siehe Personalmagazin 9/2015, Seite 28-33).

Anstelle der „subjektiven Evidenz“ bedarf die Testbeurteilung sachgerechter Kriterien. Das Diagnostik- und Testkuratorium (DTK) hat dazu das Testbeurteilungssystem (TBS-DTK) entwickelt, das zwei Bestandteile hat. Der erste Bestandteil ist eine auf Basis der DIN 33430 erstellte Checkliste, mit der geprüft wird, ob in den Verfahrenshinweisen zum Test alle Informationen zur Verfügung stehen, die notwendig sind, um die Qualität eines Tests beurteilen zu können. Diese Checkliste stellt den „Standard zur Information und Dokumentation von Instrumenten zur Erfassung menschlichen Erlebens und Verhaltens“ dar. Darüber hinaus sieht das System als zweiten Bestandteil ausführliche Richtlinien zur Qualitätsbeurteilung von Tests vor.

Ausgewählte Tests werden von Experten nach diesen Richtlinien beurteilt. Bislang liegen für rund 40 Tests solche Beurteilungen vor. Da es mehrere Tausend Tests gibt, bedeutet das umgekehrt, dass für die Mehrheit der Tests keine TBS-DTK-Rezension vorliegt. Jede Person, die sich für die Qualität eines Tests interessiert, kann die Kriterien und Richtlinien aber auf das interessierende Verfahren anwenden und beispielsweise den Testanbieter auffordern, die Checkliste auszufüllen. Alle Materialien (Checkliste, Richtlinien) sowie die bisherigen Testrezensionen stehen unter <http://www.zpid.de/Testkuratorium> zum freien Download bereit.

Jeder Test wird von zwei einschlägigen Experten rezensiert, die sich zunächst unabhängig voneinander ein Urteil bilden. Das Diagnostik- und Testkuratorium achtet auf die Einhaltung der Beurteilungsrichtlinien. Eine gekürzte Fassung der TBS-DTK-Rezension zu Precire ist im Kasten (Seite 57) wiedergegeben. Die in dem TBS-DTK formulierten Qualitätsanforderungen an die Zuverlässigkeit und Gültigkeit von Tests kann Precire demzufolge nur teilweise erfüllen. Die Experten bemängeln insbesondere die schlechte Nachvollziehbarkeit des methodischen Vorgehens und die Tatsache, dass keine ausreichenden Nachweise zur Treffsicherheit der aus Precire abgeleiteten Aussagen erbracht wurden.

Vernachlässigung der Frage der Datenqualität

Ergänzend zu der Rezension möchte ich hier auf einige Besonderheiten von Precire eingehen, zunächst auf die Datenqualität. Bei jeder datengestützten Lösung hängt die Qualität der Analyse von der Qualität der verarbeiteten Daten ab (Garbage In, Garbage Out). Es ist beunruhigend, dass die Anbieter von Precire die Qualität der Daten nicht thematisieren. Es werden überhaupt keine Anforderungen an die technische Aufnahmequalität (zum Beispiel Qualität der Telefonverbindung) gestellt. Mögliche Beeinträchtigungen der Aufnahme (zum Beispiel durch ungünstige Situationen) oder der Stimme (zum Beispiel durch Erkrankungen) werden nicht angesprochen. Dialekte sollen für die Auswertung keine Rolle spielen, solange eine Mindesterkennungsrate von 85 Prozent der Wörter vorliegt. Hierzu ist anzumerken, dass die Spracherkennungsrate deutlich von Merkmalen der Sprecher abhängt. Die Spracherkennungsrate für ein etabliertes Spracherkennungssystem beträgt für das Englisch, das in Indien gesprochen wird, 78 Prozent, für das Englisch, das in Schottland gesprochen wird, nur noch 53 Prozent (<https://hbr.org/2019/05/voice-recognition-still-has-significant-race-and-gender-biases/>). Vor allem ist die Erkennungsrate höher für die Sprachen, an denen das System trainiert wurde. Dies sind häufig Männer,

sodass die Sprache von Männern akkurater erkannt wird als die von Frauen. Anders formuliert: Die Spracherkennung hat einen „Bias“. Dabei geht es nicht nur um die Erkennungsrate. Das Gesprächsverhalten steht im Zusammenhang mit Geschlecht und Gesprächssituation (und deren Interaktion) – alles Faktoren, die bei den Precire-Analysen ignoriert werden.

