

Empirische Realität versus subjektive Repräsentation: Einfluss von Persönlichkeitsmerkmalen auf die Prozessüberwachung und -führung in Mensch-Maschine-Systemen

CHARLOTTE MEYER¹, JENS NACHTWEI² & SASKIA KAIN²

¹⁾ *Lehrstuhl Ingenieurpsychologie der Humboldt Universität zu Berlin*

²⁾ *Graduiertenkolleg prometei am Zentrum Mensch-Maschine-Systeme der Technischen Universität Berlin*

Schlüsselwörter: Persönlichkeitsmerkmale, Big-5, Leistungsmotivation, Risikoneigung, Effektstärke, Human Factor

Zusammenfassung

In der Human Factors Forschung agieren Forscher aus den unterschiedlichsten Disziplinen. Doch wie gut kennen jene sich tatsächlich mit dem Wissensbestand des Human Factors Bereichs aus? Wir sind dieser Frage für einen speziellen Bereich nachgegangen, dem Wissen über die Bedeutsamkeit von Persönlichkeitsvariablen für die Leistung von Operateuren bei der Mensch-Maschine-Interaktion. In der Literatur lassen sich empirische Belege für diese Bedeutsamkeit finden. Dennoch wird dieser Tatsache bisher in der Human Factors Forschung wenig Beachtung geschenkt. Möglicherweise verursacht durch mangelndes Wissen seitens der Forscher. Tatsächlich zeigte sich bei einer von uns durchgeführten Expertenevaluation, in der sieben Persönlichkeitsmerkmale hinsichtlich ihrer Effektstärke auf die Leistung von Operateuren in der Prozessüberwachung und -führung für den Bereich Flugverkehrsüberwachung und im Kraftwerk beurteilt werden sollten, dass die Effektstärken der Merkmale als eher gering eingestuft werden. Zudem scheint diese Einschätzung Aufgaben- und Systemunabhängig zu sein.

1. Vom Segen und Fluch eines interdisziplinären Forschungsfelds

Der Graben zwischen Messung einer Größe und der „wahren Ausprägung“ derselben spielt in jeder empirischen Wissenschaft eine Rolle. Jedwede Messung kann nur eine Annäherung an diesen ominösen wahren Wert bedeuten. So ist dieser Umstand z.B. in der Ratio zum Hypothesenprüfen (vgl. α - und β -Fehler) oder der Klassischen Testtheorie zu finden. Doch ist das, was letztlich gemessen wird und somit als empirisch real bezeichnet werden kann, auch für jeden Forscher einer Disziplin präsent? Die bisweilen kritisierte Publikationsflut (Schmidtke 2002) und die Tendenz zu interdisziplinären Forschungsansätzen haben schon lange dazu geführt, dass sogenannte Universalgelehrte nur Fiktion sein können. Meister (2001) beschreibt für den Forschungsbereich Human Factors, dass ein Forscher immer nur einen Teil des gesamten Wissenssystems seiner Disziplin verfügbar haben kann. Schmidtke (2002) mahnt zudem, dass zum Wissenschaftler auch der Mut gehört, Nichtwissen zu bekennen.

Somit ist die Interdisziplinarität im Bereich Human Factors ein zweischneidiges Schwert: die Potentiale sich gegenseitig zu befruchten, werden erkaufte weiße Flecken auf der (Wissens-)Landkarte. Kein Human Factors Forscher kann überall auf dieser Karte gewesen sein. Die Ausrichtung dieses Beitrags mag deshalb einige freuen, manche neugierig machen und bei anderen auf weniger Akzeptanz stoßen – ein Umstand, der hier in der Natur der Sache liegt, wie Salas (2008, S. 352) in seinen Reflektionen über Human Factors an der Schwelle zum 21. Jahrhundert treffend beschreibt: „*We cannot define who we are, what we represent, and who we include. Our science is everywhere and nowhere, in that our science cannot be represented in a coherent way. [...] I was surprised by how many reviewers stated, “This paper does not belong in our journal.” But another reviewer of the same paper would say, “This is an important topic in human factors.”*“ So wundert es nicht, dass auch Definitionsversuche zu Human Factors derart breit aufgestellt sind, dass vieles darin Platz hat, jedoch sicher nicht alles von jedem überblickt werden kann. Badke-Schaub, Hofinger und Lauche (2008) beziehen sich hier z.B. auf alle physischen, psychischen und sozialen Charakteristika des Menschen, die einen Handlungsbezug haben und mit soziotechnischen Systemen interagieren. Persönlichkeitsmerkmale als zeitlich und über Situationen stabile sowie verhaltensrelevante Merkmale (Asendorpf 1999) gehören sicher in dieses Konglomerat aus menschlichen Faktoren. Es stellt sich daher die Frage, ob überhaupt von *dem* Human Factors Experten gesprochen werden kann. Wir wollen an dieser Stelle nicht die Definition von Expertise diskutieren (ein Verweis auf einen gelungenen Überblicksartikel von Farrington-Darby & Wilson (2006) soll an dieser Stelle genügen). Klar ist jedoch, dass neben Fertigkeiten und Fähigkeiten eine gewisse, domänenspezifische Wissensbasis eine Grundvoraussetzung für einen Expertenstatus darstellt. Wenig diskutabel ist auch, dass es Abstufungen hinsichtlich der Expertise gibt, welche auf die Qualität der Arbeit in den meisten Fällen einen Einfluss haben. Ausgehend von empirischen Belegen über die Bedeutsamkeit von Persönlichkeitsmerkmalen haben wir die Expertenevaluation als Zugang verwendet, um für die heutige Praxis und Forschung die Frage zu beantworten, welchen Einfluss gewisse Persönlichkeitsmerkmale auf die Prozessüberwachung und -führung in Mensch-Maschine-Systemen (MMS) haben könnten und wie stark unsere Disziplin für diese Thematik sensibilisiert ist. Beantwortet haben diese Frage Experten mit unterschiedlich breitem Wissensspektrum. Während ein Promotionsstudent in seinem engen Feld sicher Experte ist, lässt er jedoch tiefgründiges Wissen in angren-

zenden Bereichen eher vermissen als ein Professor für eine der in Human Factors zusammengefassten Disziplinen. Nichtsdestotrotz ist es möglich, dass das Urteil eines Promotionsstudenten für unsere Fragestellung genauer ist, als dies eines breit aufgestellten Experten mit langjähriger Erfahrung. Haerem & Rau (2007) zeigen bspw., dass Novizen durchaus bessere Leistung als Experten zeigen können, wenn Kognition (hier Wissen zum Thema) und Aufgabe (hier Urteile im Themenbereich) eine hohe Passung aufweisen.

