Christopher Marc Schlick · Sönke Duckwitz
Frank Flemisch · Martin Frenz
Sinem Kuz · Alexander Mertens
Susanne Mütze-Niewöhner *Eds.*

Advances in Ergonomic Design of Systems, Products and Processes

Proceedings of the Annual Meeting of GfA 2016

Springer

Advances in Ergonomic Design of Systems, Products and Processes

Christopher Marc Schlick • Sönke Duckwitz •
Frank Flemisch • Martin Frenz •
Sinem Kuz • Alexander Mertens •
Susanne Mütze-Niewöhner
Editors

Advances in Ergonomic Design of Systems, Products and Processes

Proceedings of the Annual Meeting of GfA 2016

Springer

Editors
Christopher Marc Schlick
RWTH Aachen
Aachen, Nordrhein-Westfalen
Germany

Sönke Duckwitz
RWTH Aachen
Aachen, Nordrhein-Westfalen
Germany

Frank Flemisch
RWTH Aachen
Aachen, Nordrhein-Westfalen
Germany

Martin Frenz
RWTH Aachen
Aachen, Nordrhein-Westfalen
Germany

Sinem Kuz
RWTH Aachen
Aachen, Nordrhein-Westfalen
Germany

Alexander Mertens
RWTH Aachen
Aachen, Nordrhein-Westfalen
Germany

Susanne Mütze-Niewöhner
RWTH Aachen
Aachen, Nordrhein-Westfalen
Germany

ISBN 978-3-662-53304-8 ISBN 978-3-662-53305-5 (eBook)
DOI 10.1007/978-3-662-53305-5

Library of Congress Control Number: 2016960326

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer-Verlag GmbH Germany
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Editorial

The rising level of digitalization of products, processes, and services and the interconnecting computer networks are the driving factors for the increasing complexity of work systems. However, digital networked systems are not only causes of increasing complexity but at the same time they also enable us to individualize the work processes and adopt the job to fit the worker through deliberate research activities. We are just beginning to explore the potential of these new technologies in order to incorporate them in the design and evaluation of work systems.

These proceedings include a selection of papers presented at the 62nd Annual Meeting of the German Human Factors and Ergonomics Society (Gesellschaft für Arbeitswissenschaft GfA), held at the Institute of Industrial Engineering and Ergonomics of RWTH Aachen University from March 2 to 4, 2016. The conference featured more than 200 presentations and 36 posters reflecting the diversity of subject matter in the field of ergonomics, human factors, and industrial engineering. The book provides a wide range of thematic areas and deals with aspects of complexity in the design of work systems, considers digitalization in the design of products, and takes role of networks in the design of processes into account.

The contributions in this book address research questions on the interindividual as well as on the individual level. The contributions at the interindividual level regard work processes within the overall context of society as well as issues at the organizational and team level. Furthermore, contributions on the individual level deal with the characteristics of work activities and technical fundamentals of workplace design.

Considering the wide range of topics covered and the variety of scientific methods applied, it is apparent that advances in the field of human factors and ergonomics may only be achieved by a multidisciplinary approach. Thus, these proceedings address human factors and safety specialists, industrial engineers, work and organizational psychologists, specialists in occupational medicine, as well as production planners and design engineers.

The contents of this edited volume required the dedicated effort of many people. We would like to thank the authors, whose research and development efforts are published here.

Aachen, Germany

Christopher Marc Schlick
Sönke Duckwitz
Frank Flemisch
Martin Frenz
Sinem Kuz
Alexander Mertens
Susanne Mütze-Niewöhner

Contents

Best-Practice Approach for a Solution-Oriented Technology Assessment: Ethical, Legal, and Social Issues in the Context of Human-Robot Collaboration

Jochen Nelles, Susanne Kohns, Julia Spies, Christina Bröhl, Christopher Brandl, Alexander Mertens, and Christopher Marc Schlick

Abstract

Future robots will process work tasks with a high degree of complexity even for small batch sizes in collaboration with the working person, simultaneously, and within a specified workplace. Due to the transformation from robots that are spatially and temporally separated from the working person and are programmed to execute tasks in a deterministic manner towards collaborative, adaptive lightweight robots, ethical, legal, and social implications (ELSI) should be considered. The purpose of this contribution is to determine how changing human-robot collaboration impacts technology. To this end potential ELSI problems and their possible causes and effects are identified and quantitatively analyzed based on the Aachen Model of Identification, Classification and Analysis of ethical, legal, and social Implications (AMICAI). Furthermore, the impacts of technology are identified and evaluated alongside potential risks, opportunities, and potentials of human-robot collaboration. This best-practice approach describes the results of applying AMICAI based on expert workshops focusing on the application example of a human-robot collaborative workplace in manufacturing.

Keywords

Technology assessment • Ethical, legal, and social implications • ELSI • MEESTAR • AMICAI • Human-robot collaboration • Human-robot interaction • Workshop • Best-practice

This article is an extended version of the first publication by Nelles et al. (2016)

J. Nelles (✉) • S. Kohns • J. Spies • C. Bröhl • C. Brandl • A. Mertens • C.M. Schlick
Institute of Industrial Engineering and Ergonomics, RWTH Aachen University, Bergdriesch 27, 52062 Aachen, Germany
e-mail: j.nelles@iaw.rwth-aachen.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_1

1 Introduction

Recent developments in the robotics industry lead away from robots as part of highly automated manufacturing facilities for large batch sizes, through to collaborative (light-weight) robots in individualized manufacturing processes. Previously robots were spatially and temporally separated from the working person. Robots under development now are placed in a collaboration area together with the working person. This collaboration area is an area in which the robot and the working person can perform manufacturing tasks simultaneously (DIN EN ISO 10218-2). Working persons and robots work in immediate proximity and perform a work task together. This change requires ethical, legal, and social questions to be considered.

To approach these questions in a best-practice approach, (1) basic factors of technology ethics are mentioned, (2) the need to involve all stakeholders is highlighted, (3) methodological approaches to detect ethical aspects are discussed, (4) the model for ethical evaluation of socio-technical arrangements (MEESTAR) for gathering different planes of observation is explained, and (5) within the framework of the AMICAI the MEESTAR and the failure mode and effects analysis (FMEA), a qualitative as well as quantitative analysis, are combined.

As part of this investigation potential ELSI problems and their possible causes and effects are identified and quantitatively analyzed by the Aachen Model of Identification, Classification and Analysis of ethical, legal, and social Implications. Furthermore, the technological impact is identified and demonstrated alongside potential risks, opportunities and potentials of human-robot collaboration. This contribution describes the results of applying AMICAI based research to expert workshops using the application example of a human-robot collaborative workplace (Wille et al. 2016).

2 Background

The increasing use of (collaborative) robots in manufacturing companies involves advantages as well as disadvantages. On the one hand working persons can be supported in physically strenuous or monotone tasks. On the other hand, the growth of robotic systems can cause job loss and a decline of human skills and knowledge. This dualism of technology leads to the necessity of considering an ethical point of view in the development of human-robot collaboration systems. Such questions belong to the field of technology ethics, an applied and problem-centered field which concentrates on normative uncertainties when dealing with technology. It includes moral questions of equipment and tool production and the use of technology (Bendel 2015). Technology ethics reflect on design and impact of technology. Therefore it is not concerned with technology itself, but with technology in the context of application. This reflection includes three criteria: aims and purposes of technology implementation (e.g. the support of humans at work by robotics), instruments and means which are necessary for the application of technology (e.g. materials for building a robot, location of production), and consequences that arise by using the technology (e.g. risks of production for society and environment). Examining these criteria, technology ethics specifically focuses on the moral aspects involved. Besides, they also relate to scientific and technical progress. Additionally, they consider and integrate alternative options (Grunwald 2013). An important aspect concerning technology ethics is the concept of safety. The absence of unacceptable risks and dangers for humans should be guaranteed with the operation of the system (Kagermann et al. 2013) and the so called "safety" (Liggesmeyer and Trapp 2014) must be created. Safety can be subdivided into two aspects: (1) the systems must not cause any danger for human beings and (2) the systems themselves and the data they contain must be protected against misuse (Kagermann et al. 2013). Moreover, ethical questions arise with ensuing implications for substitutability of humans by machines. Thus, it depends on the question of how far human working processes will change by increasing automation. Concerning automation, two effects appear: (1) the so called down-skill effect: in this case the demand for complex skills declines while the demand for more simple skills increases; (2) the so called up-skill effect: the number of technical monitoring tasks on the activities of the robots rises to the extend that higher skills are required. Consequently, questions about the distributive justice of tasks between humans and robots arise. As a result, on one hand a high degree of automation implies a loss of workplaces, while on the other hand the manufacturing costs for the company will decrease. Additionally, risk of an employee's instrumentalization in the work place emerges (Grunwald 2013; Dworschak 2015). Further, technology ethics deal with the usage and promotion of new opportunities that are enabled by technology—for instance when a service robot reports back errors to a factory and thereby helps the maintenance team fixing them (Bendel 2015).

2.1 Ethical Constructive Technology Assessment

When developing new technologies, various aspects need to be considered. Social, legal and ethical implications must be integrated into the process of development and design. Especially ethical aspects of new technologies should gain more attention to guarantee fair, safe and usable technologies. To make it easier for developers and designers to assess all relevant ethical problems, Palm and Hansson (2006) developed a checklist considering nine ethical aspects which need to be integrated into the process of development and design of technologies. These aspects are (1) dissemination and use of information, (2) control, influence and power, (3) impact on social contact patterns, (4) privacy, (5) sustainability, (6) human reproduction, (7) gender, minorities and justice, (8) international relations, and (9) impact on human values. Palm and Hansson describe their checklist as a tool which shall identify adverse effects of new technologies at an early stage of development. To make this possible ethical implications of technology under development should be investigated continuously. This, of course, requires a close interaction with technology developers (Palm and Hansson

2006). The checklist approach is part of a project they introduce as the so called ethical technology assessment, which is based on earlier forms of technology assessment. The main aim of the TA is to gain political control over the potentially negative effects of technological development by means of early warnings. Another form is the participatory technology assessment, which demands the participation of stakeholders like politicians, scientists, journalists etc. in the developing process. The ethical technology assessment is an extension of these forms which demands detection of potential ethical problems by means of a continuous interaction. By involving relevant stakeholders all different perspectives, interests and solutions shall be considered (Palm and Hansson 2006).

A checklist approach is one way to assess ethical problems. It certainly offers an organized way to work through aspects which need to be considered. Nevertheless it does not genuinely reveal the complexity and relations between the different aspects. This is also criticized by Kiran et al. (2015), who speak of the process of technology development as co-evolution of technology and morals. Hence, both the developed technology as well as the morals and ethics are related to each other in a complex way, by which they influence each other so that they both change over time. This is why a firm checklist is not sufficient to analyze the issue. The authors therefore suggest an ethical constructive technology assessment which analyzes ethical implications in a dynamic way. Another aspect which is criticized by Kiran et al. (2015) is that it does not take into account different cultural settings and the diversity with which users appropriate new technologies. The critique of checklist approach leads the authors to four principles the constructive technology assessment should consider. First of all there is the "subject constitution": the developed technologies do not only influence moral frameworks and social processes at the macro-level, but also the everyday lives of their users. Then, the "accompaniment": the ethical Constructive Technology assessment should be framed in terms of technology accompaniment rather than assessment. Third, the "technological mediation" needs to be considered: the ethical constructive technology assessment should focus on an accompaniment that is done in such a way that design practices incorporate openness to situatedness, alternative lifeworlds and changing moral routines. And at last "non-use": the approach should take into account how technology influences users, as well as selective users and even non-users. After discussing the basic factors of technology ethics, the need to involve all stakeholders and fundamental methodological approaches to detect ethical aspects, the model for ethical evaluation of socio-technical arrangements for gathering different planes of observation is clarified.

2.2 Model for the Ethical Evaluation of Socio-technical Systems

The MEESTAR (in German: "Modell zur ethischen Evaluation sozio-technischer Arrangements") is an ethical evaluation instrument, which was constructed for investigation and assessment of ethical questions within the scope of the study "Ethische Fragen im Bereich altersgerechter Assistenzsysteme". With the help of this model, ethical problems and potential fields of conflict, as well as questions, potential opportunities and future fields of action can be identified for the use of socio-technical systems. Like this, ethical questions can be systemized and analyzed so that an overview of possible solutions for the use of socio-technical systems is gained. MEESTAR does not deliver prefabricated solution. It rather represents a heuristic instrument to ethically locate socio-technical systems in a structural dialogue (e.g. workshop) so that ethical reservations and concerns can be identified by employing the systems (Manzeschke et al. 2013).

The MEESTAR model contains three axes: on the x-axis there are seven ethical evaluation dimensions—care, self-determination, safety, equality, privacy, participation and self-conception—which were identified by a study as the core dimensions. On the y-axis there are four stages of the ethical sensibility—the range is from harmless implementation to rejecting implementation. On the z-axis there are three perspectives of observation—individual, organizational and social. In the following, the x- and z-axis shall be described more precisely.

With the help of these seven evaluation dimensions on the x-axis, evaluating people can lead to ethical questions in a concrete scenario and assign them to the respective dimension. Care in this context means that care, decisions and responsibility are assumed for needy persons if they are not be able to do this for themselves. This care thereby shall work towards the autonomy of the individual. Self-determination describes the autonomy of an individual, which, from a modern perspective is comparable to a maximum freedom of decision and action of the single person. Furthermore, safety either means—from an objective perspective—the protection against harm or—from a subjective perspective—the increase of a feeling of safety. Privacy shall shape an inviolable area around the individuals (Volkman 2003). Thereby privacy belongs to the sphere of the negative freedoms and security rights (Berlin 2002). With such rights and freedoms it shall be guaranteed that individuals can act freely inasmuch as they do not influence the rights and freedoms of other persons negatively. Privacy therefore is a condition for individual freedom and autonomy (Cooke 1999). Social equality (e.g. the access to age-appropriate assistance systems) is especially important. Here, both questions of intragenerational and intergenerational equality are relevant. Furthermore, participation means offering people accesses, rights and goods that enable them to live as an individual in a community together with other people. Eventually, self-conception describes the evaluation and perception of a subject towards itself which can be influenced through the evaluation of other persons. Self-conception of individual factors like age, illness and infirmity does also play an important role.

The three levels on the z-axis cover the respective perspectives of the responsibility of action. Individuals as well as corporate actors have to take responsibility for their actions (individually and organizationally). The responsibility on the social level can be summarized by the question how one would like to live in society and which rights and duties will be bestowed in this society (Heidbrink and Hirsch 2006). With these levels it can be ensured that possible fields of action will be considered from different perspectives.

An advantage of the model is that it provides a platform to discuss ethical questions in practice from the perspective of different stakeholders. However, when applying the model a disadvantage can be that the different evaluation dimensions on the x-axis might come into conflict. For instance, a strong focus on care might influence the autonomy and privacy of an individual negatively.

3 Consideration of ELSI with the AMICAI in the Context of Human-Robot Collaboration

In this contribution, existing models and methods were presented to assess ethical aspects of new technologies. All approaches offer advantages apart from aspects which need further improvement. The ethical technology assessment concept provides a method for assessment of all relevant aspects which need to be considered when developing new technologies via a checklist. This has the advantage of gaining a good overview on all the important factors. On the other hand it neglects the interactions and relation between these different factors. This aspect can be solved by using a model which illustrates these interactions. In the MEESTAR one can either see all important single factors playing a role in technology development and furthermore the way they are related to each other and to external factors. Besides, there is one important aspect that is still not considered in the model, namely the quantitative aspects of technology assessment.

On the basis of the failure mode and effects analysis (FMEA) according to DIN EN 60812 the quantitative aspects can be included. The FMEA is a method frequently used in constructive development processes for systematic analysis determining possible types of error conditions, their causes and effects on the behavior of the system. Using the Aachen model of identification, classification and analysis of ethical, legal, and social implications the MEESTAR and the FMEA, a qualitative as well as quantitative analysis are combined (Wille et al. 2016).

Within this best-practice approach, ethical, legal, and social issues of human-robot collaboration are considered. Here the context of observation is the implementation of human-robot collaboration workplaces. Subsequently for the systematic classification of ethical, legal, and social issues the MEESTAR was selected. The stakeholders include individual (e.g. user or working person), organizational, and societal aspects. As explained by the participatory technology assessment, possible problems caused by new technical developments need to be discussed by experts considering the ethical, legal, and social context. The inclusion of relevant stakeholders (e.g. designers, engineers, manufacturers, supplier, and customer) is recommended to involve all different perspectives, interests and solutions.

At the beginning, possible problems caused by the development of human-robot collaboration and the implementation of human-robot collaboration workplaces are identified. Furthermore the most likely causes of the problems and the problem effects for the stakeholders were discussed. Finally the methods used to detect the problems were examined. Additionally the severity S of the problem's effect is evaluated. Next is the quantification of the probability of occurrence *O*. In AMICAI the whole causal chain of cause, problem and problem effect is considered. The probability of detection *D* is to be quantified for the problem in combination with the method of detection. The risk priority number *RPN* is calculated as the product of *S*, *O* and *D*. In this manner, all problems are completed by row and respectively given their cause, effect, method of detection, severity *S*, probability of occurrence *O*, and probability of detection *D*. In conclusion, the problems can be ranked according to their risk priority number, meaning a higher *RPN* should be considered first. Finally, the solutions are supposed to be developed for the relevant problems that were identified. The earlier analysis helps, for example, by providing information about the problem causes (Wille et al. 2016).

As part of this work, ethical, legal, and social issues which might result from the implementation of human-robot collaboration workplaces and the accompanying change of work were addressed in several expert workshops. The expert committee consists of automation, robot gripper and robot manufacturers, system integrators, universities, end users of manufacturing systems and cooperative robot workplaces as well as of representatives of employers, social partners, and people with disabilities. In a kick off workshop general ethical, legal, and social questions were discussed in terms of the technical consequences of human-robot collaboration. Here, open questions and potential barriers as well as previous practical experiences were debated. In following workshops ELSI questions in the context of human-robot cooperation were identified, summarized and discussed in more depth. In addition, the collection of ethical, legal, and social problems was classified by means of the MEESTAR. Subsequently the various items have been discussed in a further workshop and analyzed with the help of the AMICAI. The aim is to identify potential ELSI problems in the implementation of human-robot cooperation workplaces early in the process.

4 Results

In the kick off workshop numerous ethical, legal, and social issues—first without differentiation between concrete problem causes, problems, and problem effects—were discussed. In the following workshops, after a consolidation, 25 ELSI aspects were identified. Finally, in a further workshop 15 ethical, legal, and social issues—differentiated between problem causes, problems, and problem effects—were analyzed. A summary of the results of the expert workshop is shown below (Table 1).

According to the evaluation of expert workshops the AMICAI (Table 1) for the observation context implementation of human-robot collaboration workplaces follows a number of potential ELSI problems and corresponding causes and effects. It is notable that about three fifth of ELSI problems concern individual aspects. In particular, a little less than half of the problem effects relate aspects regarding acceptance (sabotage, demotivation, lack of acceptance). The solution-oriented approach provides the possibility to focus on the problem causes and to create

Table 1 Documentation of the application of the AMICAI for the quantitative analysis of ELSI applied to a collaborative human-robot workplace

Observation context: Implementation of human-robot collaboration workplaces									
Classification		Problem-criticality					Problem prevention		
Plane of observation	Stakeholder	Problem cause (t/o/p)	Problem	Problem effect	*S*	*O*	Detection method (p/r)	*D*	*RPN*
Care	Organisation	Implementation of human-robot-collaboration workplaces (o)	Danger of breach of the employer's duty of care	Legal/financial consequences	9	5	Risk assessment/ analysis (p)	3	135
Equality	Society	The human being is substituted by robots (o)	Deficits in social fund	Widening gap between rich and poor	9	4	State budget (r)	4	144
Equality	Individual	Robot works with higher productivity/higher quality (t)	Fear of job loss	Sabotage	8	4	Robot downtime/ repair time (r)	4	128
Equality	Individual	Robot works with higher productivity/higher quality (t)	Fear of job loss	Demotivation	6	7	Days with inability to work (r)	4	168
Participation	Individual	Shortcomings in the input or output technique (t)	Communication gap between working person and robot	Serious injury	10	2	Days with inability to work (r)	4	80
Participation	Individual	Shortcomings in the input or output technique (t)	Communication gap between working person and robot	Demotivation	6	7	Survey after initial operation/ employee interview (p)	5	210
Participation	Organisation	Shortcomings in the input or output technique (t)	Communication gap between working person and robot	Production errors	7	5	Quality control (r)	4	140

Participation	Organisation	Shortcomings in the input or output technique (t)	Communication gap between working person and robot	Low productivity	6	7	Production figures (r)	4	168
Privacy	Individual	Sensors of the robot or workplace (t)	Working person feels monitored by sensors of the robot or workplace	Lack of acceptance	6	6	Survey after initial operation/ employee interview (p)	5	180
Safety	Individual	Technical defect (t)	Robot violates the working person	Serious injury	10	2	Days with inability to work (r)	4	80
Safety	Organisation	Technical defect (t)	Robot violates the working person	Legal/financial consequences	9	2	Legal department (r)	4	72
Safety	Individual	Robot falls below the comfort distance (p)	Working person feels harassed by the robot and is afraid	Physicall illness	9	3	Survey after initial operation/ employee interview (p)	5	135
Self-concept	Individual	Implementation of human-robot-collaboration workplaces (t)	(Social) forced to work on a human-robot collaboration workplace	Demotivation	6	7	Survey after initial operation/ employee interview (p)	5	210
Self-concept	Society	Implementation of human-robot-collaboration workplaces (t)	Gap between technologically aware generation and people who are not so technophilic	Social unrest	9	1	Public opinion poll (p)	7	63
Self-concept	Individual	Lack of acceptance (p)	Human-robot collaboration is rejected	Sabotage	8	4	Robot downtime/ repair time (r)	4	128
Self-concept	Organisation	Lack of acceptance (p)	Improper use of the robot	Low productivity	6	7	Production figures (r)	4	168

(continued)

Table 1 (continued)

Observation context: Implementation of human-robot collaboration workplaces									
Classification		Problem-criticality					Problem prevention		
Plane of observation	Stakeholder	Problem cause (t/o/p)	Problem	Problem effect	*S*	*O*	Detection method (p/r)	*D*	*RPN*
Self-determination	Individual	Robot determines work content and cycle time (o)	Demotivation, lack of tolerance	Sabotage	8	4	Survey after initial operation/ employee interview (p)	5	160
Self-determination	Organisation	Robot determines work content and cycle time (o)	Demotivation, lack of tolerance	Low productivity	6	7	Production figures (r)	4	168
Self-determination	Individual	Self-learning robot takes successively the work already done by the working person (o)	Feeling of no longer being needed	Demotivation	5	7	Survey after initial operation/ employee interview (p)	5	175
Self-determination	Individual	Self-learning robot takes successively the work already done by the working person (o)	Human-robot collaboration is rejected	Sabotage	8	4	Robot downtime/ repair time (r)	4	128
Self-determination	Individual	Robot determines work content and cycle time (o)	Overburdened	Inability to work	9	3	Empirical study (p)	2	54

Self-determination	Individual	Robot determines work content and cycle time (o)	Overburdened	Inability to work	9	3	Days with inability to work (r)	4	108
Self-determination	Individual	Robot determines work content and cycle time (o)	Overburdened	Inability to work	9	3	Survey after initial operation/ employee interview (p)	5	135
Self-determination	Organisation	Robot determines work content and cycle time (o)	Overburdened	Production errors	7	7	Quality control (r)	4	196

Problem cause t: technical, o: organizational, p: personal-related
S =1 Severity of the problem's cause from (1) insignificant to (10) catastrophic
O = Probability of occurrence of the causal chain from (1) improbable to (10) very probable
D = Probability of detection before effects of the problem occur from (1) almost certain to (10) completely uncertain
RPN = Risk Priority Number as the product of S, O and D from (1) very low to (1000) very high
Detection method r: reactive, p: preventive

reasonable measures consequently. On the basis of the requirements of an ergonomic human-robot collaboration and the TOP-model technical, organizational and personal-related/behavioral-based measures should be taken into account (Schlick et al. 2010). For example, Brandl et al. (2016) describes technical factors influencing distances that humans will accept between themselves and an approaching service robot. In the area of personal-related measures Broehl et al. (2016) derives factors for increasing acceptance of robots in production through the technology acceptance model for human-robot cooperation in production systems.

5 Discussion and Conclusion

This best practice approach shows that the solution-oriented application of the Aachen Model of Identification, Classification and Analysis of ethical, legal, and social Implications is an essential tool to detect potential ELSI issues in applied research projects at an early stage. In this contribution AMICAI is appliedon the basis of workshops in the context of the implementation of human-robot cooperation workplaces. During planning, execution, and evaluation of the expert workshops, the methods shown in the fundamentals were useful to understand to background of ELSI.

The approach of participatory technology assessment (Palm and Hansson 2006) was helpful to involve the relevant stakeholders in the development process of human-robot collaboration workplaces early in the process. The ethical technology assessment (Palm and Hansson 2006) and the ethical constructive technology assessment (Kiran et al. 2015) showed which aspects should be taken into account in establishing ethical, legal, and social implications.

The Aachen model of identification, classification and analysis of ethical, legal, and social Implications (Wille et al. 2016) used in this contribution based on the model for the ethical evaluation of socio-technical arrangements (Manzeschke et al. 2013) and the FMEA (DIN EN 60812)—leads to a systematic overview of the ethical, legal, and social implications that may arise in the implementation of human-robot collaboration workplaces. The previously described AMICAI approach allows the identification of ethical, legal, and social implications early on in an applied research project, quantify their criticality, and prioritize them based on their relevance.

In principle, the application of AMICAI in expert workshops is adopted to raise ethical, legal, and social issues in the implementation of human-robot collaboration workplaces. In a first step the needs and expectations of stakeholders are interrogated and considered during the course of the workshops. Moreover, it should be taken care of that in heterogeneous groups particularly the participants are sufficiently familiar with both the wording and the method to be used. Furthermore it should be ensured that the causal chains of cause, problem and effect are prepared consistently. During implementation it is to be remembered that the classical restrictions of the FMEA are obviously the limitations for AMICAI as well. Thus, the multiplication of the criteria S, O and D is not defined in a strict mathematical sense due to the ordinal scale. Also, on the one hand it is not guaranteed that similar risks are assigned the same *RPN*, and on the other hand it is possible that there are risks with the same *RPN* that are not equally acceptable. This stresses the fact that the *RPN* are only supposed to be an orientation and not deemed to be interval-scaled "hard facts". However, they are suitable for arranging and prioritizing ethical, legal, and social implications.

Depending on the requirements future workshops can go into more detail and focus on individual ethical, legal and, social issues. Further considerations will be needed in order to develop solutions and design recommendations based on the causal chains of problem cause, problem and problem effect and on the detection method.

Acknowledgements The best-practice approach of a solution-oriented technology assessment of ethical, legal, and social issues in the context of human-robot collaboration is being investigated by the Institute of Industrial Engineering and Ergonomics (IAW), RWTH Aachen University, within the publicly funded research and development project: "MeRoSy—Mensch-Roboter Synergie—Lernen und Adaption in der Mensch Roboter Evolution". It is supported by financial resources (FKZ: 16SV7190) of the Federal Ministry of Education and Research (BMBF) and supervised by VDI | VDE Innovation + Technik.

References

Bendel O (2015) Die Industrie 4.0 aus ethischer Sicht. HMD Praxis der Wirtschaftsinformatik 52:739–748. doi:10.1365/s40702-015-0163-z

Berlin I (2002) Two concepts of liberty. In: Hardy H, Berlin I (eds) Liberty. Oxford University Press, Oxford, NY, S.166–S.217

Brandl C, Mertens A, Schlick CM (2016) Human-robot interaction in assisted personal services: factors influencing distances that humans will accept between themselves and an approaching service robot. Hum Factors Ergon Manuf Serv Ind. doi:10.1002/hfm.20675

Broehl C, Nelles J, Brandl C, Mertens A, Schlick CM (2016) TAM reloaded: a technology acceptance model for human-robot cooperation in production systems. In: HCI International 2016 – Posters' Extended Abstracts, pp 97–103

Cooke M (1999) A space of one's own: autonomy, privacy, liberty. Philos Soc Critic 25(1):22–53

DIN EN 60812: Analysetechniken für die Funktionsfähigkeit von Systemen – Verfahren für die Fehlzustandsart- und -auswirkungsanalyse (FMEA)

DIN EN ISO 10218-2: Industrieroboter – Sicherheitsanforderungen – Teil 2: Robotersysteme und Integration

Dworschak B (2015) Kompetenzen der Zukunft in der Industrie 4.0 – Anwendungsfälle, Szenarien, Entwicklungspfade. In: Gesellschaft für Arbeitswissenschaft e.V. (GfA) (Hrsg.), VerANTWORTung für die Arbeit der Zukunft. Bericht zum 61. Arbeitswissenschaftlichen Kongress vom 25. – 27. Februar 2015, Dortmund. GfA-Press, Dortmund, pp 1–5. ISSN 978-3-936804-18-8

Grunwald A (2013) Handbuch Technikethik. J.B. Metzler, Stuttgart

Heidbrink L, Hirsch A (2006) Verantwortung in der Zivilgesellschaft. Zur Konjunktur eines widersprüchlichen Prinzips. Campus, Frankfurt am Main

Kagermann H, Wahlster W, Helbig J (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0. Forschungsunion im Stiftverband für die Deutsche Wissenschaft, Berlin

Kiran AH, Oudshoorn N, Verbeek PP (2015) Beyond checklists: toward an ethical-constructive technology assessment. J Respons Innov 2(1):5–19

Liggesmeyer P, Trapp M (2014) Safety: Herausforderungen und Lösungsansätze. In: Bauernhansl T, ten Hompel M, Vogel-Heuser B (Hrsg) Industrie 4.0 in Produktion, Automatisierung und Logistik. Springer, Wiesbaden, pp 433–449

Manzeschke A, Weber K, Rother E, Fangerau H (2013) Ergebnisse der Studie 'Ethische Fragen im Bereich Altersgerechter Assistenzsystemen'. http://www.mtidw.de/grundsatzfragen/begleitforschung/dokumente/ethische-fragen-im-bereich-altersgerechter-assistenzsysteme-1. Accessed 26 Nov 2015

Nelles J, Bröhl C, Spies J, Brandl C, Mertens A, Schlick CM (2016) ELSI-Fragestellungen im Kontext der Mensch-Roboter-Kollaboration. In: Arbeit in komplexen Systemen. Digital, vernetzt, human?!. Bericht zum 62. Arbeitswissenschaftlichen Kongress vom 03. - 05. März 2016, Hrsg.: Gesellschaft für Arbeitswissenschaft e.V. (GfA), GfA-Press, Dortmund, pp 1–6. ISBN 978-3-936804-20-1

Palm E, Hansson SO (2006) The case for ethical technology assessment (eTA). Technol Forecast Soc Change 73(5):543–558

Schlick CM, Bruder R, Luczak H (2010) Arbeitswissenschaft. Springer, Berlin

Volkman R (2003) Privacy as life, liberty, property. Ethics Inf Technol 5(4):199–210

Wille M, Brandl C, Nelles J, Mertens A, Schlick CM (2016) AMICAI: Ein methodisches Vorgehen zur quantitativen Analyse von ethischen, rechtlichen und sozialen Auswirkungen anwendungsnaher Forschungsprojekte. In: Arbeit in komplexen Systemen – Digital, vernetzt, human?! Bericht zum 62. Arbeitswissenschaftlichen Kongress vom 02. bis 04. März 2016, Hrsg.: Gesellschaft für Arbeitswissenschaft e.V. (GfA). GfA-Press, Dortmund

Human-Oriented Productivity Management as a Key Criterion for Success in the Digitalised Working World

Patricia Stock and Sascha Stowasser

Abstract
The world of work and business is constantly changing. At the moment, the digitalisation megatrend is significantly changing framework conditions for companies with a range of new requirements. Within this context, a growing desire for more flexibility, which can be achieved thanks to new developments in information and communication technology, can be seen both at companies and among employees. To ensure the long-term success of digitalisation at companies, work must be organised with a human-oriented focus in all operating areas. Human orientation has both a direct as well as an indirect impact on productivity. For efficient human-oriented productivity management, new methods and tools in industrial engineering are required, which would allow the digitalised working world to be analysed and shaped.

Keywords
Digitalisation • Industrial engineering • Productivity managament • REFA

P. Stock (✉)
REFA-Institut e.V., Emil-Figge-Str. 43, 44227 Dortmund, Germany
e-mail: patricia.stock@refa-institut.de

S. Stowasser
REFA-Institut e.V., Emil-Figge-Str. 43, 44227 Dortmund, Germany

Institut für angewandte Arbeitswissenschaft e. V., Uerdinger Straße 56, 40474 Düsseldorf, Germany

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_2

1 Digitalisation as a Driver of Change in the World of Work and Business

1.1 Transformation of the World of Work and Business

The world of work and business is currently undergoing a fundamental change initiated by various megatrends. Changing information and communication technology and the resulting digitalisation of the world of work and business is one of these megatrends [for more detailed information, see e.g. BDI (2011), Rump and Walter (2013), Spath et al. (2013), REFA-Institut (2016)]. Digitalisation unlocks new forms of flexibilisation potential for companies to create work and value-adding processes by primarily providing access to intelligent tools, instruments and automation, production and networking technologies as well as access to globally distributed information, knowledge, skills, resources, work partners and markets (Picot and Neuburger 2013). For employees, this results in new qualification requirements and changes the strain that is placed on them. In particular, the rising acceleration of processes, the increasing speed of change and the consolidation of work have resulted in new employee requirements (Rump et al. 2011).

Furthermore, not only technology is changing–employees are changing, too. Younger generations (known as "digital natives") grew up with these new technologies, which is why they are less afraid of using them and have an intuitive and rather playful approach to these technologies. On the other hand, older generations (known as "digital immigrants") first need to learn how these technologies are used [for more detailed information, see Prensky (2001a, b), Palfrey and Gasser (2008)]. This requires various strategies for using new technologies and the qualifications that are needed in order to use them.

As a result of the value shift, our perception of work has also tremendously changed. Employees also want the time they spend at work to be meaningful, fulfilling and stimulating. A person's job should not compete with his or her private life but, where possible, harmonise with it (Zukunftsinstitut 2011). As a result, employees are ever less willing to adapt to external conditions imposed by companies that do not meet their expectations, wishes and needs (e.g. Zukunftsinstitut 2011; XING 2014).

Due to the demographic change, there is growing evidence of a skills shortage that is forcing companies to actively seek out new employees and retain them through attractive working conditions, while also giving employees a new sense of confidence. In light of this, companies must always consider their employees' wishes and needs because of their growing significance for companies. Companies that pursue corporate goals exclusively and lose track of their employees in the process will face significant problems in the medium term.

1.2 The Flexibilisation of Work

Overall, a growing desire to make work more flexible can be seen both at companies and among employees, although both sides want to be able to control this flexibility. New information and communication technologies offer flexibilisation opportunities. This means that it is the company's duty to identify business potential that could be exploited in a useful and cost-effective way using digitalisation. Special attention should be paid to the requirements and needs of employees. This includes the creation of framework for human-oriented personnel deployment. The qualifications and inclusion of existing staff and the acquisition of additional, suitable skilled workers provide a basis for this. In the context of demographic change, the professional experience of older employees in particular is becoming increasingly significant as well. However, since these employees are usually unable to intuitively master the new technologies of digitalisation as "digital immigrants", they need to be given access to these technologies by means of suitable qualification measures to allow them to further contribute their knowledge and experience to the company as well.

In addition, flexible work opportunities created on the basis of digitalisation need to be implemented on a company-specific basis to allow the company, customers and employees to make use of this flexibility in the best possible way. There are various schemes for flexible work: In the case of trust-based working hours, an employee's performance is measured exclusively on the basis of the results of their work. In addition, time tracking and attendance checks do not apply. In the case of mobile work or telework, employees carry out at least part of their work outside the employer's premises–for example, in cafés, during train journeys or in co-working spaces. In the case of home-based work, employees work at home. Offices that are organised based on the open-plan principle do not provide employees with a fixed workstation. In most cases, there are various workstation types that are explicitly designed for different employee activities (e.g. rooms allowing people to concentrate on their work, creative rooms or conference rooms). This allows each employee to book a workstation for the task at hand, where they can complete this task in the most efficient way.

The relevant legal requirements must be observed for the design of flexible work. In Germany, the Works Constitution Act, the Occupational Safety and Health Act, the Federal Data Protection Act and the associated regulations are particularly relevant here as they set the framework for flexible work.

For the purpose of human orientation, both the works council and employees need to be included in the organisational process. However, new work flexibilisation schemes all require an adequate corporate culture, which is often only created by means of suitable awareness and qualification measures. Both managers and employees are called upon here to create a win-win situation as partners.

The systematic and methodical exploitation of digitalisation and flexibilisation potential can be supported using human-oriented productivity management. Human-oriented productivity management thus forms the basis for the implementation of digitalisation within the company by combining profitability and human orientation. Human-oriented productivity management therefore interprets the consideration of employee interests as a key criterion for success, which can also have a positive effect on productivity.

2 Human-Oriented Productivity Management as the Basis for the Sustainable Implementation of Digitalisation

2.1 Tasks of Productivity Management

The aim of productivity management is to increase a company's productivity. Productivity is generally defined as the ratio of output to input (Gabler Wirtschaftslexikon 2015). For instance, this can be the amount, volume or weight of the products manufactured or special performance parameters, such as gross register tonnage in the case of a shipyard. The denominators are measurable values for the input factors "workforce" (number of employees, working hours, remuneration), "operating means" (number, value) or "material" (quantities, value). The output and factors used (input) can also be valued in money at all times and used in the productivity ratio.

To determine the aggregate productivity of a company, the company's overall performance needs to be used accordingly as an output value (also known as the gross production value). Various parts of the overall performance (e.g. net value added) can also be used for the analysis of partial productivity (e.g. labour productivity). Ideally, the objective should be to consider the partial performances that are relevant to the partial productivity in question on a nuanced basis (Dikow 2006). Companies operating in differing industries or with different production structures can be compared on this "standardised" basis.

Various influential factors that have an impact on the output (revenue or added value), input (workforce, work equipment, materials) and throughput (optional factors; e.g. process organisation, quality capability) can be identified (Nebl and Dikow 2004; REFA-Institut 2016; see also Fig. 3). These influential factors can be used as a lever to increase productivity. In doing so, there are always interactions between the various influential factors. This is why it is impossible to improve all productivity factors at the same time. Productivity management generally focuses on the influential factors that can provide the best support in the pursuit of corporate goals.

A methodical and systematic approach that includes planning, controlling, implementation and monitoring measures is recommended (Dorner and Stowasser 2012; Dorner 2014; DIN EN ISO 9241–2 1992). With REFA methods, the following input factors in particular can be influenced: "workforce" (the ability to perform and its maintenance, the use of available capacity and work organisation), "operating means" (available capacity and its use) and "materials" (material flow and lead time) [for more detailed information, see REFA-Institut (2016)].

Work productivity is considered as one of the key indicators to determine a company's efficiency. In addition to internal and cross-business comparisons, time flow is also particularly significant in this context (Dorner 2014; DIN EN ISO 9241–2 1992). For example, work productivity can be specified as the number of employees involved, working hours spent or staff expenses in euros.

2.2 Human Orientation as an Operating Criterion for Success

In addition to profitability (i.e. the ratio of income to expenses), human orientation also plays a key role in the organisation of work. People are shaped by the work that they do. Well-being, health and personal development are influenced by work and can either be negatively or positively impacted. Therefore, work needs to be organised in an acceptable way on the economic, human and social levels. Both objectives–economic efficiency and a humane organisation–can be complementary, but also incompatible (Nullmeier 2011).

The necessity of human orientation is nothing new. In the 1970s, working conditions in Germany improved as part of a government-funded research programme to humanise working life (In German: "Humanisierung der Arbeit") in order to reduce the number of accidents and occupational illnesses. The amendment to the Works Constitution Act strengthened the rights of co-determination of employees and the works council [for more detailed information, see Nullmeier (2011)]. Here, the focus was on "humane work", i.e. the extent to which a job fulfils the physical, mental and social requirements and needs of an individual (Schlick et al. 2010). However, the discussion on digitalisation and Industry 4.0 usually focuses on technological aspects at present (cf. Jeske 2015) while the changes that need to be made to the structure and organisation of work are usually neglected. As a result, human orientation is not the focus of many companies at the moment.

In the area of human orientation, work is organised in line with the physical, mental and social requirements of the individual. According to Hacker (1986), there are four assessment levels in the organisation of work, which are structured hierarchically. As a result, the criteria of the lower levels need to be met before other levels can be considered. The assessment levels of humane work are (Hacker 1986):

– **Practicability:** Physical and mental requirements, such as body mass, physical strength or cognitive ability, need to correspond to the requirements of the task.
– **Safety:** Carrying out work–even over a lengthy period of time while taking breaks and holiday into account–should not result in any damage. For example, adverse health effects as a result of work-related accidents, occupational illnesses or contaminants need to be prevented.
– **Freedom from impairment:** Work-related harm is the result of strain or stress due to performance requirements–for example, after being qualitatively or quantitatively over- or underchallenged. Work-related harm can be prevented if the damage caused by high levels of stress or strain can be reversed in the short term by recovering during breaks and free time, for example.

– **Personality-building:** Work should offer opportunities to develop one's personality during work activities.

For every assessment level, there are specific operational areas and organisational fields where the criterion in question can be implemented. These are the workstation, work task, technical and social work environment and the individual. The organisation of the workstation, work task and the qualifications and development of employees have been REFA's core competences for decades. Approved REFA methods and tools are thus also suitable for supporting companies when fulfilling human-oriented requirements.

2.3 Impact of Human Orientation on Productivity

Human orientation can initially be described using the criteria provided in Sect. 2.2: practicability, safety, freedom from impairment and personality-building. The lower levels "practicability" and "safety" of this schema primarily affect the sub-factor "the ability to perform and its maintenance" of the "workforce" input factor of productivity (Fig. 1). An employee's ability to perform is primarily determined by their qualifications in addition to their physical and mental abilities and skills. This can also be ensured and promoted by means of qualification measures, training and further training in addition to occupational health and safety, the organisation of the workplace and measures to promote health.

By contrast, the levels "freedom from impairment" and especially "personality-building" have an indirect and significantly more complex and multi-faceted impact on productivity, and thus affect other input factors as well.

For example, "maintaining the efficiency" of operating means can be facilitated significantly by a well-trained employee who is willing to perform, makes use of the optimal condition of work equipment, reports potential issues to maintenance at an early stage or even rectifies them.

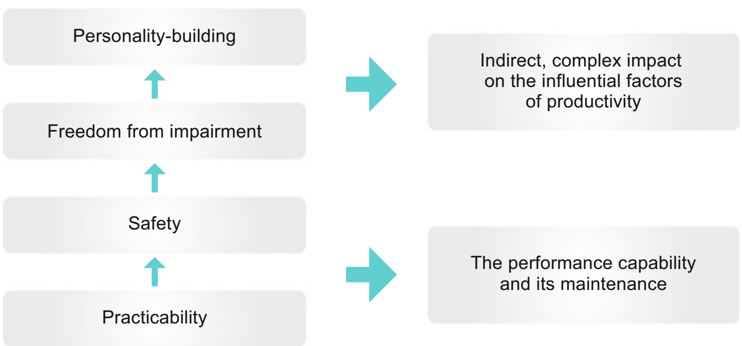


Fig. 1 Impact of human orientation on productivity (source: REFA-Institut 2016)

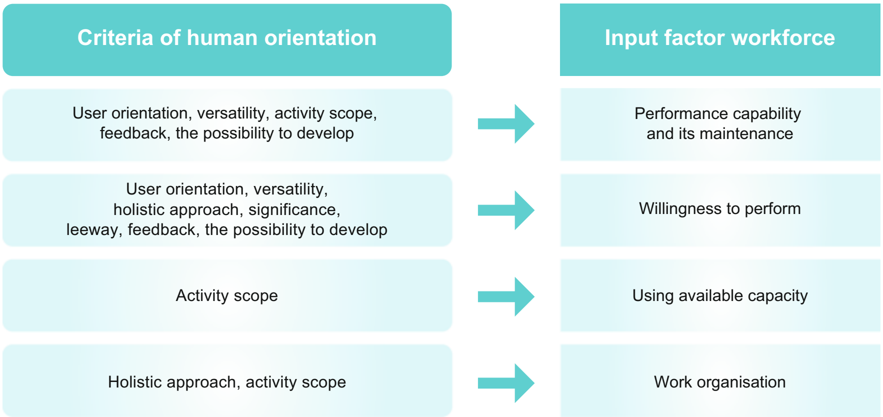


Fig. 2 Impact of human-oriented criteria on sub-factors of the "workforce" input factor (source: REFA-Institut 2016)

The assessment criteria of human orientation can be divided into other sub-criteria–for example, in accordance with DIN EN ISO 9241–2 (1992) (see Fig. 2, left). Figure 2 illustrates the direct effects of human-oriented criteria on the "workforce" input factor. For example, selected effects are [for more detailed information, see REFA-Institut (2016)]:

– **User orientation**, which considers the comparison between the employee's abilities and the requirements of the task.
– **Versatility**, which allows for the development of various skills and abilities and prevents one-sided strain.
– **The possibility to develop**, which promotes the development of existing skills and the acquisition of new ones.
– **Activity of scope**, which supports the autonomous situational use of available capacity in line with capacity requirements.
– The characteristics of a **holistic approach** and **activity of scope** can support a work organisational structure that promotes productivity.

Human-oriented work organisation has both direct and indirect effects on the individual's performance and thus on labour productivity as well. However, human orientation also has a significant impact on other influential factors in productivity. This is because these factors are directly or indirectly influenced by the performance offer and thus the employee's responsible actions which are closely associated with this (REFA-Institut 2016; cf. Fig. 3; fields marked in dark blue: performance offer; fields marked in light blue: sub-factors of productivity influenced by the performance offer). This does not include:

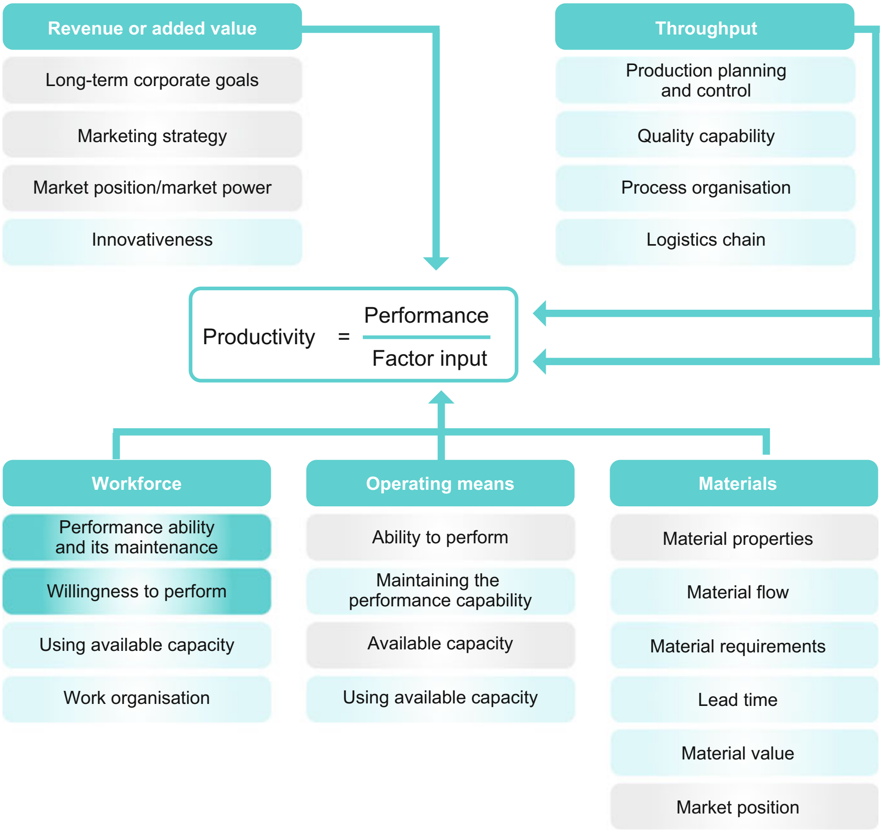


Fig. 3 Possible indirect effects of the performance offer on other sub-factors of productivity (*highlighted in colour*; source: REFA-Institut 2016; Nebl and Dikow 2004)

– Sub-factors relating to the overall performance, revenue and added value, which management establishes
– The ability to perform and the capacity of work equipment, which is essentially determined by management on the basis of investment decisions
– Material properties, which are determined by design within the context of product development
– The company's market position in the supply pyramid, which employees cannot directly influence.

In principle, all other productivity factors can be influenced by employees who are willing to perform and take responsibility. The significance of the particular impact depends on company-specific conditions.

2.4 Challenges of Human-Oriented Productivity Management

Human orientation is becoming increasingly relevant in the current transformation of the working world. This is because individuals, and therefore employees, are becoming increasingly significant for companies. Human-oriented productivity management combines company requirements regarding productivity as well as the requirements of employees with their work and work environment while harmonising them at the same time. In the 2014 autumn survey of the ifaa trend barometer, "human-oriented productivity management" indirectly came first place. This is the result of the combination of issues relating to process orientation and occupational health and safety, which were the most relevant according to those surveyed (ifaa 2015).

Human-oriented factors have indirect, long-term effects on productivity. This is why human-oriented productivity management initially needs to raise awareness of the links that exist and try to make positive effects plausible in the long run. To achieve this, productivity and selected influential factors need to be measured and observed over a longer period of time.

In this context, productivity management establishes strategically relevant influential factors and attempts to positively influence productivity by means of suitable improvement measures for these factors. Human-oriented productivity management also takes into account the impact of employees on influential factors. It also aims to have a long-term positive effect on influential factors via employees–for instance, by using knowledge and experience, qualifications and HR development, involvement, communication, appreciation or by promoting other measures, the ability to perform, the willingness to perform and personal responsibility.

The implementation of human-oriented productivity management results in specific requirements with regard to:

- the inclusion and involvement of employees
- communication
- leadership
- the development, measurement and interpretation of human-oriented key figures and indicators
- the operational inclusion of and collaboration with additional players, such as the HR department and HR development
- the development of a human-oriented corporate and trust culture

By considering the entire range of human-oriented sub-criteria, the value of work is enhanced and employees are shown appreciation, which therefore results in a positive impact on their potential and willingness to perform (REFA-Institut 2016). Last but not least, consistent human orientation can also maintain the employee's ability to work in the long run.

To remain competitive in the digitalised world of work and business in the long term, it is crucial that the willingness to perform, personal responsibility, expertise and the creativity of employees are promoted and ensured. With the same technological equipment available worldwide, this is the only way for productivity and innovation advantages to be permanently maintained and expanded in Germany (REFA-Institut 2016).

3 New Requirements in Industrial Engineering

3.1 REFA as the Architecture of Innovative Industrial Engineering

Within the context of the working world's transformation, the methods and tools of work and business organisation need to be constantly adapted or redeveloped in order to implement human-oriented productivity management. REFA makes available methods and tools that can be used to organise various structural levels within a company in a holistic and sustainable manner while fulfilling the new requirements of the working world (cf. Fig. 4). These methods and tools are aimed at finding a balance between productivity and a sustainable corporate culture that takes the interests and requirements of employees into account. The traditional strategies and methods of industrial engineering and HR management are currently being further developed by the REFA-Institut. As a link between science and practice, the REFA-Institut will continue to assist companies and employees in this turbulent day and

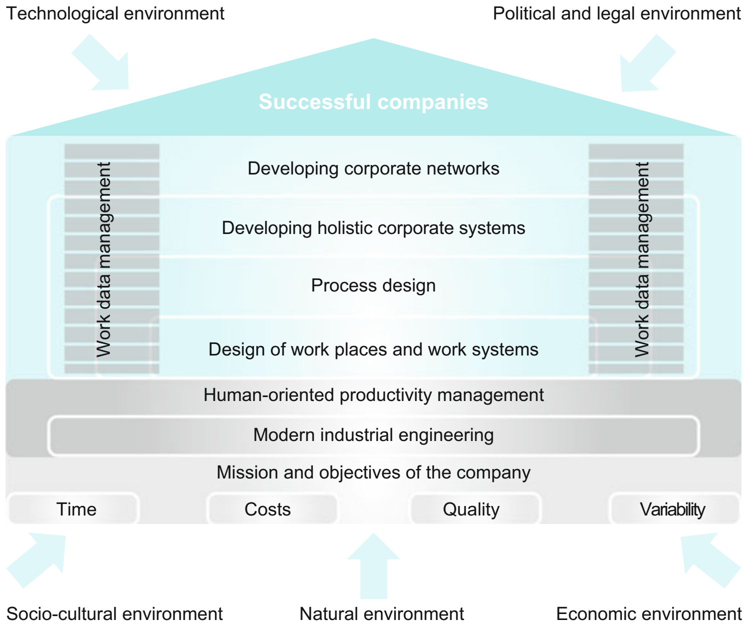


Fig. 4 REFA: methods and tools for the holistic and sustainable organisation of companies (REFA-Institut 2016)

age and provide them with the support they need to adapt to the working world's transformation.

On the basis of these methods and tools, the purpose of industrial engineering is to specifically identify the right methods and tools for a business and initiate them within the company. In particular, the objectives and success factors that are relevant to the company need to be taken into account accordingly.

Based on its original core task–the compilation of data for order planning and remuneration structuring–industrial engineering is becoming the driver of company-wide human-oriented productivity management. As a result, modern industrial engineering no longer organises work systems in production exclusively, but rather considers the company and the value-adding process as a whole, for example, with regard to the following aspects of the value chain:

– For instance, this includes packaging and shipping as well as services including maintenance and repair. The requirements of these areas are often unknown and not taken into account. For example, product development can have an impact on the types of packaging or shipping that are possible, how well products can be disassembled, the frequency of their maintenance and the tools with which they will need to be maintained and whether remote diagnosis or remote maintenance is possible, etc. These and other–often unknown–factors can be decisive for the competitiveness of products on a global market.
– Other processes and areas, such as marketing and sales, also need to be more closely linked to other departments in the product creation process in order to better fulfil the increasing number of requirements in customer orientation and individualisation. This would allow for quicker and more reliable checks in order to establish whether new customer requirements can be implemented, and if so, how. For stronger customer orientation, it could also be useful to increasingly involve customers and suppliers in product development in future and, if applicable, in the product creation process as well. For example, this could be done using digitalisation tools (known as open or crowd-sourced innovation).
– To achieve a holistic approach that does not consider only the product creation process but also the product life cycle as a whole, new requirements that have not yet been considered need to be incorporated. The reuse of components and materials and the targeted handling, preparation and disposal of toxic substances lead to requirements for disassembly that have hardly been considered to date, especially in the case of consumer goods. Quick and cost-efficient product assembly continues to be a priority, while product disassembly is not considered. Whether products can be taken apart and the amount of effort required and permitted for this purpose need to be considered in product development.
– In product development, it is also necessary to clarify the services that a company wants or needs to offer with its product in future. Options include remote diagnosis or remote maintenance, product monitoring or the compilation and interpretation of data beyond the product life cycle (big data or smart data). It is also necessary to clarify whether these services will be rendered by the

company itself or if strategic partnerships will be initiated for this purpose, if applicable.

Industrial engineering needs to expand its scope of action not only within, but also outside the company. Cross-company value chains and networks require cross-company processes that must take into account the prevailing capacity and workload situations in a quick and flexible manner. These processes need to be organised in a secure and reliable way. Their interfaces and quality requirements also need to be defined and aligned.

3.2 Interface Management Within the Company

Industrial engineering cannot solve all these problems alone. The successful management of the outlined tasks is closely related to interdisciplinary collaborations with the company's other functional areas (Hinrichsen et al. 2014). The list below shows a variety of divisions that industrial engineers have to work closely with, depending on the task (Hinrichsen et al. 2014):

– **Management** to align the organisation of work and production systems to the company's strategic requirements.
– **Marketing** to consider market requirements (delivery periods, quantities, product specifications, etc.) when developing work and production systems.
– **Product development** to ensure a product design that is suitable for the production and assembly processes and set requirements for the work and production system structure
– **Maintenance** to consider maintenance requirements in the work and production system structure and fulfil any requirements to efficiently use work equipment in production
– **Production** to fulfil the requirements for the efficient operation of work and production systems
– **Procurement** to allow raw materials, consumables and supplies as well as investment goods that significantly contribute to the efficiency of processes and investments to be procured
– **Controlling** to comply with the target budget for the work and production systems to be organised ("target costing") and create a data basis for accounting using time data
– **IT** to ensure efficient order processing

Collaborations between different business areas are not yet a matter of course. Departments frequently act more or less independently of each other and hardly exchange any information. This results in numerous interface issues. In addition, this complicates the end-to-end optimisation process along the company's value stream. To solve these problems, industrial engineering will need to cover new fields of action and develop or create new methods in future. However, it is

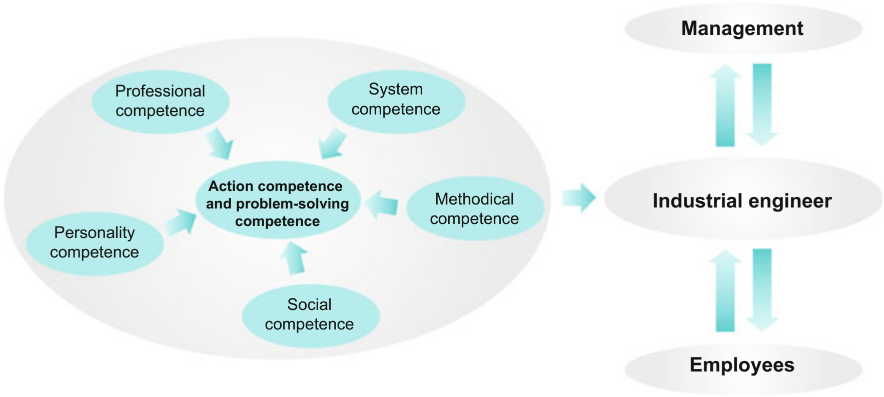


Fig. 5 Position of the industrial engineer within the company (REFA-Institut 2016)

particularly crucial that it encourages and assists or coordinates the company in order to face these challenges. Industrial engineering is particularly suitable for this task since it is performed in this case using a holistic approach. This is because it has the most comprehensive overview of the company's different departments and processes and synergies among various operating areas. This means that industrial engineers are frequently confronted with the task of having to create an appropriate communication and cooperation culture within the company. This is only possible with the company's management and managers' support.

Industrial engineers work at companies and therefore act as an interface between employees, managers and management (see Fig. 5). In their role, industrial engineers act as individuals who actively shape the process of change in the long run.

3.3 Skills of the Industrial Engineer

The profile and expertise that industrial engineers need to have are highly diverse (cf. Fig. 6). Industrial engineers must have comprehensive methodical skills in the areas of ergonomics, work preparation, work and business administration, logistics, project management and the ability to plan and implement holistic corporate systems (Stowasser 2010). In addition to professional and methodology competence, system competence–i.e. the ability to understand, control and change systems and identify risks and opportunities–is also required. Since industrial engineers occupy a key position between managers and employees, they also need to have good social and personality competences. Using these areas of competence with a problem-solving focus is key for industrial engineers in order to resolve situations and solve problems.

	Areas of competence				
	Professional competence	**Methodology comptence**	**Personality competence**	**Social competence**	**System competence**
Skills	Task-specific and activity-specific professional skills and knowledge	The ability to structure problems and reach decisions in a targeted manner	Self-assessment skills and the ability to develop autonomously within the context of work	The ability to take action in social interaction situations in a communicative and cooperative way	The ability to understand systems, effectively take action in systems and change systems
Characte-ristics	Professional training Professional skills Professional knowledge Professional commitment	Problem-solving thinking Abstract and holistic thinking Analytical skills Transferable skills Planning skills Decision-making skills Information gathering skills	Own standards and values Responsibility Creativity and rhetoric Motivation, initiative and dedication Willingness to learn and perform Flexibility and perseverance Emotional intelligence, ability to stand criticism Creating a positive work climate	Team capability and readiness to help Social responsibility Fairness, readiness to cooperate Empathy Delegation skills Tolerance Critical and self-critical skills Taking responsibility for self and others	Sound system knowledge, system control, early recognition of system options not only to be acquired by means of training, shaped by personal experience.

Fig. 6 Skills and characteristics of the areas of competence of industrial engineers (REFA-Institut 2016; Wagner 2010, supplemented)

The use and combination of various skills depending on the situation allows industrial engineers to see activities and processes from an external and specialised perspective, develop measures, and ensure that their suggestions are accepted.

4 Summary

In addition to stable and mastered processes, the organisation of work also needs to be adapted in order to sustainably introduce digitalisation. At the moment, the technological possibilities of digitalisation are largely being discussed (cf. Jeske 2015), while necessary changes to the organisation of work are still frequently neglected. In the digitalised world of work and business, a company's economic success is also based on the potential of its employees and requires skilled, healthy, motivated employees who can take responsibility, be efficient and who are willing to perform.

This means that successful companies need human-oriented productivity management that aligns the company's requirements to productivity as well as the requirements of employees to their work and work environment. Therefore, the methods and tools of work and business organisation constantly need to be adapted or redeveloped in order to implement such a form of human-oriented productivity management. The traditional strategies and methods of industrial engineering and HR management are currently being assessed and further developed by the REFA-Institut. These methods and tools are aimed at finding a balance between productivity and a sustainable corporate culture that promotes employee orientation as a key factor for success. As a link between science and practice, the REFA-Institut will continue to assist companies and employees in this turbulent day and age and provide them with the support they need to adapt to the working world's transformation.

References

BDI–Bundesverband der Deutschen Industrie e. V.; Z_punkt GmbH (Hrsg) (2011) Deutschland 2030 - Zukunftsperspektiven der Wertschöpfung. http://www.z-punkt.de/studien/studie/deutschland-2030-zukunftsperspektiven-der-wertschoepfung. Retrieved 25 June 2016

Dikow A (2006) Messung und Bewertung der Unternehmensproduktivität in mittelständischen Unternehmen–theoretische Grundlagen und praktische Anwendungen. Shaker Verlag, Aachen

DIN EN ISO 9241-2 (1992) Ergonomic requirements for office work with visual display terminals (VDTs); Part 2: Guidance on task requirements. Beuth, Berlin

Dorner M (2014) Das Produktivitätsmanagement des Industrial Engineering unter besonderer Betrachtung der Arbeitsproduktivität und der indirekten Bereiche. Institut für Arbeitswissenschaft und Betriebsorganisation des Karlsruher Instituts für Technologie, Karlsruhe. http://nbn-resolving.org/urn:nbn:de:swb:90-404079. Retrieved 25 June 2016

Dorner M, Stowasser S (2012) Das Produktivitätsmanagement des Industrial Engineering. Zeitschrift für Arbeitswissenschaft 66(2–3):212–225

Gabler Wirtschaftslexikon (2015) Stichwort: Produktivität. Springer Gabler Verlag, Heidelberg u.a. http://wirtschaftslexikon.gabler.de/Archiv/55467/produktivitaet-v6.html. Retrieved 25 June 2016

Hacker W (1986) Arbeitspsychologie: Psychische Regulation von Arbeitstätigkeiten. Huber Verlag, Stuttgart

Hinrichsen S, Jungkind W, Könneker M (2014): Industrial Engineering–Begriff, Methodenauswahl und Lehrkonzept. In: Betriebswirtschaft & Arbeitspraxis, Institut für angewandte Arbeitswissenschaft (Hrsg.), Nr. 221, S. 28–35

ifaa–Institut für angewandte Arbeitswissenschaft (Hrsg.) (2015) Pressemitteilung: ifaa-Trendbarometer "Arbeitswelt" zeigt: Industrie 4.0 ist noch nicht in der Praxis angekommen. http://www.presseportal.de/pm/82380/3014854. Retrieved 25 June 2016

Jeske T (2015) Industrie 4.0 in Deutschland. In: Bullinger AC (ed) Mensch 2020–transdisziplinäre Perspektiven. Verlag aw&I Wissenschaft und Praxis, Chemnitz, pp 16–24

Nebl T, Dikow A (2004) Produktivitätsmanagement - Theoretische Grundlagen, methodische Instrumentarien, Analyseergebnisse und Praxiserfahrungen zur Produktivitätssteigerung in produzierenden Unternehmen. Carl Hanser Verlag, München

Nullmeier E (2011) Arbeitswissenschaft–Einführung und Geschichte. Studienbrief 2-050-2301. Hochschulverbund Distance Learning, Brandenburg

Palfrey J, Gasser U (2008) Born digital–understanding the first generation of digital natives. Basic Books, New York
Picot A, Neuburger R (2013) Arbeit in der Digitalen Welt. Zusammenfassung der Ergebnisse der AG1-Projektgruppe anlässlich des IT-Gipfels-Prozesses 2013. München, Oktober 2013. http://www.bmwi.de/BMWi/Redaktion/PDF/A/arbeit-in-der-digitalen-welt,property = pdf,bereich = bmwi2012,sprache = de,rwb = true.pdf. Retrieved 25 June 2016
Presnky M (2001a) Digital natives, digital immigrants (Part 1). On the Horizon 9(5):1–6
Presnky M (2001b) Digital natives, digital immigrants (Part 2) - Do they really think differently? On the Horizon 9(6):1–6
REFA-Institut (2016) Arbeitsorganisation erfolgreicher Unternehmen–Wandel in der Arbeitswelt. Carl Hanser Verlag, München
Rump J, Walter N (eds) (2013) Arbeitswelt 2030: Trends, Prognosen, Gestaltungsmöglichkeiten. Schäffer-Poeschel, Stuttgart
Rump J, Eilers S, Wilms G (2011) Lebensphasenorientierte Personalpolitik 2.0 - Ein Leitfaden für Un-ternehmen zur Bindung und Gewinnung von Mitarbeiterinnen und Mitarbeitern. Mainz: Ministerium für Wirtschaft, Klimaschutz, Energie und Landesplanung Rheinland-Pfalz. http://www.lebensphasenorientierte-personalpolitik.de/images/stories/pdf/RLP_Brosch_2011_web.pdf. Retrieved 25 June 2016
Schlick C, Bruder R, Luczak H (2010) Arbeitswissenschaft, 3rd edn. Springer, Heidelberg
Spath D, Ganschar O, Gerlach S u.a. (Hrsg) (2013) Produktionsarbeit der Zukunft–Industrie 4.0. Fraunhofer IAO, Stuttgart. http://www.produktionsarbeit.de/content/dam/produktionsarbeit/de/documents/Fraunhofer-IAO-Studie_Produktionsarbeit_der_Zukunft-Industrie_4_0.pdf. Retrieved 25 June 2016
Stowasser S (2010) Die Verankerung des Industrial Engineerings in gegenwärtigen Organisationskonzepten. In: Zülch G (ed) Arbeitsorganisation im Zeichen wirtschaftlicher und demographischer Veränderungen. Shaker Verlag, Aachen
Wagner C u. a. (2010) Fit for Change–Der Mensch als Wandlungsbefähiger. In: wt Werkstattstechnik online, Jahrgang 100, Heft 9, S. 724
XING (ed) (2014) Pressemitteilung: Forsa-Studie zu Wünschen deutscher Berufstätiger: Frauen suchen Sinn in der Arbeit, Männer ein gutes Gehalt. https://corporate.xing.com/no_cache/deutsch/presse/pressemitteilungen/pressemitteilungen-detailansicht/article/pressemitteilungbrforsa-studie-zu-wuenschen-deu/. Retrieved 25 June 2016
Zukunftsinstitut (ed) (2011) Heidelberger Leben Trendmonitor 2011. http://www.eaf-bund.de/documents/Paare_Navipunkte/Studie_Heidelberger_Leben.Trendmonito_2011_final.pdf. Retrieved 25 June 2016

Recource-Constrained Project Scheduling Problem: Investigation of the Quality of Project Plans

Human Competitiveness of an Evolutionary Metaheuristic

Sven Tackenberg, Sönke Duckwitz, Christina Schmalz, and Christopher Marc Schlick

Abstract

This paper introduces the results whether humans are able to develop project plans with a high quality for the well-established Multi-criteria Resource-Constrained Project Scheduling Problem (*RCPSP*). To analyse this, an empirical study was conducted in which activities had to be serialized or parallelized in the plan, process steps had to be inserted or removed and durations as well as resource requirements had to be modified dynamically during planning. In contrast to this human based planning, a specific multicriteria evolutionary metaheuristic is presented that identifies human compatible plans to relatively large project management problems within a reasonable period of time. To evaluate the level of human competitiveness, a metric for measuring plan quality and the results of the empirical study are presented. The results derived from data of 100 participants and the metaheuristic show that only very few people were able to identify optimal solutions. Furthermore, humans are focusing on one target criteria when solving conflicting planning objectives.

S. Tackenberg (✉) • S. Duckwitz • C.M. Schlick
Chair and Institute of Industrial Engineering and Ergonomics, Bergdriesch 27, 52062 Aachen, Germany
e-mail: sven.tackenberg@gmail.com

C. Schmalz
Deutsche Post Lehrstuhl für Optimierung von Distributionsnetzwerken, Kackertstr. 7, 52072 Aachen, Germany

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_3

Keywords
Resource-constrained project scheduling problem • Quality of project plans • Project planning • Multicriteria optimization

1 Introduction

The development of project plans is an important task in project management. The combination of strict and weak structured precedence relations between activities as well as the limited capacity of resources make project scheduling a difficult task. In practice, project management software is used to support the development of a project plan. Thereby, a project scheduling problem can be traced back to the different variants and extensions of the Resource-Constrained Project Scheduling Problem (*RCPSP*) that have been proposed in the literature (Artigues et al. 2008; Hartmann and Briskorn 2010). While the relevant authors fail to agree on a single theory for planning a project implementation, the following characteristics are generally regarded as important for the development of a project plan (Hartmann and Briskorn 2010; Shtub et al. 2005).

A project plan in the form of a Gantt Chart or a graphical network is a prospective description of the expected approach to implement the project. Such an implementation consists of sequences of tasks/activities which are predetermined due to technical, chronometrical and organizational restrictions. Thereby, not all relations between tasks of a project are highly structured. Therefore, actors involved in a project have often an own scope of decision-making concerning which task they will execute next. Functional relations and dependencies of tasks are refined with respect to the situation. As a result, the implementation of a project heavily relies on the task selection and performance of individuals. A project plan is merely a useful resource and should support the decision making.

The critical element for developing a project plan is the valid and optimal assignment of tasks to actors with a limited capacity and to determine the point in time when the tasks are optimally processed. Based on the absence of empirical data for a quantitative statement on the quality of project plans, this paper addresses the research questions:

– To what extent is there a human competitiveness of the plans developed by an evolutionary metaheuristic for a multicriteria resource constrained project scheduling problem?
– What are the key differences between the project plans developed by humans and the solutions of the evolutionary metaheuristic?

2 Creating Project Plans

2.1 The Concept of Planning

To develop a plan which defines a sequence of future tasks and to realize them through appropriate actions is one of the more multifaceted cognitive abilities of humans (Funke and Fritz 1995). Therefore, Scholnick and Friedman try to capture the concept of planning in the following terms (Scholnick and Friedman 1987):

> Planning has been used to account for so many diverse facets of functioning that the definition of the term has become vague. Two theorists who use the term may not share the same focus. Miller, Galanter, and Pribram (1960) define planning as execution of a behavior that matches a scheme, whereas Hayes-Roth and Hayes-Roth (1979) include anticipating a course of action as well.

Most authors are therefore in agreement that the core of planning is to consider the consequences of certain actions and check whether they are able to lead to an approximation of the desired goal (Dörner 2010; Funke and Glodowski 1990). Planning can be taken as a "trial action" previous to the actual action. Therefore, planning occurs before the plan is realized, while acting means that decisions are outperformed (Dörner 2010). The result of such a planning can be defined as the predetermination of a course of action aimed at achieving some goal (Hayes-Roth and Hayes-Roth 1979). A further concept of planning is presented by Funke and Glodowski (1990). They define planning as the design of a sequence of actions, which can be regarded on different levels of resolution, while taking into account restricting constraints. In the following we will be referring to the nomenclature of Funke and Glodowski because the distinction between plan establishment and plan execution is useful to differentiate the project phases and to identify several basic competencies for the development of a project plan (Funke and Glodowski 1990).

To measure and optimize the quality of a project plan, we focus on the plan establishment phase. Our understanding of the establishment of a plan is that it covers the anticipatory ordering of partial steps that lie in future, the consideration of boundary conditions and the use of memory contents. According to Funke and Glodowski four competencies are required to develop a plan (Funke and Glodowski 1990):

- Identification of valid and efficient temporal sequences of activities,
- Evaluation and consideration of direct and indirect boundary conditions during scheduling and parameterization of an activity (e.g. duration, cost, actor assignment),
- Determination of the objectives of the desired action and transfer to planning targets. Thereby, it could be helpful to divide the scheduling task into independent building blocks,
- Development and assessment of valid scenarios of further actions which should be robust against external environmental influences.

There are numerous written references confirming the success factors of projects (Lechler 1997), but there has been little research on how well people develop project plans. Pietras and Coury even assume that humans lose track of an efficient sequence of activities and a valid resource allocation due to their limited cognitive capabilities (Pietras and Coury 1994). Therefore, not all humans are able to develop an optimal plan for a given *RCPSP*. To evaluate the degree to which a person can develop a plan with a high quality, adequate metrics to measure the difficulty of a project planning problem as well as the quality of a project plan is required. Mendling provides a metric for task difficulty and error determinants of business process models which can be adapted to a project planning problem (Mendling 2008). Despite these developments, a generally established suite for measuring the quality of project plans does not exist in project management.

In the following we will present a metric to evaluate the quality of developed project plans for different variants of the *RCPSP*. Therefore, a focus is set on the assessment of the accuracy and correctness of the initial project plans developed by humans. Furthermore, we investigate the fulfilment of the multicriteria planning objectives by comparing the initial plans of humans to the solutions of the evolutionary metaheuristic. This concept of "*plan quality*" is influenced by a variety of factors such as the complexity of a planning problem, personal factors like planning competence, domain knowledge and verbal or graphical representation of a plan. In this paper, we only investigate the planning problem related aspect.

2.2 Formal Description of a Project Planning Problem

The development of a project plan can be regarded as a standard for planning problems with sequence relationships and boundary conditions (Artigues et al. 2008; Hartmann and Briskorn 2010; Brucker et al. 1999). Due to the restricted capacity of resources, a *RCPSP* is defined as a project to be planned which consists of a set of tasks $i = \{1,\ldots, N + 1\}$ to be processed. For each task an expected duration p_i exists. The planned duration p_i and the set of immediate predecessors of a task $j \in P_i$ as well as the allocation of resources r to tasks result in a Gantt chart. Only if the predecessors j of task i within a Gantt chart are fully processed, i can start without time lag. For working on a task i, r_{ik} units of the renewable resource k during each period of processing are required. The capacity of a resource type k in a period is R_k, $k = 1, \ldots, K$. An answer for a *RCPSP* is a schedule S, which consists of a set of starting times $(S_1, S_2,\ldots, S_{N+1})$ for the tasks $i = 0,\ldots, N+1$, where $S_0 = 0$ and the precedence as well as the resource-constraints are fulfilled in a way that the schedule length $T(S) = S_{N+1}$ is minimized.

From a computational point of view, the most problem instances of a *RCPSP* are NP-hard (Bartusch et al. 1988). Therefore, for scheduling real-world project plans of reasonably large size, exact methods, such as the branch-and-bound algorithms studied by Bartusch et al. (1988) and De Reyck and Herroelen (1996) often fail to identify efficient plans within reasonable computational times, in particular when heterogeneous objectives have to be optimized simultaneously. One approach is that the multiple objectives are linearly combined into a scalar objective (Tan et al.

2005). The drawback to this approach is determining the weights of the function. Therefore, a predetermined aggregating function often has the disadvantage of missing all efficient plans. Heuristic algorithms can evaluate performances of plans regarding multiple objectives simultaneously and have been shown to be very suitable for several variants of the *RCPSP*. Heuristic approaches can be classified into two categories: the priority-based heuristics (Alvarez-Valdes and Tararit 1989; Kolisch and Drexl 1996) and metaheuristic approaches, such as simulated annealing, genetic algorithms, and tabu search algorithms (Hartmann 1997; Cho and Kim 1997; Thomas 1998).

The drawback of the *RCPSP* is that it is assumed that a task can be processed in one specific way with a deterministic expenditure of time and a fixed resource requirement. Therefore, the development of a project plan corresponds to a *Multi-Mode RCPSP* (*MM-RCPSP*) and a *Multi-Skill Project Scheduling Problem* (*MSPSP*), which are both extensions of the *RCPSP*. A *MM-RCPSP* allows heterogeneous modes $m = 1, \ldots, M_i$ to process a specific task (Kolisch and Drexl 1997). Here, each mode of a task corresponds to a specific activity m which can be regarded as sufficiently meeting the targets of the task. Such an activity is determined by the duration p_{im} as well as r_{ikm} units of the required resource k and must be processed without preemption. During a project, there are decisions to be made regarding the alternatives of further processing. These heterogeneous project scenarios can be considered in the initial plan. Therefore, the *RCPSP* is extended to include an "*exclusive or*" logical operator (Belhe and Kusiak 1995). As only exactly one successor activity is selected, the activities are connected by an "XOR" node.

The *MSPSP* considers workers with heterogeneous skills (Bellenguez-Morineau and Néron 2008; Néron and Baptista 2002). Each task of a problem requires abilities and capabilities of workers to process one of the potential activities for this task. For processing a non-preemptive task based on the required levels of skills, a subset of workers has to be identified that can process the activity (Bellenguez-Morineau and Néron 2008; Néron and Baptista 2002). A *MSPSP* is mainly determined by the restriction that one worker can fulfill only one skill requirement at a time (Li and Womer 2009). But there are also extended models which assign one worker to different skill requirements of one task (Fırat and Hurkens 2011). Therefore, the model of Firat and Hurkens demands that a set of workers and activities has to be assigned to teams and these teams must stay together during all day (Fırat and Hurkens 2011). Further restrictions of the planning problem are the limited capacity of a worker, the fixed structure of a team during a working day, and that the team members can process only one activity simultaneously.

Therefore, scheduling the tasks of a project is based on the skill levels of workers and the chosen processing strategies which lead to concrete activities. Thus, there is a 1:n relation between a task and the activities which solve this task. But, a worker (resource) can cover more than one skill/function requirement of a task. The assignment of a task i to one or a group of workers (resource) is given by a mode m. Such an assignment may have a direct effect on the duration p_{ik} and the cost c_{ik} of the activity i.

Thus, the considered scheduling problem regarding the development of a project plan is based on our previous work (Tackenberg et al. 2012) and can be formally described as follows. If several project scenarios have to be considered the

Table 1 Sets and parameters

Symbol	Definition
$\boldsymbol{A}$	Set of tasks/activities ($i, j \in \boldsymbol{A}$)
$\boldsymbol{T}$	Set of point of time ($t \in \boldsymbol{T}$)
$\boldsymbol{K}$	Set of resources ($k, l \in \boldsymbol{K}$)
$\boldsymbol{M}$	Set of modes ($m, d \in \boldsymbol{M}$)
p_{ikm}	Duration of activity i in mode m while being processed by resource k
c_{ikm}	Costs of activity i in mode m while being processed by resource k
r_{ik}	Usage of resource k to process activity i
o_{ijm}	Time units of overlapping in mode m between activity i and activity j ($o_{ijm} = 0$, no overlapping is allowed)
F	$\gg 0$
γ_{ijm}	Interval of [0,1], overlapping between activity i and activity j in mode m, 0 for no overlapping

exclusive or logical operator leads to different varieties of the *MSPSP*. All further used symbols and their definitions are presented in Table 1.

$$g_{kl} := \begin{cases} 1, \text{ if resource } k \text{ and } l \text{ are not allowed to work simultaneously} \\ 0, \text{ otherwise} \end{cases} \tag{1}$$

2.2.1 Decision Variables

$$x_{itkm} := \begin{cases} 1, \text{ if the processing of activity } i \text{ by resource } k \text{ is started at time } t \text{ in modus } m \\ 0, \text{ otherwise} \end{cases} \tag{2}$$

$$y_{ij} = \begin{cases} 1, \text{ if activites } i \text{ and } j \text{ overlap} \\ 0, \text{ otherwise} \end{cases} \tag{3}$$

$$l_{ij} = \begin{cases} 1, \text{ if there is no overlapping planned for activity } i \text{ and } j \\ 0, \text{ otherwise} \end{cases} \tag{4}$$

$$z_{ijtkm} = x_{itkm} \cdot l_{ij} \, \forall i, j \in \boldsymbol{A}, \, t \in T, \, k \in \boldsymbol{K}, m \in \boldsymbol{M} \tag{5}$$

2.2.2 Objectives

The two objectives $f_1(x)$ and $f_2(x)$ have to be optimized simultaneously:

$$f_1(x) = min \sum_{t \in T, k \in \boldsymbol{K}, m \in \boldsymbol{M}} t \cdot x_{i_{n+1}tkm} \tag{6}$$

$$f_2(x) = min \sum_{i \in A,\, t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M}} x_{itkm} \cdot c_{ikm} \tag{7}$$

2.2.3 Restrictions

$$\sum_{t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M}} x_{itkm} = 1 \forall i \in \boldsymbol{A} \tag{8}$$

$$\sum_{i \in A} r_{ik} \cdot \sum_{b \in T,\ m \in \boldsymbol{M} \left| t - p_{ikm} + 1 \le b \le t \right.} x_{ibkm} \le 1 \forall t \in T; k \in \boldsymbol{K} \tag{9}$$

$$\sum_{\substack{b \in T \left| t \le b \le t + p_{ikm} + o_{ijm} - 1 \right. \\ m, d \in \boldsymbol{M} \left| g_{kl} > 0; i \ne j; p_{ikm} \ge p_{jld} \right.}} x_{ibkm} + x_{jbld} \le 1 \forall k, l \in \boldsymbol{K}; i, j \in \boldsymbol{A}; t \in T \tag{10}$$

$$\sum_{t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M}} t \cdot x_{itkm} + p_{ikm} \cdot z_{ijtkm} \le \sum_{t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M}} t \cdot x_{jtkm} - o_{ijm} \cdot x_{jtkm} \forall (i, j) \in P \tag{11}$$

$$\sum_{t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M} \left| o_{ijm \ne 0} \right.} \gamma_{ijm} \cdot x_{jtkm} \le y_{ij} \forall (i, j) \in P \tag{12}$$

$$\sum_{t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M}} t \cdot x_{itkm} + p_{ikm} \cdot z_{ijtkm} \ge \left(\sum_{t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M}} t \cdot x_{jtkm} - o_{ijm} \cdot x_{jtkm} \right) - F \cdot \left(1 - y_{ij} \right) \forall (i, j) \in P \tag{13}$$

$$z_{ijtkm} - x_{itkm} \le 0 \forall i, j \in \boldsymbol{A},\ t \in T,\ k \in \boldsymbol{K},\ m \in \boldsymbol{M} \tag{14}$$

$$z_{ijtkm} - l_{ij} \le 0 \forall i, j \in \boldsymbol{A},\ t \in T,\ k \in \boldsymbol{K},\ m \in \boldsymbol{M} \tag{15}$$

$$x_{itkm} + l_{ij} - z_{ijtkm} \le 1 \forall i, j \in \boldsymbol{A},\ t \in T,\ k \in \boldsymbol{K},\ m \in \boldsymbol{M} \tag{16}$$

$$l_{ij} = \sum_{t \in T,\, k \in \boldsymbol{K},\, m \in \boldsymbol{M} \left| o_{ijm} = 0 \right.} x_{jtkm} \forall (i, j) \in P \tag{17}$$

$$x_{itkm} \in \{0, 1\} \forall i \in \boldsymbol{A}; t \in T; k \in \boldsymbol{K}; m \in \boldsymbol{M} \tag{18}$$

$$y_{ij} \in \{0, 1\} \forall i, j \in \boldsymbol{A} \tag{19}$$

$$l_{ij} \in \{0, 1\} \forall i, j \in \boldsymbol{A} \tag{20}$$

$$z_{ijtkm} \in \{0,1\} \forall i,j \in \boldsymbol{A},\ t \in T, k \in \boldsymbol{K}, m \in \boldsymbol{M} \tag{21}$$

Restriction (8) ensures that a task i is assigned to a worker or a resource k. Thereby, i has only one starting time. The condition (9) guarantees that the constraints of k are met. Hence, constraint (10) ensures that tasks are only simultaneously processed if valid combinations of k are scheduled. For the compliance of the precedence relations with overlapping of tasks the constraints (11) to (13) are introduced. After task i has started and a period of $o_{ijm} \neq 0$ time units have passed, the activity of task j can initially start in mode m. If $o_{ijm} = 0$, i and j cannot be processed simultaneously. Otherwise, if i and j overlap, the decision variable y_{ij} is set to 1 (12) and due to the mode m the values in restriction (11) and (13) are equal. The restrictions (14) to (16) are introduced here for a linear program. The variable z_{ijtkm} has the value of 1, if $z_{ijtkm} = 1$ and $l_{ij} = 1$. The constraint (17) defines the value of the decision variable l_{ij}. Restrictions (18) to (21) provide the binary variables.

2.3 Human Competitiveness

Based on the talk of Arthur Samuel entitled "*AI: Where It Has Been and Where It Is Going*" Koza et al. present the concept of human competitiveness for algorithm in the field of genetic programming. As we are dealing with a metaheuristic for the project scheduling problem, a plan is "*human competitive*" if it satisfies one of the following criteria (Koza et al. 2005, p. 4):

- *B) The result is equal to or better than a result that was accepted as a new scientific result at a time when it was published in a peer-reviewed scientific journal.*
- *E) The result is equal to or better than the most recent human-created solution to a long-standing problem for which there has been a succession on increasingly better human-created solutions.*
- *H) The result holds its own or wins a regulated competition involving human contestants (in the form of either live human players or human-written computer programs).*

Thereby, the criterion human competitiveness is independent of the fact that a project plan is computed by an evolutionary algorithm instead of a plan which is developed by a human. With this in mind, the multi-mode extension of the *RCPSP* is an accepted problem in the literature in the field of operations research and there are accepted results for corresponding test cases. But much more interesting is a comparison between human contestants who develop plans for a given multi-criteria *MSPSP* and an evolutionary metaheuristic. Considering the concept of planning in Sect. 2.1, the feasibility of a plan with respect to precedence relations and skill/resource requirements as well as the achievement against the planning objectives can be used as an evaluation basis.

To compute an efficient project plan for a given amount of data, the evolutionary metaheuristic is based on information, knowledge and intelligence of the human who employed the scheduling method. Using the metaheuristic, the ratio is of interest to which extend the underlying algorithms solve a project planning problem compared to the remained amount of intelligence that is supplied by the human. Koza et al. term this ratio "*artificial-to-intelligence*" (Koza et al. 2005). From this definition it is clear that human competitive plans and a high artificial-to-intelligence ratio imply less influence of a human during a planning phase of a project. Our goal is to develop an evolutionary metaheuristic that requires relatively little human effort to handle new scheduling problems and their use is classified as "routine".

3 Concept of the Multi-objective Evolutionary Metaheuristic

A Multi-objective Evolutionary Algorithm (*MOEA*) uses techniques and procedures inspired by evolutionary biology and serves here as a heuristic meta strategy to solve complex optimization problems. In this section we present a novel *MOEA* approach for the introduced formal description of the *MSPSP* with overlapping and logical operators. The goal of our approach is to successively develop human competitive plans by a systematic search within a target space.

The algorithm is based on the *Strength Pareto Evolutionary Algorithm* (*SPEA2*) and was extended to the formal description of Sect. 2.2. Before the optimization process itself is executed, the *MOEA* generates a structure of islands which is based on the characteristic of project scenarios and a global archive for nondominated individuals.

The *MOEA* starts by computing an initial population (*POP*) of individuals. Each individual represents a single project plan. Based on scheduling objectives, an individual is evaluated according to its fitness. We use a fitness value as an indicator for the dominance relations between the individuals of *POP* as well as a classification regarding their unique characteristic. The latter is based on density relations between the identified solutions and is a control parameter to distribute efficient individuals uniformly within the search space. Until the quality of plans is insufficient and the stopping criterion for the optimization process is not met, the genetic operations recombination and mutation modify the individuals of *POP*.

3.1 Definition of Individuals

The genetic representation of our *MOEA* has to reflect the scheduling problem as well as the mode assignment problem. To solve the scheduling problem, starting times have to be assigned to the tasks to be processed. The assignment of tasks to workers or resources and setting up a particular activity are part of the assignment problem. Therefore, we have developed a *MOEA* which can deal with both problems simultaneously because there is no efficient assignment of tasks to limited resources without information about starting times and makespan. Kolisch and Hartmann distinguish different codings of an individual and note the frequent use of the *activity-list* (*AL*) and *random key* (*RK*) representation (Kolisch and Hartmann 2006). Using the activity list representation, the position of an activity in the list describes the scheduling sequence as well as the predecessor relations between activities. If a random key list is used, the position of an activity does not contain any information regarding the sequence of activities.

With this in mind, we developed a new random key list for an individual by a triple $I = (\lambda, \mu, \eta)$ of a feasible list of activities λ, a list μ for taken decisions regarding different project scenarios (resulting from logical operators) as well as a list η regarding the iterative execution of activities. A list λ is feasible if the activities can be executed by the assigned resources and if all predecessors of a task are sufficiently executed before this activity starts. A decision μ to select one of several activity courses is a mapping which assigns to each logical operator $o \in \{1,\ldots,O\}$ one decision $\mu(o) \in M_o$. The outcome of a decision refers to at least one path of tasks (direct successors of the logical operator) which have to be processed. Such a decision can depend on rules or a random distribution. Similar, to each iteration operator $e \in \{1,\ldots,E\}$ (trigger of an iteration) the characteristic $\eta(e) \in M_e$ of an iteration (activities, frequency of iterated activities) is assigned. The *MOEA* uses the following notation for an individual:

$$I = \begin{pmatrix} i_0 & \ldots & i_I \\ o_1 & \ldots & o_I \\ e_1 & \ldots & e_I \end{pmatrix}. \tag{22}$$

The *MOEA* transfers each genotype $I = (\lambda, \mu, \eta)$ of the population to a unique project plan (phenotype). All tasks $(0, \ldots, I)$ of the planning problem are part of the activity list λ and each task i is characterized by several random variables. A distinction is made between different variable types of a task but each variable type must occur at least once for i:

$$\lambda_i = (s_i, e_i, n_i, w_i, I_i, d_i, r_i). \tag{23}$$

- **Relative starting time s_i**: The variable defines the earliest point of time at which the task i can be initially processed. Each value of s_i refers to the degree of completion of one of the predecessors of i, $j \in P_i$.
- **Processing time e_i**: The estimation of the makespan for a task can be associated with uncertainty. If a probability distribution is used to generate a stochastic duration, the variable e_i references to a time value to calculate p_{ikm}.
- **Number of workers n_i**: The variable refers to the number of workers which are scheduled for task i.
- **Workers w_i**: Based on the value of n_i, the value of w_i defines a single worker or a group of workers which fulfil the skill requirements. This means that w_i is one subset of all feasible combinations of workers for a given n_i.

– **Importance of a task I_i**: The variable represents the communicated importance of a task. This has an impact for the scheduling sequence of tasks.
– **Claimed date of completion d_i**: The value d_i determines the planned deviation from the claimed date of completion of task i.
– **Resources r_i**: The variable r_i references to a specific combination (type and quantity) of non-renewable and renewable resources.

A combination of workers, renewable and non-renewable resources which together allow a task processing can be regarded as one mode m_i of the formal description (11). Each λ_i is feasible with respect to the precedence relations, the workers and the renewable resources, but not necessarily with respect to the non-renewable resource constraints. Therefore, if an infeasible individual with respect to the non-renewable resources is computed, it is discarded.

3.2 Software

We close this section with a few remarks on the implemented software architecture, namely the evolutionary cycle and the island concept. The execution of the cycle (cf. Fig. 1) is repeated until the stopping criteria are met. Then the individuals on the Pareto front are selected. Here, a standardized problem definition in the form of several *XML*-files serves as an input for the metaheuristic. In the following list, the components of the evolutionary metaheuristic are described according to the sequence of execution:

– Component "*Scheduling Task Interpreter*": The scheduling task interpreter reads in the standardized description of the planning problem (*XML*-files).

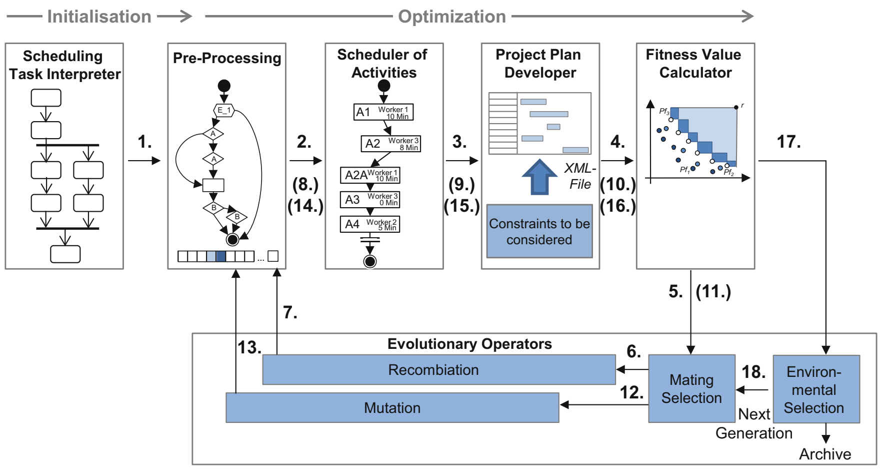


Fig. 1 Software framework

– Component "*Preprocessing*": Before the optimization of a project plan is started, the individuals for the GA are generated. Therefore, this component creates for all activities i and logical operators o and e the variables λ_i, μ_o, η_e of the individual I. Moreover, one graph of the activities is computed which describes all precedence feasible sequences of activities as well as different project scenarios (Fig. 2). The graph results from all possible decisions regarding the logical operators. Furthermore, the component computes individuals for the initial population by the following three steps which are repeated until the stopping criteria are met. First, the variables λ_i of all activities $i = 1, \ldots, N+1$ are generated by randomly selecting values which are stored in the list of the genotype $I = (\lambda, \mu, \eta)$. Thereby, each variable type of λ_i is set within the validity range. Second, the values for the logical as well as the iteration operators are randomly generated. Third, the list is checked for non-renewable resource feasibility.
– Component "*Scheduler of Activities*": For our problem specific representation, we cannot use any of the standard scheduling schemes. Therefore, we introduce a component to map the graph of the component "*Preprocessing*" to a graph which contains only the processed activities. To achieve this, the current values μ_o and η_e of an individual are taken into account. Furthermore, the variable values of all λ_i are transferred into plaintext and are assigned to the activity i of the graph. For illustration, we consider the graph given in Fig. 2 and two exemplary mapped illustrations.
– Component "*Project Plan Developer*": The graph of the component "*Scheduler of Activities*" is used to develop a detailed project plan. Based on priority rules selected by the user and the availability of the assigned workers and resources an activity is scheduled in the plan. First, the algorithm starts with the dummy

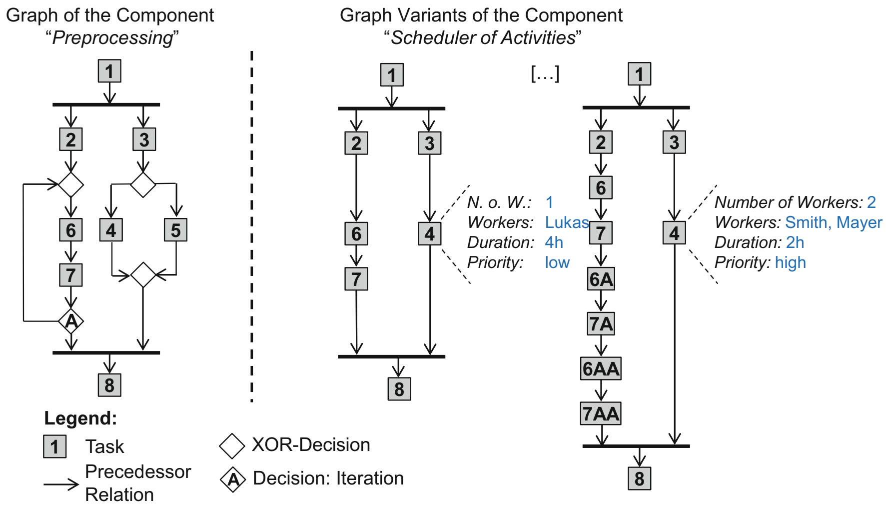


Fig. 2 Examples of graphs for a project planning problem

source activity at time 0. Then, subsequently the next unscheduled activity with the highest priority is taken from the list λ. But this activity can be only scheduled if all of its precedence constraints and resource requirements are satisfied. Otherwise, an activity with a lower priority is selected.

– Component "*Fitness Value Calculator*": To measure the quality of a developed project plan, the user can select one of the implemented metrics for the multi-criteria *RCPSP*. The well-known schemes are introduced in Sect. 4.4.
– Component: "*Evolutionary Operators*": For our definition of individuals, we cannot use any of the standard crossover and mutation operators of the *RCPSP* literature. Therefore, we have adapted the *simulated binary crossover* operator (*SBX*) (Deb and Agrawal 1995). Based on the calculated fitness values for the developed plans the operator selects two "parent" individuals for crossover, a mother $I^M = (\lambda^M, \mu^M, \eta^M)$ and a father $I^F = (\lambda^F, \mu^F, \eta^F)$. For each variable of I the operator draws a single random value u [0,1] and calculates a relative distribution factor β:

$$\beta = (2u)^{\frac{1}{\eta_c+1}} \text{ if } u < 0.5, \tag{24}$$

$$\beta = \left(\frac{1}{2(1-u)}\right)^{\frac{1}{\eta_c+1}}, \text{ else.}$$

Now the variable values x^D and x^S for two new individuals, a daughter $I^D = (\lambda^D, \mu^D, \eta^D)$ and a son $I^S = (\lambda^S, \mu^S, \eta^S)$ are produced from the parents:

$$\begin{aligned} x^D &= 0.5[(1+\beta)x^M + (1-\beta)x^F] \\ x^S &= 0.5[(1-\beta)x^M + (1+\beta)x^F] \end{aligned} \tag{25}$$

The remaining variable values of I^D und I^S which are not selected for modification are taken over from the mother respectively father. While the crossover operator combines the information of I^D and I^F the mutation operator is applied to compute newly generated child individuals $I^{Z'} = (\lambda^Z, \mu^Z, \eta^Z)$. Therefore, it can modify each variable of I. We use an adapted version of the polynominal mutation which enables us to randomly select and modify each variable of I (Deb and Goyal 1996). Furthermore, a random value v [0,1] is determined and a distribution factor δ is calculated:

$$\delta = 2v^{\frac{1}{\eta_m+1}} - 1, \text{ if } v < 0.5, \tag{26}$$

$$\delta = 1 - (2(1-v))^{\frac{1}{\eta_m+1}}, \text{ else.}$$

Next, the mutation operator modifies the variable based on the current variable value x^Z (Definition of the permissible value range: x^{max}, x^{min}):

$$x^{Z'} = x^z + \delta\left(x^{max} - x^{min}\right). \tag{27}$$

We implemented several variants of a stochastic selection operator which all follow a survival-of-the-fittest strategy. Such an operator sorts the individuals with respect to their fitness. First, a detailed project plan is developed by the components "*Scheduler of Activities*" and "*Project Plan Developer*". Subsequently, the component "*Fitness Value Calculator*" computes the fitness of an individual due to the objective function values (cf. Fig. 1). Based on probabilities which are calculated under consideration of the fitness values the component selects individuals for crossover and mutation (*mating selection*) as well as for the next generation (*environmental selection*). The newly evolved individuals computed by mutation and recombination are combined with the best individuals preserved from the previous generation. To select individuals for the next generation all of them have to be mapped to detailed project plans and evaluated again. Finally, the fittest individuals are stored in a global archive. We use such an archive to keep the best individuals over generations. If the pre-defined population and the archive size are exceeded by the number of individuals, solutions are discarded which are in crowded areas of the target space. This allows a uniformly distribution of the individuals within the search space. All other solutions are deleted from the population. The remaining individuals form the new population of the next generation and the evolutionary cycle is repeated until the stopping criterion is met.

We close this section with a few remarks on the island model. The island model is used to simultaneously optimize plans for several project scenarios. One island represents a specific characteristic of an iteratively processing of activities. The island model takes into account several sub-populations (1:1 relation between an island and an iteration characteristic of a project plan) which develop almost independently. The islands are connected by a migration of individuals. Therefore, an individual with a good fitness can swim from one island to another and will be part of this sub-population (Hartmann 2001). Triggers for such a migration are the crossover und mutation operators which modify the values of η.

4 Empirical Study

4.1 Participants

A total of 100 students of mechanical engineering, 14 female and 86 male between 20 and 27 years of age (mean: 22.51 years, SD = 2.52) participated in the study. All participants reported to have minor or no experience in project planning (minor = 36, no = 62, not specified = 2) and despite one exception, all participants indicated German as their mother tongue. All of them had no experience with software tools for project management and the experience in project management was minor (minor = 3) or non-existent (no = 96, not specified = 1).

4.2 Process

With the start of the laboratory experiment the participant's personal data were collected, i.e. age, profession, and prior experience in project management. This was followed by a paper based introduction (approximately 25 min) to the aim of the experiment, the type of planning task and the planning software to be used (Fig. 3). The introduction also included an exercise phase in which the participant had to develop a project plan of five activities together with the test supervisor. At the end of the exercise, the supervisor presented the plans which fulfilled the objectives maximally. After a brief pause, the first planning task was handed over and the participant had to develop the corresponding project plan. During processing the supervisor left the room but was always reachable by phone to answer questions. The first planning phase ended when the participant informed the test supervisor that the development of the project plan is finished. The supervisor handed out a self assessment questionnaire concerning the quality of the developed plan and the test setting. During the period of completion of the questionnaire the supervisor checked the developed plan with respect to completeness and accuracy. Once an error was detected, the participant had to make re-designs of the project plan until a valid solution has existed.

After a short break a second planning task was conducted by the participant. The further procedure was equal to the first planning task. The sequence of both planning tasks was balanced between all participants. To complete the laboratory experiment, tasks of a web based sub-version of the BIS Test Version 4 of Jäger et al. had to be answered (20 min) (Jäger et al. 1997).

