Applied Ergonomics 59 (2017) 170–181

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Human-centered design (HCD) of a fault-finding application for mobile devices and its impact on the reduction of time in fault diagnosis in the manufacturing industry

Annette Kluge [a,*], Anatoli Termer [b]

[a] Industrial, Organizational and Business Psychology, Ruhr University Bochum, Germany
[b] Cognitive and Media Science, Ruhr University Bochum, Germany

ARTICLE INFO

Article history:
Received 28 January 2016
Received in revised form
26 April 2016
Accepted 28 August 2016

Keywords:
Mobile device
Troubleshooting
Manufacturing

ABSTRACT

The present article describes the design process of a fault-finding application for mobile devices, which was built to support workers' performance by guiding them through a systematic strategy to stay focused during a fault-finding process. In collaboration with a project partner in the manufacturing industry, a fault diagnosis application was conceptualized based on a human-centered design approach (ISO 9241-210:2010). A field study with 42 maintenance workers was conducted for the purpose of evaluating the performance enhancement of fault finding in three different scenarios as well as for assessing the workers' acceptance of the technology. Workers using the mobile device application were twice as fast at fault finding as the control group without the application and perceived the application as very useful. The results indicate a vast potential of the mobile application for fault diagnosis in contemporary manufacturing systems.

1. Troubleshooting and fault finding: past and present

The reliability of machinery and technical process cells is one of the main goals in the manufacturing industry (Ortegon et al., 2014). The goals for the future of industrial production are to further enable the progress of the production process along the value chain and to lower costs (Jasperneite and Niggemann, 2012). One central concept is the computerization and digitalization of information technology.

In most cases, the user requirements of such technologies cannot be defined in detail from the very beginning of the development process. Malfunctions occurring during production can be of very high complexity due to increased connectivity and cross-linking and the coupling of different components of the production line (Kluge, 2014; Krems, 1997; Rouse, 1979). The result of higher levels of complexity is an increasing demand in terms of the maintenance and servicing of machinery and plant due to a higher set of all possible malfunctions of the production system. Higher coupling across components of a process (Hermawati et al., 2015) leads to more complex disturbance chains following malfunctions (Woods and Hollnagel, 2006). As interruptions in production lead to higher costs, it is important to keep production downtime as low as possible (Jasperneite and Niggemann, 2012; Hermawati et al., 2015).

Due to the aforementioned reasons, an improvement in the support of maintenance workers is required in order to reduce complexity, improve performance and reduce diagnosis time (Jasperneite and Niggemann, 2012). These goals raise questions with regard to design and guidelines for the usage of new information technologies. Whenever a fault appears or a fault message occurs in the display, maintenance workers are subject to time pressure as they are responsible for getting the system running again as soon as possible. The occurred faults are influenced by the complexity of the task and thus lead to a higher mental workload for the maintenance workers due to the high level of complexity of contemporary production plants (Leung et al., 2010).

The goal of the study presented here is to explore how maintenance workers can be supported while performing their tasks, including fault diagnosis and the repair of the faulty component, through the use of information technology.

Early research on fault finding Fault finding is described as a type of human problem-solving behavior, with the attempt to identify

* Corresponding author.
E-mail addresses: annette.kluge@rub.de (A. Kluge), anatoli.termer@gmail.com (A. Termer).

http://dx.doi.org/10.1016/j.apergo.2016.08.030
0003-6870/

errors in a system and repair or replace the faulty components in order to restore the system to normal functioning. It is associated with the repair of physical, mechanical or electronic systems (Jonassen and Hung, 2006), and is a cognitive task, as it involves searching for and isolating faulty components, with the fault-finder having to search or scan the system before diagnosing why the system is malfunctioning (Patrick, 1993; Sauer et al., 2013). Human abilities and possible limitations in problem solving are well documented in the literature (Henneman and Rouse, 1984; Rouse, 1978; Rouse and Rouse, 1979). For instance, it has been found that in problem-solving situations, unreliable information affects fault-finding performance (Sauer et al., 2013), with higher intermittency rates leading to longer diagnosis time and more actions taken, such as repeated tests of components and outputs (Teague and Allen, 1997).