2. Persönlichkeitsmerkmale in Human Factors: Weshalb und wie?

Welcher Bezug lässt sich nun für Human Factors Experten zur Persönlichkeitspsychologie herstellen? Insbesondere in sicherheitskritischen und produktionsorientierten MMS spielt das Wissen um Leistungsdeterminanten bei der Prozessüberwachung und -führung eine zentrale Rolle. Niederschlag fand dies u.a. in den Ansätzen zur sog. Human Reliability Analysis aus den 1980er und 1990er Jahren. Gemein ist diesen Ansätzen die Betrachtung von Performance Shaping Factors (PSF), wie sie in verschiedenen Modellen dargestellt und zur Vorhersage der Leistung von Operateuren herangezogen wurden (vgl. Swain & Guttman 1983). Diese Faktoren thematisieren insbesondere den Kontext, das technische System oder die Umwelt, vernachlässigen jedoch eine Leistungsdeterminante, die den Operateur abgesehen von seiner Expertise originär ausmacht: seine Persönlichkeit. Nach dem Scheitern der Ansätze zur Unfallpersönlichkeit aus den 1920er und 1930er Jahren wurden demnach gerade die Persönlichkeitsmerkmale nicht mehr systematisch bei der Mensch-Technik Interaktion betrachtet, so dass sich Befunde zum Einfluss der Persönlichkeitsmerkmale nur vereinzelt in der Literatur finden lassen. Huber, Kain & Nachtwei (2008) sowie Kain & Nachtwei (in Druck) geben einen Überblick über den Einfluss einiger Merkmale auf die Leistung (Bedienfehler und -zeiten) und Zustände (Situationsbewusstsein) von Operateuren in verschiedenen Kontexten. So geht z.B. Neurotizismus oft mit hochriskantem Fahren einher (Booth-Kewley & Vickers 1994); Yeo & Neal (2004) zeigten, dass es einen signifikanten positiven Zusammenhang zwischen Gewissenhaftigkeit in einer Simulationsumgebung für Fluglotsen und deren Anstrengung gibt. Doch auch, wenn sich zeigt, dass jene Merkmale durchaus einen Einfluss auf das Verhalten von Operateuren in Human Factors Untersuchungen haben, ist eher jener Ansatz von Interesse: der Einbezug der Persönlichkeitsmerkmale als Kontrollvariablen. Dabei ist eine Kontrollvariable eine spezifische Form von Moderator, welche die Beziehung zwischen unabhängiger und abhängiger Variable (uV und aV) beeinflusst; dieser Einfluss kann anschließend herausgerechnet werden (Bortz & Döring 2006). Letztlich kann der umsichtige Einbezug von Kontrollvariablen (und somit das Binden von personinhärenter Störvarianz) die Hypothesenprüfung sicherer machen und damit die Teststärke (Chance, einen tatsächlich vorhandenen Effekt auch nachzuweisen) erhöhen (Jo 2002). Die Voraussetzung dafür ist jedoch, dass diese Persönlichkeitsmerkmale in der Untersuchung (zumindest in Form von zeitökonomischen Kurzskalen) mit erfasst werden. Das wiederum ist schwierig, wenn die Sensibilisierung für dieses Themenfeld bei einigen Vertretern der Human Factors Disziplin fehlt. Doch dies ist zunächst eine Unterstellung und muss empirisch näher geprüft werden. Wie präsent sind die Befunde zum Einfluss von Persönlichkeitsmerkmalen auf die Prozessüberwachung und -führung in MMS bei Human Factors Forschern?

3. Expertenevaluation

3.1 Überblick und Fragestellung

Als Erhebungsverfahren der subjektiven Einschätzung der pragmatischen Bedeutsamkeit von Persönlichkeitsmerkmalen für die Mensch-Maschine-Interaktion aus Expertensicht wurde im Rahmen der so genannten „Frühjahrsschule 2009“ des Graduiertenkollegs prometei (Prospektive Gestaltung von Mensch-Technik-Interaktion am Zentrum Mensch-Maschine-Systeme der Technischen Universität Berlin) eine Expertenevaluation durchgeführt: Experten aus dem Human Factors Bereich sollten aufgrund ihrer Erfahrungswerte die Effektstärke von Persönlichkeitsmerkmalen in Bezug auf die Leistung von Operateuren einschätzen: Sie sollten angeben, ob und ggf. *in welcher Höhe ein bestimmtes Persönlichkeitsmerkmal einen Effekt auf die Leistung von Operateuren* in der Praxis hat. Sieben unterschiedliche Persönlichkeitsmerkmale sollten dabei in zwei konkreten Szenarien (Prozessüberwachung und -führung in einem Kraftwerk vs. in der Flugverkehrsüberwachung) in ihrer Relevanz für die Praxis und die Forschung eingeschätzt werden. Die Auswahl beider o.g. Szenarien richtete sich nach vielzitierten Mensch-Maschine-Systemen aus der Literatur: Kern empirischer Untersuchungen stellt besonders die Sicherheit von Handlungen dar, so dass Beispiele und Themen aus Hochrisikobranchen (Luftfahrt, Prozessindustrie etc.) im Vordergrund stehen (Badtke-Schaub et al. 2008). Entsprechend wurden die beiden Szenarien Flugverkehrsüberwachung und Leitwarten in der Prozessindustrie gewählt.