4.3 Scheduling Problem

We generated four test scenarios with 15 tasks each. Both were inspired by project data from an engineering service provider. All instances differ regarding the restrictions and represent a *RCPSP*, a *MSPSP*, a *MSPSP-ov* or a *MSPSP-It*. The *RCPSP* is

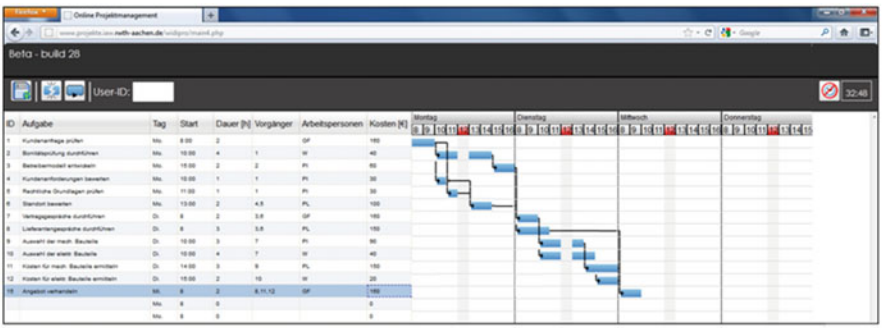


Fig. 3 Screenshot of the web-based test environment for project plan development

determined by the pre-defined assignment of workers to tasks, whereas, the *MSPSP* provides heterogeneous modes of activity-processing. Each mode of a task requires a minimum and maximum capacity of workers (e.g. Artigues et al. 2008; Hartmann and Briskorn 2010) with certain level *l* a skill *k* (e.g. skill 1 with the level: "medium": K_1[m]). The parameters of the scheduling problem can be obtained from Fig. 4 and Table 2. All four planning problems consist of three separate Building Blocks which are combined differently. Due to the identically

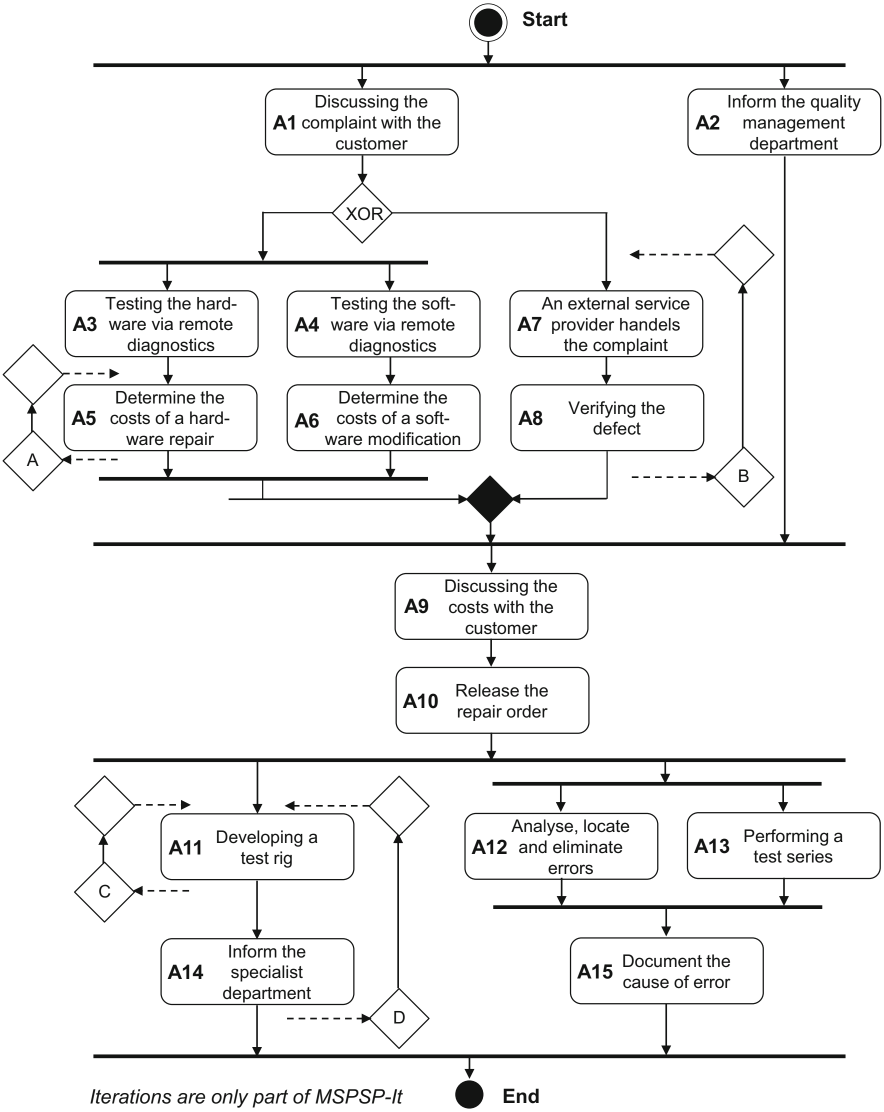


Fig. 4 Semi-formal description of the project planning problems

Table 2 Parameters for the *RCPSP*, *MSPSP* and *MSPSP-ov*

Task		RCPSP		MSPSP-II, MSPSP-II-IT				MSPSP-ov					
	P_i	p	a_{im}	M_i	AP_i^{min} AP_i^{max}	AK	a_{im}	M_i		AK	a_{im}	Overlapping	
												O_j	p: a_{im}, c_i
A1	--	1, 6	2	1 2	[2,2] [2,2]	K1[m], K5[h] K1[h], K5[h]	3 2	1 2	[2,2] [2,2]	K1[m], K5[h] K1[h], K5[h]	3 2	--	--
A2	--	2	3	1 2	[1,1] [1,1]	K1[m] K1[h]	3 2	1 2	[1,1] [1,1]	K1[m] K1[h]	3 2	--	--
A3	1	4	4	1 2	[1,1] [1,1]	K3[m] K3[h]	4 3	1 2	[1,1] [1,1]	K3[m] K3[h]	4 3	--	--
A4	1	4	3	1 2	[1,1] [1,1]	K3[m] K3[h]	4 3	1 2	[1,1] [1,1]	K3[m] K3[h]	4 3	--	--
A5	3	3	4	1 2	[1,1] [1,1]	K2[m] K2[h]	4 3	1 2	[1,1] [1,1]	K2[m] K2[h]	4 3	--	--
A6	4	2	2	1 2	[1,1] [2,2]	K2[h] K2[h]	2 1	1 2	[1,1] [2,2]	K2[h] K2[h]	2 1	--	--
A7	1	5	3	1	[1,1]	K4[h]	5	1	[1,1]	K4[h]	4	--	--
A8	7	2,4	2	1 2	[2,2] [2,2]	K1[m], K3[h] K1[h], K3[h]	4 3	1 2	[2,2] [2,2]	K1[m], K3[h] K1[h], K3[h]	4 3	1	1,3: 4, 440 2,3: 6, 480
												2	1,3: 3, 330 2,3: 4, 320
A9	2,5, 6,8	1.6	2	1 2	[1,1] [1,1]	K1[m] K1[h]	3 2	1 2	[1,1] [1,1]	K1[m] K1[h]	3 2	--	--
A10	9	2	3	1 2	[1,1] [1,1]	K1[m] K1[h]	3 2	1 2	[1,1] [1,1]	K1[m] K1[h]	3 2	1	1: 3, 240 2: 4, 200
												2	1: 3, 160 2: 4, 200
A11	10	4	3	1 2	[1,1] [1,1]	K3[m] K3[h]	4 3	1 2	[1,1] [1,1]	K3[m] K3[h]	4 3	--	--
A12	10	4	2	1 2	[1,1] [1,1]	K3[m] K3[h]	2 1		[1,1] [1,1]	K3[m] K3[h]	2 1	--	--
A13	10	3	1	1	[1,1]	K3[h]	1	1 2	[1,1]	K3[h]	1	--	--
A14	11	2	2	1 2	[1,1] [1,1]	K2[m] K2[h]	2 2	1 2	[1,1] [1,1]	K2[m] K2[h]	2 2	1	2: 3, 150 3: 3, 90 4: 3,...30
												2	2: 3, 150 3: 3, 90 4: 3,...30
A15	12,1 3	2	2	1 2	[1,1] [2,2]	K2[h] K2[h]	2 1	1 2	[1,1] [2,2]	K2[h] K2[h]	2 1	--	--

parameterized building blocks, both test scenarios of a set (*RSPCP*, *MSPSP*, *MSPSP-ov*, *MSPSP-It*) have the same level of difficulty (size, number of relations between tasks, boundary conditions by assigning a task to a worker). The blocks are only combined in different order and the designations of tasks were modified to ensure that effects of memorization can be excluded by processing both tasks in one set. The instance *MSPSP-ov* is dedicated here to the determination of the optimal simultaneous execution of tasks with predecessor constraints. Thereby, the starting times of tasks A_8, A_{10} and A_{14} are restricted to a finite number of instants corresponding to the starting time of the predecessor P_i of i (cf. Table 2). Table 2 shows the information concerning the available workers (skill level, working time).

Very heterogeneous plans can arise from the routing of activity sequences (A_3 to A_6 OR A_7 AND A_8) and the selected modes how the individual tasks are processed. To solve one of the test scenarios the participants as well the evolutionary metaheuristic had to define the starting times of the activities (day, hour) and the task modes which determine the durations and costs of the activities to be scheduled.

For the *MSPSP-It* the iterations "*A*", "*B*", "*C*" and "*D*" are integrated in the model. Therefore, the tasks *A5*, *A7*, *A8*, *A11* and *A15* have to be iterated if a specific assignment of workers to tasks is scheduled (Table 3).

4.4 Plan Quality

In order to quantitatively evaluate project planning, it is suitable to evaluate the outcome of the planning phase, the project plan, instead of evaluating the planning phase itself. The most common objectives for a project are duration, costs and quality (Hans et al. 2007). In line with widely recognized findings, Hans et al. state that it is difficult to simultaneously guarantee a short duration, low costs and high quality (Hans et al. 2007). Therefore, the criteria *optimality* of a project plan is given when all objectives are met to the greatest extent possible. However, a significant number of projects is still unsuccessful or their realization is not efficient (White and Fortune 2002). For example, many international high-profile projects have not been finished on time and/or within the given budget (Feldman 2009). Generally, the project planning and execution are listed as 'failed' if the original budget is exceeded, the project ran late or the objectives are failed (Hartman and Ashrafi 2004). Missing the expectations could be caused by plans with an inefficient scheduling of tasks and resources or an incorrect consideration of boundary conditions. Therefore, we believe that the two dimensions *optimality* and *correctness* can completely describe the quality of a project plan.

The correctness of a plan is evaluated based on the number of errors of a plan as well as the effect of an error regarding the further course of a project. This understanding is in line with Rasmussen's generic error types of omission and execution (Rasmussen 1987). An error is classified as an omission type when a required activity is not in the plan or one parameter of the standardized activity description (e.g. activity duration, cost) is missing. The second error type results from a parameter that exists but which is invalid in terms of planning constraints. To quantify the impact of an error, the minimum required numbers of adjustment steps necessary to generate a valid plan are measured.

The optimality of a plan is operationalized by the degree to which the objectives of the project are fulfilled. Therefore, we utilize three metrics: Euclidean distance

Table 3 Workers for the *RCPSP*, *MSPSP*, *MSPSP-ov*, *MSPSP-It*

	Head of Department (ID 1)		Engineer 1 (ID 2)		Engineer 2 (ID 3)		Student trainee (ID4)		External service provider (ID5)		Customer (ID6)	
	Skill	Cost [€/h]	Skill	Cost [€/h]	Skill	Cost [€/h]	Skill	Cost [€/h]	Skill	Cost [€/h]	Skill	Cost [€/h]
RCPSP	–	80	–	50	–	30	–	10	–	90	–	0
MSPSP	K1[h]	80	K1[m] K2[h]	50	K2[m] K3[h]	30	K2[h] K3[m]	10	K4[h]	30	K5[h]	0
MSPSP-ov	K1[m] K2[h]	50	K2[m] K3[h]	30	K2[h] K3[m]	10	K4[h]	40	K5[h]	0		
MSPSP-It	K1[h]	80	K1[m] K2[h]	50	K2[m] K3[h]	30	K2[h] K3[m]	10	K4[h] K4[m]	50 40	K5[h]	0

Skill: high [h]; medium [m]; low [l]; Working hours: 8:00–12:00; 13:00–17:00

(28), the dominance based hypervolume and the minimum required numbers of adjustment steps to achieve an optimal project plan.

In each case the values for all objective functions (e.g. project duration, project cost) are derived from the project plan computed by an algorithm or developed by humans. These values are used for computing the minimum Euclidean distance of an identified target vector to an efficient vector. Such a vector is efficient if it is not dominated and if it is a solution along the Pareto Front. If a minimization problem is considered, a vector $f(I)$ dominates another vector $f^*(I)$ if no component of $f(I)$ is greater than the corresponding component of $f^*(I)$ and at least one value of $f(I)$ is smaller: $f(I) \prec f^*(I)$ (Zitzler et al. 2004). Clearly, the value of the computed minimum Euclidean distance for a given plan i and an efficient vector j describes the level of optimality for each objective function k of the scheduling problem (cf. Fig. 5):

$$d_{ij} = \sqrt{\sum_{k=1}^{K} \left(\frac{f_k^{(i)} - f_k^{(j)}}{f_k^{(max)} - f_k^{(min)}} \right)^2}. \tag{28}$$

The second metric considered here, is the S-metric of Zitzler et al. (1999). It creates an exhaustive order in the quantity of vectors $f(I)$ due to the computation of a scalar fitness value for each $f(I)$. First, the known vectors $f(I)$ are divided into sets of non-dominated and dominated vectors. Two vectors $f(I)$ are incomparable if each vector contains a more advantageous component for one of the objective functions k (cf. Fig. 5). For each vector $f(I)$ the quantity of all other solutions are determined which dominate $f(I)$. This number represents the strength of a solution; a value of zero is the very optimum for the target area. Afterwards, all $f(I)$ with the same strength are sorted according to the first objective k of vector $f(I)$ (project duration). To achieve an exhaustive order, the hypervolume for all vectors with an identical strength is calculated.

For a target area with two objectives f_1, f_2, $k = 2$, the hypervolume ΔS is calculated as follows (cf. Fig. 5):

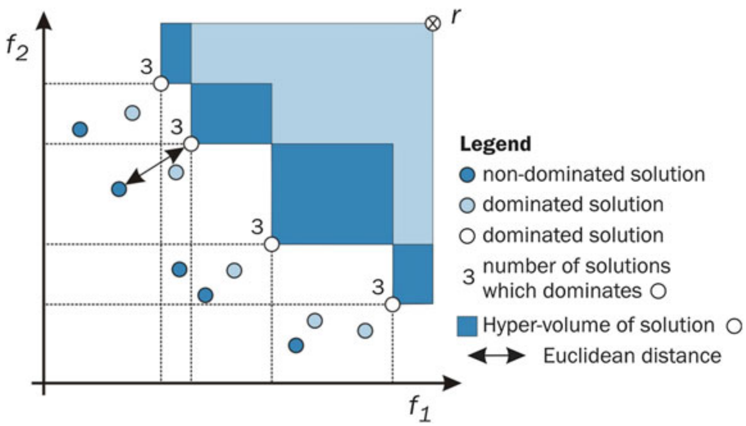


Fig. 5 Metrics for the criterion optimality of a project plan (Tackenberg et al. 2012)

$$\Delta S(x, F_k) := (f_1(I_{i+1}) - f_1(I_i))(f_2(I_{i-1}) - f_2(I_i)). \tag{29}$$

For a vector $f(I_i)$ of the individual I the distances to the next inferior individual f_1 (I_{i+1}) regarding project duration as well as the next inferior solution regarding project cost f_2 (I_{i-1}) are calculated. As we are dealing with a multicriteria scheduling problem, this metric can been deemed as one of the fairest and most useful existing methods for the evaluation of the optimality of a plan.

4.5 Variables and Hypotheses

In our experiments, the different variants of the planning problem are designated as independent variables. Clearly, the difficulty of the planning problem differs in regard to the number of tasks due to the possibility of an iterative execution of activities (*MSPSP-It*) as well as variable precedence relations due to an overlapping of acitivities (*MSPSP-ov*). All planning problems were solved by the three groups *humans*, *evolutionary metaheuristic* and *stochastic simulation*. Subsequently, the developed project plans form the basis for the evaluation of human competitiveness. Testing the quality of a plan, the required time for developing an efficient project plan becomes an independent variable. Whereas, the criteria for optimality and correctness of a project plan are interpreted as dependent variables. The following hypotheses are derived for analyzing the results of the laboratory experiment (Tackenberg 2016):

– H_1: It is expected that project plans developed by the evolutionary metaheuristic are human competitive.
– H_2: It is expected that an increase of execution time (group: participants) leads to a higher quality of the developed project plans.
– H_3: It is expected that the participants have developed project plans which are uniformly distributed along the known Pareto front.
– H_4: It is expected that the participants have developed project plans which represent their intended weighting of objectives.

4.6 Statistical Analysis and Results

We assume that the project plans developed by the evolutionary metaheuristic are human competitive for all variants of the observed *RCPSP*. In order to ensure comparability between the groups: *humans*, *evolutionary metaheuristic* and *stochastic simulation*, for each planning task the average processing time of the participants was set as a baseline. For the evolutionary metaheuristic and the simulation model it should be mentioned that this time covers the parameterization of the *XML*-files and the process of developing efficient plans.

For the investigated planning problems, the number of optimal plans of the 20 participants each is as follows: *RCPSP*: 16, *MSPSP*: 2, *MSPSP-ov*: 1 and *MSPSP-It*: 2 (Tackenberg 2016). While the distribution of the objective function

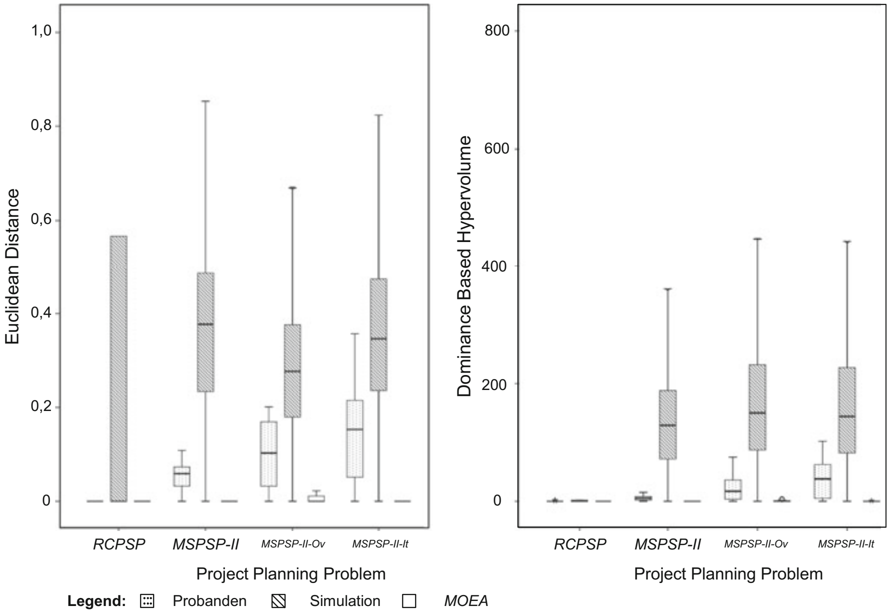


Fig. 6 Optimality of the developed project plans—Participants, Simulation Model, *MOEA*

vectors within the target space is very heterogeneous, the Box plots of the criterion for optimality are presented (cf. Fig. 6). The weak dominance of the plans developed by the *MOEA* compared to the solutions of the participants is statistically confirmed by the exact Mann-Whitney U-test ($U = 557.50$, $z = -8.840$, $\alpha = 0.01$).[1] Therefore, the null hypothesis H_{01} is rejected.

Consider the first front of each group for the planning problems given in Fig. 7. For the *RCPSP* which is characterized by a 1:1 relation between resources and tasks, all three groups fully identify the Pareto-Front. But, the extension of the planning problem (*MSPSP-II*, *MSPSP-ov*, *MSPSP-It*) leads to deviating first fronts. Clearly, the developed plans differ with regard to the distance to the known Pareto front as well as the diversity along the Pareto front. Since the efficient plans of the evolutionary metaheuristic meet all criteria for a later project implementation, the algorithm wins the regulated competition against the human players due to a higher achievement against the multicriteria objectives. Therefore, the project plans of the evolutionary metaheuristic are unreservedly human competitive (*criteria: H*).

Due to our definition of an individual and due to the fact that a schedule related to an individual is always feasible, the evolutionary metaheuristic as well as the stochastic simulation model fully comply with the criterion *correctness*. With this in mind, we evaluate the number of errors as well as their effect on the further project course only for the project plans of the participants. The Kruskal-Wallis test does not show a significant difference between the variants of the scheduling

[1]The statistical analysis was conducted with the statistical software package SPSS Version 19.0.

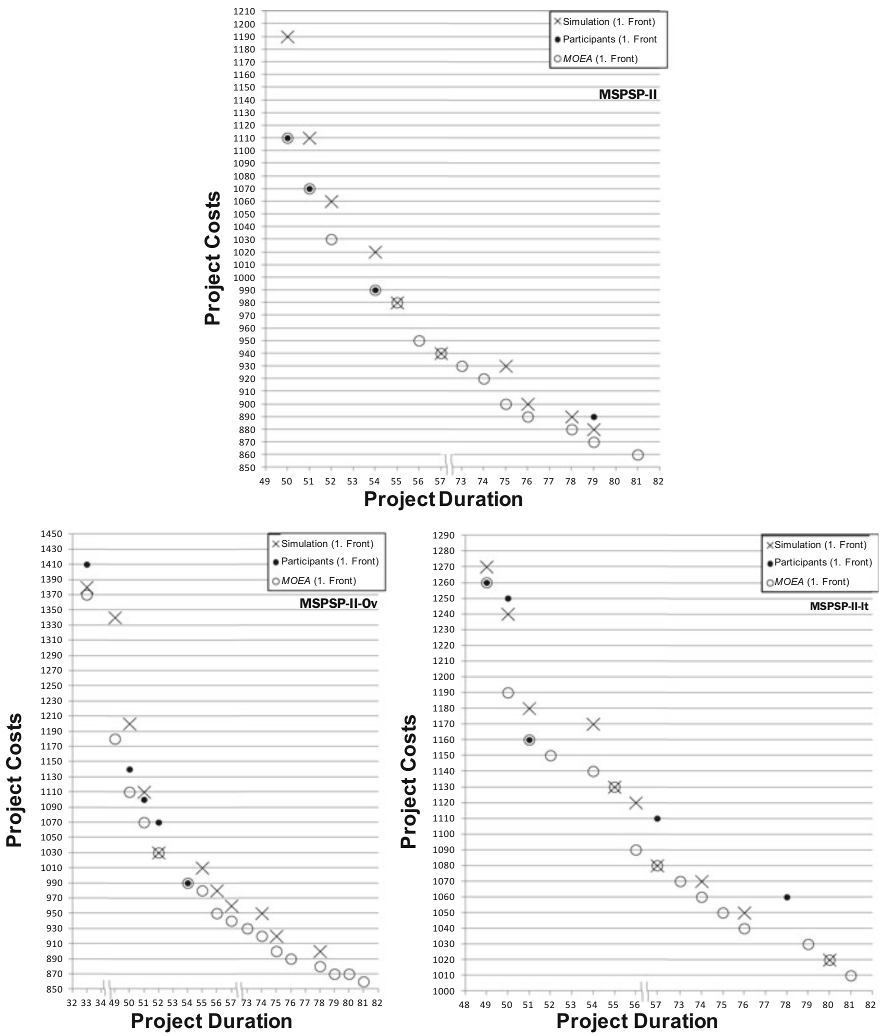


Fig. 7 First front of the three groups for the variants of the *MSPSP*

problem and the criteria for correctness; number of errors: $\chi^2 = 6.086$, $p = 0.193$, $\mathrm{d}f = 4$, effect of an error: $\chi^2 = 7.479$, $p = 0.113$, $\mathrm{d}f = 4$. But it should be noted that the developed project plans for the *RCPSP* contain the smallest number of errors.

The investigated planning problems differ regarding the extent of restrictions to be considered during plan development. It must therefore be assumed that the variant of the planning task may have an impact for the time, which is necessary for the participants to develop such a plan. First, we use the Kolmogorov-Smirnov test to show that the processing times are normally distributed. Second, an ANOVA is calculated and a significant correlation is statistically proven ($F = 16.067$, $p < 0.001$

df = 4), effect size: $\eta^2 = 0.392$. Third, we use the Games-Howell test to further investigate the correlation. It is shown that the effect in particular results from the differences between the *RCPSP* respectively *MSPSP* and the other variants of the *MSPSP*. For the *MSPSP* we also have analysed if there is a correlation between the required duration of the participants and the quality of the project plan. The results are summarised in Table 4. Even though there is a statistically provable link between the processing time and the plan quality, the observed effect must be regarded as weak $r < |0.4|$. Aside from the results of the statistical tests, the chart in Fig. 8 suggests that the hypothesis H_2 must be rejected.

One hundred participants developed project plans for the *MSPSP*. But only 13 participants were able to identify a plan which is optimal with regard to project costs and project duration and which also meets all further restrictions. Therefore, to measure the distribution along the Pareto front which consists of 14 project plans, each dominated plan was transferred into an efficient solution. This was achieved by the smallest number of necessary corrections of a participant's plan. With this in mind, we expected a uniform distribution of the 100 plans (100 plans/14 efficient plans = 7.14). The χ^2 test yields on a level of less than 0.001 a significant difference

Table 4 Correlation between the processing time of the participants and the quality of the project plans

Optimality	Euclidean distance	$r = -0.246$
		$p = 0.014$
	Dominance based hypervolume	$r = -0.229$
		$p = 0.022$
	Steps of correction	$r = -0.220$
		$p = 0.028$
Correctness	Number of errors	$r = -0.298$
		$p = 0.003$
	Steps of correction	$r = -0.336$
		$p = 0.001$

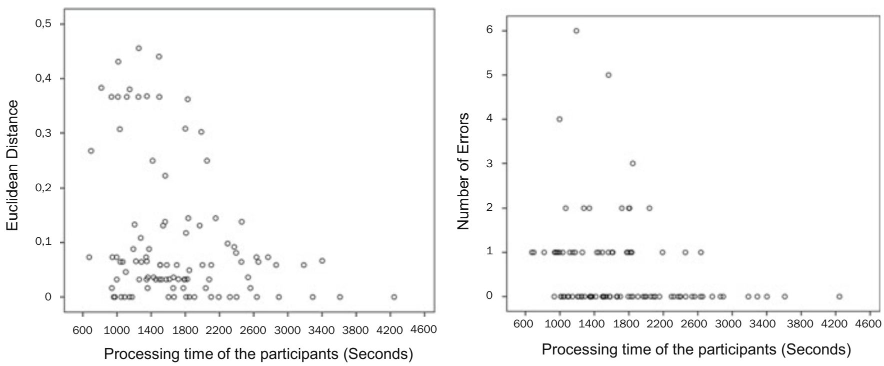


Fig. 8 Correlation between processing time and plan quality

between the observed and the expected values, $\chi^2(1, N = 100) = 167.680$, $p < 0.001$, d$f = 13$. Therefore, the hypothesis H_3 must be rejected.

Consider the distribution of the project plans given in Fig. 9, it is obvious that the participants putting more weight on the minimization of the project duration. Therefore, if a delay during a project progression occurs, there is no time buffer available. With this in mind, a variance from such a plan leads inevitably to a late project completion.

We also assume that the participants have developed plans which represent their intended weighting of the objectives. Therefore, we assigned each project plan to one of four clusters within the target space. We distinguish the following clusters according to the weighting of the two objectives (project costs, project duration):

1. significant emphasis on minimizing the project duration,
2. significant emphasis on minimizing the project costs,
3. proper balance on a simultaneous minimization of the project duration and project costs,
4. no targeted project plan development.

The graph below shows the relation between the self-assessment of the participants and the assignment to a specific cluster (cf. Fig. 10). To statistically investigate this issue the χ^2 test of Fisher is used. It verifies that there is no significant link between the intended weighting of the objectives and the planned course of the project ($\chi^2(1, N = 99) = 3.612$, $p = 0.417$, d$f = 4$) (Tackenberg 2016). Therefore, the hypothesis H_4 must be rejected. Clearly, nearly all participants who focused on a minimization of the project costs were not able to develop a plan with low cost.

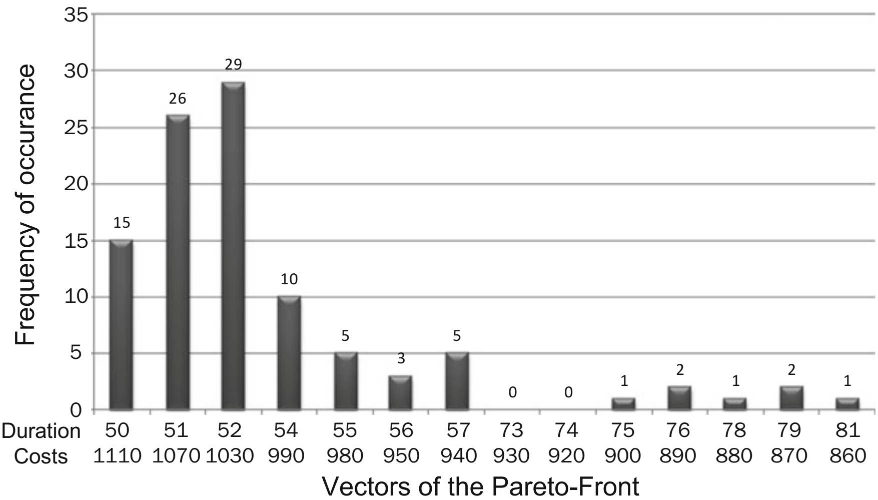


Fig. 9 Spreading of transferred plans along the Pareto front

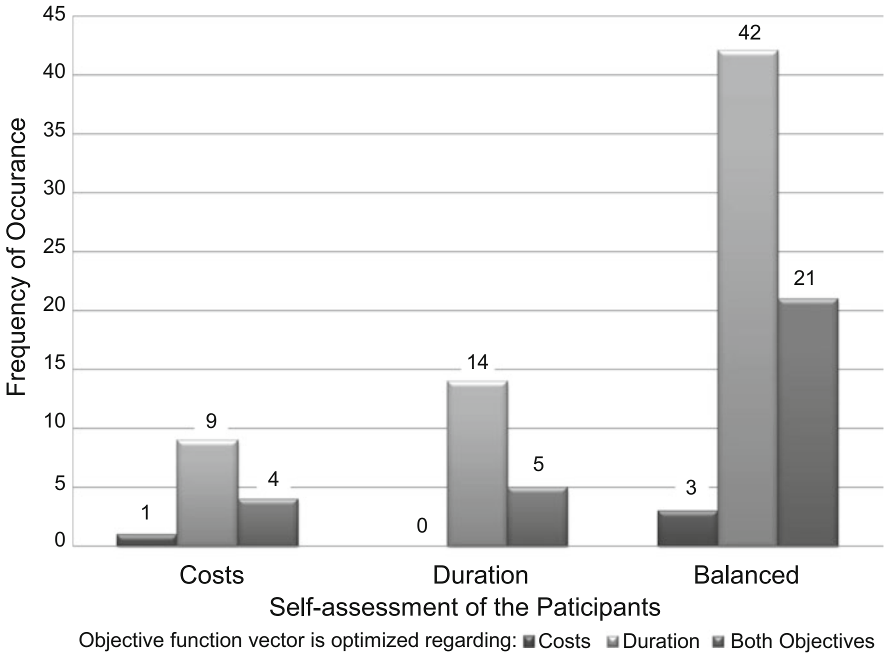


Fig. 10 Self-assessment of participants vs. location of the plan in the target space

5 Conclusion

In the previous sections we addressed the development of project plans. We introduced a project model to formalize the prospective description of a project implementation. The model enables the developed metaheuristic to identify human compatible solutions for representative cases of the multicriteria project scheduling problem. The algorithm was heavily influenced by the findings of Deb et al. (2001), Emmerich et al. (2005) and Zitzler et al. (1999).

A laboratory study was carried out to evaluate the human competitiveness of plans developed by the evolutionary metaheuristic. Thereby, all experimental groups (humans, stochastic simulation model, evolutionary metaheuristic) have received identical information regarding the scheduling problem. We investigated several scheduling problems, which represent different variants of the *RCPSP*. Comparing the plans identified by the *MOEA* and the humans, we observe that there is a considerable difference regarding the procedure of developing and optimizing the degree to which the project objectives have been met. Finally, we show that the metaheuristic generates human compatible plans for real world resource constrained project scheduling problems. Therefore, the approach offers project managers a suitable technique for the quantitative comparison of alternative project scenarios at an early planning stage. In fact, the laboratory study shows clearly that it is already difficult for humans to identify Pareto-optimal solutions for a resource constrained project planning problem with 15 activities.

In future papers we will present optimization studies of more complex resource constrained project scheduling problems. These problems will include workers with heterogeneous working hours. In addition, further project objectives will be identified and investigated to improve our statements on the quality of plans. The results to be expected will make an important contribution for the development of an assistance system for project planning and management.

Acknowledgment This work was supported by the German Federal Ministry of Education and Research according to Grant No. 01FL10011 and was administered by PT-DLR. This support is gratefully acknowledged. The project WiDiPro (www.widipro.de) proposes the optimal development of a knowledge-intensive service provision for architects and consulting companies.

References

Alvarez-Valdes R, Tararit JM (1989) Heuristic algorithms for resource constrained project scheduling: a review and an empiricial analysis. In: Słowiński R, Węglarz J (eds) Advances in project scheduling. Elsevier, Amsterdam, pp 113–134

Artigues C, Demassey S, Néron E (2008) Resource-constrained project scheduling. Models, algorithms, extensions and applications. Wiley, Hoboken, NJ

Bartusch M, Möhring R, Radermacher FJ (1988) Scheduling project networks with resource constraints and time windows. Ann Oper Res 16(1):201–240

Belhe U, Kusiak A (1995) Resource constrained scheduling of hierarchically structured design activity networks. IEEE Trans Eng Manag 42(2):150ff

Bellenguez-Morineau O, Néron E (2008) Multi-mode and multi-skill project scheduling problem. In: Artigues C, Demassey S, Néron E (eds) Resource-constrained project scheduling. Models, algorithms, extensions and applications. Wiley, Hoboken, NJ, pp 149–160

Brucker P, Drexl A, Möhring R, Neumann K, Pesch E (1999) Resource-constrained project scheduling: notation, classification, models, and methods. Eur J Oper Res 112(1):3–41

Cho J-H, Kim Y-D (1997) A simulated annealing algorithm for resource constrained project scheduling problems. J Oper Res Soc 48(7):736–744

De Reyck B, Herroelen W (1996) A branch-and-bound procedure for the resource-constrained project scheduling problem with generalized precedence relations. Onderzueksrapport No. 9613, Onderzueksrapport. Katholieke Universiteit Leuven. Leuven, Belgium

Deb K (2001) Multiobjective optimization using evolutionary algorithms. Wiley, Chichester

Deb K, Agrawal RB (1995) Simulated binary crossover for continuous search space. Complex Syst 9(2):115–148

Deb K, Goyal M (1996) A combined genetic adaptive search (GeneAS) for engineering design. Comput Sci Informat 26(4):30–45

Dörner D (2010) Die Logik des Misslingens. Strategisches Denken in komplexen Situationen. 9. Aufl. Reinbek. Rowohlt

Emmerich M, Beume N, Naujoks B (2005) An EMO algorithm using the hypervolume measure as selection criterion. In: Coello Coello CA, Aguirre AH, Zitzler E (eds) Evolutionary multicriterion optimization. Third international conference, EMO 2005. Springer, Berlin, pp 62–76

Feldman EJ (2009) Concorde and dissent. Explaining high technology project failures in Britain and France. Cambridge University Press, Cambridge

Fırat M, Hurkens C (2011) An improved MIP-based approach for a multi-skill workforce scheduling problem. J Sched 15(3):1–18

Funke J, Fritz A (1995) Über Planen, Problemlösen und Handeln. In: Funke J (ed) Neue Konzepte und Instrumente zur Planungsdiagnostik. Dt. Psychologen-Verl, Bonn, pp 1–45
Funke J, Glodowski A-S (1990) Planen und Problemlösen: Überlegungen zur neuropsychologischen Diagnostik von Basiskompetenzen beim Planen. Zeitschrift für Neuropsychologie 2:139–148
Hans EW, Herroelen W, Leus R, Wullink G (2007) A hierarchical approach to multi-project planning under uncertainty. Omega 35(5):563–577
Hartman F, Ashrafi R (2004) Development of the SMARTTM Project Planning framework. Int J Proj Manag 22(6):499–510
Hartmann S (1997) A comparative genetic algorithm for resource-constraint project scheduling. Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel. Universität Kiel, Kiel
Hartmann S (2001) Project scheduling with multiple modes: a genetic algorithm. Ann Oper Res 102:111–135
Hartmann S, Briskorn D (2010) A survey of variants and extensions of the resource-constrained project scheduling problem. Eur J Oper Res 207(1):1–14
Hayes-Roth B, Hayes-Roth F (1979) A cognitive model of planning. Cognit Sci 3(4):275–310
Jäger AO, Süß H-M, Beauducel A (1997) Berliner Intelligenzstruktur-Test (BIS), Form 4. Hogrefe, Göttingen
Kolisch R, Drexl A (1996) Adaptive search for solving hard project scheduling problems. Naval Res Logist 43:23–40
Kolisch R, Drexl A (1997) Local search for nonpreemptive multi-mode resource-constrained project scheduling. IIE Trans 29:987–999
Kolisch R, Hartmann S (2006) Experimental investigation of heuristics for resource-constrained project scheduling: an update. Eur J Oper Res 174(1):23–37
Koza JR, Keane MA, Streeter MJ, Mydlowec W, Jessen Y, Lanza G (2005) Genetic Programming IV Routine human competitive machine intelligence. Springer, New York
Lechler T (1997) Erfolgsfaktoren des Projektmanagements. Lang, Frankfurt am Main
Li H, Womer K (2009) Scheduling projects with multi-skilled personnel by a hybrid MILP/CP benders decomposition algorithm. J Sched 12:281–298
Mendling J (2008) Metrics for process models: empirical foundations of verification, error prediction, and guidelines for correctness. Springer, Berlin
Néron E, Baptista D (2002) Heuristics for the multi-skill project scheduling problem. In: International symposium on Combinatorial Optimization (CO'2002), Paris
Pietras CM, Coury BG (1994) The development of cognitive models of planning for use in the design of project management systems. Int J Hum Comput Stud 40(1):5–30
Rasmussen J (1987) The definition of human error and a taxonomy for technical system design. In: Rasmussen R, Duncan K, Leplat J (eds) New technology and human error. Wiley, Chichester
Scholnick EK, Friedman SL (1987) The planning construct in the psychological literature. In: Friedman SL, Scholnick EK, Cocking RR (eds) Blueprints for thinking: the role of planning in cognitive development. Cambridge University Press, Cambridge, pp 3–38
Shtub A, Bard JF, Globerson S (2005) Project management. Processes, methodologies, and economics, 2nd edn. Pearson Prentice Hall, Upper Saddle River, NJ
Tackenberg S (2016) Konzeption und Entwicklung einer mehrkriteriellen evolutionären Metaheuristik zur Planung wissensintensiver Dienstleistungen. Shaker, Aachen
Tackenberg S, Schmalz C, Duckwitz S, Schlick CM (2012) Multi-objective development and optimization of plans for improving the quality and productivity of service provision. In: Tăchiciu L, Ioncică M, Iosif A, Pantelescu A (eds) The 22nd RESER international conference: service and economic development, local and global challenges, 20–22 Sep 2012
Tan KC, Khor EF, Lee TH (2005) Multiobjective evolutionary algorithms and applications. Springer, London
Thomas M W (1998) A Pareto frontier for full stern submarines via genetic algorithm. Department of Ocean Engineering, Massachusetts Institute of Technology

White D, Fortune J (2002) Current practice in project management—an empirical study. Int J Proj Manag 20(1):1–11

Zitzler E, Deb K, Thiele L (1999) Comparison of multiobjective evolutionary algorithms: empirical results, TIK-Report, 70. Swiss Federal Institute of Technology (ETH), Zürich

Zitzler E, Laumanns M, Bleuler S (2004) A tutorial on evolutionary multiobjective optimization. In: Gandibleux X, Sevaux M, Sörensen K, T'kindt V (eds) Metaheuristics for multiobjective optimisation. Springer, Berlin, pp 3–38

Fields of Action for Work Design in Industrie 4.0

Tim Jeske and Frank Lennings

Abstract

The proceeding digitalization in industrial context is named as transformation to Industrie 4.0. This term has been presented to the public already 5 years ago and was chosen to point out the meaning of the corresponding changes for industrial production. Since then the term has become very present in media and is often discussed controversially. To identify its perception in small and medium enterprises—the main part of the German industry—a survey has been conducted. Queried were nearly 500 participants in management positions from the Metal and Electrical Industry. The survey's results prove that the term is very well known while its meaning is often unclear. Furthermore, it was found that Industrie 4.0 is introduced preferentially in production, planning/control, logistics, supply chain management and storage. These fields of usage and the tasks characteristically performed within, indicate fields of action for work design. Structured by informational and energetic work forms as well as technical, organizational and personal aspects, many options for work design are described in detail.

Keywords

Industrie 4.0 • Survey • Work design • Metal and Electrical Industry

T. Jeske (✉) • F. Lennings
Institut für angewandte Arbeitswissenschaft (ifaa), Uerdingerstraße 56, 40474 Düsseldorf, Germany
e-mail: t.jeske@ifaa-mail.de; f.lennings@ifaa-mail.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_4

1 Digitalization and Industrie 4.0

Digitalization is an ongoing process which already started years ago. Initial point was the development of digital technology—especially the invention of the computer. Since then the technology's advancement can be illustrated according to Moore's law: The number of transistors per unit area in microelectronic components is doubled every one and a half year (Eigner et al. 2012).

Today digital technologies are already present in private and in occupational life. For Germany this means 83 % of humans and 92 % of companies are using computers (in 2015; Statista 2016a, b). For producing companies digitalization is also described as transformation to Industrie 4.0. This term refers to the previous development of the industrial production and its major steps: The first step—Industrie 1.0—was the invention and usage of the steam machine. Afterwards the second step—Industrie 2.0—was characterized by the division of labor and followed by the third step—Industrie 3.0—which describes automation by stored program control (SPC) etc. Characteristic for Industrie 4.0 is digital information management based on networking. This means gathering digitalized information, processing them and providing them along the production process according to the requirements of every single process section, workplace and employee. Due to that, Industrie 4.0 is expected to increase productivity and lower costs. For example, costs of complexity are expected to lower up to 70 % (Bauernhansl 2014). The term Industrie 4.0 was firstly presented to the public in 2011 at Hanover Fair, the world biggest industrial fair.

Despite the expected advantages Industrie 4.0 is often discussed controversial and the transformation is developing very differently: Some companies have made large progress, while others have not yet started and are still searching for their specific approach.

2 Approach

Due to that the "Institut für angewandte Arbeitswissenschaft" (ifaa) conducted an online survey within the German Metal and Electrical Industry. The survey aimed on assessing the transformations progress and the companies' requirements for support. To do so 24 questions have been developed and assessed within a preliminary study. The introductory questions covered the characteristics of the sample, such as age and function of the participants as well as the size of their companies. The further questions can be grouped thematically and are examining the comprehension and importance of Industrie 4.0, the state of its implementation and the expected development. Some of these questions where already used in a previous study which was paper based and conducted from October 2014 to March 2015 (Jeske 2015). This enabled single comparisons in terms of a longitudinal study.

The survey was conducted all over Germany about 2 months from June to July 2015 by using the online platform SosciSurvey.

3 Results

Altogether 498 participants took part in the survey. They specified their occupational task area in their companies mostly as management board, production or personnel management. The age of the participants was predominantly between 45 and 59 years. Their companies' size was mainly small or medium (<500 employees) (ifaa 2015).

It was found that 92.1 % of the participants know the term Industrie 4.0. In contrast to that, only 32.8 % answered to have a clear (subjective) understanding of its meaning. Compared to the previous survey the terms knowledge (previously 76.2 %) as well as its understanding (previously 23.3 %) increased (ifaa 2015; Jeske 2015).

The importance of Industrie 4.0 for the future in general was assumed as "high" or "very high" by 88.4 % of the participants (in the previous survey 76.1 %). Contrasting that, a "high" or "very high" importance for the future of the own company was assumed by 69.1 % of the participants (in the previous survey 60.5 %) (ifaa 2015; Jeske 2015). Accordingly, both importances have been assumed 15 % higher than before. The difference between the general importance for the future and the importance for the future of the own company is constantly 27 %.

Queried about their preferred field of usage for implementing Industrie 4.0, most participants answered production (63.82 %, previously 62.82 %) and planning/control (55.00 %; previously 59.88 %) (Jeske 2015). These preferences were followed by logistics (50.29 %), supply chain management (36.47 %), assembly (33.53 %) and storage (32.65 %). The latter fields of usage allow no or only limited comparisons to the previous results since the options for answering were more detailed in the present survey. However, in the previous survey the fields of usage related to logistics were also among the preferred fields of usage for implementing Industrie 4.0.

The fields of usage in which most activities for implementing Industrie 4.0 take place (see Fig. 1) correspond largely with the aforementioned preferences. Application, introduction/rollout or piloting/testing area were stated particularly for the following fields of usage: production (22.3 %), assembly (17.0 %), planning/control (16.5 %), logistics (15.7 %) and storage (15.0 %). Additionally, specific planning was stated mainly for planning/control (15.8 %) and production (14.6 %). On average about all potential fields of usage in 37 % of the companies only general information are collected. On the same average 38 % of the companies state no activities. While most companies state activities at least in one potential field of usage, there are 15 % of companies which state no activity at all—those are completely inactive according to Industrie 4.0.

The activities for implementing Industrie 4.0 in all potential fields of usage have proceeded in larger companies substantially more than in smaller ones. This can be proven by significant correlations between the companies' size and their activities in every single potential field of usage ($n = 286–304$; $r_S = .120$–$r_S = .262$; $p < .000$–$p = .036$; $\alpha = .05$). This result is the same as in the previous survey: Also in this survey significant correlations between the companies' size and their activities

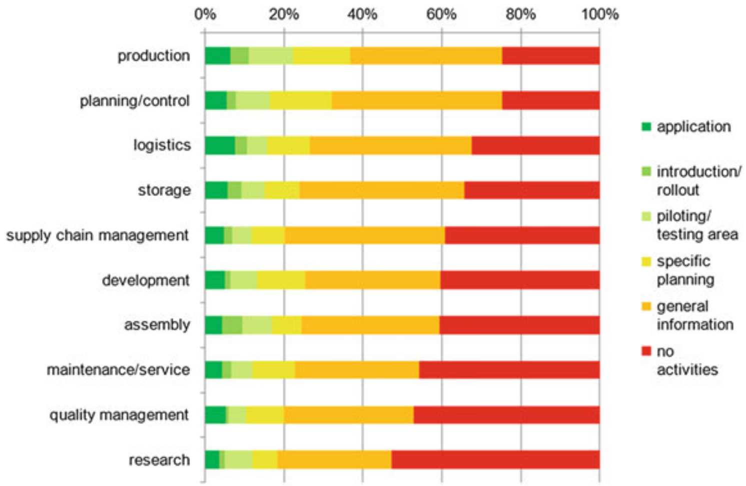


Fig. 1 Activities for implementing Industrie 4.0 by fields of usage (sorted by inactivity)

showed that smaller companies are less active in implementing Industrie 4.0, while larger companies have proceeded substantially more (Jeske 2015).

4 Fields of Action for Work Design

Since the term Industrie 4.0 is already very well known it is firstly important to improve its comparatively low degree of understanding. A clear understanding of Industrie 4.0 and its meaning for the future is necessary for being basis to explain its large potential for work design. While the term is often described and discussed very abstractly or academically, most examples of good-practice are only little known. Therefore good-practices have to be made accessible to a broad general and professional public—suitable for that are for example brochures as well as model factories. This will make the transformations effects vivid and allows discussing Industrie 4.0 referring to specific existing industrial applications instead of speculative scenarios. Thus, the advantages of digitalization as well as related risks will become more transparent. Furthermore, the transfer of good-practices between companies will be facilitated. Consequently, the understanding of Industrie 4.0 should increase while the empirically found difference between the assumed future importance in general and for the own company should decrease.

Generally a need-based work design is necessary in all departments of a company and hence in all potential fields of usage for Industrie 4.0. Due to the afore presented survey and its results some fields of usage can be treated preferentially. Those are the ones in which the transformation process has most proceeded: production, assembly and planning/control. Within these fields of usage very heterogeneous work systems are required since tasks to be performed there consist of directing as well as executing elements in different weightings. Therefore, informational and energetic work forms as well as combinations of both have to be taken into account. To do so, for both work forms a structured analysis by technical, organizational and personal aspects is conducted. Additionally, the work systems input and output have to be analyzed regarding the information transfer, which is characteristically for Industrie 4.0.

4.1 Information in Industrie 4.0

The network of digital systems—the character of Industrie 4.0—enables a high availability of information which is close to real time and leads to a high exchange of information. The handling and mastering of the related amount of information is a complex task. Therefore, specific measures and procedures have to be developed which ensure an adequate empowering of employees who manage information. For this purpose, there are two substantial technical approaches which are realizable by cyber-physical systems: modularization and decentralization. Based on cyber-physical systems information flows have to be organized to ensure that necessary information is available at the right time at the right place in an appropriate form.

From a personnel point of view, employees require an adequate competency for selecting information, for choosing communication media and for communicating to avoid misunderstandings and misinterpretations in the high number of expected exchange and coordination processes. This includes the capability for self dependent handling of high information availability and accessibility, as they are possible, for example by e-mail or smart phones.

4.2 Informational Assistance

The design potential for supporting informational work results from the advancement of technical opportunities to gather information, processing them and providing them. In consequence many approaches arise for further reducing workload by cognitive relief.

4.2.1 Technical Approaches

Technical approaches for informational assistance comprise access to and illustration of information within the work process.

Digital representations of information without any paper are advantageous when paper based documentations cannot guarantee appropriate topicality due to a high frequency of changes on product or on process specifications. Additionally, mistakes caused by using out-dated paper based documentations are avoided as well as effort on creating those documentations or proving their topicality. For clearly identifying, products or workpiece holders can be equipped with barcodes, QR-codes or RFID-tags which are assigned to the related order. This allows assigning and illustrating the orders current specifications digitally. To do so, inter alia displays or data glasses are applicable. In addition, information like task descriptions or work instructions can be illustrated depending on the current progress of an order. For this purpose also projection systems and augmented reality techniques on data glasses, tablet computers or smart phones are suitable. Thereby, the human's current field of view or the current camera images are complemented with information and virtual elements or superimposed by them. This leads to a direct relation between the reality and the associated information. For example, in the current research project motionEAP it is explored, how information can be displayed directly into a workspace by means of a combined camera and projection system. For this purpose, components to be joined shall be color-coded or the necessary directions of movement shall be illustrated by appropriate indicators (e.g. arrows). Information is therefore presented depending on the actual work situation. This also allows using such systems for quality assurance. If mistakes occur, which can be detected by the camera system, a notice can be displayed automatically and induce correction measurements. This leads to a high level of quality and allows using mistakes directly for learning.

The information and communication technology helps also supporting mobile work activities such as in the repair and maintenance area or in the field service. Besides the various possibilities for contextual presentation of information here the availability of product- or task-specific information is important. Searching for specific details can often be done with database support which reduces the necessary time and effort. Moreover, it is conceivable that information, which is based on the recognition of the contents of a camera image, is combined with historical data such as frequent malfunctions and possible solutions and is structured on a display.

These approaches require the digital availability of all necessary information. In case of location-independent use also secure data connections with sufficient bandwidth are required. The automatic situational provision of information relieves humans from searching them. By linking information with objects e.g. by the help of augmented reality or projection systems, also misunderstandings are prevented. Finally the handling effort for paper-based documents is avoided; dependent on the particular configuration of the assistance both hands can remain free to perform the actual work.

4.2.2 Organizational Approaches

Organizational approaches for informational assistance facilitate planning and coordination by taking many information into account and thereby enable employees to handle large amounts of information.

Such a planning system can be used e.g. in the context of personnel scheduling and serve to take into account not only criteria such as the presence and skills of employees, but also the ergonomic conditions of each combination of employee and workplace. In this way, a contribution can be done to the preservation of health as well as to the human working capacity and performance (INQA 2005). For this purpose, it has firstly to be determined which of the present persons can be assigned to which tasks by taking into account the required and available qualifications/ skills. On this basis, human and task profiles can be used to apply established methods for the ergonomic evaluation of stress-strain situations such as the Key Indicator Method (BAuA 2001, 2002, 2012) with computer assistance. An appropriate summary of the individual assessments' results enables to determine assignment combinations which lead to a minimal overall stress for the considered group of employees (Jeske et al. 2014). If during this procedure the existing qualifications are initially not considered, qualifying potential can be identified which enables a further reduction of the groups' overall stress.

A software tool for shift planning—often referred to as "shift-doodle"—has been developed in the research project KapaflexCy; it is especially designed for the coordination of additional shifts and is based on a participatory approach (Spath et al. 2013). If additional shifts have to be assigned specific requests are sent to all persons, which can be scheduled in principle, due to their qualification and according to the regulations of the law on working hours. If the employees own a personal smartphone and agreed to a related use, they receive these requests on their devices. For all others corresponding terminals are located in the company. For answering these requests, a company-specific adjustable period of e.g. 2 days has to be specified. After this period, fixed rules are used for selecting employees out of those who accepted the request and for notifying them. These rules can be tailored to the specific needs of the company and can consider, for example, the current balance of the working time account. In order to identify selection criteria and to develop rules for applying them a close coordination with the works council has been proven to be advantageous. In this way, operational and personal needs in the shift planning can be coordinated with the help of the information and communication technology. A successful transfer of this approach on the regular shift planning is expected to allow an improved consideration of the individual life situation of employees (childcare, care of the elderly etc.). Thus, the access to the labor market for persons who have to consider time restrictions is facilitated. In this context, also the principles of ergonomic shift design such as rapid rotation and forward shift change can be implemented in such a planning system.

Work tasks can be executed location-independently—e.g. in a home office or on the road—if they are appropriate according to the required information, materials and equipment. This is mostly the case for directing tasks, since their execution mainly requires the handling of information. As a result, location-independent work execution is only possible if a sufficient infrastructure exists, to transmit those data between companies and employees, which are basis and result of the work. Thereby, a sufficient data transmission infrastructure is characterized by availability, transmission speed (bandwidth), and data security. In contrast to directing tasks execution tasks usually require appropriate material and equipment, therefore a location-independent execution is usually impossible. Against this background, location-independent work execution is inter alia already applied in the field of remote maintenance and remote control. The flexibility related to location-independent work execution facilitates the access to the labor market for persons, who are limited in their mobility due to their life situation or due to personal reasons and thus are limited in their regional choice of a workplace.

Since both the improved shift planning as well as the location-independent work execution facilitate the access to the labor market, these approaches contribute to lessen the expected reduced number of economically active persons and the related lack of specialists which are caused by the demographic change. A combination of both approaches arises when periods of joint presence in the company are coordinated and other tasks are carried out in the home office or location-independently. In this context, in principle, also a division of the workday into periods with presence in the company and in the home office is imaginable. Thereby, also the individual circadian rhythm might find a better consideration.

The software-based combination and handling of information relieves working persons who are performing the related mostly complex planning tasks. Besides this a relief is expected due to high data availability for the described temporal and spatial flexibility of work execution.

4.2.3 Personal Approaches

Personal approaches for informational assistance aim on processing and presenting adequate information for facilitating decisions as well as for supporting creativity and learning.

Comprehensive information about a production system and the current status of the order processing, can be used in case of errors—that occur despite predictive maintenance—for developing and evaluating action alternatives based on indicators such as OEE or adherence to delivery dates. This is very important since in such cases decisions about action alternatives usually are of great consequence and have to be made under time pressure. For the necessary simulations historical data can supplement current information to improve the validity and reliability of the simulation results. Historical data are for example specific set-up times or cycle times of individual processing stations as well as statistical information about the likelihood of their adherence.

Historical data can also be used to support the work of design engineers. For this purpose, potentially usable repeat parts can be identified and proposed automatically based on a current construction progress. Equally, specific errors that occur in production or assembly of certain components can be visualized to implement appropriate changes in the current design. The effort for manual search of repeat parts and unnecessary new designs as well as the consequences of design mistakes can be reduced in this way.

A further prevention of mistakes can be realized if the systems for displaying orders and instructions which are already described in Sect. 4.2.1 show information not only in dependence of the current progress of the work, but also in dependence of the individual knowledge and experience of the employee. This allows varying the amount of information which will be displayed for the next step accordingly: Employees with little experience in the execution of a specific task receive detailed information and are thus protected from overstrain, while employees with a lot of experience in the execution of the same task are protected against underload by giving them relatively little information limited to key aspects such as a custom variation of a usual standard. Combined with the immediate feedback of mistakes (see Sect. 4.2.1) and matching assistance, a highly individualized learning within the work process is enabled and encouraged; thereby, the amount and detail of the information displayed is reduced stepwise, depending on the learning progress.

In this context, the support system for personnel scheduling described in Sect. 4.2.2 can also be extended to ensure the required knowledge and a high degree of training for specific work tasks by assigning them to each employee at the latest after a fixed period of time.

As shown by the example of simulation-based decision support the processing of existing information can contribute substantially to cognitive relief in complex decision situations. Likewise, information is useful for designing learning within the work process as well as for supporting the preservation of knowledge and a high degree of practice.

4.3 Energetic Assistance

The design potential for supporting energetic work results from the technical advancement in general combined with the advancement of information technology (especially gathering and processing). In consequence many approaches arise for further reducing workload by physical relief.

4.3.1 Technical Approaches

Technical approaches for energetic assistance comprise systems to which complete tasks can be transferred and thus allow a complete removal of the regarding workload.

For example, simple transport tasks can be transferred completely to driverless transport systems. In contrast to systems used today, the future use of such aids shall be considerably simplified and thus become more flexible. For this purpose, it is analyzed in the research project "FTF out of the box" how setup efforts can be reduced and how the assignment of tasks can be simplified. To that, those systems shall be set up without leading marks in or on the ground and shall accept transfer orders on call or through gestures (Dohrmann et al. 2014). The latter simplifies not only the daily use, but also the required training measures for their implementation.

A technical support for inventory tasks can be realized by combining the use of RFID tags with the operation of autonomous flying robots. For this purpose, within the research project "InventAIRy" storage stocks are equipped with RFID tags so that they can be identified contact-free. Thus, robots autonomously flying through a storage can combine their own position with the position of detected RFID tags which allows a systematically capturing of stock grounds and stock inventory. Assuming appropriate airworthiness and weather conditions, such flying robots can in principle also be used outside of storages and capture the inventory of mostly ample external storage areas. In this way inventories can be carried out continuously and ensure the timeliness and accuracy of booked inventory.

Both approaches show how simple tasks can be transferred to technical systems. As a result, employees are relieved from monotonous and partly physically demanding work tasks as well as overstressing of the musculoskeletal system is prevented.

4.3.2 Organizational Approaches

Organizational approaches for energetic assistance reduce workload by transferring elements of tasks to a robot by implementing a human-robot collaboration. Thereby, the specific strengths of humans and robots can be combined selectively.

For example, the welding of components requires a precisely positioning to ensure a high-quality welding seam. This element of the task is suitable for the transfer to industrial robots, since they have both a high positioning accuracy as well as a high load capacity which is required for heavy components (Busch et al. 2012). In addition, robotic systems may vary the position of the pre-aligned components, depending on the progress of the welding process. Thus, for example, the weld seam on a pipe can be performed completely in the flat position which is favorable for welding. Thereby, the employees task is to handle the welding machine on the basis of his dexterity and his specific welding experience. For example, the employee can derive measures to improve the welding process or its results based on the characteristics of the occurring arc.

To ensure the employees safety against collisions while using this form of collaboration even without enclosure or fence around the robot, a variety of information is required. This information is considered for controlling the robot and does not only affect the work task to be performed, but also the current positions of human and robot. Besides avoiding collisions also approaches are implemented to limit the consequences of collisions for humans. For this purpose, inter alia lightweight robots are suitable; they are characterized by a lower load capacity than conventional industrial robots and a lower mass which leads to a lower risk potential.

Thus, the human-robot collaboration can substantially contribute to relieve employees and their musculoskeletal system from monotonous and physically-demanding task elements. This concerns not only the described positioning of heavy components for welding, but also the provision of items from bulk containers and other transport containers or screw driving. Thus overstress and its consequences are prevented.

A further contribution for relieving the musculoskeletal system can be realized by using the already described support system for personnel scheduling with consideration of ergonomic evaluation methods (see Sect. 4.2.2). For this purpose, rotation principles are to be developed which propose task assignments by excluding the current/last assignments. This leads to a systematic and regular change of stress-strain situations.

4.3.3 Personal Approaches

Personal approaches for energetic assistance have to be chosen when the design of a task cannot be improved and neither the complete task nor elements of the task can be transferred to a technical system (mostly a robot). In these cases, exoskeletons can be appropriate.

Exoskeletons are energetic support systems that can be worn as outboard skeleton on the body and are used to hold loads in place of the human skeleton. Some of these systems use an additional energy source and are suitable to not only relieve humans, but also to expand their abilities—for example, to allow the lifting of heavy loads. An exoskeleton designed as a sitting aid is currently tested in the automotive industry (Nördinger 2015). For this purpose, the exoskeletons structure is designed to be worn on the back of the legs (also while walking little distances) and can be fixed in certain positions in order to serve as a seat.

The physical assistance by exoskeletons relieves employees and helps preserving their working capacity and performance. In contrast, an expansion of human abilities through exoskeletons with additional/external energy source does not necessarily lead to a relief; therefore, it has to be examined in each individual case—also regarding long-term effects.

5 Summary

Based on two surveys conducted with professionals and executives in the Metal and Electrical Industry it was shown that the knowledge and understanding of the concept of Industrie 4.0 have raised within approximately 6 months. Equally, the estimated importance of the topic has increased.

Since the understanding of the term is still low and the importance for the own company is estimated often lower than the general importance, there is continuously a high demand for explanation of the terms contents as well as for illustration by means of transferable practical examples. Industrie 4.0 is currently implemented mainly in the fields of production, assembly, planning/control and logistics, so that this results in priority fields of action for work design.

In light of networking and high availability of information which are characteristic for Industrie 4.0, employees require competencies for communication and media usage to proper handle large amounts of information. The handling and use of available information can be supported substantially by appropriate informational assistance systems. These systems generate solutions for specific problems based on transparent function and specified rules, which have to be checked, if necessary also changed and finally released by responsible employees. Additionally, a situational and possibly contextual provision of selected information can contribute to improving work processes. Besides this, the availability of information is also useful for supporting energetic work forms. For this purpose, information is used for controlling technical assistance systems such as automated guided vehicle systems or the human-robot collaboration.

To select context-appropriate assistance systems for supporting informational as well as energetic work form especially small and medium enterprises need appropriate guidelines. This also applies for identifying required skills necessary for the implementation and use of assistance systems.

Since digitalization respectively the transformation to Industrie 4.0 is a very dynamic process further research is necessary for designing the transformation process by developing adequate solutions which are adapted to the digitalization's actual progress.

Acknowledgements This research was supported by the ifaa's member associations which motivated participants and thus supported the survey's success substantially.

References

BAuA (2001) Leitmerkmalmethode-Heben, Halten, Tragen. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin und Länderausschuss für Arbeitsschutz für Sicherheitstechnik (ed)

BAuA (2002) Leitmerkmalmethode-Ziehen, Schieben. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin und Länderausschuss für Arbeitsschutz für Sicherheitstechnik (ed)

BAuA (2012) Leitmerkmalmethode-Manuelle Arbeit. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin und Länderausschuss für Arbeitsschutz für Sicherheitstechnik (ed)

Bauernhansl T (2014) Die Vierte Industrielle Revolution – Der Weg in ein wertschaffendes Produktionsparadigma. In: Bauernhansl T, ten Hompel M, Vogel-Heuser B (eds) Industrie 4.0 in Produktion, Automatisierung und Logistik. Springer Vieweg Verlag, Wiesbaden, S. 5–S. 35

Busch F, Thomas C, Deuse J, Kuhlenkötter B (2012) A hybrid human-robot assistance system for welding operations – methods to ensure process quality and forecast ergonomic conditions. In: Hu SJ (ed) Technologies and systems for assembly quality, productivity and customization – Proceedings of 4th CIRP conference on assembly technologies and systems (CATS). University of Michigan, Ann Arbor, pp 151–154

Dohrmann L, Podszus F, Ullmann G, Overmeyer L (2014) Mensch-Maschine-Interaktion für Fahrerlose Transportfahrzeuge – Methode zur Beauftragung von interaktiven Transportsystemen. Industrie Management 6:21–24

Eigner M, Gerhardt F, Gilz T, Mogo Nem F (2012) Informationstechnologie für Ingenieure. Springer Vieweg, Berlin

Hammer W (1997) Wörterbuch der Arbeitswissenschaft – Begriffe und Definitionen. Carl Hanser, München

ifaa – Institut für angewandte Arbeitswissenschaft (2015) ifaa-Studie - Industrie 4.0 in der Metall- und Elektroindustrie. Institut für angewandte Arbeitswissenschaft (ed)

INQA (2005) Initiative Neue Qualität der Arbeit, Thematischer Initiativkreis 30, 40, 50plus – Gesund arbeiten bis ins Alter: Demographischer Wandel und Beschäftigung. Plädoyer für neue Unternehmensstrategien. Memorandum, Dortmund

Jeske T (2015) Industrie 4.0 in Deutschland. In: Professur Arbeitswissenschaft und Innovationsmanagement (ed) Innteract Conference. Mensch 2020. Verlag aw&I Wissenschaft und Praxis, Chemnitz, pp 12–20

Jeske T, Brandl C, Meyer F, Schlick C (2014) Personaleinsatzplanung unter Berücksichtigung von Personenmerkmalen. In: Gesellschaft für Arbeitswissenschaft (ed) Gestaltung der Arbeitswelt der Zukunft. GfA-Press, Dortmund, pp 327–329

Nördinger S (2015) Exoskelette: Mechanische Muskeln stählen Werker-Rücken. Produktion 54:32

Spath D, Gerlach S, Hämmerle M, Schlund S, Strölin T (2013) Cyber-physical system for self-organised and flexible labour utilisation. In: 22nd International conference on production research (ICPR 22), IFPR (International Foundation for Production Research), Brasilien

Statista (2016a) Anteil der Computernutzer in Deutschland von 2004 bis 2015. http://de.statista.com/statistik/daten/studie/3102/umfrage/quote-der-computernutzer-in-deutschland-seit-2004/. accessed on 17.05.2016

Statista (2016b) Anteil der Unternehmen in Deutschland mit Nutzung von Computern in den Jahren 2005 bis 2015. http://de.statista.com/statistik/daten/studie/151762/umfrage/anteil-der-unternehmen-mit-nutzung-von-computern-in-deutschland/. accessed on 17.05.2016

Skilled Workers: Are They the Losers of "Industry 4.0"?

Georg Spöttl

Abstract

We are currently experiencing a re-orientation towards industrial work under the keyword "Industry 4.0". The structural change towards a service economy, proclaimed in the twentieth century, has been caught up by the industry. The share of the industry in terms of gross value added has been comparatively stable in Germany since the 1990s and today amounts to around 25 % (cf. Ahrens and Spöttl, Digitalisierung industrieller Arbeit, Edition Sigma, Baden-Baden, 2015; Bauernhansl, Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendungen—Technologien—Migration, Springer Vieweg, Wiesbaden, 2014, p. 7). "Industry 4.0" has become a new key term in the context of the "Future of the World of Work". There are a number of considerations and activities and a stepwise increasing use of Cyber-Physical Systems (CPS) in order to accelerate the implementation of "Industry 4.0". This has a considerable impact on the skilled employees, above all on skilled workers, master craftsmen and technicians. One of the central questions is whether this group of employees will still find its place in tomorrow's "intelligent" and digitalized production. This question will be discussed in this article.

Keywords

"Industry 4.0" • Skilled work • Competence development • Occupations • Change processes • Paradigm change

G. Spöttl (✉)
Universität Bremen und Steinbeis Transfer Center InnoVET, Kluesrieser Weg 74, 24939 Flensburg, Germany
e-mail: spoettl@uni-bremen.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_5

1 Introduction

Within the context of "Industry 4.0" it cannot be denied that work organization and work processes will change along with proceeding automation and real-time-oriented control. The same is true for work contents and the interaction and communication between man and technology. So far there are no empirical findings on how the work task and competence profiles are likely to change as soon as production processes will be even more digitalized and decentrally controlled and closely networked with the help of "Cyber-Physical-Systems" (CPS) and intelligent work pieces will be able to organize their way within production on their own (Ahrens and Spöttl 2015; Bauernhansl 2014, p. 7).

In spite of the currently observable bias on technological developments in the context of "Industry 4.0", the employees will be the pivot for a successful implementation. In order to identify the impact of the implementation of "Industry 4.0" on skilled workers, reference will be made to the findings of a study published by bayme vbm in spring 2016 (cf. bayme vbm Studie 2016).

2 Aims and Core Areas of the Survey

The basic aim of the survey was the identification of changed competence demands for employees on the "shop-floor" level and on the middle employment level of the manufacturing metal and electrical industry. Thus the focus was on the qualification requirements for skilled workers without an academic education and those on the middle qualification level. The results of the survey also provided recommendations to tackle the question whether a new shaping of occupations in production and/or occupational profiles should be initiated and what they should focus on. Furthermore there are recommendations for the shaping of production work in the context of "Industry 4.0" for companies and qualification institutions.

The survey started from the future role of skilled personnel. It raised the question whether an increasingly "individualized production" would also make the work tasks more demanding in terms of technology, organization and communication. Which consequences would this development have for employees below the academic employment level? The following central questions were in the center of attention:

1. What are the current and future changes triggered by the introduction of the principles of "Industry 4.0" in the manufacturing industry? What is the scope of these changes? How many employees will be affected in the future? What are the requirements for dealing with the control of intelligent production processes? Will IT- and production competences consolidate?
2. Which impact has the introduction of networked and dynamic production processes on qualification, competence and occupational profiles as well as on skilled workers, master craftsmen, and technicians of the manufacturing metal and electrical industry?

3. Which competences are necessary to work in highly automated, networked production systems? What are the consequences of these changed forms of work organization? Are new forms of "knowledge and skills" necessary?
4. Which requirements are imposed on skilled personnel by companies of the metal and electrical industry in order to be able to implement the potentials of "Industry 4.0".
5. What are the relevant occupational profiles and further training profiles that will be getting in touch with the principles of "Industry 4.0"? Which shaping principles for occupational profiles and curricula can be derived?

The survey in the described field of application had a multi-level design and was based on making use of qualitative instruments from vocational and social sciences (cf. Becker and Spöttl 2015; Windelband et al. 2012). In detail: expert interviews, case studies, and expert workshops. The survey aimed at identifying changes on production level and consequences resulting thereof for the shaping of occupational and competence profiles. The concrete background for the survey were findings of a study on the "Internet of Things" (IoT) in the fields of logistics and industrial production, funded by the Federal Ministry for Education and Research (Bundesministerium für Bildung und Forschung—BMBF) (cf. Abicht and Spöttl 2012). On the other hand the study is based on findings from a current study in the metal and electrical industry on the implementation of "Industry 4.0" and its consequences for skilled work which was conducted under the author's guidance in the years 2015 and 2016 (cf. bayme vbm Studie 2016). The epistemological interest sketched below was pursued with the applied instruments:

(a) 15 expert interviews with experts from different areas of industry, science, and vocational education and training helped to identify the state of the technological development up to "Industry 4.0" and digitalization. The interviews showed the changes of skilled work that are to be expected or are already reality and their consequences for qualification.
(b) With the aid of six case studies conducted in various enterprises and selected sectors of the metal and electrical industry, current developments in technology, changes in work organization as well as work tasks for skilled work were identified. Statements on future requirements for qualifications should be generated. The case studies were also meant to reflect an image of the "sector's" state of development in terms of "Industry 4.0".
(c) In two expert workshops changes on the companies' shop-floor levels and changes in the ordinances were discussed. It was important to open up different interpretations, to clarify real existing and different initial positions and orientations and to validate the findings of the survey. The findings generated insights which allowed for making decisions on the consequences and shaping options in vocational education and training.
(d) A third workshop with further training experts discussed how in-firm and external further training measures could ensure that skilled workers are swiftly prepared for the requirements of "Industry 4.0" in a target oriented way.

The epistemological interest aims at tangible developments on the shop-floor in order to identify the real challenges for the skilled workers. Macro-analyses have not been carried through on purpose. On the one hand they are already available. On the other hand the identification of general interconnections of social, technological and work organizational dimensions will as a rule not reveal the concrete requirements for skilled workers at the plants. The study thus aimed at discovering the requirements on a micro-level, on the shop-floor, in order to gain insight into a new shaping of vocational education and training and its curricula.

3 The State of Implementation of "Industry 4.0"

First and foremost the question shall be answered: How far have companies advanced with the implementation of "Industry 4.0"? Many research projects on "Industry 4.0" have been launched in Germany only in the last few years. Therefore a general statement across all branches is currently not yet possible.

In order to assess the diffusion depth of "Industry 4.0" and thus its presence in companies of the metal and electrical industry, the authors of the bayme vbm Studie (2016) have developed an instrument for the assessment of the diffusion of technology and work organization. The former encompasses seven technology dimensions such as sensor technology/actuating elements (networking CPS), networking (entire value added chain), radio technology (communication), Big Data (data analysis), Cloud Computing (date storage, data speed), work place intelligence CPS (share of man and technology) up to data safety (data sovereignty). Within expert workshops, each of these dimensions was assessed by experts with a view to diffusion depth of "Industry 4.0" in companies. The result (cf. Fig. 1) is the midpoint value of all experts. Thus a reference system was created which allows

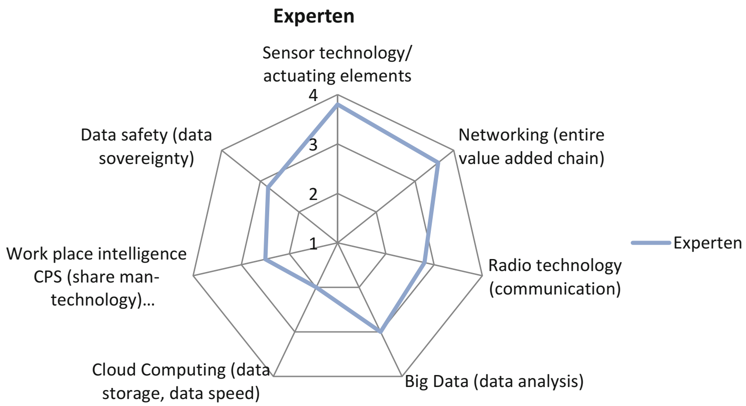


Fig. 1 Diffusion steps of technologies—assessment by experts (Source: bayme vbm Studie 2016, p. 59)

for a clear-cut characterization of the development steps towards "Industry 4.0" related to the dimensions of technology (cf. bayme vbm Studie 2016, p. 56). The result presented in Fig. 1 indicates the implementation depth of "Industry 4.0's" technology as assessed by the experts.

Figure 1 also indicates that sensor technology and actuating elements have reached the highest markedness. This means that the communication of the CPS via digital communication devices has already been networked within the value added chain and that this is highly relevant for the level of skilled work. A minor rating—markedness level 2—was assigned to the dimension of Cloud Computing. Data storage has so far mostly been taken care of by the companies themselves and still has a considerable potential for development. Both radio technology and work place intelligence were also low-rated. The reason for the low rating of work place intelligence could lie in the fact that it is just being implemented at an early stage. Data safety is an issue linked to a lot of trust which can apparently not yet be guaranteed.

The dimension of work organization—Fig. 2—is described by six dimensions, such as: process organization (across the entire value added chain), process safety (availability), process efficiency/process optimization (human-machine interface), process experience (experience knowledge), process quality (failure recording) and process understanding (within the value added chain). Similar to the technology dimensions, these dimensions were also rated by experts during expert workshops with a view to the diffusion depth of "Industry 4.0" in companies. The recorded result (cf. Fig. 2) shows the midpoint value of all experts. This result must also be conceived as a reference system which allows for a clear-cut characterization towards the development of Industry 4.0 with a view to the dimensions of

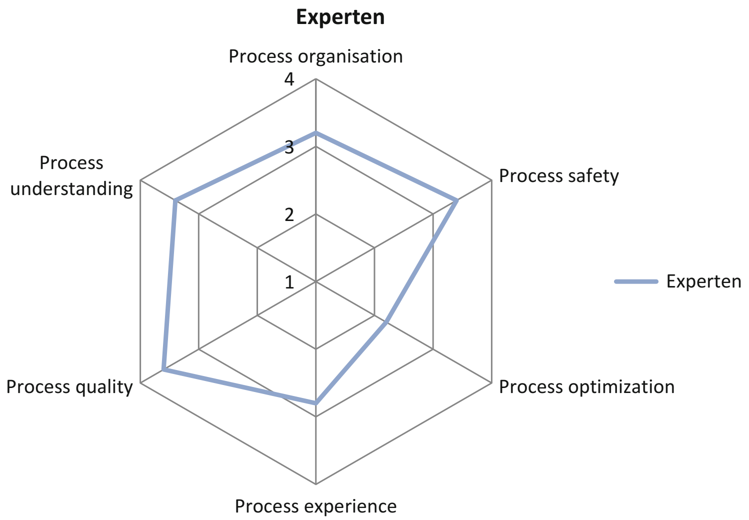


Fig. 2 Diffusion steps of forms of organization—assessment by expert (Source: bayme vbm Studie 2016, p. 70)

organization (cf. bayme vbm Studie 2016, p. 70). The result shown in Fig. 2 thus depicts the implementation depth of dimensions of work organization of "Industry 4.0" as rated by the experts.

The experts rated the dimensions of process understanding and process quality very favorably—with 3.4 and 3.6 respectively within the development steps 1 to 4. This means that the process understanding is rated to be very comprehensive and must be mastered by all skilled workers across all work processes. Process quality reveals a similar picture. Here the companies are increasingly moving towards a zero-fault production by making use of real-time data. Both dimensions are rated in a sense that they are highly relevant for the skilled workers' level and have to be mastered.

The dimension of process optimization is, however, rated lower although it is an important task for the skilled worker. Indeed the machine and the technology respectively as a supporting tool are also taken into consideration. Another conclusion would be that companies still have a long way to go for an autonomous optimization of the production processes by "Industry 4.0" technologies. This is, however, contradicted by the high rating of process organisation. The diverging rating could be explained by the assumption that organisation is seen as a general task for skilled workers whereas optimization is closer related to "Industry 4.0". A high rating is also given to process safety as a dimension of basic importance.

So far it is evident that companies are increasingly dealing with "Industry 4.0" technologies and are further developing the forms of work organisation. The development steps of the individual dimensions are, however, marked in a very different way. Companies are facing a great challenge to create a complete networking across the entire value added chain, starting with the suppliers and ending at the customers. Interface problems caused by different data formats and data security are playing a major role with the implementation of "Industry 4.0" technologies in companies. Due to the high relevance for a well-functioning production, solutions for these problems are only cautiously pursued.

4 Overview of Challenges

Based on surveys in companies (case studies) and expert interviews (cf. bayme vbm Studie 2016), the challenges resulting from the implementation of "Industry 4.0" can be summarized as follows:

1. Skilled workers, master craftsmen, technicians, i.e. persons with an occupational technical education and training and corresponding further training should be qualified for specializations relevant for "Industry 4.0". They must be able to master processes in their complexity and to safeguard a flawless operation of plants.
2. The mastering of networked systems with decentralized intelligence, the ability to deal with data and their analysis as well as the ability to safeguard a flawless operation of the plants count among the most important requirements for work

on production sites. Apart from this, it is of course expected that the still existing traditional tasks for skilled work are also still coped with.

3. The so-far setting of priorities on general questions of "Industry 4.0" have to be extended by technological priorities (CPS), by issues of work organization, by questions of work design, data security, programming techniques, trouble shooting and problem solving with the aid of assistance systems and data analysis (cf. bayme vbm Studie 2016, p. 3f).

The software technical networking and its related CPS elements are continuously becoming prevalent. Along with the increasing diffusion of "Industry 4.0", plants and machines must therefore be conceived and handled by starting with

- networking,
- CP-systems,
- the software and
- their embedding into processes.

Thus the interaction between man and machine is changing considerably. This is underlined by the more intense use of image processing, the transfer of information via visual processing in different user appliances, the use of videos and of audiovisual language etc.

In other words: Context related data are dominating which provide information on plants, production processes and process operations. With the aid of integrated sensors and actuators the behavior of machines is controlled, analysed, and documented. The collected data are processed into information for the plant operator, the workman, the skilled worker. The collected data form again the basis for tools which have to be used by skilled personnel according to the situation (cf. bayme vbm Studie 2016, p. 3f).

This amounts to the question of how "Industry 4.0" will change the organizational processes and thus the hierarchies in companies. There are not many unambiguous research findings available in literature. Therefore it still remains vague how certain decision making processes will change on the skilled-worker-level. Is there still room for co-shaping by human beings in "Industry 4.0" or will "Industry 4.0" promote a "Taylorism 4.0"? The developers and the drivers of the idea "Industry 4.0" keep on underlining that they are aiming at a cooperative interaction between all levels. Humans should be given the chance to exert influence on the shaping of their work within production. The working group "Industry 4.0" comments:

> The Smart Factory contains opportunity structures for a new work culture which is oriented towards the interests of the persons employed. [. . .]Neither technology nor technical constraints should decide on the quality of work but scientists and managers who will shape and implement the Smart Factory. What is required in this context is a socio-technological shaping perspective within work organization, further training activities, as well as technology and software architectures with a close mutual coordination. They should be "seamless" and focus on enabling intelligent, cooperative, self-organized

interaction between the persons employed and/or the technical operation systems along the entire value added chain (Kagermann et al. 2013, p. 57).

5 Changes on Skilled Worker Level

5.1 The Situation of Skilled Workers: Employment Situation

The development of “Industry 4.0” towards digitalization and its impact on the workforce has been surveyed in several studies (cf. Hammermann and Stettes 2015; Schumann et al. 2014; Wolter et al. 2015) and with different objectives. The results of the studies show that the current state of the implementation of “Industry 4.0” does not yet allow a reliable determination of the development of the need for skilled workers. However, all findings of the studies agree upon the fact that employment opportunities for low-qualified workers will decrease along with the implementation of “Industry 4.0”. A higher need of companies for trained skilled workers and academically qualified staff is predicted. The mentioned figures were, however, only rarely collected in an empirically valid way.

Based on our own surveys, the following tendential statements on the future need for skilled personnel can be summarized as follows (cf. bayme vbm Studie 2016, p. 44):

1. Trend A: Companies without “Industry 4.0”.
   No changes in direct production, however, in indirect production (e.g. work places for the future implementation of “Industry 4.0”).
2. Trend B: Companies with a low “Industry 4.0” depth.
   Stagnation or a slight increase in highly qualified skilled personnel such as skilled workers, master craftsmen, and technicians as well as a noticeable increase in productivity.
3. Trend C: Companies with high “Industry 4.0” depth.
   Increase by 20–30 % in the upper qualification level (well-qualified skilled workers, master craftsmen, technicians . . .).
   (Substantial) reduction of low-qualified workforce (semi-skilled and unskilled).

These trends show that skilled workers with a high-quality initial and further training oriented to “Industry 4.0” will probably have very good employment and career opportunities. This is contradicted by some experts who state that the skilled worker will be chanceless or even at risk due to the developments towards “Industry 4.0”. Nevertheless this group competes against academically trained persons when it comes to planning and conceptual tasks. As for other tasks such as trouble shooting and repair, putting into operation and setting-up as well as maintenance and repair, skilled workers with an occupational training in metal or electro-technology with focus on software based networking on the shop-floor cannot be replaced and are highly valued, ever more if they have 3–5 years' professional experience. The case studies underpinned that this is above all rated favorably by companies with highly complex plants. Latest with the putting into operation of plants, skilled workers with their occupational training can safeguard a continuous operation of the plants, provided they have already learned "to start thinking from the software" (bayme vbm Studie 2016, p. 44).

The statements of worsening employment opportunities of semi-skilled and unskilled workers in our survey match the statements given in other scientific studies. This group does not have the competences required for work at complex high-tech plants. This target group makes up for about 15–20 % of the production workforce in the metal and electrical industry according to region and branch and seems to be the loser of the development of "Industry 4.0".

5.2 Challenges for Occupational Profiles

Apart from results on the diffusion of "Industry 4.0" technologies and forms of work organization, further results are available regarding the changes of skilled work on the shop-floor level. These changes have a considerable impact on occupational profiles and will be sketched below.

5.2.1 Empirically Identified Changes

These changes concentrate on:

- Plant maintenance, reading and interpreting of operational data of plants, error recovery in sensor technology and actuators (mastering signal processing).

> The handicraft tasks such as a change of chucks, machine refitting, [. . .] coping with service and repair tasks will prevail (Case A).[1]
>
> There will be additional new tasks—they are strongly influenced by the electronic networking (Case E).

- Trouble shooting in highly complex plants which are networked with other machines/plants.
- Mastering of precise analysis of causes of malfunction (reading out live machine images, their assessment and deciding what has to be done).
- Acquisition of knowledge via the Internet, handbooks, data sheets, knowledge forums.

> [. . .] we are responsible for material feeding as well as for the repair of mechanical malfunctions. If the problems turn out to lie deeper, e.g. in the electrical system, we have an electrical service team on site. [. . .] This means that we as mechanics have no access to electrical control cabinets. We need the on-site service team for that. There is also service

[1] References are case studies in different companies.

personnel available for programs with camera systems. [. . .] We keep the entire machine running. Malfunctions caused by SPS are malfunctions we can repair. (skilled worker)

- Autonomous support of process optimization, safeguarding the search for information and contributing ideas for improvement based on every-day work.
- Hybrid managing of tasks by skilled workers at the plants (mechanical, electrical, and software-based tasks).

You have to know your job, which processes are dealt with at the individual stations, how work is dealt with, mechanically, electrically [. . .]. Which kind of spare parts are used. And then you need training for each malfunction you are supposed to repair—training in pneumatics, hydraulics, mechanics, electrics [. . .] (skilled worker).

- "IT competences" and especially network technology, router configurations, firewall technology, separate networks, network translation, identification of malfunctions, understanding processes, integration and real-time synchronization of processes along the product life cycle.

Increasing networking can also be witnessed in island production. The employee's role is changing towards monitoring tasks. The utilization of machines and quality are becoming more and more important. Employees are being trained for IT-control—for the new installation of production islands. (Case E)

- Data analysis and interpretation are playing a most dominant role—Hazards:
  Variety of data (Big Data) and data processing for the work process are increasing.

Sentience is suddenly shifted to the software (production manager).
Machine operators must not only know how these plants are switched on. They must also be able to read the data displayed on the monitor and to interpret them in order to draw conclusions. (Case H)

- Conducting parametrizing tasks is part of the range of tasks.
- Experience with mechatronic plants is an important prerequisite for the implementation of "Industry 4.0".

5.2.2 Generic Occupational Fields of Action in "Industry 4.0"

The sketched empirical results underpin changes in coping with tasks by skilled workers below the academic level. Based on the identified changed work requirements, nine generic occupational fields of action relevant for "Industry 4.0" were generated from the empirical fields of action and will described below.

These "generic" occupational fields of action underline "the new" relevant for metal and electrical occupations due to "Industry 4.0" and the related CPS. Below is a description of the changes in the form of occupational fields of action.

Definition: Generic Occupational Fields of Action: "Industry 4.0"
A generic field of action in Industry 4.0 is a field of tasks for skilled workers extended by "Industry 4.0" which has been identified with the help of empirical studies in companies. After the implementation of Industry 4.0, companies have partly already reached an advanced stage. A generic field of action describes new requirements and tasks within a field of tasks (e.g. plant monitoring) typical for an Industry 4.0 environment.

The overall nine generic fields of action which could be generated represent a reference system which helps to describe and characterize the new conditions based on "Industry 4.0", i.e. what has actually changed. The reference system can be used to check whether existing occupational profiles should be changed and established in order to include them into vocational education and training.

Description of the Generic Fields of Action: Plant Engineering

The planning of new automated plants is first and foremost a task for engineers. However, companies are increasingly assigning these tasks to mixed teams including specially qualified skilled workers. Together with the engineers they must select plants suitable for production. An important issue is the fact that skilled workers are confronted with the networking of plants at a very early stage. Within the framework of the planning process they can already acquire knowledge of the function of the plants and options for networking which they can later use to ensure the operation of the plant as well as for troubleshooting. Planning processes are increasingly simulated with the help of Virtual Reality. Virtual Reality allows the imaging of a realistic production system in order to simulate the behavior of the cyber-physical production system and to explore it in an interactive way.

Target perspective: Simulation of plants.

Plant Installation

The installation and refitting of automated networked plants mostly assigns coordinating tasks to engineers while skilled workers take care of the actual technical installation. With the help of the engineers, they learn the exact interrelationships and functions of the individual work steps at the plant and their integration into the overall production process.

The network structures within the value added chain and the use of sensors and actuators are exactly planned and realized. It is also important to consider all preceding and subsequent processes and to make sure that all data important for the overall process (programs, interface configurations, statistical data, data for quality check and networking) can also be processed by skilled workers.

Target perspective: Plant networking.

Set-up of a Plant and Putting it into Operation

Plants must be set up prior to their putting into operation. Tasks such as e.g. the setting of a zero point, calibration, the balancing of the start position, a disruption-free data transfer, the correct display of data on the control monitors are tasks that are taken care of by skilled workers, supported by engineers.

The putting into operation of automated plants and their networking is done in close cooperation between engineers and skilled workers. The engineers have to ensure that the necessary programming functions and interfaces within the production system are working flawlessly. The skilled workers, on the other hand, must make sure that all mechanical, hydraulic, pneumatic, electrical and electronic functions are safeguarded so that the plant can operate flawlessly and that it can be integrated into the production network.

Target perspective: Ensuring the data availability of sensor, actuator and process data in production systems (PPS, MES, SCADA, ERP, SAP).

Plant Monitoring

The monitoring of plants—also of several plants simultaneously—counts among the important tasks for skilled workers. Above all the flawless operation of plants must be safeguarded. This means that the available real-time data must be continuously monitored, analysed and evaluated and the overall operation of the plants must be monitored and corrected if necessary. The monitoring of noise generation by machines is as important as data providing information on production monitoring and production quality within the production processes and the products to be manufactured.

Target perspective: Monitoring, analyzing, and evaluation of real-time data.

Process Management (Visualization/Monitoring/Coordination/Organization)

Another important task for skilled workers in terms of plant surveillance and plant operation is to safeguard continuous process flows. This means that skilled workers must continuously monitor processes. In case of identified malfunctions they have to repair them immediately or to suppress them altogether by preventive maintenance.

One of the most important prerequisites for these tasks is an overview of the entire plant control, a sound knowledge of the functions, the flow and the operation of the plant as well as closely watching the monitoring systems. The skilled workers must be able to read, analyse, and interpret data transmitted via these systems.

Target perspective: Guarantee process safety by process monitoring and repair of malfunctions.

Data Management (Dealing with Operational Data/Access to Software/Parametrization/Programming)

Reading, analyzing, and processing of machine data and their preliminary data is another important task for skilled workers. Data are most important for the setting-up of a machine and for a quality-based operation of plants.

Skilled workers must be able to read, analyse, and interpret all relevant operational data (loads, machine and consumption modes). Deviations from standard values have to be identified and dealt with immediately in case of malfunctions. The statistical process control (SPC) in networked production systems is no longer just focused on an individual machine but more and more on entire production sections under the responsibility of the skilled workers.

Skilled workers have to focus their thinking

- by starting from the processes and
- from the software.

and thus optimize plants and their functions. This requires an entirely different understanding compared to mechanically-electrically operated plants. This other understanding is highly relevant for both the operation and the optimization of the plants and their integration into the overall production process.

An efficient plant operation calls for skilled workers to carry out parametrization tasks on their own. The correction of programming sentences and/or data analyses are also part of their field of tasks.

Target perspective: Safeguarding machine data for plant quality operation, analyzing operational data and optimizing processes.

Maintenance

The tasks for skilled workers encompass simple maintenance tasks in order to safeguard a flawless plant operation. This means that they must have access to the function of plants and must be able to cope with (metal-technological, electrical, IT-based) maintenance tasks in plants composed of different technological systems. Difficult repair tasks are exempted. The maintenance tasks also encompass preventive maintenance prepared by recording, processing, and visualization of operational and production data and should be ready to be applied at the production work places at all times. In addition skilled workers must also master virtually organized maintenance tasks and must apply assistance systems for trouble shooting, documentation and knowledge transfer.

Target perspective: Preventive, foresighted maintenance, multi-functional machines, assessment and use of different data and data formats.

Repair

Difficult maintenance and repair tasks in networked plants and individual machines are as a rule carried out by maintenance teams. These teams are specialized in such tasks including IT-tasks such as network analyses or IT-guided trouble shooting. They also master procedures for the identification of malfunctions (data analysis), causes for malfunction and their repair in complex, networked plants.

Target perspective: Considering repair interdependencies due to networking and IT-integration of machines and plants; software updates.

Trouble Shooting and Repair

As soon as malfunctions occur—e.g. caused by flawed products or process flows—skilled workers must be able to identify and remove eligible causes. This implies that they are able to master diagnostic procedures which not only encompass mechanical and electrical/ electronic basic functions but also the digital control of the production process. This is why malfunctions can no longer be identified and repaired directly at the sensors, actuating elements or cabling but rather via IT-systems or within the networking of the production.

Target perspective: Diagnosis, trouble shooting in networked plants.

6 Discussion of the Results

The discussed results of the survey on changes due to the implementation of "Industry 4.0" make it clear that there will be a change of paradigms for skilled workers in production. The so-far focus on tasks such as machine operating, setting-up, looking for minor malfunctions and their repair will still play a role. However, the access to, the analysis, the interpretation, and the assessment of data of different formats will become increasingly significant. The view on and the sense of dealing with plants will be shifted towards the software due to data-based control. This statement underpins the change of paradigm from a lengthy process of the dominance of mechanical and later electronic control towards a software-dominated operation of highly complex and efficient plants. This means that skilled workers are assigned an extended and more demanding field of tasks which often leads to more abstract challenges than before due to networking and data-based controls. Based on these developments the production managers in the surveyed companies only conceded employability for unskilled and semi-skilled workers if this target groups are ready for intensive further training. This group encompasses up to 20 % of the workforce in companies of machine building and electro technology.

Based on the diffusion depth of "Industry 4.0" shown in Chap. 3 it is obvious that "Industry 4.0" calls for not only changing occupational profiles but to enrich them by "Industry 4.0" contents. This is, however, still not enough to meet the demands of "Industry 4.0". What has to be done is a new approach with a perspective of processes and digitalization. Networking, the digitalization of processes, and the shaping of intelligent work places must be the center of interest. The current future-oriented occupational profiles allow amendments, however, no change of perspective.

This is due to the fact that structurization principles must change more intensely than in the past.

References

Abicht L, Spöttl G (eds) (2012) Qualifikationsentwicklungen durch das Internet der Dinge. W. Bertelsmann, Bielefeld

Ahrens D, Spöttl G (2015) Industrie 4.0 und Herausforderungen für die Qualifizierung von Fachkräften. In: Hirsch-Kreinsen H, Ittermann P, Niehaus J (eds) Digitalisierung industrieller Arbeit. Edition Sigma, Baden-Baden, pp 185–204

Bauernhansl T (2014) Die vierte industrielle Revolution—Der Weg in ein wertschaffendes Produktionsparadigma. In: Bauernhansl T, ten Hompel M, Vogel-Heuser B (eds) Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendungen—Technologien—Migration. Springer Vieweg, Wiesbaden, pp 6–35 bayme vbm Studie; Spöttl G, Gorldt C, Windelband L, Grantz T, Richter T (2016) Industrie 4.0—Auswirkungen auf Aus—und Weiterbildung in der M+E Industrie. Studie herausgegeben von bayme vbm. Die bayerischen Metall—und Elektro-Arbeitgeber, München. www.baymevbm.de/industrie4.0. Accessed 27 May 2016

Becker M, Spöttl G (2015) Berufswissenschaftliche Forschung. Lang Verlag, Frankfurt am Main

Hammermann A, Stettes O (2015) Beschäftigungseffekte der Digitalisierung—Erste Eindrücke aus dem IW-Personalpanel. In: Institut der deutschen Wirtschaft (eds) Vierteljahresschrift zur empirischen Wirtschaftsforschung, Jg 42, Heft 3. IW, Köln, pp 77–94

Kagermann H, Wahster W, Helbig J (eds) (2013) Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0. https://www.bmbf.de/pubRD/Umsetzungsempfehlungen_Industrie4_0.pdf. Accessed 29 July 2013

Schumann A, Assenmacher M, Liecke M, Reinecke J, Sobania K (2014) Wirtschaft 4.0—Große Chancen, viel zu tun—Das IHK-Unternehmensbarometer zur Digitalisierung. Deutscher Industrie—und Handelskammertag, Berlin

Windelband L et al (2012) Qualifikationsentwicklungen durch das Internet der Dinge in der Logistik. In: Abicht L, Spöttl G (eds) Qualifikationsentwicklungen durch das Internet der Dinge. W. Bertelmann, Bielefeld

Wolter MI, Mönnig A, Hummel M, Schneemann C, Weber E, Zika G, Helmrich R, Maier T, Neuber-Pohl C (2015) Industrie 4.0 und die Folgen für Arbeitsmarkt und Wirtschaft—Szenario-Rechnungen im Rahmen der BIBB-IAB-Qualifikations—und Berufsfeldprojektionen. In: Institut für Arbeitsmarkt—und Berufsforschung der Bundesagentur für Arbeit (eds) IAB-Forschungsbericht, Heft 8. IAB-BIBB, Nürnberg

Work Requirements and Qualifications in Maintenance 4.0

Lars Windelband

Abstract

The independence of IT systems is further gathering speed along with the development towards "Industry 4.0" and will result in dramatic changes of the future role of skilled workers. The increasing automation and digitalization of the world of work is leading to changes with regard to work organization, work processes and thus work requirements and the interaction between man and machine. By the example of the field of application "Maintenance 4.0", this article describes the potentials of the development step "Industry 4.0" as well as the challenges, the implications and consequences for all employees and their occupational initial and further training.

Keywords

Maintenance 4.0 • Industry 4.0 • Occupational initial and further training • Skilled worker

1 Introduction

The vision of "Industry 4.0" has it that all products, production facilities, tools, transportation technologies and storage systems within a value-added chain will be turned into cyber-physical systems (CPS) by the integration of computer performance and access to the Internet. This will result in the emergence of an artificial intelligence facilitating self-organisation and the interaction of cyber-physical systems with their environment. In the end, all intelligent machines will

The article is an extended version of the first publication by Windelband (2016).

L. Windelband (✉)
Institut für Bildung, Beruf und Technik, Pädagogische Hochschule Schwäbisch Gmünd, Oberbettringer Straße 200, 73525 Schwäbisch Gmünd, Germany
e-mail: lars.windelband@ph-gmuend.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_6

independently (autonomously) organize the production processes, service robots and human beings will jointly assemble new products and intelligent automated guided transportation vehicles will take care of logistic assignments on their own. Beyond the intelligent factory, production and logistic processes will be globally interlocked via the Internet in order to optimize the material flow, to avoid possible flaws in logistics and production and to be able to react to customers' wishes in a highly flexible manner.

The ever increasing transfer of more production elements into cyber-physical systems, the number of systems to be maintained is also going up. At the same time the degree of networking with an information technological infrastructure is growing. The new influence and interference factors are leading to an ever higher complexity of the production systems which have to be understood and mastered prior to maintenance (cf. acatech 2015, p. 7). The increasing quantity and quality of available data must be analysed and interpreted in order to further minimize down times and times of nonuse of production facilities.

It is to be expected that the due date for maintenance of intelligent products and/or plants will be recognized and that the maintenance measure will be initiated as soon as real-time data on loads, state of the machine, and consumption status are available thanks to "Industry 4.0" (cf. Biedermann 2014, p. 26). Thus maintenance work can be flexibly determined depending on the load and use intensity of the machine or the production facility. It is even planned that machines automatically order their own wear and tear parts/spare parts based on an analysis of previously collected machine data.

2 Research Design and Research Questions

In various case studies and expert interviews the changes in maintenance work have been identified and conclusions for new qualification requirements in the field of maintenance could be drawn. On the one hand recourse was made on results of research projects on the present and future qualification requirements for the "Internet of Things" in the fields of logistics and industrial production, funded by Bundesministerium für Bildung und Forschung (BMBF) (cf. Abicht and Spöttl 2012). On the other hand the author has evaluated current case studies and expert interviews conducted within the framework of a bayme vbm Studie (2016) and their consequences for the occupational initial and further training.

The surveys comprised several levels and are based on the use of qualitative vocational scientific instruments (cf. Windelband and Spöttl 2012):

a. Expert interviews and expert workshops with different representatives of the industry, of science and vocational education were conducted in order to obtain information on the state of the technological and work-organisational development ("Internet of Things", "Industry 4.0", digitalisation), changes in skilled work and their consequences for vocational education and training.

b. The case studies were carried out in various companies of the metal and electrical industry. With the aid of case studies, current developments in companies with regard to changes in skilled work could be identified in order to make statements on the future qualification requirements.

The following research questions are in the centre of attention:

– What are the current and future changes entailed by the introduction and implementation of new maintenance strategies within the "Internet of Things"/ "Industry 4.0"? How wide is their range in industry and in adjacent fields?
– Which competences for maintenance work within "Industry 4.0" can be derived for skilled workers on a middle employment level?
– How must production technologically relevant occupations or occupational profiles for the field of maintenance possibly further develop or change?

3 Fields of Application of "Industry 4.0"

As for the questions of the concrete changes in work processes and demands for the employees while dealing with "Industry 4.0" applications, it must be clarified which functions can actually be taken over by cyber-physical systems in a concrete case of application, which functions are likely to remain with the skilled worker and how the interaction between man and machine is changing. First consequences in terms of required competences for future skilled workers can be derived from the prospective importance of man and the degree/scope of tasks taken over by CPS within the framework of the respective "Industry 4.0" applications.

At the moment different application scenarios of the Human-Machine-Interface (HMI) for the use of CPS in production are being discussed (cf. Gorecky et al. 2014; Windelband and Dworschak 2015):

– Maintenance (i.e. servicing, inspection and repair) of production facilities by providing interactive virtual instructions,
– Monitoring of production processes as well as quality control by context-sensitively retrieving and providing information, e.g. with regard to the status of a CPS,
– Planning and simulation of production processes, e.g. by pre-determining CPS behavior,
– Use of lightweight robots (sensitive robotics) by automobile producers and suppliers in close cooperation with the employees.

Especially on the level of planning (engineering), "Industry 4.0" should lead to a further development and thus also to an optimization of the processes. Certain planning processes can, for example, be simulated by Virtual Reality or information flows could be shaped more effectively. Virtual Reality offers an image of a realistic production system in order to simulate the behavior of the cyber-physical production system and to explore it interactively (cf. Gorecky et al. 2014, p. 528).

The monitoring of production systems is often done by employees of the middle qualification level and ranges from monitoring individual production systems or production lines up to entire production halls. How far this development is likely to proceed is strongly depending on the respective branch and the kind of production applied (cf. Windelband and Dworschak 2015, p. 76).

The use of lightweight robots is currently being planned by different automobile manufacturers and their suppliers. Lightweight robots excel above all by the fact that a simple "Teach-In" for the positioning tasks the robot is about to take over can also be done manually. A simple programming by multi-functional interfaces and via simple touch screens up to control via smart phones or tablets can also be carried out by skilled workers. The lightweight robots can be operated freely (without protective fences) as the safety standards are high due to adaptive sensors and real-time calculation of collision control (ibid. p. 82) (Windelband and Dworschak 2015). Most of the companies aim at further reducing the ergonomically highly strained work places (so-called "red work places").

Based on different market and production requirements, companies choose e.g. different combinations of variants for the use of technology and different options of organization. "Maintenance 4.0" of machines allows for an improved production planning, longer operating hours and a higher availability of the machines. The field "Maintenance 4.0" will be closer looked at in four application scenarios and their consequences for skilled work and qualification.

4 State of Research on "Maintenance 4.0"

The most important task of maintenance is to safeguard the functioning of a production facility during its operating life or to restore its functional status (cf. Biedermann 2014, p. 24). According to DIN 31051:2003-06 this is defined as follows (DIN 31051 2003):

> Combination of all technical and administrative measures as well as management measures during the life cycle of a respective unit in order to preserve or to restore its functional status, thus safeguarding that the unit is able to perform the required function.

The companies have established a variety of maintenance strategies which are likely to change within "Maintenance 4.0":

a. Unscheduled maintenance: damage-based repair (repair oriented)
b. Scheduled maintenance: preventive maintenance to minimize failure risks
   - Maintenance based on actual conditions: Sensors indicate wear and tear conditions (Condition Based Maintenance). The respective wear conditions are recorded and compared with the minimum values for the safe operation of a plant.
   - Prospective maintenance: Prospective maintenance already starts during the planning phase and determines the optimal instant of time for prospective maintenance measures.

– Preventive maintenance: Preventive measures (such as inspection and maintenance) are conducted in order to initiate measures prior to the occurrence of a failure (regardless of the run time of a plant).

Down times and times of non-use of production facilities will be further minimized, mainly by monitoring systems installed during the implementation of "Maintenance 4.0". It is to be expected that the optimal time for maintenance will be recognized by intelligent products or plants and that maintenance measures will be initiated as soon as the implementation of "Industry 4.0" will provide data on loads, machine and use conditions in real time (cf. Biedermann 2014, p. 26). Thus the maintenance measures will be flexibly determined depending on the load and use intensity of the machine or the plant.

In the field of scheduled maintenance, above all the further development of sensors will have a considerably impact on the identification of the condition of plants or spare parts and the degree of their wear. This will also influence the prognostication of the remaining period of use and lifetime of a machine or plant and improve the activation of effective, preventive maintenance strategies. Operation and maintenance intervals could thus be optimized. At the same time the changes in inspection possibilities will also exert influence on the field of tasks of maintenance. The share of unscheduled maintenance (repair-oriented) is most likely to decrease compared to preventive maintenance that could become more important. On the other hand we expect that damage repair in the sense of "trouble shooting" will become more complex (cf. Hirsch-Kreinsen 2014, p. 39; Biedermann 2014, p. 26). Thus the maintenance strategies will move from a more repair-oriented maintenance to a condition-based and prospective maintenance. This is also underpinned by a survey on maintenance conducted by experts (cf. Günther et al. 2014, p. 15f.). For the next 5 years they predict a trend towards a condition-based and prospective maintenance (cf. Fig. 1). Above all Condition Based Maintenance (CBM) is considered an important development with a more target-oriented planning of maintenance tasks.

A systematic exchange of information between plant manufacturers, operators and maintenance personnel must, however, be initiated as this is still not the case today. What is currently missing are assistance systems for maintenance which

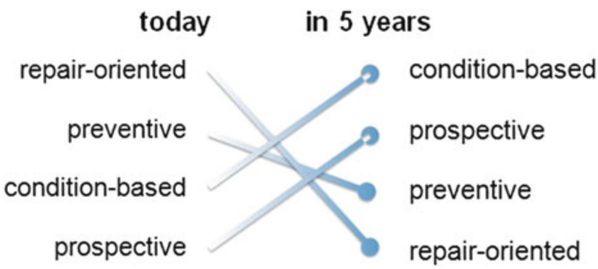


Fig. 1 Development of maintenance strategies in 5 years from today's rather repair-oriented to a condition-based maintenance (Günther et al. 2014, p. 15)

would facilitate a communication with manufacturers and other experts in the field of maintenance. The new technologies (Service Apps, portals etc.) encourage communication and the exchange of all stakeholders. They are, however, mostly not adapted to the requirements of work processes and consequently the needs of skilled workers. The knowledge and assistance systems must, however, be firmly embedded into the human work process and organizational tasks have to be coordinated respectively. Only then the skilled worker can deal with problematic cases during the work process and can ensure an exchange of knowledge.

Günther et al. (2014) think that the implementation of measures assigned to the field of maintenance within "Industry 4.0" is not only unequally distributed between the companies but also between the individual departments of a company. Only a few companies have implemented in-firm virtualization measures in pilot projects. According to Günther's et al. (2014, p. 4) surveys, an estimated 5–20 % of Austrian companies have already initiated a development towards "Industry 4.0". Above all the small and medium-sized enterprises often still lack concepts.

5 Shaping of Assistance Systems for Maintenance 4.0

So far there are just a few assistance systems which could be called cyber-physical assistance systems meant to support skilled workers. Nevertheless a clear further development of these systems is visible which help to process the huge amount of different data originating from the product development process (CAD models, process descriptions), from data sheets/operation manuals or the production process itself (consumption data, load conditions and machine status, order progress, process parameters). Networking and the digitalization of production processes will generate an enormous amount of data (Big Data). Skilled workers need situation-oriented filter mechanisms to get exactly the information on the plant condition necessary to cope with their tasks—on time and at the right place (Schlick et al. 2014, p. 62). This information will be very comprehensive and stem from different systems (MES—Manufacturing Execution System, SAP and other IT-systems). The human-machine interface will be of central importance. For mobile operation, different platforms such as tablets, smart glasses or smart phones will be available and have their own networking and computer performance. Thus work orders on maintenance can be transferred directly to the mobile device. All relevant information such as contact data, order details, operation manuals, check lists and repair lists, information on spare parts and measuring results are available online and offline and can be accessed by the employees at any time. This mobile support facilitates on-site spare parts orders, customers' signatures, clearance or follow-up orders (cf. Mathur and Weiß 2014, p. 151).