Grundsätzlich funktionieren Programme wie Precire so, dass die Mustererkennung an einem Trainingsdatensatz „gelernt“ wird. In diesem Datensatz befinden sich die Daten von Personen mit bestimmten Merkmalen (zum Beispiel Geschlecht, Bildung, Alter, Muttersprache, Dialekt). Für die Interpretation von Daten eines Teilnehmers aus einem Ernstfalleinsatz spielt es eine Rolle,

Bewertung von Precire nach dem System des DTK

Im Folgenden dokumentieren wir in stark gekürzter Form die Rezension von „Precire Jobfit“, die Prof. Dr. Lothar Schmidt-Atzert (Universität Marburg), Janina Künecke (Psychologische Hochschule Berlin) und Prof. Dr. Johannes Zimmermann (Universität Kassel) nach dem Testbeurteilungssystem des Diagnostik- und Testkuratoriums (TBS-DTK) erstellt haben. Die vollständige Fassung der Rezension kann abgerufen werden unter: <https://www.bdp-verband.de/binaries/content/assets/beruf/testrezensionen/precirejobfit.pdf>

Beschreibung des Tests und der diagnostischen Zielsetzung: Das Verfahren soll anhand der Sprachprobe eines 15-minütigen Telefoninterviews berufsrelevante Merkmale automatisiert erfassen. (...)

Bewertung des Informationsgehalts der Verfahrenshinweise:

Die Anforderungen an die allgemeinen Informationen sind teilweise erfüllt. Die Dokumentation der empirischen Studien in den Verfahrenshinweisen weist große Lücken auf. (...) In der Psychologie ist die Erkenntnis fest verwurzelt, dass Ergebnisse stets methodenabhängig sind; daher ist das Vertrauen in die Ergebnisse der unzureichend dokumentierten Studien gering.

Theoretische Grundlagen als Ausgangspunkt der Testkonstruktion: (...) „Precire Jobfit“ wertet nach eigenen Angaben 564.367 (!) sogenannte „Features“ aus, die sich aus lexikalischen, syntaktischen, prosodischen Informationen und deren Kombinationen ergeben. (...)

PRECIRE JobFit	Die TBS-DTK-Anforderungen sind erfüllt			
	voll	weitgehend	teilweise	nicht
Informationsgehalt der Verfahrensweise			X	
Objektivität		X		
Zuverlässigkeit			X	
Gültigkeit			X	
			Ja	Nein
In den Verfahrenshinweisen ist verzeichnet, wo die nach dem DTK-Testinformationsstandard notwendigen Informationen zu finden sind.				X

Objektivität: (...) Insgesamt sind die Anforderungen an die Objektivität weitgehend erfüllt.

Normierung: Die 5.201 Personen, an denen das Verfahren entwickelt wurde, fungieren als Eichstichprobe. (...) Laut Autorinnen und Autoren haben Alter, Geschlecht und Bildungsgrad keinen Einfluss auf die Skalenwerte (...). Dies ist merkwürdig, da viele Persönlichkeitsmerkmale sehr wohl von solchen Faktoren abhängen (...)

Zuverlässigkeit: (...) Insgesamt sind die Anforderungen an die Reliabilität teilweise erfüllt.

Gültigkeit: (...) Belastbare Belege zur Kriteriumsvalidität (hinsichtlich beruflicher Eignung) liegen nicht vor. (...) Die Anforderungen an die Validität sind insgesamt teilweise erfüllt.

