Aspects of Personality in highly automated Human-Machine-Teams - Development of a Questionnaire

Solveig Eschen-Léguedé[1], Katja Knappe[1] and Doris Keye[1]

Personality, Automation, Human-Machine-Interaction, Hybrid Team

Abstract

When selecting aviation personnel (air traffic controllers and pilots), cognitive abilities as well as the ability to work together with other people are taken into account. It is reasonable that working successfully in highly automated Human-Machine-Interfaces (HMI) of the future, i.e. in a "hybrid team", demands different aspects of personality and attitudes than working with a human partner. Thus the German Excellence Cluster project HYBRID is aimed at the development of a personality questionnaire which - for the first time in personality research - specifically concentrates on traits and aspects of attitude which could be relevant to "hybrid teams". Two studies encompassing each about 400 applicants for air traffic control or cockpit led to the development of an extensive research questionnaire. Broad factors of personality (Big Five) are complemented by more specific scales such as Technology Affinity, Computer Literacy, Complacency, Need for Teamwork, or Attitude towards Automation. At the same time, the research simulation HINT (Hybrid Interaction) was developed to provide a tool which can measure performance in a highly automated HMI. Work on a short version of the comprehensive research questionnaire continues.

Introduction

Working procedures in aviation become more and more automated. The European air traffic management (ATM) modernisation programme SES (Single European Sky) envisages the implementation of new automated functions in ATM. According to the SESAR's Concept of Operation "humans (with appropriate skills and competences and duly authorised) will constitute the core of the future European ATM System's operations. However, […] an advanced level of automation will be required. […] The nature of human roles and tasks within the future system will necessarily change." (SESAR Consortium, 2007).

At the same time the ability to automate ATM processes is limited and ATM will therefore continue to be a human centric process in which the responsibility and the authority for the negotiation will continue to rest on human controllers and pilots (Eißfeldt et al., 2009). This generates a situation in which future working procedures have to be performed by a human and an automated system in close interaction. Such collaboration can be understood as teamwork between human and machine: a "hybrid team" is formed (Eißfeldt, 2008). First ideas on this topic have been made by Hollnagel and Woods (1983) which stated that "through the increasing sophistication of computer applications, the man-machine interface is gradually becoming the interaction of two cognitive systems." In the literature also other terms are used, for example "human-agent teams" (Deshmukh et al., 2008) or "human-robot teams" (Hoffman & Breazeal, 2004) which all try to characterise the close collaboration of humans and machines.

The need of "hybrid teamwork" can lead to a change of ability requirements of future aviation personnel. The usual kind of collaboration between air traffic controllers (radar controller and coordinator) and pilots (pilot and co-pilot) is expected to decrease. Instead, the ability to work as the human part of a human-machine-interface (HMI) becomes more and more important. Considering these developments, the German Aerospace Center (DLR) launched a project

[1] Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Abteilung Luft- und Raumfahrtpsychologie

called HYBRID. The HYBRID project focuses on investigating the requirements on humans' personality that may change with regard to "hybrid teamwork" in cooperation with Deutsche Lufthansa AG and Deutsche Flugsicherung GmbH. HYBRID aims at surveying, if working in a "hybrid team" will demand a different profile of personality and attitude than working with a human partner. Research concerning the influence of personality on performance in "hybrid teams" is necessary to be prepared for the future. Because "although it may be difficult to anticipate how automation will affect a job, it is advantageous to anticipate job changes well in advance so that appropriate selection criteria can be identified and implemented in at the same time as operational versions of automated systems" (Manning & Broach (1992).

Method

HYBRID is organised in two main research tasks. The first task is the development of a research questionnaire which gathers information on traits and aspects of attitude that could become relevant in "hybrid teams". The second task involves the development of a research scenario which simulates anticipated demands on operators of future HMIs in aviation. After having established questionnaire and scenario, both tools will be combined in an experiment to investigate which aspects of personality are able to predict performance in highly automated working environments of "hybrid teams". This article includes two studies concerning the development of the research questionnaire and gives an outlook on the research scenario.

Development of the Research Questionnaire

In a literature review various suitable aspects of personality and attitude were identified. Established questionnaires such as for example NEO-PI-R (Ostendorf & Angleitner, 2004) or BIP (Hossiep et al., 2003) were taken into account as well as less common scales to cover more specific areas of personality and attitude which may play a role for the suitability of applicants for "hybrid teams". A list of all used questionnaires is summarized in the results section. A first version of the questionnaire was prepared for administration on a DLR in-house computer based testing system (CAT, Computer Assisted Testing). This ensured the implementation of the questionnaire in the regular selection test program for Air Traffic Control (ATC) and cockpit applicants. Using CAT, the questionnaire can be fully completed by touch-screen input, avoiding time consuming paper pencil versions and allowing computerized efficient scoring. In addition, missing values are avoided since finishing the questionnaire is only possible after all items are answered.