The shaping of these assistant systems will be crucial. This can only be successful if skilled workers are actively involved in the development and shaping process right from the start. The results of a survey on the requirements for an assistance system for maintenance, conducted by Höhnel et al. (2015) show that the highest priority is given to a simple and intuitive operation, a visual user guidance, the option of entering data via a keyboard and a robust construction for the application in an industrial environment. The results also reveal that the search for and the processing of information (detailed errors display and kinds of malfunctions; possibility to communicate; mobile access to relevant machine data; access to relevant Web applications via the Internet; display of machine plans and instructions) must be an important element in any assistance system.

6 Changes in the World of Work of Maintenance 4.0

The example of a fictitious skilled worker will show how the work tasks in maintenance could look like in the "Maintenance 4.0" era:

> The skilled worker in the maintenance department has undergone training in metal-technology and/or electrical technology. He/She is responsible for a flawless machine-to-machine communication as well as for the usage of Service-Apps and service portals (e.g. with manufacturers). Plant and machine conditions are checked with a mobile device and are presented in a virtual way (also with the help of Augmented Reality) including an enhanced integration of Online-Monitoring Systems within automation technology. The challenge for the skilled worker lies in the processing of great amounts of data for maintenance processes, partly in real-time. Interfaces must therefore be reduced or harmonized. However, the increasing complexity of the systems also entails a complexity of occurring errors. The tasks for skilled workers therefore include the correct interpretation and evaluation of the respective data. Furthermore the complex systems call for an inter-departmental—even an inter-company—cooperation. This is why the skilled worker is required to make data-based decisions within a short period of time on his or her own. Figures, sensor and operational data as well as facts are the core elements of maintenance strategies. The skilled worker thus becomes a data analyst along the value-added chain. The tasks of skilled workers are thus transforming from a condition-based service towards an integrated, prospective maintenance.

For a successful implementation of the requirements for maintenance within the framework of "Industry 4.0" companies must carry out an adaptation of the different roles their employees are playing. The management must discuss the question whether to approve of an exchange of data across the boundaries of the companies (horizontal integration). Furthermore new mobile solutions and software solutions should be developed and implemented with the help of the IT-departments (or IT-enterprises). New organizational concepts must be developed to enable skilled workers to make data-based decisions in the future and to clear obstacles. This transformation will above all change the tasks of skilled workers: Planning, data analyses and visualization will play a more important role.

These results are also underpinned by the bayme vbm Studie (2016, p. 83): Apart from the well-known tasks for maintenance and trouble shooting, those tasks are highlighted which have gained importance due to networking and digital data processing. An example: Reading out live images of the machine and their evaluation, trouble shooting with new assistance systems, trouble shooting in highly complex plants or fault repair in the field of sensor technology/actuating elements (mastering signal processing) (cf. Fig. 2).

The results underpin the intensified diffusion of networking with the aid of CPS and show that cognitive dimensions like thinking in the work process, thinking in networks, the optimization of processes and thinking starting from the software are playing an important role.

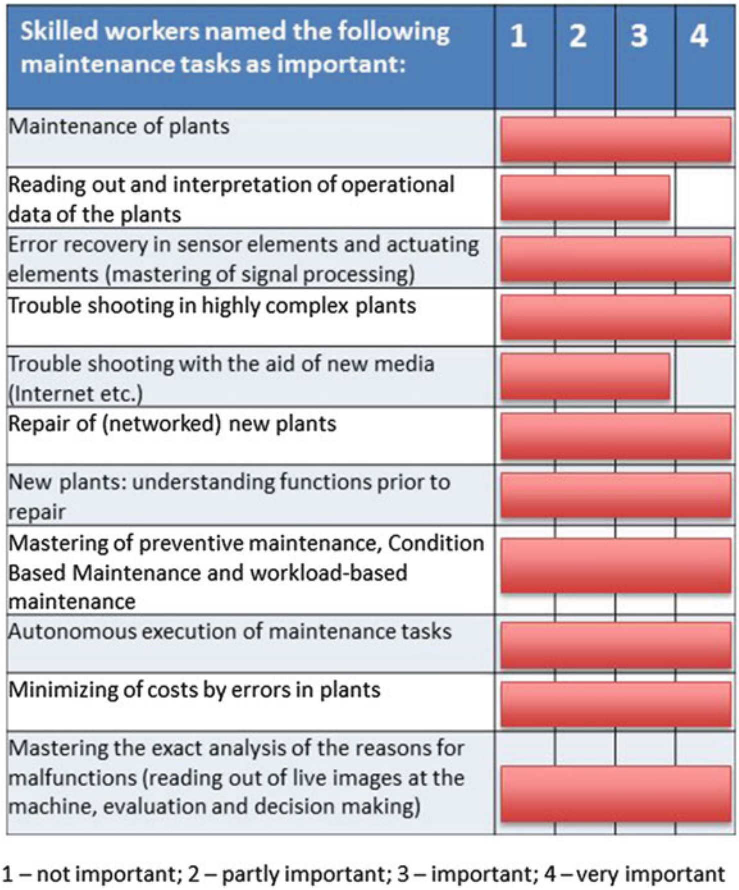


Fig. 2 Maintenance tasks in skilled work (bayme vbm Studie 2016, p. 82)

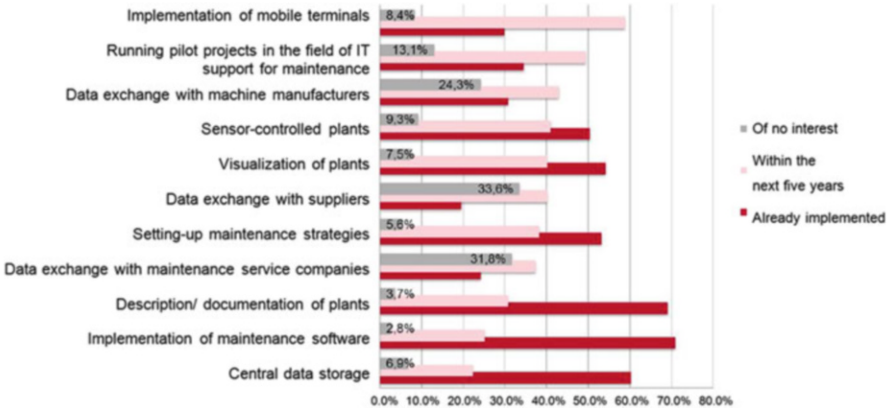


Fig. 3 Implemented and planned maintenance projects (Günther et al. 2014, p. 18)

This development is also confirmed by the results of expert surveys by Günther et al. (2014). The results are based on implemented and planned maintenance projects of the 250 interviewed company experts (Fig. 3). The “implementation of mobile terminals” is dominating the majority of the planned maintenance projects. Further projects will be carried out within the next 5 years with a focus on the implementation of pilot projects in the field of “IT-supported maintenance” and “data exchange” with machine manufacturers, suppliers and maintenance service companies. There is a slight contradiction when it comes to the issue of data exchange, as around 24.3 % of the experts indicate that a data exchange with machine manufacturers would be of no interest for them. The author, however, thinks that a hidden fear of unintentionally granting insight into the core know-how of one’s company could be one reason for this statement.

7 Competences and Qualifications for Maintenance 4.0

Mastering the complexity of the maintenance systems will be a great challenge in the future. A great amount of information and data have to be correctly interpreted and evaluated (data analysis). The shaping of the human-machine interface will play a dominant role. What will be the basis of future decision making? Today skilled workers in maintenance departments are trying to make their decisions mostly based on their experience. They make use of their intuition, their feeling and their senses. Decision making of the future will be supported by computers and a new data quality. It remains to be seen how the support for the employees will look like and which kind of human interference will prevail. The question is whether human beings dealing with these ever more automated systems will still be able to acquire the necessary knowledge in order to identify malfunctions in decisive situations and to work out solutions.

When applying the “Maintenance 4.0” technologies as an assistance system—as described in Sect. 5, skilled workers for future maintenance must have a deepened and combined knowledge of IT as well as of electronic and mechanical systems in order to react and act quickly (cf. Windelband 2014). In this context a bigger need will arise for competences in the field of control technology, data analysis as well as programming and the usage of special software. What exactly does this mean for the work process? Will we rather need generalists or specialists? Must the maintenance worker carry out programming on his or her own, will he or she “only” optimize the programs, or must he or she master a programming knowledge in order to be able to communicate with the computer scientists in terms of process optimization? How can experience knowledge be established and how can implicit experience knowledge be preserved? Already today, skilled workers have to contribute basic competences with regard to network technology, radio technology and transmission technology to identify the causes of malfunctions and to communicate with IT-experts. The skilled workers also need deepened process knowledge and a new kind of knowledge of thinking in networks. This enables them to conduct a fault analysis.

The following domain-specific fields of actions for maintenance tasks could be identified in the bayme vbm Studie (2016, p. 103ff.).

Mastering IT-Systems in Production Depending on the Situation

IT-systems include router configurations, firewall technology, error identification with the help of software, address assignment of CPS components and their integration into the MES system (Manufacturing Execution System). Employees must be able to assess the mode of operation, the error-proneness and the weak points of individual components of IT-systems in production lines and with regard to the other processes of the value-added chain. Thus they will be able to draw the correct conclusions in case of malfunctions and to repair them in a target oriented way. The operation of the plant must not be disturbed by these interventions.

Trouble Shooting and Error Recovery

Skilled workers must be able to apply standardised diagnostic procedures and to develop individual trouble shooting strategies in order to diagnose malfunctions in technically complex systems. The diagnosis of malfunctions concerns mechanical, hydraulic, pneumatic, electrical/electronical and software-technical components of networked plants. This means that the prerequisites for diagnostic tasks not only call for knowledge of physical compounds of the system elements but as well the data flows of the system. Moreover the employees should be able to access these structures. These prerequisites also enable the acting skilled worker to interpret directly perceivable information (e.g. sound qualities of plants), to detect causal interrelationships between events and to identify the causes of the failures.

Analysis of Abstract Interrelationships in Plants and Removal of Malfunctions

The point is the identification of problems. If the skilled worker is able to realize that the material feeding system is jammed, the next batch of material can be taken instead in order to continue production. What is important in this case is to first and foremost understand that the machine is not showing a malfunction. Moreover the error has to be spotted within the data link—or in the lack of it. Employees need to think in a new way. Currently they are used to look for the malfunction in the machine rather than in the logic of how things belong together.

Above all it is important to master the manifold combinations of software, to be able to read displays and to check where a sensor is not functioning, and eventually the repair of such an error.

Use of Assistance Systems in Maintenance

Skilled workers in production must increasingly work with assistance systems. These devices support the systematic exchange of information between the plant manufacturers, the operators and the maintenance staff. The assistance systems (service apps, portals etc.) have to be embedded into the work process and the organization has to be coordinated with them. Only then will it be possible that skilled workers can deal with problematic cases during the work process and thus safeguard an exchange of knowledge. Plant and machine conditions can be accessed from mobile terminals and are virtually displayed (among others with the aid of Augmented Reality). This includes an intensified integration of Online-Monitoring Systems within automation technology.

These described fields of action reveal the changes in the world of work by the ongoing digitalization in networked systems in maintenance and start-ups of production plants as well as the control of the plants by adhering to data protection and data security.

The bayme vbm Studie (2016, p. 85 f.) identified four lines of argument for the assessment of the qualification and competence levels required for skilled workers in maintenance departments for work with networked plants. The study also gives hints to contents required for qualification and competence development:

– One of the lines of argument was that it seems to be impossible to qualify all persons working as generalists to repair at least 80 % of all malfunctions. Therefore a wider and quite fundamental initial training is favored. After several years of work in the company, they can then be trained as specialists in further training courses. In this connection the thesis was formulated that generalists tend to losing importance.
– Another line of argument was that the qualification of specialists with particular IT-knowledge often has an impact on the internal hierarchy of companies. It is e.g. not unusual that a person who underwent continuous further training after initial training and who even specialized in IT-technology could be eligible for the position of a team leader rather than an academically qualified person. Master craftsmen are already well qualified for leadership tasks and thus qualify as candidates for team leader positions.
– In a third line of argument it was stated that the requirement level for skilled workers in production is usually very high and broad. The tasks can thus not be completed by individual persons or generalists. An often applied solution in the companies is e.g. to organize maintenance in teams composed of differently

qualified specialists. As a rule the technicians fill an important position as they contribute professional experience combined with high technical qualifications.
– An especially interesting line of argument was the fact that in highly automated plants the software is the interface for all technical solutions. In this case all tasks, above all service, maintenance and repair, have been structured based on software tasks. Highly qualified technicians are thus necessary to safeguard this software-technological access to plants. The programming is left to the engineers.

These remarks underpin that a sharp differentiation into mechanics and electrics is no longer made. IT-based tasks in maintenance are clearly increasing. The above mentioned task profiles are mostly an integration of partial tasks, often in connection with an interdisciplinary cooperation within a network and with a high grade of autonomy. Another indication is the ever closer interrelationship of technological units such as CPS and work organizational structures to optimize maintenance tasks.

The merger of information technological processes and maintenance processes calls for an orientation of initial and further training towards the changed technological level and above all the changed perspective. The decentralized intelligence linked to "Industry 4.0" leads to an increased availability of data which are highly process-relevant for skilled workers. Maintenance still requires traditional, handicraft abilities and skills as well as the mastering of SPS, robotics, pneumatics, hydraulics, drive technology etc. They are, however, no longer sufficient. Therefore it is no longer enough to just amend of the occupational profiles. The authors of the bayme vbm Studie et al. (2016, p. 4) rather call for a massive amendment of process orientation in occupational profiles. The perspective of maintenance processes have to be considered based on the process of informatization. Occupational and further training profiles must focus on these central developments. The above mentioned fields of action are the basis for competence development and indicate a clear path of how to meet the described demand for occupational profiles.

8 Conclusions

Changing production systems with an integration of more and more cyber-physical systems must be considered in the curricula for initial and further training. Employees can thus contribute to long-term improvement processes and process optimizations within the framework of the further development of the companies. This is not only true for occupations such as mechatronics, electronics technicians for automation technology or industrial mechanics who are currently mostly working in maintenance departments but is also relevant for other industrial metal-technological and electro-technological training occupations. All industrial occupations must intensively deal with questions of digitalization and networking within "Industry 4.0".

The area of maintenance is playing a decisive role for the flawless operation of the factory of the future. If the future development aims at the described assistance systems and if man (skilled workers on shop-floor level) is given the chance for co-shaping, "Industry 4.0" can be used as an "assistance system". The skilled worker will be able to contribute his or her individual competences for the work process. The competence requirements postulate that the necessary information for the mastering of work processes is provided and that there are adequate qualification approaches available for competence development. Skilled workers and technological applications would thus control and influence one another while man would still have the power of decision.

Skilled workers have to deal with increasing demands in terms of interpretation of system data. Above all analysis capacities and thinking in networks are required in order to deal with abstract information and to get a swift overview of the production process. One of the most important challenges of the implementation of "Industry 4.0" will be to control the flood of information (Big Data) and to just provide information required for the specific work process. If this succeeds, technology can be seen as a kind of assistance system which supports human beings with their maintenance work and which improves error and malfunction analyses. This can only succeed if the skilled workers in the maintenance departments are actively involved into the development and the implementation of CPS technology for maintenance.

Skilled workers on the shop-floor level must receive processed data of the automated control of plant conditions to support them with trouble shooting, with finding damage symptoms and with the analyses of causes. Maintenance along the value-added chain requires in-firm (vertical integration) and external cooperation (horizontal integration), transparency and trust. When using the idea of "Maintenance 4.0" as a highly complex, versatile and flexible system with man in the center—i.e. the skilled worker on shop-floor level—the system also needs skilled workers who act as decision makers, controllers, maintenance staff, co-shapers and experts. Vocational education has the decisive task to qualify skilled workers in a way that they are able to cope with these tasks.

References

Abicht L, Spöttl G (eds) (2012) Qualifikationsentwicklungen durch das Internet der Dinge. W. Bertelsmann, Bielefeld

acatech (eds) (2015) Smart Maintenance für Smart Factories. Mit intelligenter Instandhaltung die Industrie 4.0 vorantreiben (acatech POSITION). Herbert Utz Verlag, München

bayme vbm Studie; Spöttl G, Gorldt C, Windelband L, Grantz T, Richter T (2016) Industrie 4.0—Auswirkungen auf Aus—und Weiterbildung in der M+E Industrie. Studie herausgegeben von bayme vbm. Die bayerischen Metall—und Elektro-Arbeitgeber, München. www.baymevbm.de/industrie4.0. Accessed 27 May 2016

Biedermann H (2014) Anlagenmanagement im Zeitalter von Industrie 4.0. In: Biedermann H (eds) Instandhaltung im Wandel. Industrie 4.0—Herausforderungen und Lösungen. TÜV Media Gmbh, Köln, pp 23–32

DIN 31051 (2003) Norm DIN 31051:2003-06: Grundlagen der Instandhaltung

Gorecky D, Schmitt M, Loskyll M (2014) Mensch-Maschine-Interaktion im Industrie 4.0-Zeitalter. In: Bauernhansl T, ten Hompel M, Vogel-Heuser B (eds) Industrie 4.0 in Produktion, Automatisierung und Logistik. Springer Vieweg, Wiesbaden, pp 525–542

Günther G, Eckhoff R, Markus M (2014) Bedürfnisse, Anforderungen und Trends in der Instandhaltung 4.0. Projektkonsortium Instandhaltung 4.0, p. a. Salzburg Research Forschungsgesellschaft m. b. H. http://instandhaltung40.salzburgresearch.at/wp-content/uploads/IH40-Analyse-Kurzfassung-final.pdf. Accessed 24 Apr 2015

Hirsch-Kreinsen H (2014) Wandel von Produktionsarbeit—“Industrie 4.0”, Soziologisches Arbeitspapier Nr. 38, Dortmund. http://www.wiso.tu-dortmund.de/wiso/is/de/forschung/soz_arbeitspapiere/Arbeitspapier_Industrie_4_0.pdf. Accessed 10 Feb 2015

Höhnel A, Wächter M, Bullinger AC (2015) Instandhaltung von morgen. Wandlung von Instandhaltungsprozessen durch sozio-cyber-physische Systeme. Instandhaltung aktuell, 91

Mathur S, Weiß JM (2014) Werte durch Wartung—Wie Industrie 4.0 die Instandhaltung verändert und zum Mehrwerttreiber im Unternehmen wird. In: Biedermann H (eds) Instandhaltung im Wandel. Industrie 4.0—Herausforderungen und Lösungen. TÜV Media Gmbh, Köln, pp 143–153

Schlick J, Stephan P, Loskyll M, Lappe D (2014) Industrie 4.0 in der praktischen Anwendung. In: Bauernhansl T, Hompel M, Vogel-Heuser B (eds) Industrie 4.0 in Produktion, Automatisierung und Logistik. Springer Vieweg, Wiesbaden, pp 57–84

Windelband L (2014) Zukunft der Facharbeit im Zeitalter “Industrie 4.0”. J Tech Educ JOTED 2 (2):138–160

Windelband L (2016) Veränderungen in der Arbeitswelt, Kompetenzen und Lernen in der “Instandhaltung 4.0”. In: Lernen & Lehren, Heft 121, pp 16–22

Windelband L, Dworschak B (2015) Arbeit und Kompetenzen in der Industrie 4.0. Anwendungsszenarien Instandhaltung und Leichtbaurobotik. In: Hirsch-Kreinsen H, Ittermann P, Niehaus J (eds) Digitalisierung industrieller Arbeit. Nomos Verlag, Berlin, pp 71–86

Windelband L, Spöttl G (2012) Diffusion von Technologien in die Facharbeit und deren Konsequenzen für die Qualifizierung am Beispiel des “Internet der Dinge”. In: Faßhauer U et al (eds) Berufs—und wirtschaftspädagogische Analysen. Budrich Verlag, Opladen/Farmington Hills, pp 205–219

Development of a Business Game for Teaching the Kanban Method

David Brown, Sven Hinrichsen, and Melissa Paris

Abstract

Learning retention is improved when in addition to the learning methods of "classical" instruction (auditory, visual), further learning channels (discussing, doing) are addressed (Gasser, Erfolgreich lernen—Praxis-Tipps für sofortigen und nachhaltigen Lernerfolg, Books on Demand, Norderstedt, 2014). In view of this, the coursework for the degree in Industrial Engineering has been designed to include teaching formats with a high level of interaction among students. In keeping with this teaching approach, an additional business game was developed for the Production Systems module. This game provides knowledge about the structure of the logistical workflow within production in general and the Kanban method in particular. The first step in designing the game was to define learning objectives. Based on these, a basic concept was derived. This basic concept applies the guideline for the introduction of Kanban by Geiger et al. (Kanban—Optimale Steuerung von Prozessen, Carl Hanser, Munich, 2011). The detailed concept was then worked out over the course of further development, leading to the design of an assembly line for Lego cars as well as the design of tasks on the subject of Kanban capability and for calculating the number of Kanban. The business game illustrates, inter alia, the difference between the Kanban method and a conventional production planning and control system. The simulation has been tested in a university course, evaluated and, in response to the positive feedback from students, integrated in the course instruction. The conclusion of the paper provides an outlook on the further development of instruction in the context of the digitization and networking of production.

Keywords

Business game • Kanban • Didactics • Industrial engineering

D. Brown (✉) • S. Hinrichsen • M. Paris
Ostwestfalen-Lippe University of Applied Sciences, Lemgo, Germany
e-mail: david.brown@hs-owl.de; sven.hinrichsen@hs-owl.de; melissa.paris@hs-owl.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_7

1 Educational Concept at the Ostwestfalen-Lippe University of Applied Sciences

1.1 Requirements for Industrial Engineers

Industrial engineering (IE) can be regarded as a very comprehensive operational function. How this function is integrated in the structure of a company can vary considerably. Unlike large companies, SMEs rarely include IE as an independent functional department. The activities are often assigned to several functionaries (Spath et al. 2011; Hinrichsen et al. 2015). Many industrial engineering tasks are performed by the department of Work Scheduling or by executives in production, generally in addition to their day-to-day business (Hensel-Unger 2011).

In addition to the technical and methodological skills shown as an example in Fig. 1, the industrial engineer also requires social and personal skills (Hinrichsen et al. 2015). These social skills are made necessary in part by the industrial engineer's responsibility for initiating and seeing through change processes. Moreover, the increased complexity of production—caused by ever shorter product life cycles and an ever wider range of products and options—has heightened the proportion of highly qualified activities in production and production-related areas. This has led, in turn, to the increased importance of the cooperation and

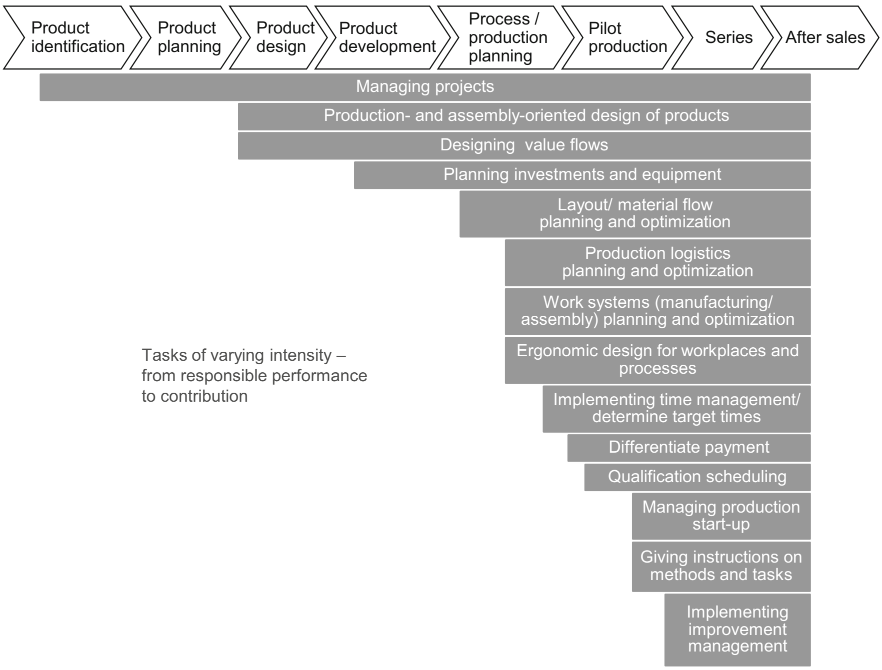


Fig. 1 Tasks of industrial engineering through the process of product creation according to Ifaa (2010), Stowasser (2010), Bokranz and Landau (2012), Hinrichsen et al. (2015)

communication skills of executives—also in IE (Erpenbeck and Heyse 1996). The relevance of these skills is clear, particularly considering that a large proportion of the work activity of an executive consists of direct communication (von Rosenstiel 2003). For the support and development of fellow employees, didactic skills are also essential. In particular, industrial engineering employees must be able to provide and implement work standards and methods (Hinrichsen et al. 2015).

An essential personal skill for executives—including industrial engineering executives—lies in the ability to detect interconnections between individual tasks and business parameters, as well as the ability to coordinate objectives, pursue them consistently and assess their achievement (Frey and Balzer 2003). Communicating one's own point of view and winning people over requires a corresponding self-awareness and the ability to engage in constructive self-criticism with the aim to learn from one's own behavior (Hinrichsen et al. 2015).

1.2 Industrial Engineering Degree

Due to the requirements facing IE experts, the faculty of production engineering and management at the Ostwestfalen-Lippe University of Applied Sciences has established a bachelor program with a main focus on industrial engineering. The curriculum was specially designed to ensure that the skills required in operational practice are taught. As a result the coursework for the degree in Industrial Engineering provides a basic education in mathematical, technical and business oriented content. Subjects include mathematical planning procedure, labor and economic law as well as quality management. The IE degree program is mainly structured around nine modules as shown in Fig. 2. These modules are linked and build upon each other (Hinrichsen et al. 2015).

The basic course Industrial Management (IBL) aims to provide knowledge about the essential functions of an industrial company and the relationships between its

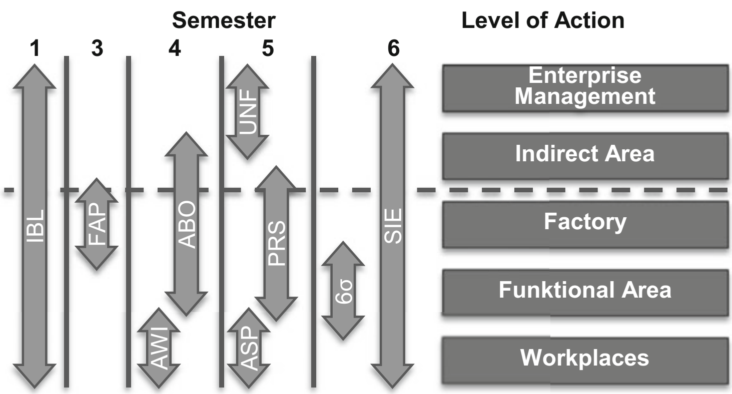


Fig. 2 Core modules of the industrial engineering degree with an assignment to levels of action (Hinrichsen et al. 2015)

functions. Furthermore, students work on particular business methods in exercises. The course Factory Planning (FAP) teaches about the planning, design and optimization of factories. Over the course of the semester, a factory planning project is carried out in which small groups of students plan the relocation of a production system, from the data analysis through to the planning and evaluation of the layout. The module Ergonomics (AWI) focuses on the physiological basics of workplace design. In small groups, students analyze a company's workplaces in self-organized projects, analyzing the ergonomic design and developing their own design recommendations. The module Work and Business Organization (ABO) considers aspects such as the organization of a production facility along with the basics of production planning and control systems (Hinrichsen et al. 2015).

The modules Work System Design (ASP), Production Systems (PRS) and Enterprise Management (UNF) include instruction on essential methods from the Toyota Production System, the REFA methodology and the main features of the MTM method. Alongside the lectures, students carry out lab exercises in 4-hour to 6-hour lesson blocks in a "Learning Factory" (IE Lab). To develop students' social and personal skills alongside their technical and methodological skills, the teaching concept of these modules is anchored within the overall business scenario of Lipper Assembly & Logistics GmbH. This virtual enterprise is engaged in the assembly of products in different variants, whereby the entire process chain (material and information flow)—from incoming goods, through the order picking and assembly, through to shipping—is considered. This involves several mobile assembly workstations and a picking area on a total of 140 square meters. In addition, a workshop area with small machines and tools is used to produce fixtures or design workplaces with the "Cardboard Engineering" method (Hinrichsen et al. 2015). Through teamwork in small groups on open problems, cooperative learning and problem-solving skills are promoted, strengthening students' social skills while simultaneously building up cognitive skills through the joint work on complex tasks (Mandl et al. 2004). The relatively high degree of freedom in self-directed processing of tasks requires the instructors to provide moderating communication, in order to promote reflection on the procedures (Hinrichsen et al. 2015).

During the sixth semester the Six Sigma method is taught in a business game (6σ) over several days. Moreover, in this semester, as part of the seminar on Industrial Engineering (SIE), students are assigned to write a paper on a scientific issue in Industrial Engineering. In addition to the curriculum, students have the opportunity to attend recognized external training courses such as REFA's basic training course, MTM-1/-UAS and basic training in SAP.

As part of the module Production Systems (PRS), students are taught the methods of the Toyota Production System, including the Kanban method, as it is of great importance for the control of the material flow in many operational production systems. For the successful implementation of Kanban in a production system, training employees in this method has high priority, because this can broaden the understanding and acceptance of this method (Dickmann 2007). For training and passing on knowledge, different methods are available. The methods differ from one another in terms of the degree of student participation as well as which learning areas are addressed (Schelten 2005; Schlick et al. 2010). The teaching and learning methods of a business game offer several advantages over conventional methods in which communication is carried out fairly unilaterally from the teacher to the student. Cooperation in a business game encourages holistic learning (addressing all learning channels), the motivation to learn and the willingness to cooperate, because complex, realistic tasks are processed independently in a team (Luttemann 2000; Gasser 2014; Schwägele 2016). The learning process is supervised by the instructors, for example through questions to the group. To teach the Kanban method, a business game with a duration of four class hours was developed. Main features of the simulation are presented below.

2 Kanban Business Game

2.1 Learning Objectives

The business game was developed with the objective of providing students with a comprehensive picture of the Kanban method as well as transferring what they have learned into operational practice. Students should be able to determine the benefits and possible risks of implementing Kanban. Moreover, they should have the ability to coordinate and plan the process of establishing the method in actual practice. To achieve these learning objectives, the differences between a conventional production planning and control system (PPC), which uses the push principle, and the basic functioning principle of Kanban should be illustrated. It is also important to help students understand the rules for Kanban, as these are required for successful implementation. Knowledge about possible use cases and limitations of the method will help students when it comes to deciding whether Kanban is suitable for controlling material flow in a particular business scenario.

2.2 Theoretical Principles

Implementing new methods in operational areas can cause various complications. Specially developed procedure models can guide the process of implementation. A process model by Geiger et al. (2011) serves as a guideline for the introduction of Kanban and consists of the following steps:

1. Checking the Kanban capability
2. Selection and definition of control loops
3. Calculating the number of Kanban
4. Selection of Kanban tools
5. Testing and implementing the Kanban system.

Checking the Kanban capability (step 1) is necessary because not all components are suitable for a Kanban control system. Based on criteria such as the consumption

Table 1 The functions of Kanban and their rules for the application of the method according to Ohno (2009)

No.	Functions of Kanban	Kanban rules
1	Provides picking or transport information	Later process picks up the number of items indicated by the Kanban at the earlier process
2	Provides information on the production process	Earlier process produces items in the quantity and sequence indicated by the Kanban
3	Prevents overproduction and unnecessary transport	No items are produced or transported without a Kanban
4	Serves as a work order, attached to Kanban containers	Always attach a Kanban to the goods
5	Prevents defective products by detecting the working process which makes the error	Defective products are not sent on to the subsequent process. The result is 100 % defect-free goods
6	Reveals existing problems and allows inventory control	Reducing the number of Kanban increases the sensitivity → waste becomes visible

flow and product characteristics, this step must be carried out for each component. During the selection and definition of the control loops (step 2), the planner first has to gain an overview of all material and information flows in order to develop the control cycles in detail (Geiger et al. 2011). Not only is the division of responsibilities in the control circuits important, but also the detailed definition of the Kanban rules in this step (Klevers 2009). The determination of the required number of Kanban cards (step 3) is accomplished through use of a specific calculation. Finally, the required Kanban tools (e.g. cards, letter boxes, signals, boards, dynamic racks) are selected (step 4) before the system is tested and implemented in step 5 (Geiger et al. 2011).

In addition to the process model for implementing Kanban, the functions of the Kanban and their rules of application shown in Table 1 are an essential part of the method. If Kanban is used in business practice without paying attention to the correct application, the method can worsen the situation (Ohno 2009). For example, when the Kanban method is correctly applied, the number of Kanban determines the maximum work-in-process (WIP). However, if components are produced without the Kanban system as a trigger, this can lead to overproduction and result in waste in the processes.

2.3 Basic Concept

The business game was modelled on Geiger's guideline for implementing the Kanban method (Sect. 2). Participants will get to know the essential criteria for introducing Kanban and simultaneously develop a basic understanding of this method.

Table 2 Objectives of the business game based on the process model according to Geiger et al. (2011)

No.	Steps for implementing Kanban	Taught knowledge
1.1	Checking the Kanban capability	Determination of Kanban capability of parts based on an ABC/XYZ analysis
1.2	Selection and definition of control loops	Teaching theoretical basics of Kanban types and rules
1.3	Calculating the number of Kanban	Calculating the number of Kanban referring to Geiger et al. (2011)
1.4	Selection of Kanban tools	Teaching theoretical basics of Kanban tools
2	Testing and implementing the Kanban system	Illustration of push and pull through a simulated production system

The business game is divided into two phases. The first phase includes the steps shown in Table 2 from number 1.1 to 1.4. Initially, the Kanban capability is determined by comparing the consumption flow with the value of the parts. In a task with a specific situation this is calculated by using the ABC/XYZ analysis. A basic knowledge of the Kanban types and rules is taught to prepare students for the assignment of selecting and defining the control loops within the method. The students are also assigned to calculate the necessary number of Kanban. Further in, over the course of the business game, the theoretical basics of Kanban tools are taught, before phase two begins.

The content of the second phase—testing and implementing the Kanban system—uses a Lego simulation to represent a production environment. At first a "conventional" PPC which uses the push system is simulated before the Kanban method is shown. Simulating the push and pull principles illustrates the difference between the two principles of material control. Furthermore, students are taught about the Kanban functions, the importance of Kanban rules as well as the advantages and disadvantages of the method. The functions of Kanban shown in Table 1 are taught through the rules of the game, since they are derived from the Kanban rules.

2.4 Detailed Concept

The Lego simulation consists of three phases and is designed for six players. The first phase simulates a conventional PPC using the push principle. As Fig. 3 shows there are three assembly stations (WP 1–3), in which three different variations of Lego cars are assembled. Furthermore, an order picker sorts the required items into containers. A store man then distributes the filled containers to the appropriate workplace. A sixth player documents the amount of assembled vehicles and disassembles them again. The phase ends when 20 cars have been assembled. After ten finished cars the WIP is determined by counting all items at the assembly stations. As soon as the push simulation is finished, the second phase begins. In this phase, the structure of the game is changed according to Fig. 3 and the Kanban method (pull principle) is simulated. The former store man assumes the role of the

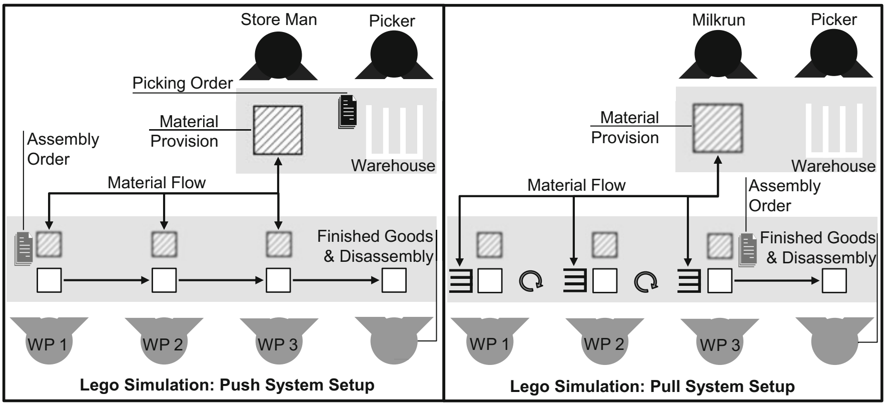


Fig. 3 Structure of the Lego simulation

Milkrun while all other players stay in the same role. A two-bin Kanban is implemented so that the amount in one container covers the replenishment time of the second one. The second phase also ends after 20 assembled cars. In the last phase of the Lego simulation (phase 3), the given data are documented and evaluated and the conventional PPC and the Kanban method are compared.

In order for the simulation to work as planned, it is necessary for the students to strictly follow the rules of the game. In the push phase the assembly orders go through production from the first to the last workstation. It is important to note that the full batch size of an assembly order must be finished before both the assembly order and the goods are given to the next workstation. In the pull phase of the simulation the rules of the game result from the Kanban rules shown in Table 1. The customer now pulls the value stream which means that the assembly orders are placed at workstation three. The next rule is that no items are made or transported without a Kanban. This means that the pre-assembly of components or emptying containers and gathering items on the assembly stations is not allowed. Furthermore, a Kanban serves as a work order. That is why the containers have to be placed on their defined spot immediately after being emptied, to trigger the production in the earlier process. The next rule is that no items are made or transported without a Kanban. This will prevent overproduction and unnecessary transport. For the Milkrun it is important to store the empty containers in the warehouse according to the FIFO (first-in-first-out) principle. This limits the variance in the replenishment time. Also important to mention is that the players must not pick items from both containers simultaneously.

After documenting the necessary data, the students analyze the data in phase three. The data should show that the WIP in phase one increases constantly whereas the WIP evens out when using the Kanban method.

After the Lego simulation is ended, the next part of the business game begins. The students have to determine the Kanban capability of the components of the Lego cars. In this assignment, the participants have to carry out an ABC/XYZ analysis of the Lego items. By comparing the value of an item with its consumption flow and taking the safety stock level into account, a statement about the Kanban capability can be made for the different items. In an additional assignment, the students calculate the necessary number of Kanban for given components. This is done using the formula to calculate the Kanban size according to Geiger.

2.5 Evaluation and Interpretation of the Results

The business game was played and evaluated together with students. The evaluation contained questions about the Kanban method as well as questions about the learning method itself. As Fig. 4 shows, the participants demonstrated a basic understanding of the Kanban method as well as the difference between the push and pull principles. Furthermore, the participants felt able to determine the Kanban capability of parts as well as draw conclusions about the advantages and disadvantages of the method.

The business game helps students in their learning process and they even express a preference to this learning method compared to a "classical" lecture. In addition, the teamwork in this simulation was positively received, indicating that

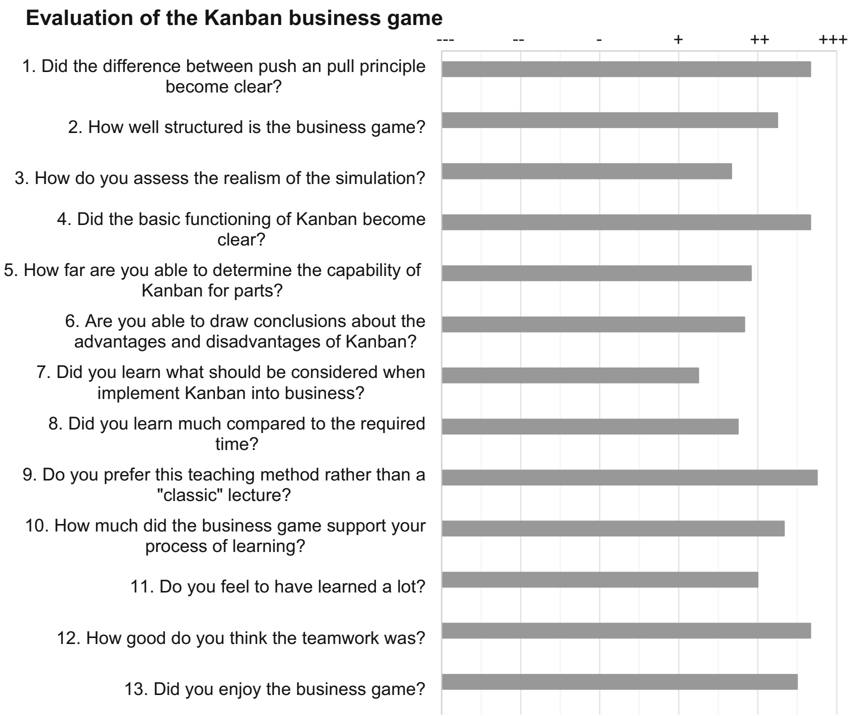


Fig. 4 Evaluation results of the Kanban business game

Fig. 5 Putting the Kanban business game into practice

communication skills are strengthened. There is wide acceptance of the teaching method, as this supports the learning process. Having been tested at the university, the business game could also have applications in private businesses. Companies could use the simulation as an educational program to help employees better understand Kanban and to support the learning process.

Figure 5 shows students performing the Kanban simulation at workstations 1–3. The picture shows the Kanban racks, assembly instructions and spaces for returning empty containers. The WIP is limited by the amount of Kanban containers in the business game.

3 Outlook

The requirements for the design of production systems are changing. Engines behind this change are ever shorter innovation and product life cycles, increasing variety, and decreasing lot sizes, down to lot-size-one. At the same time, production technology keeps on developing. With flexible, mold-free manufacturing techniques such as laser technology and 3D printing, products can be customized, making these technologies increasingly important. Additionally, over time the development of information technology is leading more and more to a fully digital process chain throughout the entire product realization process (from product development to distribution of finished products).

To ensure that research and course content address these trends, the Ostwestfalen-Lippe University of Applied Sciences has collaborated with the Fraunhofer IOSB-INA institute to found the SmartFactoryOWL. The

Fig. 6 Projection-based assistance system for manual assembly

SmartFactoryOWL is an open institute for research and demonstration. It also serves as a testing ground for SMUs from the highly industrialized region of Ostwestfalen-Lippe. The institute is located on the campus of the Ostwestfalen-Lippe University of Applied Sciences, offering an outstanding learning environment for the students. Within the factory, the Laboratory of Industrial Engineering is responsible for the research field of assembly. In cooperation with manufacturing companies and factory outfitters from the region, the lab has developed and implemented assembly systems. The different assembly systems share in common that real products from real companies are assembled with these systems. In addition, different technological developments in the context of digitization and networking of processes are demonstrated by the individual systems.

With the aid of these assembly systems, lectures can be performed with a practical orientation. Furthermore, research issues with practical relevance can be handled in close collaboration with companies. For example, a projection-based assembly assistance system for supporting employees during the assembly of complex products has been developed at the Laboratory of Industrial Engineering (Unrau et al. 2016). The system shown in Fig. 6 can be controlled by voice, gesture or touch. It supports the training process and can reduces errors.

In the upcoming semester new exercises will be designed to allow students to work with the developed assembly systems and the various assistance systems. The exercises will demonstrate the possibilities of digitally supported work processes as well as methods for designing work systems to accommodate customers’ specific requirements.

References

Bokranz R, Landau K (2012) Handbuch Industrial Engineering, Band 1, 2 Aufl. Schäffer-Poeschel, Heidelberg

Dickmann P (2007) Schlanker Materialfluss. Springer, Berlin

Erpenbeck J, Heyse V (1996) Berufliche Weiterbildung und berufliche Kompetenzentwicklung. In: Arbeitsgemeinschaft Qualifikations-Entwicklungs-Management Berlin (eds) Kompetenzentwicklung '96: Strukturwandel und Trends der betrieblichen Weiterbildung. Waxmann, Münster, pp 15–152

Frey A, Balzer L (2003) Soziale und methodische Kompetenzen—der Beurteilungsbogen smk: Ein Messverfahren für die Diagnose von sozialen und methodischen Kompetenzen. In: Frey A, Jäger RS, Renold U (eds) Kompetenzmessung: Sichtweisen und Methoden zur Erfassung und Bewertung von beruflichen Kompetenzen (Empirische Pädagogik, 17 (2), Themenheft). Verlag Empirische Pädagogik, Landau, pp 148–175

Gasser J (2014) Erfolgreich lernen—Praxis-Tipps für sofortigen und nachhaltigen Lernerfolg, 2 Aufl. Books on Demand, Norderstedt

Geiger G, Hering E, Kummer R (2011) Kanban—Optimale Steuerung von Prozessen, 3 Auflage. Carl Hanser, Munich

Hensel-Unger R (2011) Entwicklung einer Gestaltungssystematik für das Industrial Engineering (IE) unter besonderer Berücksichtigung kultureller Einflussfaktoren am Beispiel von Tschechien und Polen. Monsenstein und Vannerdat, Münster

Hinrichsen S, Jungkind W, Könneker M (2015) Industrial Engineering-Ausbildung an der Hochschule Ostwestfalen-Lippe. In: Zeitschrift für Arbeitswissenschaft, 69, 4, pp 197–201

Ifaa (ed) (2010) Produktivität steigern erfolgreich mit Industrial Engineering, Ifaa, Düsseldorf

Klevers T (2009) Kanban. FinanzBuch, Munich

Luttemann C (2000) Multimedia in der betrieblichen Weiterbildung—Integration neuer Technologien. Diplomarbeiten Agentur, Hamburg

Mandl H, Kopp B, Dvorak S (2004) Aktuelle theoretische Ansätze und empirische Befunde im Bereich der Lehr-Lern-Forschung. Ludwigs-Maximilians-Universität München, München. http://www.die-bonn.de/esprid/dokumente/doc-2004/mandl04_01.pdf

Ohno T (2009) Das Toyota Produktionssystem, 3 Aufl. Campus, Frankfurt

Schelten A (2005) Grundlagen der Arbeitspädagogik. Franz Steiner, Stuttgart

Schlick C, Bruder R, Luczak H (2010) Arbeitswissenschaft. Springer, Heidelberg

Schwägele S (2016) Planspiel—Lernen—Lerntransfer: Eine subjektorientierte Analyse von Einflussfaktoren. Books on Demand, Stuttgart

Spath D, Korge A, Krause T, Lanza G, Jondral A, Moser R (2011) Hybrides Planungswerkzeug zur adaptiven Auslegung von Lean-Methoden. Zeitschrift für wirtschaftlichen Fabrikbetrieb, Jg 106, pp 413–417

Stowasser S (2010) Produktivität und Industrial Engineering. Angewandte Arbeitswissenschaft: Zeitschrift für die Unternehmenspraxis. Ausgabe 204:7–20

Unrau A, Hinrichsen S, Riediger D (2016) Development of projection based assistance system for manual assembly. In: Proceedings of 6th International Ergonomics Conference—Ergonomics 2016, 15–18 June 2016, Zadar, Croatia, pp 365–370

von Rosenstiel L (2003) Grundlagen der Organisationspsychologie, 5 Aufl. Schäffer-Poeschel, Stuttgart

Training Interventions to Increase Innovation and Productivity in Age-Diverse Teams

K. Küper, W. Rivkin, and K.-H. Schmidt

Abstract

In multiple first world countries, demographic change is expected to lead to an increase in retirement age and associated age-diversity in work-teams. Still, despite these changes, organizations must maintain and even enhance productivity and innovation. The present chapter introduces interventions at the individual and the management level to improve innovation and productivity in organizations with teams characterized by high age-diversity. These interventions were specifically developed, implemented and evaluated in two middle-sized organizations. On the individual level, employees received a training intervention to improve cognitive abilities and another training to enhance stress management abilities. On the management level, leaders were trained in how to mitigate the adverse effects of age diversity on team performance. Results indicate that the cognitive training intervention improved abilities such as updating and maintenance of working memory and that the stress management training reduced work overload and job-related effort. In addition the leadership training reduced age-stereotypes and associated conflicts in teams. To enhance the effects of the training interventions the introduction of follow-up booster sessions is recommended.

Keywords

Age-diversity • Cognitive training • Leadership training • Stress management training • Innovation

K. Küper (✉) • W. Rivkin • K.-H. Schmidt
IfADo—Leibniz-Institut für Arbeitsforschung an der TU Dortmund, Ardeystraße 67, 44139 Dortmund, Germany
e-mail: kueper@ifado.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_8

1 Introduction

Germany and other European countries are facing substantial changes in the demographic composition of their populations, which include a significant increase in mean age and thus will probably necessitate an extension of employees' working lives to maintain productivity and associated economic growth. A side effect of this development is a much more age-diverse composition of the work force. At the same time, modern work environments are characterized by highly dynamic and complex demands that require employees to be flexible and innovative. In view of these boundary conditions, the present contribution offers a novel and holistic intervention approach, which aims at facilitating innovation and productivity of age-diverse work groups in small to medium sized organizations. The approach focuses on both the individual as well as the management level and combines individual training interventions to promote employees' cognitive functions and stress resilience as well as a leadership training intervention to enhance leaders' abilities in managing age-diverse teams.

The development and evaluation of the intervention approach was conducted in cooperation with two medium-sized companies from the Ruhr area, the Kommanditgesellschaft Deutsche Gasrußwerke GmbH & Co. and ABC-Logistic GmbH. It was funded by the Federal Ministry of Education and Research.

2 Interventions at the Individual Level

Cognitive functions are essential to the performance of a multitude of activities of daily living (including work). They are crucial for the regulation of such diverse activities as action planning, the inhibition of irrelevant information or task-inappropriate action tendencies, multi-tasking and task-switching as well as working memory, i.e. the ability to store and manipulate a limited amount of information over a limited time span (Miyake and Shah 1999). Fluid cognitive functions have been shown to have a remarkable degree of plasticity across the life span. Although they begin to decline as early as in middle age, they can also be substantially improved up to a very old age, in particular through training interventions, which usually involve the completion of computer-based tasks at an adaptive level of difficulty.

Importantly, cognitive training interventions may not only improve the trained cognitive functions themselves but can also benefit untrained functions as well as the performance of daily life activities (e.g. driving, see Cassavaugh and Kramer 2009). Particularly training interventions aimed at improving working memory have shown substantial transfer effects to attentional processes (Lilienthal et al. 2013; Salminen et al. 2012), cognitive control (Küper and Karbach 2015; Salminen et al. 2012), and general fluid intelligence (Jaeggi et al. 2008; see Au et al. 2014, for a meta-analysis).

High levels of work-related and psychosocial stress are not only detrimental to physical and psychological health and well-being, but also have a negative impact on cognition, in particular fluid cognitive functions (McEwen 2007; Stawski et al. 2006). The ability to manage stress can, however, also be trained effectively (Gajewski et al. 2012). The notions that stress can have especially harmful effects in older age (Kukolja et al. 2008; Uchino et al. 2006) and stress reactions become chronic if left untreated (Martin et al. 2001), make stress management interventions appear particularly relevant for modern work contexts.

2.1 Cognitive Training

Our computer-based cognitive training was performed in groups of up to five employees under the guidance of a certified trainer. The training involved four tasks, which required (a) the retention and reproduction of sound and colour sequences or (b) sequences of spatial positions, (c) the retention and addition of number sequences and (d) the counting of dice eyes while adhering to varying counting rules (www.freshminder.de). All tasks were thus aimed at improving working memory function with tasks (a) and (b) requiring only the maintenance and tasks (c) and (d) requiring the maintenance and manipulation of information in the working memory. Across 4 weeks, participants completed a total of eight 60-min long training sessions during which they attended to each of the four tasks in equal parts.

2.2 Stress Management Training

The stress management training was based on the 'Training emotionaler Kompetenzen' (Berking 2010) and was conducted by a certified trainer in groups of up to ten employees. The training involved educational elements about the causes and physical and psychological effects of stress as well as practical training exercises of methods aimed at reducing the subjectively perceived stress level. The practical exercises included breathing, muscle relaxation and mindfulness techniques. In the course of the training, a particular focus was put on stressful situations from the participants own work experience and the successful application of stress reduction methods in these situations. Participants completed a total of eight 60-min training sessions across a period of 4 weeks.

2.3 Participants of the Cognitive and Stress Management Training

Participants were employees of the above mentioned medium-sized companies who volunteered to be part of either a cognitive training group ($n = 21$; mean age = 35.7 years; age range = 17–57 years; 7 women), a stress management training group ($n = 38$; mean age = 42.6 years; age range = 20–64 years; 16 women) or a control group ($n = 35$; mean age = 42.2 years; age range = 19–62 years; 18 women). All participants provided written informed consent.

Participants from all groups first took part in a pre-test session. In the following 4 weeks, participants from the two training groups completed eight sessions of either the cognitive training or the stress management training. In contrast, participants from the control group did not participate in any training activities. Subsequently, all participants completed a post-test session, which was identical to the pre-test session.

3 Intervention at the Management Level

3.1 Leadership Training

Contrary to the wide-spread notion that age diversity in teams leads to more productivity and innovation, recent research has revealed adverse effects of age diversity on team performance. As causes of these adverse effects age stereotypes resulting in conflicts within the team were identified (cf., Ries et al. 2010, 2012). In order to raise the awareness of companies for the innovative potential of older employees and to avoid the adverse effects of age diversity, a leadership training was developed, implemented and evaluated.

The training is based on the model of 'servant leadership' (Ehrhart 2004; Liden et al. 2008). Servant leadership is a leadership style that focuses on serving multiple stakeholders of the organization. Hale and Fields (2007) define servant leadership as "an understanding and practice of leadership that places the good of those led over the self-interest of the leader, emphasizing leader behaviors that focus on follower development, and de-emphasizing glorification of the leader" (p. 397). Thus, servant leaders do not lead for their own or their organization's benefit, but for the benefit of multiple stakeholders, and especially their employees. Consequently, servant leaders do not lead through formal authority or charisma as proposed by other leadership styles such as transformational leadership (e.g., Burns 1978; Weber 1921/1946), but instead rely on "one-on-one" communication to understand the abilities, needs, desires, goals, and the potential of their employees (Liden et al. 2008; p. 162). Additionally, servant leaders shape their employees' views and values to encourage them to become servants and servant leaders themselves (Greenleaf 1977; Rivkin et al. 2014).

Even though servant leadership overlaps with other leadership styles such as transformational, ethical, and authentic leadership (Bass and Avolio 1990; Brown and Treviño 2006; Walumbwa et al. 2008), it also differs from these constructs in certain key aspects. First, Graham (1991) argues that leadership styles such as transformational leadership fail to consider the importance of a moral compass, which constitutes a crucial aspect of servant leadership. Second, in contrast to leadership styles that include a moral component (e.g., authentic leadership and ethical leadership), servant leadership focuses on the success of multiple stakeholders of the organization. Third, servant leadership is especially focused on the interests and competencies of followers. Thus, servant leaders aim to develop their employees, and to support their growth and success (Smith et al. 2004).

Based on the theoretical background provided above, we developed a new training for leaders of age-diverse teams (see Bröker et al. 2015). Specific training objectives included increasing awareness for age diversity and its impact on group performance (specifically how misunderstandings and problems in age-diverse teamwork can occur), developing and practicing skills, e.g. avoiding age prejudices in communication and fostering appreciation of age diversity. Furthermore, participants developed personal action plans for implementing strategies in their everyday leadership routine. The training consists of four consecutive modules.

The first training module focuses on creating awareness among the participants for the topic of demographic change and age diversity as well as providing them with the required knowledge on the effects of age diversity and its importance for teamwork. In order to deepen the understanding that it is not age per se that influences teamwork, but rather the characteristics we associate with certain age groups (Bell 2007; Harrison et al. 2002), different characteristics of young and old employees were discussed.

Already at this early point in the training, participants were asked to reflect their own images of young and old team members. Aiming at advancing the understanding of age diversity as a resource, the trainer further lectured on the advantages and disadvantages of age diversity in teams.

In the second module, the focus shifted from general principles to discussing hands-on strategies for dealing with the challenges of age diversity and drawing practical implications for leaders' everyday work life. It aimed at identifying employees' individual strengths and weaknesses with regard to their age. In order to achieve this, participants discussed their images of 'age' and received information concerning age-related changes (with regard to health, learning abilities, motivation and social skills) and the high inter-individual variability of performance across the lifespan (Baltes and Baltes 1990). On the basis of the notions obtained during these discussions, participants developed plans on how to delegate tasks, depending on the strengths of younger and older employees.

The third module focused on empowering employees in age diverse teams. Therefore, the trainer shed light on the detrimental role of conflicts and their impact on employees' motivation. Furthermore, it was clarified that establishing a culture of trust and developing possibilities for employees to autonomously determine certain aspects of work can mitigate the adverse impact of age diversity on performance and innovation.

Finally, the topic of the last module was appraisal of employee performance. Participants were trained in conducting appraisal interviews during which they focus on the positive aspects of employees work.

Participants were invited for a half-day booster session aiming at recapitulation and reflection 4 months after the training. After a brief repetition of the main training topics, the trainer introduced the method of group coaching (Lippmann 2009). Participants were encouraged to discuss problems they experienced in applying the training topics to their leadership routine. Based on the results of the group discussions, participants had the chance to go through the described situations in role plays and reflect upon their challenges in order to develop new, successful strategies for subsequent implementation. At the end of this workshop, the most promising strategies for dealing with age diversity were summarized and visualized.

3.2 Participants of the Leadership Training

Participants were 36 (13 women) leaders from the above mentioned companies. Data was collected on paid company time before the training (pre-training) and 4 months after (post-training). An additional follow-up measure in the training group was conducted 12 months later. Participation was voluntary and anonymous. To evaluate the effectiveness of the training program we obtained data from all (36) participants and their 162 (50 women) employees. The average team size was 4.5 members. Participants mean age was 39.61 years (range = 26–55 years) and employees mean age was 37.26 years (range = 19–62 years).

4 Results

4.1 Cognitive Training Intervention

Participants in the cognitive training group showed improved performance at post-test measurement compared to pre-test measurement in three of the four trained tasks: Cognitive training improved working memory updating as indexed by the (a) number addition task as well as working memory maintenance as measured in the (b) spatial sequences task and the (c) dice counting task. In the latter two tasks, the control group also showed performance benefits at post-test measurement. Paired t-tests did, however, indicate that pre-post differences in task performance were substantially more pronounced in the cognitive training group than in the control group. Cognitive training also had marginally significant positive effects on immediate memory recall and accuracy on compatible trials of the Stroop task measuring cognitive control.

4.2 Stress Management Training Intervention

In contrast to the control group and the cognitive training group, participants from the stress management training group reported reduced work overload and job-related effort at post-test measurement compared to pre-test measurement. Participants were also less prone to chronic worrying and scored marginally lower on the overall chronic stress screening scale.

The stress management training additionally benefitted working memory updating processes as indexed by reduced reaction times and marginally higher accuracy in an n-back task as well as marginally improved performance in the number addition task. The stress management training group also showed shorter reaction times in a matrix reasoning task at post-test measurement compared to pre-test measurement. This was, however, most likely due to poor performance at pre-test measurement rather than a genuine improvement in fluid intelligence. In line with these results, we observed an increased number of correctly solved matrix reasoning task matrices at post-test measurement for the control group, who also showed poor performance at pre-test measurement.

4.3 Leadership Training Intervention

Evaluating the effectiveness of the leadership training intervention, we analyzed its impact on various indicators of team functioning and performance, assessing data both from leaders and their team members 4 months and 1 year after the training, respectively. In doing so, we gained several important new insights. First, our findings indicate that the salience of age diversity was not increased by the training, even though it could have been assumed that age differences become more prevalent due to discussing such differences in our training. This result can be considered positive because high salience of age differences in teams has been observed to hinder team performance (Ries et al. 2010).

Second, we found that trained leaders showed more appreciation of age diversity after the training. The amount of age prejudices of trained leaders was also reduced and the training was found to have a long-lasting effect on age prejudices as their level remained low in the follow-up measure. Interestingly, while analyzing teams of trained leaders we found a similar effect that was however delayed. Changes in attitudes towards elderly employees and age-diverse teamwork probably have to manifest in leaders' mindsets and behavior first. After consolidation of these changes in the leaders themselves, changes in the teams can occur. This was to be expected, as many previous studies have described the relative stability of attitudes and peoples' resistance to change in well-established attitudes (e.g., Wegge et al. 2010; van Dick et al. 2008). Moreover, only small effects for attitudinal change following (diversity) training have been found in a recent meta-analysis (e.g., Kalinoski et al. 2013).

Third, we also anticipated improvements in team functioning and effectiveness. As expected, results revealed that trained leaders report fewer conflicts in their teams after the training. If we interpret conflicts as an indicator for team functioning, leaders thus experience an improvement in teamwork after the training.

Fourth, we found that the training increased innovative behavior in teams, an effect that was stable until the follow-up measure, indicating continuously increased team performance over time. However, there was no significant main effect for the identification with the team. Taken together, it could be argued that the effects of the training pointed in the right direction but were rather small. Certainly, changing attitudes and stereotypes as well as influencing team members' views is a long and complicated process. Therefore, even small effects especially on subsequent performance indicators should be worth noting.

5 Discussion

In view of substantial demographic changes and an increasing age-diverse composition of work forces, we developed, implemented and evaluated training interventions at both the individual and the management level with the aim to improve employees' (a) cognitive abilities, (b) stress management competences, and help leaders (c) to manage age-diverse work groups more effectively. All interventions were tailored to workplace settings in that they were designed to be easy-to-use and motivating as well as time- and cost-efficient.

All three training interventions could be implemented successfully and resulted in beneficial effects at the individual and management level. The participating employees showed significant improvements in several cognitive functions (in particular working memory) and improvements in their stress management strategies. Theoretically, based on the 'servant leadership' concept, the leadership training also improved team functioning and innovative behaviors as reflected in several indicators at the team leader level as well as the group members level.

One known caveat of such interventions is, however, that training effects may be short-lived if the contents of the training interventions are not refreshed in subsequent booster sessions. The companies in which we implemented the interventions are already making an effort to counteract this and sustain training and transfer benefits for their employees, e.g. by making the cognitive training tasks and stress management exercises available via the company's intranet they are trying to encourage their continued use. Furthermore, the companies offer half-day booster sessions for the leadership training in order to stabilize the effects in the long-term.

References

Au J, Sheehan E, Tsai N, Duncan GJ, Buschkuehl M, Jaeggi SM (2014) Improving fluid intelligence with training on working memory: a meta-analysis. Psychon Bull Rev 22:366–377

Baltes EB, Baltes MM (1990) Psychological perspectives on successful aging: the model of selective optimization with compensation. In: Baltes EB, Baltes MM (eds) Successful aging: perspectives from the behavioral sciences. Cambridge University Press, New York, pp 1–34

Bass BM, Avolio BJ (1990) Transformational leadership development: manual for the multifactor leadership questionnaire. Consulting Psychologists Press, Palo Alto, CA

Bell ST (2007) Deep-level composition variables as predictors of team performance: a meta-analysis. J Appl Psychol 92:595–615

Berking M (2010) Training emotionaler Kompetenzen, 2 Auflage. Springer, Berlin

Bröker L, Rivkin W, Günnewig J (2015) Training von Führungskompetenzen zur Steigerung der Innovation und Produktivität in altersgemischten Teams. In: Langhoff T, Bornewasser M, Heidling E, Kriegesmann B, Falkenstein M (Hrsg) Innovationskompetenz im demografischen Wandel. Konzepte und Lösungen für die unternehmerische Praxis. Springer Gabler, Wiesbaden, S. 90–S. 109

Brown ME, Treviño LK (2006) Ethical leadership: a review and future directions. Leadersh Q 17: 595–616

Burns JM (1978) Leadership. Harper & Row, New York

Cassavaugh ND, Kramer AF (2009) Transfer of computer-based cognitive training to simulated driving in older adults. Appl Ergon 40:943–952

Ehrhart MG (2004) Leadership and procedural justice climate as antecedents of unit-level organizational citizenship behavior. Pers Psychol 57:61–94

Gajewski PD, Falkenstein M, Stahn C, Zülch J, Haas K, Rose U, Freude G (2012) Länger geistig fit im Beruf! Abschlussbericht PFIFF 2: Umsetzung des Programms PFIFF—Training zur Förderung von Gesundheit und geistiger Leistungsfähigkeit bei älteren Arbeitnehmern. INQA-Bericht, 43

Graham JW (1991) Servant-leadership in organizations: inspirational and moral. Leadersh Q 2: 105–119

Greenleaf RK (1977) Servant leadership: a journey into the nature of legitimate power and greatness. Paulist Press, New York

Hale JR, Fields DL (2007) Exploring servant leadership across cultures: a study of followers in Ghana and the USA. Leadership 3:397–417

Harrison DA, Price KH, Gavin JH, Florey AT (2002) Time, teams, and task performance: changing effects of surface-and deep-level diversity on group functioning. Acad Manage J 45:1029–1045

Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ (2008) Improving fluid intelligence with training on working memory. Proc Natl Acad Sci 105:6829–6833

Kalinoski ZT, Steele-Johnson D, Peyton EJ, Leas KA, Steinke J, Bowling NA (2013) A meta-analytic evaluation of diversity training outcomes. J Organ Behav 34:1076–1104

Kukolja J, Thiel CM, Wolf OT, Fink GR (2008) Increased cortisol levels in cognitively challenging situations are beneficial in young but not older subjects. Psychopharmacology 201: 293–304

Küper K, Karbach J (2015) Increased training complexity reduces the effectiveness of brief working memory training: evidence from short-term single and dual n-back training interventions. J Cogn Psychol 28:1–10

Liden RC, Wayne SJ, Zhao H, Henderson D (2008) Servant leadership: development of a multidimensional measure and multi-level assessment. Leadersh Q 19:161–177

Lilienthal L, Tamez E, Shelton JT, Myerson J, Hale S (2013) Dual n-back training increases the capacity of the focus of attention. Psychon Bull Rev 20:135–141

Lippmann E (ed) (2009) Intervision. Kollegiales Coaching professionell gestalten (Intervision. Professionally designing coaching among colleagues). Springer, Berlin

Martin M, Grünendahl M, Martin P (2001) Age differences in stress, social resources, and well-being in middle and older age. J Gerontol B Psychol Sci Soc Sci 56:214–222

McEwen BS (2007) Physiology and neurobiology of stress and adaptation: central role of the brain. Physiol Rev 87:873–904

Miyake A, Shah P (eds) (1999) Models of working memory: mechanisms of active maintenance and executive control. Cambridge University Press, New York

Ries BC, Diestel S, Wegge J, Schmidt K-H (2010) Die Rolle von Alterssalienz und Konflikten in Teams als Mediatoren der Beziehung zwischen Altersheterogenität und Gruppeneffektivität. Z Arbeits Organisationspsychol 54:117–130

Ries BC, Diestel S, Wegge J, Schmidt K-H (2012) Altersheterogenität und Gruppeneffektivität: Der Einfluss von Konflikten und Wertschätzung für Altersheterogenität. Z Arbeitswissenschaft 66:58–71

Rivkin W, Diestel S, Schmidt K-H (2014) The positive relationship between servant leadership and employees' psychological health: a multi-method approach. Z Pers 28:52–72

Salminen T, Strobach T, Schubert T (2012) On the impacts of working memory training on executive functioning. Front Hum Neurosci 6:1–14

Smith BN, Montagno RV, Kuzmenko TN (2004) Transformational and servant leadership: content and contextual comparisons. J Leadersh Org Stud 10:80–91

Stawski RS, Sliwinski MJ, Smyth JM (2006) Stress-related cognitive interference predicts cognitive function in old age. Psychol Aging 21:535–544

Uchino BN, Berg CA, Smith TW, Pearce G, Skinner M (2006) Age-related differences in ambulatory blood pressure during daily stress: evidence for greater blood pressure reactivity with age. Psychol Aging 21:231–239

van Dick R, van Knippenberg D, Hägele S, Guillaume YRF, Brodbeck FC (2008) Group diversity and group identification: the moderating role of diversity beliefs. Hum Relat 61:1463–1492

Walumbwa FO, Avolio BJ, Gardner WL, Wernsing TS, Peterson SJ (2008) Authentic leadership: development and validation of a theory-based measure? J Manag 34:89–126

Weber M (1921/1946) The sociology of charismatic authority. In HH Gerth, CW Milles (trans, eds) Max Weber: essays in sociology. Oxford University Press, New York, pp 245–252

Wegge J, Van Dick R, von Bernstoff C (2010) Emotional dissonance in call centre work. J Manag Psychol 25:596–619

Change into an Ergonomic 3 Shift Model with a Reduction of the Night Shifts for All Workers

Corinna Jaeger

Abstract

Night and shift work causes specific stress and strain. According to ergonomic recommendations the period of work and sleep which is against the circadian rhythm should be short. With the aim of taking preventive measures a company in the metal and electrical industry reduced the amount of night shifts for all workers concerned and changed the flexible shift system into a fast and forward rotating system; overtime is compensated by leisure time. Employees have influence on their working time. After implementation the employees, the works committee and management vote for the retention of the ergonomic shift system.

Keywords

Working time • Shift work • Night work • Stress • Strain • Ergonomics

1 Introduction

What is to be done with an ageing work force that can bear the strain of night work less and less? This question has been asked by a company of the metal and electrical industry with a number of 150 employees at its location in Germany.

To offer a permanent night shift on a voluntary basis and to offer separately a turning early and late shift was out of the question for this company, as the relief for one would cause an additional load for others. A permanent solution for the purpose of an aging appropriate shift schedule design had been wanted.

C. Jaeger (✉)
Institut für angewandte Arbeitswissenschaft e.V., ifaa, Düsseldorf, Germany
e-mail: c.jaeger@ifaa-mail.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_9

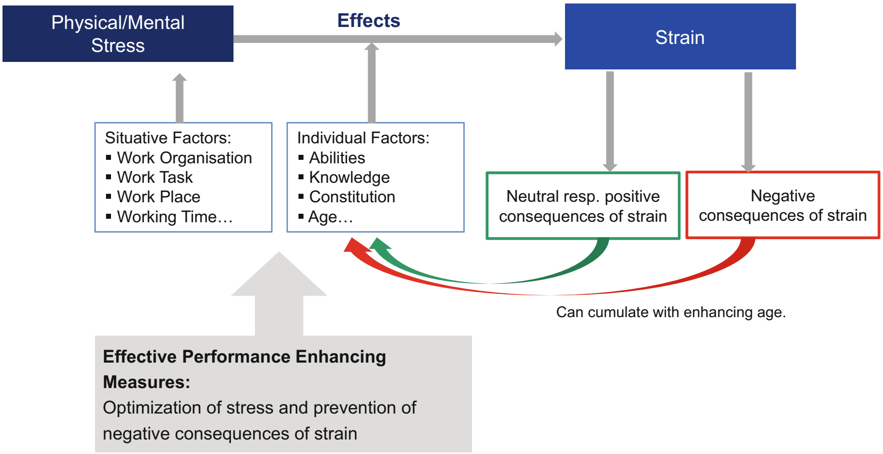


Fig. 1 Coherences of stress and strain. After Jaeger (2013: 18, 2015a: 32)

2 Ergonomic Background of Night and Shift Work

Shift work – and especially the night shift – signifies a special kind of stress and can lead to an enhanced strain (Fig. 1).

The human is inherently a diurnal being (Jaeger 2015c), at night a lot of bodily functions are reduced to a minimum in order to recover. During the day it's the opposite, the body is appointed on activity. This does apply as well. That is so for the so called early birds (the morning persons) and the night owl (the late risers), whose body clock is just postponed by a couple of hours. The bodily functions are controlled by a timing circuit such as light, heat, noise and having knowledge about the time of day (social timing circuit). Working night shift means, to perform at night against the natural sleep-wake-cycle (circadian rhythm) and having to sleep at day (Jaeger 2014a). That increases the risk of sleep deficiency, mistakes and accidents. Bodily efforts to adapt in the night shift or in the various positions of working hours are none more so than demanding. Negative consequences of the strain can gather over the course of the working life and compromise ones health. A common social impact of night and shift work is isolation, a limited family life and an increasing divorce rate.

2.1 Shift Work at an Advancing Age

The demographical change becomes apparent by an increasing average age of the population. Also the amount of middle and higher aged workers is rising. With an increasing age the night and shift work is more straining (Jaeger 2014b, 2015b). For example is the adaption to the working time getting harder. By tendency the rate for incapacitation for night shift is rising.

In 2013 a ruling of the federal court attracted a lot of attention. It had decided in the case of one nurse that she wasn't incapacitate in general if she could not work the night shift. She could demand on working during the day (BAG, Az. 10 AZR 637/13).

In order to occupy the supposedly necessary night shifts an advancing number of companies offer permanent night shift on a voluntary basis and a turning early and late shift.

In so doing it's only a temporary solution as the relief for one would cause an additional stress for the others. Particularly high additional fees for permanent night shift – in the automobile industry in Germany up to 38 % – offer an incentive that is probably at the expense of one's health.

2.2 Aging Appropriate Shift Schedule Design

Demographic change and an increasing number of night shift improper employees is calling for permanent solutions. An aging appropriate working time and shift schedule design begins with career entry and extends all over the working life (Jaeger 2014b, 2015b). It primarily focuses on prevention by optimising the stress in order to prohibit avoidable impairment. One activity is to design the night and shift work corresponding with ergonomic recommendations. Only secondary it is about relief. If there are health impairments, the stress is reduced in order to support its applicability. One of the most important steps is to reduce fielding in unfavourable positions—for example night shifts. But that is the theory.

3 Practical Implementation in a Metal and Electrical Industry Company

Let us take a look at the practice. Using the example of a metal and electrical industry company to show how steps towards an aging appropriate shift schedule design have been realised (Jaeger 2016a, b). The company produces roll bend parts for the automobile industry. Production line for the parts is carried out in series and small batch series. The German location for business has a number of 150 employees, 16 of them work in the mentioned shift system. The order situation of the automobile supplier with a standard of 35 h per week fluctuates as well cyclical as seasonal.

3.1 Initial Position

The former shift system was inflexible, expensive and from an ergonomic point of view unfavourable. In order to coping better with phases with a strong order situation, the company agreed to an average 37.5 h with the trade union IG-Metall, based upon a long backwards rotation 3 shift system with three shift groups. The

Week	1						
Group	Mo	Tue	Wed	Thu	Fr	Sa	Su
A	N	N	N	N	N	+ N	
B	S	S	S	S	S	+ S	
C	F	F	F	F	F	+ F	

N	night shift 8 hours gross
S	late shift 8 hours gross
F	early shift 8 hours gross
+ N	additional night shift
+ S	additional late shift
+ F	additional early shift

Fig. 2 Abstract of the former shift plan with additional shifts on Saturdays

Rota resumee	Evaluation			Incidents per week		
Criterion group A				1	2	3
Lenght of night shift blocks	1			6		
Lenght of late shift blocks		1			6	
Lenght of late and night shift blocks						
Length of working blocks			OK			
Length of working blocks with only 1 day in between	1			21		
Unfavourable shift sequenzes	1	2				
Free evenings per week	2			1	1	
Rest period after night shift			OK			
Weekly resumee	5	3	2			
Saturday and Sunday off			0%			
Saturday and Sunday off (excl. **)			0%			

Legend: critical, alarming, OK

Fig. 3 Ergonomical evaluation of the shift plan including additional shifts

early, late and night shift takes 8 h, including an unpaid break of 30 min. Operation hours are from Mondays, 6 a.m. to Saturdays, 6 a.m. In case of an increasing order situation, additional shifts have to be done (Fig. 2).

Weekly working hours would increase to 45 h net, which are 48 h gross. Employees felt that the long night shift blocks of six nights in a row were the most burdensome. Additional work was paid including the extra wages. Figure 3 summons the outcome of the ergonomic evaluation of the shift plans. It shows a number of critical (red) and alarming (yellow) aspects.

Six night shifts in a row ended on Sundays at 6 a.m. After that followed a backward rotation into the late shift on Mondays. As Fig. 4 illustrates the rest period in between the shifts is 32 h long. Thereby the minimum rest period of 35 h, referred to in § 11 clause 4 of the German Working Hours Act (ArbZG), is undercut. These are a combination of Sunday and Holiday resting periods of 24 h (§9 ArbZG) in addition to the daily resting period of at least 11 h (§5 ArbZG). At the long backwards rotating change from the sixth late shift at Saturdays to the early shift on Mondays the weekly rest period only amounts to 32 h as well.

In phases with a lowermost order situation all of the night shifts have been cancelled, which is very much for the purpose of an aging appropriate shift schedule design. Within the 3 weeks shift cycle resulted a whole free week. At the same time the early and late shift was extended to 10 h gross including an unpaid 45 min break (Fig. 5). This led to an uneven distribution of the lower weekly working hours of a average 30.5 h net. The employees felt that these two working blocks, with extended daily working time and a shortened rest period of 2 weeks in a row (each 45.75 h net respectively 50 h grass) are very burdensome. With increasing

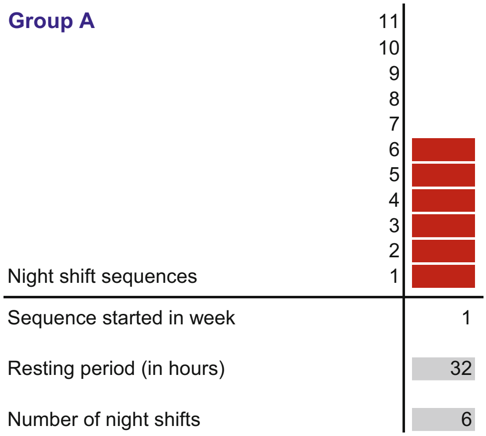


Fig. 4 Night shift sequence with a shortened weekly resting period

Week	1						
Group	Mo	Tue	Wed	Thu	Fr	Sa	Su
A							
B	S10	S10	S10	S10	S10		
C	F10	F10	F10	F10	F10		

S10	late shift 10 hours gross
F10	early shift 10 hours gross

Fig. 5 Abstract of the former shift plan reduced to 10 h early and late shift

age they needed more time to recover, whereby the part of usable leisure time was decreased.

Because additional work was paid out, a time credit was absent to compensate the deficit between actual time and estimated time in phases of low order position.

3.2 Aims and Measures

The enterprise was striving to react to variations in order quantities more reasonably and more cost-effective in the future and with that, optimize the strain on the aging employees. To achieve these goals, extra times were to be balanced by leisure time, an operational scope for flexibility upwards and downwards created, shift sequences ergonomically designed, the total of night shifts reduced for all persons involved, and employee’s requests are to be considered.

3.3 Development and Implementation of an Ergonomic Shift System

On the development of the flexible 3 shift system with reduction of the night shift for all employees, the management, the human resource manager, the production

Week	1							2							3						
Group	Mo	Tue	Wed	Thu	Fr	Sa	Su	Mo	Tue	Wed	Thu	Fr	Sa	Su	Mo	Tue	Wed	Thu	Fr	Sa	Su
A	N			F	F	F		S	SL	SL	N	N			F	FL	FL	S	S		N
B	S	SL	SL	N	N			F	FL	FL	S	S		N	N			F	F	F	
C	F	FL	FL	S	S		N	N			F	F	F		S	SL	SL	N	N		

F	early shift 8 hours gross	S	late shift 8 hours gross	N	night shift 8 hours gross
FL	early shift 9.5 hours gross	SL	late shift 9.5 hours gross		

Fig. 6 Illustration of the ergonomic 3 shift system with a reduction of the night shifts (basis plan)

manager, workers council and employees have been a part of. In order to adopt the expertise, employer and representatives of the employees together visited one of the author's seminars. Then they designed a flexible shift model. By the counsellors support it was reflected and overruled due to its lack of applicability.

The following illustrated alternative was created on the basis of the analysed requirements. Figure 6 shows the basis plan of the partly continuous shift system with 3 shift groups. The 37.5 h target weekly working hours of the former shift system, which had been agreed upon with the IG Metall, were reduced to 36.5 h. As the phases with a higher order situation predominate, the basis shift plan with an average of 39.5 h per week already is planed over the target hours. By the employees request the operation hours were accelerated to Sundays at 10 p.m. and the closure to Saturdays 2 p.m. Usually the early, late and night shift endures 8 h a time gross with 30 min unpaid break.

3.3.1 Reduction of the Night Shift for all Employees

An assessment of demands resulted in a cutback of the operating hours. To the purpose of prevention, the company used this result to find a sustainable solution for the night work. The uncommon consists in the disruption of the operation hours. Instead of leaving frame shifts aside, Tuesdays and Wednesdays operations from the night shift were transferred to the day shift and the early and late shift were extended to 9.5 h gross with 30 min unpaid break. Therefore, both night shifts were removed. That means a reduction of the night shifts from 6 down to 4 for all employees within the 3 week shift cycle. The night shift block beginning on Sundays is shortened from 4 to 2 night shifts.

3.3.2 Implementation of Ergonomic Recommendations

In terms of an aging appropriate organisation of working time (Jaeger 2014b, 2015b) with the ergonomic shift plan the company implemented further ergonomic recommendations (Beemann 2005).

Merely 2 night shifts in a row provide a short period of work and rest time against the biological sleep and awake rhythm. After both night shifts follows an 48 h working free time by request of the employees. Early, late and night shift rotate fast forward.

At the transition into a later starting shift type, the rest period is extended by 8 h, so that there is more time to recover within a working block. There are no unfavourable shift sequences. The length of each shift is adapted to the workload. Night shifts are terminated to 8 h. Only early and late shift are affected by the lengthening on 2 days. Phases with a superior high operation grade prevail, so that

Rota resumee	Evaluation			Incidents per week		
Criterion group A				1	2	3
Length of night shift blocks			OK			
Length of late shift blocks			OK			
Length of late and night shift blocks			OK			
Length of working blocks			OK			
Length of working blocks with only 1 day in between		1		8		
Unfavourable shift sequenzes		2				
Free evenings per week	1				2	
Rest period after night shift			OK			
Weekly resumee	1	3	4			
Saturday and Sunday off			33%			
Saturday and Sunday off (excl. **)			0%			

Legend:
- critical
- alarming
- OK

Fig. 7 Ergonomically evaluation of the ergonomic shift plan with a reduction of the night shifts

the basis shift plan is designed over target. Extra hours flow into working time accounts and are to be balanced by leisure time. This prevents an accumulation of working time. As a further step towards aging appropriate working time design, primarily by balancing through leisure time, the first thing that is cut off is a night shift. Same thing happens if the order situation lowers. So called frame shifts are cut off. First the night shift on Sundays, then the early shift on Saturdays. That increases the leisure time at the Weekends, saves the shift allowance and cares for a consistent distribution of the working hours. The shift plan is clearly laid out and repeats after 3 weeks. Employees have the opportunity to influence working times by taking leisure time from their working time account. Even to swap shifts is possible. Furthermore they can change into part time work by reducing their contractual operating volume.

As illustrated in Fig. 7 the ergonomic evaluation of the ergonomic shift system shows better conditions than the former shift system.

The positive effects with regard to the reduced number of night shifts and shortened night shift sequences are illustrated in Fig. 8. In the ergonomically created shift system follows a rest period of 48 h on each night shift block. And also the weekly rest period – even in the over target planned basis shift plan – 48 h at each weekend is therefore far over the time of 35 h required by law.

3.3.3 Pilot Project and Implementation

Employees and works council especially have been sceptical about the short forward rotating shift sequences. In order to evaluate the shift system developed in coorporation with the employees by concrete experiences, it has been tested for 6 months. The aging employees felt a real relief by working a smaller number of night shifts, short night shift blocks, ergonomic shift sequences and the compensation of extra hours by leisure time. As well as the consideration of requests from the employees, while developing the shift system and the influence on their own working time, cared for a high acceptance. Through the possibility to extend or shorten the operation time including management and the compensation of time credit or time deficit by working time accounts, the latest shift system covers all business related matters in a suitable and economic way. The pilot project resulted in a minor need for changes. After a corresponding modification, the flexible

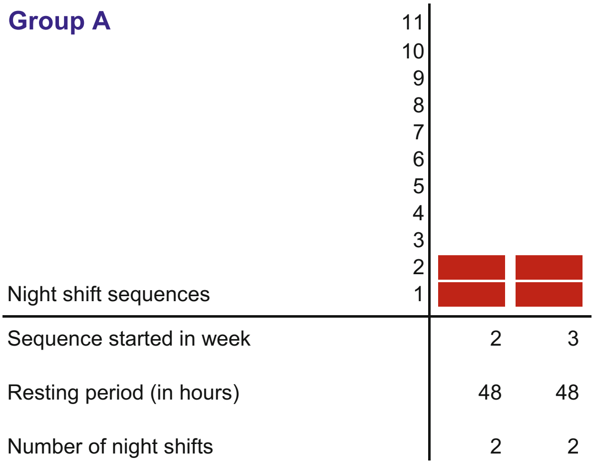


Fig. 8 Night shift sequences with an extended weekly resting period

ergonomic shift system with a reduction of night shifts was favoured by all involved parties and firmly assumed on the 01.01.2015.

4 Conclusion and Prospect

Frame conditions, which are favourable from an ergonomic point of view, as they help to optimise stress or even prevent negative consequences of strain, are often rejected or considered sceptical in operating practise. This especially happens to shortly forward rotating shift plans and the compensation of extra hours by leisure time instead of money. It's even more important to sensitise for the need of changes, to engage employees into the development process, to create compromises, to let working times—that has been changed—bring into effect during a pilot project and modify according to the responds, if necessary. In order to support the acceptance and distribution of ergonomic shift plans, the Insititut für angewandte Arbeitswissenschaft e. V. offers a 9 min movie on their website (arbeitswissenschaft.net) with the title "Schichtarbeit arbeitswissentschaftlich gestaltet—entlasted Mitarbeiter und stärkt Unternehmen". Companies, employees concerned, works council and scientific experts have word on the topic.

References

Beermann B, Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (Hrsg) (2005) Leitfaden zur Einführung und Gestaltung von Nacht—und Schichtarbeit. 9 unveränderte Auflage. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund

Jaeger C (2013) Leistungsfähig sein und bleiben—Hinweise für Unternehmen und Beschäftigte. Betriebspraxis & Arbeitsforschung 216:16–23

Jaeger C (2014a) Arbeitswissenschaftlich gestaltete Schichtsysteme kommen gut an. Betriebspraxis & Arbeitsforschung 222:36–41
Jaeger C (2014b) Arbeitszeiten alternsgerecht gestalten. Betriebspraxis & Arbeitsforschung 220: 32–37
Jaeger C (2015a) Leistungsfähig sein und bleiben. In: Institut für angewandte Arbeitswissenschaft (Hrsg) Leistungsfähigkeit im Betrieb. Kompendium für den Betriebspraktiker zur Bewältigung des demografischen Wandels. Springer, Berlin, pp 27–39
Jaeger C (2015b) Ergonomische Arbeitszeitgestaltung—Alternsgerechte Arbeitszeiten. In: Institut für angewandte Arbeitswissenschaft (Hrsg) Leistungsfähigkeit im Betrieb. Kompendium für den Betriebspraktiker zur Bewältigung des demografischen Wandels. Springer, Berlin, pp 167–175
Jaeger C (2015c) Gesundheitsförderliches Verhalten—Eigenverantwortung der Beschäftigten. In: Institut für angewandte Arbeitswissenschaft (Hrsg) Leistungsfähigkeit im Betrieb. Kompendium für den Betriebspraktiker zur Bewältigung des demografischen Wandels. Springer, Berlin, pp 393–399
Jaeger C (2016a) Nachtschichtreduktion als Maßnahme zur alternsgerechten Schichtplangestaltung—Ein Beispiel aus der betrieblichen Praxis. Betriebspraxis & Arbeitsforschung 227:19–22
Jaeger C (2016b) Umstellung auf ein ergonomisch gestaltetes 3-Schichtsystem mit Nachtschichtreduktion für alle Beschäftigten. In: Gesellschaft für Arbeitswissenschaft (GfA) (Hrsg) Arbeit in komplexen Systemen. Digital, vernetzt, human?! 62. Kongress der Gesellschaft für Arbeitswissenschaft vom 2–4 März 2016. GfA-Press, Dortmund, Beitrag C2.4, pp 1–4

Vocational Competency Management Through the Use of Serious Games: Opportunities and Risks Using the Example of the Harbour Industry

Daniela Ahrens and Sven Schulte

Abstract

The port economy oscillates between tradition and modernity: On one hand, the port continues to provide unskilled and inappropriately qualified people an entrance to dock work. On the other, the maritime industry reports high levels of specialisation and automation. This as well as the demographic change mean that questions of competence development and operational skills management are very timely. The empirical results of the research project "Work process-oriented competence development for the port of the future" prompt discussions on the innovative forms of occupational competency management based on the use of Serious Games.

Keywords

Harbour industry • Competency management • Serious games • Workbased learning

1 Introduction: Challenges in the Harbour Industry

Nearly 23 % of Germany's total freight turnover is handled in the ports of Bremerhaven and Bremen of the Free Hanseatic City of Bremen. A maritime forecast for the German seaports from the 2013's estimates an annual growth in freight traffic volume of 2.8 % until 2030 (cf. Freie und Hansestadt Hamburg Behörde für Wirtschaft, Verkehr und Innovation 2015, 26). These optimistic forecasts lead to a

D. Ahrens (✉)
University of Bremen, Bremen, Germany
e-mail: dahrens@uni-bremen.de

S. Schulte
Technical University Dortmund, Dortmund, Germany
e-mail: sven.schulte@tu-dortmund.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_10

demand for skilled workers, which is increasingly difficult to cover facing the demographic change and the increased demands on skilled workers in the maritime industry. Model calculations for the Bremen port industry expected in the age group of over-50s by 2020 an increase by around 13 %, while the share of the 35–50 year old people will be reduced by around 16 % (cf. Jürgenhake et al. 2006, 4). Among the industrial workers in the port logistics, the average age was in the year of 2006 on 41.6 years (Jürgenhake et al. 2006). The demographic change, the economic situation as well as rising labor requirements are the decisive parameters that provide the companies with the challenge of skills development. In addition, the recruitment of skilled workers in the German maritime industry is in competition with industries such as aerospace. These industries are currently more attractive to professionals than the maritime industry because they have a stronger presence and a more positive image (Freie und Hansestadt Hamburg Behörde für Wirtschaft, Verkehr und Innovation 2015).

Few studies or research have been conducted on the changing labour requirements at the ports. A comprehensive industrial sociological study for harbour work was published in the 1970s (Abendroth et al. 1979). This study deals with the increasing industrialisation and mechanisation of loading and unloading as well as the standardisation of progressive container.[1] The container handling increased not only the traffic volume, moreover through the large volume and normalised cargo units increasingly specialised handling agents (gantry cranes, straddle carriers) come to use. Abendroth et al. focus in their study the identification of requirement profiles for the profession "skilled workers port" given the fact that the majority of port workers (80 %) possessed a vocational qualification, which was however missing the occupational context, so that cost and time for intensive updating training were necessary.

A recent study on qualification in the port focused on the training of (long-term) unemployed, in order to offer them an employment perspective in the port (Mehlis et al. 2010). Regarding the increasing "professionalisation of dock work" (Mehlis et al. 2010, 9) the study aimed on the further training of unemployed people as a "Specialist in port logistics",[2] to give low-skilled persons sustainable employment perspective as well as to provide the port operators with qualified staff. The assumption was that the 2006 newly created apprenticeship "Specialist for Port Logistics" would replace the "port skilled workers" (3-month training in connection to at least 2 years of professional experience in the port) in the long term. The latest figures reveal that companies still rely on skilled port workers. Overall, only about 80 trainees for this apprenticeship in Germany were counted in the last 2 years. It seems likely that in view of the complexity of the work in the port, the companies have difficulties with a specified competence profile of a training occupation. They prefer to rely on on-the-job-teaching. Although the following definition dates back to 1979, it illustrates the complexity and variety of port work:

[1]In the mid-1960s the "Fairland" American shipping company Sea Land brought the first 99 containers to Germany. Since that time the container revolutionized shipping. About 95 % of world trade today is carried by container.

[2]In 2006, the occupation requiring formal training "Specialist for Port Logistics" was launched and replaced the previous tallyman system.

> Port work [. . .] and therefore the occupation of dockers [. . .] includes the practical port handling by means of any technique, designed by operationally organized division of labor and cooperation, and controlled by a certain entrepreneurial disposition. Among those set of tasks all kinds which refer, directly or indirectly, to the actual loading and unloading of vessels of all kinds, including the related loading and unloading of inland transport [. . .], as well as in connection with the port internally necessary transporting, registering, storing, testing and maintenance of the port, as the intersection of land and sea transport, the incoming and outgoing goods of any kind incurred (Abendroth et al. 1979, 18).

This definition is still relevant today, but the technologies and forms of work organisation have changed. Both the handling of high and heavy goods and container handling happens with a rising degree of mechanisation and automation—for example, the industrial trucks are equipped with crane machines, remote monitoring and automatic sway for the floating containers are also recorded. Although the physical strain has decreased, however experts record increasingly physical and mathematical skills, logical thinking, a quick mind, a sense of responsibility and working independently are required at the same time (Verein Hamburger Stauer von 1886 e.V. 2011). The changing nature of work in the port can not only be understood exclusively as a benefit of a stronger systematisation and a theory-driven action. The focus is rather on the relationship between knowledge and action. This work-related knowledge refers not only to systematic knowledge, but also to a kind of process knowledge. These two types of knowledge are highly interrelated. Addressed is a knowledge that refers not only to individual tasks to be performed, but to the context of individual tasks in an operational context (Fischer 2006). For example, raising the carrier structures from the ship must just happen in a straight line, the ropes have to stand in the right corner, and unequivocal view signs of the stevedore to the crane operator are essential. Bridge crane operators are responsible in a harbour for the container envelope between ship and quay. They take charge, send or load and unload ships and trucks with the assigned containers. In her mostly high mechanised job in the floating cabin under the gantry they have to deal with more than thousand-tons-arrangements precision work. The containers are placed from a distance from up to 40 m with an exactness of few centimetres. Furthermore, lightweight and extremely long blades have to be transported like a raw egg and expectant generator gondolas with high weights. The result is that just the experience and knowledge is not obsolete by automating and increasing demands on workers in terms of handling the technology. This and the growing process orientation operational workflows cause the requirements for taking over responsibility, interface capabilities, as well as social skills to rise in the maritime economy at all skill levels.

Following up on the methodological approach and the empirical results of the joint project "Work process-oriented competence development for the port of the future",[3] the first step is to outline the methodological approach for identifying the competency requirements for skilled port workers. The second step is to meet the

---

[3]The joint research project is funded by the Federal Ministry of Education and Research in the program "Vocational Competency Management and Demographic Change". It has a term from 1.12.2013 to 28.02.2017 (www.arkoh.de).

educational challenges for the implementation and the potential of Serious Games to enhance competence development. The focus lies in the didactic perspective. The aim of this paper is to show the challenges and perspectives of playful learning for skills development. The view is particularly directed to the question of how a combination of fun and learning is possible, and to identify the challenges of designing Serious Games as an innovative tool for vocational competence management in the port industry.

2 Identification of Required Skills

While educational research in the 1990s focused on competence-oriented approaches, some critical voices have taken the large number of (measuring) methods in the field of (cognitive) psychology to acquire skills to advance the beginning of a “cognitive revolution” (cf. Zlatkin-Troitschanskaia and Seidel 2011). At this point, we cannot discuss the complex debate about the concept of competence, especially as the search for a consistent theory to a concept of competence seems to be futile. Erpenbeck and Rosenstiel rather refer to the fact that the concept of competence is “theory relatively” (cf. Erpenbeck and von Rosenstiel 2003, XII). This makes it necessary to define the meaning frame of the concept of competence which was used in our research. We would like to take this opportunity to emphasise two key elements for the following underlying understanding of competence. First, the turn to an action-oriented concept of competence: The main difference to the qualification concept is that skills are “dispositions of self-organised action” (Erpenbeck and von Rosenstiel 2003). Competences become visible in concrete action while qualifications make existing knowledge skills visible as certified results. Qualifications manifest as “knowledge and skill positions” in education certificates while competencies as “self-organisation dispositions” are only accessible at the action level. Therefore, they can be acquired only in the context of the situation. Thereby, the concept of competence emphasized the level of action. Second, in a professional context an understanding of vocational skills has enforced as a unit of professional, social and human skills (Bader and Müller 2002). Though, methods skills, learning skills and communicative competence are understood as an integral part of this triad. This action-oriented concept of competence goes back to the work of Henry Roth (1971), who identified the concept of competence and made the connection between psychology and pedagogy. In his pedagogical anthropology, Roth linked the educational goal of maturity with skills development. Thereafter, maturity was understood as a competence in a triple sense: self-competence, physical and social (cf. Roth 1971, 180). For Roth, there is a mature, self-reliant act, first of a kind and expertise (intellectual maturity); secondly, a social insight and social skills (social responsibility); and, thirdly, the value of judgment and I-competence (self-determination and moral responsibility) (Roth 1971, 180). The triad of self, physical and social skills Roth brought into play is, with slight changes, fundamental to the competence debate in vocational education and training today.

To identify the competency requirements the study used in its empirical research social and vocational science methods (Gessler and Howe 2015): Concretely, a scenario analysis, work process analyses and semi-structured interviews with experts [4] of the port industry were applied. Additionally, the authors conducted a standardized survey among companies of the maritime sector. While the semi-structured interviews were aligned more exploratory and hypothesis generating to provide criteria for the work process analysis the company survey served the hypothesis validation. The work process analyses were conducted in the loading process of tower segments and support structures as well as in the container handling and stevedoring processes in the workspace of the offshore logistic.

The Scenario analysis focused the trends for the port by 2020. In four workshops with relevant stakeholders weighted contributing factors were identified and condensed into four scenarios (Duin and Thoben 2015). The scenarios are based on a matrix representing the dimensions "Human & Demography" and "Processes & Technology" [5]:

– Scenario 1 ("people first") describes that sufficient and well-qualified professionals are available, the organization of work processes and the technology used, however, remains at the present level. The focus in this scenario lies on the qualification and training level.
– Scenario 2 ("full speed ahead") indicates that the technology development and training of employees go along together, so that not only adequate and well qualified professionals, but also an innovative and effective infrastructure in the port industry is present.
– Scenario 3 ("stagnation") provides for both dimensions a pessimistic forecast. The company notices the effects of demographic change in the way that the average age of the staff increases and fewer young professionals can be recruited. The infrastructure does not develop further, so that the port is not a growth industry and therefore not an attractive workplace.
– Scenario 4 ("Smart Port") also describes a negative result of demographic change development regarding the availability and qualifications of professionals. In this scenario, however, the shortage of skilled labour is compensated by the economies of scale of digitisation.

The experts (scientists, entrepreneurs, social partners) participating in the workshops said the "Smart Port" was the scenario most likely to eventuate. It was also identified in the ArKoH research project as posing a major challenge, because it clearly demonstrates the danger of the rationalisation of jobs through digitalisation and automation.

[4] A total of nine expert interviews were conducted. These experts included representatives of training institutions for skilled port workers, a works council chairman and executives from companies in the port industry.

[5] In keeping with the socio-technical system approach, "processes" were understood to be work processes.

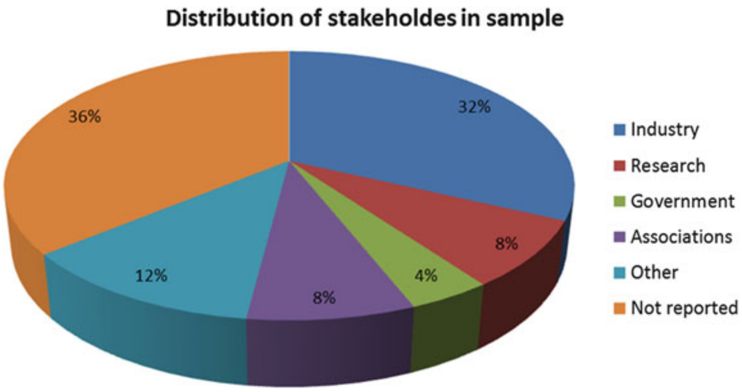


Fig. 1 Distribution stakeholders

Parallel to the scenario analysis a survey was conducted during the first half of the project. An online version and a pen-and-paper version were realised. Of the 54 representatives of companies, associations and social partners invited to participate, almost 50 % responded (Fig. 1).

The aim of the company survey was from the perspective of stakeholder assessments of the current status and the future development (a) the work processes, (b) to win the staff recruitment and training, (c) the technology and digital media, and (d) the "Port of the Future". The results show that these companies are already feeling the consequences and challenges of demographic change. The majority of respondents complained of difficulties both in recruiting skilled workers on the labour market and finding suitable trainees. With regard to how training is carried out and the degree to which work and learning are interrelated, the results show a rather conservative perspective. Still, classic seminaristic training opportunities were the dominant form at these companies. Two-thirds of respondents said that training was in the form of full-time gainful seminars held in-house. These related mainly to training by manufacturers (Fig. 2).

Media-supported learning in the work process or blended learning concepts are only used by nearly one-fifth of respondents. Comparing this survey result with the result of the scenario analysis delivers valuable notes for the ambivalent relationship between digitization and future forms of work. Company representatives primarily look at the rationalization potential of digitization. The potential of the new media is seen more from a business perspective in terms of effectiveness and efficiency and not as an opportunity for new learning processes and formats. The surveys on the one hand and the scenario analysis on the other were supplemented by structured interviews and observations of the working process. The workplace observations of the shipping of high and heavy goods were executed in half or whole days, and were documented on the basis of an observation guide. If the work process allowed, issues were resolved immediately with the respective port professionals. The results of workplace observations were weighted and validated in terms of their significance for different skill levels (professionals, master's level management personnel) at a workshop with participating company representatives.

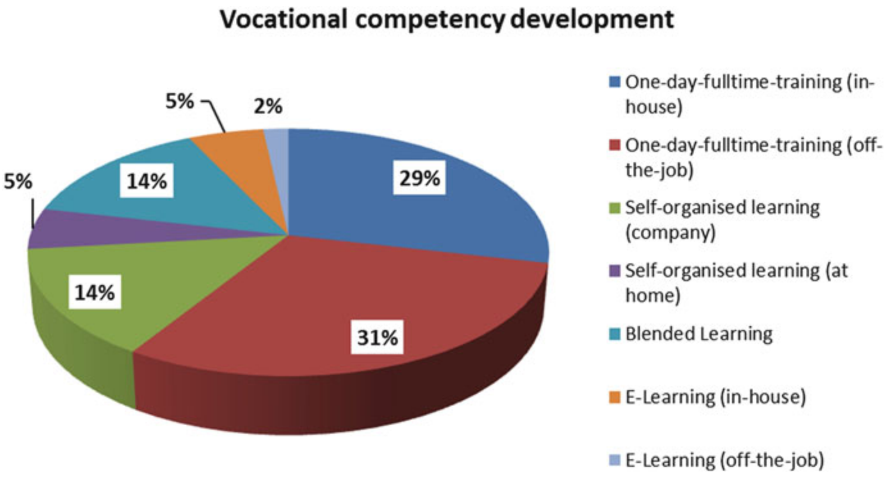


Fig. 2 Vocational competency

Triangulation of the empirical results generated the following seven learning modules for skills development:

– Maritime English,
– Selection of the appropriate abutment tableware,
– Truck selection,
– Transport safety rules,
– Knowledge of physics,
– Detecting defects in the sling, and
– Use of stop techniques.

Two aspects are emphasised in the learning modules: First, companies put a high priority on professionalism, along with the ability and willingness to deal properly with tasks and problems on the basis of technical knowledge and skills. These include the ability to recognise process contexts. This is mainly due to the fact that worker safety is of central importance in port areas. This can be seen in the methods devised to deal with the right sling to avoid accidents. For the companies, the choice and the use of the matching abutment dishes are of enormous importance, given the high risk of accidents. The second aspect concerns the training modules "knowledge of physics" and "maritime English". These learning modules underline the increasing demand for port professionals. Knowledge of, for example, lifting angle, loads and tensile forces will remain essential in the wake of increasing automation. In particular, this involves stevedoring processes and the loading of bulk goods.

3 Serious Games

The term "Serious Games" was coined by Clark C. Abt in the mid-1970s to emphasise the possible advantages of using play in learning processes:

> We are concerned with serious games in the sense that these games have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement (Abt 1975, 9).

Serious Games became popular tools for use in recruiting young adults for vocational positions. The game "America's Army" (created in 2002), for example, was the subject of much controversy. It featured a shooting game whose aim was to get adolescents excited about a military career. In the meantime, a broad range of areas of application have been found for Serious Games. The game content has been enlarged significantly and includes uses for health, the medical sector, sport as well as in playful learning of soft skills and the creation of work environments (e.g., office or factory buildings) to provide training in work processes or cooperation between skilled workers through the use of special simulations. In New York opened in 2009, the "Quest to Learn" school that aligns all their teaching and learning approach to the use of game mechanics.[6]

The increasing dissemination of mobile devices (tablets, netbooks, smartphones) is of growing importance to the education sector. Meanwhile, computer games are no longer only played by youths and young adults; they are now popular among age groups from 30 to 50 years, or even older (Goertz 2011). Although the concept of Serious Games has spread widely since the beginning of the twenty-first century, and has even spawned several established research institutes and journals (e.g., the International Journal of Serious Games), there is still a lot of scepticism about the extent to which "serious play" is possible at all. The strong connection between playing and learning is beyond doubt for children and youth, but to what extent can it be applied to work-process orientated competency development?

Presenting learning content in the form of a game in order to increase motivation and participation (especially among individuals with only a slight interest in the topic being learned) above the levels seen in seminaristic vocational training illustrates the main idea behind Serious Games. In comparison to traditional e-learning concepts, Serious Games underpin the idea of action-orientation and therefore involve intrinsically motivated learning activities (Unger et al. 2015). For the development of Serious Games it is necessary—unlike with normal games—to start with the learning objectives. The next step consists of shaping the application in an appropriately playful way while at the same time orientating the learning tasks and problems to suit the work situations of skilled harbour workers (De Witt 2012). An essential challenge for learning with Serious Games is posed precisely by the

[6]The school was founded by the American "Institute of Play" (http://www.q2l.org/).

fact that learning situations are shaped in a playful way, and not like typical learning situations (Kerres et al. 2009).