3.2 Hypothesen

Im Zuge der Expertenevaluation galt das Hauptinteresse der Effektstärke, welche die Teilnehmer den Persönlichkeitsmerkmalen zuschreiben: Mit welcher Höhe werden die Merkmale jeweils im Durchschnitt und im Vergleich zueinander eingeschätzt? Hier wird für den Durchschnitt erwartet, dass die Stichprobe die Effektstärke von Persönlichkeitsmerkmalen auf die Leistung von Operateuren in komplexen Mensch-Maschine-Systemen tendenziell als „gering“ einschätzt. Das bereits genannte spärliche Vorkommen an Forschungsbefunden zum Einfluss von Persönlichkeitsmerkmalen in Human Factors Untersuchungen lässt darauf schließen, dass Experten in diesem Bereich über entsprechend wenig (praktische) Erfahrung mit solchen Merkmalen verfügen und den Effekt von Persönlichkeitsmerkmalen im Durchschnitt zurückhaltend beurteilen. (Die Anwesenheit von Methodikern, die unter den Teilnehmern vermutet wird, sollte eine tendenziell als gering eingeschätzte Effektstärke deswegen untermauern, weil die Wahrscheinlichkeit, hohe Effektstärken in realitätsnahen Untersuchungen zu erzielen, aufgrund der hohen Variabilität in solchen Untersuchungen durch das Zusammenwirken von Person und Situation (hier auch gekennzeichnet durch Aufgabe und System) gering ist. Methodenkenner werden sich dementsprechend generell schwer tun, hohe Effektstärken zu unterstellen). Unser Fokus liegt auf einem globalen Vergleich der Merkmale, für den wir explorativ untersuchen, inwiefern sich die Effektstärkeneinschätzungen der einzelnen Merkmale unterscheiden, d.h. ob einige der sieben Merkmale als bedeutsamer eingestuft werden, als andere. Weiterhin soll zusätzlich explorativ überprüft werden, ob die Merkmaleinschätzung in Abhängigkeit des zugrunde liegenden Systems (Flugverkehrsüberwachung vs. Leitstand in einem Kraftwerk) oder abhängig von der zugrunde liegenden Aufgabenart (Prozessüberwachung vs. -führung) signifikante Unterschiede aufweist. Wir neh-

men weiterhin an, dass sich Personen mit unterschiedlichen Rollen im Human Factors Bereich (Forscher, Entwickler, Anwender), in der durchschnittlichen Effektstärkeneinschätzung der Merkmale signifikant unterscheiden. Einen Einfluss sollte zudem die selbstzugeschriebene Expertise im Bereich Human Factors haben: Analog der bereits genannten Argumentation bezüglich einer relativ geringeren Effektstärkeneinschätzung der Merkmale durch die Teilnehmer sollten Experten (Experten und Meister) aufgrund des ihnen zugeschriebenen Methodenwissens im Human Factors Bereich die Effektstärke von Persönlichkeitsmerkmalen geringer einschätzen als Nicht-Experten (Laien und Novizen).

- H1: Die Teilnehmer schätzen die Effektstärke von Persönlichkeitsmerkmalen auf die Prozessführung und -überwachung von Mensch-Maschine-Systemen für die Praxis im Durchschnitt als eher gering ein.
- H2: Es gibt signifikante Unterschiede zwischen den durchschnittlichen Effektstärkeneinschätzungen der 7 Persönlichkeitsmerkmale.
- H3: Die Merkmalseinschätzung variiert in Abhängigkeit des zugrunde liegenden Systems.
- H4: Die Merkmalseinschätzung variiert in Abhängigkeit der zugrunde liegenden Aufgabe.
- H5: Anwender vs. Forscher vs. Entwickler unterscheiden sich signifikant in der Merkmalseinschätzung.
- H6: Die Effektstärke von Persönlichkeitsmerkmalen wird von Nicht-Experten (Laien und Novizen) signifikant höher eingeschätzt als von Experten (Experten und Meister) im Human Factors Bereich.

3.3 Vorgehen bei der Datenerhebung

3.3.1 Beschreibung des Instrumentariums

Als Messinstrument für die Effektstärkeneinschätzung der Persönlichkeitsmerkmale durch die Experten diente ein von der Projektgruppe erstellter Fragebogen mit 45 Items¹. Im Rahmen der demographischen Angaben wurden die Items *Alter*, *Geschlecht*, *Studiertes Fach* sowie *Aktuelles Studienfach* und *Beruf* erfasst. Anschließend sollten sich die Teilnehmer bezüglich ihrer *Expertise in den drei Bereichen Human Factors, Persönlichkeitspsychologie und Inferenzstatistik* einer der folgenden vorgegebenen Antwortalternativen zuordnen: „Meister“, „Experte“, „Kenner“, „Laie“ und „Novize“ (Hoffman, Shadbolt, Burton & Klein 1995). Die Frage nach der *Rolle im Bereich Human Factors* sollte mittels der Zuordnung zu einer der 3 Antwortmöglichkeiten „Forscher“, „Entwickler“, „Anwender“, beantwortet und die Dauer ihrer *Tätigkeit in Jahren* angegeben werden. *Vorerfahrungen mit der Prozess-*

¹ Der Fragebogen enthielt zwei separate Itemblöcke: Die hier genannte Anzahl von 45 Items bezieht sich ausschließlich auf die Erfahrungen und Einschätzungen der Persönlichkeitsmerkmale für die Prozessführungs- und Überwachungsleistung aus der Sicht von *Operateuren* komplexer MMS. Ein zweiter Teil befasste sich mit der Perspektive eines *interdisziplinären Entwicklungsteams* komplexer MMS; dieser Teil erfasste Erfahrungen und Einschätzungen von Persönlichkeitsmerkmalen für die Entwicklung von Automatisierungen für komplexe, dynamische, technische Systeme und wird in einem anderen Kontext behandelt.