Abschlussbewertung: Bei „Precire Jobfit“ handelt es sich um ein ungewöhnliches Testverfahren, das aus dem Potenzial von Methoden des maschinellen Lernens schöpft. (...) ungewöhnlich ist die schlechte Dokumentation des methodischen Vorgehens. Das Verhältnis von untersuchten Personen (N = 5.201) zu analysierten Variablen (564.367 Features) bei der Erstellung eines Vorhersagemodells für psychologische Merkmale ist sehr ungünstig. Es ist nahezu unvermeidbar, dass dabei zufällig im Datensatz vorhandene Konstellationen von Features ausgebeutet werden. (...) Die Überprüfung der Konstruktvalidität an einer neuen Stichprobe ist schlecht dokumentiert, die Angaben zur Kriteriumsvalidität sind ungenügend.

Vor allem fehlt ein Nachweis, dass das Verfahren eine inkrementelle Validität gegenüber den Persönlichkeitsfragebogen aufweist, anhand derer es entwickelt wurde. Mit anderen Worten: „Precire Jobfit“ erfasst bestimmte Persönlichkeitsmerkmale, die ansonsten mit Fragebogen gemessen werden. Es bleibt aber offen, ob damit die berufliche Eignung besser erkannt werden kann. Es fehlt sogar der Nachweis, dass „Precire Jobfit“ diesbezüglich den Fragebogen ebenbürtig ist. Insgesamt ist die psychometrische Qualität des Verfahrens nicht hinreichend belegt. (...)

wie weit dieser konkrete Teilnehmer den Referenzpersonen im Trainingsdatensatz entspricht. Das Programm sucht Ähnlichkeiten zwischen den Werten des aktuellen Teilnehmers und Werten ähnlicher Teilnehmer in der Datenbank. Entsprechend problematisch ist die Auswertung in den Fällen, in denen der Teilnehmer den Personen im Trainingsdatensatz nicht ähnlich ist. Algorithmische Auswertungen führen regelmäßig zur Benachteiligung von Personen, die in Bezug auf den Trainingsdatensatz „Ausnahmen“ darstellen.

Problematische soziale Akzeptanz

Ein zweiter Punkt in Bezug auf die Datenerhebung betrifft die soziale Akzeptanz von Precire, die nicht hinlänglich erforscht ist. In dieser Hinsicht ist der Inhalt der vom Chatbot gestellten Fragen und die sogenannte „Subjektqualität“ zu thematisieren. Zunächst zu den Fragen. Das Interesse der (potenziellen) Arbeitgeber(innen) darf sich nur auf Informationen richten, die einen Anforderungsbezug aufweisen. Vielen Precire-Fragen fehlt der Berufsbezug. Stattdessen werden Privatleben oder gar gesundheitliche Aspekte thematisiert. Zum Beispiel wird gefragt: „Bitte beschreiben Sie den Ablauf eines typischen Sonntags“, „Wie ging es Ihnen in den letzten Wochen? Wie haben Sie sich gefühlt?“ oder „Wie laufen bei Ihnen zu Hause Familienfeiern ab, zum Beispiel Geburtstage oder Weihnachtsfeiern?“ Dies sind Themen, die für den Arbeitgeber eindeutig „tabu“ sind. Den Teilnehmern wird mitgeteilt, dass der Inhalt ihrer Aussagen nicht relevant sei und die Audiodatei und das Transkript dem Auftraggeber nicht zur Verfügung gestellt werden. Ungeachtet dieses Sachstands dürfte es psychologisch bei einigen Teilnehmern aber Befremden auslösen, zu privaten Themen befragt zu werden. Inzwischen haben die Fraktion „die Linken“ sowie einzelne Abgeordnete – unter anderem mit Bezug auf diese Art von Fragen – eine kritische Anfrage zu

Precire an die Bundesregierung gestellt (Drucksache 19/12439, 16. August 2019).