In light of the manifold fields of application and the different theoretical approaches—especially media theories, game theory and research of expertise—there is no standardised definition of Serious Games. While game-based learning favours the learning processes of the users, the term Serious Games figures out the gaming aspect itself (De Witt 2012). For the purposes of the present argument we use the term "serious games". Serious Games have been defined as "digital learning games that are used in vocational education and training to support traditional learning by using technical and multimedia elements" (cf. Feist and Franken-Wendelstorf 2011, 69). So far, playing, learning and working have been seen as unrelated activities, but now they are seen as useful in combination. Moreover, Serious Games enable learning processes that are usually not feasible for security or economic reasons (e.g., learning processes in real work situations in the fields of aviation and space flight):

> Game-based learning enables learners to undertake tasks and experience situations which could otherwise be impossible and/or undesirable for cost, time, logistical and safety reasons. Game-based learning is truly interactive, everything that the learner does, or does not, has an effect and is thus highly experimental (Corti 2006, 1).

In shaping the didactical concept that underpins Serious Games, three different approaches can be discerned (Kerres et al. 2009): First, the approach of immediate transfer of knowledge, which is gained directly throughout the game. In this case the knowledge is mainly declarative knowledge. This kind of knowledge is exemplified by simulations. Simulations—models that are as realistic as possible—aim to create learning processes that enable the learner to use the new knowledge in real (work) situations. The issue of didactic moves back in favour for the benefit of a very detailed and realistic simulation (cf. Kerres et al. 2009, 1). Furthermore, there are varying opinions on whether and to what extent simulations can be seen as games at all. Critics object on this point, referring to the requirements of simulating reality. Primarily, a simulation does not contain playful elements, but represents training and instruction material. There is an opposing view that insists that using a simulator has elements of play (e.g., launching or landing a plane via a simulator can be a very playful experience, particularly for ambitious players). Nevertheless, simulations are often perceived as an immersive experience. The authenticity of the reality makes the participants forget that this is a simulation.

The second approach emphasises the immersive experience in digital worlds. The development of role play or fantasy worlds (e.g., as an adventure or in a dangerous setting) is always to the fore. This is the opposite of a simulation, in that the strangeness is part of the concept. One criticism of this kind of playful learning is that a transfer of knowledge is nearly impossible. However, supporters of this approach point to its use of procedural knowledge (e.g., decision-making under time pressure as well as coordinating activities with teammates). Every mentioned element is clearly needed outside the gaming world. John Hagel and

John Seely Brown go so far as to recommend the MMORPG World of Warcraft for management employees in order to give them competencies in handling uncertainties and encourage creative thinking (Hagel and Brown 2009): “They [players] learn to welcome collaboration as an opportunity to learn faster by focussing on a set of individual strengths while being exposed to the diverse perspectives and experiences of those with complementary strengths” (cf. Hagel and Brown 2009, 3).

Last but not least, the third approach represents a special didactic challenge, in that learning tasks have to be prepared in a playful way. This approach has been conducive to the development of learning assignments for skilled harbour workers. An essential challenge lies in matching the game design with the workplace environment. Furthermore, recent studies have demonstrated the challenging and entertaining nature of medium-difficult tasks. The tasks may neither overstrain the player nor should the tasks be easily to handle. Thus, they have a major impact on motivation (Kerres et al. 2009). In order to create a balance between enjoyment and learning, respectively between pedagogical and game design, the development of the ArKoH-game adhered to the following criteria: aesthetically appealing illustrations, a good balance of challenge and reward, varied tasks, variable opportunities for interaction, as well as comprehensible context information and narration (cf. Breuer 2010, 28).

At the beginning of the game there is a brief self-assessment regarding one’s competency level. Depending on the level—from beginner to expert—and on the current learning module, there are specific tasks and questions that have to be solved by the player. The level of difficulty varies (e.g., in the number of possible answers to multiple-choice questions, or a time limit being imposed on a task). In this way, and also through the inclusion of an element of challenge via a quiz-duel or a high score ranking, a playful learning atmosphere is fostered (Fig. 3).

The learning modules have been implemented into a “storyline” for the ArKoH-Game.[7] This game enables problem- or task-orientated learning. While converting real work tasks into learning tasks falls into the task-orientated learning category, the problem-orientated approach focusses on the process itself (De Witt 2012). The realisation of the game design integrated several ideas on how to support the learning process: Tutorials explain the basics of the game; a non-player character (NPC) accompanies the player and provides explanations on demand, and asks questions or requests that the player explain a certain process (e.g., in order to introduce a new colleague to a work process).

In another example, in shaping the coordination of a shipping process in a playful way, the player has to start a dialogue with a new colleague. As part of the communication process the player has to explain to his new colleague the special conditions of a high & heavy shipping process. After the dialogue, questions have to be answered. By answering the questions correctly, the player reaches a higher level of the game. Wrong answers prompt text-based feedback that leads to

[7]The realisation of the “ArKoH-Game” happens on a windows-tablet, 10,1 inches, touchscreen.

Fig. 3 Starting screen of the ArKoH-game

further information or illustrations, as well as explanations of the correct response. Thus, the player does not have to leave the game. Ideally, the feedback or the follow-up information brings about the beginning of a reflection process within the player. A quiz with a linear question-and-answer format can be seen as a setting that enables knowledge learning. Since each learning module works on its own, these kinds of learning-on-demand approaches are very well suited to the idea of micro-learning and, accordingly, mobile learning (Robes 2009; Sauter and Sauter 2013). Micro-learning is not without controversy. Critics fear that learning is subordinated to economic strategies. This fear is justified. Nevertheless, successful learning processes usually take place only when own actions encounters resistance. Short learning units respond to a current problem and contain all the necessary information in order to understand them.

4 Summary and Outlook

In vocational education and training a culture of indoctrination still persists (e.g., frontal teaching in traditional seminar forms). The research results of the ArKoH project partly confirm this. In the majority the companies we surveyed are sticking to traditional forms of training. At the same time, the advancing digitisation of our working lives has led inevitably to an argument about new/digital media. This affects the discussion of new forms of vocational training, which are also supported by digital media. Through the dissemination of mobile devices (smartphones, tablets, netbooks), learning “on the way” has become more and more attractive and can easily be adapted to work processes. The younger generation of "digital natives" ("Net Generation"), for whom new media are a part of everyday life and even of working life, are not the only ones to benefit from this situation. Furthermore, Serious Games are being used to instruct people with low motivation for vocational training, as well as temporary workers. In the harbour business there are two more target groups coming along: facing the fact that cargo-handling systems and related work processes are changing permanently, Serious Games support the upgrading of skills. Through Serious Games knowledge can be refreshed and forgotten information and rules are brought into memory again. Secondly, not only skilled harbour workers, but also employees in marketing and distribution departments have been mentioned as an important target group. These employees possess trading and marketing knowledge, but they usually lack of know-how about the operative handling of cargo. In conversations with customers, or in manpower planning, knowledge of operational harbour processes (e.g., using a spreader or ground conveyer), could be a great advantage. In the present study, the surveyed companies identified three possible applications for Serious Games in the context of internal competency development:

- Use of Serious Games for in-house training: From the perspective of companies, Serious Games offer an appropriate way of developing competencies. An advantage is that such games can be played "in the meantime" during breaks and thereby can easily be integrated into the activities of fulfilling work processes. This application is particularly directed at employees that could not be motivated to take part in previous vocational education or training. That point confirms the assumption that employees are more motivated to play a game than take part in a seminar. During the evaluation of the use of the ArKoH game, another substantial difference arose between two target groups. While white-collar workers could use a Serious Game in a self-organised way in their offices, this was not possible to the same extent for skilled workers (blue-collar) because of work safety regulations in the harbour yards. Many companies have forbidden the use of mobile phones on harbour grounds. They worry that the use of Serious Games will cause attention span and concentration levels to decrease, while at the same time increasing the risk of a work accident.
- Serious games can be used in recruiting new employees to gain information about what training contents are necessary and how much time for training is needed. The companies that took part in the evaluation are discussing the extent to which serious games can be used as an innovative method for the identification of competencies. This approach does not constitute a complex and statistically determined measurement of competencies, but represents a pragmatic and effective way of identifying process- and work context-related skills.
- Self-organised learning can be promoted via Serious Games. Playful learning is learning connected with positive emotions. That point, as well as the challenging character of the games themselves (e.g., in the form of quiz duels or high-score rankings), motivates employees to embark on self-organised learning.

On the part of the companies in the harbour sector there is—one the one hand—a great interest in the use of this learning approach. On the other hand, the companies maintain a degree of scepticism based on the assumed contradiction between play and work. Since that perceived contradiction is reinforced in our industrial society, the question arises: Do we need to reformulate the relationship between work and play? Faced with the ongoing digitisation of society, it does not seem promising to categorise computer games per se as dangerous, or as being in conflict with the educational perspective. It has been 40 years since the observation was made that: "Games are not a panacea for all of the ills of the educational system today, but they do provide fast and effective relief for some of these ills" (vgl. Abt 1975, 28).

Faced with the increasing digitisation of the working world, and especially in light of the increasing importance of individual, informal and collaborative learning, greater priority must be placed on the educational potential of computer games. The research and studies conducted so far indicate that, along with mobile learning and social learning, game-based learning is getting more and more relevant. The extent to which Serious Games can serve as a tool for internal competency development is significantly influenced by the equilibration of work process-relevant learning content on the one hand and playful elements on the other. In shaping and implementing work process-orientated learning content, attention must be paid to the narration of work context and variable forms of interaction.

References

Abendroth M et al (1979) Hafenarbeit. Eine industriesoziologische Untersuchung der Arbeit und Betriebsverhältnisse in den bremischen Häfen. Campus, Frankfurt

Abt CC (1975) Serious games. Viking Compass, New York

Bader R, Müller M (2002) Leitziel der Berufsbildung: Handlungskompetenz. Die berufsbildende Schule 6:176–182

Breuer J (2010) Spielend lernen? Eine Bestandsaufnahme zum (digital) game-based learning. Landesanstalt für Medien Nordrhein-Westfalen, Band 41. Düsseldorf

Corti K (2006) Games-based learning; a serious business application. In: Pixel learning, game-based business & management skills development. https://www.cs.auckland.ac.nz/courses/compsci777s2c/lectures/Ian/serious%20games%20business%20applications.pdf. Accessed 22 June 2016

De Witt C (2012) Neue Lernformen für die berufliche Bildung: mobile Learning—social learning—game based learning. Berufsbildung in Wissenschaft und Praxis (BWP) 41(3):6–9

Duin H, Thoben KD (2015) Der Hafen der Zukunft. Szenarien für die Ermittlung von zukünftigen arbeitsprozessorientierten Kompetenzprofilen im Bereich Offshore und Containerumschlag. In: Gausemeier (Hrsg) 11. Symposium für Vorausschau und Technologieplanung, Band 347 der Verlagsschriftenreihe des Heinz Nixdorf Instituts, Universität Paderborn, pp 81–98

Erpenbeck J, von Rosenstiel L (2003) Handbuch Kompetenzmessung. Schaeffer-Poeschel Verlag, Stuttgart

Feist M, Franken-Wendelstorf R (2011) Informelles Lernen und der Einsatz von Serious Games. In: Metz M, Theis F (eds) Digitale Lernwelt—serious games. Einsatz in der beruflichen Weiterbildung. Bertelsmann Verlag, Bielefeld, pp 69–76

Fischer M (2006) Arbeitsprozesswissen. In: Rauner F (ed) Handbuch Berufsbildungsforschung. Bertelsmann, Bielefeld, pp 308–316

Freie und Hansestadt Hamburg Behörde für Wirtschaft, Verkehr und Innovation (2015) MCN. Strategie 2020: Wettbewerbsfähigkeit durch Innovation und Kooperation, Hamburg

Gessler M, Howe F (2015) From the reality of work to grounded work-based learning in German vocational education and training: background, concept and tools. Int J Res Vocat Educ Train (IJRVET) 2(3)

Goertz L (2011) Einsatzmöglichkeiten für Serious Games in Unternehmen. Personalführung 2:58–65

Hagel J, Brown JS (2009) How world of warcraft promotes innovation. Businessweek: innovation, 14 Januar 2009. http://www.businessweek.com/innovate/content/jan2009/id20090114_362962_page_2.htm. Accessed 23 June 2016

Jürgenhake U et al (2006) Demografischer Wandel in der bremischen Hafen—und Distributionswirtschaft. PortWork 05/15. Bremen

Kerres M et al (2009) Didaktische Konzeption von Serious Games: Zur Verknüpfung von Spiel—und Lernangeboten. In: MedienPädagogik. Zeitschrift für Theorie und Praxis der Medienbildung, pp 1–16

Mehlis P et al (2010) Nachhaltige Beschäftigungsperspektiven für Arbeitslose und Geringqualifizierte durch berufliche Weiterbildung. Erste Befunde aus einem Forschungsprojekt zur Qualifizierungsoffensive in der Hafenwirtschaft. In: bwp@ Berufs—und Wirtschaftspädagogik—online, Ausgabe 19, pp 1–20. http://www.bwpat.de/ausgabe19/mehlis_etal_bwpat19.pdf. Accessed 23 June 2016

Robes J (2009) Microlearning und Microtraining: Flexible Kurzformate in der Weiterbildung. In: Hohenstein A, Wilbers K (ed) Handbuch E-learning: Expertenwissen aus Wissenschaft und Praxis—Strategien, Instrumente, Fallstudien, Loseblattwerke, chap 4.36. Deutscher Wirtschaftsdienst, Wolters Kluwer Deutschland, Köln

Roth H (1971) Pädagogische Anthropologie. Band II: Entwicklung und Erziehung. Schroedel, Hannover

Sauter W, Sauter S (2013) Workplace learning. Springer Gabler, Wiesbaden

Unger T et al (2015) Digitale Serious Games. In: Blötz U (ed) Planspiele und Serious Games der beruflichen Bildung. Bertelsmann Verlag, Bielefeld, pp 157–180

Verein Hamburger Stauer von 1886 e.V. (2011) VHST 1886–2011, Hamburg. www.uvhh.de

Zlatkin-Troitschanskaia O, Seidel J (2011) Kompetenz und Erfassung—das neue “Theorie-Empirie-Problem” der empirischen Bildungsforschung. In: Zlatkin-Troitschanskaia O (ed) Stationen empirischer Bildungsforschung. VS Verlag, Wiesbaden, pp 218–233

Retirement Research Requires a Broad View and Interdiciplinarity

Results of the JPI UEP Fast Track Activity

Hans Martin Hasselhorn, Melanie Ebener, and Wenke Apt

Abstract

The *lidA conceptual framework on work, age and employment* is based on the observation that retirement is complex and therefore requires broad views and broad scientific approaches. Based on the framework, the JPI UEP project has identified three research gaps with respect to retirement research: a conceptual gap, a regional gap and a thematic gap. To understand the process of retirement in its complexity, and to provide organizational and national policies with valid evidence, research is needed that contributes to bridging these gaps. Ideally, such research should combine different methodological approaches, allow for a life course view and should make use of the potential of cross-cultural comparisons. Not least, it should be of interdisciplinary nature and bring different research communities together. *Human Engineering* as explicitly *integrated science* can not only contribute with own expertise in the field, but might adopt an integrating and moderating role here.

Keywords

Older workers • Retirement • Research needs • Human engineering

H.M. Hasselhorn (✉) • M. Ebener
Department of Occupational Health Science, University of Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany
e-mail: hasselhorn@uni-wuppertal.de; ebener@uni-wuppertal.de

W. Apt
VDI/VDE Innovation + Technik GmbH, Steinplatz 1, 10623 Berlin, Germany

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_11

1 Introduction

Since one and a half decade the number of older workers is increasing in Germany, the employment rate among the age group 55 to 64 years has increased from the year 2000 to 2014 by 25 % points to approximately 71 % among men and by 31 % points to 60 % among women (Eurostat 2015). This trend will continue in the coming years. This, however, can be concluded from projections of demographic developments in Germany which indicate that the employment potential (number of people available to the labor market) has now reached its maximum and will from now on decrease substantially. The size of this development is illustrated by estimations calculated by Fuchs et al. (2015): Assuming an annual net in-migration of 200,000 people (which equals the average net in-migration of the past 60 years in Germany) and further increasing employment participation rates among older workers they calculate a decrease of the employment potential from 45.1 Million in the year 2012 to 40.8 Million in 2035. To keep the employment potential constant until 2050 a mean net in-migration of annually 533,000 people would have to occur (Fuchs et al. 2015) which the authors consider to be unrealistic to happen. Accompanying the expected downward trend, estimates show that the German economy will continue to need about 40 million workers for many years to come.

The German Federal Government regards the decrease of the national employment potential as a threat for economic growth and innovation. In its annual economic report 2015 (Jahreswirtschaftsbericht) the Federal Ministry for Economic Affairs and Energy states that the employment potential needs to be systematically increased: by reducing early school and college dropout, by improving occupational qualification and by an increased employment participation of women, older people, migrants as well as disabled people (BMWI 2015).

If, perspectively, more and more people at higher working age have to remain in employment, the question is not only whether future generations of older workers HAVE TO DO this—but also whether they WANT and CAN do this and which social, societal and organizational framework conditions need to accompany and support this development.

The topic of employment participation of people at higher working age has attracted researchers since many decades, for example in psychology since the 1970s (Beehr 1986). In the past years, scientific attention of this topic is increasing and disciplines such as gerontology, psychology, sociology, economics, social epidemiology and occupational health are dealing with it. Yet, in their research, the disciplines tend to focus on single factors that may influence employment participation at higher working age, resulting in two-dimensional analyses such as "work factors and retirement", "financial factors and retirement" or „health and retirement". One example: Van den Heuvel and de Wind (2015) document the multitude of studies confirming the association of poor health with early exit from employment. This, however, is not consistent with data (collected in 2008/9) published by Hasselhorn and Müller (2014) indicating that in Germany about 2.7 million working people in the age group 51 to 64 years report *moderate to poor health* while about three million people of the same age group did not/no longer work and reported *good or excellent health*. Obviously, when it comes to early exit from work, more factors besides *health* must be playing decisive roles, such as individual *living conditions*, *work ability* or the *labour market*. This implies, that focusing organizational and also socio-political measures on the promotion of the health of older workers might remain insufficient if the aim is to increase employment participation of older workers. In other words: whether or not people at higher working age are still employed is usually a function of a complex interaction of influential factors.

What is often lacking in research is the adequate consideration of the multitude of further relevant determinants and their inter-relatedness (Hasselhorn et al. 2015). This limited view has raised concerns about the validity of available research findings and their relevance for policy at the organizational, industry and national level (Szinovacz 2013; Hasselhorn et al. 2015; Phillipson and Smith 2005). As a consequence, there have been increased calls for the establishment of broader views and interdisciplinary approaches in retirement research (Szinovacz 2013; Wang and Shultz 2010; Hasselhorn et al. 2015; Phillipson and Smith 2005). While only partially reflected in empirical research, the complexity of the retirement process is acknowledged in theoretical considerations within retirement research. Brown and Vickerstaff (2011) underline the "causal complexity" of the macro level or meso level in sociocultural, economic, and political factors that may interact with each other and manifest themselves within specific household structures and personal environments. The emphasis of complexity as a core feature of employment participation at higher working age is today mainly being brought forward in the course of qualitative studies (for example Brown and Vickerstaff 2011; Higgs et al. 2003; Pond et al. 2010; Loretto and Vickerstaff 2012), some quantitative studies (de Wind et al. 2014; de Wind et al. 2015) and by overview publications (Szinovacz 2013; Hasselhorn and Apt 2015; Phillipson and Smith 2005; Ekerdt 2010).

In this contribution the main findings of the JPI UEP project "Joint Programming Initiative—Understanding Employment Participation of Older Workers" (2014 to 2015) will be presented. Forty-six scientists from eleven countries representing different scientific disciplines and research communities reviewed research and outlined research needs in the field of employment participation at higher working age. More about the project and its several national and domain reports may be downloaded here: www.jp-demographic.eu/activities/exploration/fast-track-projects. JPI UEP was a project initiated in 2014 by the Joint Programming Initiative "More years, better lives—The potential and challenges of demographic change" (JPI MYBL). JPI MYBL was established in 2010 with support of the European Commission to address the major societal challenge of demographic change by coordinated research funding.

2 The JPI-UEP Project

2.1 JPI UEP: Conceptual Basis: The lidA Conceptual Framework on Work, Age and Employment

Conceptual basis of JPI UEP was the *lidA conceptual framework on work, age and employment* (Fig. 1). lidA stands for the lidA Cohort Study "lidA—leben in der Arbeit. German cohort study on work, age and health" (www.lida-studie.de, Hasselhorn et al. 2014), a longitudinal study assessing work, age, health and employment in Germany from an interdisciplinary view (occupational, social and labour market sciences). The underlying *lidA conceptual framework on work, age and employment* has been developed under the assumption that a broad view is key to gain a deeper understanding of employment participation at higher working age and to promote interdisciplinary research in the field.

The lidA framework combines eleven so called "domains" that cover core determinants known to influence retirement timing. The *domestic domain*, for example, contains factors such as marital status, partners' employment, family formation, caring obligations, informal task and role distributions, synchronization of retirement, and household income. This domain is linked to the domain *social position* (see arrow in Fig. 1). As an example, families of low social status may differ substantially from families of high social status with respect to their household income and/or their distributions of informal tasks and roles. Such diversity may influence decisions about employment participation at higher working age. According to the conceptual framework, the *domestic domain* also directly impacts the domain *motivation to continue working* by enhancing or weakening older workers' propensity to work through circumstances at home.

The *motivation* domain, in turn, is influenced by the individual's *work*, *health*, *work ability*, *finances* and the *pension system*, all indicating groupings of decision factors in the lidA model (Fig. 1). *Work* covers two domains in the framework, one

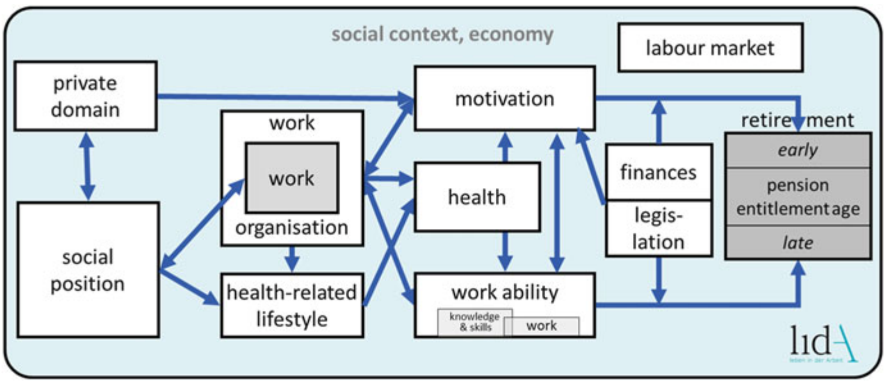


Fig. 1 The "lidA conceptual framework on work, age and employment" (Hasselhorn et al. 2015)

comprising *work organisation* (e. g. human resource management, HRM) and the other *work factors*, i.e. work content and social work environment.

While the content of the domain *health* may be obvious to the reader, its role with respect to employment participation of workers at higher working age needs to be looked at in a differentiated view as indicated earlier. *Work ability* refers to the Finnish work ability concept (Ilmarinen and Tuomi 2004) which basically defines work ability as whether someone is able to get a specific job done. The arrows in the framework indicate that this capacity is related to the work itself as well as the worker's knowledge and skills, health status, and also motivation to work. According to this concept, a worker may exhibit a high level of *work ability* despite a relatively low *health* status; what matters is whether the job fits the worker's functional, cognitive and motivational capabilities.

In the model, not only work, health, work ability and the domestic background influence the motivation to remain in employment or retire, also *financial factors* and *legislation* exert an important influence. To give an example: While in the 1990s, in many European countries, financially attractive regulations have led to the early retirement of large cohorts of older workers, today, current legislative and financial regulations may prevent the early exit even among those older workers who are highly motivated to retire early.

Four Characteristics of Employment Participation

The lida conceptual framework also illustrates four *characteristics of employment participation* at higher working age, namely that retirement is:

– complex,
– a process,
– influenced by factors on the individual level,
– affected by factors on the structural level.

Retirement Complexity

The multitude of domains in the framework and their interrelations elucidate the causal complexity of employment participation at higher working age. Above all, these interactions are occurring within a social context (Fig. 1): older workers are not making their employment / retirement considerations and decisions as stand-alone individuals, instead, they are part of a system of social relations influencing such reflections and decisions. This retirement complexity was the driver for the development of the *lidA conceptual framework* (see above).

Process of Retirement

The *lidA conceptual framework* also indicates that retirement is not a discrete event but a process that develops over a period of time.

Retirement as a *process* may either be seen as the gradual exit from a career job to complete withdrawal from the labour market. This process may be described by the term "fragmentation of retirement" indicating increasing temporal and conceptual heterogeneity in retirement, a development triggered by socioeconomic trends and policy shifts (see Hasselhorn and Apt 2015).

Several psychologists conceptualize retirement as a process referring to the process of retirement decision-making. Beehr (1986) differentiates between *thinking about retirement* (developing a preference to retire), expressing an *intention to retire* (decision), and *implementing the decision* by retiring. These concepts are covered by the *motivation domain* of the *lidA conceptual framework*.

Finally, a life course perspective of employment participation may provide another understanding of the process character of retirement. Many—if not most—determinants of early exit from work, such as *health*, *work ability* and *motivation to work*, may be seen as the result of exposures earlier in life. The *social position* domain in the *lidA conceptual framework* points to early in life factors linked to the social origin of an individual as potentially influential for employment participation at higher working age. Naegele and Sporket (2010) are following the life course approach when defining three risk factors for employment at higher working age: health risks, qualification risks and motivation risks, while emphasizing that these risks are late manifestations of early exposures in life.

This "temporal complexity" challenges the traditional connotation of retirement among policymakers, prevention advisors and—not least researchers (Hasselhorn and Apt 2015).

Individual Level

The "lidA framework" indicates that retirement is to high degree determined by factors inherent in each individual worker. Many of the factors influencing the retirement process are located within the individual and his/her personal life. This individual level view of employment at higher working age is reflected by several domains of the *lidA conceptual framework*, for example *health*, *motivation*, *domestic domain* and *financial factors*. A conclusion is that retirement is never the same for two people and that research should not neglect such individual micro level factors.

Structural Level

Finally, employment participation at higher working age is embedded in a strong structural context mainly represented by *financial factors*, *legislation* and the *labor market*. Legislative and regulatory factors, not least retirement schemes and labor market policies, effectively shape retirement behavior. Retirement researchers have to relate their findings to the relevant policy context to detect their impact and to consider their multiple effects.

3 JPI UEP: Three Research Gaps

The *lidA conceptual framework on work, age and employment* formed the structural and contextual frame of the JPI UEP project. Following its structure, experts from the JPI UEP working group collected research evidence and identified research needs across the ten *domains*, explicitly referring to the ten participating countries (Austria, Belgium, Canada, Denmark, Finland, Germany, Netherlands, Norway,

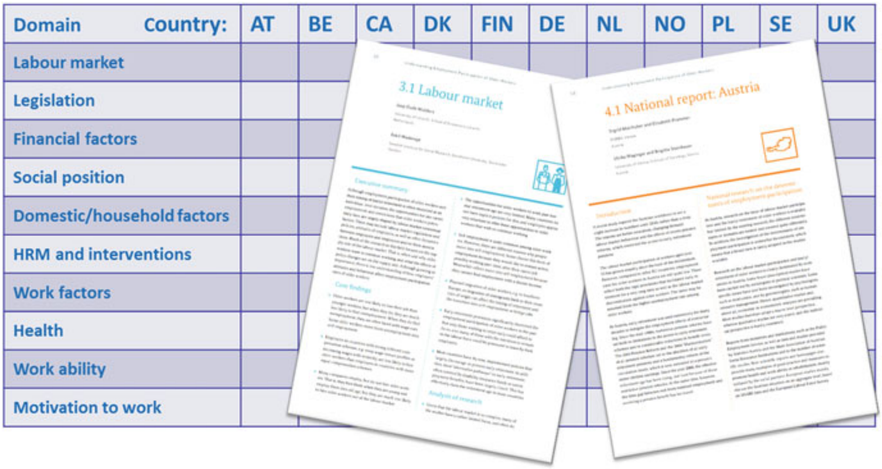


Fig. 2 JPI UEP matrix approach: ten domain chapters and 11 country reports

Poland, Sweden, United Kingdom). The ten domain chapters were complemented by eleven national reports explicitly referring to the national state of research for each of the domains and concluding with specific research needs on national level (Fig. 2).

Rather often, similar conclusions with respect to research needs were drawn across domain chapters and country reports. In summary, the research needs fall into three priority areas to advance research on employment participation of older workers, that is to

– address conceptual gaps,
– close regional research gaps, and
– fill thematic gaps (Hasselhorn and Apt 2015, P 111–114)

Conceptual Gap

A broad view of retirement requires a conceptual framework which locates retirement within the context of a multitude of determinants on individual, organizational and societal level and allows for a life course perspective. The JPI UEP working group found that most studies did not adopt such a systems view and multifactorial approach and thus bear the risk of overlooking the complexity of retirement. Also, specific longitudinal research approaches and applications of life course perspectives were frequently lacking. Furthermore, a lack of broadly conceptualized cross-national comparative research was identified. Finally, there is poor coverage of important population subgroups like older women, migrants or manual workers, who should be in the focus of retirement policy attention.

The JPI UEP working group concludes that advancement of research on the employment participation of older workers could be expected if retirement research would increase multifactorial thinking, apply more multilevel approaches and consider the multitude of employment exit routes. In addition, research should cover population subgroups and their distinct characteristics with respect to gender, social or occupational status, migration/ethnicity or age. Finally, retirement research would benefit from an openness as to the selection and combination of methodological approaches.

Regional Gap

The project findings indicate that research on employment participation of older workers is distributed very unevenly across the review countries. Highest (and high) research coverage was found in Norway and the Netherlands. Canada and Finland also showed a high level of research activity. Also Sweden and the UK have a substantial research tradition. Qualitative research was found to be more frequent in the UK and the Netherlands. Only very few research activities were identified in Poland.

The uneven distribution of retirement research across the participating countries can partly be explained by the fact that the Nordic countries and the Netherlands provide the opportunity to the research community to link individual survey data to register-based data. An explanation for the marginal research coverage in Poland could be that in the past decades the country was focused on a socially and economically painful transition from a planned to a free-market economy with high unemployment. Here, early exit pathways for older workers were opened as means for coping with transition and unemployment.

The lack of a European cross-national data structure and research coverage limits the possibility to conduct cross-national comparative studies. However, these would provide a highly valuable source for mutual learning and for identifying the main determinants that influence the transition from work to retirement. Taking the cross-national differences in the social welfare systems into account, comparative studies might be able to answer questions concerning country-specific effects and distinguish them from general, more systematic phenomena, mechanisms and settings.

Thematic Gap

The JPI UEP working group has identified a number of thematic gaps in retirement research, which should be closed by adopting a differentiated view: the role of health in the context of retirement, potential domestic and household factors, new work exposures like the increased use of technology or higher flexibility in work settings and their possible impact on retirement, the role of older women in retirement, the relation between migration and retirement, social position, the employers' perspective; the opportunities for organizational intervention, and the societal costs and gains of policy changes. To give an example:

In many countries, the group of *older women* is considered as an "untapped" potential for increasing employment participation. So far, the topic "employment participation of older women" has found scientific attention in a few countries only such as the United Kingdom, Norway, Canada and the United States. In view of the continuing gender segregation of jobs, older women's higher financial dependency in later lives and restrictions of older women's social and political autonomy, it is obvious that retirement is not the same for women and for men. Intercultural differences will also have an effect. A differentiated and up-to-date understanding of older women's multiple roles, needs, preferences and employment potential is needed. This would promote the development of supportive measures on the organizational and national policy level and by structural support in social and labour market policy. Here, cross-national and multi-methods research (including qualitative approaches) could contribute to a substantial gain in knowledge.

4 Discussion

The *lidA conceptual framework on work, age and employment* is based on the observation that retirement is complex and therefore requires broad views and broad scientific approaches. Based on the framework and its implications, the JPI UEP project has identified three research gaps with respect to retirement research: a conceptual gap, a regional gap and a thematic gap. To really understand the process of retirement in its complexity, trends and risk groups and to provide organizational and national policies with valid evidence, research initiatives are needed that contribute to bridging these three gaps. Ideally, such initiatives should combine different methodological approaches, allow for a life course view and should make use of the potential of cross-cultural comparisons. Not least, they should be of interdisciplinary nature and bring different research communities together.

Human Engineering as explicitly *integrated science* can not only contribute with own expertise in the field, but might adopt an integrating and moderating role here.

References

Beehr TA (1986) The process of retirement: a review and recommendation for further investigation. Person Psychol 29:31–55

Brown P, Vickerstaff S (2011) Health subjectivities and labor market participation: pessimism and older workers' attitudes and narratives around retirement. Res Aging 33(5):529–550

Bundesministerium für Wirtschaft und Energie (2015) Jahreswirtschaftsbericht 2015. https://www.bmwi.de/BMWi/Redaktion/PDF/J-L/jahreswirtschaftsbericht-2015.pdf. Last accessed 30 Jul 2016

De Wind A, Geuskens GA, Ybema JF, Blatter BM, Burdorf A, Bongers PM et al (2014) Health, job characteristics, skills, and social and financial factors in relation to early retirement—results from a longitudinal study in the Netherlands. Scand J Work Environ Health 40 (2):186–194

De Wind A, Geuskens GA, Ybema JF, Bongers PM, van der Beek AJ (2015) The role of ability, motivation, and opportunity to work in the transition from work to early retirement—testing and optimizing the Early Retirement Model. ScandJ Work Environ Health 41(1):24–35. doi:10.5271/sjweh.3468

Ekerdt DJ (2010) Frontiers of research on work and retirement. J Gerontol Soc Sci 65B:69–80. doi:10.1093/geronb/gbp109

Eurostat (2015) Figures from EUROSTAT Data Explorer, Employment (main characteristics and rates)—annual averages (lfsi_emp_a). http://ec.europa.eu/eurostat/data/database?node_code = lfsi_emp_a. Last accessed 30 Jul 2016
Fuchs J, Kubis A, Schneider L (2015) Zuwanderungsbedarf aus Drittstaaten in Deutschland bis 2050—Szenarien für ein konstantes Erwerbspersonenpotenzial—unter Berücksichtigung der zukünftigen inländischen Erwerbsbeteiligung und der EU-Binnenmobilität. Gütersloh: Bertelsmann. http://www.bertelsmann-stif-tung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Studie_IB_Zuwanderungsbedarf_aus_Drittstaaten_in_Deutschland_bis_2050_2015.pdf. Last accessed 30 Jul 2016
Hasselhorn HM, Apt W (2015) Understanding employment participation of older workers: creating a knowledge base for future labour market challenges. Research report. Federal Ministry of Labour and Social Affairs (BMAS) and Federal Institute for Occupational Safety and Health (BAuA). BMAS/BAuA, Berlin, 2015, 120 Seiten. http://www.jp-demographic.eu/wp-content/uploads/2015/07/JPIUEP_Brochure1.pdf. Last accessed 30 Jul 2016
Hasselhorn HM, Müller BH (2014) Die (begrenzte) Rolle der Gesundheit zur Sicherung der Erwerbsteilhabe. In: Weber A, Peschkes L, de Boer W (eds) Return to work—Arbeit für alle—Grundlagen der beruflichen Reintegration. Gentner Verlag, Stuttgart, pp 93–102
Hasselhorn HM, Peter R, Rauch A et al (2014) Cohort profile: the lidA Cohort Study—a German Cohort Study on work, age, health and work participation. Int J Epidemiol 2014:1736–1749. doi:10.1093/ije/dyu021
Hasselhorn HM, Ebener M, Müller BH (2015) Determinanten der Erwerbsteilhabe im höheren Erwerbsalter—das „lidA-Denkmodell zu Arbeit, Alter und Erwerbsteilhabe". ZSR 61 (4):403–432
Higgs P, Mein G, Ferrie J, Hyde M, Nazroo J (2003) Pathways to early retirement: structure and agency in decisionmaking among British civil servants. Ageing Soc 23:761–778
Ilmarinen J, Tuomi K (2004) Past, present and future of work ability. In: Ilmarinen J, Lehtinen S (eds) Past, present and future of work ability. Proceedings of the 1. International symposium on work ability, 5–6 September 2001, Tampere, Finland (People and work, Bd. 65, S. 1–25). Finnish Inst. of Occupational Health, Helsinki
Loretto W, Vickerstaff S (2012) The domestic and gendered context for retirement. Hum Relat 66:65–86
Naegele G, Sporket M (2010) Perspektiven einer lebenslauforientierten Ältere-Arbeitnehmer-Politik. In: Nägele G (ed) Soziale Lebenslaufpolitik. VS Verlag für Sozialwissenschaften (GWV Fachverlage GmbH), Wiesbaden
Phillipson C, Smith A (2005) Extending working life: a review of the research literature. Department for work and pensions, research report no 29. Corporate Document Services, Leeds
Pond R, Stephens C, Alpass F (2010) How health affects retirement decisions: three pathways taken by middle-older aged New Zealanders. Ageing Soc 30(3):527–545, http://dx.doi.org/10.1017/S0144686X09990523
Szinovacz ME (2013) A multilevel perspective for retirement research. In Wang M (ed) The Oxford handbook of retirement research. Oxford Library of Psychology, Oxford University Press, Oxford, New York, pp 152–172
Van den Heuvel S, de Wind A (2015) Domain: health and health-related behaviour. In: Hasselhorn HM, Apt W (eds) Understanding employment participation of older workers: creating a knowledge base for future labour market challenges. Research report. Federal Ministry of Labour and Social Affairs (BMAS) and Federal Institute for Occupational Safety and Health (BAuA). BMAS/BAuA, Berlin. http://www.jp-demographic.eu/wp-content/uploads/2015/07/JPIUEP_Brochure1.pdf. Last accessed 30 Jul 2016
Wang M, Shultz KS (2010) Employee retirement: a review and recommendations for future investigation. J Manag 36:172–206

Age-Differentiated Analysis of the Influence of Task Descriptions on Learning Sensorimotor Tasks

Francoise Kuhlenbäumer, Philipp Przybysz, Susanne Mütze-Niewöhner, and Christopher Marc Schlick

Abstract

This paper presents a study into the validity of a self-developed method to predict the learning time of sensorimotor tasks that was originally developed for young adults (age group AG I) for persons aged between 52 and 67 (age group AG II). For this purpose, a laboratory study was conducted with an age-differentiated sample of 60 participants. The participants' task was to repeatedly assemble a carburetor with the help of one of three task descriptions, which differed in regard to format (textual, text- & figure-based, animated). Execution times and numbers of assembly errors were measured to evaluate human performance. Additionally, the cumulative viewing time of the task description was measured in each trial to analyze participants' usage of the task description. Data analysis with respect to the age group and the format of the task descriptions indicates significant effects ($\alpha = 0.05$). Thus, participants who had the support of a textual task description achieved greater performance improvement than participants who used the animated task description. Concerning the age group, participants in AG I show better performance and lower observation times concerning to participants in AG II. Furthermore, nonlinear curve fittings were carry out and root mean square errors calculated in order to investigate the accuracy of the prediction method. The results show that the prediction method is less accurate for older adults.

F. Kuhlenbäumer (✉) • P. Przybysz • S. Mütze-Niewöhner
Arbeitsorganisation, Institut für Arbeitswissenschaft der RWTH Aachen, Bergdriesch 27, 52062 Aachen, Germany
e-mail: f.kuhlenbaeumer@iaw.rwth-aachen.de; p.przybysz@iaw.rwth-aachen.de; s.muetze@iaw.rwth-aachen.de

C.M. Schlick
Institut für Arbeitswissenschaft der RWTH Aachen, Bergdriesch 27, 52062 Aachen, Germany
e-mail: c.schlick@iaw.rwth-aachen.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_12

Keywords
Age-differentiated • Learning time • Task descriptions

1 Introduction

Demographic developments in Germany are posing challenges for labor organization and human resource management at manufacturing companies. In response, those companies have carried out process optimization in terms of quality, time and cost and have initiated projects how to preserve people's ability to work (Dombrowski and Mielke 2012). A couple of core tasks have been identified for managing the challenges associated with demographic change (Ackermann 2010), e.g. personnel deployment planning. It should aim at reducing physical strain through frequent changes of task and types of burden and enhancing motivation by giving workers the possibility to learn new tasks (Ackermann 2010; Stanić 2010). Since work tasks in production mostly require sensorimotor skills, these skills have to be learned specifically for each task (Rohmert et al. 1974). This means that the working person needs a certain amount of time until he or she is able to execute the new task in a way that satisfies quality and quantity requirements. This period is referred to as the "learning time" and depends on various influencing factors, which can be grouped into three classes: (1) the class of the working person, (2) the class of the work task and the class of the learning method (De Greiff 2001; Jeske 2013). When considering demographic change, special attention should be paid to learning methods. An important task for managing demographic change effectively is differentiating between learning provisions for younger and older working persons according to their knowledge and learning capabilities (Ackermann 2010). A laboratory study by Jeske (Jeske 2013; Jeske et al. 2011) investigated the influence of the format of task descriptions on learning times. Jeske found that learning time decreases with increasing information richness, according to the media richness model (Daft and Lengel 1986). Furthermore, Jeske (Jeske 2013; Jeske and Schlick 2012) developed a method to forecast the learning time of sensorimotor tasks in which the type of task description is one of the most important predictors. However, Jeske considered only young adults in his investigations. Thus, it cannot be assured that his results are valid for older persons and that his model-based predictions are therefore a valid tool for supporting the management of demographic change in manufacturing companies.

2 Background

Learning time is the period needed to acquire the ability to perform an unknown work task and subsequently to practice this task until reaching a reference performance level. Figure 1 illustrates these two phases of learning time. The first phase is characterized by a rathe passive learning process in which the learning person gets

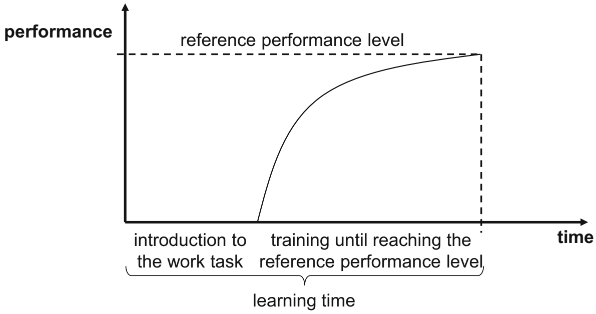


Fig. 1 Schematic representation of learning time (Jeske and Schlick 2011)

information about the new task. For instance, one might imagine the learning person observing another person executing the task or that he or she is given instruction in how to perform the new task. In contrast, the second phase is characterized by self-determined execution of the work task by the person him/herself (Jeske 2013).

The performance improvements due to development of skills and knowledge during the second phase can be very well described by means of learning curves. The first learning curve for industrial usage was formulated by Wright (Laarmann 2005; Wright 1936). According to his model, performance improves with a rising number of trials as expressed by a power function. The problem with this model is that improvement in performance becomes infinite given increasing numbers of trials. This is unrealistic, making the model unsatisfactory for most applications. Therefore, many other models have been formulated (Yaber 2011). For instance, De Jong (de Jong 1960) extended Wright's power function model (Wright 1936) by a dimension of irreducibility that is mathematically modeled as an asymptote. Other authors have used different function types like exponential or logarithm functions (de Greiff 1999). The interested reader finds more information in the literature, e.g. (Hieber 1991; Laarmann 2005; Yaber 2011; Yelle 1979). However, a mere description of learning time is seldom sufficient. In many cases, e.g. for purposes of planning and qualification, it is also important to forecast the learning time. Because previous prediction methods were inaccurate for industrial applications, Jeske (Jeske 2013; Jeske and Schlick 2012) developed a new prediction method for sensorimotor work tasks. This method estimates the parameters of the learning curve on the basis of information about the working person, the task itself, and the learning method. To describe improvement in performance following repeated executions, Jeske takes as a basis a learning curve that includes a limiting value c and provides a proportionality constant k, which he interprets as learning velocity. Furthermore, Jeske introduces a factor λ as a model parameter in his learning curve. The product of λ and c represents the time needed for the first execution t_1. Thus, the equation for the time needed for the n-th execution t_n is (Jeske and Schlick 2012):

$$t_n = c + \Bigg(\underbrace{c \cdot \lambda}_{=t_1} - c\Bigg) \cdot n^{-k} \cdot n^{-k \cdot e^{-\frac{k}{2} \cdot (n-1)}}. \tag{1}$$

Jeske assumes that the exponential function included in his learning curve indicates initial learning processes in terms of gaining the required declarative and procedural knowledge, whereas the power function represents the acquisition of motor skills. Jeske justified this assumption based on the observation that participants did not use the task descriptions after a certain amount of time. Jeske reasoned that reduced usage of task descriptions goes along with increased knowledge of the work task (Jeske 2013; Jeske and Schlick 2012). However, Jeske did not quantify the participants' usage of task descriptions.

In order to forecast learning time, Jeske suggests determining the limiting value c by means of the Methods-Time Measurement Universal Analyzing System (MTM-UAS). In regression analyses, Jeske found that λ is influenced by the difficulty of the work task H; the experience of the working person with assembly E; the type of task description D; and the gender of the working person G. Thereby, Jeske (Jeske and Schlick 2011, 2012) defines the difficulty of the work task H as an a cumulated Euclidean entropy of parts and movements according to MTM-UAS. The parameter k is determined by the parameter λ, the first and sixth factor of Fleishman F_1 and F_6 (Fleishman 1962; Fleishman and Ellison 1962), which describe the fine motor skills of the working person, and the age of the working person A. The estimation of the model parameters λ and k is carried out on the basis of the following equations (Jeske and Schlick 2012):

$$\lambda = 2.256 + 0.978 \cdot H_{UAS} - 0.755 \cdot E_{Assembly} - 0.45 \cdot D + 0.87 \cdot G \tag{2}$$

$$k = 0.141 + 0.073 \cdot \lambda - 0.008 \cdot FF_1 + FF_6 + 0.013 \cdot A. \tag{3}$$

Jeske was able to explain a large proportion of variance [$R^2 = 0.7896$ (Jeske and Schlick 2012)] in experimental validation. However, the development of the model and its validation were based on samples with participants aged between 18 and 41 years. Thus, it is not clear to what extent Jeske's method is valid for older persons.

3 Method

We evaluated Jeske's prediction model (Jeske 2013; Jeske and Schlick 2012) for participants between 52 and 67 years of age in a replication study. To do this, we conducted Jeske's study into the influence of task descriptions on learning sensorimotor tasks (Jeske et al. 2011) with an age-differentiated sample. We also analyzed participants' usage of task descriptions since Jeske (Jeske 2013; Jeske and Schlick 2012) could not quantify this in order to confirm his assumptions with regard to the structure of the learning curve (see Sect. 2). Moreover, it is expected that deeper insight will be gained into the differences between younger and older participants' learning behaviors. In this context, it is expected that older participants are more conscientious, which should result in longer observation times and fewer assembly errors than younger participants.

3.1 Participants

Thirty right-handed participants aged between 20 and 35 (age group AG I: $M = 25.17$ years, $SD = 3696$) and 30 right-handed subjects aged between 52 and 67 (age group AG II[1]: $M = 59.70$ years, $SD = 4779$) participated in this study. There were equal numbers of men and women in both age groups. Gender and age group were balanced in all experimental conditions. All participants had neither health problems nor health restrictions of their hand and arm system. Participants' binocular vision was ensured through the use of Landolt's rings at the distances 0.4 m, 0.55 m and 6 m.

3.2 Experimental Task and Task Descriptions

The experimental task in question was assembly of a Stromberg carburetor. This task requires the assembly of 32 components of different sizes. Following Jeske's formula of complexity, its degree of complexity is $H = 4.682$. According to MTM UAS, the time needed for well-trained persons is $t_{MTM\ UAS} = 146.2$ s (Jeske and Schlick 2011). To execute the task, participants were given a standardized seated workplace with all components and tools assigned to defined positions (see Fig. 2). During the assembly, participants were provided with one of three task descriptions. All task descriptions described the necessary steps for assembling the carburetor in three separate sections. However, they differed in terms of their format: (1) textual, (2) text- & figure-based and (3) animated. The task description was presented on a touch screen, which enabled the participants to switch between the three sections or, in case of the animated task description, to start and stop the animation according to their needs.

To stimulate the learning process, participants had to repeat this task ten times. After each trial a break of two minutes was provided.

[1]Compared to (Kuhlenbäumer et al. 2016a), one male participant in AG II was replaced, because he was identified as an outlier in later analyses. An analysis without this participant (59 participants) can be found in (Kuhlenbäumer et al. 2016b).

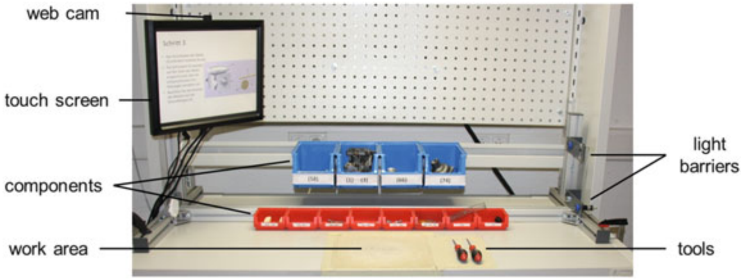


Fig. 2 Experimental setup

3.3 Experimental Variables

The independent variables are the two age groups (AG I: 20–35 years vs. AG II: 52–67 years) and three different task descriptions (textual, text- & figure-based, animated). Execution times and assembly errors in each trial were chosen as dependent variables for evaluating changes in performance. In order to analyze participants' behavior with regard to usage of the task descriptions, the duration of observation in each trial were dependent variables as well. The subjective workload at the end of the experiment was another dependent variable that was surveyed with the help of NASA Task Load Index [NASA TLX (Hart and Staveland 1988)].

3.4 Procedure

All characteristics of the participants that should have an influence on learning time according to Jeske (Jeske 2013; Jeske and Schlick 2012) were collected before the main experimental task was carried out. For this reason, participants had to complete a demographic questionnaire first. Besides questions about e.g. age and gender, the questionnaire also included a self-evaluation of participants' experience of assembly and of carburetors on a four-point scale (none, little, medium, much). The measurement of fine motor skills, which are considered in Jeske's prediction model as the first and the sixth Fleishman factors (Fleishman 1962; Fleishman and Ellison 1962), was carried out using the Motor Performance Series (Hamster 1980; Schoppe 1974). Participants' technical comprehension and retentiveness were also surveyed. To this end, a subtest of the AZUBI-TH ["Arbeitsprobe zur berufsbezogenen Intelligenz – Technische und handwerkliche Tätigkeiten" (Görlich and Schuler 2007), which can be roughly translated as "work sample on occupational intelligence – technical and mechanical jobs"] and a subtest of the German version of the Intelligence Structure Test IST-2000R ["Intelligenz-Struktur-Test 2000R" (Liepmann 2007)] were used.

Following the pretests, the main task (see Sect. 3.2), that is the repeated assembly of the carburetor started. During the assembly, participants' fixations at the task description were recorded with the help of a webcam on top of the monitor, which made it possible to determine duration of task description usage. The execution times and assembly errors of each trial were also measured. After each trial, participants were informed about their errors in the previous assembly.

The experiment ended with the NASA Task Load Index [NASA TLX (Hart and Staveland 1988)], which determines the perceived workload of each participant.

3.5 Statistical Analysis

The statistical analyses of the replication study were conducted with the statistical software package SPSS (Version 21.0). Repeated-measure analyses of variance (ANOVA) were used to analyze changes in execution time, numbers of errors, and observation time based on repeated execution. The effect size ω^2 was calculated for each significant effect and pairwise comparisons with Bonferroni correction were used as a post-hoc test. Furthermore, Peason's r was calculated to analyze the correlation between observation time and execution time or assembly errors. The chosen level of significance for each analysis was $\alpha = 0.05$ (Field 2009).

The MathWorks Matlab R2016a software was used to evaluate the model accuracy and prediction quality of Jeske's method. The analysis of model accuracy was based on best curve fits. The model parameters λ and k were approximated to the data of all participants and separately to the data of AG I and AG II. The coefficient of determination R^2 was calculated to evaluated the goodness-of-fit of Jeske's learning curve. The prediction quality of Jeske's method was analyzed qualitatively by comparing measured and individually predicted execution times. In order to quantitatively evaluate the prediction quality of Jeske's method (Jeske 2013), the root mean square error (RSME) was calculated. When calculating the RSME, the difference between measured and predicted execution times is squared to prevent overestimations and underestimations from canceling each other out.

4 Results and Discussion

4.1 Statistical Analysis of the Replication Study

4.1.1 Participants' Characteristics

The participants' self-evaluated experiences with carburetors and assembly are shown according to age group in Fig. 3. Most participants had no prior experience of carburetors. Just a few participants in AG II stated that they had medium or much experience of carburetors. Experience of assembly was somewhat greater between the participants: most participants in both age groups had none or little experience of assembly although some participants in AG II had much experience, and a few participants in AG I said that they had medium or much experience.

The average results of the motor test are shown in Table 1. Values within the range of 50 ± 10 indicate normal fine motor skills. With regard to technical comprehension, participants in AG I answered 10.63 ± 2.526 out of 16 questions

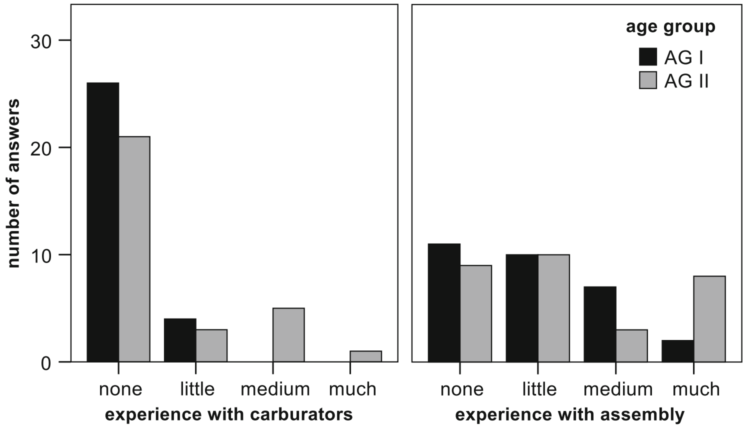


Fig. 3 Participants' experience of assembly and of carburetors

Table 1 Participants' fine motor skills (mean T-value with standard deviation)

Age group	Fleishman factor					
	1	2	3	4	5	6
AG I	51.98 (±6.183)	51.07 (±8.526)	52.50 (±7.031)	56.40 (±6.941)	51.93 (±7.187)	54.83 (±8.542)
AG II	54.92 (±3.351)	52.53 (±7.431)	48.97 (±6.552)	53.47 (±7.167)	46.70 (±6.834)	50.67 (±10.039)

Values within the range of 50±10 indicate normal fine motor skills

correctly and participants in AG II answered 9.27 ± 3.619 correctly. The average retentiveness expressed in default values is 105.40 ± 8.850 for participants in AG I and 103.067 ± 8.379 for participants in AG II. Values within the range of 100 ± 10 indicate normal retentiveness.

4.1.2 Execution Times

Average execution times and 95 % confidence intervals for participants in both age groups are shown with respect to the different task descriptions in Fig. 4. It becomes apparent that execution time decreases with the number of trials for each experimental condition. However, it is also obvious that the difference between successive execution times decreases so that execution times converge to a limit value. These observations indicate a strong learning effect that was also noted by Jeske (Jeske 2013) and is described in other literature on learning time (e.g. Hieber 1991; Rohmert et al. 1974; Yaber 2011). In statistical analyses, the repeated execution is a significant main effect with a large effect size ($F_{(3.570;\ 92.806)} = 179.664$; $p < 0.001$; $\omega^2 = 0.688$). In pairwise comparisons, it is evident that there are significant differences between successive execution times until the fourth trial. There are no significant differences between all subsequent execution times from the eighth trial on. Consequently, a learning effect is statistically proven.

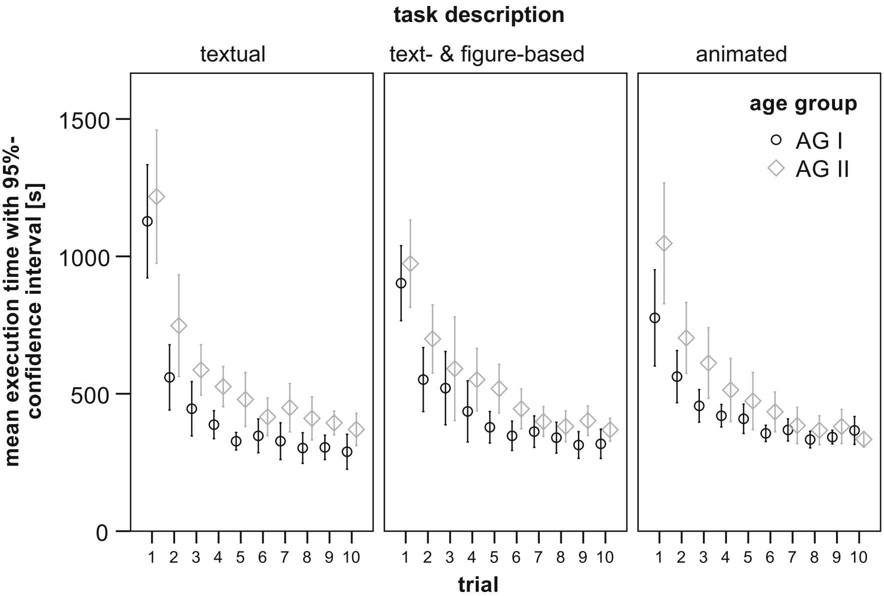


Fig. 4 Means with 95 % confidence interval of execution times of both age groups according to task description

The changes in performance due to repeated execution thereby differ depending on the different task descriptions. This is expressed in statistical analyses as significant interaction with a small effect size ($F_{(7.141;\ 192.806)} = 3.448$; $p = 0.002$; $\omega^2 = 0.016$). An analysis of improvement from trial to trial shows a significant main effect of task description with a large effect size ($F_{(2;\ 54)} = 5.066$; $p = 0.010$; $\omega^2 = 0.157$). Post-hoc analyses reveal a significant difference ($p = 0.009$) between the textual and the animated task description. Thus participants who had the support of a textual task description achieved greater performance improvement than participants who used the animated task description. There are no significant differences in level of performance improvement when using the textual vs. the text- & figure-based or the text- & figure-based vs. the animated task descriptions. These results are partially in line with Jeske's (2013) findings. Jeske (2013) also found a significant difference in the level of performance improvement when using the textual vs. the text & figure-based task description.

The difference between task descriptions is particularly noticeable in the first execution. The execution time needed by participants in both age groups for the first assembly is longest when using the textual task description. In contrast, the first execution times tend to differ between the age groups when using the other task descriptions. While participants in AG I needed the least amount of time for the first trial when using the animated task description, participants in AG II were fastest with the text & figure-based task description. The statistical analysis shows a significant difference between first execution times, with a large effect size when using the textual or animated task description, for participants in AG I ($F_{(2;\ 27)} = 5.308$; $p = 0.011$; $\omega^2 = 0.223$). For this age group, the relationship between informational richness of task descriptions and first execution time can be statistically represented as a significant correlation ($r = -0.524$; $p = 0.003$), as Jeske (2013) also noted. In contrast, the task description used by participants in AG II had no effect on the time of the first execution. However, in the statistical analyses a significant effect ($F_{(2;\ 54)} = 3.439$; $p = 0.039$) of age group and task description was determined on the tenth execution time. The associated effect size is medium ($\omega^2 = 0.074$). Thus, there is a significant difference between the age groups in the tenth trial, with a medium effect size when the textual task description is used ($T_{(18)} = -2.116$; $p = 0.049$; $r = 0.446$). In particular, participants in AG I ($M = 289.011$ s; $SD = 88.978$) needed on average less time for the tenth execution than participants in AG II ($M = 370.226$ s; $SD = 82.569$).

Figure 4 shows that participants in AG II also needed more time in nearly all trials than participants of AG I. The statistical analyses confirm this, with a significant main effect of age group ($F_{(1;\ 54)} = 20.160$; $p < 0.001$). The associated effect size is large ($\omega^2 = 0.39$).

Consequently, learning time can be seen to be significantly influenced by age group and task description.

4.1.3 Assembly Errors

Figure 5 shows the average number of participants' assembly errors in both age groups according to task description. As already discussed, a strong learning effect was found with regard to execution times. Accordingly, the number of assembly errors decreases with increasing number of trials. The statistical analysis confirms this observation with a large effect size ($F_{(3.664;\ 197,831)} = 97.477$; $p < 0.001$; $\omega^2 = 0.558$). Pairwise comparison indicates significant reductions in the number of errors for the first three successive executions. There are no significant differences between the numbers of errors from the sixth trial onward. Thus, the learning effect can also be proven on the basis of decreasing errors.

In contrast to execution times, occurrence of errors is not influenced by task descriptions. Although it seems that participants make on average fewer errors when using the textual task description than when using the others, no significant effect is revealed in the statistical analyses. However, it can be noted that the descriptive analysis reveals a relationship between the number of errors and the related execution times. Thus, the faster the participants in one age group perform their task (on average) when using a specific task description, the higher the mean number of errors is. This typical speed-accuracy trade-off conforms with Morrell and Park (1993) who, among others, investigated the effect of task descriptions on performance of assembly tasks and found that "extra time spent in assembly did result in accurate final products" (p. 397).

With regard to the influence of age, it can be seen in Fig. 5 that participants in AG II committed more errors in nearly all the trials. In the statistical analyses this observation is shown as significant ($F_{(1;\ 54)} = 4.029$; $p = 0.05$). The associated effect size is medium ($\omega^2 = 0.09$).

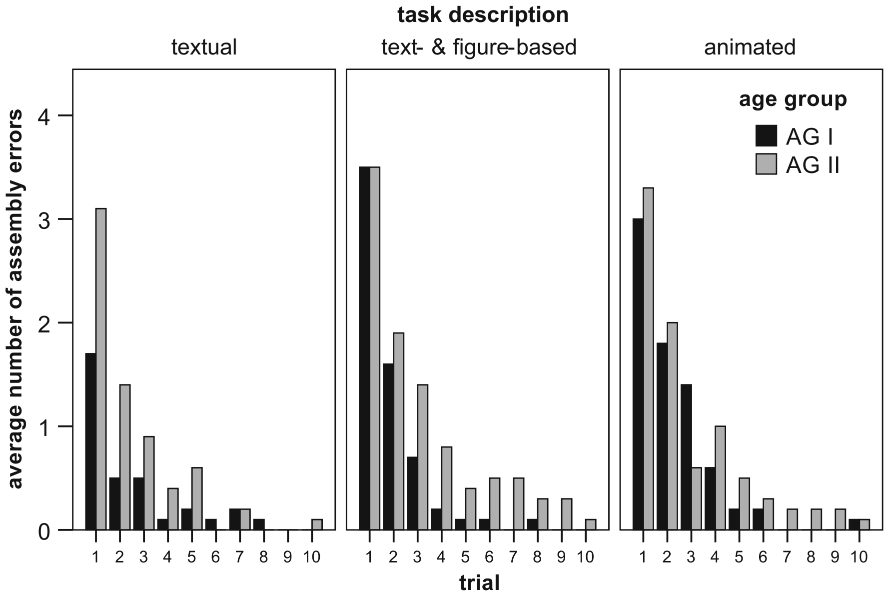


Fig. 5 Assembly errors of both age groups according to task description

4.1.4 Participants' Behavior to Use Task Descriptions

The average observation times and 95 %-confidence intervals for participants in both age groups are shown in Fig. 6. Usage of task descriptions decreases with number of trials. This decrease is statistically significant in the repeated-measures ANOVA ($F_{(2.421;\ 128.326)} = 270.079$; $p < 0.001$). In subsequent pairwise comparisons between the observation times of the trials, the decrease in usage with increasing numbers of trials is significant until the fourth trial, in the case of all trials. Scattered significant differences exist until the tenth trial.

Given the reduced number of errors with repeated execution, the data analysis suggests that participants need less and less information from the task description in order to assemble the carburetor. As a consequence, they have to spend less time reading the task description. This, in turn, has consequences for the reduction of execution times. To investigate this more closely, correlation analyses were conducted between the observation times and execution times or assembly errors in each trial. The results show that there are positive significant correlations between execution time and observation time until the seventh (AG I) or the sixth (AG II) execution. However, the results show just three significant correlations between the number of errors and the observation time in the same trial. These correlations are positive and appear in the first ($r = 0.472$; $p = 0.010$), fifth ($r = 0.426$; $p = 0.021$) and sixth trial ($r = 0.410$; $p = 0.027$) of participants in AG II. It is assumed that the correlation of the first trial can be attributed to a poor comprehension of the task description, so that longer observation times do not result in fewer errors. The occurrences of positive significant correlations in the

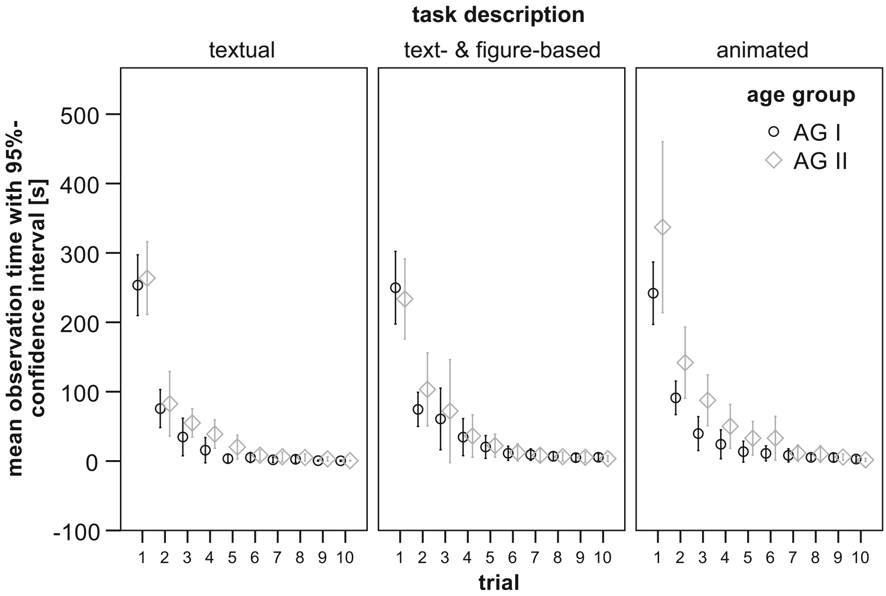


Fig. 6 Mean observation times with 95 % confidence interval for both age groups according to task description

fifth and sixth trial seem to be random and cannot be explained by human performance models.

When comparing the different task descriptions, it is apparent that the first observation times of participants in AG I are nearly the same for all types of task description. In contrast, the observation time for the first trial of participants in AG II seems to be higher when using the animated task description. However, in the statistical analysis no significant difference is found between the different task descriptions—neither for the first trial nor for all the other trials. Thus, the influence of task descriptions on execution times (see Sect. 4.1.2) is independent of observation times.

Overall, it appears that participants in AG II used the task descriptions longer than participants in AG I ($F_{(1;\ 53)} = 4.453$; $p = 0.040$).

4.1.5 Subjectively Perceived Workload

The participants' subjective perceived workload is shown in Fig. 7. A value of zero would mean *no perceived workload* and a value of 15 equals a *maximum perceived workload*. The values for AG I do not differ much for the different task descriptions. In contrast, the participants in AG II seem to perceived higher load when using the text- & figure-based task description as opposed to the other task descriptions. In this case, the perceived workload is higher for AG II than for AG I. In all other cases, the perceived workload is lower for AG II. However, there are no significant differences in subjective workload according to age group or task description.

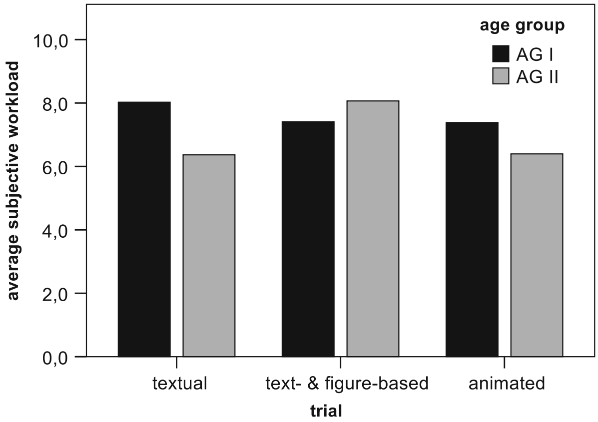


Fig. 7 Subjectively perceived workload measured by NASA TLX (Hart and Staveland 1988) for both age groups according to task description

4.2 Model Accuracy and Prediction Quality of Jeske's Method

4.2.1 Curve Fit

The investigation into the accuracy of Jeske's model is based on best curve fits. For this purpose, it was assumed that the limiting value is $c = 146.2$ s, following MTM UAS. This is the figure assumed by Jeske in his prediction method (Jeske and Schlick 2012). The model parameters λ and k are approximated for (1) all data, (2) data on AG I, and (3) data on AG II. Figure 8 shows the best curve fits according to the values of the coefficient of determination R^2 and the parameters λ and k.

The coefficient of determination R^2 is only 61.5 % when taking all participants into consideration. In age-differentiated curve fits, coefficients of determination of $R^2 = 66.8$ % (AG I) and $R^2 = 62.7$ % (AG II) were achieved. Thus, Jeske's learning curve is more suitable for AG I than AG II. This is particularly apparent from the second to fifth trials of AG II: Jeske's learning curve assumes that the reduction of execution time is larger than the measured reduction of execution time. The steep reduction in Jeske's learning curve can be attributed to the exponential term, which describes initial learning processes in terms of gaining knowledge about the task (Jeske 2013). This leads to the assumption that comprehension of the task is less important than acquisition of motor skills [which are taken into account in the power function (Jeske 2013)], in the reduction of execution times for AG II.

4.2.2 Prediction Quality of Jeske's Method

Figure 9 provides an age-differentiated comparison between measured execution times (X axis) and individually predicted execution times (Y axis) according to Jeske's prediction method (see Eqs. 1–3). Data points below/above the angle bisector indicate underestimations/overestimations by Jeske's method. Most execution times—in particular the execution times for AG II—were underestimated (Fig. 9).

The RMSE was calculated to quantify deviations between the measured and predicted execution times. The results are shown age-differentiated in Fig. 10. It

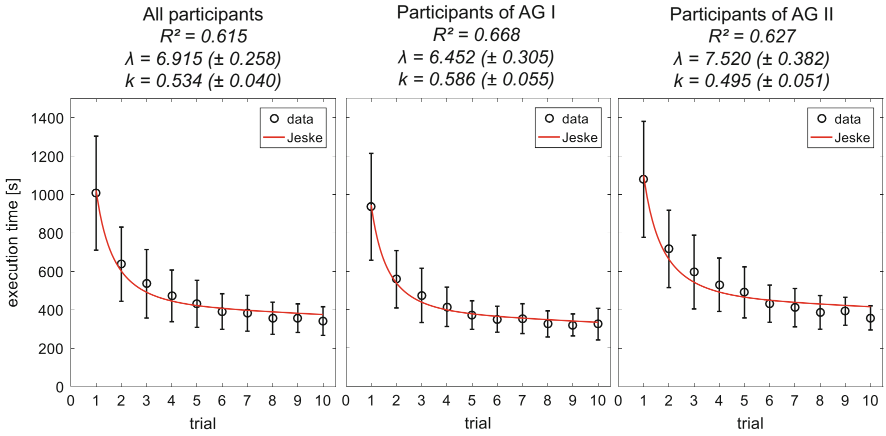


Fig. 8 Fitting of average execution times to Jeske's model for all participants and for both age groups

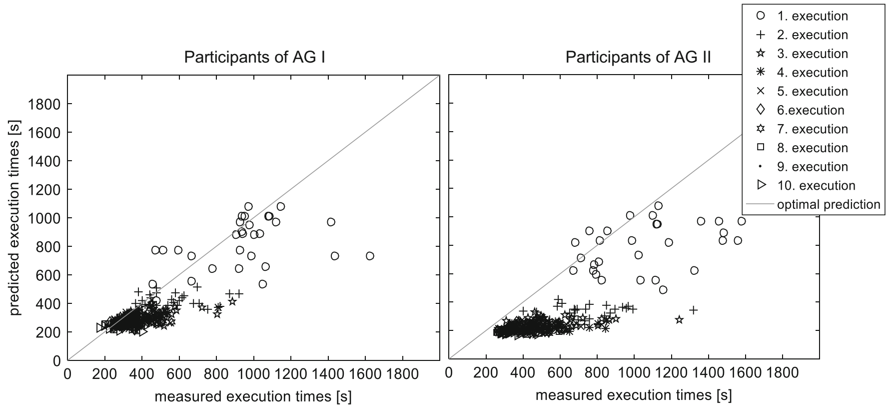


Fig. 9 Comparison of measured and individually predicted execution times

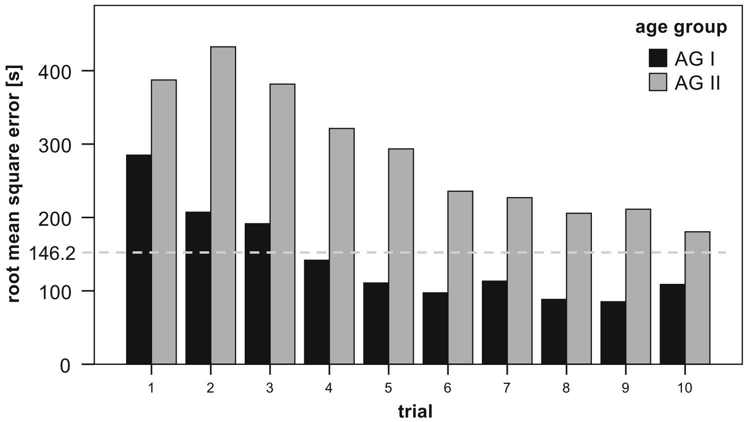


Fig. 10 RSME of predicted execution times according to Jeske (Jeske and Schlick 2012) for both age groups

gives as a reference value the standard assembly time (=146.2 s) according to MTM UAS. It is clear that the RMSE will be higher for AG II than for AG I in each trial. Furthermore, all RMSE in AG II are higher than the reference value, whereas the RMSE in AG I fall below the reference value after the third execution. Consequently, Jeske's prediction method is less suitable for predicting older people's learning times and (see Eq. 3) does not adequately consider age.

5 Conclusion and Outlook

This study showed that age group and task description significantly influence learning of a sensorimotor task. While there is a significant influence of task description on time improvement only when using the textual or the animated task description, age group significantly influences execution time and the number of assembly errors as a main effect. Furthermore, the study showed that a significant correlation between the first execution time and the informational richness (Daft and Lengel 1986) of task descriptions, which Jeske (2013) considered in his prediction method, only exists for younger participants. Thus, this correlation is not sufficient as dominant predictor for a generally valid prediction method. In further analyses it was found that the observation times on task descriptions is independent of the format. Hence, the influence of task descriptions on execution times cannot be traced back to longer observation times but rather to a easier understanding of the task when using the animated task description in comparison with the textual task description. However, in the analyses of observation times a significant difference between age groups was found. Due to the fact that a significant positive correlation was also found between usage and the number of errors in the first trial of older participants, it is assumed that older participants are more motivated to comprehend the task than young participants, even if they have difficulties completing the task. With regard to Jeske's method for predicting learning times, the analyses show that this method is less suitable for older participants. Thus, more age-differentiated investigations are necessary to identify the essential effects of age on learning time and to accordingly modify the existing prediction method. For this reason, an age-differentiated investigation is planned that considers the influence on learning times of specific instructions introducing working persons to a new task. This is expected to lead to further findings about the learning needs of younger and older persons. To modify the existing model, data will be fitted to other function processes (e.g. time series or vector autoregressive models), and advanced statistical methods will be used to increase prediction quality.

Acknowledgements This study is part of the research project "Coping with Aging in Manual Assembly Systems: Age-differentiated Analyses and Mathematical Modeling for Predicting the Time Structure of Sensorimotor-skill Acquisition for Assembly in Series Production with Numerous Product Variants" (SCHL 1805/9-1), which is funded by Deutsche Forschungsgemeinschaft.

References

Ackermann KF (2010) Alter als Kernaufgabe. Personal 06:12–14

Daft RL, Lengel RH (1986) Organizational information requirements, media richness and structural design. Manage Sci 32:554–571

de Greiff M (1999) Prognose von Lernkurven in der manuellen und teilmechanisierten Montage. REFA-Nachrichten 6:19–26

De Greiff M (2001) Die Prognose von Lernkurven in der manuellen Montage unter besonderer Berücksichtigung der Lernkurven von Grundbewegungen. VDI, Düsseldorf

de Jong JR (1960) Die Auswirkungen zunehmender Fertigkeiten. REFA Nachrichten 13 (1):155–161

Dombrowski U, Mielke T (2012) Entwicklungspfade zur Lösung des Demografieproblems in Deutschland. In: Müller E (ed) Demografischer Wandel: Herausforderungen für die Arbeits- und Betriebsorganisation der Zukunft. Gito mbH Verlag, Berlin, pp 55–80

Field A (2009) Discovering statistics using SPSS, 3rd edn. Sage, London

Fleishman EA (1962) The description and prediction of perceptual motor skill learning. In: Glaser R (ed) Training research and education. University of Pittsburgh Press, Pittsburgh, PA, pp 137–175

Fleishman EA, Ellison GD (1962) A factor analysis of fine manipulative performance. J Appl Psychol 46:96–105

Görlich Y, Schuler H (2007) Arbeitsprobe zur berufsbezogenen Intelligenz. Technische und handwerkliche Tätigkeiten; AZUBI-TH. Hogrefe, Göttingen, Vienna

Hamster W (1980) Die Motorische Leistungsserie – MLS. Handanweisung. Dr. G. Schuhfried, Mödling

Hart S, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) Human mental workload. North-Holland Press, Amsterdam, pp 239–250

Hieber WL (1991) Lern- und Erfahrungskurveneffekte und ihre Bestimmung in der flexibel automatisierten Produktion. Vahlen, Munich

Jeske T (2013) Entwicklung einer Methode zur Prognose der Anlernzeit sensumotorischer Tätigkeiten. In: Schlick CM (ed) Industrial engineering and ergonomics. Shaker, Aachen

Jeske T, Schlick CM (2011) Influence of task complexity on learning times of sensorimotor tasks in assembly systems. In: Spath D, Ilg R, Krause T (eds) Innovation in product and production – conference proceedings. 21st International Conference on Production Research (ICPR 21), Stuttgart, CD-ROM

Jeske T, Schlick CM (2012) A new method for forecasting the learning time of sensorimotor task. In: Trzcieliński S, Karwowski W (eds) Advances in ergonomics in manufacturing. CRC Press, Boca Raton, FL, pp 241–250

Jeske T, Hasenau K, Schlick CM (2011) Influence of task descriptions on learning sensorimotor tasks. In: Göbel M, Christie C, Zschernack S, Todd A, Mattison M (eds) Human factors in organisational design and management – X, vol 1. IEA Press, Santa Monica, CA, pp 343–348

Kuhlenbäumer F, Duckwitz S, Schlick CM (2016a) Altersdifferenzierte Untersuchung zur Prognose der Anlernzeit von sensumotorischen Arbeitsaufgaben. In: Gesellschaft für Arbeitswissenschaft e.V. (ed) Arbeit in komplexen Systemen – Digital, vernetzt, human?! 62nd Kongress der Gesellschaft für Arbeitswissenschaft of March 2 to 4, 2016. GfA-Press, Dortmund

Kuhlenbäumer F, Duckwitz S, Schlick CM (2016b) Age-differentiated modeling and prediction of the learning time of sensorimotor tasks. In: Schlick CM, Trzcieliński S (eds) Advances in ergonomics of manufacturing: managing the enterprise of the future: Proceedings of the AHFE2016 international conference on human aspects of advanced manufacturing, July 27–31, Walt Disney World R, Florida, USA

Laarmann A (2005) Lerneffekte in der Produktion. Deutscher Universitäts-Verlag, Wiesbaden

Liepmann D (2007) Intelligenz-Struktur-Test 2000 R. I-S-T 2000 R, 2nd edn. Hogrefe, Göttingen

Morrell R, Park DC (1993) The effects of age, illustrations, and task variables on the performance of procedural assembly tasks. Psychol Aging 8(3):389–399

Rohmert W, Rutenfranz J, Ulrich E (1974) Das Erlernen sensumotorischer Fertigkeiten. Institut für Arbeitswissenschaft, TU Darmstadt. Europäische Verlagsanstalt, Frankfurt a.M.

Schoppe KJ (1974) Das MLS-Gerät: ein neuer Testapparat zur Messung feinmotorischer Leistungen. Diagnostica 20:43–47

Stanić S (2010) Fahrzeugmontage – Herausforderungen für den demografischen Wandel. In: Sträter O, Frieling E (eds) Schriftenreihe Personal- und Organisationsentwicklung, 8

Wright TP (1936) Factors affecting the cost of airplanes. J Aeronaut Sci 3(4):122–128

Yaber MY (2011) Learning curves: theory, models, and applications. CRC Press, Boca Raton, FL

Yelle LE (1979) The learning curve: historical review and comprehensive survey. Decis Sci 10 (2):302–328

The Potential of Virtual Interactive Learning Environments for Individual and Organizational Learning: An Example for the Maintenance of Electrical Equipment

Michael Dick, Tina Haase, and Wilhelm Termath

Abstract

The contribution reflects a collaborative project between the Otto-von-Guericke University Magdeburg, the Fraunhofer Institute for Factory Operation and Automation IFF Magdeburg and an industrial partner from the field of maintenance of electrical equipment. On the basis of the experiential knowledge of the companies' experts, virtual interactive learning environments were developed and successfully proven. The development and the application of the technical system allowed the integration of individual and organizational learning by means of a systematic reflection and utilization of the expertise of the professionals in the companies.

Keywords

Virtual reality • Learning environment • Organizational learning • Maintenance • Experiential knowledge

1 Initial Situation

The starting point of the project was the interest of an energy company to improve the further education of the professionals in the field of maintenance by using new media. The application of virtual reality technologies should reveal the potential of

M. Dick (✉) • W. Termath
Institut I – Bildung, Beruf und Medien, Otto-von-Guericke Universität Magdeburg, Zschokkestr. 32, 39104 Magdeburg, Germany
e-mail: michael.dick@ovgu.de; wilhelm.termath@ovgu.de

T. Haase
Fraunhofer Institute for Factory Operation and Automation IFF/Business Unit: Measurement and Testing Technology, Sandtorstr. 22, 39106 Magdeburg, Germany
e-mail: tina.haase@iff.fraunhofer.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_13

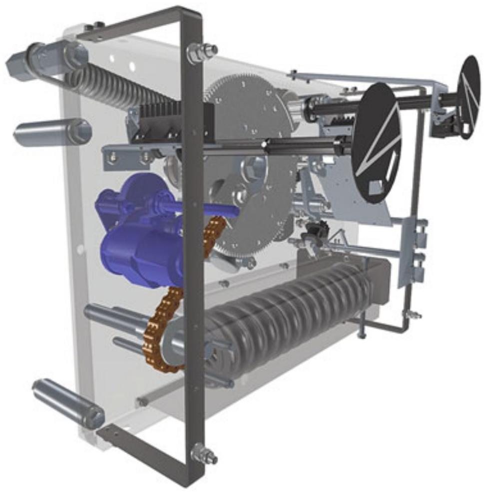


Fig. 1 3D-Visualization of the operating mechanism of a high voltage circuit breaker (Source: Fraunhofer IFF)

the visualization for the improved understanding of technical details and functional relations.

The ways of presenting technical processes and to interactively control them, should be used to create complex learning environments following the didactic guiding principle of activity-based learning (Beuting et al. 2010; Blümel 2007; Blümel et al. 2010). Figure 1 illustrates complex technical processes using the example of the operating mechanism of a high voltage circuit breaker.

In the course of cooperation it became obvious that it requires a significant effort of communication amongst the experts to describe the process of failure analysis in a way that it can be clearly visualized. Different experiences and perspectives need to be discussed and harmonized in order to design learning tasks and to develop a didactic concept for problem solving tasks.

The very fruitful discussions among the experts about promising and adequate solutions and strategies led to the demand to not only use virtual learning applications for the purpose of further education but also within the working process.

With respect to future developments the virtual models should be made available for each service technician in his working process. It was intended to give access to former successful solutions and experiences that are valuable for an efficient troubleshooting.

Additionally, new experiences and good solutions should be possible to add by using photos, videos or audio sequences, taken directly in the working situation, and to make it available online after an intra-company quality check.

Thus the thematic focus of the collaboration on the development of a learning medium for the optimization of vocational training had been extended towards a systematic collection and compilation of experiential knowledge of experts as well as the integration of vocational education and a technology based competence management.

So far, beyond the manuals of the manufacturers, the experiences and results from the maintenance tasks are only documented within reports. They usually contain the causes of failure and the results of troubleshooting, but nothing about the considerations that led to this approach of failure analysis and respective results. This means that the problem solving strategies of the experts are not documented within standardized documents of the maintenance procedures. They are more informally exchanged within consultation processes between professionals as part of oral explanations in vocational education.

With increasing technical maturity of the virtual interactive learning environment the claim was substantiated to include aspects of problem solving strategies to the design of further education and competence management with respect to content and didactics.

Of particular importance for the company was the cognition that the technology-related solutions from maintenance have partially led to technical improvements of the equipment or the development of specific tools and devices that were applied for a patent from time to time. That way, in cooperation between operators and manufacturers product improvements could be achieved.

Thus, the tension between the reference to proven certainties of professional experience and the acquisition of new knowledge was activated.

This way, the individual pool of expertise by means of “having experience” is connected with the need to integrate new findings from the work process in the sense of “making experience” into the personal range of actions.

This requirement also resulted from the permanent technical development of the new equipment that was applied. Here the professionals of the manufacturer, the operator and the maintenance companies have to learn the specifics of new technologies, materials and technical solutions and need to integrate this knowledge to theirs sphere of activity.

The companies are increasingly facing the challenge to integrate the expertise of longtime employed experts and the up-to-date factual knowledge of the junior employees in a way that equipment with a long product lifecycle can successfully be operated as well as new products and technologies can be added to the companies’ portfolio.

2 Experiences from the Industrial Use: Evaluation Method and Results

The decision to develop further technology based learning media should be based on the results of an evaluation that was conducted within the companies. So far, the learning applications could only be accessed by a small group of users. Empirical results about the acceptance, usability and the learning success that could be achieved by using the technology based learning environment, were missing.

Fig. 2 Learning situation during the evaluation process

The evaluation method and the results that could be achieved (Haase et al. 2012, 2013) will be described in the following paragraphs.

2.1 Evaluation

This section will explain the evaluation process in detail. It includes information about the participants, the measures and the procedure.

2.1.1 Participants

Twenty male trainees from two global electricity companies served as subjects in the current study. The average age was 36.6 years (SD = 10.68), ranging from 23 to 55 years of age. The study took place at a technological training center of the companies (see Fig. 2). All participants belonged to the maintenance staff of the companies. All trainees were intermediate to experienced computer users and had advanced technological knowledge in maintenance of high voltage equipment.

2.1.2 Measures

Many different training evaluation approaches have been published (Eseryel 2002). However, the best known and widely accepted model for the evaluation of training and learning was proposed by Kirkpatrick (Athanasou 1998). According to Kirkpatrick's evaluation training model (Kirkpatrick and Kirkpatrick 2006) the personal reactions of trainees as well as the resulting increase in technological knowledge are focused as training outcomes.

Examination of literature indicated enjoyment and perceived usefulness as two conceptually separate kinds of trainee's reactions (Alliger et al. 1997; Warr and Bunce 1995).

Enjoyment The extent to which trainees perceive a program as enjoyable is a central form of subjective evaluation (sometimes tapped by what organizations refer to as "happiness sheets" or "smiley sheets"). For the measurement of enjoyment a questionnaire comprising eight items was designed.

Perceived Usefulness This reaction to training focuses upon the potential applicability of the learning material in the trainee's job and thus is associated with changes in work behavior.

Learning The evaluation of learning has evolved as a global term with a variety of different methods of achievement measurement. The approaches vary as a function of educational objectives that might result from training. In the majority of cases, focus is upon the acquisition of cognitive outcomes (trainee knowledge), but training often addresses skill-based or attitudinal outcomes as well (Kraiger et al. 1993). However, in close coordination with the trainer staff it was decided that the learning should be assessed in the traditional way of a written examination. While it is appreciated that this may not be the best method of assessment, it is the method normally used within the companies, and therefore was realistic within the training environment. Consequently, a written test, focusing on the technological knowledge learnt, was given. The survey contained multiple choice questions as well as matching exercises whereby trainees had to correctly rank the answers they were provided with. The trainees were asked to complete 30 questions with an overall score ranging from 0 to 45 points. Points were not deducted for incorrect answers.

Usability With regard to trainee's educational objectives and to avoid any distraction to trainee's learning curve a summative evaluation was chosen. Accordingly, the VR application was evaluated after use by asking the VR trainees to complete the Software Usability Measurement Inventory (SUMI). SUMI is a rigorously tested and proven method of measuring software quality from the end user's point of view (Kirakowski and Corbett 1993). It yielded 50 ratings for affect, efficiency, learnability, control and helpfulness as well as giving an overall usability assessment. Several items were adapted to the needs of the investigation at hand.

2.1.3 Procedure

In order to evaluate the VR learning application a quasi-experimental pretest-posttest-follow-up control group design was chosen, with a group of trainees being taught traditionally by a trainer as the control group (TT) and a second equally sized group of trainees using the virtual reality application as the experimental group (VR). Hence, the trainees were assigned randomly to two groups of 10 persons each. In each training session participated five trainees and each one took 8 h of work. This applies for both conditions, the TT training and VR training as well.

While the TT group was taught in the usual way using Powerpoint presentation and a mockup of the operating mechanism, the VR group has used the virtual learning environment in order to learn the maintenance workflow. In the latter case the trainers' role changed from a teacher to a learning coach asking trainees to reason, integrate, and demonstrate knowledge, answering questions and giving expert feedback. For this purpose the coach received a special training to manage the new role.

Prior to the VR training session, the trainees were introduced to virtual reality. Within the tutorial the 3-D interface and the main interaction metaphors were explained.

Prior to the program (Time 1), trainees were asked to complete a questionnaire which tapped the previous technological knowledge (learning, see above) as well as demographic characteristics. Immediately after the training session (Time 2) the trainees were asked to complete a survey which addressed their initial reactions (see above) on the one hand and the acquired technological knowledge (learning) on the other hand. The items designed to tap enjoyment and perceived usefulness were intermixed in a single set of items. Additionally, participants of the VR group completed the SUMI. Approximately 3 weeks later (Time 3) the trainees were requested to answer the technological knowledge questionnaire again to assess knowledge retention over time.

2.2 Results

This section presents the results for the measurement criteria enjoyment, perceived usefulness, learning and usability.

2.2.1 Data Analysis

All reverse-worded items were recoded. For user satisfaction and perceived usefulness Cronbach's Alpha as coefficient of reliability was computed (Cronbach 1951). Afterwards, two separate two-way ANOVAs on user satisfaction and perceived usefulness with teaching method (TT, VR) and age (younger, older; median split) as factors were carried out.

To compare the learning outcome a three-way repeated measurement ANOVA with time of measurement (Time 1, Time 2, Time 3) as the repeated-measures factor and teaching method (TT, VR) as well as age (younger, older; median split) as between factors was estimated. When main effect was significant, Bonferroni corrected p-values are reported for post hoc comparisons (t-test, two-tailed).

Additionally, an independent t-test was used to determine whether there was a difference in usability rating between younger and older trainees.

The alpha level for tests of statistical significance was set to .05. For all ANOVAs, p-values were Greenhouse-Geisser corrected. For descriptive data mean and standard deviation are reported.

2.2.2 Enjoyment

For user satisfaction an internal reliability of $\alpha = .891$ was observed. A main effect of teaching method ($F_{(3,16)} = 5.38$, $p = .034$) was found. Thus, trainees who participated in VR are more satisfied with the seminar than members of the TT group.

2.2.3 Perceived Usefulness

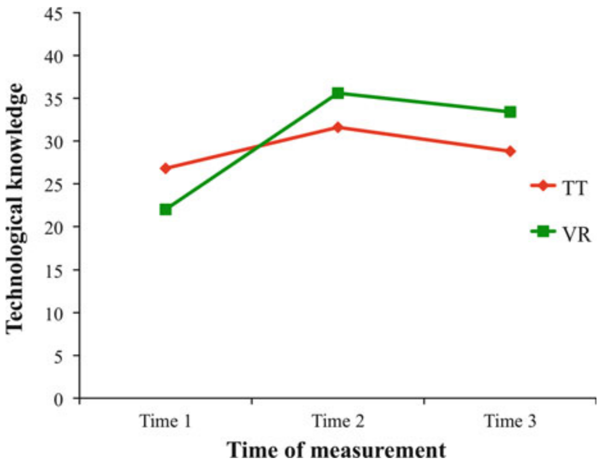
For perceived usefulness an internal reliability of $\alpha = .864$ was observed. An interaction between teaching method and trainees age emerged from the analysis, $F_{(3,16)} = 6.19$, $p = .024$. Age biases differences between the perceived usefulness of TT and VR. Younger trainees rated VR more useful for subsequent job performance, while older trainees assessed TT as more valuable.

2.2.4 Learning

A main effect of time of measurement, $F_{(2,32)} = 25.32$, $p < .001$ was revealed, due to improved technological knowledge at Time 2 ($M = 35.40$, $SD = 6.40$) compared to Time 1 ($M = 24.45$, $SD = 7.44$), $t_{(18)} = 6.2$, $p < .001$ and Time 3 ($M = 31.30$, $SD = 4.29$), $t_{(18)} = 3.11$, $p = .02$. The comparison between Time 1 and Time 3 is significant as well, $t_{(18)} = 4.43$, $p = .001$.

Besides, an interaction between time, teaching method and age emerged from the analysis, $F_{(2,32)} = 3.76$, $p = .039$. In detail, as Fig. 3 demonstrates, younger trainees benefited more from TT while older trainees achieved improved results due to VR (Time 2). The investigation of learning outcome over time (Time 3) shows enhanced performance for VR compared to TT. This applied to both older trainees (TT: $M = 28.80$, $SD = 4.27$; VR: $M = 33.40$, $SD = 3.56$) and younger trainees (TT: $M = 28.20$, $SD = 3.49$; VR: $M = 34.80$, $SD = 1.92$). Even though younger trainees learned more from TT at Time 2, the retrieve of technological knowledge (Time 3) was advanced for VR (see Figs. 3 and 4).

Fig. 3 Learning outcome of younger trainees depending on teaching method



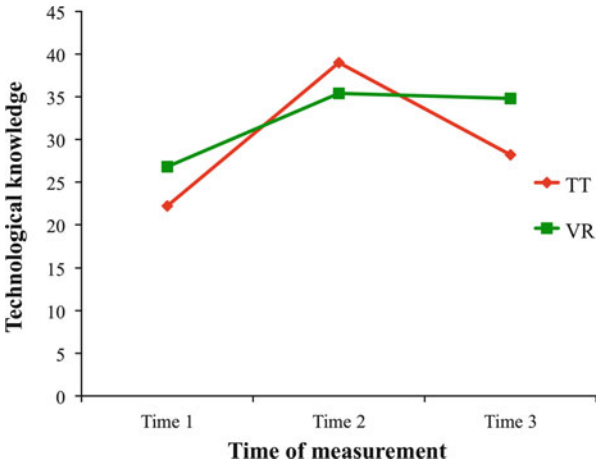


Fig. 4 Learning outcome of older trainees depending on teaching method

2.2.5 Usability

A comparison of usability as a function of age revealed no difference between the groups, $t_{(18)} = .661$, n.s. Both, younger trainees (M = 3.85, SD = .53) and older trainees (M = 3.67, SD = .27) assessed the usability of the virtual learning environment moderate positive.

2.3 Discussion of the Evaluation Results

The high acceptance and perceived usefulness among the experienced technicians can be explained by the connectivity to their professional experience. Technicians are aware of situations in which such a technology based training environment would have been of good help for solving a complex task or for a deeper understanding of the devices behavior. Within the training unsolved situations could be re-experienced and explanations could be found. This made the learning environment somehow more attractive for the more experienced technicians compared to the younger ones.

The novice and younger users are familiar with new media from theirs everyday life experience. Computer games, smartphones and 3D-cinema, as some examples, prove that 3D content is very common for younger users. The missing technological experience for the maintenance of high voltage equipment explains the fairly lower rating for the perceived usefulness.

3 Individual and Organizational Learning

In the description of the initial situation it was illustrated how on the one hand, the professionals need to acquire new knowledge and on the other hand, companies have to adapt their structures, processes and if necessary their strategies according to the changing conditions. The learning processes for the professionals that are required refer on the one hand to the technical characteristics of the equipment and the performance criteria and maintenance instructions that have changes. This kind of factual learning content can be explained in seminars using classical media like powerpoint slides, texts, images and photos or videos that are didactically designed.

On the other hand, maintenance is characterized by unforeseen and unpredictable situations that occur during the troubleshooting. Meeting these challenges requires from the professionals in the sense of a job profile as a "problem solver" (Fischer 2005, p. 312; Pahl and Herkner 2007) beyond well-founded specialist systematic knowledge for the assessment of operating conditions, measurement data or failure images also specific knowledge of peculiarities, irregularities or weaknesses in the systems. Due to the interconnectedness of systems, an interdisciplinary and holistic understanding of the upstream and downstream components and processes is also required (Drescher 2002).

In case of challenging maintenance tasks that require problem solving competence, the professionals integrate their systematic knowledge with the experience from the processing of difficult tasks they have solved before (Schaper and Sonntag 1997).

That means that further insights that rely on specific problem solving strategies and the cooperation with external partners are gained within the working process. In addition to a qualified analysis the correct conclusion from failure and damage reports or maintenance protocols have to be drawn. The professionals are requested to make decisions depending on the current situation and to make use of the respective allowance at discretion. Beyond the factual knowledge this kind of responsible and problem solving acting requires the action leading internalization of operational values, norms and rules. These influencing factors develop and solidify as tacit knowledge in particular in concrete work situations, regardless of any prepared learning settings. They have significant behavior-influencing effects, even if they are not consciously perceived or made explicit in any case (Neuweg 2004).

The following figure (Fig. 5) illustrates this relationship with respect to the development of vocational competence. Competency development takes place in professional working situations, mediated by the more or less reflection of experience.

In an article within a handbook of the sociology of work, Fritz Böhle refers to the relevance of experience-based and situated handling and experiential knowledge by using a number of empirical evidences. (Böhle 2010) He turns out that within the labor and industry sociological research the concept of experience-guided subjectifying work provides evidence for the role of sensual experience and the independent generation of experience as the basis of knowledge acquisition as well as the cognitive and practical interpretation of subjective factors like feelings and emotions (Böhle 2010, p. 160).

Especially in the context of current discussions about technological developments for networking and automation under the headline of "Industry 4.0", the conclusion

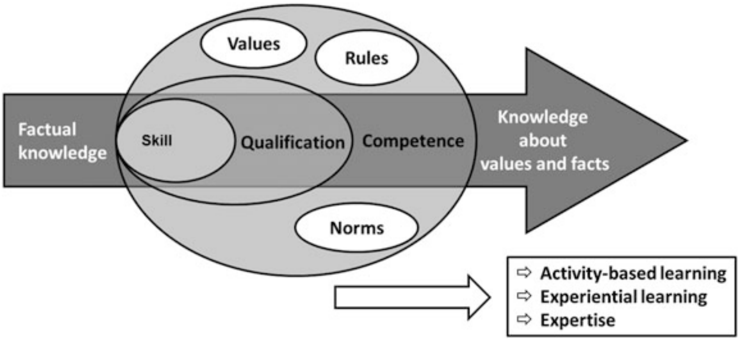


Fig. 5 Competence development in activity-based learning (Source: Own drawing following Erpenbeck and Sauter (2007), p. 69)

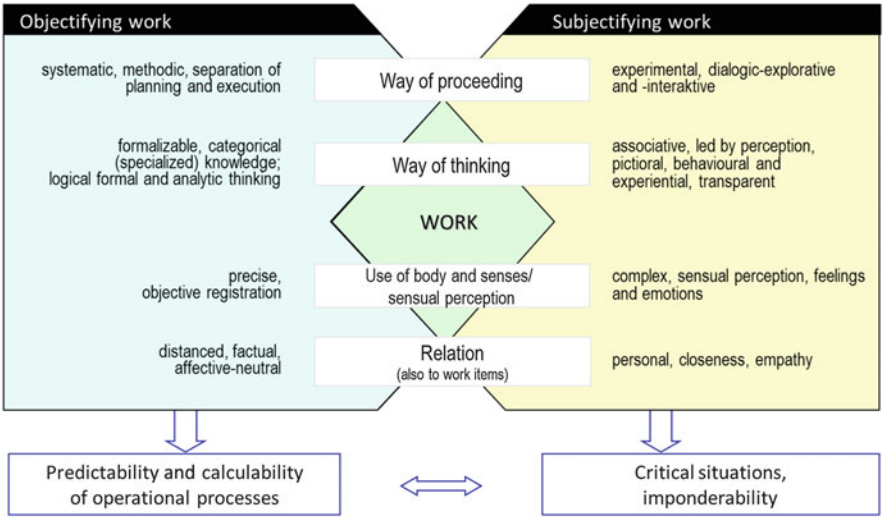


Fig. 6 Objectifying work and subjectifying work (Source: Own drawing following Böhle et al. (2011), p. 21)

that within this concept work items and systems are perceived as not completely calculable and controllable (Böhle 2010, p. 161).

Figure 6 provides an overview of the relation described.

This understanding of work is in substantial parts compatible with research results from ergonomics (Ahrens 2002) and vocational education (Schulze et al. 2002). Coping with problem situations or new tasks in operational practice is a boost for the acquisition of work process knowledge (Fischer 2005, p. 313). Therefore not only existing knowledge is used for the task accomplishment but during the process of problem solving, new insights and knowledge are generated (Fischer 2005).

These processes, conceived as learning, include both the individual acquisition processes and the individual growth of knowledge as well as the competence to apply this knowledge and—when meaningful—make it available in the organization by means of communicating and finally to anchor it within rules, instructions and routines.

Argyris and Schön (2008) formulate their understanding of this dynamic: "*Organizational knowledge occurs when individuals within an organization experience a problematic situation and inquire into it on the organization's behalf. They experience a surprising mismatch between expected and actual results of action and respond to that mismatch through a process of thought and further action that leads them to modify their images of organization or their understandings of organizational phenomena and to restructure their activities so as to bring outcomes and expectations into line, thereby changing organizational theory-in-use. In to become organizational, the learning that results from organizational inquiry must become embedded in the images of organization held in its members' minds and/or in the epimestological artifacts (the maps, memories, and programs) embedded in the organizational environment*." (p. 16).

In this sense Michael Dick understands organizational learning as the development of the ability of an organization to change appropriately with their environment (Dick 2008, p. 459). This means that the individual learning processes also set incentives and impetus to the corresponding development of operational products, structures, processes or routines. As indispensable feature of organizational learning, Dick emphasizes that this learning process is deliberately and in anticipated direction (Dick 2008, p. 459). In this respect he differs between organizational learning and organizational change. With the concept of learning a subject of learning that defines the intention and the direction of the learning process is always connected (Dick 2008, p. 459).

Insofar as learning processes develop "new cultural activity patterns" or "new forms of work activity", Engeström speaks of "expansive learning at work" (Engeström 2008). However, in the analysis of case studies he emphasizes that the significant triggering action for the expansive learning process can be found in the conflictive questioning of the usual practice (Engeström 2008, p. 85).

This conflict-related aspect has been confirmed in the operational practice that formed the basis for this paper and will be presented in the following discussion.

4 Discussion

The intention of improving vocational education by using technology based learning media could be confirmed with the first evaluations that were conducted.

In the course of the professional development of learning tasks for the virtual learning environments improvements of work processes could be identified in cooperation with the company experts. In some cases, rules and regulations of work safety have been adjusted.

The analysis of failures led to technical improvements of the equipment and their integration to the learning media several times.

In the development of the training concept for the training and support center the seminar offers were extended through new forms of systematic learning within the work process.

After participating at the seminars, trained technicians will be accompanied by "teaching mechanics" in the form of a supervisor. These professionals have conducted the specialized practical exercises during the seminars and will take the technical responsibility in the subsequent working process. After a successful "on-the-job-training" the responsibility will be transferred to the assigned technician. The disciplinary liability remains unaffected.

For the seminars including practical training, internal certificates are granted. Thus, the quality assurance and the technical and methodological standardization of tenders will be guaranteed.

According to the specific requirements for the problem solving competence of maintenance technicians, great value is placed on not only teaching factual knowledge but also the understanding of the overall technical context is certified.

The virtual reality based learning environments are constantly updated by adding new equipment and technical details. They serve in particular for the visualization of abstract and hard to see components, assembly groups and processes as well as for the documentation of complex troubleshooting.

They are on the one hand used as (self-) learning media in the seminar and on the other hand the documented troubleshootings and work procedures are used directly in the work process in case of unclear situations or for the preparation of work. Thus, they have also developed to assistance systems which are used by technicians with mobile systems.

The positive impetus for using virtual learning and working environments for vocational education have been initiated by the local specialists and executives of the company. The training activities in total could be brought to the external market, so that as a result a new division was established, together with the new building of an education center.

To that extent, one can speak of successful processes for both individual and organizational learning. In the operational practice, the ratio of individual and organizational learning processes is characterized by different, even divergent, influences and interests. The actors are required to communicate and negotiate the individual and group related developments with the interests of other divisions within the company as well as the specifications of the competent authorities. This development has not been without conflict.

In summary it can be stated that the partners included in this development process may have had in mind the now reached level of development as a vague possibility, or more or less a realistic vision.

In organizational theory processes and structures are called "emergent" that describe diachrone emergence phenomena whose results are unpredictable because the structure that influences the result is developed through the project (Schreyögg 2003, p. 417).

It was not a clearly formulated objective that was pursued rigorously and systematically. Instead, over the years several options have been arisen that could be realized successfully with respect to the current interests and power constellations of dedicated persons.