Subjects

403 applicants for pilot or ATC training participated in the first questionnaire-study. Testing took place in June and July 2010 at the German Aerospace Centre (DLR) in Hamburg. 198 candidates were applying for pilot training, 205 for ATC training. 67% of the candidates were male, 33% were female. 81% were aged 18-21 years, 16% were aged 22-25 years, and 3% were aged 26-28 years. Participation was voluntary. Applicants were asked to take part in a study called "Demands on Future Aviation Personnel". The participants worked on the questionnaire after finishing the regular test program of the first level of the selection procedure. 48 applicants could be tested at a time and duration of testing varied between 90 and 120 minutes including a break.

Based on the results of the first study, the questionnaire was edited by readjusting items according to item analysis. Additionally, a short version of the questionnaire was developed. A second study was conducted utilizing the edited questionnaire encompassing a new sample of 409 applicants (194 applicants for pilot training and 215 applicants for ATC training). The second study took place in March 2011 under parallel conditions as the first but with duration

of 60 minutes including a break. 76% of the candidates were male, 24% were female. 89% were aged 18-21 years, 9% were aged 22-25 years, and 2% were aged 26-28 years.

Analyses

Item analysis regarding reliability, item difficulty and item discrimination were performed with data from the first study. Based on these results, items were selected displaying sufficient variance and adequate item characteristics.

Analyses of the data from the edited questionnaire also included an item analysis. The results of the item analysis supplemented the evaluation of the short version of the questionnaire. The short version questionnaire was correlated with the complete scales of the edited questionnaire.

Results

Table 1 shows an overview of the used questionnaires in study 1 and study 2. Most are established scales and were used in accordance with the authors. Only the last three scales (BnT, EgA and BIO) are in-house developments.

Table 1: Overview of the used scales in study 1 and study 2

Test Initial	Test Name	Test Authors
BFI-10	Big-Five-Inventory-10	Rammstedt, B. & John, O.P. (2007)
BIP	Bochum Inventory of job-related Personality	Hossiep, R. & Paschen, M. (1998).
CFQ	Cognitive Failure Questionnaire	Lumb, P.L.K. (1995)
CNFB	Computer Usage Questionnaire	Schroeders, U. & Wilhelm, O. (2010)
FEcA	Computer-specific Attribution Questionnaire	Dickhäuser, O. & Stiensmeier-Pelster, J. (2000)
FIT	Individual Attitude towards Teamwork Questionnaire	Mohiyeddini, C. (2001).
INCOBI-R	Computer Literacy Inventory	Richter, T., Nauman, J. & Horz, H. (2010)
KUT	Locus of Control when Interacting with Technology	Beier, G. (2004)
NEO-PI-R	Revised NEO Personality Inventory, German Version	Ostendorf, F. & Angleitner, A. (2004)
SWE	Generalized Self-Efficacy scale	Schwarzer, R. & Jerusalem, M. (1995)
TA-EG	Technology Affinity - Electronic Devices	Karrer, K., Glaser, C., Clemens, C. & Bruder, C. (2009)
CaP	Complacency as Potential	Feuerberg, B., Bahner, E. & Manzey, D. (2005)
BnT	Need for Teamwork	HYBRID Project
EgA	Attitude towards Automation	HYBRID Project
BIO	Biographical Questionnaire	HYBRID Project

Analyses of the first questionnaire resulted in satisfying characteristics for almost all administered items and scales. Apart from very few items, participants used the whole range of the scale from 1 ("strongly disagree") to 5 ("strongly agree"). Standard deviation, difficulty of items, item discrimination, and internal consistency (Cronbach's α) showed acceptable till good values. An exception constituted some scales from the INCOBI-R and the complete FEcA which had to be eliminated because they did not meet the methodological requirements. Concerning the FEcA, it appeared that participants misunderstood the items systematically leaving the data uninterpretable. A reason for the failure could rest in the response format of the questionnaire that was originally designed for paper pencil testing and proved inapplicable for administration on CAT.

The second study aimed at replicating the findings from study 1. In addition, three scales were analysed that had not been utilized in the first study. Two of the new scales, both belong-

ing to CaP, showed low internal consistency and need to be reworked before being used in future experiments of the project. Table 2 displays all internal consistencies (Cronbach's α) of the scales from study 1 and 2.