In their early, but nevertheless still valid, review and evaluation in this field of study, Morris and Rouse (1985) stated that “human troubleshooting performance has been found to degrade as systems increase in size or complexity” (p. 526), and further identified the skills that are necessary for successful troubleshooting (Chavaillaz et al., 2016). Relevant for the present study are (1) the skill to perform functional tests, and (2) the skill to employ a strategy in addressing the problem. Such strategies include narrowing down the location of a fault; theoretical understanding of the functioning; and transfer of knowledge which is required for the application or adaptation of existing fault-finding strategies and for the development of new ones (Patrick, 1993). This is achieved, for example, through the observation of an abnormal situation, identification of symptoms, search for possible causes, a gradual narrowing down of possible causes and the initiation of corrective action (Jonassen and Hung, 2006; Kluge, 2014). Finally, the degree of fault-finding performance depends on the use of procedural knowledge (diagnostic rules) and principle knowledge (theoretical knowledge), especially in fault-finding situations with high interconnectivity. Thus, it is necessary to design diagnostic strategies which combine the two types of knowledge into a unified framework (Ham and Yoon, 2007; Linou and Kontogiannis, 2004).

For the development of an effective support system, it is crucial to gain an understanding of how relationships between relevant variables (syntactic cues; experience-driven associations) are perceived and used by fault-finding experts to enhance their fault-finding performance via performance-supporting systems (Lindgaard, 1995; Wilson, 2000).

In order to ensure that the troubleshooter will employ an appropriate strategy, it is necessary to proceduralize the testing sequence and to provide guidance (Morris and Rouse, 1985; Patrick et al., 1989). For the proceduralization of the testing sequence, training is important, as improvements in problem-solving accuracy may only be achieved through extended practice of sequences and in particular through the availability of information about the quality of moves during problem solving (Duncan, 1985). If training time is limited, the use and application of a fault-finding strategy can be enhanced by traditional and dynamic job aids (Kluge et al., 2013; Kluge, Greve, Borisov & Weyers, in press). However, the advantage of job aids depends on the quality of the guidance provided, which in turn depends on the identification of the relevant information required for successful fault finding.

In summary, in previous research, the emphasis was placed on the development and retention of fault-finding skills via training (Gramopadhye et al., 1997; Henneman and Rouse, 1984; Hochholdinger and Schaper, 2013; Linou and Kontogiannis, 2004; Patrick et al., 1996; Rouse, 1978) and on the cognitive process of fault-finding (Morris and Rouse, 1985; Patrick et al., 1989).

Current developments and technological advancements for technology-supported fault finding

Despite the vast technological developments in the past few decades, so far, little research has been devoted to the question of how technology can best be employed to support guidance using diagnostics technologies in order to detect and correct failures in the future. While the vast majority of the research was conducted in the 1980s (Rouse and Rouse, 1979; Henneman and Rouse, 1984; Duncan, 1985; Patrick et al., 1989), in the early years of the 21st century, few studies have been carried out in the research field of fault finding supported by technology. This dearth of studies can be attributed, among other things, to the fact that expert systems have lacked user acceptance for several reasons, such as an underlying deficiency in the cognitive coupling between the human and machine subsystems. Human users often have to adopt the machine's solution, which can lead to errors in decision making, especially under uncertainty and risk (Fitter and Sime, 1980; Langlotz and Shortliffe, 1983). As the organization of human factors and machine components determines performance, technology in fault diagnosis should aid the user in the process of decision making, rather than making decisions or recommending solutions (Woods, 1985). It is unlikely that new computer systems will find acceptance if those who carry the cost in terms of workload do not receive benefits from the new technology (Woods and Hollnagel, 2006).

Driven by technological change, mobile media usage behavior is evolving rapidly. This fast technological change opens the door to new, enriching opportunities as media usage habits can be adapted to the context of work. Wendemuth et al. (2014) argue that an infiltration of different techniques in industrial diagnosis, such as mobile devices (tablet-PCs) and modalities like gesture, haptics or sign language, is to be expected in the future.

How can industrial diagnosis quality be enhanced accordingly? A possible strategy (solution) for achieving such an objective is the employment of information systems (Jasperneite and Niggemann, 2012), which should be optimized for mobile use and serve both as cognitive relief and as an information technology that is relevant to the task (Wendemuth et al., 2014).

The development of efficient information systems for mobile use requires the involvement of the future users in order to match the fault finders' needs. Therefore, the existing knowledge concerning technology, workflow processes and production has to be ascertained and then formalized for the purpose of achieving a systematic approach towards the knowledge structure of the diagnostic process (Ham and Yoon, 2007). Subsequently, the appropriate diagnostic knowledge has to be prepared in the most adequate and expert manner in order to be accessible for the work staff. Limited human processing ability leads to the necessity of a human-centered approach, as it is vital to incorporate the cognitive strategies and ideas of the personnel into the development process (Kluwe, 1997). In order to design effective information systems, it is necessary to identify the decision-making/problem-solving requirements and to improve cognitive performance (Woods, 1985).