Darüber hinaus erfordert die Aussage, dass der Inhalt der Aussagen nicht relevant sei, einen Beleg, den Precire schuldig bleibt. Aussagen umfassen Worte. Bei Precire wird ausgewertet, welche Worte vom Sprecher genutzt werden. Diese Worte werden zu Wortkategorien geordnet. Personen mit Kindern werden, nach ihrem Privatleben gefragt, wahrscheinlich Wörter aus dem Themenfeld „Kinder“ häufiger nutzen als Personen ohne Kinder. Dies könnten – das müsste man untersuchen – beispielsweise mehr „Emotionswörter“ sein. Wäre das so, wäre es – entgegen der Behauptung – nicht irrelevant, wonach gefragt wird.

Verletzung der Subjektqualität

Zur sozialen Akzeptanz des Verfahrens gehört auch die „Subjektqualität“. Es sollen „keine Daten erhoben und verwendet werden, welche der willentlichen Steuerung der Betroffenen grundsätzlich entzogen sind“ (Richtlinie des Ethikbeirats HR-Tech). Es gibt viele Daten, die Auskunft über die Psyche des Menschen geben, die wir als Mensch nicht (gut) kontrollieren können. Hierzu zählen insbesondere Biosignale wie beispielsweise die Reaktionen des Kreislaufes, der Atmung, der Muskulatur und der Haut. Bewerber und Mitarbeiter müssen davor geschützt werden, unbeabsichtigt oder unwillkürlich Informationen an die (potenziellen) Arbeitgeber weiterzugeben.

Ich thematisiere hier nicht die juristischen, sondern die psychologischen Aspekte. Es ist davon auszugehen, dass eine Verletzung der „Subjektqualität“ der Teilnehmer die soziale Akzeptanz des Verfahrens mindert. In diesem Kontext wurde schon diskutiert, dass es auch vorkommen kann, dass ein Bewerber im Interview stottert oder im Assessment Center kurzzeitig die Kontrolle über seinen Gesichtsausdruck verliert. Geschenkt. Subjektqualität ist kein „Null-Eins“ Kriterium. Aber es dürfte außer Frage stehen, dass Precire wesentlich stärker als andere Verfahren die Subjektqualität infrage stellt. Precire informiert die Teilnehmer in der Einverständniserklärung explizit, dass es „um in der Sprache zum Ausdruck kommende unbewusste Aspekte des Sprechenden“ geht und wirbt damit, dass „die Testteilnehmer die Ergebnisse kaum gezielt in eine gewünschte Richtung steuern können“ (Handbuch Precire, 2016, Seite 49). Verfälschbarkeit und Subjektqualität sind zwei Seiten derselben Medaille. Unverfälschbarkeit eines Verfahrens erkaufte man sich immer durch Intransparenz und die Verletzung der Subjektqualität. Die mit der Unverfälschbarkeit einhergehende Intransparenz verhindert außerdem sachlogisch die für die Eignungsdiagnostik notwendige „informierte Einwilligung“ in das Verfahren.

Das größte Hindernis auf dem Weg zur Anwendung von Precire stellt die Intransparenz der Auswertung dar. Nach Lektüre des Testmanuals kann man eigentlich nur sagen: Aus einer 15-minütigen Sprachprobe erstellt Precire „irgendwie“ eine Aussage über die „kommunikative Kompetenz“, über „stabile Persönlichkeitseigenschaften“, über „persönliche Skills“ und über die „allgemeine Passgenauigkeit der Person auf die zu besetzende Stelle“. Basis dieser Aussagen sind laut dem Testmanual Korrelationen zwischen Bestandteilen der Sprachprobe und psychologischen Fragebogen sowie – nicht näher beschriebenen – Einschätzungen im Assessment Center. Diese Vorgehensweise ist in vielerlei Hinsicht überaus problematisch. Zunächst ist es