führung und -überwachung in Mensch-Maschine-Systemen wurden fünfstufig mit den Antwortmöglichkeiten „keine“, „wenig“, „durchschnittlich“, „viel“, „sehr viel“ erhoben; ergänzend sollte dabei jeweils die *Art des Mensch-Maschine-Systems* angegeben werden. Der nachfolgende Befragungsabschnitt enthielt die eigentliche Merkmalseinschätzung, in welcher die Effektstärke von sieben Persönlichkeitsmerkmalen (Emotionale Stabilität, Extraversion, Verträglichkeit, Gewissenhaftigkeit, Offenheit nach Rammstedt und John (2005), Leistungsmotivation nach Lang und Fries (2006) und Risikoneigung nach Nachtwei (2008); die hier dargestellte Reihenfolge entspricht der Reihung im Fragebogen) auf die Leistung von Operateuren in komplexen Mensch-Maschine-Systemen einzuschätzen waren. Diese Einschätzung erfolgte jeweils getrennt für die Prozessführungs- und überwachungsleistung und weiterhin jeweils für das Szenario einer Flugverkehrsüberwachung und das Szenario eines Leitstands in einem Kraftwerk entsprechend der Grobunterteilung von Effektstärken nach Cohen (1988) auf einer vierstufigen Ratingskala mit „kein Effekt“, „niedrig“, „mittel“ oder „hoch“. Im letzten Abschnitt Reflektion wurden die Teilnehmer gebeten, auf einer fünfstufigen Skala mit „gar nicht“, „wenig“, „mittelmäßig“, „überwiegend“, „völlig“ anzugeben, *wie schwierig sie die Evaluation generell* empfunden hatten. Ergänzend sollte *angegeben und begründet* werden, *bei welchen Merkmalen* sie diese Schwierigkeit erlebt hätten. Die Instruktion zur Bearbeitung des Evaluationsbogens erfolgte mündlich durch die Leiterin der Expertenevaluation sowie schriftlich durch ein Vorwort im Fragebogen selbst, in dem das Potenzial von Persönlichkeitsmerkmalen im Bereich Human Factors dargelegt wurde. Auf drei weiteren Instruktionseiten wurden alle vorkommenden Persönlichkeitsmerkmale, der Begriff „Effektstärke“ sowie die beiden genannten Mensch-Maschine-Szenarien „Flugverkehrsüberwachung“ und „Leitstand in einem Kraftwerk“ ausführlich erläutert. Eine Einleitung wies darauf hin, dass die Angaben der Teilnehmer vertraulich behandelt würden.

3.3.2 Beschreibung der Stichprobe

Soziodemographische Daten: Insgesamt 28 Teilnehmer der Frühjahrsschule nahmen an der Evaluation teil. Das Durchschnittsalter der Stichprobe betrug 33 Jahre (SD: 9.6). Die jüngste Person war 25 Jahre alt, die älteste 60 Jahre. Der Männeranteil betrug 42,9 %, der Frauenanteil 57,1 %. Mit 53,6 % hat der Großteil das Studium der Psychologie absolviert, 32 % haben das Studium im Bereich Technik (Ingenieurwesen, Informatik) und 7,1 % ein nicht Human Factor bezogenes Studienfach, z.B. Kommunikationswissenschaft, absolviert. Zwei Personen hatten keine Angabe gemacht. *Verteilung der Expertise:* In der Angabe zur Expertise bezeichnen sich im Bereich Human Factors 71,4 % der Stichprobe mindestens als Kenner, im Bereich der Persönlichkeitspsychologie stufen sich 46,4 %, und in der Inferenzstatistik die Hälfte der Stichprobe (50 %) mindestens als Kenner ein. Hinsichtlich der Rolle im Bereich Human Factors nimmt die deutliche Mehrheit der Stichprobe (78,6 %) die Rolle des Forschers ein. 14,3 % verstehen sich als Entwickler und 7,1 % als Anwender. Die durchschnittliche Jahresanzahl der Tätigkeit im Bereich Human Factors beträgt 5,95 Jahre (SD: 9.39). Zu beachten ist jedoch, dass der Großteil der Stichprobe keine oder nur wenige Jahre Tätigkeit in diesem Bereich vorweisen kann. 42,9 % verfügen über maximal ein Jahr Erfahrung; 1,4 % geben hier an, noch kein Jahr im Bereich Human Factors tätig zu sein. Mit der Prozessführung und -überwachung von Mensch-Maschine-Systemen haben 64,3 % der Teilnehmer noch keine und 21,4 % nur wenig Erfahrung machen können. Bezüglich der Art der Systeme nennen 17 % Systeme aus der Prozessindustrie und 10,7 % Systeme aus dem Verkehrswesen.

3.3.3 Beschreibung des Untersuchungssettings

Die Datenerhebung erfolgte lokal mit allen Teilnehmern der prometei Frühjahrschule. In einem Tagungsraum erhielten die Teilnehmer den Evaluationsbogen im Anschluss an eine kurze Einweisung in die Thematik durch die Leiterin der Evaluation. Eine ausführliche Präsentation zur thematischen Einbettung der Evaluation wurde im Anschluss an die Erhebung gehalten, um verfälschte Angaben der Teilnehmer im Fragebogen zu verhindern. Die tatsächliche Bearbeitung des Bogens durch die Teilnehmer wurde zeitlich nicht begrenzt, jedoch gemessen. Die durchschnittliche Bearbeitungszeit betrug 35 Minuten (SD: 0,05 Minuten). Die kürzeste Bearbeitungsdauer betrug 25, die längste 44 Minuten.

4. Datenanalyse

4.1 Vorgehen bei der Auswertung

Die Werte aus allen Fragebögen wurden nach Abschluss der Evaluation in Excel und SPSS überführt. Fehlende Einträge wurden in der Berechnung durch gesonderte Kodierung berücksichtigt. Mehrfachgestufte Antworten wurden kodiert. Bei zwei Items wurden die offenen Antworten Clustern zugeordnet und diese Cluster dann ebenfalls kodiert: Die Angaben zum „Studierten Fach“ wurden den Bereichen Psychologie, Technik oder einem Rest, die Nennung der Systemart wurde entweder dem Bereich „Verkehrswesen“ oder dem Bereich „Prozessindustrie“ zugeordnet. In beiden Fällen wurden die Cluster dann kodiert. Zur Überprüfung von Hypothese 1 wurden deskriptiv Mittelwerte und Standardabweichungen berechnet. Für den mit der Hypothese 2 geforderten globalen Merkmalsvergleich wurde eine einfaktorielle ANOVA für verbundene Messungen, zur Überprüfung der Hypothesen 3 und 4 zweifaktorielle ANOVAs für verbundene Messungen gerechnet. Hypothese 5 war aufgrund einer ungünstigen Verteilung der Teilnehmerzahlen auf die drei Gruppen (lediglich 2 Personen bezeichneten sich als Anwender, 4 Personen als Entwickler und 22 Personen schätzten sich als Forscher ein) inferenzstatistisch nicht prüfbar. Hypothese 6 legt aufgrund der unterschiedlichen Gruppengrößen von Experten vs. Nicht-Experten die Berechnung eines U-Tests nahe.