References

Ahrens D (2002) Erfahrungsbasiertes Wissen und experimentelles Lernen: Die Macht “unscharfen Wissens”. In: Jenewein K, Knauth P, Röben P, Zülch G (Hg) Kompetenzentwicklung in Arbeitsprozessen. Beiträge zur Konferenz der Arbeitsgemeinschaft gewerblich-technischer Wissenschaften und ihre Didaktiken in der Gesellschaft für Arbeitswissenschaften am 23./24. September 2002 in Karlsruhe. 1. Aufl. Nomos-Verl.-Ges., Baden-Baden (Bildung und Arbeitswelt, 9), pp 363–371

Alliger G, Tannenbaum S, Bennet W, Traver H, Shotland A (1997) A meta-analysis of the relations among training criteria. Personnel Psychol J 50:341–358

Argyris C, Schön DA (2008) Die Lernende Organisation. Grundlagen, Methode, Praxis. Aus dem Amerikanischen von Wolfgang Rhiel, 3rd edn. Schäffer-Poeschel, Stuttgart

Athanasou J (1998) A framework for evaluating the effectiveness of technology-assisted learning. Indus Commerc Train 30(3):96–103

Beuting J, Haase T, Termath W (2010) Qualifizierung von technischen Fachkräften in der Instandhaltung von Hochspannungsbetriebsmitteln mit Methoden der Virtual Reality. Lernen Lehren 25(97):26–33

Blümel E (2007) Stand und Entwicklungstrends des Einsatzes von VR/AR-Techniken für Qualifizierung und Training im Arbeitsprozess. In: Schäper M (Hg) Kompetenzentwicklung in realen und virtuellen Arbeitssystemen. Bericht zum 53. Kongress der Gesellschaft für Arbeitswissenschaft vom 28. Februar – 2. März 2007. [an der Otto-von Guericke-Universität Magdeburg, Institut für Berufs- und Betriebspädagogik und dem Fraunhofer Institut für Fabrikbetrieb und Automatisierung (IFF)]. GFA-Press, Dortmund, pp 241–244

Blümel E, Jenewein K, Schenk M (2010) Virtuelle Realitäten als Lernräume: Zum Einsatz von VR-Technologien im beruflichen Lernen. Lernen Lehren 25(97):6–13

Böhle F (2010) Arbeit als Handeln. In: Böhle F, Voß G, Wachtler G (Hg) Handbuch Arbeitssoziologie. Unter Mitarbeit von Anna Hoffmann. 1. Aufl. VS Verlag für Sozialwissenschaften, Wiesbaden, pp 151–176

Böhle F, Bolte A, Neumer J, Pfeiffer S, Porschen S, Ritter T et al (2011) Subjektivierendes Arbeitshandeln – “Nice to have” oder ein gesellschaftskritischer Blick auf “das Andere” der Verwertung? Arbeits- und Industriesoziologische Studien 4(2):16–26

Cronbach LJ (1951) Coefficient Alpha and the internal structure of tests. Psychometrika 16:297–334

Dick M (2008) Organisational learning. In: Rauner F, Maclean R (eds) Handbook of technical and vocational education and training research. Springer, Dordrecht, pp 458–466

Drescher E (2002) Arbeitsprozesswissen in der betrieblichen Instandhaltung. In: Fischer M, Rauner F (Hg) Lernfeld: Arbeitsprozess. Ein Studienbuch zur Kompetenzentwicklung von Fachkräften in gewerblich-technischen Aufgabenbereichen. 1. Aufl. Nomos-Verl.-Ges., Baden-Baden (Bildung und Arbeitswelt, 6), pp 255–272

Engeström Y (2008) Expansives Lernen in der Arbeitswelt. Für eine Neukonzeptionierung der Tätigkeitstheorie. In: Rosa L (Hg) Entwickelnde Arbeitsforschung. Die Tätigkeitstheorie in der Praxis. Unter Mitarbeit von Hartmut Giest und Georg Rückriem. 1. Aufl. lehmanns media, Berlin (International cultural-historical human sciences, 25), pp 61–89

Erpenbeck J, Sauter W (2007) Kompetenzentwicklung im Netz. New Blended Learning mit Web 2.0, 1st edn. Luchterhand - Wolters Kluwer Deutschland GmbH, Köln

Eseryel D (2002) Approaches to evaluation of training: theory & practice. Educ Technol Soc J 5 (2):93–98
Fischer M (2005) Arbeitsprozesswissen. In: Rauner F (ed) Handbuch Berufsbildungsforschung. Bertelsmann, Bielefeld, pp 307–315
Haase T, Termath-Bechstein W, Martsch M (2012) Virtual reality-based training for the maintenance of high voltage equipment. In: Urban B (Hg) eLearning Baltics 2012. Proceedings of the 5th International eLBa Science Conference in Rostock, Germany, June 21–22, 2012. Fraunhofer Verl., Stuttgart
Haase T, Termath W, Martsch M (2013) How to save expert knowledge for the organization: methods for collecting and documenting expert knowledge using virtual reality based learning environments. Proc Comput Sci 25:236–246
Kirakowski J, Corbett M (1993) SUMI: the software usability measurement inventory. Br J Educ Technol 24:210–212
Kirkpatrick DL, Kirkpatrick JD (2006) Evaluating training programs, 3rd edn. Berrett-Koehler Publishers, San Francisco, CA
Kraiger K, Ford J, Salas E (1993) Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. J Appl Psychol 78(2):311–328
Neuweg GH (2004) Könnerschaft und implizites Wissen. Zur lehr-lerntheoretischen Bedeutung der Erkenntnis- und Wissenstheorie Michael Polanyis. 3. Aufl. Waxmann, Münster. Online verfügbar unter http://www.worldcat.org/oclc/180002978
Pahl JP, Herkner V (2007) Rahmencurriculum Instandhaltung. 1. Aufl. Bertelsmann W. Verlag, s.l. Online verfügbar unter http://ebooks.ciando.com/book/index.cfm/bok_id/22569
Schaper N, Sonntag K (1997) Kognitive Trainingsmethoden zur Förderung diagnostischer Problemlösefähigkeiten. In: Sonntag K, Schaper N (Hg) Störungsmanagement und Diagnosekompetenz. Leistungskritisches Denken und Handeln in komplexen technischen Systemen. vdf, Hochsch.-Verl. an der ETH, Zürich (Mensch Technik Organisation - MTO, 13), pp 193–210
Schreyögg G (2003) Organisation. Grundlagen moderner Organisationsgestaltung; mit Fallstudien. 4., vollst. überarb. und erw. Gabler, Wiesbaden. Online verfügbar unter http://www.worldcat.org/oclc/66778502
Schulze H, Rose H, Witt H (2002) Erfahrungsgeleitete Arbeit in der Fertigung mit computergesteuerten Werkzeugmaschinen. In: Fischer M, Rauner F (Hg) Lernfeld: Arbeitsprozess. Ein Studienbuch zur Kompetenzentwicklung von Fachkräften in gewerblich-technischen Aufgabenbereichen. 1. Aufl. Nomos-Verl.-Ges., Baden-Baden (Bildung und Arbeitswelt, 6), pp 195–215
Warr P, Bunce D (1995) Trainee characteristics and the outcomes of open learning. Personnel Psychol 48:347–375

Designing Rooms for Virtual, Informal Communication: Reciprocal Awareness as a Central Criterion

Corsin Flepp, Michael Imhof, Gregory Meier, Thomas Ryser, Roger Burkhard, Hartmut Schulze, and Andreas Simon

"You See Me, When I Also See You"—Reciprocal Awareness as a Design Criterion for Rooms Providing Virtual, Informal Communication, using a Virtual Café as an Example

Abstract

The trend towards decentralized collaboration in companies leads to challenges for informal communication because spatial proximity is missing. This is a problem since informal communication is considered to be key for successful collaboration. Telepresence systems, which connect distant places, are potential solutions. However, little is known about which conditions are beneficial and which ones detrimental to informal communication. In this qualitative study, conditions which further informal communication, were examined in different virtual café settings. A method was developed which combined participatory design with a qualitative experiment. In the Usability Lab of the University of

C. Flepp (✉) • G. Meier • T. Ryser • R. Burkhard • H. Schulze
Institute for Research and Development of Collaborative Processes (ifk), School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, 4600 Olten, Switzerland
e-mail: corsin.flepp@fhnw.ch; gregory.meier@fhnw.ch; thomas.ryser@fhnw.ch; roger.burkhard@fhnw.ch; hartmut.schulze@fhnw.ch

M. Imhof
Institute for Higher Education in Applied Psychology (ZfA), School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland, Riggenbachstrasse 16, 4600 Olten, Switzerland
e-mail: michael.imhof@students.fhnw.ch

A. Simon
Institute of Experimental Design and Media Cultures (IX DM), School of Art and Design, University of Applied Sciences and Arts Northwestern Switzerland, Freilager-Platz 1, 4023 Basel, Switzerland
e-mail: andreas.simon@fhnw.ch

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_14

Arts and Sciences Northwestern Switzerland (FHNW), 19 people (N = 19) tried out various virtual café settings, analyzed requirements for optimization and subsequently tested them. At the same time, 20 group interviews were conducted and analyzed according to the principles of heuristic-detecting social research. Three subcategories which influence each other were identified as key results (awareness, privacy and control). These three subcategories need to be balanced when a virtual café (room and technology) is designed. Furthermore, encouraging (reciprocal) awareness could also be a possible solution.

Keywords
Informal communication • Participatory design • Qualitative experiment • Reciprocal awareness • Decentralized communication • Decentralized cooperation • Virtual communication • Computer-mediated communication

1 Introduction

According to Hrastinski (2010), informal communication is a key success factor for company performance. Between 80 and 90 % of all interpersonal interaction in the work place is informal (Kraut et al. 1990) and deals mostly with work related topics (von Bismarck et al. 1999b). Pentland's (2012) study outcomes show that cross-sector informal communication is recognized a key criterion for team performance. It is a significant factor which conveys organizational culture and knowledge and also encourages the loyalty and good will of employees (Fish et al. 1993). Kraut et al. (1990) as well as Coradi and Boutellier (2013) substantiate a clear connection between spatial proximity and the frequency of interaction. However, due to the ever expanding globalization, the trend towards decentralized collaboration is impeding the important factor for informal communication, spatial proximity (Schulze et al. 2014). The economic relevance of this impeded informal communication is shown by the fact that already in 2008, 47 % of the 4.8 million people in the Swiss workforce worked in companies with decentralized locations, and this trend is still rising (Brändle et al. 2010). In order to compensate for this decentralized distance, organizations have been forced to support their cooperative processes through cost-intense, internet-based media (e.g. Telepresence systems[1]) (Rack et al. 2011). The disadvantage is that such systems usually support the formal aspects of collaboration and not the informal ones (von Bismarck et al. 1999a).

Using technical media to satisfy the demands of informal communication in decentralized settings is challenging and, at the same time, solution potential. The use of audio- and video technologies to permanently connect geographically separate places can potentially solve the dilemma between the trend to cross-company, virtual cooperation and the importance of informal communication. The research

[1]Alternatively video transmission systems.

project, "The Development and Implementation of Places for Virtual, Informal Communication (OviK[2])", also deals with this topic which will be described in more detail in the following section.

2 Research Project "The Development and Implementation of Places for Virtual, Informal Communication (OviK)"

In this section, the overall project OviK will be presented. The current study, a sub-project of OviK, is explained in Sect. 3 and the following.

2.1 Project Description

OviK is a research project promoted by the Swiss Commission for Technology and Innovation (KTI) and by industrial partners (application partners). This project has run for 2 years and ended in November, 2016.

The project looked at how and under which conditions informal, virtual communication can be established or promoted among a company's different sites. Therefore, the application partners (see Sect. 2.3) set up a "room for virtual, informal communication" in two different sites of individual companies. In the pilot phase, the practical usage of the OviKs were introduced, tested, evaluated and optimized through an iterative developmental process (modifications included).

Before the begin of the project's pilot phase, inspections, interviews, diary studies, and observations (see Sect. 2.3) at companies of the application partners were conducted to assess the needs and to differentiate and define the demands on the OviKs. With this information, the first initial settings or rather, pilot scenarios (see Sect. 2.1.1ff.) were conceptually devised and prototypes developed.

In the following the various pilot scenarios will be briefly described. The three developed conceptual pilot scenarios were based on complex combinations of the aspects *motion*, *encounter possibilities* and *intimacy*, each of which was weighted differently in each case.

2.1.1 Pilot Scenario "OviK—Virtual Piazza"

The "OviK—Virtual Piazza" (see Fig. 1) is similar to a public place with the strong aspect of *motion* and transience, or rather they are transient places to linger with limited proximity. Such a place can be, for example, an entrance hall or transit area in organizations or companies with a welcoming- or lingering-areas. Such areas are being continuously expanded into places for informal encounters.

[2]The abbreviation for Places of Virtual, Informal Communication.

Fig. 1 Illustration of a Virtual Piazza, © CCTP Competence Centre for Typology and Planning in Architecture, 2015

2.1.2 Pilot Scenario "OviK—Virtual Extended Office"

The "OviK—Virtual Extended Office" (see Fig. 2) is a shared room for a group of employees who work and communicate virtually and decentralized (e.g. project work). Colleagues in the one place are meant to work with their colleagues in the other place as if they all sat together in the same room.

2.1.3 Pilot Scenario "OviK—Virtual Café"

The "OviK—Virtual Café" (see Fig. 3) is a semi-public room to increase the chance of serendipitous informal cross-company encounters. This is in contrast to the more formal video conferences. In addition, it is legitimized through a socially accepted attractor (cafeteria). Places where employees can relax, such as a cafeteria or something similar, can be used in this way.

The current study focuses on this pilot scenario "OviK—Virtual Café" (see Sect. 3ff.).

2.2 Project Goals

A central goal of the whole project is the initial development and testing of OviKs to promote informal virtual communication between geographically separate sites. The OviKs should have coordinated furnishings and video communication technologies. A further goal was to have the different companies and the application partners implement and use the developed OviKs throughout the project. Based on project experience, progress reports were written explaining the necessary requirements of informal, virtual communication.

Fig. 2 Illustration of a Virtual Extended Office, © CCTP Competence Centre for Typology and Planning in Architecture, 2015

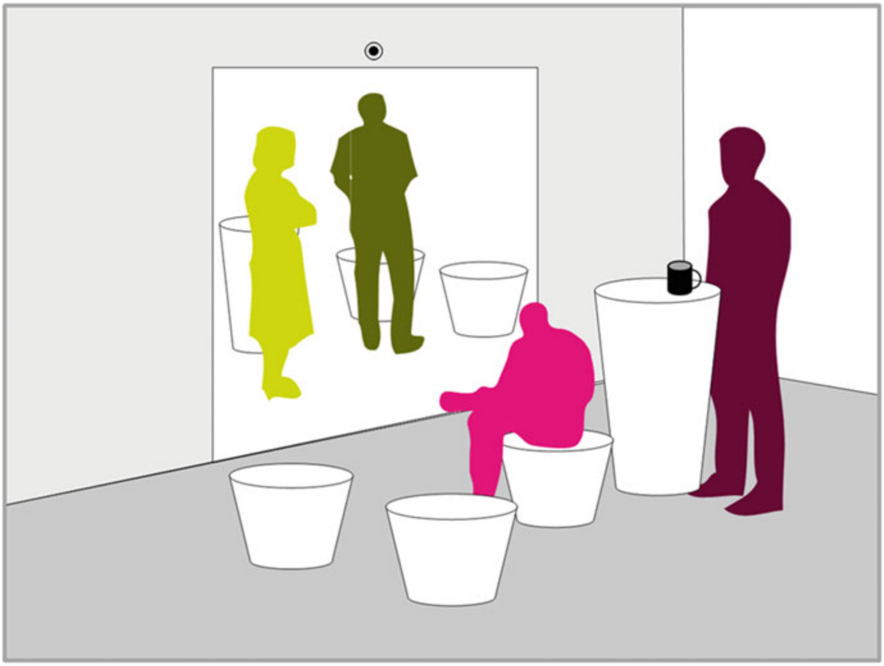


Fig. 3 Illustration of a Virtual Café, © CCTP Competence Centre for Typology and Planning in Architecture, 2015

After the project's end, a descriptive model of the evaluated OviKs is going to be conceptualised that will contain the relevant data and the parameters for design and equipment. In addition, an advisory concept for the *implementation*, *introduction* and *operation* of OviKs is planned.

2.3 Project Members

Several research partners participated in this project: the Institute for the Research and Development of Collaborative Processes (ifk) from the School of Applied Psychology and the Institute of Experimental Design and Media Cultures (IXDM) from the Academy of Art and Design, which are both parts of the University of Applied Sciences and Arts Northwest Switzerland (FHNW). Prof. Dr. Hartmut Schulze (Institute Director ifk) was the project manager and Andreas Simon (member of IXDM) the co-project manager. A further research partner was the Competence Centre for Typology and Planning in Architecture (CCTP), which is a part of the Lucerne University of Applied Sciences and Arts—the Lucerne School of Engineering and Architecture. These research partners were primarily responsible for the development and evaluation of the OviKs in close consultation with their implementation partners who were Vitra AG (home and office furnishings), Cisco Systems (Switzerland) GmbH (IT/Telecommunication). They provided the various video technologies and also the design components and furniture to develop the OviKs. In addition, so-called application partners were also involved in the project who were willing to use and actually test the first virtual initial settings. These application partners were: POST CH AG (transport and logistics), Trivadis AG (informatics) and SKAN AG (isolation and clean room technologies). Vitra AG and Cisco System (Switzerland) were also application partners.

In summary, it can be said that the project plan OviK seamlessly connected the furnishings with the video technological components, something which was new and innovative. The permanent audio-visual connection was the key medium linking the rooms. This connection could be used at any time by anyone without any further active interaction with the video technological components to get in touch with people in the other place. Another prerequisite was the integration of furniture appropriate for an OviK. The development of such interconnected components during the project was made possible thanks to the close cooperation of the two project partners, Vitra AG and Cisco System (Switzerland) GmbH.

Further information about the OviK project can be found on the following website: http://www.fhnw.ch/aps/ifk/projekte

3 Theoretical and Empirical Foundation

The following section provides the theoretical and empirical foundation to understand what is meant by informal- and by computer-mediated informal communication. In addition, a brief overview of other, studies related to virtual cafés will be given. Moreover, key success factors are presented which promote (computer-mediated) informal communication. The *situation awareness* model is also introduced to embed the analyzed subject in the literature.

3.1 Informal Communication

Kraut et al. (1990) describe informal communication as something that happens spontaneously. It is interactive and has many subjects ranging from work related to more private communications. They also worked out dimensions to differentiate informal from formal communication. Hrastinski (2010) complemented these dimensions, but focused more on computer-mediated informal communication. According to these dimensions, as seen in Fig. 4, computer-mediated informal communication is *unscheduled, interactive, spontaneous, optional, participant organized* and *experience-focused*. Further, it is characterized by *informal language* and generates *few costs* (Hrastinski 2010).

Now that computer-mediated informal communication has been explained, the virtual café and related literature will be presented.

3.2 Pilot Scenario "OviK—Virtual Café"

As stated before, a virtual café is a semi-public place that enhances informal cross-company encounters and can be differentiated from formal video conferences (see Sect. 2.1.3). Two cafés are connected virtually by means of a video broadcasting system (see Fig. 5) so that both cafés are linked together constantly with sounds and images (Tollmar et al. 1999).

communication:	formal	vs.	informal
dimension:	scheduled	vs.	unscheduled
	one-way	vs.	interactive
	present agenda	vs.	emergent agenda
	mandatory	vs.	optional
	authority-organised	vs.	participant-organised
	content focus	vs.	experience focus
	formal langage	vs.	informal language
	high cost	vs.	low cost

Fig. 4 The formal and informal dimensions of computer-mediated communication (adapted from Hrastinski (2010), p. 26)

Fig. 5 Virtual Café. *Left-hand side*: Illustration of a Virtual Café, © CCTP Competence Centre for Typology and Planning in Architecture, 2015. *Right-hand side*: Example of a Virtual Café in the Usability Lab of the School of Applied Psychology in Often, Switzerland, 2015

Since the 1990s occasionally studies have been carried out that have dealt with virtual cafés (e.g. Kraut et al. 1990; Tollmar et al. 1999). However, due to the technological shortcomings, the idea could not be then implemented (Schulze et al. 2014).

In addition to the knowledge what a computer-mediated communication and a virtual café is, research has shown that there are six success factors which promote computer-mediated informal communication. They will now be explained in the following section.

3.3 Success Factors for Computer-Mediated Informal Communication

Kraut et al. (1990) defined five prerequisites, respectively, success factors that promote computer-mediated informal communication. Schulze et al. (2014) later added a sixth factor, confidentiality regulation.

1. *Low personal cost:*

   The first of the six factor is *low personal cost* (cf. Kraut et al. 1990) which means that certain conditions have to be fulfilled which ease informal communication. Therefore, according to Kraut et al. (1990), as little effort as possible should be required to establish communication.
2. *Social Presence:*

   Social presence is the second factor (cf. Schulze et al. 2014). Informal communication is based on the certainty that someone is there. However, there is a difference between knowing about the others activities and the perception of the others in the same room (co-presence/telepresence) (cf. Kraut et al. 1990, quoted from Schulze et al. 2014). With computer-mediated informal communication, it is important to be able to judge whether potential communication partners are inside the other room and whether or not they want to communicate (cf. Kraut et al. 1990, quoted from Schulze et al. 2014).

3. *Abundance of the transferred Information:*

   The third factor is called the *abundance of the transferred information* (cf. Schulze et al. 2014). Kraut et al. (1990) concluded that in informal communication various senses should be involved and that the visual channel plays an important role when a conversation starts. In addition, the combination of a visual- with an auditory channel then ensures that communication can be conducted (cf. Kraut et al. 1990).

4. *Common working environment:*

   The *common working environment* and shared objects is the fourth factor (cf. Schulze et al. 2014) both of these elements are very important for information communication according to Kraut et al. (1990). Informal communication needs a common theme (cf. Al-Zubaidi and Stevens 2004).

5. *A concentration of suitable partners:*

   The next factor, the fifth, is *concentration of suitable partners* (cf. Kraut et al. 1990). A certain spatial proximity is necessary as well as a concentration of potential communication partners so that meeting someone spontaneously becomes possible (cf. Schulze et al. 2014). Computer-mediated informal communication could be supported by a system, which would find *appropriate communication partners* or initiate contact with other people (cf. Kraut et al. 1990, quoted from Schulze et al. 2014).

6. *Confidentiality Regulation:*

   The sixth and last factor is *confidentiality regulation* (cf. Schulze et al. 2014). Users of communication systems should have the possibility to regulate the level of confidentiality. This could be achieved by having a spatial separation: one zone to make contact and another one into which people could withdraw and have a private talk (cf. Schulze et al. 2014).

Having explained the six success factors, respectively, prerequisites for computer-mediated communication, the *situation awareness* model and its elements (e.g. *activity*) will now be presented. In this study, computer-mediated informal communication can be seen as an *activity*.

3.4 Theoretical Framework Model: Situation Awareness

Endsley's (1995) framework model of *situation awareness* (SA) is looked upon as a theoretical background and as an application-oriented concept within the context of complex working worlds (Schaub 2012). According to Schaub (2012), this model chiefly includes aspects of attention and consciousness. It must be added though, that in the SA model, the interaction between person and environment takes centre stage (cf. Endsley 1995). According to Schaub (2012), the model integrates

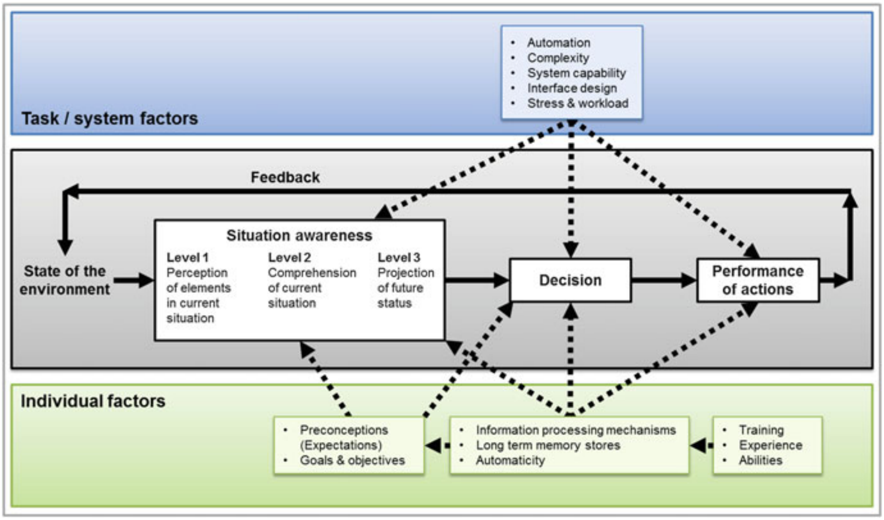


Fig. 6 Framework of situation awareness (adapted from Endsley (1995), p. 35)

different perspectives from "action theory" (cf. Hacker 1986), "natural decision making" (cf. Klein 1997; Orasanu and Salas 1993) and "complex problem solving" (cf. Dörner 1989; Funke 2003; Schaub 2009). For Endsley (1995), SA is to be understood as a level of knowledge which determines how much a person knows about a current situation. This knowledge level is composed of *perception, comprehension* and *projection* (see Fig. 6).

SA's three levels (*perception, comprehension* and *projection*) (see Fig. 6) can also be understood as higher cognitive[3] functions respectively processes, which, occur before a *decision* and can ultimately turn into concrete actions (Schaub 2012). According to Schaub (2012), the three levels can be described as follows (cf. Fig. 6):

1. The objects in the environment must first be *perceived* by the particular person. This level thus includes the *perception* of the condition, the characteristics and the dynamics of the relevant situational elements.
2. In a further step, the meaning of the situational elements must be understood by the particular person. This level describes the integration of the situational elements into a whole picture, which allows the person to *comprehend* the actual situation. Therefore, this integration causes the *comprehension* of the meaning of the individual situational elements.
3. In the third step, changes in the environment or in a future condition of the object can be *predicted,* respectively, *projected* by a person for a certain period of time.

[3]Cognition can be understood as a general term that involves conscious and unconscious mental processes (e.g. perception or thinking) (Wirtz 2013).

Expectations about the future behavior of the situational elements can be generated that are based on the achieved *comprehension* of the situation in the second level.

However, these three levels or rather, these higher cognitive functions can be influenced from two directions (see Fig. 6): from the particular person (individual factors) and from environmental factors (task and system factors). Additionally, fundamental cognitive resources, for example, *information processing mechanisms* as well as *long term memory stores* and *automaticity*, further influence more complex subprocesses such as an individual's *goals, expectations,* and *hypotheses* (Schaub 2012).

In summary, in order to informally communicate the making of a *decision,* respectively, *the performing of an action*, the *perception* of the situational elements is key. Due to the *perception* of the situational elements, their meaning can be understood. This is important because the particular person can then predict, respectively, *project* the development of the situation and the future condition of the actual surroundings elements, which in turn is key for *decision* making and taking *action* to accomplish *goals* (Schaub 2012).

4 Relevance and Question

Based on research (cf. Schulze et al. 2014), it is assumed that the technical infrastructure, such as monitor resolution or audio quality, and spatial aspects, such as room partitioning, influence the experience and the users' communication behavior in a virtual café.

With the current level of knowledge, it is not known yet which conditions are responsible for the initiation of informal communication in a virtual café. This points to a research gap, which should be addressed within this study. Therefore, the following research question has been formulated:

Question: Which conditions promote and/or hinder informal communication in a virtual café?

Investigating this question should bring new scientific knowledge and also be of practical use. As mentioned before (see Sect. 1), there is a trend today towards a cross-company collaboration, whereby this geographic separation of the employees endangers informal communication (cf. Schulze et al. 2014). This study should provide practical indications for the design and introduction of virtual cafés to reinforce the use computer-mediated communication.

In the following section, the study's methodology is explained.

5 Methodology

In spite of all efforts to the contrary, the study of informal communication in a virtual café can be considered to be unchartered territory. Therefore, a qualitative procedure was chosen (cf. Lamnek 2010). The reason for this decision was the following: The virtual café as a research object had to be first constructed by the users as there was no comprehension about it. Therefore, the users had to first experience a virtual café themselves and then try it out to become experts in regard to the subject.

Based on this reason or rather, this presupposition, a specifically modified iterative and qualitative procedure was developed: The approach of Kleining's (1986) *qualitative experiment* and the three steps of Spinuzzi's (2005) *participatory design*, *1. initial exploration of work*, *2. discovery process*, and *3. prototyping*, were adapted and combined. An overview of developed process will now be described (adapted from Spinuzzi 2005):

1. *Initial exploration of work:*

   The purpose of the first step was to let the authors of this study become comfortable with the way users collaborate, which included researching (literature research, field observations and expert interviews), analysis of the used technologies (e.g. various audio-visual conferencing systems) and the process of designing, introducing and using OviKs (e.g. behavioural routines: initiation of informal communication).
2. *Discovery process:*

   In the next step, the goals and the desired results of the users were clarified (e.g. through literature research, field observations and expert interviews). This was done based on the technical possibilities and on the interests and experiences of the partners who were involved in the research project. With this gained knowledge, a list with key dimensions was deduced. Based on the list different initial settings of virtual cafés were conceptualised. In the *qualitative experiment,* the dimension list together with the initial settings built the foundation for the systematic variation of the different aspects (e.g. put the coffee machine either inside or outside of the camera area, etc. see Fig. 7).
3. *Prototyping:*

   In this last step, the developed initial settings of a virtual café were implemented in the Usability Lab of the School of Applied Psychology in Olten. As defined by the *qualitative experiment* (cf. Kleining 1986), the object of research was observed from different perspectives and as a result, new psychological phenomena could be discovered. This was possible due to the characteristic variations of the various dimensions.

   The implementation of the *prototyping* took place within the framework of three 1-day workshops. 19 employees (users) of the research-, application- and implementation partners tested different complex virtual café settings (see Fig. 7). In a participatory way, the users could also revise the café setting requirements, modify them in order to try them out again. Each setting had a contact- and/or a retreat zone (see Fig. 7). At the same time, 20 group interviews were conducted.

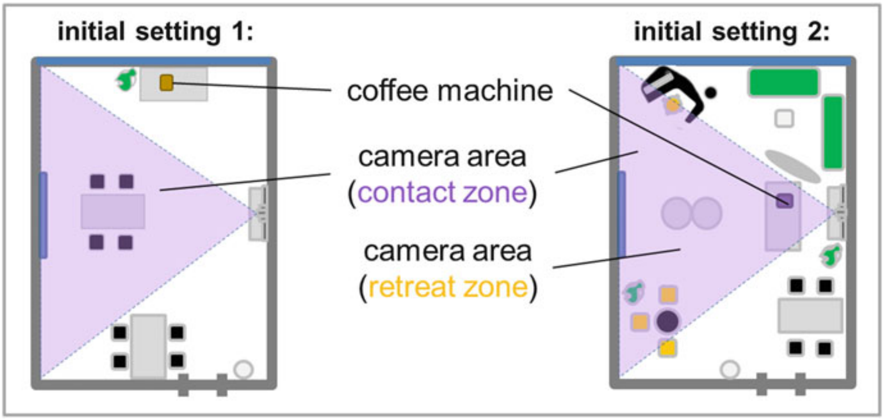


Fig. 7 Variations of initial settings, Usability Lab of the School of Applied Psychology in Olten, Switzerland, 2015. Initial setting 1 shows a possible variation of a virtual café. This setting has only a contact zone. In comparison, the more complex initial setting 2 has a retreat zone in addition to a contact zone for its users

Furthermore, the strategy of maximum variation was used for the sampling of participants (cf. Flick 2010). During the selection process, attention was paid to gender, sample size, hierarchical levels, relationships between the participants, organizational background and also experience with virtual cafés in order to have as many different perspectives as possible. The participants worked in the field IT, Sales, or Product Management, for example.

The results were analysed with the heuristic and detecting principles of Hagemann (2003). All coded statements were checked for double coding (those statements which could be assigned to several categories). Furthermore, the connection, respectively, the relationship between the identified categories was defined (cf. Hagemann 2003) (see Fig. 8).

6 Results

A psychological phenomenon (overarching construct, respectively, supercategory) showed itself to be the most relevant result for the virtual café. In this study, it is called *situational sense* (see Fig. 8) and consists of three subcategories:

1. *Privacy (37 mentions):*
   The feeling of being protected, of not being observed or monitored.

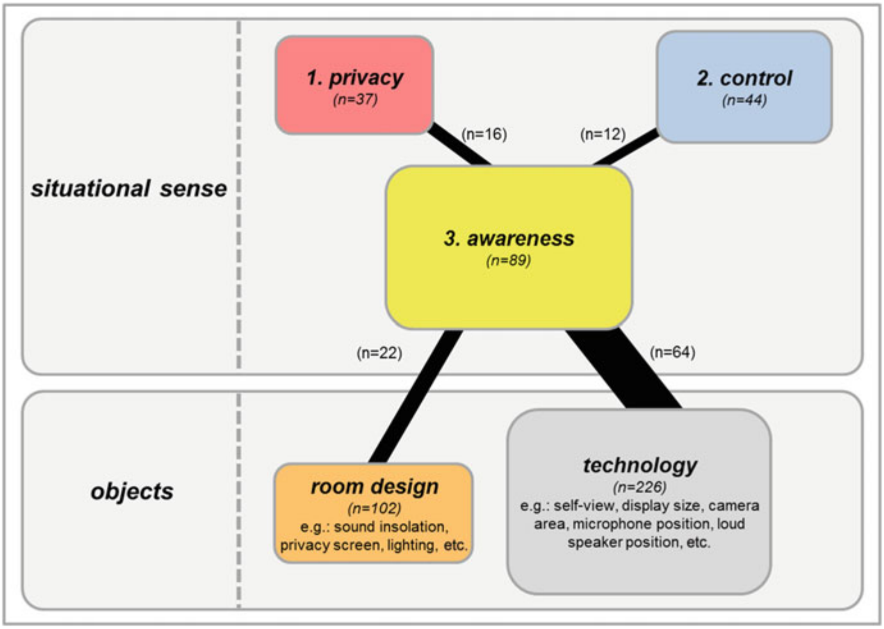


Fig. 8 Illustration of results. Frequency of connections (double coding) between the category awareness and the other categories as well as all mentions of the categories itself

2. *Control (44 mentions):*

   The feeling of being able to control the activities in a room and to use the technology and the room according to one's individual needs (e.g. being able to have a coffee, while not having to participate in the virtual café).
3. *Awareness (89 mentions):*

   Consciousness (also understood as a level of knowledge) of the own and the other room, containing the present people and the available objects. Moreover, it contains the consciousness of the visual and auditory area of present people as well as of the video transmission system.

Awareness develops because the users constantly revise their knowledge which includes these three different levels of knowledge:

a. *Knowledge of the other person (presence, identity, message conveyance) in one's own room and in the other one:*

   With this knowledge it becomes apparent who (identity) is in one's own room and who is in the other one and to what degree the conveyed message in the communication (i.e. intention, interest, and attention) can be understood as such by both parties.
b. *Knowledge of the object (attendance and use) in one's own room and in the other one:*

   The user knows which objects are in one's own room and in the other one and how to use them.

c. *Knowledge of the visual and auditory areas of one's own room and of the other room (present people and video transmission system):*

   The user knows how far the visual and the auditory area of present people and of the video transmission system extend in one's own room and in the other room (room height, width and depth). This includes the knowledge of the limits of the visual and the auditory area.

In addition, it is clear from the results that the supercategory (see Fig. 8) *situational sense* and the supercategory *objects* as well as its subcategories are frequently connected (double coding).

7 Discussion and Design Recommendations

Concerning the subcategories *privacy*, the use of the virtual café is being hindered if the users cannot do what they want (activities) without being observed (e.g. making or drinking a coffee). On the other hand, it is an advantage when the users feel they are unobserved and unheard. As for the subcategories *control*, things which could enhance computer-mediated informal communication are beneficial, especially when, for example, the volume of a loud speaker can be regulated in order to better hear the users in the other room. However, it can be a disadvantage if, the technical infrastructure in the other room can be changed. For example, if someone remotely changes the camera angle in the other room, then the people in that room could feel that their *privacy* has been invaded. Regarding the subcategories *awareness*, it is a disadvantage for the computer-mediated communication if the users do not have the possibility to find out who is in their room and who is in the other room and if people outside the visual area can still follow the conversation. In contrast, the knowledge of how far the visual and/or the auditory area stretches into one's own room and into the other room is helpful. Knowing how people perceive each other in their own room and in the other room also enables informal communication.

According to Hrastinski (2010) (see Sect. 3.1) computer-mediated communication differentiates itself from formal communication by the *lower costs*. This is also a success factor (Kraut et al. 1990). Based on the results (cf. Sect. 6), it can also be assumed that *low costs* are not possible if the users of both cafés do not develop *awareness*. This is because establishing a conversation is related with more effort (e.g. calling the attention of users of the other room). In reverse, an existing *awareness* of the users (v.i.z. users of both cafés develop *awareness*, thus it is *reciprocal*), more likely result in *lower costs* and in an informal communication.

Referring to Endsley's (1995) framework model of *situation awareness* (see Sect. 3.4), the authors suppose there is a large similarity to the phenomena *awareness* (see Sect. 6) since both represent a level of knowledge. The key difference between the two phenomena lies in the fact that *awareness* represents a specific level of knowledge in the context of virtual cafés, respectively, computer-mediated informal communication. Further, *reciprocal awareness* does not concern itself with the knowledge level of a single person, but with that of several people and their connected reciprocal interactions. For example, there would be no contact between people who do not know each other.

Furthermore, when the success factors are considered, some overlapping of computer-mediated communication (cf. Kraut et al. 1990; Schulze et al. 2014; see Sect. 3.3) and *awareness* can then be detected. For example, both *social presence* and subcategory a. of *awareness* (knowledge of the other person) (see Sect. 6) mention knowledge of other people's presence. The latter one differentiates itself from *social presence* in such a way that not only the presence, but also the identity and the message of the other person must be considered.

Awareness is the most important category of this study. When it was compared to other psychological phenomena, respectively, subcategories, *awareness* had the most connections (double codings, see Fig. 8) of all categories. Furthermore, *reciprocal awareness* facilitates computer-mediated informal communication. Therefore, we suggest adding a seventh factor, *reciprocal awareness*, to the existing prerequisites, respectively, success factors of computer-mediated informal communication (Kraut et al. 1990; Schulze et al. 2014; see Sect. 3.3). Based on the discovered connections (double codings), it can be concluded that the three subcategories (of the category *situational sense*) reciprocally influence each other and can indirectly promote or hinder computer-mediated informal communication. It was also shown that users would like a good deal of *control* over the knowledge and the presence of people in the other room (thus subcategory a. of *awareness*; see Sect. 6), which, on the other hand, limits the perceived *privacy* of the people in that room.

Therefore, when a virtual café is designed, the subcategories *awareness, privacy* and *control* should to be balanced by the spatial and technological aspects of a virtual café. For example: "You can see or hear me in the other room only when I can see or hear you". If the balancing is successful, an adequate perception of *privacy* and *control* can be achieved.

Furthermore, spatial and technical aspects should be adequately designed in order to support *reciprocal awareness* (knowledge categories a., b., c., see Sect. 6).

Despite the knowledge gained by this study more research is needed to better understand conditions which promote or hinder informal communication in distributed settings (e.g. virtual cafés). Interesting avenues for research could be the generalisation of the findings in regard to *reciprocal awareness* to other cross-company computer-mediated communication setting. Further research might examine the optimal level of *awareness*, *privacy* and *control* to promote computer-mediated informal communication.

References

Al-Zubaidi K, Stevens G (2004) CSCP at work. In: Keil-Slawik R, Selke H, Szwillus G (eds) Mensch & computer 2004: Allgegenwärtige Interaktion. Oldenbourg, Munich, pp 137–146

von Bismarck W-B, Held M, Maslo J (1999a) Anforderungen an ein System zur Unterstützung der informellen Kommunikation. Zeitschrift für Arbeitswissenschaft 1:10–17 von Bismarck W-B, Held M, Schütze H-J, Alex A (1999b) Die Verbreitung, Nutzung und Wahrnehmung von Medien in Organisationen. Mannheimer Beiträge zur Wirtschafts- und Organisationspsychologie 1:68–89

Brändle N, Etzensperger C, Künzi D, Vautier C (2010) Swiss Issues Branchen. Die Struktur der Schweizer Wirtschaft 1998–2020. Credit Suisse Zürich. http://www.ub.unibas.ch/digi/a125/sachdok/2010/BAU_1_5312089.pdf. Accessed 23 July 2016

Coradi A, Boutellier R (2013) Treffpunkt Kaffeemaschine: Begegnungsräume fördern kreative Interaktionen. Zeitschrift für Führung und Organisation zfo 82:322–328

Dörner D (1989) Die Logik des Misslingens. Rowohlt, Reinbek

Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors 37:32–64

Fish RS, Kraut RE, Root RW, Rice RE (1993) Video as a technology for informal communication. Communications of the ACM 36:48–61, http://www.comm.ucsb.edu/faculty/rrice/A42FishKrautRootRice1993.pdf. Accessed 23 July 2016

Flick U (2010) Qualitative Sozialforschung: Eine Einführung, 3rd edn. Rowohlt Taschenbuch Verlag, Reinbek

Funke J (2003) Problemlösendes Denken. Kohlhammer, Stuttgart

Hacker W (1986) Arbeitspsychologie. VEB Deutscher Verlag der Wissenschaften, Berlin

Hagemann O (2003) Qualitativ-heuristische Methodologie im Lehr-Dialog. In: Hagemann O, Krotz F (eds) Suchen und Entdecken: Beiträge zu Ehren von Gerhard Kleining. Rhombos-Verlag, Berlin

Hrastinski S (2010) The informal and formal dimensions of computer-mediated communication: a model. Int J Netw Virtual Organ 7:23–38

Klein G (1997) The current status of the naturalistic decision making framework. In: Flin R, Salas E, Strub M, Martin L (eds) Decision making under stress: emerging themes and applications. Ashgate, Aldershot, pp 11–28

Kleining G (1986) Das qualitative Experiment. Kölner Zeitschrift für Soziologie und Sozialpsychologie 38:724–749

Kraut RE, Fish RS, Root RW, Chalfonte BL (1990) Informal communication in organizations: form, function, and technology. In: Oskamp S, Spacapan S (eds) Human reactions to technology: the Claremont symposium on applied social psychology. Sage, Beverly Hills, CA, pp 145–199

Lamnek S (2010) Qualitative Sozialforschung, 5th edn. Beltz Psychologie Verlags Union, Weinheim

Orasanu J, Salas E (1993) Team decision making in complex environments. In: Klein G, Orasanu J, Calderwood R, Zsambok CE (eds) Decision making in action: models and methods. Ablex, Norwood, NY, pp 327–346

Pentland A (2012) The new science of building great teams. Harv Bus Rev 90(4):60–70

Rack O, Konradt U, Clases C (2011) Virtuelle und globale Kooperation. Editorial zum Themenheft. Wirtschaftspsychologie 13(3):41–51

Schaub H (2009) Der Soldat als Human Factor - Störfaktor moderner Rüstungsprojekte? (Gneisenau Blätter: „Soldat und digitales Schlachtfeld“, vol 8). Gneisenau-Gesellschaft der OSLw e.V., o. O.

Schaub H (2012) Wahrnehmung, Aufmerksamkeit und „Situation Awareness“ (SA). In: Badke-Schaub P, Hofinger G, Lauche K (eds) Human factors: Psychologie sicheren Handelns in Risikobranchen. Springer, Berlin, pp 64–81

Schulze H, Burkhard R, Knöpfli D, Mateescu M, Ryser T (2014) Das virtuelle Café. Ein Ansatz zur Förderung computer-vermittelter informeller Kommunikation. In: Skerlak T, Kaufmann H, Bachmann G (eds) Lernumgebung an der Hochschule. Auf dem Weg zum Campus von morgen, vol 66, Medien in der Wissenschaft. Waxmann, Münster, NY, pp 237–260

Spinuzzi C (2005) The methodology of participatory design. Tech Commun 52:163–174

Tollmar K, Chincholle D, Klasson B, Stephanson T (1999) VideoCafé–virtual Espresso-Cafés and semi-located communities. *Technical report, TRITA-NA-D9905, CID Report CID-47*. KTH, Stockholm

Wirtz MA (ed) (2013) Dorsch - Lexikon der Psychologie, 16th edn. Hans Huber Verlag, Bern

Analysis of the Stress and Strain of Repetitive Assembly Tasks

Markus Jürgen Heidl and Alexandre Boespflug

Abstract

Short cyclical work tasks occur with associated risks such as: overload through repetition, movement frequencies being too high, handling excessive weights or excessive forces. Therefore, we will thoroughly examine the workplaces of a German industrial company where repetitive assembly tasks are carried out. This study, with 249 questionnaires, shows that the highest perceived stresses lie mostly with posture during work and work organization. Pain in the upper extremities, which are used all the time, is less frequent. Therefore, the focus of assessment for such working tasks must be shifted to assessment of the whole body.

Keywords

Repetitiveness • Assembly • Stress • Strain • Assessment

1 Introduction and Aims

There are several methods, with different approaches and itemization, for the assessment and evaluation of physical workload and stress in manual tasks. But there are restrictions: none of the existing methods have yet been validated for short cycle activities of 30 s or less.

The outcome of two surveys (see state of research for further details) reveals that the results of the different methods of assessment vary. Hence it is unclear which of the existing methods should be chosen in practice to get conclusive results regarding the intensity of the workload and the implications of this intensity for the analyzed workplace.

M.J. Heidl (✉) • A. Boespflug
C/MPS3, Robert Bosch GmbH, Postfach 30 02 20, 70442 Stuttgart, Germany
e-mail: Markus.Heidl@de.bosch.com; Alexandre.Boespflug@de.bosch.com

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_15

But to be able to validate those existing methods for repetitive assembly tasks the level of stress and strain of the employees working regularly on such workplaces must have been quantified. To do so, as part of this study and as a first step, 249 associates have been interviewed with a questionnaire and 66 different workplaces have been analyzed. The results are presented in this paper. As a next step, the statistical analysis of these results will show potentials in the methods as well as suggestions as to which method is the best one to use for the assessment of repetitive assembly tasks.

2 State of Research

A significant increase in productivity and a reduction in costs was achieved in recent years through great progress in the field of automation of production process. Yet automation is often not cost effective because of short innovation cycles and high investment demand. The advantage of manual production in manufacturing lies in higher flexibility and skills of employees. However, in order to plan long term with the staff, the assembly tasks must be executable long into an employees' career without causing long-term health conditions.

In order to detect and prevent improper biomechanical stress in time, an explicit ergonomic evaluation of the workplaces is needed, preferably in the planning phase. There are a variety of methods available, for instance OCRA checklist (CL), ISO 11228-3, DIN EN 1005-5, RULA or EAWS, most of which have been theoretically and generally compared by Takala et al. (2010) and Steinberg et al. (2011). The assessment methods show a high convergent validity in these theoretical precomparisons. In particular, it was found earlier (Steinberg et al. 1998) that the different methods show similar results in average conditions. The results differ increasingly only in an extreme expression of individual characteristics.

Within the manual assembly tasks analyzed in the study, the repetitive work provides an extreme expression of individual characteristics when the activity rate is elevated. Silverstein et al. (1986) propose that manual activity be described as highly repetitive when a continuously repeating cycle takes less than 30 s. This limit also appears as an approximate assessment value in the evaluation of machine related repetitive work in the standard DIN EN 1005-5 (2007). Studies in which the results of the methods are compared with each other based on real, repetitive (with cycle times of 30 s or less) workplaces, show large differences in the assessments (Bruder and Kaiser 2012; Heidl 2013). These differ so much, that the same activity is classified in a green zone in a first method, in a yellow in a second and in a red zone in a third method.

Although effects from warming up (Debitz 2005) and training (Joiko et al. 2010) can be observed, experts say that frequent similar movements in the hand/arm area can lead to a variety of ailments. Hartmann et al. (2013) mention styloiditis, carpal tunnel syndrome and tendonitis as examples of diseases that result from repetitive actions. It makes sense in theory, but is there a difference in practice? Employers must be keen to ensure that the workload can be assessed realistically during all manual assembly tasks.

Wittig et al. (2013), in their study of the effects of repetitive jobs, provide statistics regarding the occurrence of complaints and pain sensations, as they collected data of the complaints in the back, neck, shoulders, arms, hands and fingers of volunteers of different professions. Kotzab (2014) and Steinberg et al. (2011) also investigated the complaint perception of production employees and included the back, neck, shoulders, elbows, forearms and hands. The exact values are given in the discussion.

3 Method

To create a database for stress and strain in repetitive assembly work, 66 workplaces in five different plants of a German automotive supplier were recorded and evaluated. Therefore, all default values, such as exercise frequency, weights, distances, breaks etc. were taken.

In addition, a total of 249 subjects, who regularly work in the selected workplaces, were asked about their subjective perceived exertion and their complaints. For this, a questionnaire was composed, which included the following matters:

– General information on the person, such as age, gender, handedness, working hours, work experience and additional sport load on the upper extremities
– A BORG-scale (Borg 2004) for the self-assessment of stress and strain of the daily workload and the specific load of the selected tasks
– Selected features of the questionnaire for the subjective assessment of workplace exposure of Slesina (1987, 2009) (see Figs. 1 and 3)
– An expanded form of the Nordic questionnaire (following Kuorinka (1987), see Fig. 2) including a detailed form of the upper extremities

<table>
<tr><th colspan="8">Fragebogen nach Slesina</th></tr>
<tr><td colspan="8">Bitte schätzen Sie die Arbeitsbelastungen Ihrer derzeitig ausgeübten beruflichen Tätigkeit ein; kreuzen Sie hierzu die entsprechenden Felder an. Wichtig ist dabei ausschließlich Ihr persönliches Empfinden!</td></tr>
<tr><td rowspan="2">Bitte überlegen Sie, ob folgende Merkmale oder Belastungsfaktoren an Ihrem Arbeitsplatz vorkommen!</td><td colspan="4">Wie häufig oder wie stark trifft dieses Merkmal oder der Faktor bezogen auf eine Schicht auf Ihre Arbeit zu?</td><td rowspan="2"></td><td colspan="2">Fühlen Sie sich selbst dadurch körperlich oder geistig belastet oder beansprucht?</td></tr>
<tr><td>oft (>50%)</td><td>mittel (25-50%)</td><td>selten (0-25%)</td><td>nie</td><td>ja</td><td>nein</td></tr>
<tr><td>Beispiel: Lärm</td><td>X</td><td></td><td></td><td></td><td></td><td>X</td><td></td></tr>
</table>

Fig. 1 Header of the questionnaire for the subjective assessment of workplace exposure

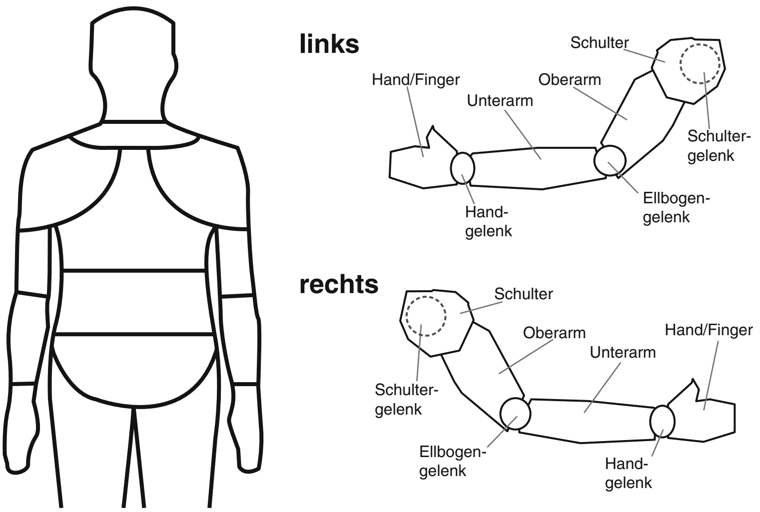


Fig. 2 Excerpt of the Nordic questionnaire, expanded by more detailed upper extremities

The response to the questionnaire was voluntary and anonymous. After a brief introduction the forms were filled in independently by the subjects. The investigator was nearby, in case of questions.

The mental aspects were left aside in the subjective assessment of workplace exposure. The questionnaire, as well as the study, was limited to the assessment of physical stress and strain, therefore the questionnaire of Slesina was reduced to a total of 29 features. Interviewing remained the same, both the occurrence as well as the subjective stress/strain was requested.

Once the analysis is completed, the selected workplaces will be evaluated with the assessment methods and the results of those will be compared. In recent studies (Heidl 2013) the methods EAWS (Ergonomic Assessment Worksheet), OCRA checklist (Occupational Risk Assessment), LMM MA (Key Indicator Method for assessing physical workload during Manual Handling Operations), HARM (Hand Arm Risk Assessment) and SI (Strain Index) have shown the best results with repetitive assembly tasks and therefore were selected to analyze the workplaces in this study.

4 Results

The sample of surveyed employees consists of 249 production staff from five different plants of a German automotive supplier. The proportion of female subjects is 58.6 %. Furthermore, 60.6 % of the respondents are younger than 50. On average, the interviewed employees have worked in their current roles on the assembly lines for at least 5.54 years (SD: 5.64) and have already been working for 13.16 years (SD: 10.28). 31.7 % of the subjects, in their spare time, perform sport or exercise such as weight training, martial arts, handball, etc. that involved the upper body. In addition to the physically strenuous work, 18.9 % did some upper body exercise or sport in the past.

A total of 66 different work tasks with cycle times between 5 and 64 s were examined. There are six workstations with cycle times of more than 30 s because several work pieces (up to ten) are handled within one cycle. The time per piece therefore is much shorter. The overall average of cycle time amounts to 19.68 s (SD 11.78).

The occurrence and the perceived strain of the 29 items selected for the questionnaire for the subjective assessment of workplace exposure are shown in Fig. 3. The items are sorted by the perceived strain which is shown in red. The posture that causes the most perceived strain is marked as "unfavorable body posture" (73.5 %) and "forced body posture" (59.3 %), and these are the items with the highest scores. Work organization also seems to have an impact on the perceived strain as "shift work" and "weekend working" are rated at 58.2 % and 57.4 % respectively. Next is "standing" (55.7 %), followed by "monotonous work" and "cycle-dependent work" which all are rated over 50 % as causing "strain" by the surveyed employees.

The frequency of the occurrence of strain symptoms can also be seen in Fig. 3. Where employees are concerned the most prevalent is "shift work" (91.4 %) and "cycle-dependent work" (85.4 %). Other frequent items are "dependent on the unit speed", "accurate viewing of details", "standing" and "weekend working" with 77.7 %, 76.4 %, 74.8 % and 74.7 % respectively. "Hand skill" (72.9 %), "time coping" (70.4 %) and "concentration" (70.2 %) were also rated as occurring over 70 % of the working time.

As explained in Sect. 3, the employees have also been questioned about the occurrence of complaints and pain in the upper extremities as well as neck and back. Only 18.07 % of the subjects stated that they were free of pain. The majority

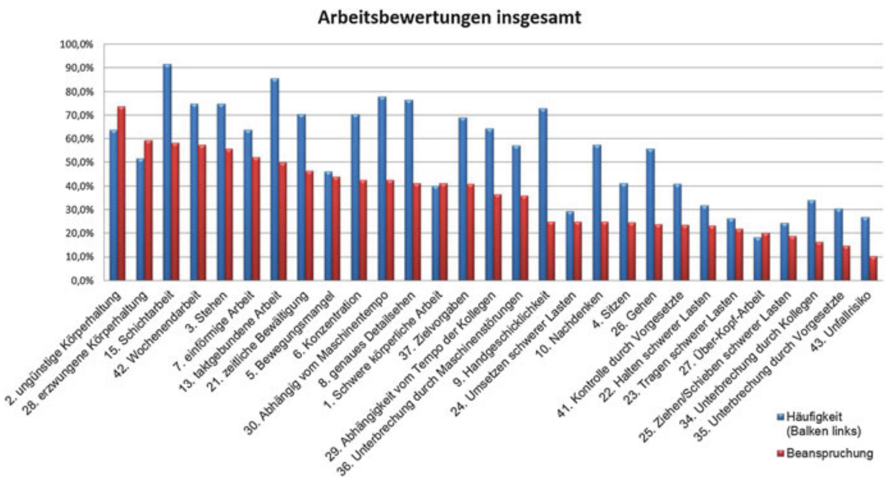


Fig. 3 Occurrence and perceived strain of the 29 selected items of the questionnaire for the subjective assessment of workplace exposure

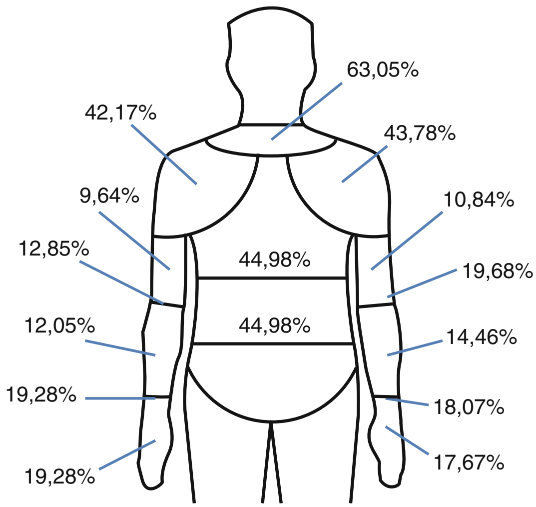


Fig. 4 The occurrence of complaints and pain in the different body regions, n = 249

had complaints in the neck, the back and/or the shoulders. A summary is shown in Fig. 4: 63.05 % of the subjects complain about the neck, another 44.98 % indicate pain in the upper as well as the lower back. Over 40 % of the employees feel pain in the left (42.17 %) and right (43.78 %) shoulder. The upper extremities are mentioned much less. The complaints and pains range from 9.64 % (left upper arm) to 19.68 % (right elbow).

The subjective rating of stress and strain via the CR10 scale of BORG provides mean values of low 1.33 to high 8. In fact, 5 workplaces are rated between "nothing at all" and "weak (light)" (0–2), 22 workplaces were considered "weak" to "moderate" (>2–3), 16 workplaces were classified between "moderate" and "somewhat strong" (>3–4), 15 places were categorized as "somewhat strong" to "strong (heavy)" (>4–5), and 8 places were rated higher than "strong" (>5–10). The overall average is calculated to be 3.64 (SD 1.50).

Finally, the selected assembly tasks were analyzed using five different assessment methods: EAWS, OCRA CL, LMM MA, SI and HARM. EAWS was used in two different versions (1.3.2c and 1.3.3) in order to see which implications are caused by the changes in praxis. Each assembly task was analyzed by two different assessors to reduce the influence of personal perception. The results can be seen in Table 1.

As you can see from the results above the assessors came to similar results in each area (green, yellow and red). It is only in the Strain index that the assessors' results are different. The small differences in results for the other methods can be explained by rounding and subjective interpretation (e.g. number of actions, grip conditions).

Furthermore, the new version of EAWS seems to be stricter than version 1.3.2c as less assessment results tend to be in the green zone and more tend to the red. Nevertheless, EAWS rates most assembly tasks, as well as LMM MA and HARM,

Table 1 Ratings of the five different assessment methods for the 66 different assembly tasks

	EAWS 1.3.2c		EAWS 1.3.3		OCRA CL		LMM MA		SI		HARM	
	A	B	A	B	A	B	A	B	A	B	A	B
Green	41	41	35	39	11	18	45	41	11	19	25	26
Yellow	18	22	22	20	13	13	21	25	23	26	38	39
Red	7	3	9	7	42	35	0	0	32	21	3	1

Each task has been analyzed by two different assessors: A and B

in between green and yellow. It is only OCRA CL and SI which are stricter. With OCRA, most work tasks are considered as red.

5 Discussion and Outlook

As already assumed in Sect. 2 (state of research), this study confirms the diverse results of the different assessment methods. It is, therefore, not obvious which method is best to evaluate the stress and strain of repetitive assembly tasks. Nor is it clear which one to choose in praxis.

Moreover, the assessment methods for "repetitive work" focus on overload through repetition, movement frequencies and/or on grip conditions. Evidenced by the study, this shows that this is an incomplete perspective.

As this study (with 249 subjects) was able to show, the highest perceived strain was not in the arms, but rather in the entire body. Particularly the neck, the back and the shoulders are stressed, a load that also then extends to the arms, hands and fingers.

Equally unexpected were the results of the 29 selected items of the questionnaire for the subjective assessment of workplace exposure. The perceived items causing the most strain were by far the different forms of body posture. "Standing" and "lack of exercise" were also rated as highly straining. Next are items of work organization such as "shift-" and "weekend-working", followed by items typically associated with repetitive work such as "monotonous work", "cycle-dependent work" and "time coping".

There are similar studies regarding the occurrence of complaints and pains in the area of environment assembly tasks. The comparison can be seen in Table 2. It shows the percentile values of complaints and pains as well as the difference between them and this study (bracketed).

The results from complaints and pains of the upper back are lower, but results from the lower back are higher in comparison. Furthermore, in Kotzab's study (2014) the complaints in the neck and shoulders are smaller, but higher for the wrists. Again, the pains in the elbows are in the same range.

Wittig et al. (2013) did a survey on behalf of the Federal Institute for Occupational Safety and Health (BAuA) with 15,239 employees of the working population.

Table 2 Comparison of complaint- and pain-values of different body parts of three different studies

	Heidl and Boespflug (2016)	Wittig et al. (2013)	Steinberg et al. (2011)	Kotzab (2014)
Upper back	44.98 %			32.37 % (−12.61)
Lower back	44.98 %	46.37 % (+1.39)	35.19 % (−9.79)	57.81 % (+12.83)
Neck	63.05 %		55.84 % (−7.21)	43.16 % (−19.90)
Shoulders	55.42 %		41.30 % (−14.12)	41.22 % (−14.20)
Neck and shoulders	73.90 %	47.52 % (−26.38)		
Arms	44.58 %	20.95 % (−23.63)		
Elbows	21.69 %		18.95 % (−2.74)	22.71 % (+1.02)
Wrists	26.10 %			36.07 % (+9.97)
Hands/fingers	26.10 %	15.47 % (−10.63)		
Forearms and hands	43.37 %		27.76 % (−15.61)	

The subjects were distributed over all working sectors with the proportion of interviewees from "manufacturing" at 32 %. Compared to the subjects of our study, who are working continuously in repetitive assembly tasks, complaints of the lower back are approximately the same, but pains in neck and shoulders, and in the arms and the hands/fingers are considerably less.

Another comparison comes from the data of the researches of Kotzab (2014) and Steinberg et al. (2011), who recorded complaints and pains of production staff. Therefore, the comparison to these two studies should even be more interesting.

Steinberg et al. (2011) interviewed 1293 employees, of which approximately half perform manual assembly tasks. The other half of the interviewees do office work. The prevalence of symptoms in the musculoskeletal system is specified for the entire study population. The complaints of the population are continuously lower than those of this study. Complaints centered around the elbows are in approximately the same range, but for the other body regions there are major differences. In this study, complaints and pains of the back (by 10 %) and neck (by 7 %) are considerably higher than for the interviewees of Steinberg et al. (2011). The difference for the shoulders, forearms and hands is even higher, as they differ by as much as 15 %.

Kotzab (2014) on the other hand studied three different assembly lines with cycle times of 25–84 s and interviewed 621 employees of these lines. Therefore, Kotzab's study is most similar to this one and, therefore, the results mostly comparable. The results from complaints and pains of the upper back are lower, but results from the lower back are higher in comparison. Furthermore, in Kotzab's study (2014) the complaints in the neck and shoulders are smaller, but higher for the wrists. Again, the pains in the elbows are in the same range.

In contrast with the results of this study and the comparison to other studies, the previous approach of the assessment methods focuses on the movement frequency in combination with force and grip conditions, as well as the joint position during the assembly tasks. Even though in some assessment methods the duration of the tasks and the breaks are already established, however there is more that should be done.

In future, repetitive tasks should be approached holistically. There are, by way of example, numerous work places where the employee has to be standing, which leads, as has been proven (amongst others by Garcia et al. 2015), to discomfort in the legs and back. But standing is only considered in EAWS, and when considering standing they are only considering the strain on the whole body, which lies in a different area than the one (Sect. 4) regarding repetitive loads. Moreover, the neck, which causes complaints in most of the interviewees, is not considered in any of the assessment methods. Those and other stressors of the whole body should be considered in the assessment of repetitive working tasks.

As a next step, all the selected working tasks will be analyzed in detail. They will be statistically evaluated as to which method is best for the assessment of manual repetitive working tasks. Thereupon suggestions will be worked out as to how to adapt this method to optimize the results.

References

Borg G (2004) Anstrengungsempfinden und körperliche Aktivität. Dtsch Arztebl 101:A1016–A1021 [Heft 15]

Bruder R, Kaiser B (2012) Abschlussbericht der Studie Arbeitsplatzgestaltung bei repetitiven Tätigkeiten. Robert Bosch Fahrzeugelektrik Eisenach GmbH, Institut für Arbeitswissenschaft der TU Darmstadt, Darmstadt, 03.12.2012

Debitz U (2005) Die Gestaltung von Merkmalen des Arbeitssystems und ihre Auswirkungen auf Beanspruchungsprozesse. Schriften zur Arbeits-, Betriebs- und Organisationspsychologie. Band 16. Verlag Dr. Kovac, Hamburg

DIN EN 1005-5 (2007) Sicherheit von Maschinen – Menschliche körperliche Leistung – Teil 5: Risikobeurteilung für kurzzyklische Tätigkeiten bei hohen Handhabungsfrequenzen; Deutsche Fassung EN 1005-5:2007

Garcia M-G, Läubli T, Martin BJ (2015) Long-term muscle fatigue after standing work. Hum Factors 57:1162–1173

Hartmann B, Spallek M, Ellegast R (2013) Arbeitsbezogene Muskel-Skelett-Erkrankungen: Ursachen, Prävention, Ergonomie, Rehabilitation. ecomed MEDIZIN, Verlagsgruppe Hüthig Jehle Rehm GmbH, Heidelberg

Heidl MJ (2013) Untersuchung der Anwendbarkeit des Verfahrens OCRA zur Bewertung repetitiver Tätigkeiten. Masterthesis, TU Darmstadt, 26.11.2013

Heidl MJ, Boespflug A (2016) Untersuchung der Belastung und Beanspruchung repetitiver Montagetätigkeiten. In: Gesellschaft für Arbeitswissenschaft e. V. (Hrsg) Arbeit in komplexen

Systemen - Digital, vernetzt, human?! Dokumentation des 62. Arbeitswissenschaftlichen Kongresses Aachen 02.03.-04.03.2016. GfA Press, Dortmund

Joiko K, Schmauder M, Wolff G (2010) Psychische Belastung und Beanspruchung im Berufsleben. Erkennen—gestalten (5. Aufl.). Dortmund-Dorstfeld: Bundesanst. für Arbeitsschutz u. Arbeitsmedizin. Zugriff am 12.05.16. Verfügbar unter http://www.baua.de/cae/servlet/contentblob/673898/publicationFile/

Kotzab D (2014) Vergleich ergonomischer Bewertungsmethoden mit dem subjektiven Befinden der Mitarbeiter. Kassel university press GmbH, Band 12, Kassel 2015

Kuorinka I, Jonsson B, Kilbom A, Vinterberg H, Biering-Sørensen F, Andersson G, Jørgensen K (1987) Standardised Nordic questionnaires for the analysis of musculoskeletal symptoms. Appl Ergon 18(3):233–237

Silverstein BA, Fine LJ, Armstrong TJ (1986) Hand wrist cumulative trauma disorders in industry. Br J Ind Med 43:779–784

Slesina W (1987) Arbeitsbedingte Erkrankungen und Arbeitsanalyse – Arbeitsanalyse unter dem Gesichtspunkt der Gesundheitsvorsorge. Enke, Stuttgart

Slesina W (2009) FEBA: Fragebogen zur subjektiven Einschätzung der Belastungen am Arbeitsplatz. ASER-Institut. Zugriff am 13.05.2016. Verfügbar unter http://www.rueckenkompass.de/out.php?idart=18

Steinberg U, Caffier G, Mohr D, Liebers F, Behrendt S (1998) Modelhafte Erprobung des Leitfadens Sicherheit und Gesundheitsschutz bei der manuellen Handhabung von Lasten. 1. Auflage. Wirtschaftsverlag NW Verlag für neue Wissenschaft GmbH, Bremerhaven

Steinberg U, Liebers F, Klußmann A, Gebhardt HJ, Rieger MA, Behrendt S, Latza U (2011) Leitmerkmalmethode Manuelle Arbeitsprozesse 2011. Bericht über die Erprobung, Validierung und Revision, Forschung Projekt F 2195. Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund/Berlin/Dresden 2012

Takala EP, Pehkonen I, Forsman M, Hansson GA, Mathiassen SE, Neumann WP, Sjøgaard G, Veiersted KB, Westgaard RH, Winkel J (2010) Systematic evaluation of observational methods assessing biomechanical exposures at work. Scand J Work Environ Health 36 (1):3–24. PMID: 1995321

Wittig P, Nöllenheidt C, Brenscheidt S, Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (Hrsg) (2013) Grundauswertung der BIBB/BAuA-Erwerbstätigenbefragung 2012 – Männer/Frauen in Vollzeit. Dortmund, Berlin, Dresden

Analysis and Evaluation of Physical Workload During Long-Cyclic Tasks as a Prerequisite for Ergonomic Work Design

Dorothee Müglich, Karlheinz Schaub, Bastian Kaiser, Steffen Rast, Lukas Bier, Katharina Rönick, Andrea Sinn-Behrendt, Peter Kuhlang, and Ralph Bruder

Abstract

Evaluation of physical workload such as the screening tool Ergonomic Assessment Worksheet (EAWS) were developed and validated for use in industrial manufacturing with short cycles (up to 3 min). If this tool is used for work stations that have significantly longer cycles or no cycle at all, its use is limited. On the one hand, measuring the workload intensity becomes increasingly expensive. On the other hand, the workload intensity for these scenarios has to be examined as well. The following contribution presents an instance of how the EAWS can be used for long-cycle tasks on the example of a maintenance workplace at Deutsche Bahn. This work is focused on methods for adapted data collection from which indicators for possible adjustments to the evaluation procedure can be concluded.

Keywords

Physical workload • Long-cyclic tasks • Ergonomic Assessment Worksheet (EAWS)

D. Müglich (✉) • K. Schaub • L. Bier • K. Rönick • A. Sinn-Behrendt • R. Bruder
Institute of Ergonomics and Human Factors (IAD), Technische Universität Darmstadt, Otto-Berndt Str. 2, 64287 Darmstadt, Germany
e-mail: d.mueglich@iad.tu-darmstadt.de

B. Kaiser
Kaiser Production Consulting, Remchinger Straße 16a, 75236 Kämpfelbach, Germany

S. Rast • P. Kuhlang
German MTM Association, Elbchausse 352, 22609 Hamburg, Germany

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_16

1 Introduction

The German regulations for health and safety at work (89/391/EEC) (Arbeitschutzgesetz) as well as the regulation of safety on machine and products (2006/42/EC) (Geräte- und Produktsicherheitsgesetz), which are the basis of safety and risk evaluation, demand profound ergonomic knowledge for their implementation as well as for work place design. Consequently ergonomic assessment methods and tools have been developed to meet the demands of legislation (e.g. Ergonomic Assessment Worksheet EAWS; Key Indicator Method KIM). These methods became widespread in the risk evaluation of predominately assembly tasks.

In order to evaluate the risks that result from the physical workload of a task, the EAWS screening tool has been utilised and proven worldwide. EAWS was mainly developed for the evaluation of short-cyclic tasks (duration of each cycle less than 10 min) and requires an even and recurring distribution of workload over the course of one shift, respectively a full work day. Long-cyclic tasks (cycle > 10 min, Fig. 1) are predominantly found in manufacturing companies such as aircraft manufacturing and rail-vehicle manufacturing. Lately there has been an increased demand to evaluate those long-cyclic tasks in the maintenance and repair sector. Due to health and safety regulations at work and demographic constraints the ergonomic risks for these tasks have to be determined and minimized as well.

In addition to these conditions and because of the length of the cycles, the diversity of the tasks and the divergence in influencing factors, data collection and processing for long-cyclic tasks is currently only possible with time consuming effort.

Adapting the EAWS process to evaluate long-cyclic tasks is no trivial matter, as there are significant differences in data collection and evaluation compared to short-cyclic tasks. Furthermore, the boundaries between short- and long-cyclic tasks have not been sufficiently defined.

One of the research goals of the co-operation between the Institute of Ergonomics and Human Factors at TU Darmstadt (IAD) and the German MTM Association (MTM), is to develop a method based on EAWS in order to adequately evaluate long-cyclic as well as non-cyclic tasks. The first step is to review the approach of the workload analysis with regard to its practicality. The second step is to present an exemplary workload evaluation for long-cyclic workplaces using EAWS.

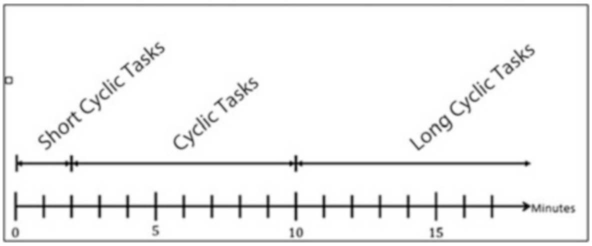


Fig. 1 Definition of the duration of each cycle (Kuhlang et al. 2015)

2 Ergonomic Assessment Worksheet (EAWS)

The “Ergonomic Assessment Worksheet” (EAWS) is a screening tool to evaluate the physical workload on the human body in different workplaces. It was developed for production processes and for production planning in the automotive industry and similar industries. By using EAWS physical stress can be evaluated with a high level of detail. Aspects of successive stress superposition can be greatly simplified for short-cycle tasks. The results of the evaluation are the basis for the communication between production and the development department (Schaub et al. 2012). Due to the detailed results of the evaluations, EAWS has significant advantages compared to other established evaluation procedures (Savino et al. 2016).

Official regulations determine the application and continuous development of evaluation procedures to a great extent. With the introduction of the EU-Framework Directive 89/391/EEC of the council on the 12th June 1989 and the EU-machinery directive 98/37/EC, a legal framework for occupational health and safety was determined (Schaub and Landau 2004). As new ergonomic knowledge influences the norms, the further development of the EAWS has to be dynamic in order to meet the demands of occupational safety.

The EAWS scheme for the evaluation of physical risk factors is in accordance with the machinery directive EN 614-1:2006/prA1:2008, consisting of a “three zone” evaluation systematic. The zones are divided into a traffic light pattern (Fig. 2). Occupational conditions that fall into the “green zone”, are characterised by a safe construction and a safe operation and tasks that require longer and continuous use fulfil ergonomic principles. Task evaluations that fall into the “yellow zone” only fulfil ergonomic principles under certain conditions (limited use, short duration). The “red zone” includes tasks that do not meet the requirements of ergonomic principles and are conditions that lead to an unsafe operation prevail (EN 614-1:2006/prA1:2008).

This evaluation scheme is used throughout EAWS, whereby the physical risk factors are separated and assessed into four sections and then, in accordance with the superposition of the sectional stress values, an overall stress value is determined. The sections are structured as follows:

0-25 Points	Green	Low risk: recommended; no action is needed
>25-50 Points	Yellow	Possible risk: not recommended; redesign if possible, otherwise take other measures to control the risk
>50 Points	Red	High risk: to be avoided; action to lower the risk is necessary

Fig. 2 Traffic light scheme of the EAWS-worksheet

- Section 0—Extra Points:
  In this area general evaluations of the task are determined, that cannot be included in the following areas. These include rebound forces, impulses and vibrations.
- Section 1—Body Posture (torso and arm position):
  This area includes the evaluation of static and/or high frequency movements of the torso and the arms. Only static postures of at least 4 s are taken into account. Two forward bends or lifting the arms higher than 60° ten times per minute are counted as high frequency movements. Weights under 3 kg and action forces in the range of 30–40 N are included.
- Section 2—Action Forces:
  This area considers action forces of more than 40 N. For this the maximum force for the combination of body position and the direction of the force is selected from the table. Also the maximum forces acting on the fingers that correspond with hand movements are determined. In conclusion the corresponding point values are selected using the level of the force and duration or frequency of the action.
- Section 3—Manual Material Handling:
  This area is focused on evaluating the manual handling of loads. First of all the load points for shifting (lifting/setting down), carrying and holding as well as pushing and pulling are determined, which can be taken from the table using the weight of the load and the type of execution. After that the contact points, execution points as well as frequency, duration and distance points are necessary. These points are then added according to the given formula which results in the overall score for this area.
- Section 4—Upper Extremities (during repetitive tasks):
  In that area the stress of upper extremities during highly repetitive tasks is evaluated. First of all the force, frequency and condition, under which an object is grabbed, have to be determined. Following that the hand/forearm/shoulder joint position is identified using the highest pressure on the hand, forearm or shoulder. After that, additional functions that impede the execution of a task are checked. In conclusion the duration of the repetitive movement has to be recorded and the overall rating for the stress on the upper extremities during repetitive tasks has to be calculated using the given formula.

Sections 1–3 refer to workload on the whole body and follow a "traditional ergonomic evaluation approach". Findings on the general energy consumption, load distribution along the spine and local muscle fatigue are the basis for the evaluation criteria. A valuation from more than one section is not permitted which further increases the validity of the procedure. According to the required force or the load weight during a task, the workload is rated as body posture (less than 30–40 N or 3–4 kg), as action force or as manual material handling (Schaub et al. 2012). Section 4 refers to the upper extremities on the basis of medical and epidemiological data. It is referenced by the OCRA-method (Schaub et al. 2012). The overall result is calculated by adding the points from posture, force, load and extra points.

Aside from this are the results from the upper extremities that are a direct result from the formula to determine the overall result of the stress on the upper extremities. These two results are compared to the EAWS-valuation table and the risk of the workplace is derived.

3 Initial Situation

There are a few different methods to analyse and evaluate the physical workload during short-cyclic tasks with different focuses, that all simplify the evaluation of the workplace considerably. Aside from the EAWS, the Key Indicator Method (KIM), RULA- (Rapid Upper Limb Assessment), the OCRA- (Occupational Risk Assessment) or NIOSH-method are common tools.

There have been only few studies conducted to evaluate the workload of long-cyclic or non-cyclic tasks. In 2011, working together with a manufacturer of utility vehicles, the IAD examined approx. 100 long-cyclic assembly work stations in bus manufacturing. During this examination 12 model workstations were recorded on video and analysed in great detail to evaluate the physical workload situation of the workers using EAWS.

A former study carried out by MTM with the Deutsche Bahn (DB) presents a promising new possibility to minimize the analytical effort for long-cyclic tasks. It combines the advantages of a work sampling method based on the PATH approach (Buchholz et al. 1996) with an EAWS analysis. This procedure uses descriptions and a record of the different processes to generate the ergonomically influencing variables. During this study 8509 observations of approx. 20 workstations in set intervals were carried out. Furthermore, long-cyclic workplaces in the food industry and the metal working sector were observed (Kuhlang et al. 2015).

The PATH method (Posture, Activity, Tools and Handling) is based on the work sampling method and it was developed by Buchholz et al. (1996) to characterize the ergonomic hazards of non-repetitive work especially of heavy highway construction work. The posture evaluation in the PATH method is based on the Ovako Work Posture Analysing System (OWAS), with other codes including worker activity, work equipment use, manual material handling and grasp type.

4 Methodology

For the ergonomic evaluation of long-cyclic tasks detailed information about the physical workload are needed along the whole duration of the task. This information can only be collected during the observation of the task, as most tasks are not repeated over the course of the observation. Different documentation aids were used to record the observations.

The IAD used a "physical workload matrix" to record the data. MTM used a "data record sheet" to record the data.

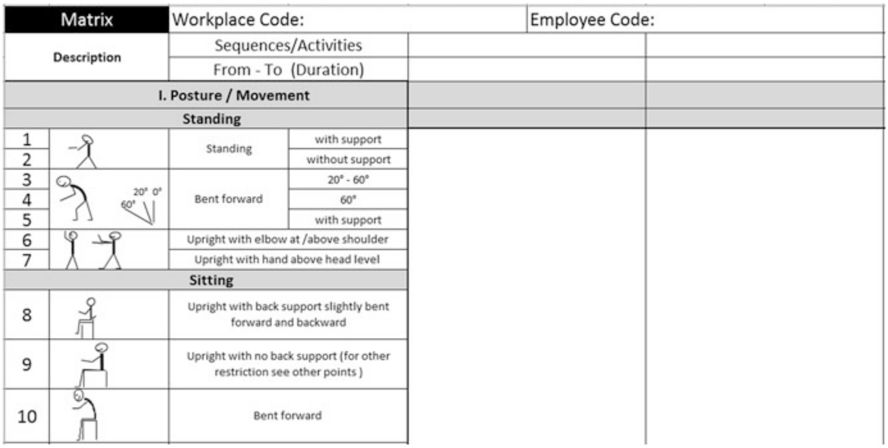
4.1 Physical Workload Matrix

With the support of the physical workload matrix the data needed for the evaluation of long-cyclic tasks can be recorded while the task is being carried out. The EAWS sheet is the basis of the new developed physical workload matrix. Using this process the workload variations body posture, action forces and the manual material handling can be evaluated in different sectors.

The physical workload matrix consists of two data entry forms. The first form is used to record the body posture and movement, the asymmetric incline and rotation of the torso as well as the arms being held far from the body. Figure 3 is a detail of this form. The second form is used to record action forces, the manual material handling as well as special characteristics of the activity such as particular joint positions or surrounding conditions (e.g. noise). Each partial activity as well as its duration is noted in the header of each form. With this, each partial activity has its own column into which the observed stress situations are entered. In order to simplify and quicken the use of the form, each workload situation is allocated a number. This number is entered into the column and the duration and frequency are recorded. When recording the handling of loads and action forces, the height of the load, respectively the force is noted.

A pre-structuring of the matrix into sequences (Wells et al. 1997) is also provided. Later, the overall workload can be determined through a time weighting of each sequence.

The physical workload matrix is used to record an observation chronologically over a longer period of time where there is no possibility of using video cameras or other equipment to record the activity. Measurements of the workplace have to be done separately. It is advisable to hold a briefing with the worker who is responsible for the workplace to obtain a work flow, so that parts of the matrix can be filled out beforehand.



Matrix		Workplace Code:		Employee Code:	
Description		Sequences/Activities			
		From - To (Duration)			
I. Posture / Movement					
Standing					
1		Standing	with support		
2			without support		
3	20° 0° 60°	Bent forward	20° - 60°		
4			60°		
5			with support		
6		Upright with elbow at /above shoulder			
7		Upright with hand above head level			
Sitting					
8		Upright with back support slightly bent forward and backward			
9		Upright with no back support (for other restriction see other points)			
10		Bent forward			

Fig. 3 Detail of the physical workload matrix (1. from)

Data Record Sheet **Body postures and arm movements**		Sequence No. Start: End:	Sequence No. Start: End:
2D favorable	% estimated		
2D unfavorable			
2D / 3D			
Special Force (frequency)			
Finger-/Hand	normal		
	difficult		
Arm-/Shoulder/Whole Body Focus	normal		
	high		
	extreme		
Remarks			
Load handling (frequencly)			
Shifting	normal		
	difficult		

Fig. 4 Detail of the data record sheet

4.2 Data Record Sheet

With the data record sheet (Fig. 4) all timely and ergonomically influencing factors that are necessary for a workload evaluation of long-cyclic tasks can be recorded efficiently.

The partial activity that has to be evaluated as well as its duration is recorded in the header. Each sequence has its own column into which the observed workload situations are entered separately. Each workload situation is separated into "normal", "high" and "extreme". The frequency or percentage of each workload is entered into the appropriate level.

Using the weight, the number and the frequency the work items are assigned to different ways of handling (i.e. shifting of a load). In order to record the time of an action the main body posture (symmetrical or asymmetrical) is noted in words. By recording the number and frequency the influential factors of each sequence are determined for further evaluation. The work sequences are recorded in chronological order. This methodology facilitates the systematic transfer of data into the MTM software EAWS*digital*.

4.3 Comparison of Data Collection Methods

The different data collection instruments have proven to be valuable in everyday practice. There are advantages of both methods. The main advantage of the physical workload matrix is the detailed and chronologic recording of ergonomically relevant data while the advantage of the data record sheet is the combination of frequency and ergonomic aspects, which has proven to be very useful. Both methods are compared in the Table 1.

Table 1 Comparison of data collection methods

Physical workload matrix	Data record sheet
Process-oriented data collection	Object-oriented data collection
Detailed recording of ergonomically relevant data	Combination of frequency and ergonomic aspects

Table 2 Time calculation for the disassembly/assembly of annular spring segments

Pre calculation following data card		Actual value (min)
10	Feed work status (segments)	2.96
20	Disassemble annular spring segment	30.98
30	Refurbish annular spring segment (used part)	21.68
40	Check annular spring segment	2.94
50	Weigh annular spring segment	3.86
60	Annular spring segment cleaning segment seat	7.03
70	Annular spring segment cutting of thread	30.03
80	Mounting of annular spring segment	60.00
	... for 6 segments	159.48
	... for 12 segments = 1 wheel set	319

To summarize, the comparison of both methods showed that a skilful combination of both data collection methods would be an optimal basis for the collection of data in long-cyclic tasks.

5 Exemplary Application of the Methodology

For this pre-study we selected a long-cyclic workplace in the DB maintenance facility in Dessau (Germany), where segments of annular springs used in rail vehicles are refurbished. The segments are taken apart, cleaned and refurbished before being put back together.

This workplace can be characterised by its high physical workload that is a result of handling heavy loads (for each segment of an annular spring 13 kg), a multitude of constrained postures and critical working heights as well as having to use high action forces (i.e. bending of locking plates and using a clamping device).

The duration of a cycle for this activity is approx. 5 h (319 min) per axis, which MTM has already examined in 2011 in terms of time aspects (Table 2). Thus the activity in disassembly/assembly is representative for a long-cyclic task.

Four people observed the activities at the selected work station (Fig. 5) over the course of a complete cycle from different perspectives. Aside from the observation protocols, they wrote a procedure protocol and took photos of individual work situations. In addition, ergonomics experts, health managers and workers carried out a subjective stress evaluation of the workplace.

Fig. 5 Example of the evaluated workplace

6 Results

As anticipated, this study proved that the data collection and evaluation for long-cyclic activities is not only time consuming but also significantly more difficult than data collection for short-cyclic activities.

An observation of the execution of each activity was proven to be an important basis for the evaluation of long-cyclic activities from an ergonomic point-of-view. It is important that all important influential factors that are relevant for an ergonomic evaluation are considered in their time frame. Due to the long duration of each movement (duration of each cycle) it is necessary to note each observation in a detailed and reliable manner. The methods (physical workload matrix and data collection sheet) that were used proved to be practical and valuable.

Furthermore, it is has become clear, that it is sensible to carry out the data collection with at least two people. This ensures that all information can be recorded completely and accurately. Before the data can be used for an evaluation all observations have to be compared and checked for their plausibility.

In the following we present for the three phases—before, during and after the on-site data collection—practical hints and valuable knowledge.

6.1 Lessons from the Preparation for Data Collection

To ensure an effective and efficient data collection it is important that the observers know the task they are going to evaluate. This includes the order and extent of the partial tasks as well as the working materials and component parts that are used. All the following (available) documents should be viewed, checked for timeliness and evaluated as a preparation for data collection if available:

– A description of the task and instructions
– An analysis of the work sequence (i.e. MTM-UAS or MTM-MEK)

– Movement graphs
– Part lists and equipment descriptions

Another aspect that should be considered is that the planned work method often diverges from the actual working method (Bokranz and Landau 2012). This can be due to less strict definitions of the work method or due to the complexity of long-cyclic tasks.

In order to record these differences and be able to take them into account, the following points have to be considered

– After viewing the work methods in the description of the task they have to be reviewed with experienced workers or foremen and checked for differences
– Formulating a description of the work system on the basis of statements by experienced workers in order to recognize influences that disrupt the work method

Long-cyclic tasks often include repetitive partial tasks. It is especially important to identify and describe these partial tasks in preparation for the data collection as well as to pay attention to any divergence during the repetitive partial tasks.

It is common that tasks with complex objects require different variations of specific parts that can lead to diverging ergonomic influences and different stress factors. In order to portray stress situation realistically, the main product variations (top sellers) have to be discussed beforehand. Less frequent product variations (exotics) should not be included in the data collection.

6.2 Lessons for the Execution of the Data Collection

The data collection methods (physical workload matrix and data collection sheet) are the most important instruments for data collection.

In addition, the following should be seen as valuable advice for the observation of long-cyclic tasks:

– Due to larger work areas and changing work places of long-cyclic tasks, the observers have to be increasingly mobile.
– It is advisable to collect data from different perspectives, which involves at least two people taking notes at the same time. To be able to conflate and match the data the same collection method should be used and the observers should have similar experience.
– Video recordings are a useful tool to capture a complex work situation as they allow for a review of unclear situations after the actual observation has been completed. All legal requirements needed for video recordings should be clarified beforehand.

6.3 Lessons for the Post Processing of the Data Collection

Even with extensive preparation and consideration of all the advice concerning the data collection, the detailed recording of a complex long-cyclic task is a difficult challenge, often made harder by mistakes and variations in the collection. To avoid severe discrepancies, a plausibility check is advisable. The following methods can be used:

– Aligning of how often an activity was carried out (i.e. moving of objects) with part lists and task descriptions
– Checking the work chronology or the partial activities with experienced workers or foremen

With this study we have shown that it is possible to collect solid data for the evaluation of physical workload in long-cyclic tasks using EAWS. An EAWS evaluation was carried out on the basis of the data collected in this pre-study of the work place "disassembly/assembly of annular spring segment".

This calculation was done on the basis of the existing EAWS-catalogue and calculation algorithm. Any occurring spikes in workload were not taken into account, as the basis of the evaluation is an equal distribution of workload.

In an additional questionnaire the determined workload perception of the workers and the assessment of the experts was confirmed with the result of the evaluation. This is relevant not only for the overall result but for the duration of the task as well as the evaluation of individual partial tasks.

7 Discussion

The data collection methods and the subsequent EAWS evaluation are a valuable option to record all workload variations that occur during long-cyclic tasks in their intensity, duration and chronology. Nonetheless, it would be helpful to examine the advantages of both collection methods in order to find out where one is more suitable than the other or where they can complement each other. It is evident that the physical workload matrix is taking the process into consideration and mainly records body positions and the duration of the stress factor, while the data collection sheet is object orientated and more suitable to record the frequency of body forces as well as the handling of loads. In comparison with the evaluation of short-cyclic tasks it must be noted that the data collection and analysis is significantly more complex.

Long-cycle activities might have less time pressure than short-cycle activities and it might be possible to avoid lengthy, static forced postures. On the other hand, lengthy static workload can occur in long-cyclic activities that do not happen in short-cycle tasks. Furthermore, the stress factors are often not equally distributed throughout the working day but occur as stress spikes during peak times (Finke

2011). Assumptions made for the development of stress evaluation methods for rest and fatigue cannot be applied (Schulte 1987).

The following criteria should be met to be able to apply an EAWS evaluation correctly (Kuhlang et al. 2015):

- No occurring spikes in workload
- No concentrated or cumulated stress effects
- No long periods of exertion that require rest periods
- No surrounding influences that impede physical performance

For the application to long-cyclic tasks the following adjustments to the EAWS procedure should be made (Finke 2011):

Peak Loads

For short-cycle activities the stress loads occur in regular intervals. The stress is distributed evenly over the course of a shift. For long-cyclic activities the stress is typically not distributed evenly and peak loads might occur during specific phases. For example, the rating for the body posture is distributed following a linear principle. If longer static body postures occur, then this assumption is not correct. For this case a progressive course of rating is needed.

Similarly, while handling loads in long-cyclic tasks different variations of the same load handling type can occur during one cycle. In this case a weighted mean value has to be calculated as the EAWS method only allows one rating per "moving", "carrying", "holding" and "pulling & pushing". This means that dangerous peak loads during the handling of loads are not indistinguishable or are not included in the evaluation.

Effects of Fatigue and Rest

It can happen that the process of long-cyclic tasks can lead to longer waiting times or longer postures in less strenuous body positions. Long-cyclic tasks are sometimes characterized by less time pressure. Recuperation effects are possible, that are not taken into account in short-cyclic tasks. On the one hand, long static forced postures can be avoided. On the other hand, long static forced postures can occur, which do not happen in short-cyclic tasks due to time limits. Assumptions, formulated for evaluation methods for rest and fatigue (Schulte 1987) are not applicable.

If, in any case, it becomes apparent that the tiring static body positions and the ensuing difficulties in individual body parts are not compensated in the following sections, the scaling should be altered. One solution would be to determine certain body segments that experience difficulties, which cannot be compensated in the following activity. The point values of the body positions that are stressful could be increased.

Minimum Evaluation

A further point of criticism of the evaluation with the EAWS procedure is that the evaluation of the body posture begins at 5 %. For long-cyclic durations these small percent values can mean a significant duration of a static posture. This makes it necessary that the minimum evaluation has to be lowered.

Repetition

There is no repetitive character identified in long-cyclic tasks, as many different activities are carried out over a long period of time. That is why the stress on the upper extremities has not been recorded in this pre-study.

In summary, the focus of the adjustments to the EAWS for long-cyclic tasks is concerned with peak loads and potentially with the ensuing bottlenecks. One helpful approach to improve the analysis and the evaluation of long-cyclic tasks would be to divide the cycle into smaller sections with partial tasks (Wells et al. 1997). Hence, the overall workload can then be determined by the time weighting of each sequence.

8 Outlook

The physical workload matrix is not only used for the DB maintenance in the facility in Dessau but is also used to record physical workload data in non-cyclic tasks in construction. After the practical tests of the entry sheets, the next step should be a complete validation of the process. For this, long-cyclic respectively non-cyclic tasks should be recorded with entry sheets as well as video recordings and the necessary additional information. Subsequently an evaluation on the basis of the video and the accompanying data as well as with the help of the new approaches should be carried out. Based on these results the entry sheets should be further secured and refined. Simultaneously they should be checked for their practicability.

A further adaption of the EAWS to evaluate long-cyclic or non-cyclic tasks would mean more extensive research, possibly even the development of a completely new method would be necessary. The work described here can serve as groundwork for such a development.

References

Bokranz R, Landau K (2012) Handbook industrial engineering. Schäffer-Poeschel, Stuttgart

Buchholz B, Paquet V, Punnett L, Lee D, Moir S (1996) PATH: a work sampling-based approach to ergonomic job analysis for construction and other non-repetitive work. Appl Ergon 27(3): 177–187

EN 614-1:2006/prA1:2008. Safety of machinery – Ergonomic design principles – Part 1: Terminology and general principles

Finke K (2011) Studie zur Analyse und Bewertung von langzyklischer Montagetätigkeit. Diplomarbeit, IAD, TU Darmstadt

Kuhlang P, Rast S, Liebig S, Lüben A, Finsterbusch T, Mühlbradt T (2015) Methoden und Perspektiven zur ergonomischen Bewertung und Gestaltung langzyklischer Tätigkeiten in der Fahrzeuginstandhaltung. In: Biedermann H (ed) Praxiswissen Instandhaltung, Intelligente, lernorientierte Instandhaltung – Smart Maintenance. TÜV Media GmbH, Köln, pp 181–193

Savino M, Mazza A, Battini D (2016) New easy to use postural assessment method through visual management. Int J Indus Ergon 53:48–58

Schaub K, Caragnano G, Britzke B, Bruder R (2012) The European assembly worksheet. Theor Issues Ergon Sci 14:616–639

Schaub K, Landau K (2004) Ergonomie und Prävention in der betrieblichen Praxis. Angewandte Arbeitswissenschaften 180:52–70

Schulte B (1987) Praxisorientiertes arbeitswissenschaftliches Instrumentarium zur analytischen Ermittlung von Erholungszeiten. RKW-Verlag, Eschborn

Wells R, Norman R, Neumann P, Andrews D, Frank J, Shannon H, Kerr M (1997) Assessment of physical work load in epidemiologic studies: common measurement metrics for exposure assessment. Ergonomic 40:51–61

Walking “Normally” vs. “Sideways” in Simulated, Simple Assembly Operations: Analysis of Muscular Strain in the Legs

Jurij Wakula, Dorothee Müglich, and Ralph Bruder

Abstract
The muscular strain of the lower extremities when walking “normally” and “sideways” was analysed using the simple, simulated U-assembly line in the Laboratory of the Institute for Ergonomics and Human Factors in Darmstadt (IAD). Test subjects executed their assembly operations in different scenarios in two studies. The U-line in the first study consisted of three work stations and five work stations in the second one. Electrical activities (EA) in six leg muscles on each leg (left and right) were measured and analysed by using EMG method. Ten test subjects without experience in assembly work took part in both studies. The results in the first study show that walking “sideways” puts lower extremities under more stress than walking “normally” does. We were able to record higher electrical activities values (especially dynamic EA-shares) in four out of six analysed leg muscles. EMG-results in the second study show that when “walking sideways counter-clockwise”, three muscles on the right leg are under greater stress than the muscles on the left leg.

Keywords
Simulated U-line • Walking “sideways”/“normally” • Walking stress and muscular strain on legs

J. Wakula (✉) • D. Müglich • R. Bruder
Institut für Arbeitswissenschaft (IAD), Technische Universität Darmstadt, Otto-Berndt-Str. 2, 64287 Darmstadt, Germany
e-mail: wakula@iad.tu-darmstadt.de; bruder@iad.tu-darmstadt.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_17

1 Introduction

Nowadays the workflow principle of "One-Piece-Flow" is used in assembly, to ensure a high degree of flexibility in production and at the same time the highest possible quality of the product at lowest cost in a very short time. This principle originated from the Toyota production system and is based on the Just-In-Time principle of the "Lean Production"-approach (Kummer 2009). Using this method it is possible to process a product on multilevel workstations without having to place the product into temporary storage. Thus a flexible output quantity can be maintained, despite having a high variance. To avoid long distances between the first and the last station, the workstations are typically laid out in a U-shape. The whole workplace is referred to as an U-assembly line.

When working in an U-assembly line a great amount of stress is put on the upper extremities due to short movement cycles that include highly dynamic hand-arm movements, as well as stress on the lower extremities that results from frequently standing and walking between different work stations. In addition, movements that require the worker to walk "sideways" are frequent. In order to make a statement concerning the strain that selected leg muscles are exposed to when walking counter-clockwise "normally" and "sideways" in an U-line assembly, two laboratory studies were conducted at the IAD (Institute of Ergonomics and Human Factors at TU Darmstadt) (Wakula et al. 2016a). This contribution presents the results of these studies.

2 Methodology and Experimental Setup in Both Studies

2.1 Surface Electromyography (EMG)

The objective strain data was measured on six selected leg muscles (left and right half of the body), using the method of surface electromyography (EMG, Strasser 1996).

For the EMG recording a portable TeleMyo 2400 G2 device by Noraxon (Noraxon, USA, Fig. 1) and a video recorder by Panasonic NV-DS28EG was

Fig. 1 TeleMyo 2400 G2 system from Noraxon

used. The EMG was recorded using a continuous trace, a frequency of 1500 Hz, a time interval of 20.0 s, and disposable surface electrodes (Noraxon Dual Electrodes, USA and Ambu Blue Sensor P, Denmark). Since the Noraxon EMG-device can only record a limited number of muscular activities, the number of muscles that were to be recorded in the experiment had to been reduced. In an experiment with one subject the 15 surface muscles on each leg (left and right) that are active during work on a U-line, were recorded and compared. The following muscles proved to be most relevant for the studies on U-lines: *Musculus (m.) tibialis anterior*, *m. gluteus medius*, *m. peroneus brevis*, *m. biceps femoris* as well as *m. gastrocnemius (medialis)* (Fig. 2). As the *m. peroneus brevis* is very small and the electrodes are difficult to place, this muscle was excluded from further studies and analysis. In addition to the five muscles, we chose to include the *m. quadriceps femoris (vastus lateralis)* for some analysis, as it is essential for the extension of the knee joint and is therefore activated when walking. Functions of the analyzed muscles are presented in Table 1.

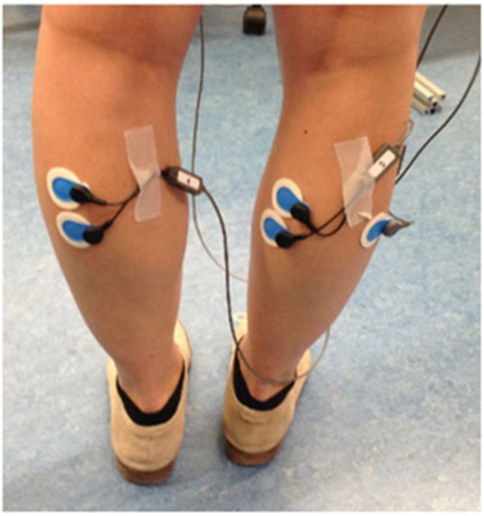


Fig. 2 EMG-electrodes placed on the legs

Table 1 Function of the analyzed muscles

Muscles	Function
m. gluteus medius	Stabilization of the pelvis as well as abduction and inward rotation of the hip joint. A contraction of the ventral fibers results in a flexion and inward rotation. The dorsal fibers perform an extension and outward rotation.
m. quadriceps femoris (vastus lateralis)	Extension of the knee joint. It keeps the knee from buckling when standing.
m. gastrocnemius (medialis)	A plantar flexion of the foot as well as a flexion of the knee joint.
m. tibialis anterior	Dorsalflexion and supination (tilt inward) of the foot.
m. biceps femoris	It bends the knee joint and laterally rotates the knee joint (when knee is flexed). The long head of the m. biceps femoris also extends the hip joint.

For the measuring of the resting activity the subject was in a recumbent position to allow the muscle to relax. The additional testing of the maximal voluntary contraction (MVC) in the muscle is necessary for the EMG analysis to be able to compare the results of individual subjects. For the study we measured the MVC for each muscle twice at begin and at the end of measurements. The MVC-measurements were done according to the literature (AWMF online 2013).

2.2 Experiment Layout and Scenarios in the Study 1

The work on a simulated U-line can be divided into two situations with two variations in execution. In situation A the subject stood at an assembly table for 15 s and in situation B for 30 s while doing simple assembly tasks. The two variations when changing stations were described as "walking normally" (variation 1) and "walking sideways" (variation 2).

The U-line with three work stations was 2 m long and 1.4 m wide (Fig. 3). The starting position is identical in both situations. The feet are parallel to each other and the subject is standing in the centre in front of station 1 and moving on counter-clockwise.

"Walking normally" requires seven steps and "walking sideways" needs nine steps.

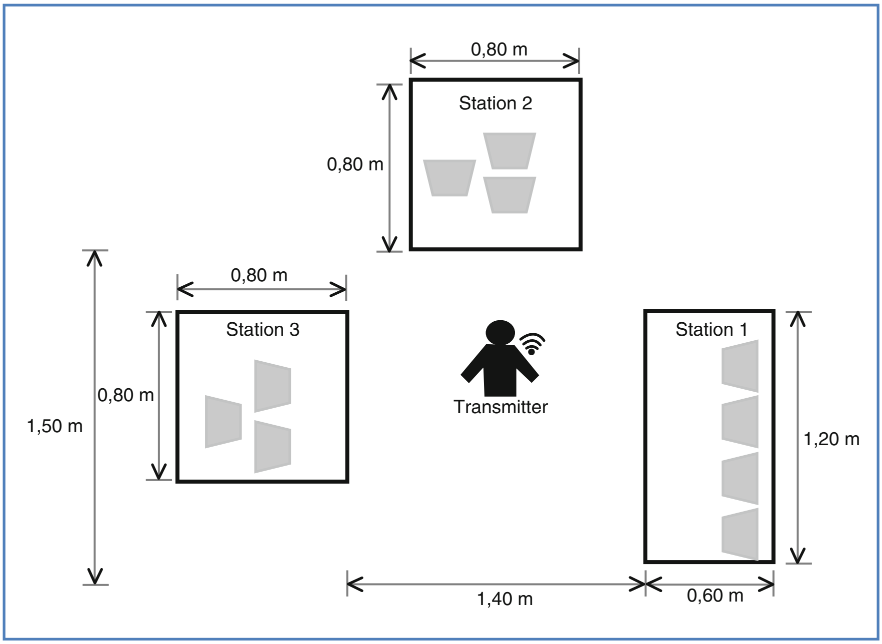


Fig. 3 Simulated U-line with three work stations in the study 1

Table 2 Analysed muscles in study 1

Left leg	Right leg
m. gluteus medius	m. gluteus medius
m. quadriceps femoris (vastus lateralis)	m. biceps femoris
m. gastrocnemius (medialis)	–
m. tibialis anterior	–

We identified and analysed the following scenarios in this study:

– Scenario A1: assembly of small pieces while standing for 15 s at each of the three stations and "walking normally" between the stations;
– Scenario A2: same as A1 but "walking sideways" (1–2 steps) between three stations;
– Scenario B1: assembly of small pieces while standing for 30 s at each station and "walking normally" between them;
– Scenario B2: same as B1 but "walking sideways" between the stations.

In between each of the four scenarios there was a break of approx. 5 min. Four young subjects (three female and one male), being 24–26 years old and 170 cm tall on average, took part in the experiment. The average body weight was 58 kg.

We measured and analysed muscular strain in the following six leg muscles (two on the right and four on the left leg) on four subjects (Table 2):

2.3 Setup of the U-Line and Subjects in the Study 2

In the second study we focused on a further analysis of the effects on the muscular strain of "walking sideways counter-clockwise" in the right and the left leg. Six test subjects were examined while working on the simulated U-line with five work stations in the IAD-laboratory with short-cycle tasks (approx. 80 s). In this study the U-line was set up out of five work stations and was 2.2 m long and 1.2 m wide.

The tables were set up close to each other (0.15–0.4 m distance between each) to allow the workers to change between the tables/stations with two steps ("*walking sideways, counter-clockwise*") (Fig. 4). There were double more "sidesteps" in this U-line compare to first study. The standing time at each station was about 15 s. During this time the subject had to execute simply assembly tasks similar to the first study with the material provided, using both hands.