Table 2: Internal consistency (Cronbach's α) of the used questionnaires in study 1 and study 2

Scale	Cronbach's α (Study 1)	Cronbach's α (Study 2)
BIP: Flexibility	.74	.77
BIP: Teamorientedness	.84	.84
CFQ	.82	.84
CNFB	.86	.86
FIT: Disposition for Teamwork	.66	.72
FIT: Reservation towards Teamwork	.59	.53
INCOBI-R: Self-confidence in Using Computers	.85	.80
INCOBI-R: Computer Related Attitude, Scale B	.83	.82
INCOBI-R: Computer Related Attitude, Scale F	.84	.83
INCOBI-R: Computer Related Attitude, Scale G	.78	.72
INCOBI-R: Computer Related Attitude, Scale H	.84	.79
KUT	.89	.87
SWE	.73	.79
TA-EG: Enthusiasm about Technology	.78	.78
TA-EG: Subjective Competency in using Technology	.74	.75
TA-EG: Negative Impacts of Technology	.71	.71
TA-EG: Positive Impacts of Technology	.64	.69
CaP: Uncertainty and Risk Tolerance		.45
CaP: Trust in Technology		.51
BnT		.77
EgA		.70

Additionally, a short version of the research questionnaire encompassing 50 of the original 200 items was developed in order to safe time, cost, and applicant's effort. For this purpose, one up to six items (depending on the length of the original scale) had been chosen to represent each scale. Some of these items are original items of a scale and some combine the information of several items. The part-whole correlations of the short scales with the full scales were all highly significant but varied from .22 to .59.

Outlook: Research Scenario

The research scenario HINT (Hybrid Interaction) was conceptualised based on the anticipation of "hybrid teamwork". The purpose of HINT is the simulation of relevant processes in a future interaction of a human operator and an automated system in aviation. HINT is a possible anticipation of the future but cannot claim to be an exact forecast of the future reality. However, HINT tries to display simplified central requirements on future operators. Among other, these requirements imply system monitoring or "operational monitoring" (Eißfeldt et al., 2009) and exception or request handling.

The conceptual design of the research scenario HINT incorporates a simplified simulation of air traffic in two separated sectors. Participants take the role of the operator and have to manage the air traffic in the best possible way while being supported by automatic mechanisms distributing the air traffic in their sector. Both sectors are supposed to fulfil their target values as good as possible (see Figure 1). With automated functions, the operator has to stay in the loop to be able to react appropriately on requests of the other partner sector as well as putting own requests to optimise the target values. The design of the HINT includes the second sector being either handled by another human operator or by a fully automated system.

Performance in HINT can be measured by various variables, for example correct/false inputs, response time, fulfilment of target values etc.

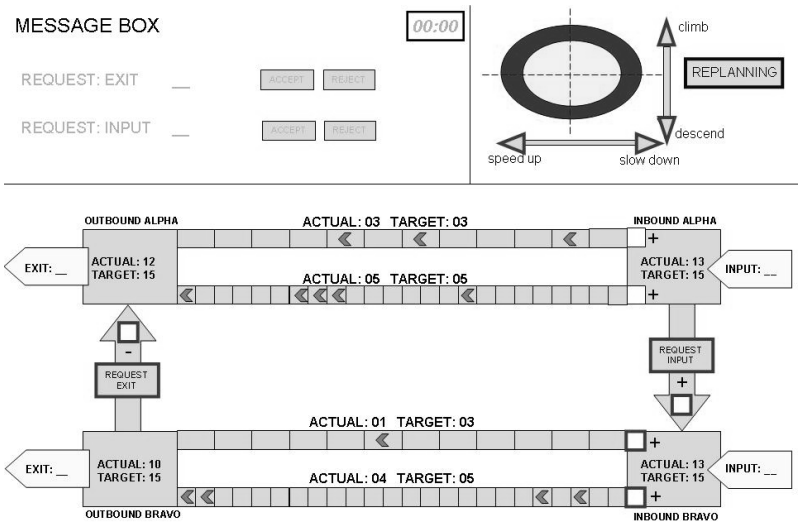

Figure 1: HINT Simulation

Discussion and Conclusion

The results of the questionnaires are a promising first step to provide a basis for a solid and comprehensive final version of a research questionnaire that can be used in the forthcoming experiments of the HYBRID project. For the measurement of personality traits and attitude scales, an adequate number of items fulfil the methodological requirements, i.e. a broad variety of traits can be included and appropriately measured.

For the first time, a personality questionnaire was specifically tailored to traits which can be relevant in "hybrid teams". Next steps include the survey of the questionnaire's predictive validity of performance in "hybrid teams", i.e. if aspects of personality predict performance differences in "hybrid teams". Before that, pilot studies with the research scenario have to be conducted to determine influential experimental variables and adjusting its difficulty. Tying up with the research questionnaire, the combination of both tools in an experiment will be the core of the HYBRID project in order to measure personality's influence on performance in HINT.

An additional field of work is the short version of the research questionnaire, because the present version does still not meet methodological requirements. The correlations between the short scales and the complete scales are to low to be able to represent the complete questionnaire. Therefore, the first experiments with the research questionnaire and the research scenario will include the complete version only. However, the development of a short version will continue.