4.2 Ergebnisse

Die durchschnittliche Einschätzung aller Merkmale für die Prozessführungs- und Überwachungsleistung verdeutlicht Abbildung 1. Sie ist gekennzeichnet durch einen Mittelwert von 1.5 (SD = 0.65), der damit genau zwischen einer geringen und einer mittelhohen Einschätzung liegt und damit der Hypothese 1 einer tendenziell geringen Effektstärkeneinschätzung der Persönlichkeitsmerkmale auf die Leistung von Operateuren in komplexen Mensch-Maschine-Systemen entspricht. Durchschnittlich am höchsten eingeschätzt wurde das Merkmal Gewissenhaftigkeit (M = 2.4, SD = 0.77). Ihm folgen die Merkmale Risikoneigung (M = 2.2, SD = 0.75), Emotionale Stabilität (M = 1.9, SD = 0.69) und Leistungsmotivation (M = 1.6, SD = 0.79). Im niedrigen Bereich liegen die Merkmale Verträglichkeit (M = 1, SD = 0.84), Extraversion (M = 1, SD = 0.62) und Offenheit (M = 0.65, SD = 0.81).

Hinsichtlich der Frage, ob die Teilnehmer die Effektstärken der einzelnen Persönlichkeitsmerkmale für die Leistung in komplexen MMS unterschiedlich hoch ein-

schätzen (H2), weist das Ergebnis einer einfaktoriellen ANOVA für verbundene Messungen (Faktor: Merkmal) auf signifikante Unterschiede in der Merkmaleinschätzung hin. $F [6, 162] = 28.69, p < .01$. Post-hoc-Vergleiche mit Bonferroni-Korrektur ergaben, dass die Merkmale Emotionale Stabilität, Gewissenhaftigkeit und Risikoneigung in ihrem Einfluss auf die Leistung in komplexen MMS jeweils signifikant höher eingeschätzt wurden ($p < .01$) als die Merkmale Extraversion, Offenheit und Verträglichkeit. Die Effektstärke von Gewissenhaftigkeit schätzen die Teilnehmer auch gegenüber den Merkmalen Emotionale Stabilität ($p < .05$) und Leistungsmotivation ($p < .05$) als signifikant höher ein. Die Effektstärke von Leistungsmotivation wird signifikant höher eingeschätzt als die von Extraversion ($p < .05$) und Offenheit ($p < .01$).

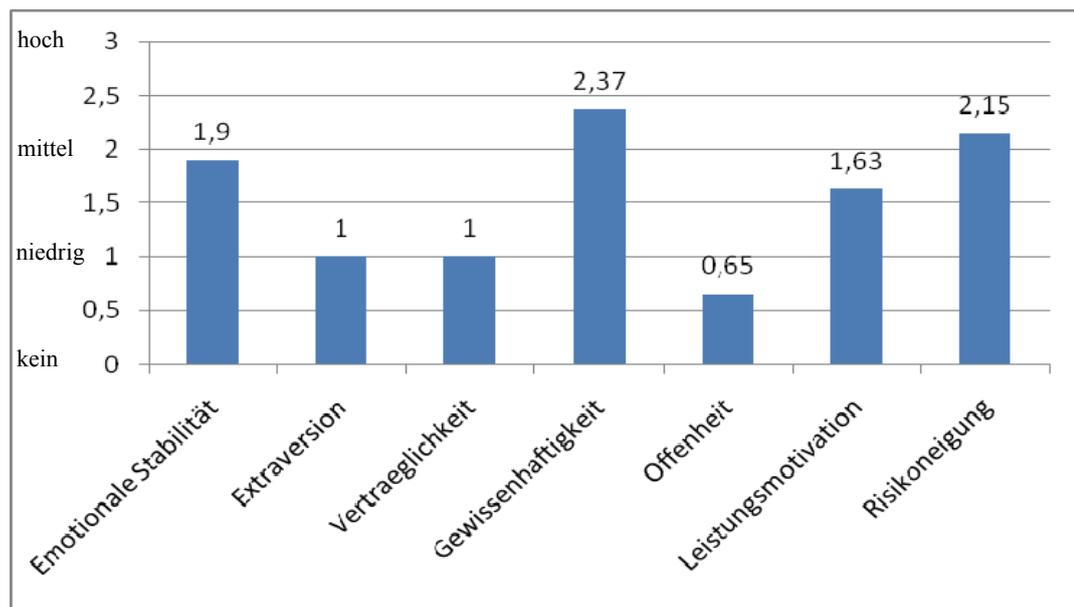


Abbildung 1: Mittlere Effektstärkeneinschätzungen von 7 Persönlichkeitsmerkmalen auf die Prozessführungs- und Überwachungsleistung von Operateuren in komplexen Mensch-Maschine-Systemen

Als erweiterter Ansatz sollte geprüft werden, ob das zugrundeliegende System (Flugverkehrsüberwachung vs. Leitstand in einem Kraftwerk (H3)) oder die zugrunde liegende Aufgabe (Prozessführung vs. -überwachung (H4)) zu einer unterschiedlichen Einschätzung der Merkmale führen würde. Dieser Ansatz basiert auf der Frage nach der Generalisierbarkeit von Systemen und Aufgaben, die ein Mensch-Maschine-System kennzeichnen. Wir nehmen an, dass es für die Betrachtung des Einflusses von Persönlichkeitsmerkmalen einen Unterschied macht, welches System und welche Aufgaben des Operators betrachtet werden. Für die Untersuchung des Einflusses durch das System bzw. die Aufgabe wurde pro Merkmal eine zweifaktorielle ANOVA für verbundene Messungen (Faktoren: Aufgabe und System) durchgeführt. (Die Mehrfachberechnung einer zweifaktoriellen ANOVA legt eine Alpha-Adjustierung nahe; die in diesem Zusammenhang herangezogenen p-Werte für das Signifikanzniveau für die Hypothesen 3 und 4 sind daher jeweils Bonferroni-korrigiert zu $p < .007$ bei sieben Einzelvergleichen²). Diese Berechnung ergab für das System ausschließlich im Merkmal Leistungsmotivation einen signifikanten

² “There are several ways in which the familywise error rate can be controlled. The most popular (and easiest) way is to divide α by the number of comparisons, thus ensuring that the cumulative Type I error is below .05. Therefore, if we conduct 10 tests, we use .005 as our criterion for significance. This method is known as the Bonferroni correction after Carlo Bonferroni” (Field 2005).