The approach used in the study presented here was to employ the methodology of a human-centered design process (ISO 9241-210:2010). Once the need for human-centered design (HCD) has been identified, the process is characterized by four activities:

1. User analysis: Who will use the product, what will they use it for, and under what conditions will they use it?
2. Task analysis: Business requirements, user goals that must be met for the product to be successful.
3. Prototyping: Create a design solution
4. Evaluation: Usability testing with actual users

The conception of mobile applications in the manufacturing industry has to take into account Human Factors as one integral part of the design process (Wendemuth and Biundo, 2012). The cognitive demands of the production environment need to be addressed by software developers. In order to be accepted by the employees, a new information system must be able to enhance the motor, cognitive and associative skills of human beings and further ease their work (Usher and Kaber, 2000).

Two theoretical constructs have been utilized for the assessment of technology acceptance. The Technology Acceptance Model (TAM) describes how users of an information technology come to accept and (voluntarily) use a technology. Empirical research on the TAM has shown that the model is highly predictive of the adoption and use of information technology (Davis et al., 1989; Adams et al., 1992; Venkatesh and Davis, 2000; Venkatesh and Morris, 2000).

The Task-Technology Fit Model (TTF) identifies the technology characteristics and the task characteristics as relevant factors that influence the impact of an information technology on individual performance (Goodhue and Thompson, 1995). Goodhue and Thompson (1995) suggest that technologies must be utilized and fit the task they support in order to have a performance impact, thereby determining the success of an information technology (Goodhue and Thompson, 1995). The TTF has been shown to be generally relevant for mobile information systems (Gebauer and Shaw, 2004).

It can be concluded that how the innovation is perceived by the work staff is of essential importance. Only if there is an added value, e.g. in the form of enhanced individual performance and/or relief from workload, will new technologies be accepted and used by the employees. In general, we assume that fault-finding performance can be improved by information technology that is custom designed for maintenance purposes.

The hypotheses of the current study are as follows:

1.. The experimental group (which uses the mobile application in troubleshooting) will solve the fault scenarios in a shorter period of time.

2.. The experimental group will make fewer errors compared to the control group.

Finally, we assume that the development of the mobile application on the basis of the cognitive strategies of the respective experts will have a positive effect on the technology acceptance and on the task-technology fit.

Based on the illustrations outlined above, an integral part of human-centered design (HCD) is the consideration and integration of the target groups of maintenance workers as well as Subject Matter Experts (SMEs) who come to use the mobile application. Therefore, factory workers were engaged in the development process of the application. The development process is described in detail in the following paragraphs.

2. The implementation of principles of human-centered design in a mobile fault-finding application

The mobile application was developed in four distinct steps:

1. In-depth interviews with maintenance workers
2. Development of fault scenarios by Subject Matter Experts (SMEs)
3. Selection of the layout and concept design
4. Evaluation and testing in the field

2.1. In-depth interviews with maintenance workers

In-depth interviews were conducted with eight maintenance workers of a large manufacturing company in an Eastern European country. The age of the interviewees varied, and their length of experience in the work field ranged from one to 18 years. The interview questions were developed in order to access information about the employees' approach to a given failure scenario (see Appendix 1). We asked the maintenance workers about the diagnostic process, e.g. what the procedure looks like, how faults can be classified, and which information is accessible and usable. The appendix contains the main questions from the interview guide.

The purpose of these in-depth interviews was to assess the situational conditions as soon as a fault arises. We wished to gain information about (i) the communication (e.g. via telephone or walkie-talkie) between the maintenance workers, (ii) how potential causes of the fault might be identified (e.g. by using the manual or reading the instructions provided by a display), (iii) which information is relevant and/or accessible, and (iv) how a fault is fixed (e.g. how the necessary parts are taken out and replaced). This methodology is viewed as an integral part of the HCD approach. The method allows the active involvement of actual users as part of the user requirements and task analysis, an approach which improves the understanding of user and task requirements towards the system. Iterative design, usability evaluation, task analysis, informal expert review and field studies have been identified as key HCD methods (Mao et al., 2005).