Provided that the research on scenario and questionnaire establishes personality aspects relevant for performance in "hybrid teams", the future selection of aviation personnel or personnel for other workplaces which include "hybrid teamwork" will be in need of adequate and efficient measurement tools. A shortened questionnaire could meet these requirements in the future.

Literature

Beier, G. (2004). *Kontrollüberzeugungen im Umgang mit Technik: Ein Persönlichkeitsmerkmal mit Relevanz für die Gestaltung technischer Systeme.* Berlin: dissertation.de.

Deshmukh, A.V., McComb, S.A. & Wernz, Ch. (2008). Agents as Collaborating Team Members. In: Letsky, M.P., Warner, N.W., Fiore, S.M.& Smith, C.A.P. (Eds.), *Macrocognition in Teams.* Aldershot: Ashgate

Dickhäuser, O. & Stiensmeier-Pelster, J. (2000). Entwicklung eines Fragebogens zur Erfassung computerspezifischer Attributionen. *Diagnostica*, 46 (2), 103-111.

Eißfeldt, H. (2008): Projektantrag Hybrid: *Entscheidungsverhalten in hybriden Teams*. (unveröffentlichtes Dokument). Hamburg: DLR

Eißfeldt, H., Grasshoff, G., Hasse, C., Hoermann, H.-J., Schulze Kissing, D., Stern, C., Wenzel, J. & Zierke, O. (2009). *Aviator 2030 - Ability Requirements in Future ATM Systems II: Simulations and Experiments. DLR-FB-2009-28*. Köln: DLR

Feuerberg, B.V., Bahner, J.E. & Manzey, D. (2005). Interindividuelle Unterschiede im Umgang mit Automation - entwicklung eines Fragebogens zur Erfassung des Complacency-Potentials. In: L. Urbas & C. Steffens (Hrsg.), *Zustandserkennung und Systemgestaltung. 6. Berliner Werkstatt Mensch-Maschine-Systeme.* (Reihe 22, Fortschrittberichte VDI, Nr. 22) (S. 199-202). Düsseldorf: VDI.

Hoc, J.-M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. *International Journal of Human–Computer Studies*, 54(4), 509–540.

Hoffman, G. and Breazeal, C. (2004). Collaboration in Human–Robot Teams. In *Proc. AIAA 1st Intell. Syst. Tech. Conf.*, Chicago, IL, Sep. 2004.

Hollnagel, E. & Woods, D. D. (1983). Cognitive Systems Engineering: New Wine in New Bottles. *International Journal of Man-Machine Studies*, 18, 583}600.

Hossiep, R., Paschen, M. & Mühlhaus, O. (2003). Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (BIP). Göttingen: Hogrefe.

Karrer, K., Glaser, C., Clemens, C. & Bruder, C. (2009). Technikaffinität erfassen – der Fragebogen TA-EG. In: A. Lichtenstein, C. Stößel und C. Clemens (Eds.), *Der Mensch als Mittelpunkt technischer Systeme. 8. Berliner Werkstatt Mensch-Maschine-Systeme.* (ZMMS Spektrum, Reihe 22, Nr. 29, pp. 196-201). Düsseldorf: VDI Verlag GmbH.

Lumb, P.L.K. (1995). Cognitive failures and performance differences: validation studies of a German version of the cognitive failures questionnaire. *Ergonomics*, 38 (7), 1456 - 1467.

Manning, C. A., & Broach, D. (1992). Identifying ability requirements for operators of future automated Air Traffic Control Systems. In Office of Aviation Medicine (Eds.), *DOT/FAA/AM-92/26.* Washington, D.C.: FAA.

Mohiyeddini, C. (2001). Fragebogen zur Erfassung individueller Einstellungen zur Teamarbeit (FIT). In W. Sarges & H. Wottawa (Eds.), *Handbuch wirtschaftspsychologischer Testverfahren.* Lengerich: Pabst Science Publishers.

Ostendorf, F. & Angleitner, A. (2004). *NEO-Persönlichkeitsinventar nach Costa und McCrae: NEO-PI-R*. Göttingen: Hogrefe.

Richter, T., Naumann, J., & Horz, H. (2010). *Eine revidierte Fassung des Inventars zur Computerbildung (INCOBI-R).* Zeitschrift für Pädagogische Psychologie, 24, 23-37.

Schroeders, U., & Wilhelm, O. (2010). Computer usage questionnaire: Structure, correlates, and gender differences. Conditionally accepted for publication in *Computers in Human Behavior*.

Schwarzer, R. & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright & M. Johnston (eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON.

SESAR Consortium (2007): The ATM Target Concept. In *DLM-0612-001-02-00a.* Brussels: European Commission.