Haupteffekt mit $F [1, 3.94] = 11.42, p < .007$. Das heißt, die Einschätzung der Effektstärke von Leistungsmotivation für die Prozessführung und -überwachung durch Operateure variierte abhängig von der zugrunde liegenden Art des Mensch-Maschine-Systems (Flugverkehrsüberwachung oder Kraftwerk). Die Effektstärkeneinschätzung war dagegen bei keinem Merkmal beeinflusst von der Art der Aufgabe, das heißt, davon, ob das Merkmal für die Prozessführung vs. Prozessüberwachung eingeschätzt werden sollte; der F-Wert für den Einfluss der Aufgabe überschritt bei allen 7 Merkmalen das geforderte Signifikanzniveau von $p < .007$.

Hypothese 6 basierte auf der Annahme, dass Experten des Bereiches Human Factors die Effektstärke von Persönlichkeitsmerkmalen geringer einschätzen würden als Nicht-Experten. Ein U-Test (Mann-Whitney) konnte jedoch für kein Merkmal signifikante Unterschiede ausweisen ($p > .05$): Experten unterscheiden sich hinsichtlich der Merkmalseinschätzung nicht signifikant von Nicht-Experten im Human Factors Bereich.

5. Diskussion

Wie wir Eingangs erläuterten, kennzeichnet den Forschungsbereich Human Factors eine breite Interdisziplinarität. Entsprechend stammt das Wissen dieses Forschungsbereichs aus einer Vielzahl an Disziplinen. Für die Aneignung einer gewissen Expertise im Human Factors Bereich muss neben Fertigkeiten, Fähigkeiten und Erfahrungen auch spezifisches Wissen aus diesen verschiedenen Domänen vorhanden sein. Dazu gehört auch das Wissen um die praktische Bedeutsamkeit von menschlichen Faktoren, wie Persönlichkeitsvariablen, für die Leistung von Operateuren.

Um zu klären, ob dieses Wissen über die praktische Bedeutsamkeit von Persönlichkeitsvariablen tatsächlich unter den Human Factors Experten verbreitet ist, haben wir die Effektstärke von sieben verschiedenen Persönlichkeitsvariablen für die Leistung von Operateuren bei der Prozessüberwachung und -führung in der Flugverkehrsüberwachung und im Kraftwerk durch Personen mit unterschiedlichem Expertisegrad einschätzen lassen. Es zeigte sich insgesamt eine eher geringe Effektstärkeneinschätzung. Dies könnte die Unkenntnis über den Einfluss von Persönlichkeitsmerkmalen auf die Ergebnisse von Studien zur MMI widerspiegeln, verursacht durch die bisher geringe Anzahl an Literatur zu dieser Thematik. Ebenso könnte dies darauf hinweisen, dass im Human Factors Bereich tatsächlich ein spezifisch ausgerichtetes Expertentum vorherrscht und die Forscher hinsichtlich der Bedeutsamkeit von Persönlichkeitsmerkmalen für ihren Forschungsbereich bisher ungenügend sensibilisiert sind. Genauso könnte es daran liegen, dass 50 % der Teilnehmer Experten/Kenner der Inferenzstatistik sind und somit Wissen über die vielfältigen Einflussmöglichkeiten auf Untersuchungsergebnisse in realen oder realitätsnahen Settings besitzen. Wodurch Persönlichkeitsvariablen für die Human Factors Experten nur einen kleinen Teil der Einflussmenge ausmachen könnte. Dies wiederum könnte eine niedrige Beurteilung der Bedeutsamkeit der Persönlichkeitsvariablen zur Folge haben. In jedem Fall besteht hier ein Aufklärungsbedarf in der Human Factors Domäne.

Trotzdem wurden die einzelnen Persönlichkeitsmerkmale durch die Befragten differenziert betrachtet, d.h. ihnen wurde eine unterschiedlich hohe Bedeutsamkeit zugeschrieben. Als stärkste Einflussgrößen auf die Leistung der Operateure wurden die Merkmale Gewissenhaftigkeit, emotionale Stabilität, Risikoneigung und Leistungsmotivation eingestuft.

Bei der Prüfung, ob diese Merkmalseinschätzung abhängig von der Art des Systems oder von der vom Operateur auszuführenden Aufgabe ist, zeigte sich eine Aufgaben- und eine Systemunabhängigkeit (mit Ausnahme des Merkmals Leistungsmotivation). Die nachgewiesene systemabhängige Bewertung des Merkmals Leistungsmotivation deutet darauf hin, dass nicht alle MMI undifferenziert betrachtet werden können. Erst durch weitere Untersuchungen auch über den aktuellen Untersuchungskontext hinaus bzw. durch Untersuchungen mit höherem Stichprobenumfang und damit einer höheren Teststärke für Einzelvergleiche kann die Frage nach der Systemunabhängigkeit der Einschätzung der praktischen Bedeutsamkeit von Merkmalen aufgeklärt werden. Zudem sollte die Beschreibung der Systeme und Aufgaben zukünftig detaillierter erfolgen. Einige der Teilnehmer gaben im anschließenden Reflektionsteil des Fragebogens an, aufgrund der unzureichenden Aufklärung über die Aufgaben des Operateurs oder die Systeme zu wenig Kenntnis besessen zu haben, um die Bedeutsamkeit der Merkmale einschätzen zu können. Ebenso sollten die einzelnen Persönlichkeitsmerkmale selbst noch ausführlicher definiert werden, da einzelne Befragte die Passung zwischen Oberbegriff und Facetten oder zwischen den Facetten eines Persönlichkeitsmerkmals als zu gering empfanden.