The test procedure included six test series with different duration of "walking sideways, counter-clockwise". The first two test series took about 10 min, each with a break of 5 min in between. After that four test series with duration of 20 min each and a 10-min break in between were executed. Before and after the tests, resting activity and MVC measurements of the analysed muscles of each test subject were recorded.

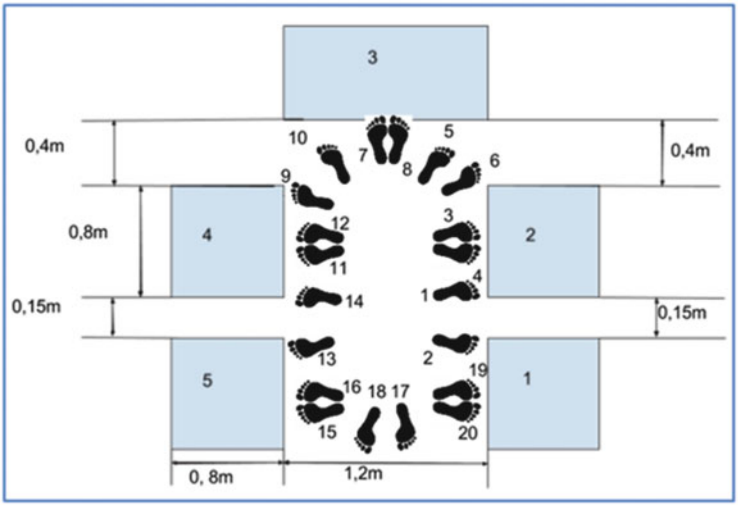


Fig. 4 Simulated U-line with five work stations in study 2

Table 3 Test subjects

Subjects	Age	Weight [kg]	Body height [cm]	BMI-index
1	29	63	1.85	18.4
2	19	60	1.80	18.5
3	19	85	1.87	24.3
4	30	87	1.85	25.4
5	29	85	1.87	24.3
6	27	83	1.87	23.7

We measured and analysed muscular strain in following six leg muscles (three in the right and three in the left leg):

– m. tibialis anterior
– m. gastrocnemius lateralis
– m. gastrocnemius medialis

Six male young subjects took part in the tests. The data concerning the test subjects can be found in the Table 3:

3 Data Processing and Results

In order to process the muscular strain data the software Myo Research XP Master Edition in the version 1.08 was used. Markers were set in the EMG raw data to define the exact start and finish of the test procedure. Afterwards, rectification and smoothing of the raw data were completed. The data of all subjects for each muscle was entered into a separate excel-spreadsheet. On this basis we created a diagram showing the static EA-parts, the dynamic EA-parts and the mean EA-values for each EMG test series. Detailed information to each parameter can be found in the EMG guideline (AWMF online 2013).

3.1 Results Study 1

For the muscle *m. quadricaps femoris (vastus lateralis)* in the left leg we found the results presented in Fig. 5a, b. When comparing the muscular activity while standing, it became clear that the mean values of the two time variables (15/30 s) were relatively constant. We recorded the high dynamic EA-parts of the EMG activity while the subject was walking "sideways" (scenario A2). Overall, the *m. quadriceps femoris (vastus lateralis)* is exposed to only a low strain level for scenarios A1, B1 and the two time variables.

The data, recorded while walking (Fig. 5b) shows, that the EA mean value and the EA dynamic value for "walking sideways" (scenarios A2 and B2) are significantly higher than they are for "walking normally" (scenarios A1, B1).

These results show, that the *m. quadriceps femoris (vastus lateralis)* is under higher stress when "walking sideways" (scenarios A2 and B2) than when "walking normally" (scenarios A1 and B1).

When comparing the activities of the muscle m. quadriceps femoris (vastus lateralis) in the left leg, it becomes clear that the EA-mean values and the dynamic EA-parts are more lower for "standing" than they are for "walking". Thus, "standing" puts less pressure on this muscle than "walking" does.

Results in the second analysed muscle *m. tibialis anterior* in the left leg show that the high dynamic value for "walking sideways" (scenarios A2 and B2) are significantly higher than they are for "walking normally" (A1 and B1) (Fig. 6b).

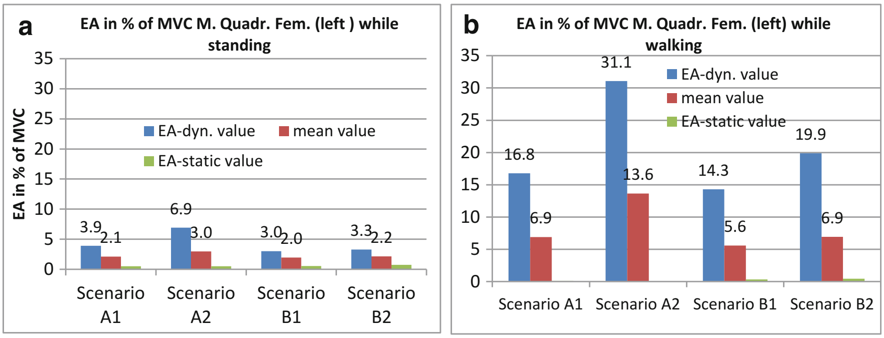


Fig. 5 (**a**, **b**) Comparison of standardised EA-parts of *m. quadriceps femoris* while standing (**a**) and "walking sideways/normally" (**b**) in the left leg over four subjects

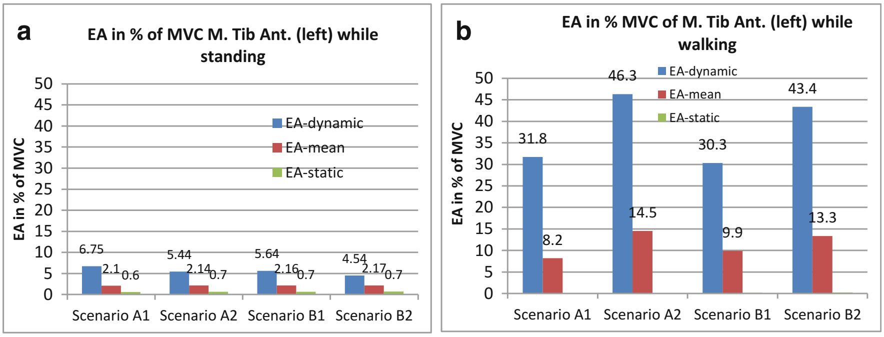


Fig. 6 (**a**, **b**) Comparison of standardised EA-parts of *m. tibialis anterior* (left) while standing (**a**) and "walking sideways/normally" (**b**) in the left leg over four subjects

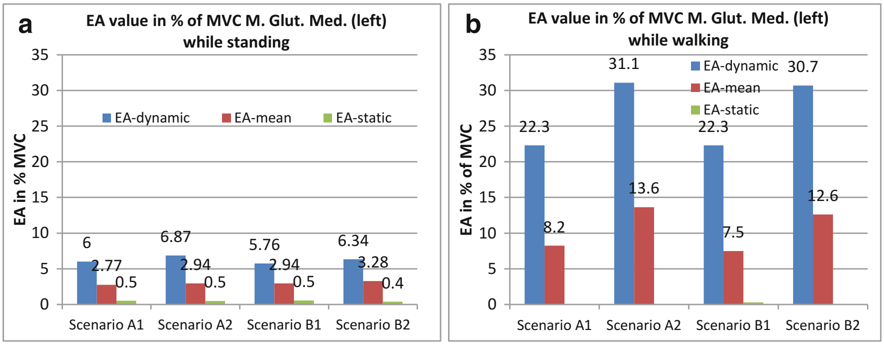


Fig. 7 (**a**, **b**) Comparison of standardised EA-parts of *m. gluteus medius* while standing (**a**) and "walking sideways/normally" (**b**) in the left leg over four subjects

There are only minor differences in electrical activities in this muscle and in *m. quadriceps femoris (vastus lateralis)*.

A comparison of the data of the *m. gluteus medius* on the left leg while "standing", showed no significant differences in the four analysed scenarios (Fig. 7a, b). The EA-mean values as well as the dynamic EA-parts stayed at a consistent, low level (<10 % of MVC). The diagrams in Fig. 7b demonstrate that when "walking normally" (scenarios A1 & B1), muscle *m. gluteus medius* has the same EA-level in the left leg. Minor EA-differences were measured when "walking sideways" (scenarios A2 & B2).

As we have measured muscle activities data of the *m. gluteus medius* in four scenarios on the right and on the left leg, we compared standardised EA-activities while "walking" on the left vs. the right leg (Fig. 8a, b).

The diagrams in Fig. 8a, b illustrate that:

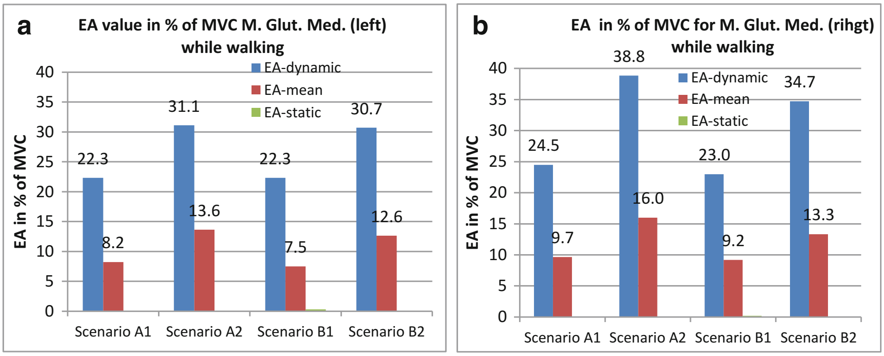


Fig. 8 (**a**, **b**) Comparison of standardised EA-parts of *m. gluteus medius* in four scenarios while "walking" in the left (**a**) vs. the right leg (**b**) over four subjects

– when "walking sideways", muscle *m. gluteus medius* of the right leg are under greater stress than the muscles in the left leg and
– when "walking normally", muscle *m. gluteus medius* has approx. the same EA-level in the left and in the right leg

3.2 Results Study 2

Initially the standardised EA-data of each muscle were analysed separately and thereafter the results for each muscle for all subjects were outlined.

Figures 9 and 10 show the EMG results of the *m. tibialis anterior* (left and right leg opposite each other) of subject 1 and 2 in the six test series.

The diagrams show clearly that the activity "walking sideways counter-clockwise" contribute to higher strain in both muscles of the right leg than in the muscles of the left leg.

These tendencies are also present in the analysed EA-data (EA-mean values) for the two muscles *m. tibialis anterior* and *m. gastrocnemius lateralis*, averaged over six test subjects (Figs. 11a, b and 12a, b).

An analysis of the mean EA-values of *m. tibialis anterior*, that are all in the range between 1.5 % and 2.7 % of the MVC, shows that these vary only slightly.

When comparing the EA-mean values of the 10- and 20-min test series, the differences are in the range of 1–2 % of MVC. The dynamic EA-parts are in the range of approx. 11–15 % of MVC and differ about 3 %. Overall, this muscle is more exposed in the right leg than it is in the left.

The muscle *m. gastrocnemius lateralis* also shows higher electrical activities (EA-mean values and EA-dynamic) in the right leg than in the left leg.

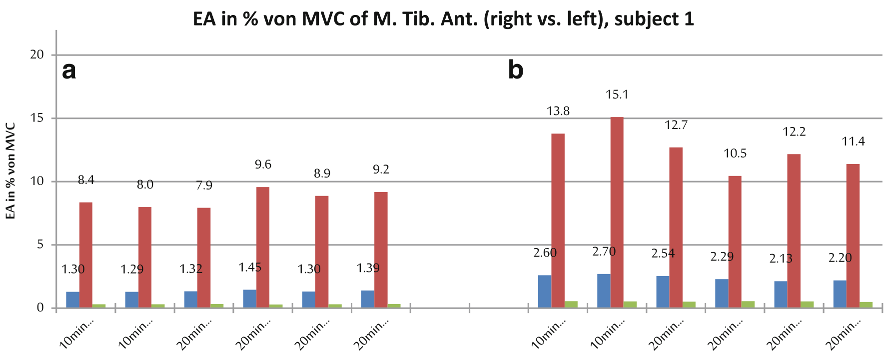


Fig. 9 Comparison of standardised EA-parts of *m. tibialis anterior* while "walking sideways" in the left (**a**) vs. the right leg (**b**) of subject 1

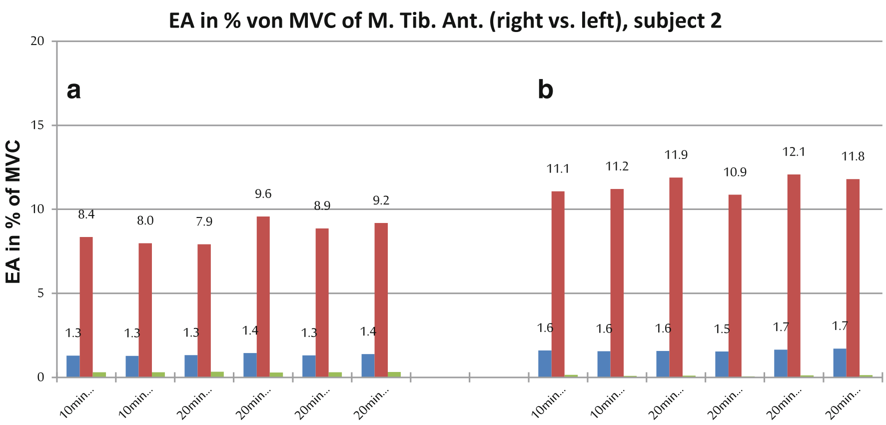


Fig. 10 Comparison of standardised EA-parts of *m. tibialis anterior* while "walking sideways" in the left (**a**) vs. the right leg (**b**) of subject 2

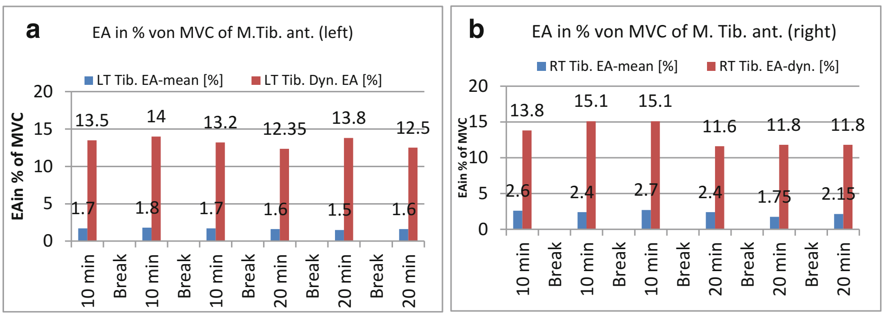


Fig. 11 (**a**, **b**) Standardised EA-parts of *m. tibialis anterior* while walking sideways in the left (**a**) vs. the right leg (**b**), averaged over six test subjects

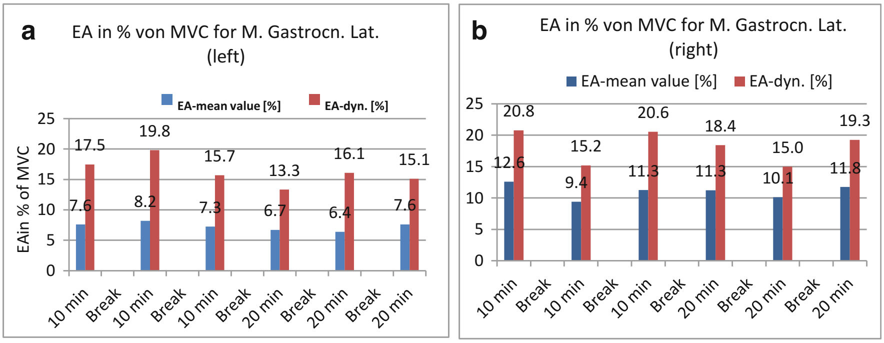


Fig. 12 (**a**, **b**) Standardised EA-parts of *m. gastrocnemius lateralis* when "walking sideways" in the left (**a**) vs. the right leg (**b**), averaged over six test subjects

4 Discussion

The muscles we analysed in the study 1 were evaluated according to different criteria, which allowed us to formulate an overall evaluation for further field studies. For this we evaluated the preparation phase in the clarity of the electrode placement, the quality of the measurements taken during resting activity as well as the performance of the MVC-testing. Additionally the muscular activity while standing and walking (mean values, dynamic and static EA-parts) was included as an important criteria. The criteria from the category "preparation" were evaluated qualitatively and the results of the EMG-Analysis during the study were used for a quantitative evaluation of the muscles (Abakumov et al. 2014).

The following muscles proved to be especially valuable for EMG-Analysis in further field studies:

- *m. quadriceps femoris (vastus lateralis)*
- *m. tibialis anterior*
- *m. gastrocnemius (medialis)*
- *m. gastrocnemius (lateralis)*

These muscles are quickly and easily prepared for measuring and show obvious EMG-activities (Wakula et al. 2016b).

Similarly, the *m. gluteus medius* proved to be relevant for field studies as it showed promising values for electrode placement, MVC-measuring and the recording of muscular activity. The main issue with this muscle is that an artery is very close to the muscle and that the pulse will superimpose the resting measurement.

The *m. biceps femoris* showed the highest differences in the individual evaluation. During preparation it proved to be very challenging as it is very hard to find.

Especially when walking “sideways counter-clockwise”, which is the main stressor in an U-line, the muscles *m. quadriceps femoris* (vastus lateralis), *m. tibialis anterior and m. gluteus medius* show a significant muscular activity. The activity in these muscles is lower for “normal” walking and standing.

Overall, the six muscles in the study 1 show higher EA-values for walking (“normally” as well as “sideways counter-clockwise”) than they do for standing.

The results of the EMG recording in the lab study 2 prove that when “walking sideways counter-clockwise”, some muscles of the right leg are under greater stress than the muscles in the left leg.

5 Conclusions

5.1 Main Conclusions from Lab Study 1

A comparison of the four scenarios when “walking” showed that the increase in EA-mean values and the dynamic EA-parts have approximately the same value for 15 and for 30 s. From this we conclude that the muscle is under significantly greater stress when walking “sideways” than when walking “normally”.

When comparing the data for “standing” and “walking” it becomes evident, that the mean values and dynamic parts for “walking” are significantly higher than the EA-values for “standing”. This leads to the conclusion that the muscle is under greater stress when “walking” than when “standing”. The static EA-parts are low and remain at the same level of activity as they do for *m. quadriceps femoris*.

There can be derived the following four main conclusions from this lab study (Wakula et al. 2016a):

- “Walking sideways counter-clockwise” is often practiced by workers working on assembly lines that have work station that are very close together;
- “Walking sideways counter-clockwise” produce a higher stress factor than “walking normally counter-clockwise” does. We measured higher EA-values (especially dynamic EA-values) in four of six analysed leg muscles;
- When standing for 15 or 30 s there are slight differences in the strain level of the analysed muscle;
- The static EA-values were low for all analysed muscles (<3 % of MVC) and showed only very slight differences in the four scenarios.

5.2 Main Conclusions from Lab Study 2

The analysed three muscles of each leg show higher electrical activities (EA-mean values and EA-dynamic) in the right leg than in the left leg when “walking sideways counter-clockwise” in small simulated U-line. The detailed examinations are based on a laboratory study with six young subjects. In order to verify the presented findings statistically, further lab and especially field studies with additional and more diverse subjects are necessary.

Acknowledgement The authors thank the Employer's Liability Insurance Association wood and metal (Berufsgenossenschaft Holz und Metall, BGHM) and the Employer's Liability Insurance Association electrical engineering (Berufsgenossenschaft Elektrotechnik, BG ETEM) for their support of the project "U-line assembly systems", that made these studies possible.

References

Abakumov N, Behnisch P, Friedrich A, Gebauer C, Röth J, Stolz N (2014) Analyse der Muskelaktivität in den Beinen beim Stehen und Gehen bei U-Linie-Montagetätigkeiten. Tutorium am IAD, Technische Universität Darmstadt; J. Wakula und D. Müglich, Betreuer

AWMF online (2013) Oberflächen-Elektromyographie in der Arbeitsmedizin, Arbeitsphysiologie und Arbeitswissenschaft - Arbeitsmedizinische S2k-Leitlinie der Deutsche Gesellschaft für Arbeitsmedizin und Umweltmedizin (DGAUM) und der Gesellschaft für Arbeitswissenschaft (GfA). AWMF online. http://www.dgaum.de/fileadmin/PDF/Leitlinien/Leitlinienentwurf_Oberflaechen_Elektromyographie.pdf (Zugegriffen: 05.07.2016)

Kummer S (2009) Grundzüge der Beschaffung, Produktion und Logistik. Pearson Deutschland GmbH, Hallbergmoos

Strasser H (1996) Beanspruchungsgerechte Planung und Gestaltung manueller Tätigkeiten – Elektromyographie im Dienst der menschengerechten Arbeitsgestaltung. Ecomed Verlag, Landsberg/Lech

Wakula J, Müglich D, Bruder R (2016a) „Normales" vs. „seitliches Gehen" bei simulierten einfachen Montagetätigkeiten – Analyse der muskulären Beanspruchungen in den Beinen. In: Gesellschaft für Arbeitswissenschaft (Hrsg) Arbeit in komplexen Systemen. Digital, vernetzt, human?! GfA-Press, Dortmund

Wakula J, Fichtner K, Bruder R (2016b) Analyse der physischen Belastungen und der muskuläre Beanspruchungen an der simulierten U-Montagelinie mit unterschiedlichen Mechanisierungsgrad in der Prozesslernfabrik der TU Darmstadt. In: Gesellschaft für Arbeitswissenschaft (Hrsg) Arbeit in komplexen Systemen. Digital, vernetzt, human?! GfA-Press, Dortmund

Analysis of a Multimodal Human-Robot-Interface in Terms of Mental Workload

Marc Schneider and Barbara Deml

Abstract

The trend away from highly automated processes using heavy industrial robots towards HRI-systems (human-robot-interaction) using light-weight robots, is leading to a strong increase concerning the technical complexity of future industrial work systems. This article reports the results of an experimental evaluation of a multimodal human-robot-interface. Two typical application scenarios from the robot programming were examined, each performed with three different modalities of control. The mental workload of the operators was operationalized by subjective, objective and physiological indicators. There were significant differences in mental workload between the different types of control in both application scenarios. The present results demonstrate the benefits of adaptive system design in human-robot-interaction.

Keywords

Human-robot-interaction • Workload assessment • Adaptive system design

1 Introduction

The development of modern industry is characterized by decreasing product life cycles accompanied by an ever-growing variety of product variants. Future working places in the industrial sector will provide simultaneously increasing demands on the flexibility and repeatability (Neudörfer 2014). To cope with this growing dynamic, the high repeat accuracy of the robot is to be combined with the flexibility of the human in human-robot-cooperation (Haddadin 2014). Besides the benefits, such a combination provides an enormous increase in the complexity of future work

M. Schneider (✉) • B. Deml
ifab, KIT, Kaiserstraße 12, 76131 Karlsruhe, Germany
e-mail: Marc.schneider3@kit.edu; Barbara.deml@kit.edu

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_18

systems. The development of adaptive, human-centered systems can sustainable support the development of individual skills and knowledge, lead to well-being and assist workers of different age (Spillner 2015). An exemplary process for such a HRI-system is the teaching of a welding track. The individual coordinate points of the welding path can be taught to the robot, so that these can be followed, modeled or corrected afterwards. The present study analyzes two sub-processes with respect to the mental workload: path planning and trajectory correction. Each thread is in this case designed with three modalities of control to identify differences between the modalities of control depending on the current working process. Such a task dependent comparison of control modalities is an important and necessary condition for the user-centered design of adaptive interfaces for the complex HRI-work systems of the future.

The reminder of the article is structured as follows: First, the theoretical background with regard to the mental workload is explained before the used methods of workload assessment are described. After an overview of the experimental design, the data analysis will be presented and discussed afterwards.

2 Theoretical Background

The *mental workload* can be defined as the cognitive response of the *human information processing system* to external task load as a function of the personal prerequisites and the individual coping strategies (Packebusch 2003). The construct of mental workload is widely analyzed theoretically as well as empirically for many years until today (Manzey 1998). Before having a closer look at the construct of mental workload underlying the study of this article, it is necessary to explain the fundamental aspects of human information processing. One current approach is the *information processing loop* developed by Sträter (2005). This model is particularly suitable for explaining the physiological reactions caused by mental workload (e. g. blink rates, pupil dilations), as processing times and response times can be derived directly. The processing loop by Sträter (2005) is based on the world model. On the one hand there is an external world, including among others the working system and the environmental factors. The external world provides changeable information for the human and may require actions and responses. On the other hand there is the internal world which contains the individual representation of the external world build up on experiences, concepts and goals of the individual person. In the interaction of a human with a working system during a work task both worlds are continuously updating and enhancing each other with respect to the current state of the system and the actual goals of the human (Neisser 1976). This mutually adjustment is often conscious, but usually unconscious. In the case that the external and internal world fit well together, no adaptions are necessary and the system of both worlds is balanced. Humans are acting unconsciously until there is a stimulus-related disorder or change in the external world, so that no suitable representation in the inner world can be found. This condition is called *cognitive dissonance*. It is clear that cognitive processes equally refer to both worlds and there is an iterative

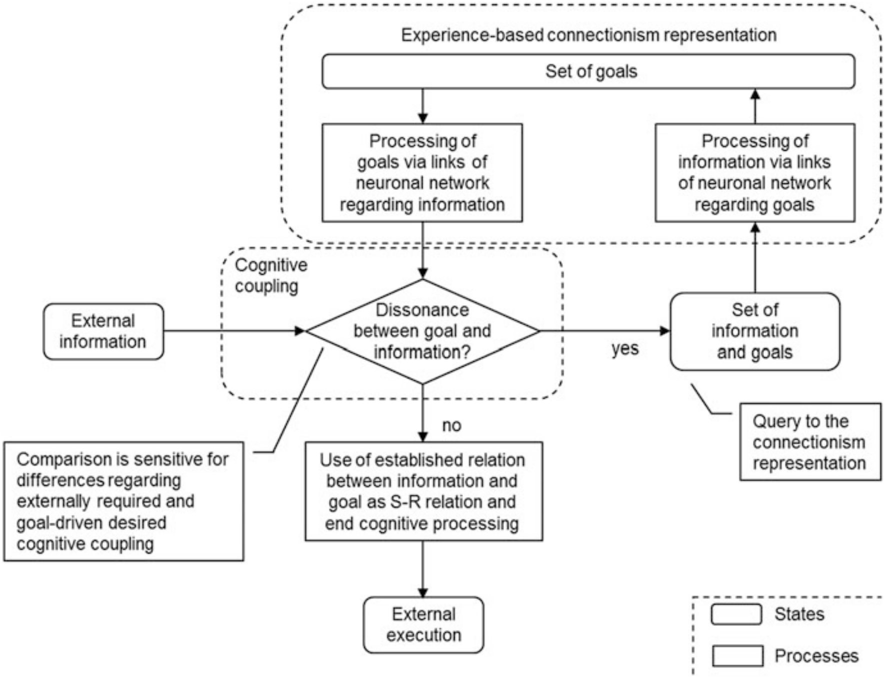


Fig. 1 The cognitive information processing loop (Sträter 2005, S. 79)

process of restoring the balance after a dissonance (Arenius 2009; Sträter 2005). The processing loop is built up on that continuous interaction between both worlds (Fig. 1).

External information that enters the processing loop becomes a task that is processed based on the existing internal information. Humans are able to act and react *skill based* in case of no, or only a very slight, dissonance between the external and internal world. If there is an incorrect adjustment of the worlds due to a large dissonance, the human needs to search *rule based* for similar internal information that can minimize the resulting dissonance. If the dissonance cannot be resolved properly more loops (*knowledge based*) are required (Rasmussen 1986). The processing loop allows a consideration of the dynamic changes of system parameters or other environmental influences which can lead to errors. The central comparator (the limbic system) selects the information and can be influenced by intentions and objectives (decision-making). The continuous adjustment of both worlds, their mutual influence and the compensation of any possible dissonances are important processes needed to be considered in the description and detection of mental workload (Sträter 2005). Over the duration of a task the behavior can adjust due to changes in the goals of humans or the environmental factors. Accordingly, the mental workload can change over time, even in the absence of obvious changes in task demands or task difficulty. Furthermore the model distinguishes between cognitive levels of information processing. On the one hand the information can be processed skill based (often unconscious). On the other hand greater dissonances require rule based or knowledge based processing of information (often conscious). In particular the unconscious cognitive processes make enhanced use of existing mental models, which may differ individually very strong. In addition to the variability of information processing over time it may be highly individual. These circumstances further complicate the detection of mental workload at a general level including complete samples of subjects. Before explaining the approach of a differential consideration of mental workload assessment used in this article the *resource theory* as one of a multiplicity of models trying to describe cognitive processing and mental workload is specified (Wickens et al. 2013). The capacity of mental resources is the upper limit of the information processing capability of human like modelled in the processing loop. The capacity and availability of resources are underlying certain elasticity depending on the task loads, the individual constitution and the level of training. An increase in task requirements may lead to an increase in resource mobilization up to a certain individual degree (Wickens et al. 2013). The ratio of resources provided and the performance of a task is often assumed to be linear until all resources are invested (Kahneman 1973). At that individual point the performance decreases with further increases of task demands. In the case that not all resources were needed to work on a primary task, the further resources can be used for the processing of additional tasks (Wickens 1984). If several tasks were processed simultaneously, they infer to different degrees. A variety of experiments examined the relationship between the demand level of a primary task and the performance in a secondary task. Many results indicate that interference between the tasks cannot be attributed to the level of requirements in the primary task (Wickens 1984). These results are consistent with the findings based on information processing loop—the model of multiple resources is exactly at this point. Much of the performance variation in dual-task situations can be attributed neither to the difficulty of the task, nor to the procedure of activation of resources (of the favored tasks). Differences in the efficiency of time sharing between two tasks are routed in differences concerning the requirements for different cognitive processing structures of the human. Various structures of information processing are supported by independent resources. The efficiency of the time sharing between two tasks is better when different structures are recruited. An example is the information perception by the eyes or ears, or the required response to a change of the information (Wickens 1984). For further descriptions of these four dichotomous dimensions within the cube model it should be made at this point to Wickens (1984). Based on the mentioned theoretical considerations, this article uses a differential approach for assessing mental workload. By the use of a general approach on sample level too much information about the individual workload characteristics is lost. However, considering the individual to raise conclusions for the design of systems is accompanied by great effort concerning analysis and implementation. The differentiate approach is a trade-off between cost and accuracy, at which sample is divided into groups based on various characteristics (Driver and Mock 1975). The study results of Fels et al. (2015) demonstrate that such a differentiation contributes reliable predictions in the field of mental workload assessment. The number of adaptive cognitive processes is used to as criterion for the discrimination of the clusters. Through the mutual updating and changing of the internal and the external world, the human-machine-interaction and thus also the mental workload changes during time on task (Sträter 2005). The adaption processes take place both consciously and unconsciously. Both types are differentiated and take into account for the classification of the clusters. In the following section with used methods the physiological indicators of mental workload are distinguished with respect to their sensitivity to both types of information processing.

3 Methods

To measure the construct of mental workload a tri-modal approach is used in this article. Within this approach subjective, physiological and performance-based parameters are measured in order to conclude the mental workload.

The theoretical chapter already explains a part of the diverse inter- and intra-individual factors that influence the mental workload during a working task. To account for the greatest possible proportion of the influencing factors, it is necessary to consider mental workload from three perspectives (Fig. 2). For example negative changes in performance can be routed back to increased mental workload, but also to a decline in effort of the subjects. In this case, the performance data can be compared to the subjective and physiological data. The following sections will explain the used indicators and methods in detail.

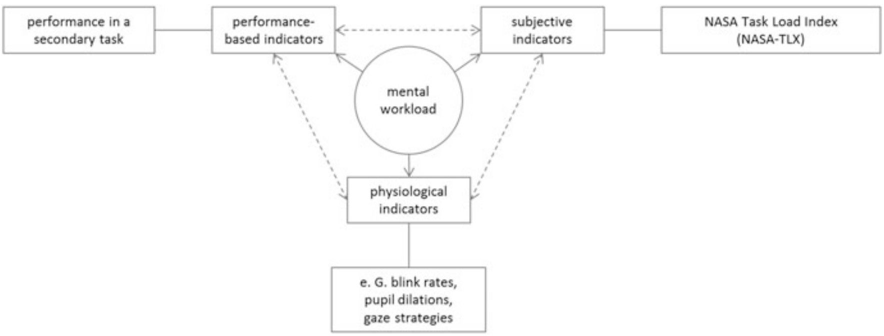


Fig. 2 Tri-modal approach for mental workload assessment

3.1 Subjective Perceived Workload and Performance-Based Measurement

Depending on the research object different measurement instruments for both subjective perceived workload and the performance can be the most reliable and sensitive choice. According to Hart and Staveland (1988) the subjective workload assessment with a questionnaire or interview is the only method having access to the essentials of workload. Subjective indicators are the only way to find out the subjective perceived effects of task loads. Besides this advantage there are several factors biasing the subjective indicators. Unexperienced participants often have problems verbalizing or assessing their mental workload correctly. For this reason, there are some criteria to evaluate subjective measurement instruments (sensitivity, Diagnosticity, validity, intrusiveness, reliability, implementation, acceptability). Following Rubio et al. (2004) the NASA Task Load Index (*NASA-TLX*) has a very high validity, reliability as well as acceptability. The NASA-TLX developed by Hart and Staveland (1988) is rather widespread and economical, well-structured and easy to understand. Rubio et al. (2004) point out weaknesses of the NASA-TLX regarding the sensitivity and Diagnosticity. In former studies it could be shown, that the questionnaire is able to reveal significant differences in task difficulty with respect to mental workload measured by physiological and performance-based measures (Fels et al. 2015; Schneider and Deml 2016). The NASA-TLX consists of six items (mental demand, physical demand, temporal demand, performance, effort and frustration). For this study the questionnaire is used without the item of physical demand, because it is expected to be not relevant for the assessment based on the results of a former study (Schneider and Deml 2016). After each task the subjects rate all five items on 20-stage scale ranging from low to high.

Based on the resource theory by Wickens (1984) a secondary task is used in this study to assess the performance of the subjects. The subjects were instructed to focus on the main task, the interaction with the robot, and edit the secondary task only if they have spare capacity. According to the resource model the secondary task measures the remaining, available resources. The available resource capacity and the resource requirements of the main task can be concluded by the failure rate of the secondary task (Wickens et al. 2013). As secondary task a so called n-back task is used. Within this task there is a code, consisting of a letter and a number (e.g. A9 or C3), visually presented on an additional screen every 4 s. Each time there is a new code presented the subjects have to represent the previous one verbally.

3.2 Physiological Indicators

Physiological indicators use the reactions of the human body to detect mental workload continuously and with high temporal resolution (Manzey 1998). In this study only ocular parameters, recorded by a binocular head mounted eye tracking system, should be regarded. The ocular parameters have the advantages that the recording with a head mounted system has a very high user acceptance. The subjects cannot be seen or identified from the recordings made by the field camera and both eye cameras. Furthermore there is a very short latency between the cognitive system and the eyes due to the anatomical arrangement (Sträter 2005). For cross validation and the control over other constructs which may influence the ocular parameters (e.g. fatigue or sleepiness), there are six parameters assessed in this study. Following these six parameters should be explained in detail and distinguished by the sensitivity to either conscious or unconscious mental processes.

3.2.1 Unconscious Parameters

The following parameters can be classified as particularly sensitive to unconscious cognitive processes as they may not be influenced deliberately during time on task. After a task they cannot be reported or verbalized by the subject (Martens and Wyble 2010; van Gaal and Lamme 2012). The first parameter in this category is the *blink rate*. Commonly, an eye blink takes about 70–100 ms (Marshall 2000) and the blink rate describes the number of eye closures in a pre-defined period of time (here: 30 s). Wilson (2009), for example, showed that there is a significant negative relationship between the blink rate and the increasing difficulty of certain temporal phases of a simulated flight. However Cardona and Quevedo (2014) could not find any significant correlation between the blink rates of car drivers and their performance at different complexity levels. This may also be due to the fact that blink rate is not only influenced by the mental workload, but it is also altered by other user states, such as the degree of sleepiness (Watling and Smith 2013). As shortly mentioned above the *blink duration* can take between 70 and 500 ms and can also be used to assess mental workload. Benedetto et al. (2011) suggest that the blink duration is even more sensitive to the cognitive workload than the blink rate. They tested both parameters in a lane change task with two conditions, control condition and dual task condition. In a study by Martins and Carvalho (2015) the parameter *PERCLOS* (percentage of eye closure) was derived, which is a rather well established parameter to predict fatigue and which is defined as the accumulated blink duration on a pre-defined time interval. The PERCLOS is also a sensitive measure for the mental workload like, among other, the study by Brookhuis and Waard (2010) showed in a driving scenario. The last parameter sensitive to unconscious cognitive processes is the *pupil dilations*. The pupil of the human eye may be compared to an optic lens, which changes its size in order to regulate the incidence of light into the eye. The pupil dilations are controlled by two antagonistic muscles surrounding the pupil: the Sphincter Pupillae, which scales down the pupil size para-sympathetically, and the Dilator Pupillae, which scales up the pupil size sympathetically (Beatty and Lucero-Wagoner 2007). Besides changes in pupil size due to an adaption to the light conditions or the distance of the regarded object, there is also the psycho-sensory reflex (Loewenfeld 1993). This reflex is triggered by neurocognitive activities over different muscles-depending paths, indicating that the fluctuation in pupil size is also related to mental workload (Steinhauer et al. 2004). However, the separation of the individual reflexes from the signal of pupil diameter is a challenging task. Marshall and Sandra (2000) introduced the *Index of Cognitive Activity* (*ICA*), which succeeded the separation of the psycho-sensory reflexes. In the presence of psycho-sensory reflexes the signal, created from the continuous recording of pupil size, is characterized by abrupt discontinuities. In the presence of effortful cognitive processing, the pupil responds rapidly with a reflex reaction (Marshall 2002). Marshall (2000) uses the wavelet analysis which allows analyzing the frequency of a signal without losing the reference of time. There are several recent studies among others which show explicit the significant relation between the ICA and subjective as well as objective reference measurements (Marshall 2007; Marshall et al. 2003; Schwalm 2009).

3.2.2 Conscious Parameter

The parameters described below can be consciously influenced by subjects during a task. They can be placed in a direct relationship with the strategy used for information perception. Due to the fact that changes in search behavior or gaze strategy, which affect these parameters, can mostly be verbalized by the participants after the task, these processes can be assigned to the conscious cognitive processes (Goodale and Milner 1992). Due to the reason of a restricted area of sharp seeing, eye movements are needed, to ensure that the retinal image of the interesting object is mapped completely and correctly on the retina. The alignment of the fovea, the "gaze" on an object, is called fixational eye movement. It could have been shown that in high workload situations the *fixation duration* is directly related to the time, which is needed to either gather information or to understand a certain problem at hand (Backs and Walrath 1992). In other words, if the mean fixation duration is short, a higher amount of time may be dedicated to searching behavior. In this study, too, the fixation duration is considered, whereby the mean value of all fixations in a defined period of time (30 s) is regarded. Further the *Nearest Neighbor Index* (*NNI*) is a spatial statistics algorithm, which in the context of eye-tracking expresses the proximity of each fixation relative to all other surrounding fixations. In other words, it assesses the randomness of fixation patterns (Di Nocera et al. 2007). Thereby a ratio between the average of the observed minimum distances between fixations and the mean distance, which would to be expected for a random distribution of fixations, is derived. In consequence, for a random distribution a ratio that is equal to 1 is to be observed. For values less than 1 the mean of minimum observed distances is smaller than the random one, which indicates a clustering of fixations and which stands for so called informational gaze patterns. For values larger than 1 a regular, non-random gaze behavior is given (Di Nocera et al. 2007). According to Di Nocera et al. (2007) the value range for NNI lies between 0, which corresponds to maximum clustering, and 2.1491, which accounts for a strictly regular hexagonal pattern. It is assumed that under high stress conditions the gaze behavior does not follow a clear strategy and that it is characterized by searching for information randomly (NNI = 1). There is a lot of empirical evidence that NNI reveals a significant correlation with both subjective as well as objective performance parameters of cognitive workload (Halverson et al. 2012).

4 Experimental Setting

For the evaluation of the multimodal human-robot-interface, two typical scenarios from the robot programming have been studied with three input modalities. In the first scenario the subjects had to plan and train a welding path to the robot whereas in the second part they had to control and correct the trained path. Right before the experiment the subjects were instructed in detail with training sequences until they feel comfortable with the controls and the handling of the robot. In the first application scenario the subjects had to teach a predefined welding path to the robot. The path here was composed of seven contour points that should be taught sequentially. The second application scenario involved the control and correction of the taught welding path. For this, the subjects were instructed to follow the path until they reach the fourth of seven points and delete this one. Afterwards they had to check the path with now six points again. The used welding path was a modified version of typical welding processes of body parts in the automotive manufacturing (Fig. 3).

Both scenarios have to be performed three times by each subject at which the sequence of used control modalities was randomized. With voice and gesture control, the subject used defined voice or gesture commands for saving and deleting points or for driving to already memorized points. In the case of the path planning the robot has been moved to the new points by a motor torque sensor and they were then stored with a voice or a gesture command. The third control modality was a tablet PC with the path planning and correction performed using a graphical interface. While the subjects work on both application scenarios they wore a head-based eye tracking system (Ergoneers Dikablis Professional) and continuously processed also the secondary task (mental rotation). After each trial the

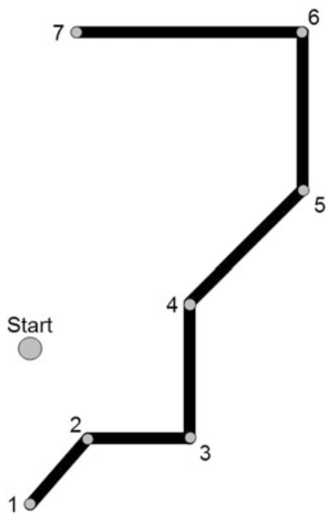


Fig. 3 Welding path of application scenario 1 and 2

subjects were asked to fill out the NASA-TLX, so that questionnaire was completed six times (2 tasks × 3 control modalities) by each subject.

5 Data Analysis

The study based on a predominantly student sample with $N = 19$ subjects who were 24.14 years old on average ($SD = 3.25$). There were 16 male and 3 female subjects with a total of 5 subjects wore glasses. The data were analyzed using analysis of variance with respect to measurement repetitions as well as correlation analysis.

5.1 Planning of Welding Path

There were significant main effects between the control procedures concerning the path planning—both for the mean values of NASA-TLX ($F(2, 36) = 7.08$, $p < 0.01$) as well as for the error rate in the secondary task ($F(2, 36) = 3.99$, $p < 0.05$). The tablet control induced significantly lower workload than voice or gesture control in both measurements (Fig. 4).

The analyses of the individual items of the NASA-TLX offer further interesting facts for the evaluation of the controls. There were no significant differences concerning the item "temporal demand" between the control modalities, so it can be assumed that this parameter did not affect the overall assessment. Moreover the item "frustration" was very low on average ($M = 7.95$, $SD = 1.70$), which indicates that isolated non-recognitions of voice or gesture commands did not increase the experienced workload.

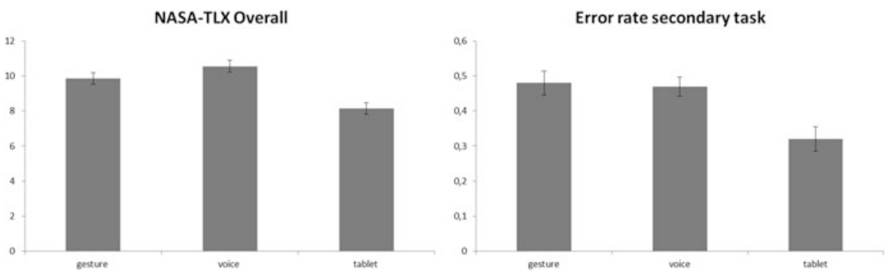


Fig. 4 Path planning: NASA-TLX values and error rate of secondary task

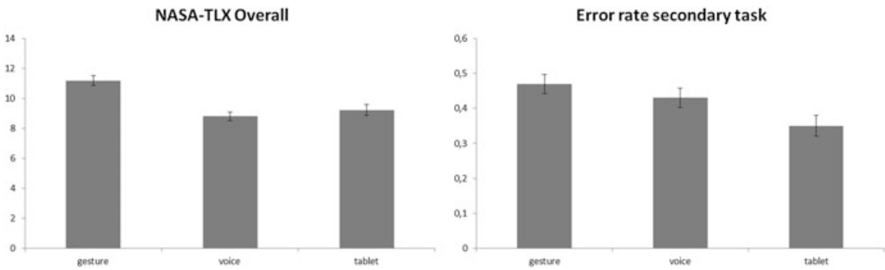


Fig. 5 Path correction: NASA-TLX values and error rates of secondary task

5.2 Trajectory Control and Correction

There were also significant differences between the control modalities for the control and correction of the trained path. Both the NASA-TLX ($F(2, 34) = 7.47$, $p < 0.01$) and the performance in the secondary task ($F(2, 34) = 3.89$, $p < 0.05$) showed these main effects (Fig. 5). In this application scenario the subjective perceived workload differed from the performance in secondary task. The subjective evaluations showed only that the workload of the gesture control was experienced significantly higher than both other modalities. The error rate, however, was significant lower using the tablet control in contrast to both other modalities. In the case of the tablet control it was not necessary to recognize any command on the feedback-screen which led to increased fixation durations on secondary task. Therefore the error rate was not representative with the tablet control for path correction. According to the NASA-TLX scores the voice control provided the lowest workload in this application scenario.

5.3 Analysis of Physiological Data

By using a new approach of Schneider and Deml (2016) for the grouping of a sample to workload-related adaptation strategies (consciously and unconsciously) based on theoretical background mentioned in the beginning of this article, there could be found highly significant correlations between all physiological parameters and the NASA-TLX values on the differential (grouped) level (Table 1). The sample was classified based on the characteristic of the subject's individual adaption strategies. Conscious adaptations during time on task, for example, changed the view strategy, which directly impacts the NNI values, but also the fixation duration. A subject changed the strategy not only aware of his behavior, but also unconsciously. Unconscious adaptations showed, among others in the temporal change of the pupil diameter or the blink rate as well as the blink duration (Schneider and Deml 2016). These different adaptive strategies of each cluster can be described highly reliable through different physiological parameters depending on the degree to which the subjects of each cluster adapt either conscious or unconscious. A

Table 1 Correlation coefficients physiological and subjective data

	Blink rate	Blink duration	PERCLOS	Pupil dilations	NNI	Fixation duration
NASA-TLX	$r = -0.62^{**}$	$r = -0.58^{**}$	$r = -0.63^{**}$	$r = 0.66^{**}$	$r = 0.70^{**}$	$r = 0.60^{**}$

$^{**}p < 0.01$
The mean values of the correlation coefficients base on the Fischers' z-transformation

standardized evaluation on the level of the whole sample was not possible due to the individually different behavior. With these results the study confirmed the theoretical implications of the cognitive information processing loop (Sect. 2).

6 Discussion

All three analyzed workload indicators of the tri-modal approach showed a consistent result with respect to the assessment of the control modalities. For the path planning the tablet control showed the lowest workload values ($M_{TLX} = 8.14$, $SD = 3.06$), whereas in the case of trajectory correction the voice control induced the lowest workload ($M_{TLX} = 8.80$, $SD = 2.62$). The results of this study showed clearly that an adaptive and multimodal interface design could lead to an optimized workload situation in human-robot-interaction. However, further studies are needed to combine the control modalities in one application scenario and to verify that a dynamic change of the modalities may lead to an optimization of the overall workload situation. Concerning the mental workload assessment, the results of a clustered evaluation confirm the theoretical considerations. It is difficult to describe a whole sample with a uniform physiological parameter configuration due to the internal world, which influences individually the perceptions and actions. There are significant changes in information processes during time on task. For example the prioritization between main and secondary task varies greatly between subjects and also time-dependent. In upcoming studies, the classification of the sample clusters should be further investigated in order to incorporate the theoretical and practical insights into the future development of sustainable systems.

References

Arenius M (2009) Getting the feeling – "Human Error" in an educational ship-handling simulator. Magisterarbeit. University Linköping, Linköping. Institut of Computer Science, checked on 1/24/2016

Backs R, Walrath L (1992) Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. Appl Ergon 23:243–254

Beatty J, Lucero-Wagoner B (2007) The pupillary system. In: Cacioppo JT, Tassinary LG, Berntson GG (eds) Handbook of psychophysiology, 3rd edn. Cambridge University Press, Cambridge, pp 142–162

Benedetto S, Pedrotty M, Minin L, Baccino T, Re A, Montanari R (2011) Driver workload and eye blink duration. Transport Res 14:199–208

Brookhuis K, de Waard D (2010) Monitoring drivers' mental workload in driving simulators using physiological measures. Accid Anal Prev 42:898–903

Cardona G, Quevedo N (2014) Blinking and driving: the influence of saccades and cognitive workload. Curr Eye Res 39:239–244

Di Nocera F, Camilli M, Terenzi M (2007) A random glance on the flight deck: pilots' scanning strategies and real-time assessment of mental workload. Edited by University of Rome, Rome

Driver MJ, Mock TJ (1975) Human information processing, decision style theory, and accounting information systems. The Account Rev 50:490–508, checked on 2/3/2016

Fels M, Bauer R, Gharabaghi A (2015) Predicting workload profiles of brain–robot interface and electromyographic neurofeedback with cortical resting-state networks. Personal trait or task-specific challenge? J Neural Eng 12(4):46029, DOI: 10.1088/1741-2560/12/4/046029

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15:20–25

Haddadin S (ed) (2014) Towards safe robots. Approaching Asimov's 1st law. Springer, Heidelberg (Springer tracts in advanced robotics, 90)

Halverson T, Estepp J, Christensen J, Monnin J (2012) Classifying Workload with eye movements in a complex task. Proceedings of the Human Factors an Ergonomics Society Annual Meeting 58:168–172

Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) Human mental workload. North Holland Press, Amsterdam

Kahneman D (1973) Attention and effort. Prentice-Hall, Englewood Cliffs, NJ, (Prentice-Hall series in experimental psychology), checked on 1/20/2016

Loewenfeld I (1993) The pupil: anatomy, physiology, and clinical applications. Iowa State University, Ames

Manzey D (1998) Psychophysiologie mentaler Beanspruchung. In: Rößler F (ed) Enzyklopädie der Psychologie. Ergebnisse und Anwendungen der Psychophysiologie, vol 5, 5th edn. Hogrefe, Göttingen, pp 799–864

Marshall SP (2000) Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity on 7/18/2000. Patent no. 6.090.051

Marshall SP (2002) The index of cognitive activity: measuring cognitive workload. IEEE Human Factors Meeting, 7, Scottsdale

Marshall SP (2007) Identifying cognitive state from eye metrics. Aviat Space Environ Med 78:165–175

Marshall SP, Pleydell-Pearce CW, Dickson BT (2003) Integrating psychophysiological measures of cognitive workload and eye movements to detect strategy shifts. In: Proceedings of the thirth-sixth annual Hawaii international conference on system sciences. IEEE, Los Alamitos, CA

Martens S, Wyble B (2010) The attentional blink: past, present, and future of a blind spot in perceptual awareness. Neurosci Biobehav Rev 34:947–957

Martins R, Carvalho JM (2015) Eye blinking as an indicator of fatigue and mental load – a systematic review. In: Arezes PM (ed) Occupational safety and hygiene III. Selected extended and revised contributions from the International Symposium on Safety and Hygiene, Guimarães, Portugal, 12–13 February 2015. CRC Press/Balkema, Leiden, The Netherlands, pp 231–235

Neisser U (1976) Cognition and reality. Principles and implications of cognitive psychology. W.H. Freeman, San Francisco, CA

Neudörfer A (ed) (2014) Konstruieren sicherheitsgerechter Produkte. Methoden und systematische Lösungssammlungen zur EG-Maschinenrichtlinie. 6., aktual. Aufl. Springer Vieweg, Berlin (VDI-Buch)

Packebusch L (2003) Psychische Belastung und Beanspruchung - Normung für die Praxis. In: Wirtschaftspsychologie aktuell. Arbeit und Gesundheit, vol 4, 4th edn. With assistance of R. Decker. Deutscher Psychologen Verlag GmbH, pp 32–36

Rasmussen J (1986) Information processing and human-machine interaction. An approach to cognitive engineering. North-Holland, New York (North-Holland series in system science and engineering, 12)
Rubio S, Díaz E, Martín J, Puente JM (2004) Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX, and workload profile methods. Appl Pyschol 53:61–86
Schneider M, Deml B (2016) An integrated approach to mental workload assessment. In: Deml B, Stock P, Bruder R, Schlick C (eds) Advances in ergonomic design of systems, products and processes. Proceedings of the Annual Meeting of GfA 2015, vol 1, 1st edn. Springer, Berlin, pp 191–208
Schwalm M (2009) Pupillometrie als Methode zur Erfassung mentaler Beanspruchungen im automotive Kontext. Dissertation, University of Saarland, Saarbrücken
Spillner R (2015) Einsatz und Planung von Roboterassistenz zur Berücksichtigung von Leistungswandlungen in der Produktion. Utz, Wiss, München (IWB-Forschungsberichte, 296)
Steinhauer S, Siegle G, Condray R, Pless M (2004) Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. Int J Psychophysiol 52:77–86
Sträter O (2005) Cognition and safety An integrated approach to systems design and assessment. Ashgate, Aldershot
van Gaal S, Lamme VAF (2012) Unconscious high-level information processing: implication for neurobiological theories of consciousness. Neuroscientist 18:287–301
Watling C, Smith S (2013) In the blink of an eye: the circadian effects on ocular and subjective indices of driver sleepiness. Australasian Road Safety Research, Policing & Education Conference, Brisbane
Wickens CD (1984) Processing resources in attention. In: Parasuraman R, Davies DR (eds) Varieties of attention. Academic Press, London, pp 63–102
Wickens CD, Hollands JG, Banbury S, Parasuraman R (2013) Engineering psychology and human performance, 4th edn. Pearson, Boston, MA
Wilson GF (2009) An analysis of mental workload in pilots during flight using multiple psychophysiological measures. Int J Aviat Psychol 12:3–18

How to Evaluate the Usability of Smart Devices as Conceivable Work Assistance: A Systematic Review

Jan Terhoeven and Sascha Wischniewski

Abstract

The ongoing digitalization of work and working environments suggests that the application of smart devices as a work assistance will increase. In this context, particularly it is important to consider an ergonomic design, especially the usability of mobile interactive systems before using it in the working environment. Therefore, a specific significance will be attributed to the usability evaluation of smart device applications. Consequently the Federal Institute for Occupational Safety and Health (BAuA) implemented a study by means of a systematic review, which examined how to evaluate the usability of smart devices. The aim of the review was to investigate, if there are specific methods to evaluate the usability of smart devices or how far consisting methods including their criteria from the area of conventional screen work can be adopted. Results indicate that the majority of existing studies used evaluation criteria from the area of conventional screen work or recently developed kits for mobile usability evaluation, which are built on similar criteria. Actually the established methods and criteria are still adequate, but they have to be adjusted to the rapid development of new technologies with respect to new influence factors like innovative interaction concepts, smaller displays or mobile operations.

Keywords

Mobile devices • Usability • Mobile user interface design

J. Terhoeven (✉) • S. Wischniewski
Federal Institute for Occupational Safety and Health (BAuA), Dortmund, Germany
e-mail: terhoeven.janniklas@baua.bund.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_19

1 Introduction

The increasing performance of computer systems and digitalization of the working environments result in significant impacts on the organization, contents and requirements of future work. Especially against the background of Industry 4.0, a decline of manual work and routine work with a simultaneous increase of knowledge-intensive work is expected. Employees will assume more and more monitoring and controlling tasks, job content are becoming more and more complex and previously rigid workplaces are becoming mobile with a decentralization of decisions (Spath et al. 2013). Considering the demographic development and an imminent shortage of skilled professionals, the changes, which occur in the course of the digitalization, lead to large challenges for the work organization and design. A popular approach of the practice-oriented research and development to meet these challenges is to integrate “smart devices” (e.g. smart glasses, smartphone, tablet-pc) as a work assistance system for an appropriate provision of information as well as for the interaction between humans and production systems (Schuh et al. 2014; Ittermann et al. 2015).

From the perspective of occupational safety, it is crucial that health and well-being of employees profit from the implementation of technological innovations as an important key factor for increasing the competitiveness of companies. Considering that, there are a lot of different factors to verify from the ergonomic point of view before using smart devices as work assistance. In this process not only the physical and mental stress during the use (Terhoeven et al. 2015; Grauel et al. 2014), but especially the usability are important. According to § 3 of the Ordinance on Industrial Safety and Health (BetrSichV), the employer has to assess all occurring hazards prospectively, which arise with the use of work equipment (hazard assessment). In particular, this comprises the consideration of the usability, including an ergonomic design according to the age and ageing of employees.

In this paper, an actual review of the Federal Institute for Occupational Safety and Health (BAuA) regarding the usability evaluation of smart devices is presented. Therefore, it was investigated by means of a systematic review, if there are specific methods to evaluate the usability of smart devices or how far consisting methods including their criteria from the area of conventional screen work can be adopted. For this reason, initially the state of the art of usability evaluation as well as the research aim of the review will be described. Subsequently the methodical approach of the systematic review will be described. In the following the representation and discussion of the results take place together with an outlook for future research needs.

2 The Systematic Reviewing Process

2.1 State of the Art and Research Question

The requirements related to the usability of an interactive system and the criteria for evaluating this are clearly defined by the DIN EN ISO 9241-11 (1998). Thereby, the design of the interaction and the presentation of information have the strongest impact on the usability. For these influence factors common methods in the usability or software engineering exist, mostly in form of usability heuristics. These common heuristics contain different criteria to evaluate the usability of software, but the criteria also have to be considered prospectively in the development phase. In case of the interaction design exist common evaluation kits like the usability heuristics of Jakob Nielsen (1993), the eight golden rules by Shneiderman and Plaisant (2005) or the DIN EN ISO 9241-110 (2008). All of them give advices how to design the human-machine-interface and which criteria can be used to evaluate it. In this process appears a high content-related congruence regarding the individual criteria of the different evaluation kits. The heuristics of Nielsen are the most detailed criteria, but the other methods are summarizing Nielsen's individual criteria on a higher level and are not less precise in doing so. A comparing overview of the criteria of the three described references contains Table 1.

As well there are established evaluation kits for the designing and evaluating the presentation of information to ensure the usability of interactive systems. The DIN EN ISO 9241-12 (2000) and the requirements concerning information design by

Table 1 Comparison between the usability criteria of three common evaluation kits

#	Nielsen (1993)	Shneiderman and Plaisant (2005)	DIN EN ISO 9241-110 (2008)
1	Visibility of system status	Offer informative feedback	Self-descriptiveness
2	User control and freedom	Support internal locus of control	Controllability
3	Consistency and standards	Strive for consistency	Conformance to expectations
4	Recognition rather than recall	Reduce short-term memory load	Conformance to expectations
5	Error prevention	Offer simple error handling	Error tolerance
6	Help users recognize, diagnose and recover from errors	Permit easy reversal of actions	Suitability for learning
7	Flexibility and efficiency of use	Enable to use shortcuts	Possibility of individualization
8	Aesthetic and minimalist design	Design dialog to yield closure	
9	Match between system and real world		Appropriateness for the task
10	Help and documentation		

Wickens and Hollands (2000) are given as an example at this point. However, all these evaluation kits for ensuring the usability of interactive systems by an appropriate interaction and information design are dealing with conventional screen work instead of using smart devices with significant smaller screens and other control systems like touch-screen.

For this reason, the Federal Institute for Occupational Safety and Health (BAuA) initiated a systematic review to sight the current state of studies about evaluating the usability of smart devices, especially against the background of a potential work assistance system. It was questioned how existing criteria and procedures for the evaluation and consideration of the usability in designing interactive systems can apply for Smart Devices or whether they have to be adapted to current technologies.

2.2 Methodical Approach

In order to collect the current state of studies about evaluating the usability of smart devices, especially in the context of a potential use as work assistance, a systematic review was chosen as the fundamental method. The procedure was inspired by Mattioli et al. (2010) and Arksey and O'Malley (2005), who describe a systematically performed review based on appropriate search strings. The literature search took place inside the five big and established databases EBSCOhost, PubMed, ScienceDirect, Scopus and Web of Science as well as their embedded smaller databases.

For the subject orientation in the systematic review process three categories were chosen to classify the search terms. The first category contained all search terms around the core issue of usability ("usability", "perceived use", "user-friendliness", etc.), against which the second category dealt with search terms related to the topic smart devices, i.e. potential technologies and conflating terms for technology groups (e.g. "smartphone", "mobile device", "head-mounted display"). By the third category, using terms such as "study", "measurement" or "test", the review aspired publications, which deals with the evaluation of different applications of smart devices regarding the usability. In all categories English as well as German search terms were used. In Table 2 below all search terms are aggregated, which were used in the systematic review.

After defining all terms and their classification into one of the three categories, a scheme for an iterative identification of relevant publications was set up. For this purpose initially the different possible and relevant search fields for the review process were transferred on a truth table based on propositional logic. As relevant search fields the title (T) were selected on the one hand and abstract or keywords (A) on the other hand. Consequently search classes for clustering the results were defined depending on whether the particular search term category is located in the title or abstract/keywords. In this process the search term category "usability" was set of top priority, the category "smart device" out in the first place, the category of "smart devices" behind it and the category "evaluation" on the third spot, from which the search classes TTT, TTA, TAT etc. resulted (Table 3). The defined

Table 2 Allocation of search terms to the search term categories

Terms of usability	Terms of smart devices	Terms of evaluation
"usability"	"smart device"[a], "smart-device"	"study", "studies"
"usability", "utility"	"wearable device", "wearable-device", "wearable"	"evaluation", "evaluate", "evaluating"
"perceived use"	"mobile device", "mobile-device"	"test", "testing"
"perceived usefulness"	"head mounted display", "head-mounted display"	"measurement", "measure", "measuring"
"user-friendliness"	"smart glass", "google glass", "HMD"	"Studie"
"serviceability"	"smart phone", "smartphone"	"Bewertung"
"Gebrauchstauglichkeit"	"smartwatch", "smart watch"	"Untersuchung"
"Nutzerfreundlichkeit"	"tablet", "tablet pc", "tablet-pc"	"Evaluierung"
"Bedienfreundlichkeit"	"tablet computer", "tablet-computer"	"Untersuchung"
"Bedienerfreundlichkeit"	"Mobiles Endgerät", "Mobile Endgeräte"	

[a]Depending on the databases the plural of each search term was added separately

Table 3 Search classes and rating score depending on the search term categories

Usability	Smart devices	Studies	Rating score
T[a]	T	T	5
T	T	A[b]	4
T	A	T	4
A	T	T	3
T	A	A	2
A	T	A	1
A	A	T	1

[a]Search term category is located in the publications title
[b]Search term category is located in the publications abstract or keywords

search classes were not only appropriate to assort the detected publications, but conduced to rate the results for a focused screening later on. For this purpose, each search class got assigned an individual rating score (Table 3). If search terms from each category were found in the title of a search hit (TTT), the publication got five points as rating score. If there were only two of the three search term categories existent in the title (e.g. TAT), the publication got three points as rating score. For only one search term category in the title, the search hit was assessed with one point (e.g. ATA). Because the systematic review in particular dealt with the topic usability, a publication with one or three points got an extra point in the rating score, if the search term category "usability" was detected in the title (e.g. TAA).

Using the previously described preliminary works, it was possible to prepare search strings for each of the five databases in order to collect the population of relevant publications. Based on the individual syntactical rules of the databases, a separate search string was defined for each of the seven search classes. The combination of five databases and seven search classes resulted in 35 search strings in total, which ensured the systematic process of organizing the hit list of each database. The search strings were such structured, that initially the title level and subsequently the level of abstract and keywords have been integrated. Exemplary for the structure, the search string for the database Scopus in the search group TTA with the search terms 1 to n appears as follows:

> TITLE ((usability 1 OR usability 2 . . . OR usability n) AND (smart device 1 OR smart device 2 . . . OR smart device n) AND NOT (evaluation 1 OR evaluation 2 . . . OR evaluation n)) AND (ABS (evaluation 1 OR evaluation 2 . . . OR evaluation n) OR KEY (evaluation 1 OR evaluation 2 . . . OR evaluation n))

The actual process of the review was divided into four phases. The first phase represented the identification of the overall study sample, aggregated across all databases using the search strings and an additional hand search. In the systematic review only studies as from 2007 were considered, because this was the formative year with the iPhone as supposedly first smartphone coming on the consumer market. Firstly, it was necessary to eliminate all duplicates out of the hit list as well as outliers, which came about because of dual word meanings (e.g. the search term “tablet” leads to many results in the field of drug studies). Based on an initial screening on the title-abstract level more matches were sorted out subsequently in the second phase of the review process, because they differed thematically from the analysis focus. Thus, a first evaluation basis for a more detailed screening with regard to the analysis focus was build up. This basis was assessed in the third phase of the review process using the rating system described above (Table 3). By means of the rating system an appropriate amount of hits could be filtered out for the in-depth screening on abstract and full-text layer in the fourth phase. For this purpose all studies with five or four rating points, about one-third of the matches with three or two rating points and randomly scattered publications with one rating point have been considered furthermore. Due to the in-depth screening the effective amount of relevant publications has been determined, of which the contained explanations regard the topic and the full-text was be analysed in detail concerning the objective of the review.

3 Results

Using the defined search strings across all databases in the first phase of the systematic review, a population of n = 3523 publications dealing with at least one of the search term categories was identified. However, an amount of n = 2551 hits had du be excluded because of being duplicates or outliers due to dual word meanings. Thus, n = 962 hits were taken over in the second phase of the review.

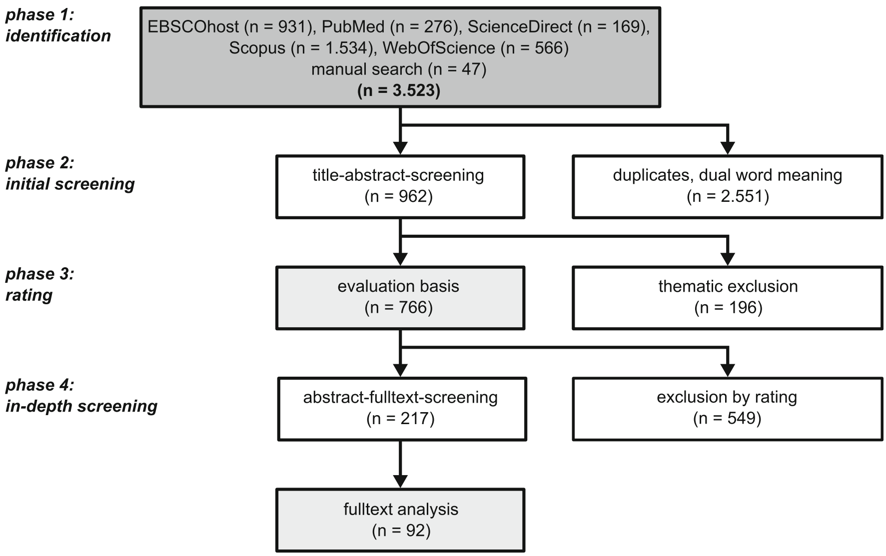


Fig. 1 Review process and results of the individual phases

Based on the initial screening on title-abstract level another n = 196 matches were thematically excluded. Consequently an evaluation basis of n = 766 matches remained, which could be thematically classified as publications dealing with the evaluation of usability in respect of smart devices. By means of the subsequent rating process in the third phase of the review additional n = 549 publications were excluded as described in Sect. 2.2. In the in-depth screening on abstract and full-text level ultimately n = 92 publications were filtered out for the full text analysis. An overview of the systematic review process including the results of each review phase shows Fig. 1.

3.1 Initial Results

Based on the result of the initial screening in the second phase of the review process (n = 766) different analyses were performed, which shows potential trends and focal points considering the topic usability and smart devices. Table 4 shows as an example the distribution of the evaluation basis in total but also with regard to different smart devices across the individual years of the review period. The first line of all studies evaluating base are to be considered. Overall it represents an expected increase in the number of publications from 2007 to today.

In the other lines of Table 4 the results are clustered per smart device technology. Thereby, the number of studies, which deal with smartphones, is still well above those, which are investigating tablet-PCs. Investigations in the field of smart

Table 4 Distribution of individual focused search results across the review period

Year	2007[a]	2008	2009	2010	2011	2012	2013	2014	2015
overall	40	50	63	62	101	77	109	134	130
smartphone	17	17	25	28	48	46	63	90	91
tablet-pc	6	4	6	3	10	15	31	26	28
smartwatch	0	0	0	0	0	0	0	0	2
smart glasses	1	2	2	2	2	0	1	2	4

[a]The first iPhone as supposedly first smart device sets the beginning of the review period

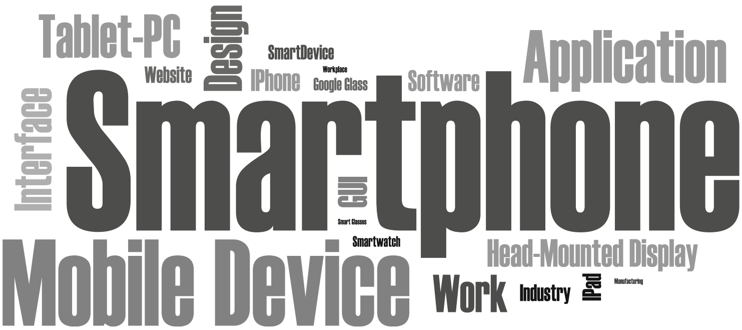


Fig. 2 Tag Cloud of different search terms on title level concerning the topic usability

watches, head-mounted displays or smart glasses are limited to a minimum. Prospectively there is a lot of action concerning these innovative technologies required.

The described results are represented as well in Fig. 2, which illustrates in the form of a Tag Cloud that the focus within the current publication and study situation concerning the usability of smart devices is lying on smartphones. The Tag Cloud displays the ratio of the amount of publications about an individual topic with the size of the respective imaged search term, which is contained in the publications title in combination with the topic usability in title, abstract or keywords.

It should be noted that the majority of studies are dealing with the evaluation of the usability of individual applications, without specifically addressing the effects, which result due to using it within a special context of use. Usability studies of smart devices and corresponding applications in the context of an insert in an industrial work environment can be only sporadically assumed on abstract level. The majority of usability studies are evaluating individual applications, using a high frequency of use as the sole criterion for the user-friendliness and thereupon conclude from this on the usability of the application (Hongu et al. 2015).

3.2 Results of the Full-Text Analysis

Within the amount of $n = 92$ publications for the full-text analysis a small number of studies (<25) were identified, which deal with the use of different (heuristic) methods or criteria for the evaluation of usability of mobile devices. An amount of n = 14 studies are using criteria, which are directly related to the methods described in Sect. 2.1. Exemplarily for that, the references Huang and Wang (2011), Inostroza et al. (2012) and Varsaluoma (2009) can be mentioned at this point. There can be found a high correlation between the criteria of usability for interactive systems in the field of conventional screen work (cf. Sect. 2.1) and the criteria applied for the evaluation of the usability regarding smart devices or the considered application in the studies. In particular the usability heuristics of Jakob Nielsen (1993) are mostly applied there to evaluate the usability of the specific application. In some cases there are differently named criteria, which are very similar to the common methods. Some references like Inostroza et al. (2012) or Gielkens and Wetzel (2012) developed their own methodology to evaluate mobile applications, but they uses the common criteria again. The same applies for some publications like Billi et al. (2010) or Conde et al. (2010), where other developed kits for usability evaluation are used, which are again built on Nielsen (1993). Because nearly every criterion in the investigated studies can be assigned to one of the heuristics of Nielsen, the following analysis refers to those. Based on the full-text analysis it was possible to filter out, how much and which criteria was used individually in the studies. For this reason it was worked out for every criteria, which reference is using it for evaluating the usability inside the specific application (cf. Table 5). Overall, it can be noted, that within the analysed studies all usability criteria of the common evaluation kits are referenced for the mobile usability evaluation by the full or nearly the full amount of studies. Furthermore, eight of ten criteria are referenced by an amount of at least thirteen studies. Only two criteria are referenced less, but still by an amount of at least ten studies.

The highest accordance get the usability criteria, that there has to be a good match between the system and the real world or rather a self-descriptiveness as well as an aesthetic, minimalist, easy design. These criteria are immediately followed by a group of six criteria, which are referenced by an amount of thirteen studies. There can be found related requirements of a continuous visibility of the system status, the user control and freedom as well as the flexibility and efficiency of use. Further on, there are two associated criteria with the need for a design, which promote the users recognition and helps to diagnose and recover from errors. Finally the requirement of help and documentation within the system got referenced on the same level. The criteria of error prevention by the system (n = 11 studies) and the consistency of the software design (n = 10 studies) are behind on a lower level of references.

4 Discussion and Outlook

Against the background of the ongoing digitalization of work and working environments, it seems conceivable, that the application of smart devices as a work assistance will increase. In this context, particularly it is important to consider and to ensure the usability of the system in the individual application before using it. Therefore, a specific significance will be attributed to the evaluation of the

Table 5 Allocation of studies to usability heuristics

Heuristics	Used by reference
Match between system and real world	Barricelli et al. (2013), Billi et al. (2010), Conde et al. (2010), Fetaji and Fetaji (2011), Gielkens and Wetzel (2012), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
Aesthetic and minimalist design	Barricelli et al. (2013), Billi et al. (2010), Conde et al. (2010), Fetaji and Fetaji (2011), Gielkens and Wetzel (2012), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
Visibility of system status	Barricelli et al. (2013), Billi et al. (2010), Conde et al. (2010), Fetaji and Fetaji (2011), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
User control and freedom	Barricelli et al. (2013), Billi et al. (2010), Conde et al. (2010), Fetaji and Fetaji (2011), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
Recognition rather than recall	Barricelli et al. (2013), Billi et al. (2010), Conde et al. (2010), Fetaji and Fetaji (2011), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
Flexibility and efficiency of use	Barricelli et al. (2013), Billi et al. (2010), Conde et al. (2010), Fetaji and Fetaji (2011), Huang and Wang (2011), Inostroza (et al. 2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
Help users recognize, diagnose and recover from errors	Barricelli et al. (2013), Conde et al. (2010), Fetaji and Fetaji (2011), Gielkens and Wetzel (2012), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
Help and documentation	Barricelli et al. (2013), Conde et al. (2010), Fetaji and Fetaji (2011), Gielkens and Wetzel (2012), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg

(continued)

Table 5 (continued)

Heuristics	Used by reference
	(2008), Soomro et al. (2013), Varsaluoma (2009), Wahab et al. (2010), Zaibon and Shiratuddin (2010)
Error prevention	Barricelli et al. (2013), Conde et al. (2010), Fetaji and Fetaji (2011), Gielkens and Wetzel (2012), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Varsaluoma (2009), Wahab et al. (2010)
Consistency and standards	Barricelli et al. (2013), Conde et al. (2010), Fetaji and Fetaji (2011), Huang and Wang (2011), Inostroza et al. (2012), Karahoca et al. (2010), Marcus et al. (2013), Schönfelder and Schmalstieg (2008), Varsaluoma (2009), Wahab et al. (2010)

usability of smart device applications. Consequently the described systematic review was implemented, which examines how to evaluate the usability of smart devices. The objective of the review was to investigate, if there are specific methods to evaluate the usability of smart devices or how far consisting methods including their criteria from the area of conventional screen work can be adopted.

Initially the review showed that the majority of publications are dealing with smartphones instead of other innovative technologies with high potential as work assistance. Furthermore, the amount of studies increased in the last years, but only a small number of studies describe a potential use of innovative technologies for work applications. An optimal task-technology-fit stays the basic prerequisite before using smart devices as work assistance (Terhoeven et al. 2015).

Concerning the usability evaluation of smart devices the review revealed that all in all the established methods and criteria of usability from the area of conventional screen work are used as well to design and evaluate the usability of mobile user interfaces. Recently developed mobile usability evaluation kits can be found in limited cases (Gielkens and Wetzel 2012), but are still built on the common criteria. Especially the heuristics of Jakob Nielsen (1993) were applied in the majority of the studies, but a content-related delimitation of the other established methods (DIN, Deutsches Institut für Normung 2008; Shneiderman and Plaisant 2005) is often not possible. In particular, the most referenced criteria for evaluating the usability of smart devices were the match between system and real world and an aesthetic, minimalist design. This becomes clear with regard to the small display size of smart devices. The information design for smart devices should be simple and quickly manageable based on the user's experiences. This criteria correlate with the following requirements of recognition or helping the user to recognize, diagnose and recover from errors. These criteria relieve the working memory and promote learning effects. The fact that the criteria of error prevention and consistency are referenced by a smaller amount of studies is a bit slightly unexpected. Consistency helps the user to recognize and control the system more easy and thus to relieve the working memory as well. The error prevention should be assured to protect the user from unexpected consequences. It is assumed that these criteria were included in the other considered criteria or prospectively provided with the development of the application.

With regards to the progressive digitalization of work, it becomes increasingly important to put smart devices, especially the usability of it, in the focus of the research and development. Actually the established methods from the area of conventional screen work for evaluating the usability are still adequate, but they have to be adjusted to the rapid development of new technologies. Different interaction concepts like touch screen or gesture control as well as distractions because of unsuitable information design are not respected in conventional criteria. Furthermore, there are not enough studies about a potential use of smart devices at work. Mostly the studies are looking at consumer or gaming applications.

In order to continue the research in the area of smart devices further studies with different issues will follow. Based on the results of this review, the BAuA addresses different concepts of information and interaction design for smart devices, which take into account the usability, an optimal technology tasks-fit, an optimization of cognitive load and other ergonomic requirements regarding the use as work assistance.

References

Arksey H, O'Malley L (2005) Scoping studies: towards a methodological framework. Int J Soc Res Methodol 8:19–32

Barricelli BR, Devis Y, Abdelnour-Nocera J, Wilson J, Moore J (2013) MANTRA: mobile anticoagulant therapy management. In: 7th International conference on pervasive computing technologies for healthcare, IEEE, Venice, pp 278–281

Billi M, Burzagli L, Catarci T, Santucci G, Bertini E, Gabbanini F, Palchetti E (2010) A unified methodology for the evaluation of accessibility and usability of mobile applications. Univers Access Inform Soc 9(4):337–356

Conde A, Lopez de Ipina K, Larranaga M, Elorriaga JA, Lopez JM, Irigoyen E, Garay-Vitoria N, Ezeiza A, Rubio J (2010) An intelligent tutoring system oriented to the integration of people with intellectual disabilities. In: Demazeau Y, Dignum F, Corchado JM, Bajo J, Corchuelo R, Corchado E, Fernandez-Riverola F, Julian VJ, Pawlewski P, Campbell A (eds) Trends in practical applications of agents and multiagent systems. Springer, Heidelberg, pp 639–647

DIN, Deutsches Institut für Normung (1998) Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten. Teil 11: Anforderungen an die Gebrauchstauglichkeit. EN ISO 9241-11

DIN, Deutsches Institut für Normung (2000) Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten. Teil 12: Informationsdarstellung. EN ISO 9241-12

DIN, Deutsches Institut für Normung (2008) Ergonomie der Mensch-System-Interaktion. Teil 110: Grundsätze der Dialoggestaltung. EN ISO 9241-110

Fetaji M, Fetaji B (2011) Comparing developed MLUAT (mobile learning usability attribute testing) methodology with qualitative user testing method and heuristics evaluation. In: Proceedings of the 12th international conference on computer systems and technologies, ACM, Vienna, pp 516–523

Gielkens C, Wetzel R (2012) A framework for usability evaluation of mobile mixed reality games. In: Herrlich M, Malaka R, Masuch M (eds) Proceedings of the 11th international conference on entertainment computing. Springer, Heidelberg, pp 401–404

Grauel BM, Terhoeven JN, Wischniewski S, Kluge A (2014) Erfassung akzeptanzrelevanter Merkmale von Datenbrillen mittels Repertory Grid Technik. Zeitschrift für Arbeitswissenschaft 68:250–256

Hongu N, Pope BT, Bilgic P, Orr BJ, Suzuki A, Kime AS, Merchant NC, Roe DJ (2015) Usability of a smartphone food picture app for assisting 24-hour dietary recall: a pilot study. Nutr Res Pract 9:207–212

Huang CH, Wang CM (2011) Usability analysis in gesture operation of interactive e-books on mobile devices. In: Marcus A (ed) Design, user experience and usability, Pt I, HCII 2011. Springer, Heidelberg, pp 573–582

Inostroza R, Rusu C, Roncagliolo S, Jimenez C, Rusu V (2012) Usability heuristics for touchscreen-based mobile devices. In: 9th international conference on information technology: new generations, IEEE, Las Vegas, pp 662–667

Ittermann P, Niehaus J, Hirsch-Kreinsen H (2015) Arbeiten in der Industrie 4.0. Trendbestimmungen und arbeitspolitische Handlungsfelder. Hans-Böckler-Stiftung, Düsseldorf

Karahoca A, Bayraktar E, Tatoglu E, Karahoca D (2010) Information system design for a hospital emergency department: a usability analysis of software prototypes. J Biomed Informat 43 (2):224–232

Marcus A, Abramowitz S, Abulkhair MF (2013) Heuristic evaluation of iCalamityGuide application. In: Marcus A (ed) Design, user experience, and usability: web, mobile, and product design, Proceedings of the Second International Conference, DUXU 2013. Springer, Heidelberg, pp 130–139

Mattioli S, Zanardi F, Baldasseroni A, Schaafsma F, Cooke RMT, Mancini G, Fierro M, Santangelo C, Farioli A, Fucksia S, Curti S, Violante FS, Verbeek J (2010) Search strings for study of putative occupational determinants of disease. Occup Environ Med 67:436–443

Nielsen J (1993) Usability engineering. Morgan Kaufmann, San Francisco, CA

Schönfelder R, Schmalstieg D (2008) Augmented reality for industrial building acceptance. In: Lin M, Steed A, Cruz-Neira C (eds) Proceedings of the IEEE virtual reality 2008. IEEE, Reno, pp 83–90

Schuh G, Potente T, Wesch-Potente C, Weber AR, Prote JP (2014) Collaboration mechanisms to increase productivity in the context of industrie 4.0. In: Windt K (ed) Procedia CIRP 19, 2nd CIRP robust manufacturing conference 2014. Elsevier Procedia, Amsterdam, pp 51–56

Shneiderman B, Plaisant C (2005) Designing the user interface: strategies for effective human computer interaction. Addison Wesley, Boston

Soomro S, Wan Ahmad WF, Sulaiman S (2013) Evaluation of mobile games using playability heuristics. In: Badioze Zaman HB, Robinson P, Olivier P, Shih TK, Velastin S (eds) Advances in visual informatics, Proceedings of the third international visual informatics conference, IVIC 2013. Springer, Heidelberg, pp 264–274

Spath D, Ganschar O, Gerlach S, Hämmerle M, Krause T, Schlund S (2013) Produktionsarbeit der Zukunft—Industrie 4.0. Fraunhofer Verlag. Stuttgart

Terhoeven J, Grauel B, Wille M, Wischniewski S (2015) Head Mounted Displays als Arbeitshilfen der Zukunft—Gestaltung eines beanspruchungsoptimalen Einsatzes. In: Schenk M (ed) Digitales Engineering zum Planen, Testen und Betreiben technischer Systeme, 18. IFF-Wissenschaftstage, 24–25. Juni 2015. Fraunhofer IFF. Magdeburg, pp 125–129

Varsaluoma J (2009) Scenarios in the heuristic evaluation of mobile devices: emphasizing the context of use. In: Kurosu M (ed) HCD 2009. Springer, Heidelberg, pp 332–341

Verordnung über Sicherheit und Gesundheitsschutz bei der Verwendung von Arbeitsmitteln (Betriebssicherheitsverordnung—BetrSichV), vom 03. Februar 2015

Wahab NA, Osman A, Ismail MH (2010) Engaging children to science subject: a heuristic evaluation of mobile learning prototype. In: Proceedings of the second IEEE international conference on computer engineering and applications, ICCEA 2010. IEEE, Bali Island, pp 513–516

Wickens CD, Hollands JG (2000) Engineering psychology and human performance. Prentice Hall, Upper Saddle River, NJ

Zaibon SB, Shiratuddin N (2010) Heuristics evaluation strategy for mobile game-based learning. In: Proceedings of the 6th IEEE international conference on wireless, mobile and ubiquitous technologies in education, WMUTE 2010, IEEE, Kaohsiung, pp 127–131

Menu Styles of Mobile Devices and Their Influence on Gaze Behavior While Walking

Jessica Conradi, Bjoern Nord, and Thomas Alexander

Abstract

Mobile IT-devices (Smartphones, Tablet-PCs, etc.) are often used while performing other tasks in parallel, e.g. while walking. However, mobile device and environment often compete for the users' attention. Binding too much attention on the mobile device will reduce attention on the environment. Especially in risky environments like road traffic, this might trigger substantial danger for users and third parties. Therefore, graphic user interfaces (GUIs) have to be adapted to it. Yet, lightweight mobile devices have small displays and only a limited number of objects can be displayed. Content with multiple subunits has to be arranged, e.g. by forming subcategories. Hierarchical structured menus facilitate this. In our survey, we compared the effect of different menu concepts on gaze behavior while walking. Menus containing 4–8 icons per level required the lowest number of gazes. In single interactions, the shortest visual distraction was found for the least number of objects on the screen.

Keywords

Mobile HCI • Mobile devices • Distracted walking • Menu styles • Gaze behavior

J. Conradi (✉) • B. Nord • T. Alexander
Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Bonn, Germany
e-mail: jessica.conradi@fkie.fraunhofer.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_20

1 Introduction

Our modern information society asks people to be permanently "on-line" and to have access to and to be in exchange with the World Wide Web. So, information technology has to be—and is—used constantly. Size and weight of modern mobile devices stimulate this.

However, using mobile IT-devices always and everywhere proved to have disadvantages. It provokes usage not only exclusively but also during performing additional activities. This takes place in different environmental contexts, e.g., while walking. In this setting, the user's attention is drawn to the device and therefore away from the environment. This leads to distracted walking. A survey founded by the American Academy of Orthopedic Surgeons found 78 % of the adults in the U.S. believing that this is a "serious" issue, 74 % say other people are almost always or sometimes distracted while walking. Just 29 % of the Americans admitted to do so themselves. Using a cell phone is a common means of distraction to walkers (e.g. 90 % talking on the phone; 85 % using a smartphone). Many pedestrians are aware of the danger of being distracted as they either witnessed a distracted walking incident (38 %) or have been involved in one themselves (26 %) (Ipsos 2015).

This puts pedestrians at risk of accident, injury and even death. In recent years, numbers of accidents due to smartphone usage have multiplied and are likely to continue rising (Nasar and Troyer 2013). Besides demanding attention, walking influences body motions. Different parts of the body move in different ways. Head as well as the device-holding hand move independently from each other, each step sends a small shockwave through the body (Pozzo et al. 1990; McDonald et al. 1997). The vestibular reflex of the visual system as well as the hand-arm-system compensates these effects only partly (Grossman et al. 1989). This hampers interaction with mobile devices regarding visual as well as motion issues. For example, visual acuity while walking is lower than while standing. This is due to effects of dynamic visual acuity caused by the relative motion of device and eye. A walking speed of 5 km/h causes a loss of about 20 % of visual acuity compared to standing. Signs like letters or icons have to be enlarged to compensate this loss (Conradi and Alexander 2014). Hitting a button while walking proved to be harder than while standing. For a low error rate buttons have to be larger in walking settings (Conradi et al. 2015). So, due to the walking setting, visual as well as movement issues require larger interaction objects. However, enlargement of relevant objects leads to a reduced number of displayed objects and items.

Therefore, walking calls for specially adapted graphic user interfaces, taking the limited display space into consideration. Modern mobile devices are powerful and provide a vast number of applications, which can handle a huge amount of information. Therefore, information objects cannot be presented at once but have to be structured in some way. This calls for interaction metaphors meeting the requirements, e.g., by arranging large numbers of items in hierarchical menus. They provide a given number of interaction options on each level (hierarchy breadth) resulting in a number of levels (hierarchy depth). Therefore, considering a given number of items, a large hierarchy breadth triggers a small hierarchy depth and vice versa. For example, a database of 64 items could be arranged in 2, 3, or 6 equally sized levels, resulting in level of different broadness, depicting 8, 4 or 2 options. The broadest hierarchy depicts eight options per level in two levels resulting in $8^2 = 64$ items. A medium-broad hierarchy provides four options per level, which results in three levels ($4^3 = 64$). A slimmer hierarchy of two options per level demands six levels ($2^6 = 64$) (Paap and Cooke 1997).

In a hierarchical menu structure comprising multiple levels the user has to learn to navigate between any given levels. The learning effort increases with the number of levels. Furthermore, each level calls for at least one interaction. So, broader hierarchies demand a low number of interactions to execute the item in question. However, a higher number of options per level increase the searching effort for the option in question. In mobile devices, display size limits the number of scrolling-free presentable options. The so-called "crowding" of objects on the display is a constraint imposed by the amount of available space. However, crowding hampers interaction and should be avoided. When this space is exceeded by the options, hierarchical depth has to be provided. Another point favoring deeper hierarchies is funneling. Funneling refers to a reduction in the total number of options processed by the user. This is achieved by designing a less broad but deeper menu structure. Then, an efficiency gain occurs for long processing times per option. This processing time includes the time demand for encoding options, comparing encoded representations with the target and deciding to terminate or to continue search. A simple example of the benefits of funneling is presented with a database of 64 items: in case of crowding all options on a single screen, an exhaustive search would lead to processing each of the 64 options. Yet, only two options are shown at once in a minimum-breadth hierarchical menu with a depth of six levels. In this case, the user has to search only two options per level in six levels; this results in processing only 12 options altogether. However, while the search demand decreases, the number of responses by the user (e.g., touch button) and the system (e.g., show next level) increase (Paap and Cooke 1997).

Menu hierarchies in desktop applications have been a research topic for years. A study focused on the optimal relation of breadth and depth for 64 items. Experimental task was finding a target word in the lowest hierarchy level. Level breadth was varied between 2 and 64 options. Experimental error rate as well as search time hinted at an optimal menu breadth of 4–8 items per level (Miller 1981). Other surveys showed broader hierarchies to be more effective as interaction time and accuracy decrease in deeper hierarchies and disorientation increases (Chae and Kim 2004; Jacko and Salvendy 1996).

Besides menu structure, there are other factors influencing the user performance. An influence of user expertise was shown. Novices worked faster with broad menu structures, but experts showed shorter time on task in slimmer, deeper structures. This was attributed to the experts' knowledge about the navigation to the target item (Dray et al. 1981). However, time pressure triggers different performance in different menu structures. Other authors compared broad and deep menu structures with novices. Time pressure slowed down time on task and increased error rate in all of the menu structures (Wallace et al. 1987).

In some ways, requirements of mobile devices correspond to those of stationary computers with larger displays. Crowding and funneling have to be considered, especially in respect to the limited space of the displays. While some authors recommend broader hierarchies (Barbaria et al. 2001), others found slimmer hierarchies to be beneficial for small displays (Geven et al. 2006). The hierarchical structure effects navigational behavior and users' perception and therefore should be adapted to the display size (Chae and Kim 2004). According to a meta-analysis, slim hierarchies should be used. These should utilize the whole screen space; scrolling is to be avoided. In designing, it is important to keep the application's objective in mind, which includes the usage of different kinds of menu structures for different applications (Kim et al. 2011). This demand can be applied to the setting of walking and interaction as secondary task.

Besides hierarchical menus, other kinds of menus are frequently used for smartphones, e.g. fisheye-menus or 3D-menus. For these menu types, the size of depicted options is variable. While selected options are bigger, other options are represented smaller. Changing the selected option also changes the size of the options. This type of menu is limited by the small screen space of mobile devices. Furthermore, it involves smaller buttons and often scrolling. As scrolling as well as small buttons are to be avoided in a walking setting, these types of menu structures were not considered in this study.

However, other menu structures can be used depending on the kind of database. In graphical composition tasks, graphical editors can be used. They consist of a set of buttons, which allow for changing special attributes of a graphical element. The current status of the element is depicted continually. In case of a limited number of attributes of the graphic element, all options can be depicted in one screen. Dependent on the starting point of the displayed graphic element and on the target element, different numbers of user responses are needed. However, this kind of editor demands high mental effort in regard to changing the graphical element. Compared to a hierarchical menu, which leads the user through the choosing process, it results in intricate interaction.

Touch displays trigger movements of the finger und subsequently of the arm to reach the button in question. Fitts' Law addresses this kind of movements (Fitts 1954). It describes the relation of the distance to the target button, the size of the button and time on task. Influence of the distance is obvious. Target size affects the precision of the movement. Movements that are more precise demand re-adjustments which results in longer time on task. According to Fitts' Law, distance and target size determine the index of difficulty of a particular target. For low indices of difficulty, a fast, solely ballistic movement occurs. Higher indices of difficulty induce corrections of the movement, and therefore result in longer lasting movements. We also proved the applicability of Fitts' Law for walking settings (Alexander et al. 2008).

To decide upon an action and to direct the movement, the user requires visual feedback. Therefore, he or she has to gaze upon the smartphone. For the time of this gaze, the visual focus is on the display and for this time at least, the user is distracted from the environment. The longer this distraction lasts, and the longer the surrounding scenario evolves without the users' attention, the higher the probability, that the user misses relevant developments. Therefore, errors are likely to occur and result in dangerous situations.

Gaze duration can be described in different ways. One is the duration of a single gaze on the device. This may be influenced by the complexity of the GUI. A simple GUI, presenting few objects, might need only short gazes to take in the meaning. So, the visual focus has to be distracted only for a short moment. GUIs with more objects might trigger longer visual distractions. Besides this, the number of glances on the smartphone is relevant, because it provides information about how often the user is distracted. Furthermore, the duration of fulfilling a particular action, a single interaction in a particular level, is of interest. It is determined by the glance duration as well as by the number of glances. This may give hints for basic interaction recommendations. Another important measure for distraction is the overall time for fulfilling a desired action, e.g., selecting an item of a database. This overall duration time is influenced by the single interactions as well as by the number of screens. In hierarchical menus, the hierarchy depth or the number of levels determines the number of screens, and the breadth of the hierarchy determines the complexity of one single screen.

In previous studies, walking has been addressed scarcely. Therefore, in this study we focus on the influence of menu structures in hierarchic menus and a graphical editor, respectively, on the gazing behavior in the special setting of walking in an attention-demanding surrounding.

2 Method

We assumed that the menu structure influences the gazing behavior in smartphone usage, especially while walking. Therefore, we carried out a study with 16 male volunteers aged 27 ± 6.7 years (MW ± SD). A design with repeated measures for all conditions was chosen, conditions were permutated according to Latin square (Tabachnick and Fidell 2007).

Participant's task included the interaction with a smartphone while walking on a treadmill. In the main task, a set of 64 different symbols was used. These symbols differed in fringe, form, color and filling-element and were easily distinguishable. They were selected to be items of 2^6, 4^3 and 8^2 hierarchical menus as well as for a symbol editor. As soon as a symbol randomly showed up in the environment, interaction process started. The symbol was visible during the whole process until the participant quitted the interaction. The first 32 symbols were used for practice; data of the second set of 32 symbols was used in the analysis. Time to fulfill the task in one condition was about 15 min.

A projection-wall (2.3×3.6 m) with a virtual environment was used (see Fig. 1). The participants walked through a rural virtual landscape. The landscape was realized by means of the gaming-engine CryEngine 2 (Crytec®, Frankfurt a.M.,

Fig. 1 Test environment including large-scale projection wall and treadmill

Germany). This environment allowed for administering an attention-demanding secondary task. Human-shaped distractors appeared randomly to draw the participant's attention to the environment. Participants had to observe the environment and react to the distractors.

As device for facilitating almost natural walking we used a treadmill. During the experiment, the treadmill determined participant's walking speed. The treadmill was an H/P/Cosmos pulsar (h/p/cosmos sports & medical GmbH, Nussdorf, Germany).

The mobile interaction device was a Samsung Galaxy S2 (I9100, see Fig. 2, left). A head-mounted eye-tracking system (Dikablis®, Ergoneers GmbH, Germany), was used to estimate the number and duration of glances directed at the smartphone (see Fig. 2, right). The eye-tracking system consisted of two cameras, first of them aiming at the participant's eye. Based on this, the position of the pupil was calculated. The second camera aimed forward and recorded the visual field of the participant. The environment was fitted with markers. These markers allowed for the calculation of the position of the smartphone at any moment. The footage of the two cameras was matched and the analyzation software calculated the duration and the number of glances directed at the smartphone.

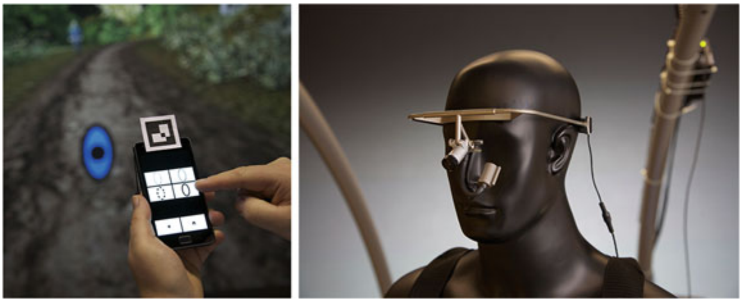


Fig. 2 Interaction device (*left*) and eye-tracking system (*right*)

The independent variables of the experiment were treadmill walking speed and menu type. Walking speed was 2.5 and 5 km/h, respectively. The second independent variable, menu type, consisted of four levels, i.e. three hierarchical menus types and an intricate symbol editor. The realization of the GUIs is depicted in Fig. 3.

Walking induces a relative motion of the hand and therefore hampers interaction. To exclude any influence of this on the interaction, button size was chosen deliberately to meet the conditions of Fitts' law for ballistic movements (Fitts 1954; Alexander et al. 2008). Therefore, display height allowed a maximum of five buttons and display width allowed two buttons. In consequence, a total of ten buttons was depicted. Two of them were necessary for the "back" and "home" buttons. Therefore, a maximum number of eight options per level were applicable. The resulting height of the buttons was 15 mm. This size was used uniformly for all buttons, regardless of the number of buttons used in a particular GUI.

The independent variable "menu type" was varied on four factor levels, three hierarchical levels and a graphical symbol editor. The hierarchical menus differed in depth and breadth. The slimmest hierarchy consisted of two options per level (see Fig. 4 left) which required 6 levels (2^6) to lead to 64 items. The medium hierarchy level is depicted in Fig. 4 (second from left), it shows four options on one level in 3 levels (4^3). The broadest hierarchy is shown in the Fig. 4 (second from right), it has 8 items per level (8^2). The right-hand picture in Fig. 4 shows the symbol editor (SE). By touching one of the six buttons, the attributes of the resulting symbol change accordingly. The currently composed symbol is shown continuously at the bottom on the screen. The minimum number of interactions needed to compose a particular symbol was determined by the pre-set symbol and the target symbol. In case of high similarity, only few adjustments are necessary. Maximum number of demanded adjustments was 6, medium adjustment number was 3. In all conditions, it was possible to correct input by either using the "back"-button on the bottom (hierarchical menus) or by a second interaction with the accidentally used button (SE).

Dependent variables of the experiment were derived from the eye-tracking system. The system provided automated, marker based analysis of the raw data

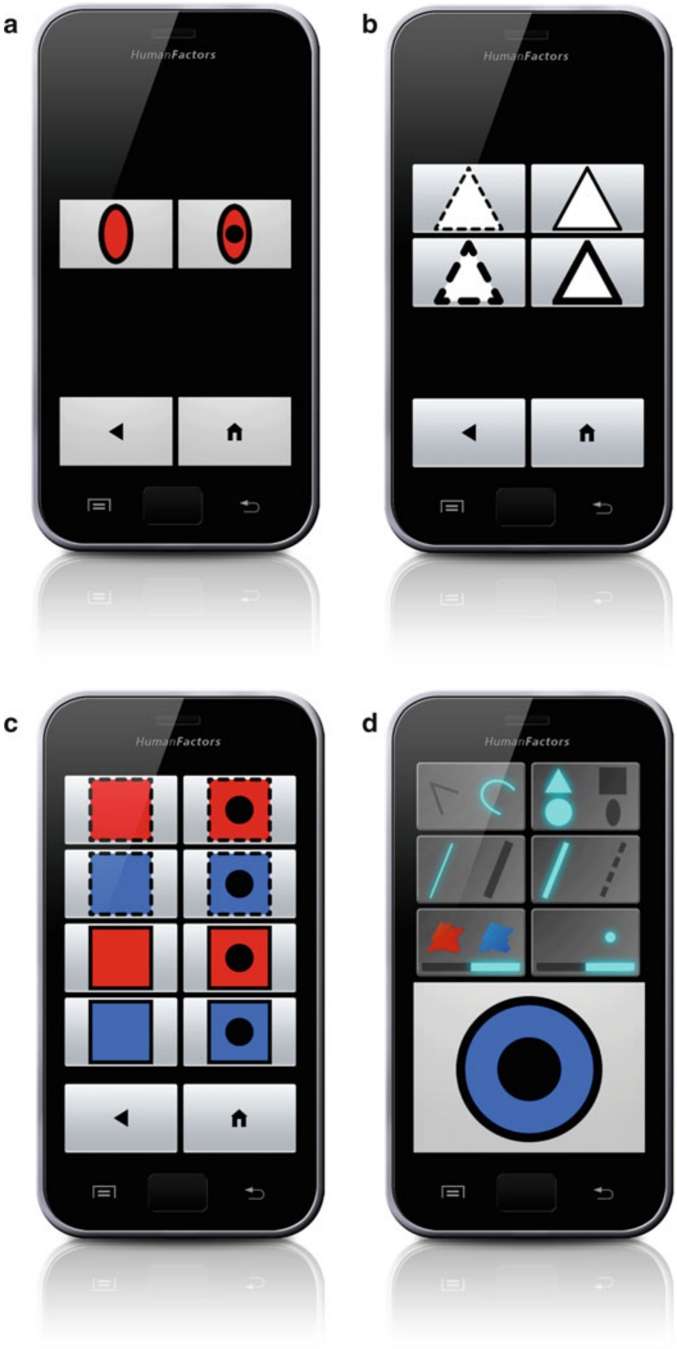


Fig. 3 Independent variable menu type and its realization on a smartphone

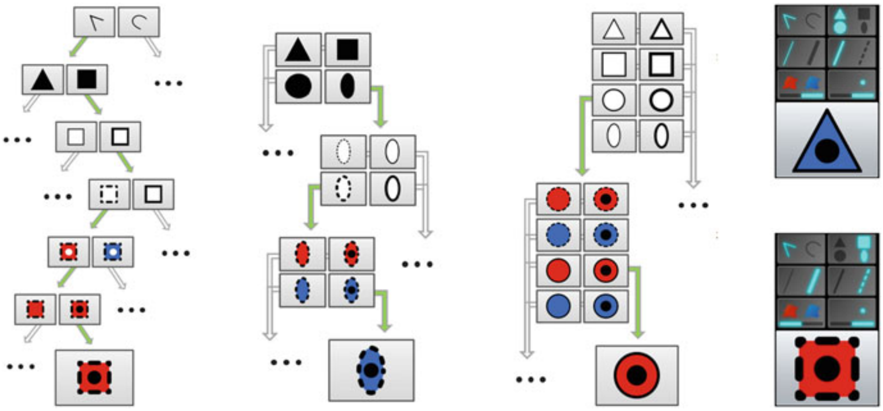


Fig. 4 Examples for operation paths of the menu types: 2^6: 6 single interactions (*left*); 4^3: 3 single interactions (second from *left*); 8^2: 2 interactions (second from *right*); SE: 0–6 interactions (*right*)

and provided number of glances and mean glance duration. Based on these, glance time per single interaction was calculated. Only hierarchical menus were included in this part of analysis, as each input triggered the change of hierarchy level but not in the single-screen symbol editor. Finally, overall glance time per interaction cycle was computed, including all four menu types. An interaction cycle covers the entire time to register a symbol, a variable number of single interactions and finally the confirmation of the selected symbol. It also includes mistakes and corrections. To measure this time on cycle, we used the percentage of smartphone glance time compared to entire time for finishing the task.

Besides glance behavior, performance in main task and secondary task, as well as workload was registered and analyzed (Conradi et al. 2016).

A two-factored multivariate analysis of variance was used. In advance, all data were checked for normality and sphericity (Mauchly-Test). In case of significant results, pairwise comparisons using the Bonferroni-correction were administered. Significance level was 5 %. We used SPSS 20 for analysis.

3 Result

Number of glances on the smartphone is depicted in Fig. 5. They were found to be influenced by the menu type ($p < 0.01$, $F(3,45) = 11.49$) with an effect size (partial eta squared) of $\eta 2 = 0.434$. The post-test showed significant differences for the menu type combinations 2^6 and 4^3, as well as for 2^6 and 8^2 (each $p < 0.01$). A statistical trend was found for the combination 4^3 and SE. Other combinations showed no significant difference. Estimated marginal means for the factor levels were M(2^6) = 134.8; M(4^3) = 90.9; M(8^2) = 82.8 und M(SE) = 113.4. The menu types 4^3 and 8^2 demanded a lower number of glances on the smartphone

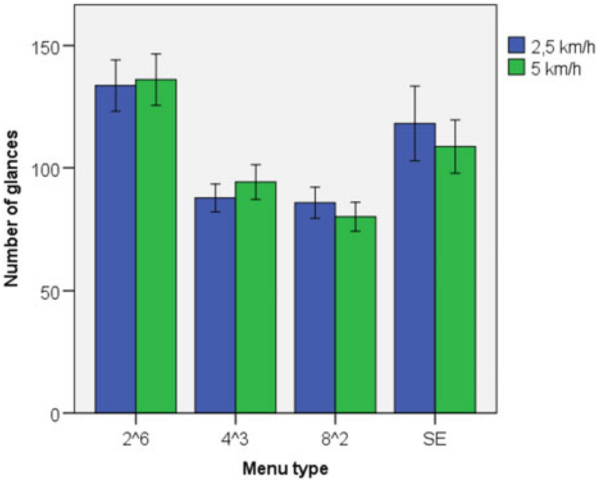


Fig. 5 Mean and ±1 standard error of number of glances

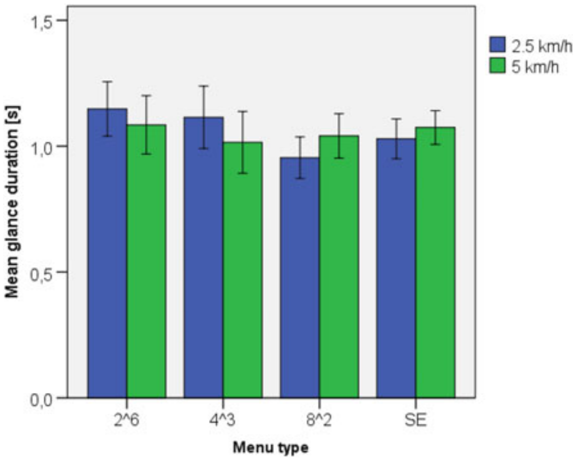


Fig. 6 Mean and ±1 standard error of mean glance duration

than in the menu type 2^6. The statistical trend hints to a difference between 4^3 and SE, with fewer glances in the 4^3 condition.