As a result, maintenance workers reported that in the case of a fault which is indicated by the plant operator, who calls the maintenance personnel by phone, the plant operator is asked for important information about the symptoms of the fault, e.g. what the operator has already checked. The maintenance worker then checks possible fault-finding steps which the operator has already conducted. The maintenance worker is reliant on the quantity and quality of the information given by the operator. Work experience is reported as a very important asset, as an experienced maintenance worker is able to predict the effectiveness of specific solutions (e.g. in 85% of cases, solution No. 3 out of 5 is the correct solution). Experienced maintenance workers have the solution “in their head”, and in the case of unfamiliar faults, error code lists are used. Experience is stated at the most important factor for reducing time in fault diagnosis, as the successful fault-finding steps are part of the experts' knowledge.

As concrete solutions for optimizing the fault-finding process, maintenance workers suggested placing manuals and error code lists closer to the manufacturing station or robot systems. In many cases, the manuals are located too far from the robot station, and one needs to take a bicycle to reach them (because the manufacturing hall is some 100 m long). It could be helpful to have some hints especially when faults have no visual symptoms, because one does not have the correct solutions in mind. Some robot systems do have instructions which can be accessed via the touchscreens at the robot station; however, these are not very user-friendly, requiring many screens to be clicked through before reaching the information needed, which costs a lot of time. Other robot stations show only the error message, but do not provide any help about possible causes or repair instructions. Currently, some maintenance workers carry private handwritten notebooks with them, in which they have noted the most frequent error codes and instructions. Others have downloaded the App provided by the supplier (which they paid for themselves) on their private mobile phones, and therefore use their private mobile phones even though this is not allowed. In most cases, the paper version of error codes is not used because searching for error descriptions and instructions is too laborious. Maintenance workers summarized that a mobile device which they can carry with them while walking through the plant would be very helpful, as a lot of time is currently wasted due to the long distances that need to be covered when they walk (or cycle) to the manufacturing system to identify the fault or read the error message; walk or cycle back to the desktop computer located in the work room in a different location in the factory building to see what the error code means; and then walk or cycle back to fix the problem.

2.2. Development of fault scenarios by SMEs for evaluation

On the basis of the information gained from the interviews, three fault scenarios (technical communication fault, gate system fault, misalignment of robot axis), representative for many of the faults occurring, were developed by three respective SMEs who had not been interviewed at an earlier stage.

Scenario 1 simulated a problem with the communication of several components of the plant. The cause was a defect optical fiber.

Scenario 2 simulated a fault at the gate system of a manufacturing plant. The energy supply system was manipulated in order to simulate a low voltage, caused by a defect cable. Thus, the system generated an error code and the gate could not be closed or opened either way.

Scenario 3 simulated a misalignment of a robot axis. The misalignment was caused by a defective connector in the communication module of the industrial robot.

The SMEs had several years of work experience and were chosen by their supervisors for the development of the fault scenarios. Also in this stage of development, structured interviews were conducted in the field using the Critical Decision Method (CDM, Crandall et al., 2006) in order to elicit information about the SME's decision-making and problem-solving processes during non-routine critical incidents, which are part of fault diagnosis (Crandall et al., 2006). The SMEs were urged to explain the sequence of the respective fault scenario and were asked about the following aspects:

a) the strategy used once a fault occurs;
b) the information which is available and relevant;
c) the symptoms which can appear;
d) the probabilities associated with a particular cause of the occurred fault
e) the estimated time for the steps of diagnosis

The interviews had a duration of approximately 20–30 min and were part of an iterative design process. The analysis of the user and task requirements is essential in order to develop usability goals and objectives.

As a result, the general requirements of the system were clarified, fault scenarios were documented, and user and task requirements were identified. The requirements included a navigation through the support system of the technology, a search function, and visualization of faulty components. On the basis of the validation interviews, the cognitive strategy of the SMEs was documented and used for the concept draft of the mobile application.

2.3. Selection of the layout and concept design

The procedure of the fault diagnosis was modeled to determine the logical structure of the fault diagnosis process (see Fig. 1). Subsequently, the modeling of the fault scenarios was used for the