Die Annahme, dass Experten des Human Factors Bereichs die praktische Bedeutsamkeit von Persönlichkeitsvariablen aufgrund ihrer fundierteren Kenntnis über das Konzept Effektstärke geringer einschätzen als Nicht-Experten, konnte nicht bestätigt werden. Experten unterscheiden sich hinsichtlich der Merkmalseinschätzung nicht signifikant von Nicht-Experten im Human Factors Bereich. Die deskriptive Auswertung ergab für die Expertengruppe, dass 14 von 20 Human Factors Experten sich auch als Experten im Bereich Inferenzstatistik bezeichnet hatten. Auf der anderen Seite hatten sich alle 8 Nicht-Experten im Bereich Human Factors auch als Nicht-Experten im Bereich Inferenzstatistik eingestuft. Dass die Effektstärkeneinschätzungen der Persönlichkeitsmerkmale sich zwischen beiden Gruppen trotzdem nicht unterscheidet, könnte auf unterschiedliche Gründe für dieselbe Einschätzung hinweisen: Die Experten haben aufgrund ihres Methodenwissens eine geringe Einschätzung vorgenommen; die Nicht-Experten haben aufgrund ihres fehlenden fachlichen Wissens im Bereich Human Factors die Effektstärken der Persönlichkeitsmerkmale aus Zurückhaltung gering eingeschätzt. Dieser Schluss sollte jedoch ebenfalls in vertiefenden Forschungsansätzen und mit größeren Stichprobenumfängen überprüft werden.

Die Ergebnisse dieser Arbeit weisen darauf hin, dass zukünftig in der Human Factors Forschung mehr Augenmerk auf die Bedeutsamkeit von Persönlichkeitsvariablen für MMI gelegt werden muss. Persönlichkeitsvariablen sollten in Studien der Human Factors Forschung als Kontrollvariablen mit erhoben werden. Um dies in der Praxis als gängigen Ansatz zu verbreiten, bedarf es aber noch einiger missionarischer Tätigkeiten. Dieser Beitrag soll der Sensibilisierung der Forscher im Bereich Human Factors für die praktische Bedeutsamkeit von Persönlichkeitsvariablen für die MMI dienen.

Danksagung

Wir danken den Teilnehmern der Frühjahrsschule 2009 für ihre Teilnahme an der Expertenevaluation.

Literatur

- Asendorpf, J.B. (1999). *Psychologie der Persönlichkeit* (2. Auflage). Berlin: Springer.
- Badke-Schaub, P., Hofinger, G. & Lauche, K. (Hrsg.) (2008). *Human Factors. Psychologie sicheren Handelns*. Heidelberg u.a.: Springer
- Booth-Kewley, S. & Vickers, R.R. (1994). Associations between major domains of personality and health behavior. *Journal of Personality*, 62, 281–298.
- Bortz, J. (2005). *Statistik: Für Human- und Sozialwissenschaftler* (6. Auflage). Berlin: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Auflage). Berlin: Springer.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Fehr, T. (2006): „Big Five“: Die fünf grundlegenden Dimensionen der Persönlichkeit und ihre dreißig Facetten. In: W. Simon (Hrsg.), *Persönlichkeitsmodelle und Persönlichkeitstests: 15 Persönlichkeitsmodelle für Personalauswahl, Persönlichkeitsentwicklung, Training und Coaching*. S. 113-135. Gabal: Offenbach.
- Farrington-Darby, T. & Wilson, J.R. (2006). The nature of expertise: A review. *Applied Ergonomics*, 37, 17-32.
- Field, A. (2005). *Discovering statistics using SPSS*. London: Sage.
- Haerem, T. & Rau, D. (2007). The Influence of Degree of Expertise and Objective Task Complexity on Perceived Task Complexity and Performance. *Journal of Applied Psychology*, 92(5), 1320-1331.
- Hoffman, R., Shadbolt, N., Burton, A. & Klein, G. (1995). Eliciting knowledge from experts: a methodological analysis. *Organizational Behaviour. Human Decision Making*, 62, 129–158.
- Huber, S., Kain, S. & Nachtwei, J. (2008). Effekte sicherer nachweisen: Persönlichkeitsmerkmale als Kontrollvariablen in der Human Factors Forschung. In M. Grandt & A. Bauch (Hrsg.), *Beiträge der Ergonomie zur Mensch-System-Integration*, 143-159. Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt e.V. (DGLR-Bericht 2008).
- Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods*, 7, 178-193.
- Kain, S. & Nachtwei, J. (in Druck). Die Rolle von Kontrollvariablen in der Human Factors Forschung - Ein bewährtes Konzept in einem modernen Anwendungsfeld. In: B. Krause & P. Metzler (Hrsg.). *Empirische Evaluationsmethoden Band 12*, ZeE Verlag: Berlin.
- Lang, J.W.B. & Fries, S. (2006). A Revised 10-Item Version of the Achievement Motives Scale - Psychometric Properties in German-Speaking Samples. *European Journal of Psychological Assessment*, 22 (3), 216-224.

- Meister, D. (2001). Basic Premises And Principles Of Human Factors Measurement. *Theoretical Issues in Ergonomics Science*, 2(1), 1 - 22.
- Nachtwei, J. (2008). *Testhandbuch zum Fragebogen zur Erfassung von Strategie und Expertise in Experimenten, revidierte Fassung (FESE-R)*. Humboldt-Universität Berlin: unveröffentlichtes Manuskript.
- Ostendorf, F. & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar nach Costa und McCrae: NEO-PI-R.*, 11-68. Göttingen u.a.: Hogrefe.
- Rammstedt, B. & John, O.P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203-212.
- Salas E. (2008). At the turn of the 21st century: reflections on our science. *Human Factors*, 50(3), 351-353.
- Schmidtke, H. (2002). Vom Sinn und Unsinn der Messung psychischer Belastung und Beanspruchung. *Zeitschrift für Arbeitswissenschaft*, 56, 4-9.
- Swain, A. D. & Guttman, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. Washington, DC: U.S. Nuclear Regulatory Commission.
- Yeo, G. & Neal, A. (2004). A multilevel analysis of effort, practice and performance: Effects of ability, conscientiousness and goal orientation. *Journal of Applied Psychology*, 89, 231-247.