Walking at 2.5 and 5 km/h respectively, showed no significant influence on the number of glances ($p = 0.696$; F(1,15) = 0.159, M(2.5) = 106.30 und M(5) = 104.73).

Mean glance duration on the smartphone showed no significant influences, neither for the menu type ($p = 0.629$; F(3,45) = 0.583, M(2^6) = 1.12 s; M (4^3) = 1.065 s; M(8^2) = 0.997 s; M(SE) = 1.052 s) nor for walking speed ($p = 0.814$; F(1,15) = 0.058; M(2.5) = 1.061 s; M(5) = 1.054 s). A graphical representation of this data is given in Fig. 6.

Glance time per single interaction or screen is depicted in Fig. 7. Only the three hierarchical menu types were used in the analysis, as the method does not apply to

Fig. 7 Mean and ±1 standard error of glance time per single interaction

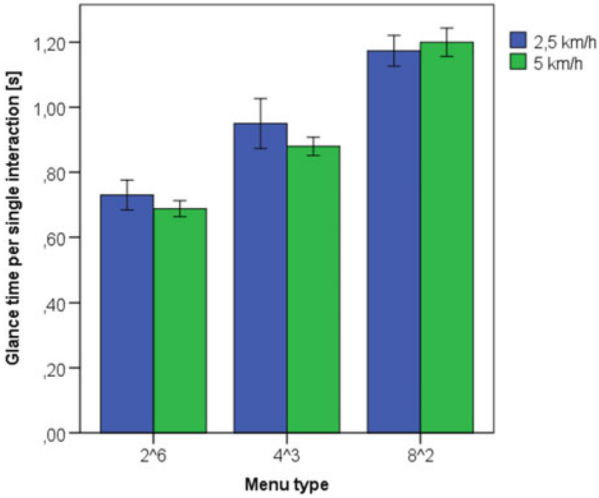
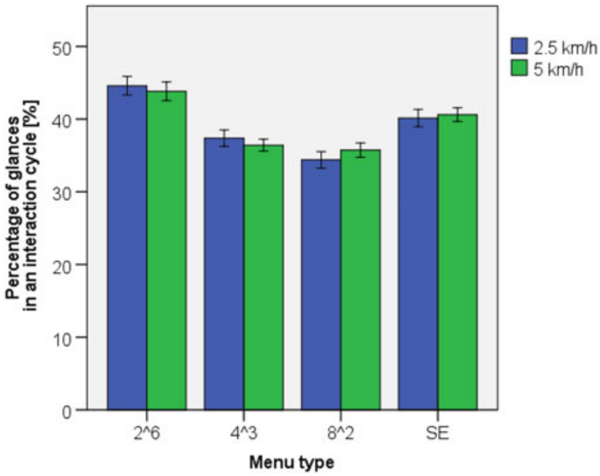


Fig. 8 Mean and ±1 standard error of glance time percentage in all interaction cycles



the symbol editor. Menu type had a highly significant influence on it ($p < 0.01$, F(2,30) = 68.1, η^2 = 0.819). Post-tests showed high significance for all combinations, with $p < 0.01$ for each. Marginal means were M(2^6) = 0.71 s; M(4^3) = 0.92 s and M(8^2) = 1.19 s. So, the glance time per interaction rose with the number of buttons. Compared to 2^6, in 4^3 the time on task increased by 30 % and in 8^2 by 68 %.

Walking speed showed no significant influence on glance time per single interaction ($p = 0.268$; F(1,15) = 1.223, $M_{2,5} = 0.95$ s und $M_5 = 0.92$ s).

The percentage of smartphone glance time compared to the entire time for finishing the task was analyzed to estimate the glance time consumption of all interaction cycles (see Fig. 8). They showed a highly significant influence for menu type ($p < 0.01$, $F(3,45) = 20.34$, $\eta^2 = 0.576$). Posttests proved highly significant differences for the combinations 2^6 and 4^3, 2^6 and 8^2, as well as for 8^2 and SE (each $p < 0.01$). Marginal Means were M(2^6) = 44.2 %; M(4^3) = 36.9 %; M (8^2) = 35.1 %; M(SE) = 40.4 %.

For walking speed on the percentage of glances on the smartphone no significant influence was found ($p = 0.969$; $F(1,15) = 0.02$, M(2.5) = 39.12 und M(5) = 39.15).

4 Discussion

We conducted a study concerning menu designs for smartphones used while walking. The results indicate an influence of menu type (hierarchical menus and graphical editor) on gazing behavior.

The focus of our study was on menu hierarchies based on the same overall set of items. The menu designs differed in number of options per level, resulting in different numbers of levels and therefore different number of single interactions. Additional, a non-hierarchic menu type was included, i.e., an intricate graphic symbol editor.

To evaluate the menu types, we considered an entire interaction cycle. It consisted of registering a symbol and selecting a target item from a sized item pool, including several single interactions. We found deeper hierarchies with a small number of options to lead to a high number of glances at the smartphone as well as to a high overall interaction time. This was true for hierarchies containing two options per level. We expected this outcome as the slim hierarchy demands a high number of levels and therefore a high number of single interaction steps.

For hierarchies with four and eight options per level we found no differences for number of glances as well as for overall interaction. We expected a difference between those two conditions just as between the 2-option per level hierarchy and the broader hierarchies. However, the higher number of options and the resulting heightened searching time in the broader menu type seems to have used up the advantage of the low number of single interactions in the broadest hierarchy. Similar results for desktop applications were found by other authors (Chae and Kim 2004; Jacko and Salvendy 1996).

The graphical symbol editor provided an interaction process, which demanded three interactions per cycle on average. This equals the interaction demand of a 4-options-per-level hierarchy. However, a statistical trend indicated the number of glances to be higher in the symbol editor than in the medium hierarchy level. So, the editor had no advantages over the other levels. Furthermore, the graphic editor has shown to trigger a high number of interaction errors (Conradi et al. 2016). Therefore, in the walking setting, this menu type was found to be disadvantageous.

Menu types did not influence single gaze duration; more complex GUI screens did not trigger longer glances. So, the heightened effort to figure out the meaning of a more complex screen was compensated by an increased number of similarly long gazes.

Walking condition showed no influence on gazing behavior. However, only slow (2.5 km/h) and fast (5 km/h) walking were covered in the study. The vestibular reflex of the visual system as well as the hand-arm-system seems to have compensated the walking-induced body movements in slow and fast walking in a similar way. However, standing and very fast walking or running should be considered in further studies.

Additionally, we looked into single interactions involving only a single level of a menu hierarchy. Glance time for the interaction in this case was the lowest for the interface offering the least number of options. Doubling the number of options from two to four resulted in an increased time on task (130 %). Increasing the number of options to 8, the interaction time rose to 168 % compared to the two-option condition. This clearly shows the relevance of the number of options in selection screens for application design in attention-demanding environments. It matches other results also recommending slim hierarchies (Geven et al. 2006). While necessary information has to be given, all kinds of less important or redundant information or interaction options should be questioned. Furthermore, crowding should be avoided. This result is in accordance with findings deriving from desktop computers (Paap and Cooke 1997). So, especially for applications designed for walking settings, all depicted information and options have to be scrutinized thoroughly and included only in case of high relevance. This could result in smaller sets of items and therefore facilitate interaction.

To sum up, in a given set of items, we found a medium broad menu hierarchy of 4–8 options per level to be the most advantageous. An intricate graphical editor was less applicable. However, the most beneficial way of improving interaction for a walking setting, is to reduce interaction options to the relevant minimum.

References

Alexander T, Schlick C, Sievert A, Leyk D (2008) Assessing human mobile computing performance by Fitts' law. In: Lumsden J (ed) Handbook of research on user interface design and evaluation for mobile technology. IGI Global, Hershey, pp 830–846

Barbaria K, Giacoppo S, Kutur U (2001) Single item search and selection in handheld devices: a pilot study on the effects of font size & menu design. www.cs.umd.edu/class/fall2001/cmsc838s/rust.doc. Accessed 7 Mar 2016

Chae M, Kim J (2004) Do size and structure matter to mobile users?: an empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones. Behav Inform Technol 23(3):165–181. doi:10.1080/01449290410001669923

Conradi J, Alexander T (2014) Analysis of visual performance during the use of mobile devices while walking. In: Hutchison D, Kanade T, Kittler J et al (eds) Engineering psychology and cognitive ergonomics, vol 8532. Springer International, Cham, pp 133–142

Conradi J, Busch O, Alexander T (2015) Optimal touch button size for the use of mobile devices while walking. Proc Manuf 3:387–394. doi:10.1016/j.promfg.2015.07.182

Conradi J, Nord B, Alexander T (2016) Interaction dialog design for the use of mobile devices while walking. In: Schlick C, Trzcieliński S (eds) Advances in ergonomics of manufacturing: managing the enterprise of the future. Springer, Heidelberg. doi:10.1007/978-3-319-41697-7

Dray S, Ogden W, Vestwig R (1981) Measuring performance with a menu-selection human-computer interface. In: Proceedings of the 25th annual meeting of the Human Factors Society, pp 746–748

Fitts PM (1954) The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 47(6):381–391

Geven A, Sefelin R, Tscheligi M (2006) Depth and breadth away from the desktop. In: Nieminen M, Röykkee M (eds) MobileHCI '06 proceedings of the 8th conference on human-computer interaction with mobile devices and services, p 157

Grossman GE, Leigh RJ, Bruce EN, Huebner WP, Lanska DJ (1989) Performance of the human vestibuloocular reflex during locomotion. J Neurophysiol 62(1):264–272

Ipsos (2015) Distracted walking study: topline summary findings. http://www.anationinmotion.org/wp-content/uploads/2015/12/AAOS-Distracted-Walking-Topline-11-30-15.pdf. Accessed 7 Mar 2016

Jacko J, Salvendy G (1996) Hierarchical menu design: breadth, depth, and task complexity. Percept Motor Skills 82:1187–1201

Kim K, Jacko J, Salvendy G (2011) Menu design for computers and cell phones: review and reappraisal. Int J Hum Comput Interact 27(4):383–404

McDonald PV, Bloomberg JJ, Layne CS (1997) A review of adaptive change in musculoskeletal impedance during space flight and associated implications for postflight head movement control. J Vestib Res Equilib Orient 7(2–3):239–250

Miller DP (1981) The depth/breadth tradeoff in hierarchical computer menus. In: Proceedings of the Human Factors Society 25th annual meeting, pp 296–300

Nasar JL, Troyer D (2013) Pedestrian injuries due to mobile phone use in public places. Accid Anal Prev 57:91–95. doi:10.1016/j.aap.2013.03.021

Paap KR, Cooke N (1997) Design of menus. In: Helander MG, Landauer TK, Prabhu PV (eds) Handbook of human–computer interaction, 2nd edn. Elsevier, Amsterdam

Pozzo T, Berthoz A, Lefort L (1990) Head stabilization during various locomotor tasks in humans. Exp Brain Res 82(1):97–106

Tabachnick BG, Fidell LS (2007) Using multivariate statistics, 5th edn. Pearson/Allyn & Bacon, Boston

Wallace D, Anderson N, Shneiderman B (1987) Time stress effects on two menu selection systems. In: Proceedings of the 31st annual meeting of Human Factors Society, pp 727–732

An Age-Differentiated Perspective on Visualizations of Personal Health Data

Sabine Theis, Peter Rasche, Alexander Mertens, and Christopher Marc Schlick

Abstract

The digitalization of medical human-machine system equips data with an essential role during processes supported by digital information systems. Here, data and information visualizations are able to reduce complexity only if influencing variables on human understanding, insight, and decision-making can be controlled, quantified and ergonomically designed. Present chapter provides an review on previous work in computer sciences, engineering, psychology and medicine in order to descriptively summarize human aspects which are relevant for the design of data and information visualizations in healthcare settings. The second part of this chapter builds upon the outcome of this review by working out current challenges of information and data visualization for consumer healthcare systems and introducing three studies which serves to tackle those challenges.

Keywords

Visualization • Data validation • Health care • Human factors • Research agenda • Literature study

1 Introduction

Health is a fundamental social value and is a valuable commodity. From an economical point of view, the benefit of *healthcare* is comprised of the prevented health-dependent failures and reduced disease spending. The healthcare industry in

S. Theis (✉) • P. Rasche • A. Mertens • C.M. Schlick
Institute of Industrial Engineering and Ergonomics, Ergonomics and Human-Machine-Systems, Bergdriesch 27, 52062 Aachen, Germany
e-mail: s.theis@iaw.rwth-aachen.de; p.rasche@iaw.rwth-aachen.de; a.mertens@iaw.rwth-aachen.de; c.schlick@iaw.rwth-aachen.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_21

Germany is growing faster than the overall economy and has generated eleven percent of the German gross value in 2015. In this regard, demographic change constitutes vital challenge for the healthcare sector. With progressive aging of western population, the number of those who need medical care as well as the morbidity of individuals increases. Age patterns, but also the quantitative ratio between men and women, the proportion of a countries residents, foreigners and naturalized citizens are expected to change.

At the same time the medical domain evolves from a reactive to a predictive and preventive one, where an early intervention strives to prevent or at least detect illnesses early, as to start treatment earliest and as individualized as possible. Changing the medical approaches and processes from general to individual and stratified is expected to improve well-being during course of life and especially in old age. Side effects of individualized medicine include less health expenditures and constant quality of care compensating the implications of *demographic change*.

The collection of large, heterogeneous personal data is referred to as one way to implement individualized and preventive medical processes. Data required for personal medicine include clinical data, census data, epidemiological data and data from imaging methods, but on the other hand miniaturization of technology and sensors also foster collection of personal data in home settings. The applications of digital medical sensors in patients and elderlies homes range from smart-homes to active implants to intelligent prostheses to small wrist gauges activity tracking devices measuring steps and blood pressure to glucose meters smartphone extensions.

Analyzing personal health-related data by data-mining and deep learning technologies generates new knowledge for pathogenesis, prevention and individualized diagnosis, prevention, mentoring and monitoring. Besides tailored medical treatment and efficient resource usage, individual medicine also opens up new perspectives for individuals. Especially, data collected and displayed on smartphone and wearables-applications provide consumers with an overview about their own nutrition and exercising behavior or their mental capability. A greater leeway in decision-making reveals to individuals regarding a detection of health risks, selection of preventive measures and therapy method. But safe, efficient and effective transmission of data characteristics only accrues by understandable, easy-to use data representations. If data and information is represented by graphic features, data characteristics and relationships can be understood more easily by a larger target group. Attaching graphical features to data, -types,-characteristics and -transitions puts the highest bandwidth channel to mediate between data and human. The original definition of visualization implies a “formation of mental visual images”, a second definition refers to “putting something into visual form”. Though, by externalizing the mental visualization it becomes a cognitive tool whose virtue is not automatically given by the fact that it is a visible. Instead different factors have to be taken into account during the creation or selection of visualization. Contextual factors such as optics, lightning, display size and resolution, lightness, brightness, contrast, but also more important human aspects including perceptual, attentional and cognitive processes but also and equally important the users tasks, goals corresponding with individual information need and behavior.

2 Digital Consumer Health Products and Visualizations

In order to get a first impression of existing personal digital health products which come together with data or information visualizations, preceding subchapter describes personal health applications and individual perspectives on visualization. Here an emphasize is put on mobile monitoring of private health-related aspects and on medical test results and personal health records which make frequent use of *data visualization* and *information visualization*. In some cases the borders between these two areas already overlap. This could be interpreted as hint to future development of both application areas towards cooperative care systems where private and professional medical data is merged together in one database and individually shown to the user depending on his/her role and tasks.

In order to establish a narrative overview, several research papers and reviews dating from 1980 to 2015 were considered using APA PsycNET (psychology) and ACM Digital Library (computer science). For the search in all libraries we used the following search terms: ((((“visualization technique”) OR “information visualization”) OR “data visualization”) AND “health”).

2.1 Mobile Personal Health Monitoring

Medication adherence is one of the most important tasks for consumer or personal health with incremental amount of digital information systems support. Pillboxes and bottles which automatically connect and send data in form of messages to patients or caregivers phones (Theng et al. 2013). Other than this pure medication reminder another application includes visualizations of diabetes patient’s continuous and discrete data plotted for a mobile health feedback system with the purpose to monitor medication ingestion events and physiological measures simultaneously with sensors, adhesive patches and a smartphone (Browne et al. 2015). Visualizations of this app included medication ingestion and activity, heart rate, sleep, rest, and glucose data over varying time intervals and supported insights into individual longitudinal patterns of medication adherence and self-management in the natural setting. The visualizations were accessible to healthcare professionals as well as patient consumers by preserving accuracy of the data and being understandable at the same time. Time-dependent line graphs were able to augment communication (interviews) between patients and medical professionals. Frost and Smith (2002) developed a system which users could apply to keep track of parameters important for a treatment of their diabetes. Data formats included behavioural and sensor data. Besides medication adherence, wellness or fitness applications are a domain in which lay users and consumers are provided with behavioral, physiological or affective information which poses certain relevance for their personal health. The quantified-self (QS) movement (Swan 2012) uses on-body sensors which provide input for traditional bar-chart und line graph visualization primarily displayed on web applications and mobile phones. By nature, people have a vague sense of their activities and often find it difficult to keep track of their progress over time. Furthermore of comparing themselves to others, a social element that helps many people sustains motivation. Ali-Hasan et al. (2006) found for example that visualizations showing pedometer data of activity encouraged students to set personal goals, activity planning, analyzing. Furthermore the insight into own data lead to an increase of motivation through virtual competitions and teamwork. The same effect was observed for a mobile system using photos and manual entries instead of sensor data. Users could take photos of their food, attach a category to it and manually enter their exercise behaviour. Photos and entries were displayed on a horizontal timeline. The authors found that it raised the awareness of health concerns and motivated users to engage a healthier behavior (Brown et al. 2006). Additionally, Goodman and Foucault found within their literature review that visualizing weight loss and fitness activities creates structure, accountability and motivates the users (Goodman and Foucault 2006). Subramonyam (2015) implemented an innovative solution to display data on a mobile phone for quantified self-monitoring data. Here, the so called "MagicMirror" maintains conjunction between data and its context by visualizing health data with the body as reference frame. The systems detect gestures and projects heart-rate, calories, steps and sleep data onto a virtual image of the user. By touching ones heart e.g., the mirror shows a numerical representation of the current heart rate value. Instead Epstein et al. (2014) analyzed the specific tasks of individuals using self-tracking devices and which data visualizations they appreciate. In total 14 participants regularly wearing self-tracking devices were interviewed about their tasks and goals. Results suggest that participant's main goal is to get an overview on their physical activity and an insight into their workout behavior serving the long-term goal of maintaining or increasing the personal level of physical activity. Subjective interviews with all participants revealed that visualization preferences are closely related with individual goals. All participants preferred to view workout-patterns, eleven participants mentioned their work schedules influenced their activity. Sensor data are the basis for many wellness visualizations and air quality plays an important role for personal health and wellbeing (Jones 1999) a lot of wellness devices monitor and visualize corresponding parameters. Within the "inAir" application as within most wellness applications time-based line graphs where the vertical axis being air quality level and the horizontal axis being time. A line graph representing particle counts over four hours within the main part of the screen. There are two modes of data visualization: a single-user and sharing mode. In single-user mode, the graph displays only its own data, and in sharing mode a user can choose to display either a single line from the local sensor or multiple lines in different colors from all participants. Additionally, the display renders a grey area in sharing mode, which represents the range of data from all participants. The visualization of air quality helped to improve air quality by fostering the behavior and awareness of the users regarding the measured (Kim et al. 2013). The psychological needs of users with chronic pain during fitness had concerns regarding visual appearances of fit instructors within exercise videos and preferred an abstract or realistic representation of the instructor (Swann-Sternberg et al. 2012). An interesting approach to investigate visual preferences of users was conducted by Jang et al. (2013) who focused on investigating the visual communication of pain. Eight Individuals drew pen-and-paper diagrams to communicate about their pain. Four subjects used color to describing surface-level symptoms and they tried to match the color to the symptom. Symptoms below the surface were never described with colors. Marks like crosses, dots and lines were used to indicate the location of pain while shape outlines and shading were used to refer to an area. Users preferred the body diagram and drawing as easier, accurate and more specific than textual representations. The happiness counter is on example for a digital health application and encouraged users to be happy by monitoring time based smiling and visualized in with the help of a calendar having one smiling icon each day the user smiled (Tsujita and Rekimoto 2011) . An affective health system for coping with stress and stressful situations was built by Vaara et al. (2009) logging data about a person's arousal (galvanic skin response), pulse (heart rate), movement (accelerometer) and contextual data from the device itself like photos, text messages, but also surrounding devices Bluetooth signal. All data were visualized for the user to detect patterns. The iteratively built visualization contained geometrical shapes as representations for places. Positions and activities were ordered linearly according to time on a horizontal axis and with a circular representation of time. Besides a participatory design the authors reported no evaluation.

Preceding summary illustrates the increasing amount of application domains where digital consumer health applications communicate aggregate health-related data to people interested in factors influencing their health by means of visualizations. The major part of mentioned applications relates to a less private setting where data are interpreted without professional expertise or additional clinical data sources.

2.2 Medical Test Results and Personal Health Records

Personal also professional medical records become increasingly intervened with personal data and are more and more being made available to patients. Non-experts and consumers primarily employ computers and smart phones to access their *personal health records* while initially personal health records (PHRs) include professional medical information. Personal health records contain personal health data administered by an individual often referred to as patient. PHR support people in medical care, but they also encourage healthy individuals to maintain their health, improve their well-being and prevent illnesses (Tang et al. 2006). This implies a promotion of autonomous and impartial handling of individual's own health. Just as electronic health records (EHR), PHRs refer to digital instance of patient data. Patients are able to maintain, aggregate, and analyze their own data which then can support caregivers and health care providers. PHR differ among providers and allow patients to directly enter information on personal diagnoses, medications, laboratory tests, diagnostic studies, and immunizations into an information system. A review on PHR systems literature (n = 130) by Archer et al. (2011) stated that due to the accessibility of health information on the internet, individuals increasingly incorporate online information to make health-related decisions for themselves and others. As care givers and physicians play a pivotal role in people's personal health life the authors identify a certain overlap and expect PHRs and EHRs to be brought together at some point within the digitization of healthcare services. The kind of health-related information which needs to be included into PHRs remains contentious. Showing individual's data which a healthcare provider keeps about them might be as reasonable and useful as providing data and information an individual's keeps for personal use to clinicians. Information collection, information exchange/sharing, and information self-management count among important PHR systems activities. Relevant data include demographic information, lists of health problems, health procedures, illnesses, allergies, home monitoring data, family (health) and social history, lifestyle/wellness data like data generated by fitness trackers, health recommendations, immunizations, examinations, medications, laboratory test results and appointments as well as physicians and personal notes. Mentioned data can be provided by the patient his/herself, the caregiver, physicians, clinical EHR systems or insurances. Functionalities include sending and receiving electronic messages to and from doctors' offices, making completing prescription renewal forms, appointments, and referral authorizations; viewing lists of current medications and allergies; and accessing health and practice information. Decision support can also assist patients in managing chronic illnesses, based on monitoring data.

PatientsLikeMe is one example for a PHR application. Within a web-platform, patients could view and understand their personal health data and exchange personal disease-specific information with each other based on data visualizations (Frost and Massagli 2008). Initially PHR pointed to an analogous medical file which was kept at home and regularly updated by the patient including various health information. At the beginning digital personal health records were limited (Kim and Johnson 2002) but nowadays the digital personal access to health information is becoming increasingly important (Archer et al. 2011; Tang et al. 2006; Frost and Massagli 2008).

While patients are experts regarding their personal behavior, environment or their perceived symptoms they are assumed to have less knowledge about medical tasks like diagnoses medication or laboratory test result interpretation as well as less numeracy. Therefore a user-oriented perspective on data visualizations has an even greater relevance. If confronted with medical data in patient portals patients perceived the information storage, the time-based view of lab results and reminder notifications as a major benefit. Concerns relate to privacy and a fear of reduced face-to-face interaction. Much older patients felt that it would be too costly learning how to use digital healthcare information systems though the authors hypothesizes that interest in technology correlates with age and low educational level. They summarized that a lack of ability, confidence and interest of older adults but also system complexity and a lack of usability, prevents elderly from using ICT supported systems. Elderlies said that looking at laboratory test results, schedule appointments, sending questions to the healthcare providers and functions to online renew their prescriptions are tasks they want to perform with online health platforms (Latulipe et al. 2015).

In order to provide patients with context information for the more autonomous interpretation and handling of their *medical records* Solomon et al. (2016) designed meaningful representations of test results within patient portals in a user-centered way. The authors found that color or text conveying contextual information can be effective, but may conflict with graphic related goals such as eliciting emotional responses or supporting clear communication between doctors and patients. The LifeLines system for example visualizes personal time-dependent data such as medical histories (Plaisant et al. 1996). Its evaluation included qualitative non-statistical approach to improve the software. Another system, InfoZoom, displays data sets as highly compressed as an interactive table. Both systems visualize raw data without temporal conclusion or domain knowledge. The KNAVE visualization platform instead visualizes time-oriented clinical data of a single patient. Clinicians can query, visualize and explore those data with the help of context-sensitive time data interpretations, as well as the dynamic visual exploration of the raw data (Goren-Bar et al. 2004). All three examples suggest, that time-oriented data plays an important role in health information visualization.

In comparison with the computer science domain, literature from psychological journals relevant for an age-dependent visualization design, do not focus on the visualization artifacts themselves but on perception and cognition related to it. Given our previous query results only (Breslow et al. 2009; Salthouse 2010, 2015) apply the term “visualization” explicitly and investigate data visualization relevant objectives. From the seven results of our initial search query only four appeared to be relevant for data and information visualizations. (Breslow et al. 2009) replicated the results from Merwin and Wickens (1993) showing a task x scale type interaction where the performance is higher-ranking using multicolored scales on identification task and superior using brightness scales on relative comparison task. Their results suggested that people use different search strategies during an identification task and a comparison task with color and brightness scales which are a parallel search strategy for comparison tasks and a serial search for identification tasks. The results of Breslow et al. (2009) implicate that a legend within a visualization increase execution time of visualization tasks. They furthermore emphasize the importance of color distances for the visual search in multicolored displays which require a color-difference standard to guide designers of color scales to produce easy to use color scales. While (Padilla et al. 2015) concentrate on non-expert decision making. Understanding how people interpret and use visually presented uncertainty data is an important yet seldom studied aspect of data visualization applications. The authors found that there was a difference in how decisions were made with spatial versus nonspatial glyphs, but no difference among the spatial glyphs themselves. Furthermore, the effect of different glyph types changed as a function of the variability of the distributions. Hesse et al. (2015) found through their review that life-sensitive, personal health record (PHR) can be used to support persistent conversations and situational awareness for teams involved into cancer care, while *mobile health apps* make it easier to support healthy behavior through self-regulatory feedback. Spence et al. (1999) found out that color codings for quantitative coding, which are close to perceptual linearity, were best for both simple and complex judgments in data visualization.

In summary we conclude, based on the literature we found, that visual representations of private health data and personal clinical data are an important field of research in computer sciences and in medical informatics. Research concentrates on the development of computational systems, data handling, system architecture and interaction methods. User-centered development and usability engineering methods are largely applied to build interactive health systems. Recent systems use mobile and wearable input and output devices. Data visualizations are used in the major part of all systems to provide healthy and ill subjects as well as physicians and clinical experts with health-related information. What appears is, that data visualizations are most of the time considered as just one part within the system and that age-differentiated recommendations regarding an ergonomic design merely exist.

3 Problem Statement

Human information processing is central for the perception and understanding of data visualizations. During the design and evaluation of data visualization human information processing has to be taken into account, especially as the fast aging of many western and eastern societies and their increasing reliance on information technology create a compelling need to reconsider older users' interactions with computers. Changes in perceptual and motor skill capabilities that often accompany the aging process bring important implications for the design of information and communication technologies.

The importance of an ergonomic perspective on data and information visualization research for digital healthcare information systems becomes apparent especially in the light of age-related changes. Many visual functions deteriorate slightly with age. This age-related slowing of visual function has been interpreted as reduced processing efficiency or effectiveness. Data visualizations thus need to consider age-related changes in order to be usable for the elderly who can be considered as one major target group. While general recommendations for age-related design of graphical user interfaces exist sparse can be found for data visualizations.

So far, the human factor has mainly been involved at an a posteriori stage of novel visualizations techniques or tools. As tasks strongly influence a certain evaluation their analysis needs to precede the visualization development and its further evaluation. The problem of such task-oriented research evolves from the different abstraction levels tasks have. Either they are domain dependent or domain independent. Domain-independent tasks which are used as basis for further evaluation studies might not be relevant for domain specific activities. Therefore, a detailed task analysis on the relation between various task abstraction levels and their interrelationship are necessary in order to generate generalizable results for all tasks relevant to the domain of digital healthcare information systems.

User-centred design for digital health-care systems prevalently starts after the main purpose or application area has been defined. Research and development activities are driven by a project idea or an initial hypothesis. Only few development activities consider the actual information needs of users as starting point for the development of a system.

4 Efforts Towards an Age-Differentiated Design of Data Visualizations

While the Tech4Age project (www.tech4age.de) focuses age-differentiated evaluation of digital health care systems in order to built a general pattern language summarizing design recommendations for system designers and engineers, given article presents the projects efforts on visualizations of personal health data. In total three studies will be described. Initially Sect. 4.1 describes results of a context analysis investigating health-related information need of people older than 60 years in Germany. Subsequently Sect. 4.2 describes efforts to formalize health-related tasks and abstract visualization tasks and according data types. This formalization will be the input for a study described in Sect. 4.3. This study will examine age-differences in performance and insight of data visualization in proportion to feature and conjunct search task performance. A conclusion and summary is provided in Sect. 5.

4.1 Elderlies Health Information Need and Behavior as Context of Visualizations of Personal Health Data

Deep understanding of user needs is crucial for building successful digital services and technology. At the beginning of research activities it is thus important to understand and specify the context of use, specify the user goals and activities in order to achieve quality in use. This is especially important for health-related systems with end users having exceptional requirements: elderlies, for example, are less familiar with modern information and communication technology (ICT) while visual acuity (Elliot et al. 1995), spatial vision starts declining (Sekuler et al. 1982) and visual sensitivity alters (Zhang and Sturr 1995). Also mental models of elderlies are characterized by enormous life experiences and thus strongly differ from users who are familiar with ICT systems. While we see information technology as a tool to which humans assess information, the analysis of user's *health information need* and *health information-seeking behavior* can motivate technology usage as it presides over engineering processes.

4.1.1 Research Questions and Background

Present context analysis aims at investigating which information has a meaning for elderly's personal health. Knowing what a patient needs to know about her/is health and which sources s/he applies to find those information, creates the basis for the design and the development of artifacts conveying health information to the elderly patient. This leads to the questions (1) how do elderly currently gain relevant health information, (2) what health-related information do elderly need (3) and how and why to they elderly approach certain information sources?

Information behaviour as human behaviour in relation to sources of information includes both active and passive information seeking, and information use. It includes direct communication between people, but also the passive information perception. Information seeking behaviour instead is an active and directed behaviour, where the user purposive seeks for information as a consequence of a need to satisfy some goal. In the course of seeking, the individual may interact with digital or non-digital information systems. Information behavior in general and especially within the healthcare context is determined by intrinsic and extrinsic factors which should also be captured during our investigation. Concerning information needs in healthcare Miller and Mangan (1983) showed that individual coping strategies have an impact on information needs. People avoiding to handle their illness are more aroused by a large quantity of illness-related information, while they are less if provided information quantity fits their coping strategy. Wilson and Walsh (1996)$_s$ model of information seeking behaviour illustrates the influence of personal and environmental factors. Besides, coping strategies as part of activation mechanisms, psychological, demographic, role, environmental and source characteristics have an impact on a person's information need (Auerbach 2001). Activity-theory (Miettinen 2009; Nardi 1996) and Wilsons model (Wilson 2000, 2005) on information seeking behavior build the foundation for questions on health-related activities, required information and confounding aspects. Validated tools captured the confounding concepts of coping strategies (Endler and Parker 1999) and need for cognition (Cacioppo et al. 1984; Verplanken et al. 1992). The Need for Cognition Scale is an assessment instrument that quantitatively measures the tendency for an individual to engage in and enjoy thinking. Results indicate that need for cognition was related weakly and negatively to being close minded, unrelated to social desirability, and positively correlated with general intelligence.

4.1.2 Method

Aim of this study was to investigate which health-related information need and behaviour do older adults in Germany have and how is this influenced by individual coping strategies and demographic variables. In order to investigate the context of health information systems, qualitative interviews provide rich and detailed information. Our aim was to generate insights and hypotheses rather than generalizable results. Respectively, we conducted a mixed-method study—involving structured surveys and in-depth interviews—with people about their health information sharing routines and preferences for different information sources. Participants (n = 20) older than 60 years were acquired via adult education institutions. At the beginning of each session, the participant first filled in a questionnaire about demographic and role-related parameters. Then the semi-structured interview investigating health-related information needs and behaviour was completed. All interviews were probing questions participants had about health, personal health and related concepts (medication, prevention, health and vital data, health insurance, hospitals or institutions) and regarding activities, goals and tools they use, share, and access health information with. The interview was based on existing interview guidelines on information needs of urban residents (Warner et al. 1973), information behaviour and needs related to information source/technology usage (internet, VT, print, smartphone, doctor, family and pharmacist). A standard questionnaire on coping strategies (Coping Inventory for stressful situations, CISS, Cosway et al. 2000) helped to post hoc divide the sample into groups. Finally, audio recordings were transcribed by means of the T4 transcription software (www.audiotranskription.de/english/f4.htm). Then two independent analysts iteratively developed an open and a theory-oriented coding theme with the Dedoose Software (www.dedoose.com).

4.1.3 Results

Results for n = 10 showed that the information the *elderly* require to stay healthy most frequently consists of a diagnostic assessment of an observed symptom along with cause estimations and treatment recommendations (n = 9). Three participants reported that they were completely satisfied with the health information they get, while six participants reported that they were partly satisfied with the information available to them. Four participants described problems getting information about examination results and related procedures. Six other participants also reported having trouble contacting their physicians or exchange vital data from laboratories or monitoring activities. In addition, four participants described information-sharing between medical experts as cumbersome. One participant suggested "*a solution that documents the content of each appointment, diagnosis and treatment as patient history that could be shared between physicians and viewed or even edited by patients*." Last but not least, four participants described information needs regarding health insurance services. Details about pricing and availability of chargeable health services were considered insufficient, especially if covered by private health insurance. Four participants desired greater transparency regarding billed services versus performed services. Sixty percent of the preliminary sample believed that an excessive preoccupation with health-related information could trigger a disease. As a result, they avoid devoting any more attention than necessary to health issues. In contrast to the 60 % group that believes an excessive preoccupation with health-related information could trigger a disease and so they avoided engaging in health information-seeking behavior, we identified a second group of 40 % that actively engages in health-related information behavior. They (1) put effort into quantifying and documenting personal health data in order to monitor their health, (2) strive to improve health-relevant behaviour and (3) cooperatively use the data they gathered to communicate health-related information to stakeholders. Elderly patients perceive their physician as a competent professional authority to whom they outsource information processes and decisions so as to not burden themselves with information searches and decision-making in addition to dealing with their disease. Participants rated their family doctor or a specialist as their most important health information source (n = 9).

4.2 Formalizing Health Visualization Tasks and Data Types With Regard to Generalizable Evaluation Results

Investigating ergonomic aspects of data and information visualizations requires a solid model of relevant tasks in order to use these as experimental tasks during evaluation. Tasks per se differ in domain relevance and abstraction level. To our knowledge no information exists about user-centred general task analysis for the digital information systems domain. Furthermore the relevance of individual abstract visualization tasks, and corresponding data-types, for domain-specific health tasks remains unclear. Brehmer and Munzner's (2013) differentiated different perspectives of visualization tasks based on the concept of cognitive task analysis (Vicente 1999). Unfortunately, healthcare and telemedicine taxonomies predominantly try to differentiate ambiguous terms representing the concept of IT supported medical processes. Bashshur et al. (2011) provides a conceptual context of the terms e-health addressing tasks as functionality dimensions: consultation, diagnosis, monitoring and mentoring. His research remains vague when it comes to the origin of his classification. Therefore, we want to verify it from the user's perspective and extend it. Additionally we want to know which abstract visualization tasks are relevant for the tasks in digital healthcare systems (telemedical systems).

In order to find an answer to the previously mentioned research question we set up an online questionnaire consisting of fifteen questions. Bashshur's et al. (2011) classification of telemedical functionality dimensions, Brehmer and Munzner (2013) multi-level model of abstract visualization tasks and Shneiderman (1996)'s task-by-data-type taxonomy for visualizations build the basis for the questionnaire (see Table 1).

We sent the link to the online questionnaire to N = 400 health-care systems experts and elderlies while getting N = 48 answers from experts and N = 50 from non-expert elderly (older than 60 years). Preliminary results based on a sample size of n = 10 experts suggest that their answers reflect parts of Bashshur's telemedicine tasks. Experts agreed that one-dimensional data, group data, single values, tree structures and distributions are relevant for medical consultation tasks, while quantitative, time dependent, two-dimensional data as well as data organized in a net structure, single values, outliers and nominal data are the most important ones over all tasks. Concerning relevant data types for data visualizations during mentoring, time dependent, ordinal, three-dimensional data are together with anomalies tree data and distributions the most important ones. Quantitative, time dependent and ordinal data are instead together with group dependent and distributions the most important ones.

Table 1 Questions investigating general health tasks in relation to abstract visualization tasks and according data types

No.	Question	Answer option
1	What medical tasks and activities can be supported by digital health systems?	Text field
2	Which data play an important role for digital health systems?	Text field
3	What data are required for medical consultation?	Text field
4	What data are required for medical diagnosis?	Text field
5	What data are required for medical mentoring?	Text field
6	What data are required for medical monitoring?	Text field
7	Specify what data is required for the following medical tasks. [horizontally: consultation, diagnoses, mentoring, monitoring; vertically: datatypes (Shneiderman 1996)]	Text field
8	Indicate which abstract visualization tasks (vertically) are required for given medical tasks (horizontally) [horizontally: consultation, diagnoses, mentoring, monitoring, vertically: abstract tasks from Brehmer and Munzner (2013)]	Text field
9	Assess the benefit of digital health systems for the following tasks. (Horizontally (tasks): consultation, diagnoses, mentoring, monitoring, vertically (benefit assessment): very high, high, moderate, low, very low)	Text field
10	What medical domains benefit from digital health systems?	Text field
11	Which diseases can be treated better with digital health systems?	Text field
12	For which places of care are digital health systems suitable?	Text field
13	What treatment methods can be supported by digital health systems?	Text field
14	How important do you think are the following application dimensions for digital health systems? (Horizontally: medical domain, symptoms, location of care, method of treatment, vertically: very important, important, neutral, unimportant, very unimportant)	Checkbox matrix
15	Assess your knowledge in digital health systems.	Likert Scale

4.3 Investigating the Relation Between Age-Dependent Visual Search Task Performance and Visualization Task Performance and Insight

Older adults often report difficulties when searching for items within cluttered visual scenes (Kline et al. 1992). Often feature search characteristics and pre-attentive visual features are reported to have an influence on visualization task performance. The efficiency of an individuals search is known to vary with age which leads us to the assumption that visualization benchmark and insight tasks might vary as well. This subchapter will therefore describe background, problem and planned experiment to investigate age-dependent performance on feature and conjunct search in relation to visualization performance and insight tasks.

Visual information acquisition and processing can be described by the paradigm of *feature search* and *conjunct search*, where a target must be identified within a number of distractors. The process of search and the search time in visual search tasks depend on the characteristics of the target and distractor objects. Treisman and Gelade (1980) constituted in their feature integration theory (FIT) that certain fundamental features exist which are perceived in parallel (shape, color, texture). When an object is different from other objects in one of these characteristics, this object is pre-attentively perceived (feature search). This is done automatically and unconsciously within a few milliseconds (200 ms) and is independent of the number of other objects. The top-down-driven perception of objects with conjunct properties (conjunct search) instead corresponds to a sequential process that requires more time and depends on the number of distractors.

The model of *directed search* (Wolfe et al. 1989; Wolfe 1994) however, assumes that the pre-attentive search, is processed at a large part of the visual field. Wolfe and his colleagues assumed that information that has been detected in parallel during the pre-attentive phase are used to direct the sequential process of perception, which are then proceeded in a smaller area of the field of with more attention. This can also be observed by means of the symmetry factor during conjunct search tasks. Wolfe and Friedman-Hill (1992) have shown that the visual search can be facilitated by a vertically symmetrical arrangement of distractors. This suggests that the distractor symmetry is processed in parallel and facilitates to detect a target among an array of distractors. One question that arises regarding the sequential search is to which extent the working memory is involved and whether once detected distractors are stored as such in memory and are not considered in the further search again, or whether the visual search corresponds to a non-cognitive process. By means of empirical and inferential studies following the dual-task paradigm, visual search task were associated with a task of visual working memory, Woodman et al. (2001) showed that an occupation of the working memory (through remembering one, two or four colors as well as forms had to be kept in mind) does not affect the performance in visual search tasks. Oh and Kim (2004) conducted a similar study, but they differentiated between the availability of visual working memory and spatial working memory. With regard to the visual working memory this results are similar with those of Woodman et al. (2001). The simultaneous execution of a task of spatial working memory however interferes with visual search processes. Woodman and Chun (2006) came based on the literature to the conclusion that the spatial working memory for storing position properties is involved to visual search and that a once fixated item is supposed to be memorized there. The visual working memory for storing object properties, however, is not required for visual search. But additionally, the eccentricity of the target object has an effect on the visual search: the further a target is away from the place of fixation, the later it is detected (Carrasco et al 1995;. Wolfe 1998). Rayner and Fisher (1987) assume that there are two parts or areas within the search for a target letter: a central area by which all the information about the target object are available, and a preview area by which some information about the target can be detected, which is not in a recognizable range. The more the size of both areas increases, the more similar target and distractors are. Wolfe et al. (1998) were able to provide evidence that it is the eccentricity effects not a purely visual process, but that attentional process plays a role in here. Their studies show that attention is drawn more quickly to central/foveal items than on peripheral ones. In addition to the eccentricity effects of the target object the position of distractors has an additional impact on the visual search.

Already from the age of 30 on, longer reaction times are observed during (Hommel et al. 2004). Age-differences occur only slightly within the pre-attentive search tasks (feature search), while they are more prominent during attentive top-down-driven conjunct search tasks. However, the strong age-differences of conjunct search tasks add by the number stimuli (Hommel et al. 2004; Plude and Doussard-Roosevelt 1989) and by heterogeneity of the distractors (Madden et al. 1996). Nothing is known about age-dependent feature number thresholds from which on the performance in conjunct search tasks declines. Furthermore, an age-related shift of the speed-accuracy trade of was found in the visual search. Elderly require typically more time but make fewer mistakes than younger participants (Strayer and Kramer 1994; Rogers and Gilbert 1997). In this context, it could further be shown that older adults stronger tend to fixated a stimuli a second than younger participants (Veiel et al 2006; Mitzner et al. 2010) and elderly tend to fixate the target stimuli longer than younger participants (Veiel et al. 2006). The results of Maltz and Shinar (1999) also revealed that older people have significantly longer search times, more fixations and shorter saccades than younger people. Studies of Scialfa et al. (1994) and Ball et al. (1988) suggest that the area in which stimuli can be detected without fixation change during visual search tasks is smaller in elderly subjects than in younger subjects.

As pre-attentive and conjunct features have been cited to have an impact on humans handling and understanding of data visualizations, and as age-related differences have been reported there, our first objective of this study is to investigate age-related performance in visualization tasks involving standard time based visualizations of health data. A second objective will be to investigate the feature search performance of a large sample of participants across adulthood and to which extend their feature search performance influences the performance in visualization benchmark and insight tasks. Additionally, the influence of different attentional and perception parameters on visualization task performance will be tested.

For the identification of age-related factors we will use a mixed design in which participants in each age group (20-50, 50-90 years) will perform benchmark and insight-tasks by using different visualization of health data. The data type and kind of benchmark tasks will be defined by the output of the study on task-data-type-taxonomy. Only health-relevant tasks and data types will be considered. After the visualization tasks participants will complete 30 trials searching for a target item on each of 5 different visual search displays. Three search displays will be used to assess feature, double and triple conjunction search (target present trials only). Two search displays will be used to assess exhaustive searching versus stopping in feature and conjunction searches (target absent versus present trials). Preliminary perceptual and cognitive tests include intensity and selectivity aspects of individual attention, like general attention, vigilance and short-term, early and long-term attentional activation measured by means of the TAP 2.3 (Zimmermann und Fimm 2002). To understand mental processes during feature search and during visualization tasks more objectively, the participant's eye movement will additionally be captured. In addition, contrast sensitivity and individual visual acuity will be recorded with the Optovist vision test equipment. The planned study is used to determine the relationship between perception- and attention parameters and performance in dealing with data visualization quantitatively. Potential age-differences will finally be examined with analysis of variance; relationships between individual factors will be tested by means of correlation and regression analysis.

5 Summary and Conclusion

A review on existing computer science and psychology literature revealed the importance of visualizations for personal health information systems. It became clear that digitalization of health processes and services increasingly incorporate medical experts and elderlies into cooperative care processes where personal health data is exchanged between medical experts and private users. While a lot of health applications incorporate data visualizations their definition is ambiguously and ranges from graphical representation of abstract statistical data to graphical representations of realistic physical objects. Data visualizations are treated as a part a technical system and are seldomly evaluated and designed age-dependently. Consequently a lack of knowledge regarding an age-differentiated design of health data visualizations was identified. A description of three individual studies addresses this lack of knowledge was precedingly presented. Preliminary results from a qualitative user study on health information need and behavior of older adults in Germany substantiated findings from the initial literature review and identified the need for cooperative health care and data exchange between patients and medical experts. Furthermore, a task-dependent study which we see as basis for generalizable results of ongoing visualization studies was described. So far time-dependent data are the most important data types, while diagnosis and symptom evaluation are important medical tasks for experts as well as novices. Finally the research design of an evaluation study aiming at general recommendations for the age-differentiated evaluation of health data visualizations was presented. Here, the relation of age-dependent feature and conjunct search performance on visualization benchmark and insight tasks will provide general recommendations for the age-dependent design of health data visualizations.

References

Ali-Hasan N, Gavales D, Peterson A, Raw M (2006) Fitster: social fitness information visualizer. In: CHI'06 extended abstracts on human factors in computing systems. ACM, New York, NY, pp 1795–1800. doi:10.1145/1125451.1125792

Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE (2011) Personal health records: a scoping review. J Am Med Inform Assoc 18(4):515–522, http://doi.org/10.1136/amiajnl-2011-000105

Auerbach SM (2001) Do patients want control over their own health care? A review of measures, findings, and research issues. J Health Psychol 6(2):191–203, http://doi.org/10.1177/135910530100600208

Ball KK, Beard BL, Roenker DL, Miller RL, Griggs DS (1988) Age and visual search: expanding the useful field of view. J Opt Soc Am 5(12):2210–2219

Bashshur R, Shannon G, Krupinski E, Grigsby J (2011) The taxonomy of telemedicine. Telemed J E-Health Off J Am Telemed Assoc 17(6):484–494

Brehmer M, Munzner T (2013) A multi-level typology of abstract visualization tasks. IEEE Trans Vis Comput Graph 19(August):2376–2385

Breslow LA, Trafton JG, Ratwani RM (2009) A perceptual process approach to selecting color scales for complex visualizations. J Exp Psychol Appl 15(1):25–34. doi:10.1037/a0015085

Brown B, Chetty M, Grimes A, Harmon E (2006) Reflecting on health: a system for students to monitor diet and exercise. In: CHI'06 extended abstracts on human factors in computing systems. ACM, New York, NY, pp 1807–1812. doi:10.1145/1125451.1125794

Browne SH, Behzadi Y, Littlewort G (2015) Let visuals tell the story: medication adherence in patients with type ii diabetes captured by a novel ingestion sensor platform. JMIR Mhealth Uhealth 3(4), e108, http://doi.org/10.2196/mhealth.4292

Cacioppo JT, Petty RE, Kao CF (1984) The efficient assessment of need for cognition. J Pers Assess 48(3):306–7

Carrasco M, Evert DL, Chang I, Katz SM (1995) The eccentricity effect: target eccentricity affects performance on conjunction searches. Percept Psychophys 57(8):1241–1261

Cosway R, Endler NS, Sadler AJ, Deary IJ (2000) The coping inventory for stressful situations: factorial structure and associations with personality traits and psychological health. J Appl Biobehav Res 5(2):121–143

Elliot DB, Yang KCH, Whitaker D (1995) Visual acuity changes throughout adulthood in normal, healthy eyes: seeing beyond 6/6. Optom Vis Sci 72(3):186

Endler NS, Parker JDA (1999) CISS: coping inventory for stressful situations. Multi-Health Systems, Toronto

Epstein D, Cordeiro F, Bales E, Fogarty J, Munson S (2014) Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In: Proceedings of the 2014 conference on designing interactive systems. ACM, New York, NY, pp 667–676. doi:10.1145/2598510.2598558

Frost J, Massagli M (2008) Social uses of personal health information within patientslikeme, an online patient community: what can happen when patients have access to one another's data. J Med Internet Res 10(3), e15, http://doi.org/10.2196/jmir.1053

Frost J, Smith B (2002) Visualizing health practice to treat diabetes. In: CHI'02 extended abstracts on human factors in computing systems. ACM, New York, pp 606–607. doi:10.1145/506443.506506

Goodman E, Foucault BE (2006) Seeing fit: visualizing physical activity in context. In: CHI'06 extended abstracts on human factors in computing systems. ACM, New York, NY, pp 797–802. doi:10.1145/1125451.1125609

Goren-Bar D, Shahar Y, Galperin-Aizenberg M, Boaz D, Tahan G (2004) KNAVE II: the definition and implementation of an intelligent tool for visualization and exploration of time-oriented clinical data. In: Proceedings of the working conference on advanced visual interfaces. ACM, New York, NY, pp 171–174. doi: 10.1145/989863.989889

Hesse BW, Beckjord E, Finney LJ, Rutten AF, Cameron LD (2015) Cancer communication and informatics research across the cancer continuum. Am Psychol 70(2):198–210

Hommel B, Li KZ, Li S-C (2004) Visual search across the life span. Develop Psychol 40(4):545–558

Jang A, MacLean D, Heer J (2013) Designing a prototype interface for visual communication of pain. In: CHI'13 extended abstracts on human factors in computing systems. ACM, New York, NY, pp 427–432. doi: 10.1145/2468356.2468432

Jones AP (1999) Indoor air quality and health. Atmos Environ 33(28):4535–4564, http://doi.org/10.1016/S1352-2310(99)00272-1

Kim MI, Johnson KB (2002) Personal health records: evaluation of functionality and utility. JAMIA 9(2):171–180, Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11861632

Kim S, Paulos E, Mankoff J (2013) inAir: a longitudinal study of indoor air quality measurements and visualizations. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, NY, pp 2745–2754. doi:10.1145/2470654.2481380

Kline DW, Kline TJ, Fozard JL, Kosnik W, Schieber F, Sekuler R (1992) Vision, aging, and driving: the problems of older drivers. J Gerontol 47(1):27–34

Latulipe C, Gatto A, Nguyen HT, Miller DP, Quandt SA, Bertoni AG et al (2015) Design considerations for patient portal adoption by low-income, older adults. In: Proceedings of the 33rd annual ACM conference on human factors in computing systems. ACM, New York, pp pp. 3859–3868. doi:10.1145/2702123.2702392

Madden DJ, Pierce TW, Allen PA (1996) Adult age differences in the use of distractor homogeneity during visual search. Psychol Aging 11(3):454–474

Maltz M, Shinar D (1999) Eye movements of younger and older drivers. Hum Fact J Hum Fact Ergon Soc 41:15–25

Merwin DH, Wickens CD (1993) Comparison of eight color and gray scales for displaying continuous 2D data. In: Proceedings of the human factors and ergonomics society… annual meeting human factors and ergonomics society. Meeting. pro.sagepub.com. http://pro.sagepub.com/content/37/19/1330.short

Miettinen R (2009) Dialogue and creativity: activity theory in the study of science, technology and innovations. Lehmanns Media, Berlin

Miller SM, Mangan CE (1983) Interacting effects of information and coping style in adapting to gynecologic stress: should the doctor tell all? J Pers Soc Psychol 45(1):223–236

Mitzner TL, Touron DR, Rogers WA, Hertzog C (2010) Checking it twice: agerelated differences in double checking during visual search. In: Proceedings of the human factors and ergonomics society annual meeting 2010, vol 54, pp 1326–1330

Nardi BA (1996) Context and consciousness: activity theory and human-computer interaction. MIT Press, Cambridge, MA

Oh S-H, Kim M-S (2004) The role of spatial working memory in visual search efficiency. Psychon Bull Rev 11(2):275–281

Padilla LM, Hansen G, Ruginski IT, Kramer HS, Thompson WB, Creem-Regehr SH (2015) The Influence of Different Graphical Displays on Nonexpert Decision Making under Uncertainty. J Exp Psychol Appl 21(1):37–46

Plaisant C, Milash B, Rose A, Widoff S, Shneiderman B (1996) LifeLines: visualizing personal histories. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, New York, pp 221–227. Retrieved from http://drum.lib.umd.edu/bitstream/handle/1903/5670/TR_95-88.pdf?sequence = 1&isAllowed = y

Plude DJ, Doussard-Roosevelt JA (1989) Aging, selective attention, and feature integration. Psychol Aging 4(1):98–105

Rayner K, Fisher DL (1987) Letter processing during eye fixations in visual search. Percep Psychophys 42(1):87–100

Rogers WA, Gilbert DK (1997) Do performance strategies mediate age-related differences in associative learning? Psychol Aging 12(4):620–633

Salthouse TA (2010) Influence of age on practice effects in longitudinal neurocognitive change. Neuropsychology.doi:10.1037/a0019026

Salthouse T (2015) Test experience effects in longitudinal comparisons of adult cognitive functioning. Develop Psychol. doi:10.1037/dev0000030

Scialfa CT, Thomas DM, Joffe KM (1994) Age differences in the useful field of view: an eye movement analysis. Optom Vis Sci 71(12):73–742

Shneiderman B (1996) The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings of IEEE symposium on visual languages, 1996. IEEE, Boulder, CO, pp 336–343

Sekuler R, Owsley C, Hutman L (1982) Assessing spatial vision of older people. Am J Optom Physiol Optic 59(12):961–68

Solomon J, Scherer AM, Exe NL, Witteman HO, Fagerlin A, Zikmund-Fisher BJ (2016) Is This good or bad?: Redesigning visual displays of medical test results in patient portals to provide context and meaning. In: Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems. ACM, New York, NY, pp 2314–2320. doi:10.1145/2851581.2892523

Spence I, Kutlesa N, Rose DL (1999) Using color to code quantity in spatial displays. J Exp Psychol Appl 5(4):393–412

Strayer DL, Kramer AF (1994) Aging and skill acquisition: learning-performance distinctions. Psychol Aging 9(4):589–605

Subramonyam H (2015) SIGCHI: Magic mirror—embodied interactions for the quantified self. In: Proceedings of the 33rd annual acm conference extended abstracts on human factors in computing systems. ACM, New York, NY, pp 1699–1704. doi: 10.1145/2702613.2732884

Swan M (2012) Sensor Mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0. J Sens Actuator Netw 1(3):217–253, http://doi.org/10.3390/jsan1030217

Swann-Sternberg T, Singh A, Bianchi-Berthouze N, Williams A (2012) User needs for technology supporting physical activity in chronic pain. In: CHI'12 extended abstracts on human factors in computing systems. ACM, New York, pp 2441–2446. doi:10.1145/2212776.2223816

Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ (2006) Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 13(2): 121–126

Theng Y-L, Fernando ONN, Deshan C, Goh LYQ, Hong J, Sen A et al (2013) CuePBox: an integrated physical and virtual pillbox for patient care. In: CHI'13 extended abstracts on human factors in computing systems. ACM, New York, NY, pp 433–438. doi:10.1145/2468356.2468433

Treisman AM, Gelade G (1980) A feature-integration theory of attention. Cogn Psychol 12(1): 97–136

Tsujita H, Rekimoto J (2011) Happinesscounter: smile-encouraging appliance to increase positive mood. In: CHI'11 extended abstracts on human factors in computing systems. ACM, New York, NY, pp 117–126. doi:10.1145/1979742.1979608

Vaara EK, Höök K, Tholander J (2009) Mirroring bodily experiences over time. In: CHI'09 extended abstracts on human factors in computing systems. ACM, New York, NY, pp 4471–4476. doi:10.1145/1520340.1520685

Veiel LL, Storandt M, Abrams RA (2006) Visual search for change in older adults. Psychol Aging 21(4):754–762

Verplanken B, Bas V, Hazenberg PT, Palenéwen GR (1992) Need for cognition and external information search effort. J Res Pers 26(2):128–36

Vicente KJ (1999) Cognitive work analysis: toward safe, productive, and healthy computer-based work. CRC Press

Warner ES, Murray AD, Palmour VE (1973) Information needs of urban residents. Final report on contract no OEC-0-71-4555 to US Office of Education

Wilson TD (2000) Human information behavior. Inform Sci 3(2):49–56

Wilson TD (2005) A re-examination of information seeking behaviour in the context of activity theory. Inform Res 11(4). InformationR. net: 1

Wilson TD, Walsh C (1996) Information behaviour: an inter-disciplinary perspective: a review of the literature. British Library Research and Innovation Centre London

Wolfe JM (1994) Guided search 2.0: a revised model of visual search. Psychon Bull Rev 1(2): 202–238

Wolfe JM (1998) Visual search. In Pashler H (ed) Attention. Erlbaum/Taylor and Francis, London, pp 13–73

Wolfe JM, Friedman-Hill SR (1992) On the role of symmetry in visual search. Psychol Sci 3: 194–198

Wolfe JM, Cave KR, Franzel SL (1989) Guided search: an alternative to the feature integration model for visual search. J Exp Psychol Hum Percept Perform 15(3):419–433

Wolfe J, O'Neill P, Bennett S (1998) Why are there eccentricity e_ects in visual search? visual and attentional hypotheses. Percept Psychophys 60(1):140–156

Woodman GF, Chun MM (2006) The role of working memory and long-term memory in visual search. Vis Cogn 14(4–8):808–830

Woodman GF, Vogel EK, Luck SJ (2001) Visual search remains efficient when visual working memory is full. Psychol Sci 12(3):219–224

www.audiotranskription.de/english/f4.htm. Retrieved 27 Jul 2016

www.dedoose.com/. Retrieved 27 Jul 2016

Zhang L, Sturr JF (1995) Aging, background luminance, and threshold-duration functions for detection of low spatial frequency sinusoidal gratings. Optom Vis Sci 72(3):198–204

Zimmermann P, Fimm B (2002) A test battery for attentional performance. In: Leclercq M, Zimmermann P (eds) Applied neuropsychology of attention. Theory, diagnosis and rehabilitation. Psychology Press, London, pp 110–151

How the Duration of Automated Driving Influences Take-Over Performance and Gaze Behavior

Anna Feldhütter, Christian Gold, Sonja Schneider, and Klaus Bengler

Abstract

The take-over of the driving task in highly automated vehicles at system limits is subject to latest research in ergonomics and human-machine-interaction. Most studies focus on driving simulator studies, examining the take-over performance mainly after short periods of automated driving, although take-over requests may not occur such frequently in future automated vehicles. This study tries to close this gap and compares driving performance and reaction times of a take-over after 5 and 20 min of automated driving. Further, the gaze behavior in the beginning and in the end of the 20 min period is compared. While the duration of automated driving did not show to influence the take-over performance, gaze behavior changed within the 20 min of automated driving. The SuRT and the 20 min automation period induced slower reactions, but no significant changes regarding accelerations and time to collision.

Keywords

Take-over • Automated driving • Automation effects • Gaze behavior

1 Introduction

In highly automated vehicles (Gasser 2012) or conditional automation (SAE International 2014), the driver is not required to monitor the system, but be available to take over control at system limits. The performance of the driver in such take-over situations is a crucial aspect when considering controllability of highly automated

A. Feldhütter (✉) • C. Gold • S. Schneider • K. Bengler
Lehrstuhl für Ergonomie, Technische Universität München, Boltzmannstraße 15, 85747 Garching, Germany
e-mail: feldhuetter@lfe.mw.tum.de; gold@lfe.mw.tum.de; schneiders@lfe.mw.tum.de; bengler@lfe.mw.tum.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_22

vehicles. Within the last years, several studies focused on this take-over and the influencing factors on take-over performance, like non-driving related tasks, performed while driving automated (Gold et al. 2015; Neubauer et al. 2012; Radlmayr et al. 2014) drivers' age (Körber et al. 2016; Petermann-Stock et al. 2013) or complexity of the situation (Gold et al. 2016). Next to those factors, a long duration of non-interrupted automated driving previous to a take-over seems to influence the take-over performance of drivers (Neubauer et al. 2012). This is the only study known to the authors, where automation effects of longer automated driving periods are considered and deterioration of driver performance, likely caused by monotony or fatigue, became apparent. Drivers engaged in different phone tasks showed quicker brake reactions compared to drivers without non-driving related task in a take-over that occurred after 25 min of automated driving. This is likely to be an important finding, as other studies include only short contacts with the automated system and periods of automated driving previous to the take-over request (TOR) of less than 5 min or only a few seconds (Louw et al. 2015). These short periods could be an unrealistic scenario, as a system frequently requesting a take-over may lack acceptance and may be rated as unsafe and would not make it to series production or frequent use. Therefore, this study tries to replicate results of the simulator study of Neubauer et al. (2012) and measure take-over performance, based on drivers' input and gaze behavior in take-over situations after different durations of automated driving and under consideration of two different non-driving related task (NDRT) conditions.

2 Method

2.1 Participants

In total, 31 participants completed the simulation drive whereas one person had to be excluded due to technical problems. The remaining sample of 30 participants consisted of 14 females (46.67 %) and ranged in age from 21 to 28 years with a mean age of 24.17 years ($SD = 2.09$). Mean driving experience was 7.2 years ($SD = 2.19$). Of the participants, 25 (83.33 %) have never experienced highly automated driving before.

2.2 Driving Simulator

The present study is based on a high fidelity static driving simulator of the Chair of Ergonomics which consists of a full vehicle mockup. Six projectors create a 180 degree field of view and allow the use of the side mirrors as well as the rearview mirror (see Fig. 1).

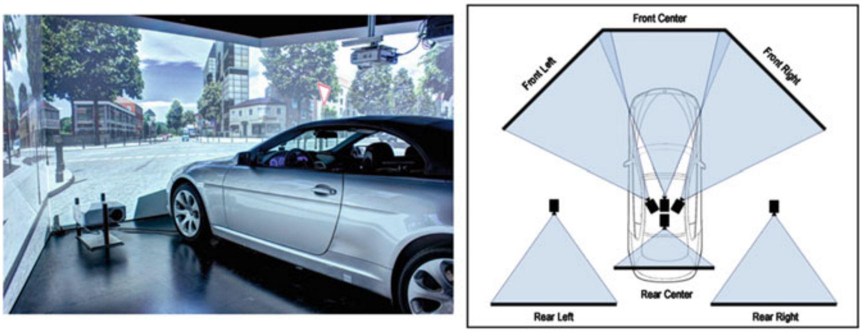


Fig. 1 The high fidelity static driving simulator of the Chair of Ergonomics

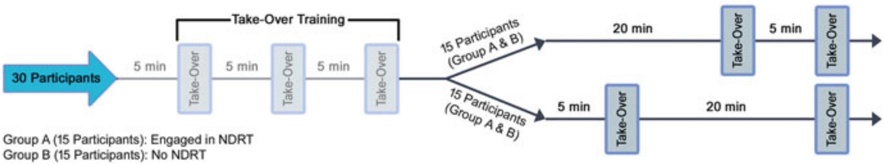


Fig. 2 Experimental design

2.3 Experimental Design

During the 45 min simulated drive on a three-lane highway the participants experienced in total five take-over situations. The first three take-over situations followed on 5 min of automated driving served as training and are not considered for examining the proposed research question. To investigate the effect of the automation duration on take-over performance and gaze behavior, the participants drove 5 and 20 min before the fourth and fifth take-over situation. The sequence of the two different automation durations was counterbalanced between participants (see Fig. 2). Furthermore, it is assumed that a non-driving related task could compensate possible deteriorations due to underload during longer automated drives by maintaining a suitable vigilance level. Therefore, there were two different NDRT conditions during the experiment. In the non-underload condition, half of the participants were engaged in the visually distracting standardized Surrogate Reference Task (SuRT (ISO/TS 14198)) before the take-over occurred (Group A). The SuRT was presented on a display in the middle console (see Fig. 3). To reduce monotony and avoid underload during the time period of automated driving, the SuRT was repeatedly offered for a short time (between 0.5 and 2.5 min). To generate the underload condition, the other half of the participants did not have any task to perform during the entire time of automated driving (Group B).

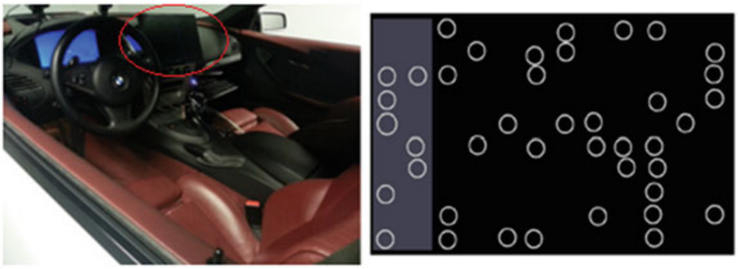


Fig. 3 The Surrogate Reference Task presented on a display in the middle console

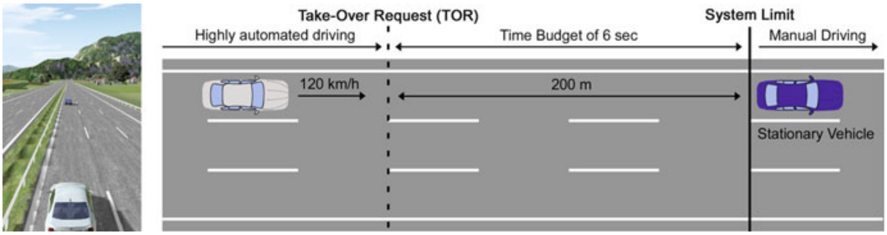


Fig. 4 *Left*: The simulated take-over situation on the left lane without additional traffic. *Right*: Illustration of the take-over procedure (cf. Damböck et al. 2012, modified)

2.4 Take-Over Situation

The take-over situation was identical for both automation durations (5 and 20 min). The ego-vehicle was located on the left lane of the three-lane highway with no other traffic on the remaining two lanes (see Fig. 4). The take-over due to a system limit was initiated by a suddenly appearing stationary vehicle on the lane of the ego-vehicle and was requested by an auditory alert. The time budget for taking over the driving task was set to 6 s which corresponds to a distance to the collision object of 200 meters at a velocity of 120 km/h. Figure 4 shows an illustration of the take-over situation. To ensure that each participant had the same time budget for taking over control, the stationary vehicle appeared in front of the ego-vehicle at the same time as the TOR sounded. Especially in the underload condition (i.e. without SURT), this is crucial as without visual distraction the participant could monitor the surrounding traffic situation the entire time and could therefore discover the obstacle too early.

2.5 Dependent Variables

In the present study, the effect of different automation durations on the take-over performance and the gaze behavior is investigated. In order to assess the take-over performance, the reaction time (RT), the take-over time (TOT), the maximum longitudinal (Acclong) and lateral acceleration (Acclat) of the ego-vehicle that occur during the take-over and the time-to-collision (TTC) are considered as dependent variables. The RT represents the time that the participants need for directing the first gaze away from the SuRT after the TOR. This measure is only applied in the SuRT condition. The TOT is the time that the participants need to start a conscious maneuver as a reaction to the TOR. A driver input is considered a conscious maneuver as soon as the steering wheel angle exceeds 2 degrees or the braking pedal position exceeds 10 % (Gold et al. 2013). The time that theoretically remains until a potential collision with an obstacle assuming constant speed of both, the ego-vehicle and the obstacle, is considered as the minimum occurred TTC within the take-over situation. Together with the longitudinal and lateral acceleration the surrogate safety measure TTC corresponds to the criticality of the take-over. These dependent variables are measured in the two take-over situations with different previous automation durations. The sample size for comparing the RT after 5 and 20 min diminishes from 30 to 12 participants as only half of them performed the SuRT during the TOR and three had to be removed from the analysis due to technical reasons.

In order to assess changes in the gaze behavior due to automation duration, the participants wore a head mounted eye tracking system (Dikablis 2.5) during the experiment. The driving scene which corresponds to the area of the windshield is defined as an area of interest (AOI). Here, the cumulative duration of the gazes, the average duration of one gaze, the maximum duration of one gaze and the number of gazes towards this AOI within 60 s are set as dependent variables. They are compared for the 3rd and 18th minute of the 20 min uninterrupted automated driving. Within these periods, no SuRT was presented. In order to assess the impact of underload, the dependent variables of the take-over performance (TOT, Acclat, Acclong and TTC) and of the gaze behavior (the cumulative duration of the gazes, the number of gazes, the average duration of one gaze as well as the maximum duration of one gaze towards the AOI within 60 s) are compared the conditions with SuRT to without SuRT.

2.6 Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics 22 software and consisted of two mixed analyses of variance (ANOVA). The first one focused on take-over performance and included task condition (SuRT vs. without SuRT) as between-subjects factor and duration of automated driving prior to take-over (5 vs. 20 min) as within-subjects factor. The second one analyzed the impact of task condition (between-participants factor) and the duration of previous uninterrupted automated driving (3rd minute vs. 18th minute, within-factor) on drivers' gaze behavior. For the comparison of the RT in the SuRT group, a two-sided t-test for paired samples was conducted. Throughout the whole analysis, a significance level of $\alpha = 5\ \%$ was set.

3 Results

Figure 5 presents the results regarding driver performance following the take-over request. The duration of automated driving prior to the take-over had no significant main effect, whereas the introduction of the SuRT led to an extended take-over time ($F(1, 28) = 12.19$, $p = 0.002$, $\Delta\ M_{TOT} = 0.36$ s). Consistently, the group engaged in the SuRT also showed a non-significant tendency towards a shorter minimum time-to-collision ($F(1, 28) = 3.62$, $p = 0.067$, $\Delta\ M_{TTC} = 0.47$ s). No differences were observed concerning lateral and longitudinal acceleration. There were no significant interaction effects.

Regarding the reaction time in the SuRT condition, participants were significantly faster after 5 min ($M = 0.55$ s, $SD = 0.08$ s) when compared to the take-over after 20 min ($M = 0.64$ s, $SD = 0.09$ s; $df = 11$, $p = 0.014$, $\Delta\ M_{RT} = 0.09$ s; see Fig. 5).

In Fig. 6 the drivers' gaze behavior in the 3rd and 18th min of uninterrupted automated driving is compared. Towards the end of the 20 min period, the time spent looking at the driving scene decreased ($F(1, 28) = 5.37$, $p = 0.028$, $\Delta\ M_{CDG} = 5.25$ s). Gazes towards the driving scene tended to be shorter but more frequent, although values for the average duration ($F(1, 28) = 3.56$, $p = 0.070$, $\Delta\ M_{ADG} = 1.12$ s) and the number of gazes ($F(1, 28) = 3.60$, $p = 0.068$, $\Delta\ M_{NG} = 3.2$) did not reach significance. The maximum duration of single gazes decreased ($F(1, 28) = 9.70$, $p = 0.004$, $\Delta\ M_{MDG} = 3.74$ s). The SuRT led to an increase in the number of gazes ($F(1, 28) = 9.48$, $p = 0.005$, $\Delta\ M_{NG} = 5.4$) but did not affect any other variable. Again, no significant interaction effects were observed.

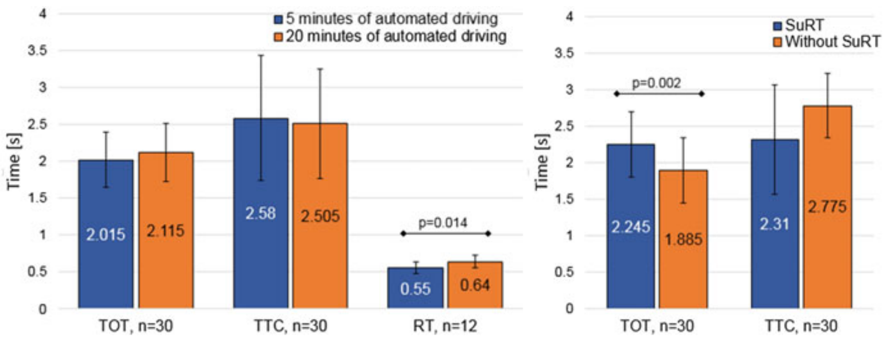


Fig. 5 *Left*: Take-over time (TOT), minimum time to collision (TTC) and reaction time (RT) compared after 5 and 20 min of uninterrupted automated driving. *Right*: TOT and TTC compared in the two NDRT conditions (SuRT vs. Without SuRT), error bars represents the standard deviation

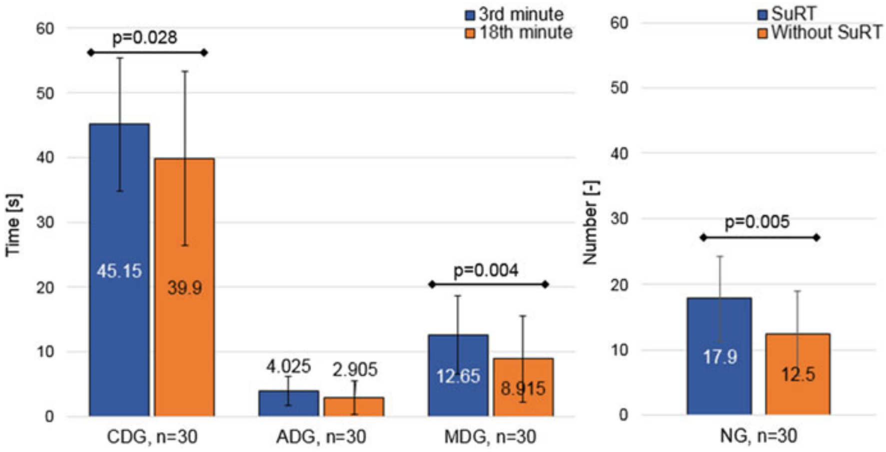


Fig. 6 *Left*: Cumulative duration of all gazes (CDG), average duration of one gaze (ADG) and maximum duration of one gaze (MDG) towards the driving scene within 60 s compared in the 3rd and 18th minute of uninterrupted automated driving. *Right*: Number of gazes (NG) towards the driving scene within 60 s compared in the two NDRT conditions. Error bars represent the standard deviation

4 Discussion and Limitations

In the present study, the effect of different automation durations prior a take-over request (5 vs. 20 min) on the take-over performance and the gaze behavior during the automated drive was investigated. Furthermore, it was examined whether the engagement in the SuRT during the automated drive has an effect on these parameters.

Concerning the take-over performance, the time for averting the first gaze from the SuRT towards the driving scene after a TOR significantly increased after 20 min of automated driving in the group with SURT. Longer reaction times may indicate a decrease in vigilance level and an increase in fatigue (Graw et al. 2004). Here, a duration effect seems to occur but it cannot be clarified whether it is due to automation or long term engagement in the SuRT. Concerning the remaining parameters of the take-over performance, no differences were observed comparing the 5 and the 20 min period. This lack of an evident deterioration due to fatigue is consistent with the findings of Neubauer et al. (2012) and Saxby et al. (2013). It is conceivable that due to an insufficient length of the chosen intervals no fatigue-related or hypovigilance-related deteriorations occurred. Additionally, it is also possible that after prior sufficient training, as it was provided in this study, deterioration effects on the take-over performance occurred already after 5 min of automated driving. Thus, no differences between the 5 and 20 min regarding the take-over performance could be found.

Regarding the task condition, results showed a significant increase of the TOT in the SuRT group. It is assumed that the extended TOT is caused by the visually and cognitively distracting SuRT. However, the remaining parameters for take-over performance did not show significant differences due to the task condition. Thus, it can be supposed that the extended TOT is induced by the time the participants needed to redirect the gaze from the visually distracting SuRT towards the driving scene. Therefore, Neubauer et al.'s (2012) finding that being engaged in a non-driving related task during automated driving enhances subsequent response speed could not be replicated.

Comparing the task conditions regarding the gaze behavior, only in the number of gazes towards the driving scene a significant difference was found. As none of the other gaze parameters showed any effect or tendency, it is assumed that the SuRT group frequently controlled the availability of the task, which was presented on a display in the center console of the vehicle and thereby increased the number of gazes to the driving scene.

Towards the end of the 20 min period the percentage of time spent watching the driving scene diminished. This might be caused by more frequent or prolonged blinking, which is an indicator for fatigue (Schleicher et al. 2008; Schmidt et al. 2009). However, there are more suitable metrics, like PERCLOS (Knipling and Wierwille 1994), for assessing fatigue by eye-tracking data which was not possible to be evaluated by the Dikablis system. A different approach to explain the diminished gaze duration towards the end of the 20 min period is that the participants did not spend the same attention to the driving scene. An increase in glance durations and glance numbers away from the AOI may be indicators for visual distraction (ISO 15007-1:2014). This may indicate that the participants showed more self-initiated distraction after longer duration of automated driving by averting the eyes from the driving scene and letting the gaze wander due to monotony and boredom. However, the number of gazes increased with longer automation duration. The assumption is that despite more visual distraction the participants still controlled the traffic and the driving scene by regular but short glances. As the glances towards the driving scene became shorter the participants tried to compensate the thereby reduced absorption of visual information by more frequent short glances (Damböck 2013). The more frequent control glances may also indicate that the participants did not trust the system enough to reduce checking the driving scene (Moray 2000) with being simultaneously visually distracted by letting the gaze wander. This is feasible as due to the experimental design the participants have previously experienced at least three take-over requests and therefore probably expected another take-over situation.

5 Summary

In this study, an effect of the duration of automated driving on the reaction time was found, while the remaining parameters for take-over performance did not show any significant differences. The gaze behavior was also affected by the duration of the automation. It is assumed that the drivers let their gazes wander and therefore showed self-initiated visual distraction due to monotony after 20 min of automated driving. In this experiment the non-driving related task, which was the visually distracting SuRT, did not show an effect on the take-over performance except for the take-over time. The longer times are presumably caused by the additional time the participants needed for redirecting the gaze from the SuRT to the driving scene after the take-over request.

Acknowledgement The opportunity to conduct this study was granted by the research program ASHAD (Automation in Society—The Case of Highly Automated Driving) by the Munich Center for Technology and Society, Technical University of Munich.

References

Damböck D (2013) Automationseffekte im Fahrzeug—von der Reaktion zur Übernahme. Dissertation, Technische Universität München, München

Damböck D, Farid M, Tönert L, Bengler K (2012) Übernahmezeiten beim hochautomatisierten Fahren. 5. Tagung Fahrerassistenz

Gasser TM (2012) Rechtsfolgen zunehmender Fahrzeugautomatisierung: Gemeinsamer Schlussbericht der Projektgruppe. Bremerhaven. Retrieved 09 Jul 2012

Gold C, Damböck D, Lorenz L, Bengler K (2013) Take over! How long does it take to get the driver back into the loop? Proc Hum Factors Ergon Soc Annu Meet 57(1):1938–1942

Gold C, Berisha I, Bengler K (2015) Utilization of drivetime—performing non-driving related tasks while driving highly automated. In: Proceedings of the Human Factors and Ergonomics Society (HFES) 2015. Annual Meeting, Los Angeles, CA

Gold C, Körber M, Lechner D, Bengler K (2016) Taking over control from highly automated vehicles in complex traffic situations—the role of traffic density. Hum Factors. doi:10.1177/0018720816634226

Graw P, Kräuchi K, Knoblauch V, Wirz-Justice A, Cajochen C (2004) Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. Physiol Behav 80(5):695–701

SAE International (2014) SAE surface vehicle information report: taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (No. J3016). Retrieved 12 Nov 2014

ISO 15007-1:2014 (11.2014): ISO International Organization for Standardization

ISO/TS 14198 (11.2012): ISO International Organization for Standardization.

Knipling R, Wierwille W (1994) Vehicle-based drowsy driver detection: current status and future prospects. In: Paper presented at the IVHS America fourth annual meeting, 17–20 April 1994

Körber M, Gold C, Lechner D, Bengler K (2016) The influence of age on the take-over of vehicle control in highly automated driving. Transport Res F Traffic Psychol Behav 39:19–32. doi:10.1016/j.trf.2016.03.002

Louw T, Merat N, Jamson AH (2015) Engaging with highly automated driving: to be or not to be in the loop? In: 8th International driving symposium on Human factors in driver assessment, training and vehicle design, Salt Lake City, UT. Retrieved 13 Nov 2015

Moray N (2000) Are observers ever really complacent when monitoring automated systems? Proc Hum Factors Ergon Soc Annu Meet 44(6):592–595

Neubauer C, Matthews G, Saxby D (2012) The effects of cell phone use and automation on driver performance and subjective state in simulated driving. Proc Hum Factors Ergon Soc Annu Meet 56(1):1987–1991

Petermann-Stock I, Hackenberg L, Muhr T, Mergl C (2013) Wie lange braucht der Fahrer?: Eine Analyse zu Übernahmezeiten aus verschiedenen Nebentätigkeiten während einer hochautomatisierten Staufahrt. In 6. Tagung Fahrerassistenzsysteme, München, 2013

Radlmayr J, Gold C, Lorenz L, Farid M, Bengler K (2014) How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. Proc Hum Factors Ergon Soc Annu Meet 58(1):2063–2067

Saxby DJ, Matthews G, Warm JS, Hitchcock EM, Neubauer C (2013) Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. J Exp Psychol Appl 19(4):287–300

Schleicher R, Galley N, Briest S, Galley L (2008) Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? Ergonomics 51(7):982–1010

Schmidt EA, Schrauf M, Simon M, Fritzsche M, Buchner A, Kincses WE (2009) Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. Accid Anal Prev 41(5):1087–1093

Uncanny and Unsafe Valley of Assistance and Automation: First Sketch and Application to Vehicle Automation

Frank Flemisch, Eugen Altendorf, Yigiterkut Canpolat, Gina Weßel, Marcel Baltzer, Daniel Lopez, Nicolas Daniel Herzberger, Gudrun Mechthild Irmgard Voß, Maximilian Schwalm, and Paul Schutte

Abstract

Progress in sensors, computer power and increasing connectivity allow to build and operate more and more powerful assistance and automation systems, e.g. in aviation, cars and manufacturing. Besides many benefits, new problems occur e. g. in human-machine-interaction. In the field of automation, e.g. vehicle automation, a comparable, metaphorical design correlation is implied, an unsafe valley e.g. between partially- and highly-automated automation levels, in which due to misperceptions a loss of safety could occur. This contribution sketches the concept of the (uncanny and) unsafe valley of automation, summarizes early affirmative studies, gives first hints towards an explanation of the valley, outlines the design space how to secure the borders of the valley, and how to bridge the valley.

F. Flemisch (✉)
IAW Institut für Arbeitswissenschaft, RWTH Aachen, Aachen, Germany

FKIE Fraunhofer Institut für Kommunikation, Informationsverarbeitung und Ergonomie, Fraunhoferstr. 20, 53343 Wachtberg, Germany
e-mail: frank.flemisch@fkie.fraunhofer.de

E. Altendorf • Y. Canpolat • G. Weßel
IAW Institut für Arbeitswissenschaft, RWTH Aachen, Aachen, Germany

M. Baltzer • D. Lopez
FKIE Fraunhofer Institut für Kommunikation, Informationsverarbeitung und Ergonomie, Fraunhoferstr. 20, 53343 Wachtberg, Germany

N.D. Herzberger • G.M.I. Voß • M. Schwalm
ika Institut für Kraftfahrzeuge, RWTH Aachen, Aachen, Germany

P. Schutte
Aviation Development Directorate AMRDEC/RDECOM, US Army, Brooklyn, NY, USA

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_23

Keywords
Automation • Assistance • Robotics • Human-machine systems • Uncanny unsafe valley

1 Introduction: Assistance, Automation and Robotics

Enabled by technical advancements in the field of sensors, computers and connectivity, as well as motivated by cost-pressure along with ever-increasing performance requirements, the complexity of information systems has steadily grown in the last decades (cf. Hollnagel 2007). A part of this complexity can be compensated with assistance systems and automation, however, unwanted side effects such as "Operator/pilot out of the loop" or "Mode confusion" (cf. Billings 1997) are reported in a variety of domains like aviation, nuclear power plants and automotive.

Rather than speaking about over-automation in an undifferentiated manner, Norman (1990) points out that not over-automation is the problem, but inappropriate feedback and interaction.

There is a concept in robotics, which can be considered as a specific form of automation, known as "The Uncanny Valley", where robots showing high, however imperfect similarities to humans are perceived by humans as uncanny and disconcerting (Mori 1970; Mori et al. 2012). Conscious of the Uncanny Valley, research and development in robotics is focusing on cooperative robotics, where to a certain extent humans and highly automated robots work together in the same work spaces, instead of fully-automated robots (cf. Mayer 2012; Kuz et al. 2015).

A similar development regarding cooperative assistance and automation is currently emerging in the area of ground vehicles ensuing from the aviation domain (e.g., Flemisch and Onken 1998; Schutte 1999; Goodrich et al. 2006). It became increasingly clear, through basic concepts such as Levels of Automation, that assistant and automated systems are related and, (a) should be discussed holistically and (b) could be depicted on a scale—that is, a spectrum of assistance and automation, (cf. Flemisch et al. 2003, 2008). This point of view was later applied in the standard categorization of vehicle automation (cf. BASt 2012a; SAE 2014), which differentiates between assisted, partially- and highly-automated systems. Figure 1 shows a simplified scale of assistance and automation related to the control

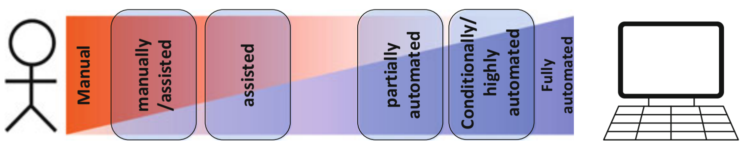


Fig. 1 Control distribution between the human and automation represented as an assistance- and automation-scale, here with explicit automation-levels/modes (inspired by Sheridan 1980; Flemisch et al. 2003, 2008, 2012, 2014, 2015a, b; Gasser et al. 2012b; SAE 2014)

distribution between the human and the automation in the assistance- and automation-levels including manual, assisted, partially-, highly- and fully-automated/autonomous.

A possible unsafe valley of automation can be found in the right half of the scale between partly- and highly-automated, which could be rather uncanny for the user, and more importantly, rather unsafe, as described further down.

2 Early Indicators for the Existence of an Unsafe Valley

There is a good chance that the metaphor of an (uncanny and) unsafe valley can be applied to automation in all kinds of domains. Early systematic explorations within the area of partly- and highly-automated vehicle control have been conducted for ground and air vehicles since 2003 (NASA-H-Mode) and since 2004 for ground vehicles as part of DFG-H(orse)-Mode-Projects.

These were inspired by the H-metaphor, a design metaphor that takes the rider-horse interaction as a blueprint for vehicle automation (Flemisch et al. 2003, 2015a, b; Bengler and Flemisch 2011; Altendorf et al. 2015). The initial base research sparked a series of national and EU-projects, introduced the term highly-automated driving (e.g. Flemisch et al. 2006; Hoeger et al. 2008, 2011) and inspired the development of partially-automated "piloted" driving e.g. of Volkswagen, Audi, Mercedes and the more chauffeur inspired Tesla.

At the beginning of this research and development in 2000, it was debated as to whether one or multiple modes between assisted- and fully-automated automation levels would be advisable and how they should be designed, especially regarding the involvement-degree of the operators, here the drivers, and the extent of the automation's intervention and the required safety measures, e.g. by operator monitoring.

With accumulating research, it became clear that there are combinations of partially- as well as highly-automated modes that are functional, while other implementations are not. An example of a well-functioning implementation of a lower automation level **in the car domain** is presented by Ma and Kaber (2005). In their implementation automation only supports longitudinal control. They revealed that an Advanced Cruise Control (ACC) system is able to enhance system performance in terms of lane deviations and speed control in tracking a lead vehicle and increase drivers situation awareness, even when drivers are distracted by a non-driving related task.

In the early decade of this century, research explored whether assistive (i.e., not fully autonomous) solutions beyond ACC, with coupled longitudinal and lateral control could be successful. Starting from the H-Mode projects, a series of studies (e.g. Schieben and Flemisch 2008; Schieben et al. 2008; Petermann and Schlag 2009), have shown that designs differentiating between partly- and highly-automated systems and using well-arranged transitions could succeed. Rauch et al. (2010) demonstrated that an integrated driver-state detection may improve the safety and acceptance of the automation systems.

Furthermore, Merat et al. (2012) presented a safe implementation of a highly-automated vehicle, which takes longitudinal and lateral control and can perform gentle maneuvers itself. They conducted an experiment in which the automation had complete control of the automobile in normal situations and drivers were warned when approaching an obstacle and manual control had to be resumed. The results showed that, in their implementation of highly-automated driving, no negative effects on system performance emerged.

Besides these successful examples, there are clear hints that areas between well-functioning modes and their variants exist, which are clearly less safe: For example, simulator-based studies conducted at DLR showed that partially-automated driving designs, where drivers did not have to apply steering torques any more, can cause problems in compensating system failures. However, when the driver still needed to apply some of the required steering torques, these failures could be absorbed (Schieben and Flemisch 2008). Furthermore, additional studies, (e.g., Damböck 2013; Schwalm et al. 2015; Schwalm and Ladwig 2015; Voß and Schwalm 2015) have shown, that highly automated designs could lead to reduced take-over capabilities of drivers. This correlation is proven, e.g., by a lack of compensatory reaction in terms of reducing activity in non-driving-related tasks that would be needed for an appropriate preparation of a take-over situation (Voß and Schwalm 2015; Schwalm et al. 2015; Schwalm and Ladwig 2015). Moreover, a real-world indicator for the existence of an uncanny/unsafe valley might be the first fatal crash of an automated ground vehicle, a partially automated Tesla, which has become public in 2016, just shortly after the first publication of the unsafe valley in February 2016 (Flemisch et al. 2016a).

As of 2016, everyday- and user-experience with highly automated ground vehicles is sparse, however additional indicators for the existence of an unsafe valley can be derived from the **aviation domain**. As the applicability of the H-Mode has shown, ergonomic principles for cooperative guidance and control can be applied to both domains. Moreover, Schutte et al. (2016) argue for a similar phenomenon in the aviation area. They present examples of systemic failure resulting from highly automated flying (National Transportation Safety Board 2010; et d'Analyses 2012) and argue that these incidents were partly due to bad system design instead of pilot failure alone. They presented an alternative cockpit design, which keeps the pilot in the loop. In terms of the unsafe valley they avoid falling into it by staying on the left side of the abyss. During scientific discussion in the context of this publication it turned out that also other researchers already had resembling ideas like the valley of automation. Thus e.g. Maurer (2016) describes an u-shaped curve for the relation between user-transperancy and driver assistance systems. Furthermore Maurer (2016) and Eckstein (2016) describe the grand canyon of the driver-assistance systems, which refers especially to the differences in effort and protection systems between partially and highly automated vehicles.

3 A First Glimpse: What Happens in an Unsafe Valley?

Are there scientific phenomena that could determine the existence and the characteristics of such an unsafe valley? Figure 2 pictures an example of a potential "crash" into an uncanny valley, here applied to the car domain. Within that example, a system failure with an adjacent take-over occurs in front of a narrow curve. In one condition (which here is named partially-/highly-automated, because it is was supposed to be operated as partially automated, but drivers used it in a way highly automated systems are used) this leads to an unsafe driving performance, i.e. a "crash" into an uncanny valley. This does not occur in another condition (partially-automated). This suggests that there might be a correlation between the control repartition—which ranges from manual to fully automated driving—and an output quantity, such as performance or driving safety. Figure 2 conceptualizes this issue: While a mode with a low degree of automation (M3.1: e.g., partly-automated) still provides sufficient safety, a higher automation level (M3.2) seems to be rather unsafe. However, passing this unsafe valley, an even higher automation level (M4) could be safe again.

An important element, which could lead to the safety drop in M3.2, is the connection between the automation performance and the operator's performance in case of a take-over situation: A potential correlation between the operator's involvement and his take-over capabilities could be that, if the former is too low, then the latter will not happen in time. A nominally highly reliable and capable automation that infrequently is insufficient and incapable will induce a pure monitoring role for the operator, i.e. "supervisory control", for which humans are not well prepared (e.g., Sheridan 1976; Endsley 1995). A pertinent approach from Schwalm et al. (e.g., Schwalm et al. 2015; Schwalm and Ladwig 2015; Voß and Schwalm 2015) postulates that operators, here drivers, abandon a continuous control and regulative process in case of a (too) high automation level. Due to this, they supposedly are no longer capable of applying regulatory measures in terms of a functional situation management. Instead, if the drivers are fully involved in the driving task, i.e. manual driving, this assumed regulatory process

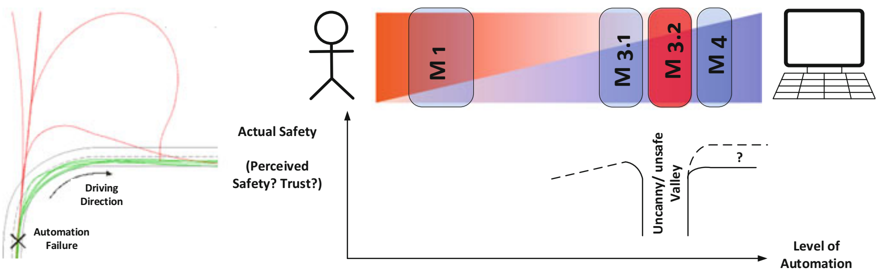
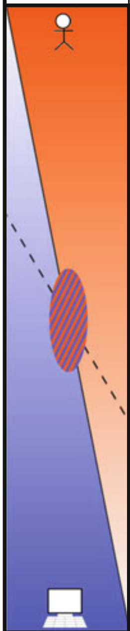


Fig. 2 An example of a "crash" in an Uncanny Valley: Intercepting a system failure prior to a curve (*green*: system variant partially-automated; *red*: system variant partially-/highly automated)

would allow them to analyze and anticipate the driving situation, and to adequately distribute available cognitive resources.

In order to conceptualize this idea of reduced driver's involvement as a risk factor for the appearance of an unsafe valley, a schema was developed for the context of driving, in which the different automation levels were combined with possible states of driver's involvement (cf. Fig. 3: Herzberger et al. 2016). These states depict minimal requirements for each level of automation regarding the driver's involvement which is necessary to guarantee safe driving performance. These minimum requirements are separately reported for the three levels of a driving task [navigation, guidance and stabilization; see Donges (2015)] in order to provide a more detailed sketch of the driver-system interaction. In Fig. 3, the orange shaded area represents driver-sided task fulfillment, while the blue shaded area represents system-sided task fulfillment (Herzberger et al. 2016).

In total, five driver states were defined on the basis of the SAE Levels (SAE International 2014). In the following, these five driving states will be described regarding the required extent of involvement, presented from high to low. The state Fully (F) requires an active performance of the entire driving task. The second state,



BASt	SAE		Navigation	Guidance	Stabilisation
Driver Only	0		F	F	F
			-	-	-
Assisted	1		F	I1	I1
			-	A	A
Partially automated	2		F	I2	I2
			-	L	L
	3		F	R	R
Highly automated			-	L	L
	4		-	-	-
Fully automated			F	F	F
	5		-	-	-
-			F	F	F

- - - minimal requirements for driver involvment
—— potential distribution of control between human and machine

Fig. 3 Minimal requirements for driver states. Note: The crosshatched area depicts the unsafe valley

Involvement (I), is divided into Involvement 1 (I1) and Involvement 2 (I2). State I1 still requires an active performance, either longitudinal or lateral control, and a cognitive monitoring of the entire driving task. Moreover, drivers in state I1 have to be ready to take over the driving task at any time without a take-over request (TOR). In contrast to I1, drivers in the state of I2 do not actively perform a driving task. However, they are still required to perform the cognitive monitoring. Still, drivers in state I2 have to be ready to take over the driving task at any time without a TOR. Regarding the state of Retrievable (R), a cognitive monitoring of the driving task is no more required. Nevertheless, the driver has to be able to resume the driving task after an appropriate TOR. In the final state, Non-Retrievable (NR), the driver cannot take over control of the vehicle. This state is reached, if the driver e.g. sleeps or if a transition from Retrievable to a higher level of involvement is not possible. These five states only define the minimal requirements for a driver's involvement for the safe management of the driving task. Any requirements for a possible intervention by the driver are not considered by now.

Two different vehicle conditions were specified: Firstly, in the condition of assisted driving (A) the vehicle either takes over longitudinal or lateral control during the driving task. Secondly, the condition of lead driving (L) which posits that a vehicle controls any longitudinal or lateral actions to a full extent. On top of that, the earlier introduced driver state Fully (F) can be transferred to the vehicle. It requires an active performance of the entire driving task by the vehicle.

Based on this concept of states of driver's involvement in the context of automated driving, it is possible to derive an explanation for a "crash" into the uncanny/unsafe valley: For drivers, the SAE Levels 2 and 3 seem to pose the same requirements in normal traffic of automated driving, as both lateral and longitudinal guidance are carried out by the system. Nevertheless, the actual requirements of a Level 2 system to the driver are clearly higher due to the missing TOR in case of a system failure. In Level 2 drivers have to constantly monitor the driving task (I2), while in Level 3 it is not required (R). In other words, drivers tend to misconceive a Level 2 (I2) state with a Level 3 (R) state (cf. orange crosshatched area in Fig. 3). Only if system limits are reached, the differences between the two (required) states of involvement will appear and the driver potentially crashes into the uncanny/ unsafe valley.

Another advantage of the classification of driver states is the timely recognition of a decreasing driver involvement. If this risk is recognized, possible consequences can be rebalanced and/ or reduced. In the following, a time course of occurrence probability of an uncanny/ unsafe valley is presented.

Figure 4 highlights the temporal course of a "crash" into an unsafe valley: At first, everything is working properly, because the automation's capabilities are sufficient for handling the vehicle control situation and the operator is still able to take over (T1). Over time, habituation effects can emerge (T2). If in that case a system failure occurs, a human operator's take-over capability would be insufficient for the safe management of required take-over maneuver (T3). Within that scenario, the unsafe valley metaphorically resembles a crevasse (ice crevice) that seems to be crossable over a stable bridge. This bridge, however, will break down as soon as it is

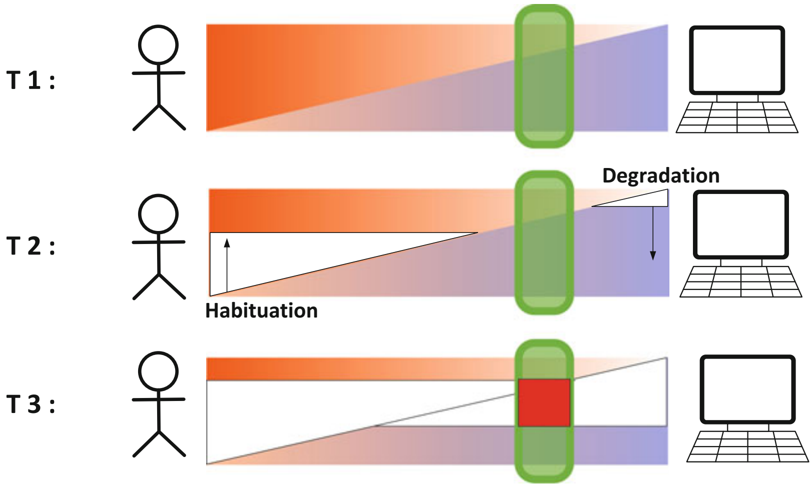


Fig. 4 A possible connection between control-distribution and certainty

used and will devour the all too trustful operator. Another possible metaphor depicts a pair of scissors: One blade represents the operator's take-over capabilities, the other the automation's availability. The first blade is slowly closing when the operator's take-over capability decreases with increasing trust. As soon as the automation availability decreases temporarily, the movement of the two blades is cutting off the operator from the control of the process, e.g., of the vehicle.

The deeper mechanism of the "scissor" could be located between the operator's confidence in the system's performance capability and the operator's take-over capability, as sketched, e.g., by Manzey and Bahner (2005). Here, the relationship between the perceived/attributed automation capacities on the part of the operator and the actual capacity of the automation seems to be relevant. First, there is the ideal scenario in which the expectations with regard to the technical system's abilities are realistic and fit the actual capacities. In that case, it is to be expected that even after critical events, which will outline the system's boundaries, no adjustments to the attributions are necessary. The trust in the system is on an appropriate and constant level. However, there is a second scenario, in which the actual capacities do not meet the (too high) expectations. According to Lee and See (2004) this inadequate calibration might be traced back to an insufficient precision and specification of the judgment. In the literature, this phenomenon is discussed under the notions of overtrust/overreliance (e.g., Inagaki and Itoh 2013; Lee and See 2004) and automation bias/complacency (e.g., Bahner 2008; Mosier and Skitka 1996). While overtrust and automation bias might be understood as cognitive components that postulate the exaggerated system trust, overreliance and complacency depict the behavioral component. The latter two terms often are used complementary. They might be conceptualized by means of an insufficient or infrequent operator-sided system monitoring compared to what actually would be required with regard to the system's capacities (Bahner 2008). Only in those critical situations that evince the automation limits, an adaptation of the attributions with regard to the system abilities takes place. Subsequently, a strong loss of trust occurs, which is—according to existing literature (Hoffman et al. 2013; Lee and See 2004)—in most cases only difficult to compensate. Payre et al. (2016) affirm this role of complacency with regard to the emergence of an unsafe valley. They postulate that distinct complacency will lead to difficulties (e.g. slow reactions) in the course of take-over maneuvers between automated and non-automated driving. Here it also becomes clearer why "unsafe" is the better name for the valley, as the drivers even feel too comfortable, at least before the incident or accident.

With regard to that phenomenon, there seems to be an additional irony in the progress of automation: At the beginning of a technological development process, an automated system still has various errors that an operator can react on by means of an increased readiness for take-over. Due to the technical system's increasing availability over time, the user's first experience with a system error or failure will be postponed in time. As such, an undue confidence will build up. Under these conditions, the same error that could be compensated beforehand will have a higher impact on road safety and user expectations. Likewise, the loss of trust will be stronger, the more difficult a mistake might be compensated and the more costs it provokes (Bahner 2008) as well as the bigger it is and the more difficult it is to predict (Lee and See 2004).

Choi and Ji (2015) provide a more holistic model of trust on adopting an autonomous vehicle: Within their approach, they combined different relevant concepts (e.g., trust, perceived risk and personality traits) within one model and took it under empirical examination. They found that the concepts of trust (defined via system transparency, technical competence and situation management) and perceived usefulness have major influence on the behavioral intention to use a system. Other factors such as perceived risk and ease of use as well as locus of control and sensation seeking only have minor or no influence at all. According to these results, it might be assumed that exaggerated trust and perceived usefulness lead the naive driver into the uncanny and unsafe valley.

4 A Sketch of the Design Space: The Solution's Dimensions for Safeguarding an Unsafe Valley

Justified by the successful implementation of automated systems such as in Goodrich et al. (2006), Hoeger et al. (2012) or Altendorf et al. (2015), the assumption is that not automation or higher automation per se is unsafe, but that there are unsafe regions around safe automation designs and combination of different assistance and automation levels and transitions between levels or modes. How can the unsafe regions be safeguarded in order to utilize the safe regions? Decisive dimensions that form the design space of safeguards could be:

a. Abilities of the human and automation: The human capability might depend on a selection process e.g. in domains like aviation, or on the distributions of a general population e.g. in the driving domain. The abilities of the automation depend on their interplay with the environment, and might be structured according to normal operations, systems limits and system failure. Increasingly important will be the (meta-) ability to describe its own ability to sense and act, e. g. in case of a sensor degradation due to changes in the environment like bad weather.
b. The distribution of tasks, authority, responsibility and a minimum of autonomy for both the automation and the human, as described e.g. by Flemisch et al. (2011).
c. Combining (a) and (b), the control-distribution of the corresponding automation-level, which interact with the human's involvement, and which could be organized in clear modes (Fig. 1.5). It is an open research and engineering question as to how many and which modes are needed and/or wanted at all, but there are clear hints that to many modes can lead to mode confusion especially in time critical situations. It is especially not clear whether partially and conditionally automated levels are needed at all, even if there are hints that a well designed level of partial automation might improve the take-over ability and might also be fun. Another open research and engineering question is how many different modes can be differentiated and operated safely, and how potential migration paths could look like in order to ensure upwards and downwards compatibility.