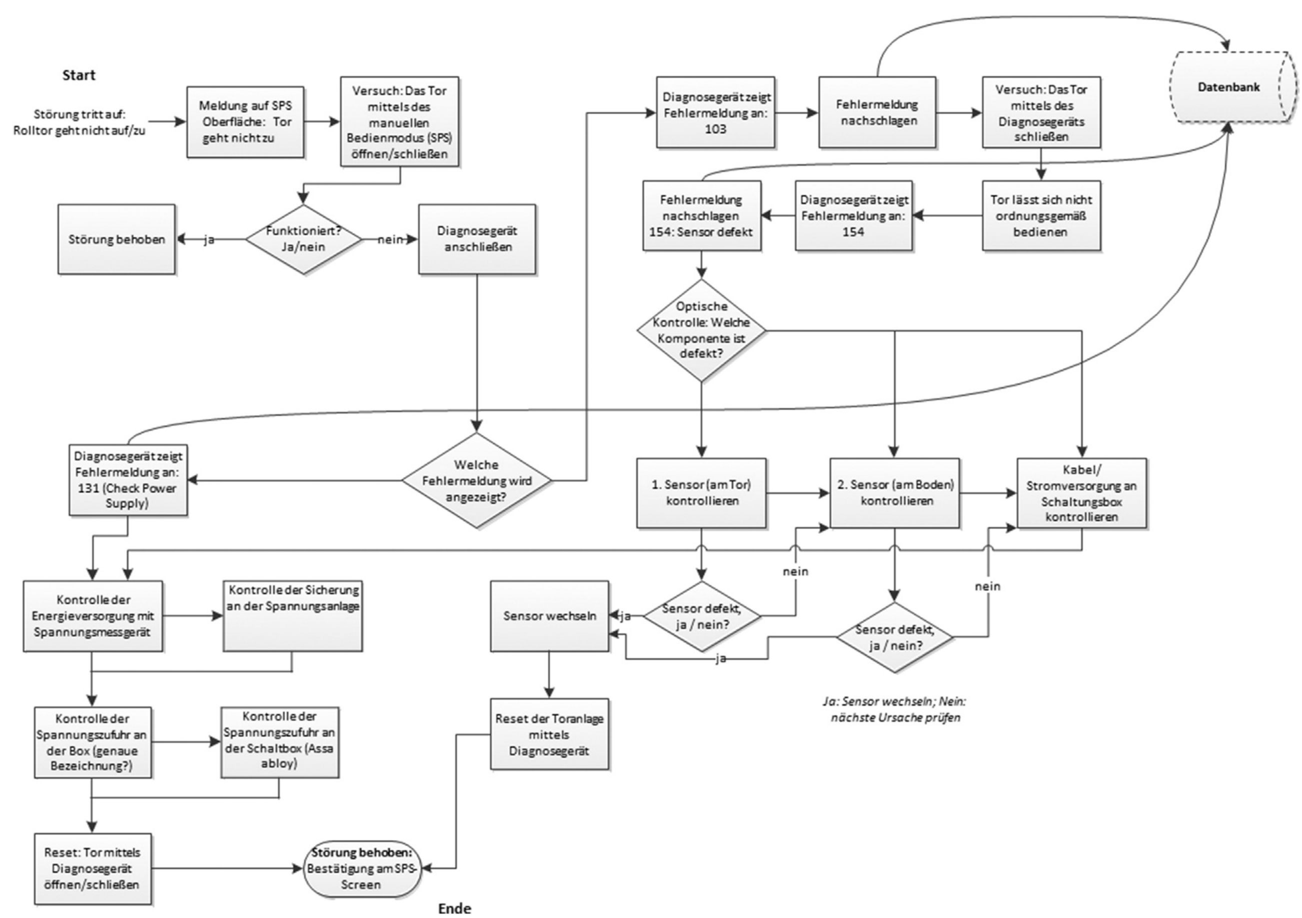

Fig. 1. Modeling of fault diagnosis procedure of scenario 1.

design of a mobile application using a web-based wireframing and prototyping design framework. The criteria of the TAM and TTF were considered in the development process of the application. Therefore, the mockups were presented and discussed with the SMEs. Based on the feedback of the target group, changes in the application were integrated into the existing concept.

2.4. Evaluation and testing in the field

In order to measure the effects and potential benefits of the usage of the mobile application, a field study (with an experimental and a control group) was conducted. The study had a quasi-experimental two-group design with a total of $n = 42$ participants from the maintenance department of the manufacturer. The experimental group (EG) worked with the mobile application while a control group (CG) worked without the application.

Each group consisted of 21 participants with differing ranges of age and work experience. The participants in the experimental group had a mean age of $M = 29.00$ ($SD = 8.48$) years with a mean work experience of $M = 2.21$ ($SD = 1.78$) years, while the participants in the control group were $M = 29.00$ ($SD = 8.60$) years old with a mean work experience of $M = 3.76$ ($SD = 3.54$). Both groups were subdivided into groups of seven participants for each of the three scenarios. The experiment was discontinued if a participant was unable to diagnose the fault within a twenty-minute time frame.