Anhang

1. Auszug Instruktionen:

Vorwort

Dieser Evaluationsbogen wurde im Rahmen des seit April 2008 bestehenden Projekts „Persönlichkeitsmerkmale als Kontrollvariablen in der Human Factors Forschung“ (PeKoHF) erstellt von Charlotte Meyer, Saskia Kain und Jens Nachtwei. Das Projekt ordnet sich ein in das Graduiertenkolleg *prometei* (am Zentrum Mensch-Maschine-Systeme der Technischen Universität Berlin). Es setzt an empirischen Befunden an, die zeigen, dass *Persönlichkeitsmerkmale die Leistung von Operateuren bei der Prozessüberwachung und -führung von komplexen Systemen beeinflussen* können. In der Forschung wird diesen Tatsachen bisher dennoch wenig Beachtung geschenkt, unter anderem auch deshalb, weil die Erfassung von Persönlichkeitsmerkmalen zur Aufklärung zusätzlicher Varianz häufig mit zeitaufwendigen Fragebögen assoziiert wird. Das Projekt hat sich daher zum Ziel gesetzt, den Einsatz valider Kurzskalen (zur Erfassung von Persönlichkeitsmerkmalen) zur Verbesserung der Datenqualität in empirischen Untersuchungen zu etablieren (Huber, Kain und Nachtwei 2008). Als ein Baustein des Projekts wird dabei auch die Frage verfolgt, wie Human Factors Experten die Effektstärken von Persönlichkeitsmerkmalen bei der Vorhersage der Leistung Operateuren beurteilen. Das heißt konkret, sieben **unterschiedliche Persönlichkeitsmerkmale** sollen aus Expertensicht in den folgenden **zwei Szenarien** in ihrer Relevanz für die Praxis und die Forschung eingeschätzt und die Ergebnisse dann mit den Befunden aus der Literatur und den Ergebnissen eigener empirischer Untersuchungen verglichen werden:

1. Prozessüberwachung und -führung in der Flugverkehrsüberwachung durch Operateur
2. Prozessüberwachung und -führung im Kernkraftwerk durch Operateur

Sie sind dafür unsere Experten! Wir werden Sie im Folgenden bitten, die jeweiligen Persönlichkeitsmerkmale nach ihrer Wichtigkeit für die Leistung von Operateuren einzuschätzen. Dazu sollen Sie jeweils die **Höhe der Effektstärke** angeben, die Sie aufgrund Ihrer Expertise und Erfahrung bei einem Merkmal für die Leistung eines Entwicklers bei der Konzeption von Assistenzsystemen und Automaten für die Prozessüberwachung und -führung respektive die Leistung eines Operateurs bei der Prozessüberwachung und -führung in dem jeweiligen Szenario vermuten.

2. Beschreibung der Szenarien im Evaluationsbogen:

Szenario 1: Operateure der Flugverkehrsüberwachung. An einem Fluglotsenarbeitsplatz arbeiten in der Regel 2 Lotsen (Operateure), der Radarlotse und der Koordinationslotse. Sie müssen ihre Arbeit aufeinander abstimmen. In solchen Situationen gemeinsamer Aufgabenbewältigung ist geteiltes Situationsbewusstsein entscheidend, welches das Team dazu befähigt, seine Handlungen zu koordinieren und sich an neue Situationen anzupassen. Eine Hauptaufgabe der Flugverkehrsüberwachung bzw. der Fluglotsen besteht vor allem in der teilweise simultanen Überwachung bzw. in dem Vergleich unterschiedlicher Kurse und Flugbahnen mittels entsprechender Bild-

schirme. Weiterhin ist die Kommunikation mit unterschiedlichen Cockpitbesatzungen, mit dem Management der eigenen Fluglinie und evtl. mit technischen Experten zu nennen. Durch Nutzung von Instrumenten, Vorschriften und Dokumenten, die dazu eingesetzt werden, zunächst unvollständige oder inakkurate Information zu validieren, muss ein an die jeweilige Situation angepasstes Situationsbewusstsein entwickelt werden. (Badtke-Schaub et al. 2008)

Szenario 2: Operateure der Leitstelle eines Kraftwerkes. In einer Leitstelle eines Kraftwerkes liegt eine fast unüberschaubare Menge an Informationen vor, die potenziell bedeutsam sein können und somit Teil des geteilten Situationsbewusstseins sein müssen (nach Badtke-Schaub et al. 2008). Kernkraftwerke sind komplexe Anlagen, ausgestattet mit zahlreichen mechanischen und elektronischen Systemen. Schon bei ihrer Entwicklung werden mögliche Schadensfälle einkalkuliert, um größere Gefährdungen auszuschließen (wie das Vermeiden von Überbeanspruchungen durch Sollbruchstellen/Sicherungen). Operateure in einem Kraftwerk müssen unterschiedliche Systeme (z.B. Reaktorregelung inkl. Überwachung des Reaktorkühlkreislaufes) bei funktionsfähigem Zustand regeln (Normalbetrieb), Fehlfunktionen von Anlagenteilen oder Systemen (Anormaler Betrieb) und Zustandsveränderungen bzw. Störfälle (z.B. Druck- oder Volumenveränderungen) erkennen und erforderliche Maßnahmen in ihrem jeweiligen Aufgabenbereich einleiten (Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit 1996).

3. Beispiel für die verwendete Definition eines der Persönlichkeitsmerkmale im Evaluationsbogen:

Emotionale Belastbarkeit

Bedeutung:

Jemand ist selbstsicher, gelassen, unbesorgt und kommt bei Stress nicht leicht aus dem Gleichgewicht (Ostendorf & Angleitner 2004).

Facetten (Fehr 2006):

Unbesorgtheit (furchtlos, entspannt)

Gelassenheit (ruhig, nicht leicht reizbar)

Optimismus (hoffnungsvoll, zuversichtlich)

Unbefangenheit (selbstsicher, ungezwungen, soziales Wohlbefinden)

Beherrschtheit (fähig, seine Begierden zu kontrollieren)

Stressresistenz (unabhängig, ruhig)

4. Verwendete Definition von Effektstärke im Evaluationsbogen:

Die Effektstärke ist ein Schätzer für die Größe des Effekts (Zusammenhang/ Unterschied), in diesem Fall von Persönlichkeitsmerkmalen und Leistung des Entwicklers/ Operateurs. Sie wird berechnet, um einen einheitlichen Maßstab zu erzeugen. Mit diesem können Unterschiede (Differenzen, Verhältnisse) oder Zusammenhänge aus unterschiedlichen Untersuchungen mit unterschiedlichen Stichproben und Messinstrumenten verglichen werden. Damit stellt die Effektstärke (oder Effektgröße) ein Maß für die Bedeutung bzw. die praktische Relevanz der Unterschiede oder Zusammenhangsmaße dar (Bortz 2005).