**Das größte
Hindernis auf
dem Weg zur
Anwendung von
Precire stellt die
Intransparenz der
Auswertung dar.**



PROF. DR. MARTIN KERSTING ist Professor für psychologische Diagnostik an der Justus-Liebig-Universität Gießen (www.kersting-internet.de)

aus inhaltlichen Gründen heikel, sich auf Korrelationen zu verlassen, da eine Maschine nicht zwischen Scheinkausalität und Kausalität unterscheiden kann. Die auf eine Drittvariable (Stadt versus Land) zurückzuführende Korrelation zwischen der Anzahl der Storchpaare und der Geburtenrate ist für eine Maschine gleichwertig zu einer Korrelation zwischen zwei kausal miteinander verbundenen Variablen. Korrelative Zusammenhänge liefern keine Aussagen über Kausaleffekte. Um einen Zusammenhang kausal interpretieren zu können, müssen zusätzlich zur Kovariation der beiden Variablen zwei weitere Bedingungen erfüllt sein: Die verursachende Variable muss der verursachten Variable zeitlich vorausgehen. Und der Einfluss der verursachenden auf die verursachte Variable darf nicht auf eine dritte Variable zurückführbar sein. Ob diese Bedingungen gegeben sind, muss (aktuell) noch ein Mensch prüfen.

Blackbox Precire: Warum Personalarbeit im Dunkeln nicht funktioniert

Auch zum Schutz vor Vorurteilen muss man die „Muster“, die der Rechner nutzt, inhaltlich betrachten und interpretieren. Dazu müssen diese offengelegt werden. Maschinen übernehmen ungeprüft alle Muster, die sie im Trainingsdatensatz entdecken – auch wenn es sich um „Scheinzusammenhänge“ oder um „unerwünschte“ Zusammenhänge handelt. Entgegen der Annahme, dass eine Maschine „objektiv“ sei, gilt, dass die Maschine Vorurteile übernimmt, die in den bisherigen Daten stecken. Persönlichkeitsfragebogen und Assessment Center, die Datenbasis von Precire, sind überaus anfällig für Verfälschungen und weisen insgesamt nur eine geringe Kriteriumsvalidität auf. Ein Programm, das an diesen Daten trainiert wird, übernimmt die Beschränkungen dieser Verfahren. Nur durch eine inhaltliche Bewertung der den Daten zugrunde liegenden „Muster“ kann sichergestellt werden, dass mit dem Einsatz des Programms keine unerwünschten Nebenwirkungen verbunden sind. So liefert die Sprache beispielsweise auch Hinweise auf eine depressive Erkrankung – eine Information, die in der Eignungsdiagnostik keine Rolle spielen darf. Solange die Auswertung eine Blackbox bleibt, solange der Zusammenhang zwischen dem von dem Rechner hergestellten (sprachlichen) Verhalten der Teilnehmer und den Verfahrensergebnissen im Dunkeln bleibt, bleiben zentrale Voraussetzungen für den Einsatz in der Eignungsdiagnostik unerfüllt.

Hier korrespondiert die Perspektive der Psychologie mit der juristischen Perspektive. Inhaltliche Bewertungen und Entscheidungen, die für den Betroffenen eine rechtliche Folge nach sich ziehen oder ihn erheblich beeinträchtigen, dürfen nach dem Bundesdatenschutzgesetz (BDSG) nicht „ausschließlich auf eine automatisierte Verarbeitung personenbezogener Daten gestützt werden“. Sie müssen durch eine „natürliche Person“ vorgenommen werden. Das BDSG sieht außerdem weitgehende

Auskunftspflichten vor, wobei den Betroffenen auch der logische Aufbau der automatisierten Verarbeitung der ihn betreffenden Daten zu erläutern ist. Grundsätzlich spricht nichts gegen die Nutzung von automatisiert verarbeiteten Informationen in der Eignungsdiagnostik, es muss aber gewährleistet werden, dass – wie es in der dritten Richtlinie des Ethikbeirats HR Tech formuliert wird – „Menschen entscheiden“. Und diese Menschen müssen zumindest prinzipiell verstehen, wie die Maschine zu ihren Vorschlägen kommt.