Another dimension of the design design space are the transitions between the modes, that can be initiated either by the human (Fig. 5 red) or by the automation (Fig. 5 blue). It is an open research and development question as to how the

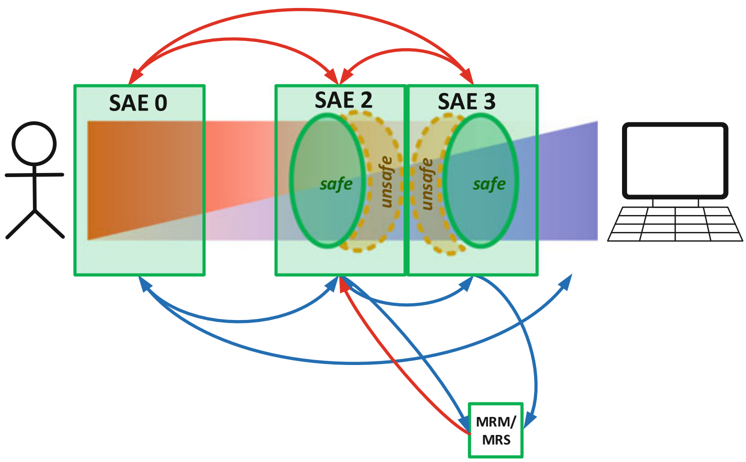


Fig. 5 Uncanny Valleys: Modi and Transitions

transitions are balanced and secured against false and failed transition. A safeguard, e.g., for a transition from right side of the automation scale, crossing the valley towards the left side of the scale, could be an interlocked transition, where the control is only handed over to the operator if it is really clear that the operator has taken over control. This interlocked transition was for cars and trucks successfully implemented and tested in the HAVEit project. This safeguard works best if the right rim is also secured with a transition to a minimum risk state, e.g., via a minimum risk maneuver, as is described in Hoeger et al. (2011).

A safeguarding measure, e.g., of the left rim of the unsafe valley can be, for example, that an insufficient involvement of the human is determined by monitoring attention and reacting accordingly e.g. through prompts, as described in e.g. Rauch et al. (2010) and in Schwalm et al. (2015) and Schwalm and Ladwig (2015).

Another safeguard on the left rim of the valley is the communication of the ability of the automation, as shown by Beller et al. (2013) based on Heesen et al. (2010) as a concept to communicate an uncertainty value of the automation.

A similar direction of safeguarding against falling into an unsafe valley is provided by different authors who promote a so-called "trust management": First, systems could already be adapted within the design phase (e.g., a system's transparency could be highlighted, see Bahner 2008). Alternatively, it could be tried to improve users' system perception. Last, system capabilities and limits should be communicated more overt (cf. Muir 1994). In this context, Payre et al. (2016) postulate that intensive practice is required, in order to bridge the uncanny valley. According to them, the outline of the system's working and its boundaries is indispensable in order to avoid safety-critical automation effects.

An additional concept that could help prevent people from falling into the unsafe valley could be nudging, the art of promoting certain behaviors through small changes in the environment, which nevertheless leave the individual to decide for herself (e.g. Thaler and Sunstein 2014). Nudging could be used for influencing humans to avoid the rims of the unsafe valley, for example through promoting more involvement (Flemisch et al. 2016b). A new concept which links nudging even stronger to self-determination is currently being developed at the Institute for Industrial Engineering and Ergonomics at RWTH Aachen University.

Putting those safeguards together, a holistic picture or concept of human-machine resilience should be derived how the human machine system acts in normal operations, reacts to disturbances that try to push the system to system limits and how it reacts to system failure. Figure 6 shows such transitions between normal operations, system limits and system failure. The upper part of Fig. 6 shows a degradation of the machine that might result in a situation with a control deficit, where the human might be somehow prepared (upper part) or not enough prepared (middle). The lower part of Fig. 6 shows a degradation of the human that might result in a complete dropout of the human, where the machine might only be able to handle the situation for a limited time (lower part). A minimum risk maneuver that hopefully results in a state with a minimum risk or maximal possible safety might be the last resort in order to keep the human-machine system safe. From a minimum

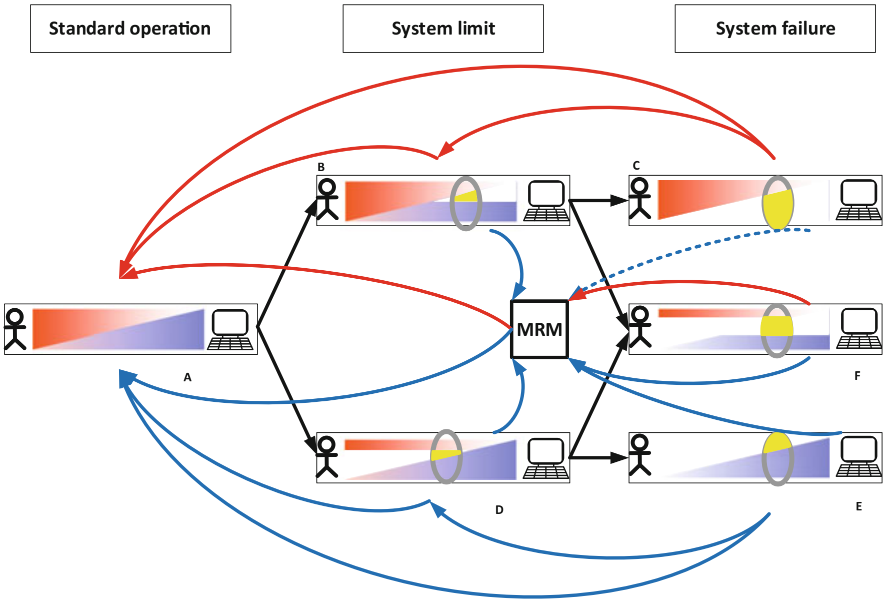


Fig. 6 Resilience through recovery or graceful degradation between normal operations, system limits or system failure

risk state the human might be able take over again to normal operations if possible. The benefit of thinking in a layered approach of normal operations, system limits and failure could be in the flexibility to stabilize the human-machine system at the limits and, instead of going into a complete, unrecoverable failure, recover or gracefully degrade from system limits and failures.

Putting those safeguards together systematically could result in a highly cooperative human-machine system, that could be inspired by natural examples of cooperation in movement, for example the H(orse)-Metaphor (Flemisch et al. 2003), it's generalization of shared and cooperative guidance and control, complemation (Schutte 1999) or cooperative automation (e.g. Flemisch et al. 2014; Bengler et al. 2014), and it's instantiation H-Mode (e.g., Goodrich et al. 2006; Altendorf et al. 2015) or Conduct-by-Wire (Winner and Hakuli 2006). The key seems to be to combine partially and highly automated automation levels in a way, that in partially automated level, the operator is being involved and prevented from losing her situation and mode awareness, e.g., with continuous haptic feedback and attention monitoring, while in highly automated, the human-machine system is so secured that it keeps safe even in the extreme case that the operator cannot come back into the loop again.

5 Outlook: Balancing Risks and Chances of Assistance and Automation by Securing the Unsafe Valley

Although the causative reasons for the emergence of such a valley might not be finally deduced yet, it becomes increasingly clear that there is at least one unsafe valley on the scale of assistance and automation, with clear applications at least in air and ground vehicle automation, and a good chance that the concept might also be valuable in other domains. It is undoubtedly necessary to secure the boundaries of an uncanny valley.

There is justified hope that in most cases, the unsafe valley can be well structured and therefore comparably well secured. We should nevertheless be aware that there might be cases for which the structure of the unsafe valley could also be more complex, more like a mountain landscape. In those cases the unsafe valley(s) should be reasonably mapped at first, before the boundaries are secured and viable bridges could be built over it.

In order to increase or maintain safety, and to harvest many chances of automation, the risks of automation have to be controlled. The systematic mapping and development of safeguarding measures down to safe combinations will require an interdisciplinary research and development, but will hopefully prevent naive and probably too trustful operators, system designers and engineers from falling into the unsafe valley, sometimes even an abyss of automation.

Acknowledgments Thanks to the initial supporters at NASA, as well as the DFG und its referees for the support in the H-Mode projects and in the research program "Kooperativ Interagierende Fahrzeuge", the EU for the support in the projects HAVEit and InteractIVe. The colleges at DLR, the TU Darmstadt, TU München and RWTH Aachen for the rich discussion.

References

Altendorf E, Baltzer M, Heesen M, Kienle M, Weißgerber T, Flemisch F (2015) H-Mode, a Haptic-multimodal interaction concept for cooperative guidance and control of partially and highly automated vehicles. In: Winner H et al (eds) Handbook of driver assistance systems. Springer, Cham

Bahner JE (2008) Übersteigertes Vertrauen in automation: Der Einfluss von Fehlererfahrungen auf complacency und automation bias. Dissertation, TU Berlin

Beller J, Heesen M, Vollrath M (2013) Improving the driver–automation interaction an approach using automation uncertainty. Hum Factors 55(6):1130–1141

Bengler K, Flemisch F (2011) Von H-Mode zur kooperativen Fahrzeugführung—Grundlegende ergonomische Fragestellungen. In: 5. Darmstädter Kolloquium: kooperativ oder autonom?. Darmstadt

Bengler K, Dietmayer K, Farber B, Maurer M, Stiller C, Winner H (2014) Three decades of driver assistance systems: review and future perspectives. IEEE Intell Transport Syst Mag 6(4):6–22

Billings CE (1997) Aviation automation: the search for a human centered approach. Lawrence Erlbaum Associates, Mahwah, NJ

Choi JK, Ji YG (2015) Investigating the importance of trust on adopting an autonomous vehicle. Int J Hum Comput Interact 31:692–702

Damböck D (2013) Automationseffekte im Fahrzeug—von der Reaktion zur Übernahme. Disseration, TU München
Donges E (2015) Fahrerverhaltensmodelle. In: Handbuch Fahrerassistenzsysteme. Springer Fachmedien, Wiesbaden, pp 17–26
Eckstein L (2016) Personal communication
Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors 37 (1):32–64
et d'Analyses BDE (2012) Final report on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF 447 Rio de Janeiro–Paris. BEA, Paris
Flemisch FO, Onken R (1998) The cognitive assistant system and its contribution to effective man/machine interaction. NATO RTO MP-3, Monterey, CA
Flemisch FO, Adams CA, Conway SR, Goodrich KH, Palmer MT, Schutte PC (2003) The H-metaphor as a guideline for vehicle automation and interaction (No. NASA/TM—2003-212672). NASA, Langley Research Center, Hampton
Flemisch F, Kelsch J, Schieben A, Schindler J (2006) Stücke des Puzzles hochautomatisiertes Fahren: H-Metapher und H-Mode; 4. Workshop Fahrerassistenzsysteme; Löwenstein
Flemisch F, Kelsch J, Löper C, Schieben A, Schindler J (2008) Automation spectrum, inner/outer compatibility and other potentially useful human factors concepts for assistance and automation. In: de Waard D, Flemisch FO, Lorenz B, Oberheid H, Brookhuis KA (eds) Human factors for assistance and automation. Shaker, Maastricht, pp 1–16
Flemisch F, Schieben A, Temme G, Rauch N, Heesen M (2009) HAVEit Public Deliverable D33.2 "Preliminary concept on optimum task repartition for HAVEit systems", Brussels
Flemisch F, Schieben A, Strauss M, Lüke S, Heyden A (2011) Design of human-machine interfaces for highly automated vehicles in the EU-project HAVEit. In: Proceedings of 14th international conference on human-computer interaction
Flemisch F, Heesen M, Hesse T, Kelsch J, Schieben A, Beller J (2012) Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations. Int J Cognit Tech Work 14(1):3–18
Flemisch F, Bengler K, Bubb H, Winner H, Bruder R (2014) Towards cooperative guidance and control of highly automated vehicles: H-mode and conduct-by-wire. Ergonomics Special Issue Beyond Human-Centred Automation 57(3). Online 24.2.2014
Flemisch F, Schwalm M, Deml B (2015a) Systemergonomie kooperativ interagierende Fahrzeuge. Projektantrag an die DFG
Flemisch F, Winner H, Bruder R, Bengler K (2015a) Cooperative guidance, control and automation. In: Winner H et al (eds) Handbook of driver assistance systems. Springer, Cham
Flemisch F, Altendorf E, Baltzer M, Rudolph C, Lopez D, VOß G, Schwalm M (2016a) Arbeiten in komplexen Mensch-Automations-Systemen: Das Unheimliche und unsichere Tal (Uncanny Valley) der Automation am Beispiel der Fahrzeugautomatisierung; 62. GfA-Frühjahrskongress "Arbeit in komplexen Systemen – Digital, vernetzt, human?!" Aachen
Flemisch F, Altendorf E, Weßel G Canpolat Y (2016b) Personal Communcation
Gasser TM, Arzt C, Ayoubi M, Bartels A, Buerkle L, Eier J, Flemisch F, Haecker D, Hesse T, Huber W, Lotz C, Maurer M, Ruth-Schumacher S, Schwarz J, Vogt W (2012a) Rechtsfolgen zunehmender Fahrzeugautomatisierung—Gemeinsamer Schlussbericht der Projektgruppe. Fahrzeugtechnik F 83, Bundesanstalt fur Straßenwesen (BASt)
Gasser TM et al (2012b) Rechtsfolgen zunehmender Fahrzeugautomatisierung - Gemeinsamer Schlussbericht der Projektgruppe Bundesanstalt für Straßenwesen (BASt), (F 83)
Goodrich K, Flemisch F, Schutte P, Williams R (2006) A design and interaction concept for aircraft with variable autonomy: application of the H-Mode. In: Digital Avionics Systems Conference, USA
Heesen M, Kelsch J, Löper C, Flemisch F (2010) Haptisch-multimodale Interaktion für hochautomatisierte, kooperative Fahrzeugführung bei Fahrstreifenwechsel-, Brems- und Ausweichmanövern; Gesamtzentrum für Verkehr Braunschweig (Hrsg.): Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel AAET, Braunschweig

Herzberger ND, Voß GMI, Schwalm M (2016) Personal communication
Hoeger R, Amditis A, Kunert M, Hoess A, Flemisch F, Krueger H-P, Bartels A, Beutner A (2008) Highly automated vehicles for intelligent transport: Haveit approach. ITS World Congress, New York
Hoeger R, Zeng H, Hoess A, Kranz T, Boverie S, Strauss M et al (2011) Final report, Deliverable D61. 1. Highly automated vehicles for intelligent transport (HAVEit). 7th Framework programme
Hoeger R, Wiethof M, Rheker T (2012) Complexity measures of traffic scenarios: psychological aspects and practical applications. In: International conference on driver behaviour and training 2011, Paris
Hoffman RR, Johnson M, Bradshaw JM, Underbrink A (2013) Trust in automation. IEEE Intell Syst 28(1):84–88
Hollnagel E (2007) Keynote zur 7. Berliner Werkstatt "Prospektive Gestaltung von Mensch-Technik-Interaktion" 2007, Berlin
Inagaki T, Itoh M (2013) Human's overtrust in and overreliance on advanced driver assistance systems: a theoretical framework. Int J Veh Technol 2013:1–8
Kuz S, Bützler J, Schlick CM (2015) Anthropomorphic design of robotic arm trajectories in assembly cells. J Occup Ergon 12(3):73–82
Lee JD, See KA (2004) Trust in automation: designing for appropriate reliance. Hum Factors 46 (1):50–80
Ma R, Kaber DB (2005) Situation awareness and workload in driving while using adaptive cruise control and a cell phone. Int J Ind Ergon 35(10):939–953
Manzey D, Bahner JE (2005) Vertrauen in Automation als Aspekt der Verlässlichkeit von Mensch-Maschine-Systemen. Beiträge zur Mensch-Maschine-Systemtechnik aus Forschung und Praxis. Festschrift für Klaus-Peter Timpe 93–109
Maurer M (2016) Personal communication
Mayer MP (2012) Entwicklung eines kognitionsergonomischen Konzepts und eines Simulationssystems für die robotergestützte Montage, Dissertation. Shaker, Aachen
Merat N, Jamson AH, Lai FC, Carsten O (2012) Highly automated driving, secondary task performance, and driver state. Hum Factors 54(5):762–771
Mori M (1970) The uncanny valley. Energy 7(4):33–35 (in Japanese)
Mori M, MacDorman KF, Kageki N (2012) The uncanny valley [from the field]. IEEE Robot Autom Mag 19(2):98–100
Mosier KL, Skitka LJ (1996) Human decision-makers and automated decision aids: made for each other? In: Parasuraman R, Mouloua M (eds) Automation and human performance: theory and applications. Lawrence Erlbaum Associates, Mahwah, NJ, pp 201–220
Muir BM (1994) Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. Ergonomics 37(11):1905–1922. doi:10.1080/00140139408964957
National Transportation Safety Board (2010) "Loss of Control on Approach, Colgan Air, Inc. Operating as Continental Connection Flight 3407 Bombardier DHC-8-400, N200WQ, Clarence Center, New York, February 12, 2009". NTSB/AAR-10/01. National Transportation Safety Board, Washington, DC
Norman DA (1990) The problem with automation. Philos Trans R Soc Lond B 327:585–593
Payre W, Cestac J, Delhomme P (2016) Fully automated driving: impact of trust and practice on manual control recovery. Hum Factors 58(2):229–241
Petermann I, Schlag B (2009) Auswirkungen der Synthese von Assistenz und Automation auf das Fahrer-Fahrzeug System. Paper presented at the 11. Braunschweiger Symposium Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel (AAET) 2011. Braunschweig
Rauch N, Kaussner A, Krueger H-P, Boverie S, Flemisch F (2010) Measures and countermeasures for impaired driver's state within highly automated driving. Transport Research Arena, Brussels

SAE J 3016:2014 Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. Society of Automotive Engineers

Schieben A, Flemisch F (2008) Who is in control? Exploration of transitions of control between driver and an eLane vehicle automation. VDI/VW Tagung Fahrer im 21. Jahrhundert 2008. Wolfsburg

Schieben A, Damböck D, Kelsch J, Rausch H, Flemisch F (2008) Haptisches feedback im spektrum von fahrerassistenz und automation. In: 3. Tagung Aktive Sicherheit durch Fahrerassistenz, Garching

Schutte PC (1999) Complemation: an alternative to automation. J Inform Tech Impact 1 (3):113–118

Schutte P, Goodrich K, Williams R (2016) Synergistic allocation of flight expertise on the flight deck (SAFEdeck): a design concept to combat mode confusion, complacency, and skill loss in the flight deck. In: Stanton NA, Landry S, Di Bucchianico G, Vallicelli A (eds) Advances in human aspects of transportation. Springer, Berlin, pp 899–911

Schwalm M, Ladwig S (2015) How do we solve demanding situations—a discussion on driver skills and abilities. In: 57th Conference of experimental psychologists 2015, Hildesheim

Schwalm M, Voß GMI, Ladwig S (2015) Inverting traditional views on human task-processing behavior by focusing on abilities instead of disabilities–a discussion on the functional situation management of drivers to solve demanding situations. In: Engineering psychology and cognitive ergonomics. Springer, Cham, pp 286–296

Sheridan TB (ed) (1976) Monitoring behavior and supervisory control. Springer, Berlin

Sheridan TB (1980) Computer control and human alienation. Technol Rev 83(1):65–73

Thaler RH, Sunstein CR, Balz JP (2014) Choice architecture. In: The behavioral foundations of public policy. Princeton University Press, Princeton, NJ

Voß GMI, Schwalm M (2015) 1. Kongress der Fachgruppe Verkehrspsychologie 2015. Braunschweig

Weßel G, Altendorf E, Flemisch F (2016) Self-induced nudging in conditionally and highly automated driving. Working Paper, Aachen

Winner H, Hakuli S (2006) Conduct-by-wire–following a new paradigm for driving into the future. In: Proceedings of FISITA world automotive congress, Oct 2006, vol 22, p 27

Man-Robot Collaboration in the Context of Industry 4.0: Approach-Avoidance Tendencies as an Indicator for the Affective Quality of Interaction?

Gerhard Rinkenauer, Adrian Böckenkamp, and Frank Weichert

Abstract

Collaborative work of man and machine in close proximity is considered as an enabling feature of the Industry 4.0 concept. Acceptance and performance are crucial aspects for the success of such working environments. To enhance acceptance it is necessary to understand the effects of robot behavior on human's behavior. The aim of the current study was to investigate the effects of robotic motion on human expressive behavior during the interaction in close distance. As a theoretical framework the concept of approach-avoidance behavior was adopted. For this reason we analyzed human motion trajectories during the interaction with an industrial robot in a real setup. Our findings suggest that an active robot affects the movement behavior of interacting participants considerably. Interestingly, an active robot seems to be more positively evaluated than a non-moving robot. In contrast to our expectations the approach-avoidance behavior of our participants suggest that human-like movement patterns of the industrial robot were evaluated less positively than robot-like mechanical behavior. In general assessing approach-avoidance behavior seems to be a promising method for an implicit measure of the affective quality of human-robot collaboration.

Keywords

Human-robot collaboration • Approach-avoidance conflict • Human behavior analysis • Industry 4.0

G. Rinkenauer (✉)
Leibniz Research Centre for Working Environment and Human Factors (IfADo), 44139 Dortmund, Germany
e-mail: rinkenauer@ifado.de

A. Böckenkamp • F. Weichert
Department of Computer Science VII, TU Dortmund University, 44227 Dortmund, Germany
e-mail: adrian.boeckenkamp@tu-dortmund.de; frank.weichert@tu-dortmund.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_24

1 Introduction

A current issue of the research field "robotics" is to solve questions about the development of safer robots in closer human–machine interaction systems at the manufacturing shop floor level. Because of the close collaboration between humans and robots many problems of robotics research are not only related to technical issues but are also framed by social aspects (Moniz 2015). Therefore, social sciences approaches are of high relevance in order to investigate human-centered aspects in collaborative environments. Such collaborative environments on the shop floor level are especially in the focus of the Industry 4.0 paradigm. However, human-robot interaction with collaborative robots (cobots) present particular challenges for interdisciplinary research. Therefore, for the current study we combined research methods from robot research, computer science and psychology to study human-robot interaction. The aim of the study was to assess how robot behavior affects human behavior and to what extent changes in human behavior may serve as an implicit measure of acceptance. Therefore, we adopted the so called approach-avoidance concept from motivation psychology as a theoretical and methodological framework for the experimental design and the analysis of human behavior.

2 Collaborative Working Environments

A collaborative working environment (CWE) supports human workers in their individual cooperative work. The future vision of CWEs is that human workers and cobots populate the shop floor. The cobots autonomously assist human workers and are accepted as a kind of mechanical colleague. This vision is especially put forward by the Industry 4.0 paradigm.

2.1 Industry 4.0

Industry 4.0 paradigm generally promotes the computerization of manufacturing and the combination of several major innovations in digital technology. These technologies include for example intelligent sensor concepts, big data analytics, artificial intelligence, the Internet of Things and advanced sensor and actor networks. Whereas such already existing technologies thought rather separately in current industrial design concepts they will be joined together within Industry 4.0 to integrate the physical and virtual worlds, the concept of Cyber-Physical (Production) Systems (Schwab 2016). The idea that an industrial robot can take the role of a cooperative and supportive tool for the workers is part of the "Industry 4.0" paradigm. However, due to safety reasons and missing research in human-robot collaboration, robot-based assistive systems are still not widely spread in the manufacturing industry. A current research challenge are for example the design of safe autonomous robots which are able to collaborate and are accepted by the human collaborator.

2.2 Human-Robot Collaboration

Until recently, the use of robots was restricted to specific industrial environments and tightly controlled tasks. Today, however, robots are employed over a wide range of sectors and a variety of different tasks. A close collaboration between humans and robots is an aspired goal and to achieve this objective robots have to become more adaptive and flexible. Advances in sensors are enabling robots to understand and respond better to their human collaborators and CWEs will enable human workers and robots to interact next and even close to each other. Collaborative working offers many advantages compared to the separate employment of man and machine. First, by combining the capability of perception and a well-marked somatosensory system of humans with the precision, resilience and availability of robots, flexibility in manufacturing related tasks is maximized. Second, the employment of robots reduces the workload of humans and facilitates ergonomic designs of CWSs. This specifically applies to decentralized working zones within the upcoming scope of Industry 4.0 (Brettel et al. 2014). Reducing the distance between a robot and a worker (their "interaction zone") is desirable as it saves space, reduces handling times and workloads for humans, i.e., the smaller the interaction zone, the more the aforementioned advantages carry weight. As a consequence the need for acceptance of the robot's proximity to an individual is enhanced. Acceptance is an essential requirement for increased productivity of cooperative working and therefore in the focus of the current study. Apart from that, safety issues become more important (Vasic and Billard 2013), albeit beyond the scope of this work.

2.3 Anthropomorphic Robots

One approach to enhance acceptance is to design robots appearance and behavior more human-like. Thus, it is suggested that anthropomorphism might be an important design principle in order to achieve a safe, effective, and efficient cooperation between humans and robots (Kuz et al. 2013). There is converging empirical evidence that human users can predict human-like robot movements better and anthropomorphic behavior also enhances the acceptance to collaborate with the robot. For example, Huber et al. (2008) showed that participants responded faster to a handover task with a robot when the robot moved with a human-like than a mechanical (trapezoid) trajectory. Furthermore, Kutz and colleagues (Kuz et al. 2013; Kuz and Schlick 2015) showed that participants performed better in the prediction of robot movement destinations when the robot performed a human-like trajectory than a constant speed trajectory. Additionally, Huges et al. (2016) evaluated the perception of robot movements in a virtual reality environment to assess what makes robot movements human-like. Different levels of anthropomorphic robot movements were generated and rated from participants by means of questionnaires. The authors identified characteristics in the robot movements which led users to attribute the movements for example either as human-like or as aggressive. Such characteristics of robot movements were essential in the participants' impressions, viz. a human-like and non-aggressive robot movement helped the users feel safer, less stressed, and more willing to work with the robot.

3 Human Movement Behavior

Understanding the effects of robots on human behavior plays a central role in this study. Therefore, we focus on the analysis of changes of human's motion behavior when interacting with an industrial robot.

3.1 Approach-Avoidance Concept

The theoretical background for our empirical study is based on the approach-avoidance concept within the context of motivation psychology. Lewin (1938) characterized the human as moving in a space replete with psychological and physical objects possessing attractive or repulsive qualities that would tend to draw the person toward or away from an object (Townsend and Busemeyer 1989). Objects could also have both positive and negative qualities at the same time, resulting in ambivalence on the part of the person. The positive and negative "charges" were called "valences." Miller (1959) postulated some criteria about approach-avoidance based on Lewins ideas. For example, it was assumed that the tendency to approach a goal as well as the tendency to avoid a feared stimulus is stronger the nearer the subject is to it. The strength of avoidance increases more rapidly with nearness than does that of approach. In other words, the gradient of avoidance is steeper than that of approach. Miller (1959) pictured the gradients as linear, although explicit formulae were not given. In Fig. 1 we depict Millers assumptions in the context of human-robot interaction.

According to that concept, trends of approach or avoidance in the motion behavior are expected depending on the affective evaluation of objects and situations by human beings (Eerland et al. 2012; Van Dantzig et al. 2008; Puca et al. 2006). Basically, it is assumed that actions which are more forceful (e.g., with a higher movement velocity) towards an object or situation or which are decreasing the distance (e.g., by leaning forward) can be rated as an approach behavior and, thus, indicate that approach forces predominates avoidance forces. In contrast, actions whose force direction points away from an object or situation or which are increasing the distance can be rated as avoidance behavior and rather indicate negative avoidance forces outweighs the positive approaching forces. This study aims at assessing to which extent approach and avoidance behavior can be observed during the interaction with a robot.

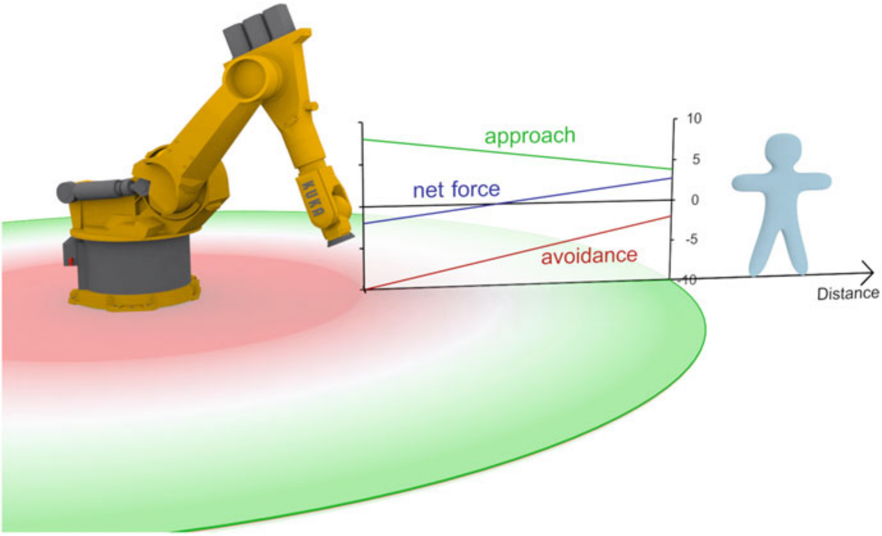


Fig. 1 The approach-avoidance concept based on an attracting force (*green graph*) and a repulsive force (*red graph*). The gradient of the repulsive force is assumed to be steeper than the one of the attracting force. As a net force field there is a range with attracting and a range with repulsive effect on behavior depending on the distance. In the transition area between the two force fields an unsystematic varying or oscillating behavior is expected

3.2 Aiming Movements and Posture

Numerous researchers have studied how affective stimuli trigger approach and avoidance reactions by using directed arm movements towards or away from negative or positive stimuli as the behavioral measure of interest. For example Chen and Bargh (1999) instructed participants to perform lever movements towards or away from them in response to positive and negative words. Rinck and Becker (2007) employed a similar paradigm but participants (with and without spider phobia) had to respond to spider pictures by means of a joystick. In those studies reaction time measures were used to infer approach and avoidance effects. Puca et al. (2006) used an extended experimental setup in which participants moved their right forearm towards or away from positive or negative words. The forearm was fixed on a moveable platform which allowed to measure distance as well as response force. In the current study we employed single aiming movements which were performed either towards or away from a robot. Aiming movements are composed of two distinct phases (e.g. Woodworth 1899; Elliott et al. 2010). Initially, there is a ballistic, preprogrammed phase that directs the limb towards the target. This initial distance-covering portion of the movement is followed by a homing phase. During this second phase, visual and proprioceptive feedback is used to reduce any divergence between the limb and the target position. In this landing phase also correcting movements are taking place to lead the hand to the target

(Meyer et al. 1988). Thus, by analyzing different phases of aiming movement it may be possible to infer about several processes of motor programming, movement performance and sensory feedback (cf. Elliott et al. 2010) which are related with movements towards (approach) or away (avoidance) from the robot. Because the full analyses (including the consideration of correcting movements) would go beyond the scope of this study, we will only report results for movement amplitudes of the performed aiming movements. In addition, analyses of posture during the performance of the aiming will be presented because there is evidence from recent studies that posture provides a relevant indicator for distance regulation and thus for approach-avoidance behavior (e.g. Eerland et al. 2012; Stins and Beek 2011).

4 Empirical Study

To assess to what extent human aiming movements are affected by robot behavior we conducted an empirical study in which participants interacted with an industrial robot. The results of the study were interpreted in the context of the approach-avoidance concept.

4.1 Experimental Design

A full factorial design with the factors ROBOT CONDITION, MOVEMENT DIRECTION and TARGET DISTANCE was employed. Each factor consists of discrete levels and the experimental conditions took on all possible combinations of these levels across the three factors. Such a design allows us to study the effect of each factor on the dependent measures (movement kinematics, posture), i.e., the behavior of the participants, as well as interaction effects between the factors. The factor ROBOT CONDITION consists of the levels “mechanical movement”, “biological movement” and “passive” (no movement). The latter level was used as baseline condition to assess participants’ behavior which is unaffected by robot movements. The factor MOVEMENT DIRECTION consists of the levels “forward movement” and “backward movement”. These levels were employed to study approach and avoidance behavior explicitly. For example, if participants lean more forward towards the robot in a certain robot condition A in comparison to a robot condition B, such a result would interpreted as an enhanced approach behavior for robot condition A. Finally, the factor TARGET DISTANCE consists of the levels short and long distances. These levels were used to assess if fine and coarse movements are influenced differently by the robot conditions, but will not be analyzed in closer detail in this study. On the basis of the anthropomorphic design concept and the theoretical considerations based on motivation psychology, we would expect, for example, that participants show a higher approach and a lower avoidance behavior in the biological than in the mechanical robot condition.

4.2 Materials and Methods

Participants Nine participants (6 females) took part in the experiment. All of them were students of educational sciences at TU Dortmund University and did not have any experience in the interaction with robots. Average age was 25.4 (SD = 3.1) and all participants were of normal or corrected to normal sight. Each of them reported to be right handed. The experimental session took about 45 min.

Experimental Setup An industrial robot of type KUKA KR-125/3 located in a safety cell was programmed to perform different kinds of pointing movements on a custom-designed 9-field pattern placed on a table in front of the robot. The industrial robot has six degrees of freedom and a repeatability of ±0.2 mm (KUKA Roboter GmbH). Participants were placed in front of the robot with a distance of about 3.3 m. There was a similar 9-field pattern in front of the participants and they were instructed to reproduce the robots pointing motions in mirror-image. For safety reasons, there was no overlap of the motion ranges of the human and the robot. The minimum distance between the robot's end-effector and the participant's hand was 1.6 m. An infrared (IR) based optical motion tracker (Optotrak Certus by Northern Digital Inc.) was used to record the participant's movements in real-time (cf. Fig. 2). The motion tracker had an accuracy of 0.1 mm and a resolution of 0.01 mm. Active IR markers were attached to the participant's right hand and shoulder. The motion tracker was placed so that all markers were always inside the detectable volume of the system. Figure 3 exemplary shows resulting trajectories for a marker attached to the hand. As already mentioned above, in addition to the different robotic motion types, a "passive robot condition"

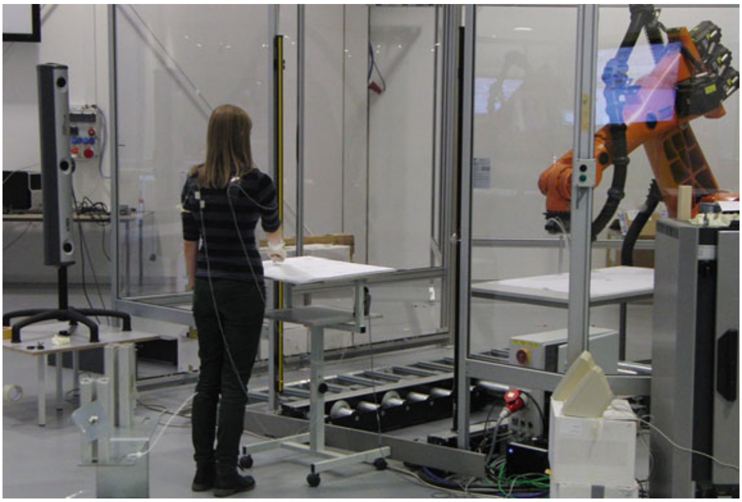


Fig. 2 Experiment setup composed of the robot (*right*), a participant, the motion tracker (*left*), and the two 9-field tables in front of the robot and participant respectively

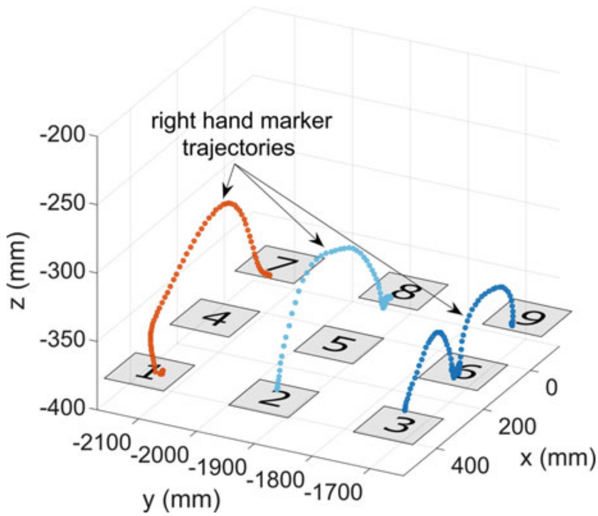


Fig. 3 Exemplary human trajectories of the right hand during movements from start to destination fields with long and short movement amplitudes

was added in which the participants were instructed by synthetic voice commands only. The voice commands have been generated by using Google's Speech Synthesis Engine (Google Inc.).

Robot Programming Motion programming in robotics typically provides three basic motion commands, namely linear (LIN), point-to- point (PTP), and circular motions (CIRC) (Brecher et al. 2013). For our experiment we used a series of PTP commands to create the robotic motions. PTP commands are typically used to move the robot's end-effector to a certain destination position because they provide the highest possible speed. More specifically, we distinguish between two types of movements: "mechanical" and "biological" movements. The latter tries to approximate the properties of human motion behavior (anthropomorphism) by using a Gaussian trajectory. However, due to technical restrictions, the velocity profile of biological movements only roughly approximated a Gaussian (cf. Böckenkamp et al. (2016) for more details about robot programming).

Procedure At the beginning of an experimental session, participants were informed about the aim of the experiment and signed an informed consent form. After instruction, each participant started with a short training session of about 5 min to familiarize with the setup. The first half of the training gave them an understanding of the 9-field table and the interaction with the robot at 65 % of the maximum speed. The second half of training demonstrated the instructions by voice commands. The actual experiment was conducted similarly. In the first half (20 min, active robot condition), the robot presented pointing movements from certain start- to endpoints and in the second half (10 min, passive robot condition), the movement destinations were given by voice commands. During both parts, participants had to perform the same aiming movements (physically showcased by the robot in the first half and given by voice commands in the second half).

Data Analyses Analyses of variance (ANOVAs) (Rutherford 2001; MacKenzie 2013) and t-tests (Dunn 2001) were conducted for the movement amplitudes of the hand movements and the shoulder position. The mean amplitude of short and long movements was 23.6 (SD = 1.2 cm) and 45.7 cm (SD = 1.9 cm) respectively. Averages for each participant and experimental condition were calculated for each hand movement. For each measure, separate ANOVAs were conducted with the factors ROBOT CONDITION (mechanical, biological, passive) × MOVEMENT DIRECTION (forward, backward) × TARGET DISTANCE (short, long). If there was no significant difference between the mechanical and biological robot conditions, the two conditions were collapsed into the condition "active", i. e., the factor levels of ROBOT CONDITION were reduced to active vs. passive. A similar ANOVA was performed with the averaged measurements of the shoulder positions of each participant. In the context of the approach-avoidance concept, interactions between the factors ROBOT CONDITION and MOVEMENT DIRECTION especially allows to interpret results in this theoretical context and we will therefore mainly focus in our analyzes on these two factors.

5 Results

We present our results beginning with the posture measures during movement performance followed by the movement amplitudes.

Posture As an indicator for changes in posture shoulder position was tracked. Figure 4 depicts the mean shoulder positions as a function of MOVEMENT DIRECTION and ROBOT CONDITION. There was a significant main effect of MOVEMENT DIRECTION, $F(1,8) = 6.4$, $p < 0.05$, viz. shoulder position was closer to the robot in forward than for backward movements (40.3 vs. 42.0 cm). The interaction between ROBOT CONDITION and MOVEMENT DIRECTION just failed to attain significance, $F(2,16) = 3.4$, $p = 0.06$, however separate ANOVAs revealed that there is a significant interaction for the passive and biological robot conditions, $F(1,8) = 7.7$, $p < 0.05$, but again just failed for passive and mechanical conditions, $F(1,8) = 4.9$, $p = 0.06$. ANOVA for biological and mechanical robot condition did neither show a main effect of ROBOT CONDION, $p > 0.2$ nor and interaction between ROBOT CONDITION and MOVEMENT DIRECTION, $p > 0.06$. Collapsing data for biological and mechanical robot conditions into the active condition revealed a significant interaction between ROBOT CONDITION and MOVEMENT DIRECTION for the active and passive robot condition, $F(1,8) = 8.9$, $p < 0.05$ but no main effects of the two factors, $\forall p > 0.05$. In sum, the patterns of results for posture suggests that in forward movements the shoulder is moved closer to the active than to the passive robot

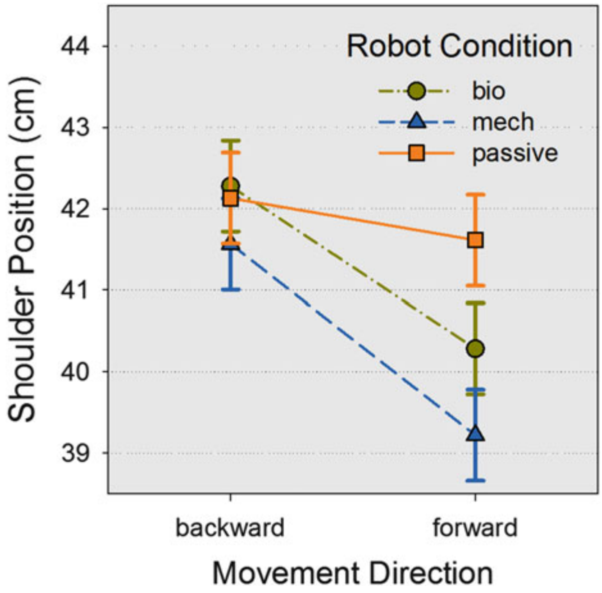


Fig. 4 Shoulder position (posture) as a function of MOVEMENT DIRECTION (backward, forward) and ROBOT CONDITION (biological, mechanical, passive)

whereas backward movements did not differ. Analyzes furthermore suggest, that shoulder position in the forward movement seems to be closer to the robot for the mechanical than for the biological in comparison to the passive robot condition.

Movement Amplitude Movement amplitude informs us if the distance traveled by the hand varies as a function of the experimental conditions. Figure 5 depicts the average movement amplitude as a function of MOVEMENT DIRECTION and ROBOT CONDITION. The ANOVAs reveal a significant main effect of ROBOT CONDITION, $F(1,8) = 4.2$, $p < 0.05$, viz. averaged movement amplitude was lowest for passive (34.3 cm), increased for biological (34.7 cm) and highest for mechanical (35.1 cm) robot behavior. Main effect of MOVEMENT DIRECTION did not attain significance, $F(1,8) = 4$, $p > 0.07$. However, ANOVA revealed a significant interaction between ROBOT CONDITION and MOVEMENT DIRECTION, $F(2,16) = 3.7$, $p < 0.05$ (Fig. 5). Separate ANOVAs showed that the this interaction is caused by the interaction for the passive and biological ROBOT CONDITION, $F(1,8) = 6.1$, $p < 0.05$. Additional paired t-Tests show that movement amplitude does not differ for passive and biological condition in forward movements ($p = 0.61$) but in the backward movements ($p = 0.02$), viz. a bigger distance was moved away from the robot in the biological condition in comparison to the passive control condition. Separate ANOVAs also revealed that there was no significant interaction of the two factors for the passive and mechanical moving robot, $p > 0.3$ but a significant main effect of ROBOT CONDITION, $F(1,8) = 8.6$, $p < 0.05$. In sum the patterns of results for movement amplitude suggest that in comparison to the passive control condition biological robot movements cause a

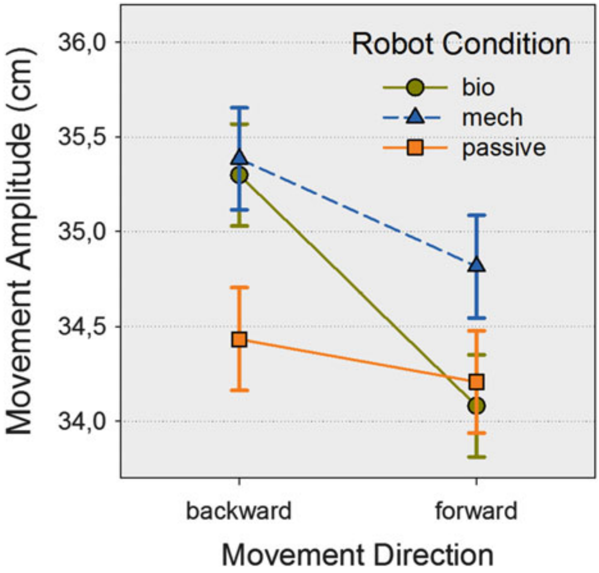


Fig. 5 Movement amplitude as a function of MOVEMENT DIRECTION (backward, forward) and ROBOT CONDITION (biological, mechanical, passive)

larger distance in backward movements whereas the mechanical condition causes a general enhancement of the movement amplitude independently from the direction of the hand movements.

6 Discussion

In the near future, robots and humans will interact rather closely. Thus, the robot's workspace will overlap with the human's personal space. The aim of this study was to asses to what extend human behavior is affected by the different movement behaviors of an industrial robot. As a theoretical framework we adopted the approach-avoidance concept from motivation psychology. According to the basic theoretical assumptions of this framework, movement patterns which are more forceful towards the robot or decrease the distance between the participant and robot are rated as approach behavior. Movements, however, whose force direction is pointing away from the robot or increase the distance between a participant and the robot, are rated as avoidance behavior. Generally, the pattern of results of this study can be interpreted as approach-avoidance behavior in response to the different robot conditions. The systematic differences in the posture measures between the active and passive robot conditions suggest that there seems to be a stronger approach behavior to the active than to the passive robot. More detailed analyzes of posture suggest that the mechanical robot conditions elicited a stronger approach behavior than the biological robot condition. In addition the findings for movement amplitude revealed larger amplitudes away from the robot for the biological robot condition in comparison to the passive condition. Such a finding may be interpreted as a stronger avoidance behavior for the biological condition. The findings for both measures (posture and movement amplitude) suggests that in our study biological movement patterns of the robot had a less positive valence than mechanical movements and this finding, however, does not agree with our expectations.

One explanation for this finding may be that the movement patterns of the biological condition did not agree with the expectations of the participants about the movement behavior of an industrial robot. A frequent explanation of the negative rating of an anthropomorphic design is the so called "uncanny valley". This concept was first introduced by Mori (1970). He put forth a thought experiment in which a robot became more human-like. It was assumed that there would first be an increase in its acceptability, however, as it approached a nearly human state there would be a dramatic decrease in acceptance. The reason for this drop in acceptance may be that at a certain degree of similarity the robot becomes rather uncanny for humans and therefore the drop in acceptance was called uncanny valley. Especially motion may intensify an uncanny situation (Pollick 2010). Human actions consist of many different sensory cues. If these cues are not mutually consistent then reconciling the differences among cues in the perception of humans might lead to a state of unease. There is empirical evidence that the recognition of affect and emotion from human movement is represented by dimensions of activity and valence (Pollick et al. 2001). The dimension of activity is supported by the speed of a movement and the valence dimension (positive or negative) appears to be related to structural relations among the body parts. The form and motion of a robot might contain different cues to human activity and these cues may depend on the movement characteristics of the robot. For example Hughes et al. (2016) suggested that the inertia on the end-effector was of importance for a robot movement to be perceived as a rather human-like or an aggressive machine-like movement. Thus, our aim to enhance the similarity to human-like movements in the biological robot condition may have failed in some way because of the non-adequate choice of movement parameters.

In sum, our findings suggest that an active robot affects the movement behavior of interacting participants. Interestingly, an active robot seems to be more positively evaluated than a passive robot. Furthermore, our attempt to equip an industrial robot with a human-like movement behavior obviously failed, because the implicit rating responses of our participants suggest that the biological behavior was rated less positively than the mechanical robot-like movement behavior. This issue will be addressed in future studies. In general, however, more research is needed to replicate, clarify, and extend our results. For example, it needs to be seen if only the movements towards or away from the robot are affected or if the movement behavior is more generally affected. Therefore, in our current research we evaluate not only for- and backwards movements but also lateral movements. Furthermore, as the theoretical concept of Miller (1959) suggests, the distance between robot and human will have a considerable impact on the approach-avoidance behavior. Therefore, the distance will be systematically varied in future studies. Additionally, we will take into account individual differences analogous to the research of Puca and colleagues (2006). The consideration of the personal dimension may be useful to better adapt human-robot interaction to the individual needs and preferences.

In conclusion, the Industry 4.0 paradigm suggests that in the near future, robots and humans will work hand in hand and for this collaborative interaction it is mandatory that the robot's workspace will overlap with the human's personal space. Approach and avoidance behavior may be a useful implicit measure of the affective quality of human- robot interaction (e.g. in addition to questionnaires) by either assessing whether human workers rate the robot's interaction or the social relation between them rather positive or negative. Furthermore, this method in general may be relevant for the study of social robotics. Analogously to concepts of social psychology (Kennedy et al. 2009) it may be possible to estimate to what extent spatial violations of the personal space of the human collaborators affect their behavior during the interaction with robots.

References

Böckenkamp A, Weichert F, Rinkenauer G (2016). Investigating the effects of robotic motion on worker's behavior in cooperative working environments. In: RO-MAN 2016 - The 25th IEEE international symposium on robot and human interactive communication. IEEE, New York, August 2016, pp 670–675

Brecher C, Müller S, Kuz S, Lohse W (2013) Towards anthropomorphic movements for industrial robots. In: Digital human modeling and applications in health, safety, ergonomics, and risk management. Human body modeling and ergonomics. Springer, Berlin, pp 10–19

Brettel M, Friederichsen N, Keller M, Rosenberg M (2014) How virtualization, decentralization and network building change the manufacturing landscape: an industry 4.0 perspective. Int J Sci Eng Tech 8(1):37–44

Chen M, Bargh JA (1999) Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. Pers Soc Psychol Bull 25(2):215–224

Dunn D (2001) Statistics and data analysis for the behavioral sciences. McGraw Hill Higher Education, New York

Eerland A, Guadalupe TM, Franken IHA, Zwaan RA (2012) Posture as index for approach-avoidance behavior. PLoS One 7(2):e31291. doi:10.1371/journal.pone.0031291

Elliott D, Hansen S, Grierson LE, Lyons J, Bennett SJ, Hayes SJ (2010) Goal-directed aiming: two components but multiple processes. Psychol Bull 136:1023–1044. doi:10.1037/a0020958

Google Inc. Text-to-Speech (TTS): speech synthesis engine. https://translate.google.com/translatetts?ie = UTF-8&q = query text. Accessed 16 Feb 2016

Huber M, Lenz C, Rickert M., Knoll A, Brandt T, Glassauer S (2008) Human preferences in industrial human-robot interactions. In: Proceedings of the international workshop on cognition for technical systems, Munich, Germany

Hugues O, Weistroffer V, Paljic A, Fuchs P, Larim AA, Gaudin T, Buendia A (2016) Determining the important subjective criteria in the perception of human-like robot movements using virtual reality. Int J Human Robot 13(2)

Kennedy D, Gläscher J, Tyszka J, Adolphs R (2009) Personal space regulation by the human amygdala. Nat Neurosci 12(10):1226–1227

KUKA Roboter GmbH. Technical Data: KR 125/3. http://www.gspnexus.com/schedetecniche/kr125 3.pdf. Accessed 16 Feb 2016

Kuz S, Schlick C (2015) Anthropomorphic motion control for safe and efficient human-robot cooperation in assembly system. In: Lindgaard G, Moore D (eds) Proceedings of the 19th triennial congress of the International Ergonomics Association (IEA), Aug 2015

Kuz S, Mayer M, Müller S, Schlick C (2013) Using anthropomorphism to improve the human-machine interaction in industrial environments (Part I). In: Duffy VG (ed) Digital human modeling and applications in health, safety, ergonomics, and risk management. Human body modeling and ergonomics. Springer, Berlin, pp 76–85

Lewin K (1938) The conceptual representation and measurement of psychological forces. Contributions to psychological theory. Duke University

MacKenzie I (2013) Human-computer interaction: an empirical research perspective. Morgan Kaufmann, Amsterdam

Meyer DE, Abrams RA, Kornblum S, Wright CE, Smith JEK (1988) Optimality in human motor performance: ideal control of rapid aimed movements. Psychol Rev 95:340–370. doi:10.1037/0033-295X.95.3.340

Miller NE (1959) Liberalization of basic S-R concepts: extensions to conflict behavior, motivation and social learning. In: Koch S (ed) Psychology: a study of a science, Study 1, vol 2. McGraw-Hill, New York, pp 196–292

Moniz A (2015) Intuitive interaction between humans and robots in work functions at industrial environments: the role of social robotics. In: Vincent J, Taipale S, Sapio B, Lugano G, Fortunati L (eds) Social robots from a human perspective. Springer, Heidelberg, pp 67–76

Mori M (1970) Bukimi No Tani (the Uncanny Valley). Energy 7:33–35

Northern Digital Inc. Optotrak Certus: overview. http://www.ndigital.com/msci/products/optotrak-certus/. Accessed 16 Feb 2016

Pollick FE (2010) In search of the uncanny valley. In: Daras P, Mayora O (eds) User centric media. Springer, New York, pp 69–78

Pollick F, Paterson H, Bruderlin A, Sanford A (2001) Perceiving affect from arm movements. Cognition 82:B51–B61

Puca R, Rinkenauer G, Breidenstein C (2006) Individual differences in approach and avoidance movements: how the avoidance motive influences response force. J Pers 74(4):979–1014

Rinck M, Becker ES (2007) Approach and avoidance in fear of spiders. J Behav Ther Exp Psychiatr 38:105–120

Rutherford A (2001) Introducing ANOVA and ANCOVA: a GLM approach. SAGE, London

Schwab K (2016) The fourth industrial revolution. World Economic Forum, Cologny, Geneva

Stins JF, Beek PJ (2011) Organization of voluntary stepping in response to emotion inducing pictures. Gait Posture 34(2):164–168

Townsend JT, Busemeyer JR (1989) Approach-avoidance: return to dynamic decision behavior. In: Izawa C (ed) Current issues in cognitive processes: the Tulane Flowerree symposium on cognition. Erlbaum, Hillsdale, NJ, pp 107–133

Van Dantzig S, Pecher D, Zwaan RA (2008) Approach and avoidance as action effects. Q J Exp Psychol 61(9):1298–1306

Vasic M, Billard A (2013) Safety issues in human-robot interactions. In: IEEE International Conference on Robotics and Automation (ICRA), May 2013, pp 197–204

Woodworth RS (1899) The accuracy of voluntary movement. Psychol Rev 3:1–119

Joint Angle Depending Representation of Maximum Forces in Digital Human Models: Investigating Multivariate Joint-Torque Polynomials for Elbow Flexion and Elbow Extension

André Kaiser, Michael Spitzhirn, and Angelika C. Bullinger

Abstract

Joint angle-dependent calculations of human action forces constitute a significant advantage of digital human models, but further knowledge of maximum allowable joint torques has to be acquired. This paper describes a method to calculate polynomials (functions) for maximum torque of elbow flexion and extension that depend on various joints of elbow and shoulder to increase the quality of prediction.

First, the anatomical influences are described to explain which joint angles affect the determination of elbow flexion torques and elbow extension torques. Following, the used experimental design with 948 measurements and the procedure to conduct the polynomials is shown in detail. Here, different ways to determine polynomials are investigated. The results are polynomials for elbow flexion and extension torques which depend on the elbow-flexion angle (EF), the elbow-rotation angle (ER), the shoulder-flexion angle (SF), the shoulder-abduction angle (SA) and the shoulder-rotation angle (SR). The polynomials explain 81.5 % of the variance of the measured values for flexion and 76.5 % of the variance for extension, respectively. Overall, the developed model gives an accurate prediction for maximum flexion and extension torques in regard to the described joint angles. It can therefore help to improve the quality of ergonomic planning using digital human models.

Keywords

Digital human models • Human joint angle • Maximum isometric forces • Maximum isometric torque • Shoulder • Elbow

A. Kaiser (✉) • M. Spitzhirn • A.C. Bullinger
Professur Arbeitswissenschaft und Innovationsmanagement, TU Chemnitz, Chemnitz, Germany
e-mail: andre.kaiser@mb.tu-chemnitz.de

C.M. Schlick et al. (eds.), *Advances in Ergonomic Design of Systems, Products and Processes*, DOI 10.1007/978-3-662-53305-5_25

1 Introduction

Force and load analyses are important functions in digital human models (DHM) (Chaffin 2005; Mühlstedt and Spanner-Ulmer 2009). Five approaches are widely used to determine human forces depending on available information (Rohmert et al. 1994; Günzkofer 2015). Figure 1 shows five different approaches to calculate the maximum force level.

Firstly, task-related or task-specific forces provides a means to calculate forces. Here, forces are computed for specific tasks such as turn the crank depending on task-related influencing factors (DIN 33411-3 1986). Secondly, forces can also be related to body postures. For that purpose, a specific general posture (such as standing upright) and the force directions are needed (DIN 33411-5 1999; Wakula et al. 2009). The third approach calculates the maximum force in relation to the posture, the force direction and the range. The result can be for example 'Isodyne', stated in DIN 33411-4 (1987). The fourth approach enables the calculation of maximum force related to the joint angles (Schaefer et al. 1997; Schaefer et al. 2000; Bubb et al. 2006). The fifth method, the muscle approach, is the most sophistic approach to calculate maximum force, but requires deep and detailed information about the muscle structure and muscle recruitment. Under this consideration the angle-related approach seems to be an adequate method to achieve precise data under reasonable effort. According to Engstler (2012) joint angle related calculations of human forces offer significant advantage compared to classical load evaluation procedure and can be integrated in DHM easily. That is because the available information about the joint angles and the force direction in regard to different human positions and anthropometrics are available for each timeframe. Using this approach could allow to design workplaces more accurate for certain people with different anthropometry.

Using the angle-related approach (cp. Sect. 2.1) knowledge about the maximum torques within the main movement directions (e. g. flexion/extension, supination/ pronation) in regard to joint angles is required. Studies that describe the torque of the elbow flexion and extension in regard to all joint angles of elbow and shoulder could not be found.

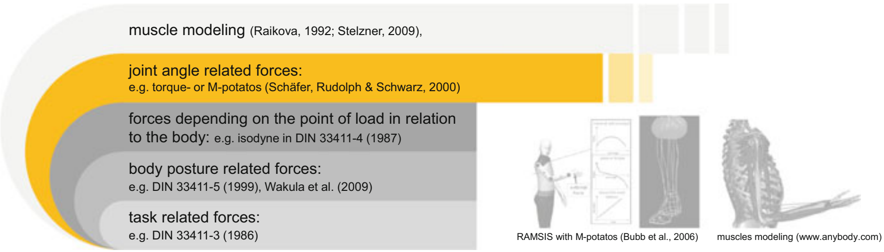


Fig. 1 Different approaches to compute body forces depending on available information according to Günzkofer (2015) and Rohmert et al. (1994)

Hence, in this article a method to calculate the maximum torque of elbow flexion and extension in regard to all of these joint angles is described. Here, multivariate polynomials are calculated by using the independent variables of the degrees of freedom of the shoulder and elbow. The angle of the shoulder abduction and upper arm rotation are also included to increase the accuracy of the prediction model.

The article consists of five main sections. Firstly, information about existing ‘joint depending torques’ models with their limitations as well as basic principles to functional anatomy of muscles and the calculation of joint angles are presented. Subsequently, the experimental design and measurement procedure as well as the procedure to conduct the polynomial regressions are described to state the used data and presumptions for the determination of the polynomial regressions. After this the polynomial functions are presented and discussed in detail. The article finished with a summary and a perspective.

2 State of Art

2.1 Ellipse and Ellipsoids for Prediction of Torques

To calculate forces without complex muscle modeling the approach of “Torque or M-potatoes” by Schaefer et al. (2000) can also be used. Figure 2 shows the approach using the example of shoulder and elbow torques. In this model the subsystem elbow has two degrees of freedom with the four main directions flexion/extension and supination/pronation. The possible maximum torques produced by the muscles between these main directions are predicted by using an ellipse with different quadrant semi-axes. These semi-axes present mathematical functions of the maximum torque in the main directions. The shoulder consists of three degrees of freedom with the six main directions (flexion/extension; adduction/abduction; internal rotation/external rotation). The maximum forces depending on the main directions are displayed by an ellipsoid. The shapes of the ellipse and ellipsoid, also

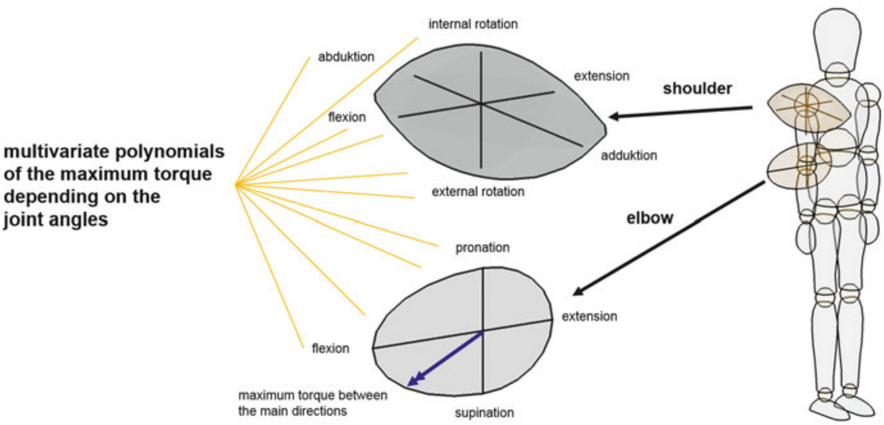


Fig. 2 M-potatoes of the shoulder and the elbow

known as 'M-potatoes', change in relation to the received posture (Günzkofer 2013). Günzkofer (2013) validated this approach for the elbow.

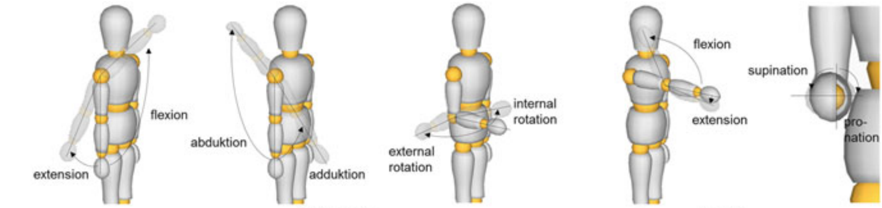
Core of the approach are multivariate polynomials that determine the joint torques of the main directions depending on various joint angles. Günzkofer (2013) computed the flexion and extension force of the forearm by using a polynomial formula which depends on the elbow flexion angle (EF), the elbow angle of rotation (ER) and the shoulder flexion angle (SF). Here, 20 test persons were involved in an experiment and 126 measurements were conducted for each person (72 measurements per flexion, 54 measurements per extension). The abduction of the shoulder or the rotation of the upper arm were not considered in the experiment; notwithstanding an influence is stated in the literature (Kapanji 2001; Schünke et al. 2007).

2.2 Elbow Strength Curves

Elbow strength curves for flexion and extension are investigated in different studies. Some experiments focus on the difference between eccentric, concentric and isometric contractions (Doss and Karpovich 1965; Sato and Sakai 1968; Singh and Karpovich 1968). Other focus on differences between isokinetic and isometric contractions (Osternig et al. 1977). Many studies show that isometric strength curves depend on EF (Elkins et al. 1951; Williams and Stutzman 1959; Doss and Karpovich 1965; Sato and Sakai 1968; Singh and Karpovich 1968; Osternig et al. 1977; Petrofsky and Phillips 1980; Knapik et al. 1983). The EF depending curve is generally accepted as ascending-descending curve with a maximum around 90° EF, depending on how EF is described. Others display the differences of the EF depending curves in different ER positions (Downer 1953; Provins and Salter 1955; Rasch 1955; Jorgensen and Bankov 1971; Kulig et al. 1984). Kapanji (2001) states that elbow flexion torques are higher in a pronated than in a supinated position. However, there are no strength curves displayed. Other authors emphasize that EF depending flexion-curves are higher in neutral and supinated ER position than in pronated position (Downer 1953; Rasch 1955; Jorgensen and Bankov 1971). Kulig et al. (1984) and Günzkofer (2013) summarize most of these studies and show that only a few studies observe the influence of the shoulder-flexion-angle SF (Winters and Kleweno 1993; Davidson and Rice 2010; Günzkofer et al. 2011; Günzkofer 2013). Günzkofer (2013) find out that EF, ER and SF have an influence on flexion- and extension-torques. Their results show that the effect of SF is influenced by EF for both flexion and extension. It is also demonstrated that supinated and neutral ER produce more flexion-torque than a pronated ER. No study was found that observed the elbow flexion and extension with respect to all elbow and shoulder angles (EF, ER, SF, SA, and SR).

2.3 Functional Anatomy of Shoulder and Elbow

The shoulder girdle consists of three joints and two accessory joints. The central joint Art. humeri (Humeroskapularjoint) is a ball and socket type of synovial joint



	shoulder						elbow			
	extension	flexion	abduktion	adduktion	external rotation	internal rotation	flexion	extension	supination	pronation
AAOS (1965)	60°	180°	180°		90°	70°	150°		80°	80°
AMA (1993)	50°	180°	180°			90°	140°		80°	80°
Kapandji (2001)	50°	180°	180°	45°	80°	110°	145°	0° - 10°	85°	90°
Schünke et al. (2007)	40°	170°	180°	40°	90°	70°	150°	0° - 10°	90°	90°

Fig. 3 Range of motion of shoulder and elbow in relation to AAOS (1965), AMA (1993), Kapanji (2001) and Schünke et al. (2007)

formed by the articulation of the proximal head of the humerus and the glenoid cavity of the scapula. The shoulder is capable of motion in a variety of axes and layers (Amell 2004; Schünke et al. 2007).

The elbow is a hinge type of synovial joint formed by the articulation of the distal humerus and the proximal ends of the ulna and the radius. The distal radioulnar joints are formed from the distal aspects of the radius and ulna. The elbow joint is capable of flexion and extension. On the other side, the radioulnar joint is capable of the supination and pronation of the forearm (Amell 2004).

The separation of the movements and the torque to special joints is not simple, therefore the human body can only be approximated in digital human models and biomechanics (Stelzner 2009). For that reason, Mühlstedt (2012) recommends the usage of physiological functional units. Subsequently, the unit elbow consists of the radioulnar joint and the elbow joint. The unit shoulder includes the whole shoulder girdle. Figure 3 shows the range of motion of both units in relation to AAOS (1965), AMA (1993), Kapanji (2001) and Schünke et al. (2007).

Ten different muscles are needed to flex or extend the forearm. Following, muscles with low influence on the movement are presented in italics. Underlined muscles pass by at least two joints. The following muscles are included in flexion: m. biceps, m. brachialis, m. brachioradialis, *mm. extensores carpi radialis longus* and *brevis*, *m. pronatores teres*, *m. flexor d. superficial*, *m. palmaris longus* and for extension: m. triceps and m. anconeus (Schünke et al. 2007).

The forearm muscles (flexor and extensor that are underlined and written in italics) extend from the humerus towards their approaches in hand and fingers (Schünke et al. 2007). This means that a change in hand position also leads to changes in pre-extension and thus changes forces of the forearm muscles. Because of their low influence on the respective movements, the hand position as independent variable is not considered.

Biceps and triceps, which extend from the scapula to the ulna or radius, have a major influence on the movements of the forearm (Schünke et al. 2007). As a result, all angles of the shoulder and the elbow change the pre-extension of both muscles. Therefore, they need to be considered as independent variables.

2.4 Calculation of Joint Angles and Definition of Main Directions of Movements

In our approach the shoulder and elbow are considered in the force prediction model. Due to different calculation methods can lead to different joint positions, which has influence on the polynomials, the mathematical approach to calculate joint angles has to be determined.

In this study, the coordinates of points are computed by three successive rotations to the spatial axes (x, y, and z). Each rotation around one of the spatial axes is done by multiplication with a rotation matrix (Stelzner 2009). Two approaches, the Bryant-angles (also gimbal angles) and Euler-angles, which differ in the order of the rotations around each spatial axis, can be distinguished. Bryant-angles use the order x, y′, z″ while Euler-angles use the sequence z, x′, z″. Matrix multiplications are not commutative, so that different rotation sequences lead to varying positions. Both Bryant- and Euler-angles are used in robotics and in DHM (Stelzner 2009; IfM 2016). It should be noted that there are singularities at different points in both calculation-methods. Among other things that is a reason why the rotation sequence is discussed in literature (Senk and Cheze 2006; Hill et al. 2008; Lempereur et al. 2014). Lempereur et al. (2014) come to the conclusion that a task-specific sequence results in increased precision and less problems with the singularities.

As a result it needs to be defined which sequence should be used as well as the neutral position of the body parts. As neutral position of the arm, the downwards directed stretched arm, with the thumb to the front and the flexion-axis of the elbow directing along the transverse axis is defined. The left side of Fig. 4 illustrates the neutral position. Figure 4 also shows the rotation sequence for the elbow and the shoulder movement. The elbow with its two degrees of freedom allows a rotation first around the xe-axis (elbow-flexion-angle EF) and then around the ze’-axis (elbow-rotation-angle ER). If this order were vice versa, the rotation of the ze-axis would also rotate the flexion-axis (xe-axis) of the elbow, which would have the same effect as a rotation of the upper arm.

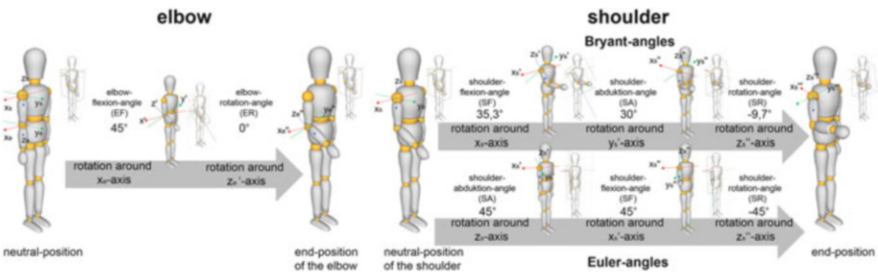


Fig. 4 Neutral position and rotation sequences

The right side of Fig. 4 shows the difference between a Bryant-angle sequence and an Euler-angle sequence for the shoulder. Both sequences end in the same position but have different rotation angles. Figure 4 also shows which angles are used to describe the shoulder-flexion-angle (SF), the shoulder-abduction-angle (SA) and the shoulder-rotation-angle (SR). While every rotation of the Bryant-angles nearly matches the real movement of a flexion, an abduction or a rotation, the final position is easier assigned by the Euler-angles. For this reason, Euler-angles are used to calculate the joint position.

3 Methods

3.1 Experimental Design

The authors use the 'arbeX' test stand developed by their research team and tested in many studies (Mühlstedt and Spanner-Ulmer 2011; Mühlstedt 2012; Kaiser et al. 2016) to measure the maximum forces. 'ArbeX' consists of a seat and a force generator. With the force generator, based on the 'lat pulldown machine' Body-Track Hi-Professional BT530, it is possible to measure the effective torque (Mühlstedt 2012). When the wire rope hoist is locked, maximum forces can be measured. A line-force dynamometer and an inclinometer as well as a goniometer measure the force in the rope and the EF angles. The position of ER, SF, SA and SR are mechanically arrested (Kaiser et al. 2015). Figure 5 shows the 'arbeX' with seat and power generator, the compensation of the gravitational force and the mechanical arrest of the SF angle. The noraxon myoReasearch software and sensors are used.

3.2 Measurement Procedure

Three methods are mainly used to measure maximum isometric forces (Kroemer 1977). Using the ramp method (method 1) the subject needs to increase their force until exhaustion. The momentum method (method 2) instructs the test person to apply the maximum force instantly. Using the plateau method (method 3) the subject needs to increase their force within 1 s and hold it for another 4 s. Within

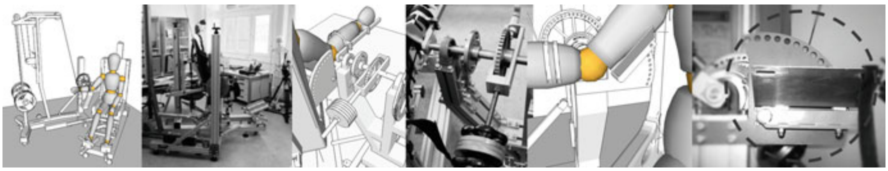


Fig. 5 ArbeX test stand (*from left to right*): test stand with seat and power generator (3D Model, real), gravitational force compensation (3D Model, real), mechanical arresting of the SF angle (3D Model, real)

these 4 s an average value over a 3 s interval is defined as the maximum force (Caldwell et al. 1974; Chaffin 1975; Kroemer 1977; Kumar 2004). However, different intervals are common in literature. Wakula et al. (2009) used a 1.5 s period to require maximum forces and Rohmert et al. (1994) used a 2 s interval. This study uses the plateau method of 4.0 s. Within these 4 s an average value over a 1.5 s interval is defined as the maximum force. This procedure is used because of the scientific importance of the studies of Wakula et al. (2009).

Gallgher et al. (2004) describe the test-retest variability (TRV) of the plateau method with 10 %. To fulfil the requirement, they recommend that subjects perform two repetitions. If the two measurement data are not within the 10 % range, an additional trial has to be carried out until the two highest values are within 10 %. Stobbe and Plummer (1984) require an average of 2.43 iterations per test parameter. Therefore, three repetitions per test variable are subsequently defined for this experiment.

A rest period is defined as 2 min between each repetitions (Gallgher et al. 2004; Wakula et al. 2009). Stobbe (1982) shows that such a rest period is adequate to prevent localized muscle fatigue (Engstler 2012). The maximum number of repetitions per muscle group is often not mentioned (Kroemer 1977; Rohmert et al. 1994; Gallgher et al. 2004; Wakula et al. 2009). Therefore, a pretest with the subject helps to define the number of repetition. After three series of 20 repetitions in one position, it is found that after the 18th repetition the subject cannot achieve a value within the 10 % TRV. 12 repetitions per day and per muscle group is therefore conservatively defined.

Nine measurement points for the EF (5 positions for flexion, 4 positions for extension) are defined to determine the polynomials for the flexion and extension. In each EF position three positions of the ER, three positions of the SF and three positions of the SA were measured. In result, this leads to 243 measurement points and additional 81 measurement points in modified SR. This results in 324 test points, which have to be repeated three times. In sum 972 measurements have to be conducted to determine the multivariate polynomials for the maximum torques of elbow extension (432 measurements) and the elbow flexion (540 measurements).

The number of measurement points increases the resolution and possible degree of the polynomial. It also ensures a more accurate prediction of existing forces. Figure 6 shows the measurement points.

Kaiser et al. (2015) describes subject characteristics such as age, gender, motivation, fitness or favorite body side and further important factors, which can influence the test results. Because of this high measuring effort only one test person is measured in detail. The test subject is a 30 year old male, active in sports, but not a weight sportsman or an endurance athlete. The right side of the subject is measured. As a control instance his stress is subjectively measured with the Borg-scale CR-10 (Borg 1982, 1998). He was not externally motivated.

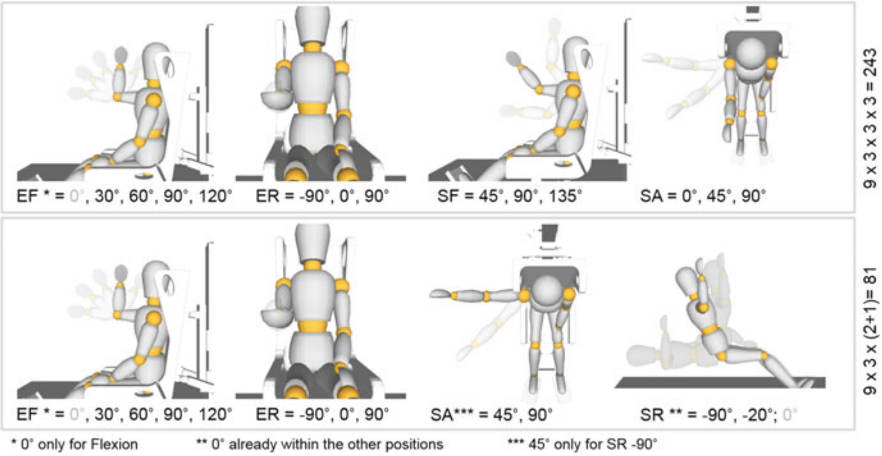


Fig. 6 Overview of the measurement points

3.3 Procedure for Polynomial Regression

Müller (2006) describes a procedure to conduct a polynomial regression which is used subsequently. The first step contains in a data correction. After this, the polynomial regression has to be done, followed by a residual analysis.

There are different ways to determine the maximum forces from the three repetitions of each measurement point. Gallgher et al. (2004) according to Caldwell et al. (1974) and Chaffin (1975) suggest that the two highest iterations, which are within 10 % TRV, should be used (method 1). Rohmert et al. (1994) use the average of all three repetitions (method 2). Further scopes are the use of the average of the two highest values within 10 % TRV (method 3) or the use of all repetitions which are arranged within 10 % TRV (method 4). Figure 7 illustrates the four approaches with the chosen measuring repetitions. P1 to P4 are separate measuring points; each with three iterations.

For the polynomial regression different strategies can be used to compute the regression. Possible ways are the forward selection, where additional variables are added; or the backward elimination, where single variables are removed. The results are assessed with the mentioned quality criteria (SD, R^2, adj. R^2, pred. R^2) and verified by a graphical analysis (Müller 2006).

The backward elimination of variables is used. Here, the polynomial with the highest possible degree is defined. Engstler (2012) for example uses for the calculation of the flexion and extension formula a square function for EF and SF. Günzkofer (2013) deploys a square exponent for EF and ER and a linear component for SF. In general, strength curves are divided into ascending, descending and ascending-descending curves (Kulig et al. 1984). Hence polynomials of first or second degree are integrated. Cubed or higher degree can also be used if the curves have only one local maximum within the range of motion.

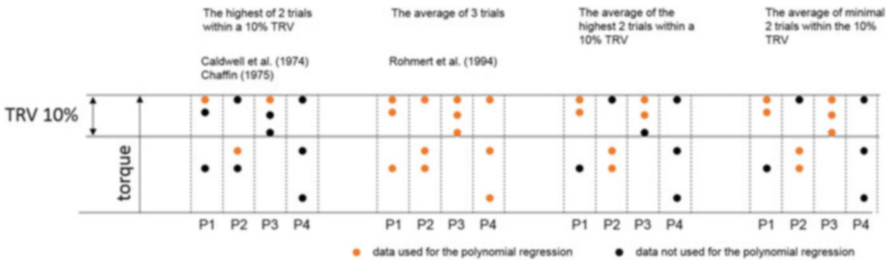


Fig. 7 Four approaches to determine the maximum force from three repetitions

Because of the measured points for ER, SF, SA and SR only square functions are allowed. For EF a cubic function term can be deployed because only one local maximum within the range of motion is stated by the graphical analysis.

4 Results

4.1 Measurement Results

948 measurements (516 for elbow flexion and 432 for elbow extension) are done. 24 measurement points could not be measured due to sense of pain. A TRV between 4.70 % (flexion repetition two and three) and 7.56 % (extension repetition one and three) is obtained. Figure 8 shows the aggregated results of the measurements.

Each boxplot contains all values with the shown angle. For example the first boxplot on the left-hand side of the elbow flexion includes all values with EF matching 0°. Because no measuring point with EF equal to 0° is measured for the elbow extension, there is no boxplot on the left-hand side for the extension. As it is seen in the graphical overview, the EF has a major influence on the torque values. An angle of 135° in SF seems to result in lower torque values, especially in elbow flexion, compared to other SF angles. These aggregated boxplots do not show the influence between the different angles. Therefore, a polynomial regression is conducted.

The method for the data correction is selected using a first polynomial regression (cp. Fig. 9 using F1 and E1). Here, the following four methods are compared due to the quality criteria: the standard deviation (SD), the coefficient of determination (R^2), the adjusted R^2 (adj. R^2) as well as the predicted R^2 (pred. R^2).

Method 1 (Caldwell et al. 1974), achieves the highest result for R^2 for extension torque (flexion: SD 5.954, R^2 0.817, adj. R^2 0.791, pred. R^2 0; extension: SD 7.424, R^2 0.776, adj. R^2 0.737, pred. R^2 0.686). Method 2 (Rohmert et al. 1994) leads to lower standard deviation (SD) compared to the first method, but has higher SD than method 3 (flexion: SD 5.712, R^2 0.812, adj. R^2 0.804, pred. R^2 0.796; extension: SD 7.313, R^2 0.747, adj. R^2 0.734, pred. R^2 0.68). With method 3, using the average of the highest two repetitions within a 10 % TRV, the best R^2 for the flexion torque can be gained (flexion: SD 5.664, R^2 0.822, adj. R^2 0.806, pred. R^2 0.797; extension: SD

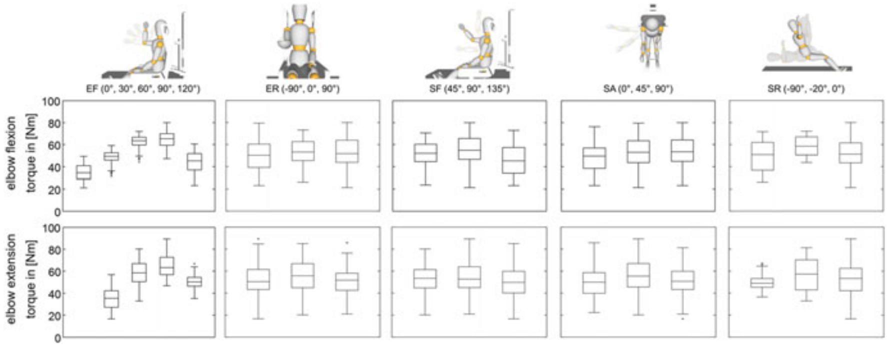


Fig. 8 Boxplots of the measured values

nr.	polynomials	SD	R²	adj. R²	pred. R²	elimi-nate	F	P
F1	A B C D E AB AC AD AE BC BD BE CD CE DE AA BB CC DD EE AAA	5.59	0.816	0.808	0.805	CE	0	1
F2	A B C D E AB AC AD AE BC BD BE CD DE AA BB CC DD EE AAA	5.584	0.816	0.808	0.797	BC	0.055	0.815
F3	A B C D E AB AC AD AE BD BE CD DE AA BB CC DD EE AAA	5.578	0.816	0.809	0.798	E	0.132	0.717
F4	A B C D AB AC AD AE BD BE CD DE AA BB CC DD EE AAA	0.573	0.816	0.809	0.799	EE	0.046	0.83
F5	A B C D AB AC AD AE BD BE CD DE AA BB CC DD AAA	5.567	0.816	0.81	0.8	AC	0.295	0.587
F6	A B C D AB AD AE BD BE CD DE AA BB CC DD AAA	5.563	0.816	0.81	0.801	B	0.729	0.394
F7	A C D AB AD AE BD BE CD DE AA BB CC DD AAA	5.562	0.816	0.81	0.802	BD	1.145	0.285
F8	A C D AB AD AE BE CD DE AA BB CC DD AAA	5.563	0.815	0.81	0.802	BE	1.065	0.303
F9	A C D AB AD AE CD DE AA BB CC DD AAA	5.563	0.815	0.81	0.803	-	-	-
E1	A B C D E AB AC AD AE BC BD BE CD CE DE AA BB CC DD EE AAA	6.885	0.765	0.752	0.739	CE	0	1
E2	A B C D E AB AC AD AE BC BD BE CD DE AA BB CC DD EE AAA	6.876	0.765	0.752	0.739	AC	0.011	0.917
E3	A B C D E AB AD AE BC BD BE CD DE AA BB CC DD EE AAA	6.867	0.765	0.753	0.74	CD	0.042	0.838
E4	A B C D E AB AD AE BC BD BE DE AA BB CC DD EE AAA	6.858	0.765	0.753	0.741	AD	1.902	0.169
E5	A B C D E AB AE BC BD BE DE AA BB CC DD EE AAA	6.866	0.763	0.753	0.74	AAA	2.167	0.142
E6	A B C D E AB AE BC BD BE DE AA BB CC DD EE	6.877	0.762	0.752	0.74	-	-	-

Fig. 9 Backward elimination of the polynomial variables

7.09, R^2 0.768, adj. R^2 0.75, pred. R^2 0.734). Using method 4, where the average of all values is within a 10 % TRV, the lowest SD, the highest adj. R^2 and the highest pred. R^2 for flexion and extension can be created (flexion: SD 5.59, R^2 0.816, adj. R^2 0.808, pred. R^2 0.805; extension: SD 6.885, R^2 0.765, adj. R^2 0.752, pred. R^2 0.739). Subsequently, method 4 is used for the data correction in the model.

In using this approach, 15 maximum values and 15 minimum values from 432 measurements are excluded from the iteration for the extension movement. No measuring point is completely deleted. For the flexion movement, 5 maximum and 16 minimum values from 516 measurements are eliminated. No measuring variables are deleted.

4.2 Derived Polynomial Regression

Figure 9 shows the different polynomials with the used quality criteria and the backward eliminated variables. The polynomials are shown in a shorted way. The 'A' stands for the variables EF, 'B' for the variable ER, 'C' for the variable SF, 'D' for the variable SA and 'E' for the variable SR. In conclusion the expression 'AA' stands for EF times 2. All shown polynomials achieve for the F-Test an overall p value lower than 0.05.

The numbers 'F1' to 'F9' show the backwards elimination process for the elbow flexion polynomials and 'E1' to 'E6' for extension polynomials. The variables with the highest p value respectively lowest F-value are eliminated and therefore are not part of the polynomial in the next row. If no variables are eliminated, no variable will have a p value higher than 0.05. Because the polynomials are used for maximum torque prediction, the pred. R^2 has to be maximized and the SD has to be minimized.

In result, Fig. 9 shows the selected polynomials in an orange labeling. For the prediction of the torque value the formula 'F9' has the second lowest SD and the second highest pred. R^2. In addition each variable has a significant effect ($p < 0.05$) in 'F9'. E4 for extension has the highest pred. R^2 with the lowest SD, but one variable (AD) has a p-value greater than 0.05. An elimination of AD would result in a lower pred. R^2 and a higher SD. That is why F9 and E4 are chosen to predict the maximum torque of flexion and extension. In contrast to F9 and E4, polynomials without SA and SR describe only 69.5 % of the variance of the measured values for flexion (SD = 7.094 [Nm], $R^2 = 0.695$, adj. $R^2 = 0.691$, pred. $R^2 = 0.684$) and 61.2 % of the variance of the measured values for extension (SD = 8.689 [Nm], $R^2 = 0.612$, adj. $R^2 = 0.604$, pred. $R^2 = 0.595$). As a result the torque of the elbow flexion (TEF) is calculated in [Nm] with the variables EF, ER, SF, SA and SR in [°] by using the following formula:

$$\begin{aligned}\mathrm{TEF}(\mathrm{EF},\mathrm{ER},\mathrm{SF},\mathrm{SA},\mathrm{SR}) = {} & 0.21 \times EF + 0.443 \times SF - 0.086 \times SA \\ & + 1.635 \times 10^{-4} \times EF \times ER + 1.103 \times 10^{-3} \\ & \times EF \times SA + 2.13 \times 10^{-3} \times EF \times SR \\ & + 1.866 \times 10^{-3} \times SF \times SA - 8.932 \times 10^{-4} \\ & \times SA \times SR + 7.812 \times 10^{-3} \times EF^2 - 1.652 \\ & \times 10^{-4} \times ER^2 - 3.18 \times 10^{-3} \times SF^2 - 1.403 \\ & \times 10^{-3} \times SA^2 - 7.141 \times 10^{-5} \times EF^3 \end{aligned} \tag{F9}$$

The torque of the elbow extension (TEE) in [Nm] is calculated by the next formula:

$$\begin{aligned}\mathrm{TEE}(\mathrm{EF},\mathrm{ER},\mathrm{SF},\mathrm{SA},\mathrm{SR}) = {} & 2.262 \times EF - 0.073 \times ER + 0.426 \times SF \\ & + 0.361 \times SA - 0.377 \times SR + 3.67 \times 10^{-4} \\ & \times EF \times ER + 4.282 \times 10^{-4} \times EF \times SA \\ & + 2.642 \times 10^{-3} \times EF \times SR + 4.702 \times 10^{-4} \\ & \times ER \times SF - 3.059 \times 10^{-4} \times ER \times SR \\ & - 5.045 \times 10^{-4} \times ER \times SR - 1.302 \times 10^{-3} \\ & \times SA \times SR - 0.018 \times EF^2 - 4.439 \times 10^{-4} \\ & \times ER^2 - 2.412 \times 10^{-3} \times SF^2 - 4.412 \\ & \times 10^{-3} \times SA^2 - 3.959 \times 10^{-3} \times SR^2 + 3.195 \\ & \times 10^{-5} \times EF^3 \end{aligned} \tag{E4}$$

Figure 10 shows the 3D-plots for F9 and E4. The black points represent the measured values and the orange surface points visualize the polynomials. Within the grey highlighted area the polynomials represent the interpolated values; while outside the polynomials are extrapolating. Here, the extrapolation leads to problems. An example can be seen for elbow extension in Fig. 10. When EF is lower than 30° a value below 0 Nm can be predicted for TEE.

Figure 11 shows results of the residual analysis. The boxplots on the right side give an overview for the residuals for elbow flexion and extension. The plots next to the left show the residuals in regard to the EF angle. Especially for elbow flexion the range of the residuals is higher for EF nearly 0° and 120°. Figure 7 illustrates that the range of the measurement values around 0° and 120° is not much higher than for

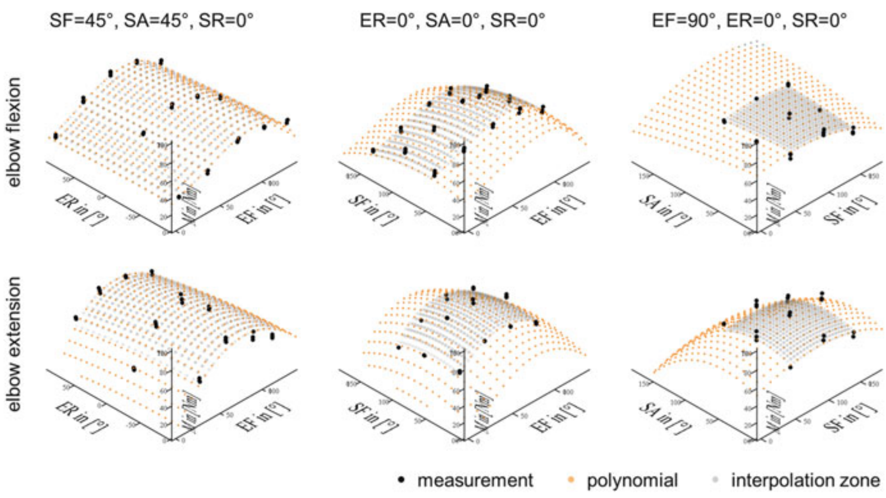


Fig. 10 Graphical representation of F9 and E4

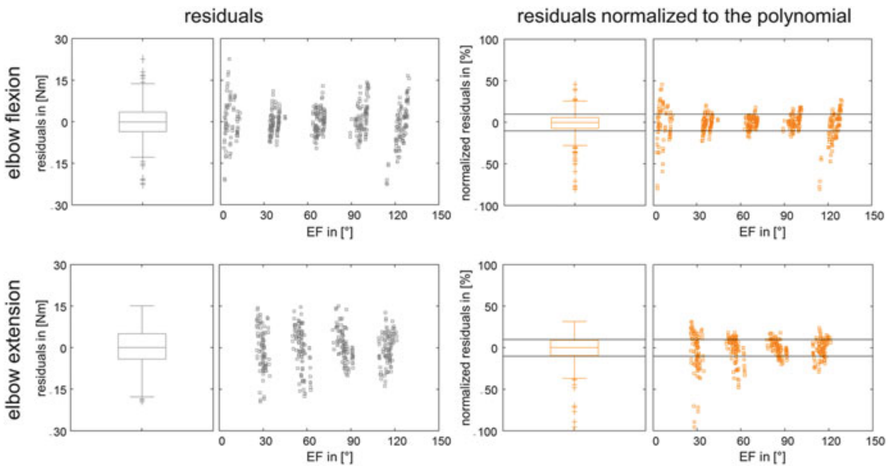


Fig. 11 Residuals and normalized residuals as boxplots and values depending on EF

other EFs. But it has to be considered, that the values for EF 0° and 120° are lower. Therefore, the residuals where normalized to the polynomials.

The orange boxplots give an overview. Over 50 % of all values are within the 10 % TRV according to their polynomial values. However, there are values with more than 75 % differences. The plots next to the right show more details. For flexion the normalized range is high for EF around 0° and around 120°. For Extension the normalized range is higher for EF 30°.

5 Discussion

The residual analysis demonstrates that the polynomials computation achieves adequate results for the torques prediction. Over 50 % of all measuring values are within a 10 % TRV. According to Gallgher et al. (2004), this would be a normal variability. The other values spread until more than 75 % difference. This accuracy of the polynomials is influenced by the used joint angles. The residual analysis shows a higher deviations for joint angle combinations with low torques.

The influence of angle calculation has to be considered. Other methods such as ‘Bryant angles’ lead to results which do not represent the same positions as the measurement data. However, all discussed methods, even the approach of Euler-angles, generate singularities. These can lead to different torques according to the polynomials which can result in discontinuities. For the defined rotation sequence, singularities exist for positions with SF of 0° respectively for SF of 180°. These positions are not within the measured range of motion and have therefore to be extrapolated.

As already described the extrapolation leads to probably inaccurate results. There are different ways to solve such problems. One way consist in the use of splines which may be increase the accuracy within the range of motion, but do not allow extrapolation. Another possibility is the adaption of artificial zero points to fix the problem with zero-crossing within the range of motion. The last solution are new measurements to extend the range. Until such a solution is defined it is recommended to use the polynomials within the measured joint angles.

The used way for the data correction is selected due to the first polynomial regression. A backward elimination is not conducted for each of the four methods. Therefore, it can be possible that other methods could lead to better results which has to be tested as a next step.

All used polynomial regressions show an overall p value lower than 0.05. Using the backward elimination leads to better results in the used quality criteria. In result, the conducted formula differ from other polynomial functions described by Engstler (2012) or Günzkofer (2013). In contrast to Engstler (2012), but similar to Günzkofer (2013), ER^2 has a significant effect for elbow flexion [$F(1,492) = 6.248$, $p = 0.013$]. Compared to Günzkofer (2013) and Engstler (2012) the polynomial for elbow flexion shows no significant interaction between EF and SF, but a significance for EF and SA is stated [$F(1,492) = 35.02$, $p = 0.000$]. The polynomial for the elbow extension shows as against Engstler (2012) and Günzkofer (2013) that ER has also a significant effect [$F(1,401) = 14{,}814$, $p = 0.000$]. The influence of SA and SR on the force prediction accuracy are proved by comparing the polynomial regressions with and without both variables SA and SR. The polynomials with SA and SR describe the variance of the measured values for flexion and for extension more precise and have a higher prediction accuracy of the polynomials (cp. Sect. 3.3). At this stage, it has to be mentioned that the difference between the various polynomials could also be a result of the used test subjects. In this experiment only one person is observed in comparison with 10 in Engstler (2012) and 20 in Günzkofer (2013). The characteristics of the test subjects can influence the result significantly such as stated in Kaiser et al. (2015).

Overall, the computed polynomials show that the integration of the shoulder angles SA and SR increase the observed accuracy of the torque prediction. It can be stated that all observed angles influence the maximum torque significant and therefore have to be integrated in the polynomial regression to achieve a high accuracy.

6 Summary and Perspective

The authors have described a method to calculate functions (polynomials) for maximum torque of elbow flexion and extension that depend on the joints of elbow and shoulder to increase the quality of prediction. In consideration that all shoulder angles might have influence on the outcome of the measurements, the polynomials of Günzkofer (2013) and Engstler (2012), are extended. Thus, the Research Target of increasing the accuracy of torque prediction has been reached.

For that purpose, a literature review was conducted to determine the joint angle-dependent characteristic of the maximum torque based on functional anatomy of the shoulder and the elbow as well as maximum force strength curves. In addition, a review of different measuring methods for isometric maximum forces had to be done to define the experimental design. The measured ranges of motion were determined by EF 0° until 120°, ER from −90° to 90°, SF from 45° to 135°, SA from 0° to 90° and SR from −90° to 0°.

For computation of the polynomials 948 measurements for one test subject were conducted at the arbeX test stand. Here, different ways to determine the maximum torques values as well as to compute polynomials are discussed.

In result, final polynomials show that all angles of the elbow and the shoulder have significance influence on the predicted values of the flexion and extension torques of elbow. The functions describe 81.5 % of the variance of the measured values for the maximum torque of the elbow flexion respectively 76.5 % for the maximum torque of the elbow extension. The aim to improve the accuracy of torques models could be achieved due to the integration of should angles, because without SA and SR the functions describe only 69.5 % respectively 61.2 % of the variance of the measured values.

To validate the M-Potatoes for the upper arm, the defined polynomials are only the first step. As next step polynomials for all main movement directions of the shoulder as well as the supination and pronation of the elbow have to be computed. The calculation of different percentiles is a main task to extent the described approach to a broader area of application.

References

AAOS (1965) American Academy of Orthopaedic Sugeons. Joint motion: method of measuring and recording, 4th print. E. & S. Livingstone, Edinburgh

AMA (1993) American Medical Association guides to evaluation of permanent impairment, 4th edn. American Medical Association, Chicago

Amell T (2004) Shoulder elbow and forearm strength. In: Kumar S (ed) Muscle strength. CRC Press LLC, Boca Raton, FL, pp 227–247. ISBN 0-415-36953

Borg G (1982) A category scale with ratio properties for intermodal and interindividual comparisons. In: Geissler HG, Pethold P (eds) Physchophysical judgement and the process of perception. VEB Deutscher Verlag der Wissenschaft, Berlin, pp 25–34

Borg G (1998) Borg's perceived exertion and pain scales. Human Kinetics, Champaign, IL

Bubb H, Engstler F, Fritzsche F, Mergl C, Sabbah O, Schaefer P, Zacher I (2006) The development of RAMSIS in past and future as an example for the co-operation between industry and university. Int J Hum Factors Model Simulat. doi:10.1504/IJHFMS.2006.011686

Caldwell LS, Chaffin DB, Dukes-Dobos MD, Kroemer KHE, Laubach LL, Snook SH (1974) A proposed standard procedure for static muscle strength testing. Am Ind Hyg Assoc J 35 (4):201–206

Chaffin DB (1975) Ergonomics guide for the assessment of human static strength. Am Ind Hyg Assoc J 36(7):505–511

Chaffin DB (2005) Improving digital human modelling for proactive ergonomics in design. Ergonomics 48(5):478–491. doi:10.1080/0002889738506892

Davidson AW, Rice CL (2010) Effect of the shoulder angle on the activation pattern of the elbow extensors during submaximal isometric fatiguing contraction. Muscle Nerve 42(4):514–521. doi:10.1002/mus.21717
DIN 33411-3 (1986) Körperkräfte des Menschen. Maximal erreichbare statische Aktionsmomente männlicher Arbeitspersonen an Handrädern. Beuth Verlag GmbH, Berlin
DIN 33411-4 (1987) Körperkräfte des Menschen. Maximale statische Aktionskräfte, Werte. Beuth Verlag GmbH, Berlin
DIN 33411-5 (1999) Körperkräfte des Menschen. Maximale statische Aktionskräfte (Isodynen). Beuth Verlag GmbH, Berlin
Doss WS, Karpovich PV (1965) A comparision of concentric, eccentric and isometric strength of elbow flexors. J Appl Physiol 20(2):351–353
Downer A (1953) Strength of the elbow flexor muscles. Phys Ther Rev 33(2):68–70
Elkins E, Leden U, Wakim KG (1951) Objective recording of the strength of normal muscles. Arch Phys Med Rehabil 32(10):639–647
Engstler F (2012) Perzentilierung maximaler Gelenkmomente des Menschen. Verlag Dr. Hut, München
Gallgher S, Moore JS, Stobbe TJ (2004) Isometric, isoinertial, and psychophysical strength testing: devices and protocols. In: Kumar S (ed) Muscle strength. CRC Press LLC, Boca Raton, FL, pp 129–156. ISBN 0-415-36953
Günzkofer F (2013) Elbow strength modelling for digital human models. Verlag Dr. Hut, München
Günzkofer F (2015) Kräfte. In: Bubb H, Bengler K, Grünen RE, Vollrath M (eds) Automobilergonomie. Springer, Wiesbaden. ISBN 978-3-8348-1890-4
Günzkofer F, Engstler F, Bubb H, Bengler K (2011) Isometric elbow flexion and extension joint torque measurements considering biomechanical aspects. In: First international symposium on digital human modeling, Lyon, France
Hill A, Bull AMJ, Wallace AL, Johnson GR (2008) Qualitative and quantitative description of glenohumeral motion. Gait Posture 27:177–188
IfM (2016) Dynamicus 6.0. Modellierung und Simulation von mehrkörperdynamischen Menschmodellen. Referenz- und Handbuch. Institut für Mechatronik e. V. www.ifm-chemnitz.de
Jorgensen K, Bankov S (1971) Maximum strength of elbow flexors with pronated and supinated forearm. In: Jokl E (ed) Biomechanics II. Karger, Basel, pp 174–180
Kaiser A, Wimmer P, Bullinger AC (2015) Gelenkwinkelabhängige Untersuchung von Maximalkraft und Dauerleistung zur Definition ergonomisch effizienter Arbeitsbereiche im digitalen Menschmodell. In: GfA e.V. (Hrsg.), VerANTWORTung für die Arbeit der Zukunft, 61. Kongress der GfA. GfA-Press, Dortmund
Kaiser A, Spitzhirn M, Dittrich F, Bullinger AC (2016) Gelenkwinkelabhängige Kraftmodellierung für den Einsatz in digitalen Menschmodellen—Untersuchung der unanhängigen Variablen zur Erzeugung multivariater Gelenkmomentpolynome am Beispiel der Ellenbogen Flexion und Extension. In: GfA e. V. (Hrsg.) Arbeit in komplexen Systemen. Digital, vernetzt, human? 62. Kongress der GfA, GfA-Press, Dortmund
Kapanji IA (2001) Funktionelle Anatomie der Gelenke. Schematisierte und kommentierte Zeichnungen zur menschlichen Biomechanik, 3rd edn. Hippokrates, Stuttgart. ISBN 3-7773-1941-4
Knapik JJ, Wright JE, Mawdsley RH, Braun JM (1983) Isometric, isotonic, and isokinetic torque variations in four muscle groups through a range of joint motion. Phys Ther 63(6):938–947
Kroemer KHE (1977) Die Messung der Muskelstärke des Menschen—Methoden und Techniken Hrsg: Bundesanstalt für Arbeitsschutz und Unfallforschung. Forschungsbericht Nr. 161. Dortmund. Wirtschaftsverlag NW GmbH, Bremerhavem
Kulig K, Andrews JG, Hay JG (1984) Human strength curves. In: Terjung RL (ed) Exercise and sport science reviews, vol 12. The Collamore Press, Lexington, MA, pp 417–466
Kumar S (2004) Muscle strength. CRC Press, Boca Raton, FL
Lempereur M, Brochard S, Mao L, Remy-Neris O (2014) Validity and reliability of 3D marker based scapular motion analysis: a systematic review. J Biomech 47(10):2219–2230

Mühlsted J, Spanner-Ulmer B (2011) Modeling muscular strain induced by movements for enhancing digital human models. Human factors in organisational design and management

Mühlstedt J (2012) Entwicklung eines Modells dynamisch-muskulärer Arbeitsbeanspruchung auf Basis digitaler Menschmodelle. Universitäts verlag, Chemnitz

Mühlstedt J, Spanner-Ulmer B (2009) Homo Sapiens Digitalis: Über den Praxiseinsatz digitaler Menschmodelle

Müller M (2006) Angewandte statistische Regression. Züricher Hochschule Winterthur. https://stat.ethz.ch/education/semesters/WS_2006_07/asr/Stat.pdf. Zugegriffen: 15 July 2016

Osternig LR, Bates BT, James ST (1977) Isokinetic and isometric torque force relationship. Arch Phys Med Rehabil 58(6):254–257

Petrofsky JS, Phillips CA (1980) The effect of elbow angle on the isometric strength and endurance of the elbow flexors in men and woman. J Hum Ergol 9(2):125–131

Provins KA, Salter N (1955) Maximum torque exerted about the elbow joint. J Appl Physiol 7 (4):393–398

Rasch PJ (1955) Effect of the position of forearm on strength of elbow flexion. Res Q 27:333–337

Rohmert W, Berg K, Bruder R, Schaub KH (1994) Kräfteatlas—Teil 1 Datenauswertung statischer Aktionskräfte. In: BAuA (ed) Schriftenreihe der Bundesanstalt für Arbeitsmedizin—Forschung Fb 09.004. Wirtschaftsverlag NW, Bremerhafen. ISBN 3-89429-524-4

Sato M, Sakai Y (1968) Isometric and eccentric strength of elbow flexors and extensors. J Anthropol Soc Nippon 76:183–190

Schaefer P, Rudolph H, Schwarz W (1997) Variable force limits for optional target populations. In: Proceedings of the 13th IEA conference, Tampere, pp 533–535

Schaefer P, Rudolph H, Schwarz W (2000) Digital man models and physical strength—A new approach in strength simulation. Proceeding of the SAE DHMC (Digital Human Modeling Conference), SAE 2000-01-2168

Schünke M, Schulte E, Schumacher U (2007) Prometheus—LernAtlas der Anatomie—Allgemeine Anatomie und Bewegungssystem. 2. Auflage

Senk M, Cheze L (2006) Rotation sequence as an important factor in shoulder kinematics. Clin Biomech 21:S3–S8

Singh M, Karpovich PV (1968) Strength of forearm flexors and extensors in men and women. J Appl Physiol 25:177–180

Stelzner G (2009) Zur Modellierung und Simulation biomechanischer Mehrkörpersysteme. Schriftenreihe des Instituts für Technische Mechanik Band 11. Universität Karlsruhe. Universitätsverlag Karlsruhe

Stobbe T (1982) The development of a practical strength testing program for industry. Unpublished dissertation

Stobbe TJ, Plummer RW (1984) A test-retest criterion for isometric strength testing. In: Human Factors and Ergonomics Society annual meeting proceedings 28

Wakula J, Berg K, Schaub KH, Bruder R (2009) Der montagespezifische Kraftatlas (BGIA-Report, 3/2009). BGIA; Technische Informations-u. Universitätsbibliothek, Sankt Augustin, Hannover

Williams M, Stutzman L (1959) Strength variation through the range of motion. Phys Ther 39:145–152

Winters JM, Kleweno DG (1993) Effect of initial upper-limb alignment on muscle contributions to isometric strength curves. J Biomech 26(2):143–153