Precire ist kein selbstlernendes Programm

In diesem Kontext möchte ich auch darauf hinweisen, dass – entgegen der weitverbreiteten Meinung – Computerprogramme in der Eignungsdiagnostik nicht ohne Weiteres „lernen“, sich nicht selbstständig optimieren. „Lernen“ setzt ein Feedback über „richtig“ und „falsch“ voraus. Wenn ein Rechner Empfehlungen zu Konsumprodukten ausspricht („Kunden, die sich für x interessieren, interessieren sich auch für y“) werden die Reaktionen (zum Beispiel „Klicks“) auf die Empfehlung ausgewertet. Läuft die Empfehlung ins Leere, bewertet der Rechner die ausgesprochene Empfehlung als „falsch“ und optimiert den Algorithmus. Precire ist kein „intelligentes“ Programm in diesem Sinne. Da der Rechner nicht erfährt, ob seine Diagnosen und Empfehlungen richtig oder falsch waren, lernt er daraus nicht. Die Precire zugrunde liegenden Annahmen werden nicht überprüft.

Fazit und Bewertung

Ich bin kein Technikfeind, ich bin überzeugt, dass uns die künstliche Intelligenz und Big Data große Fortschritte in der Personaldiagnostik ermöglichen. Ich wäre Feuer und Flamme für Precire, wenn es ein Forschungsinstrument wäre. Leider wurde die Methodik aber nicht ausreichend erforscht, sondern als „Schnellschuss“ sofort in die Praxis getragen. Der „Ernstfalleinsatz“ in der Personalpraxis setzt aber eine Nachvollziehbarkeit in der Auswertung voraus und lässt sich nur über empirische Nachweise der Aussagekraft rechtfertigen. Das Verfahren müsste gegenüber weniger invasiven Verfahren einen solchen Zugewinn an Erkenntnis liefern, dass die Einschränkungen der Subjektqualität hingenommen werden könnten. Bislang ist aber bestenfalls gezeigt, dass das Verfahren einem herkömmlichen Persönlichkeitsfragebogen – unter dem Gesichtspunkt der Validität eines der schlechtesten Verfahren der Personaldiagnostik – ebenbürtig ist. Etwas Neues erschließt Precire nicht. Leider. Denn die Herausforderung, an der Precire gescheitert ist, bleibt: Eine Personaldiagnostik, die die neuen Möglichkeiten von Big Data und künstlicher Intelligenz im Rahmen der gesetzlichen und ethischen Rahmenbedingungen nutzt, um uns neue Einsichten zu eröffnen, die Menschen und Organisationen so zusammenbringen, dass beide sich positiv entwickeln.

Aus Verantwortung müssen wir dem „prometheischen“ Impuls widerstehen und die Büchse der Pandora so lange geschlossen halten, bis wir den Inhalt verstanden haben und die Büchse kontrolliert öffnen können. Dazu ist unabhängige wissenschaftliche Forschung notwendig. Vielen ist eine wesentliche Pointe der Geschichte von Pandoras Büchse in Vergessenheit geraten: Die Büchse enthielt nicht nur Übel, sondern – ganz unten in der Büchse versteckt – auch die Gabe der Hoffnung. ■

Hochleistungsteams
Was sie ausmacht und welche Rahmen-
bedingungen sie brauchen

Eignungsdiagnostik
Experten streiten über die
Tauglichkeit von Precire

Fachkräfteeinwanderung
Die wichtigsten Fragen und
Antworten zum neuen Gesetz

personal. magazin

Impulse zur Gestaltung der Arbeitswelt



11/2019

personalmagazin.de 13.50 €



HR Innovation Award
Erntezeit für gute Ideen