2.5. Procedure

The field study was conducted in the training center of a large manufacturer in an Eastern European country. The training center is usually used for the training of novice maintenance workers in basic skills, fault finding and repair of several different technologies such as electricity, welding technology, and robotics. For the purpose of the study, the training center was used to implement the three fault scenarios in order to provide a controlled experimental setting for all participants and in order to avoid disturbing the regular manufacturing process (Fig. 2). The quasi-experiment was led by a researcher of the University of Duisburg-Essen and one experienced supervisor who attended all 42 trials as an observer (see below).

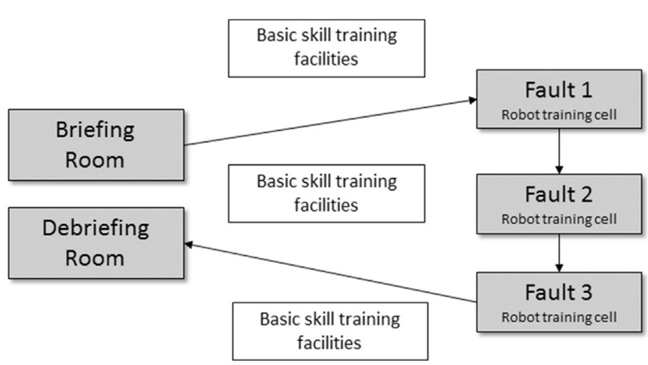

Fig. 2. Schematic figure of the training center.

The participants were individually briefed about the purpose of the study in a separate room near the training center of the manufacturer. Next, the participants were individually asked to handle one of the three fault scenarios, either with (experimental group) or without (control group) the mobile application. The time for successful diagnosis ranged from 4:12 min to 18:56 min with a given maximum time frame of twenty minutes for each participant. The study was conducted on three consecutive days in July 2014. After the fault-finding exercise, the participants remained in a separate room and were instructed not to talk to their peer workers about the purpose and contents of the study. After fault diagnosis, participants of the control group were able to interact with the mobile application for a few minutes to gain a first impression, enabling them to rate the application. The procedure for each participant lasted for approx. 30 min including briefing and debriefing (see Fig. 3).

2.6. Independent variable: description of the fault-finding application

The development of the mobile application was based on the information gained from the interviews with the SMEs, which were completed in the second phase of the development process. The diagnostic process, which was stated verbally by SMEs, was transformed into decision trees (see Fig. 4) as a visual representation of the procedural knowledge. Based on the decision trees, the next step was the development of prototypes for the mobile touch interface of the application. Therefore, a wireframing tool was used. The mockups formed the basis for the development of a working application running on the Apple iPad.

Fig. 4 illustrates the mockup development process of fault scenario 2. The fault-finding process begins with a first manual test of the gate functionality (1). Second, the maintenance worker is asked to connect a device for fault diagnosis (2), which displays an error code. The maintenance worker is able to use the dynamic search of the application (3). The application displays an instruction text for

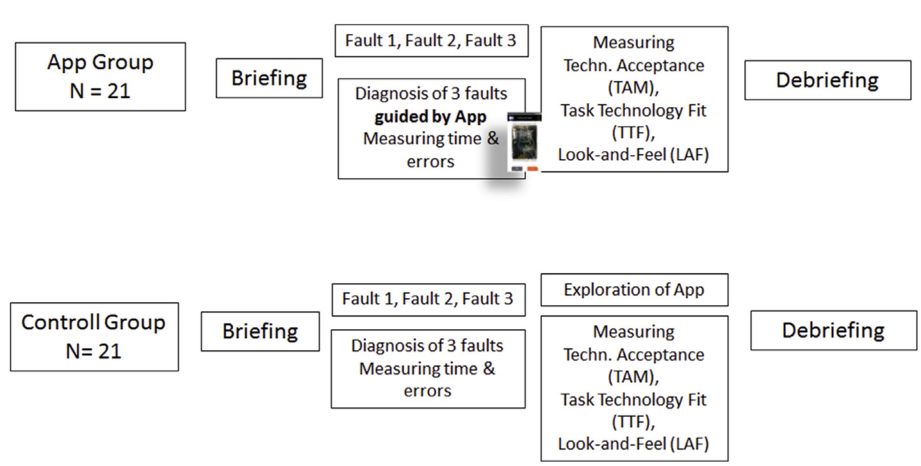

Fig. 3. Procedure of the field study for experimental App group (upper part) and control group (lower part).

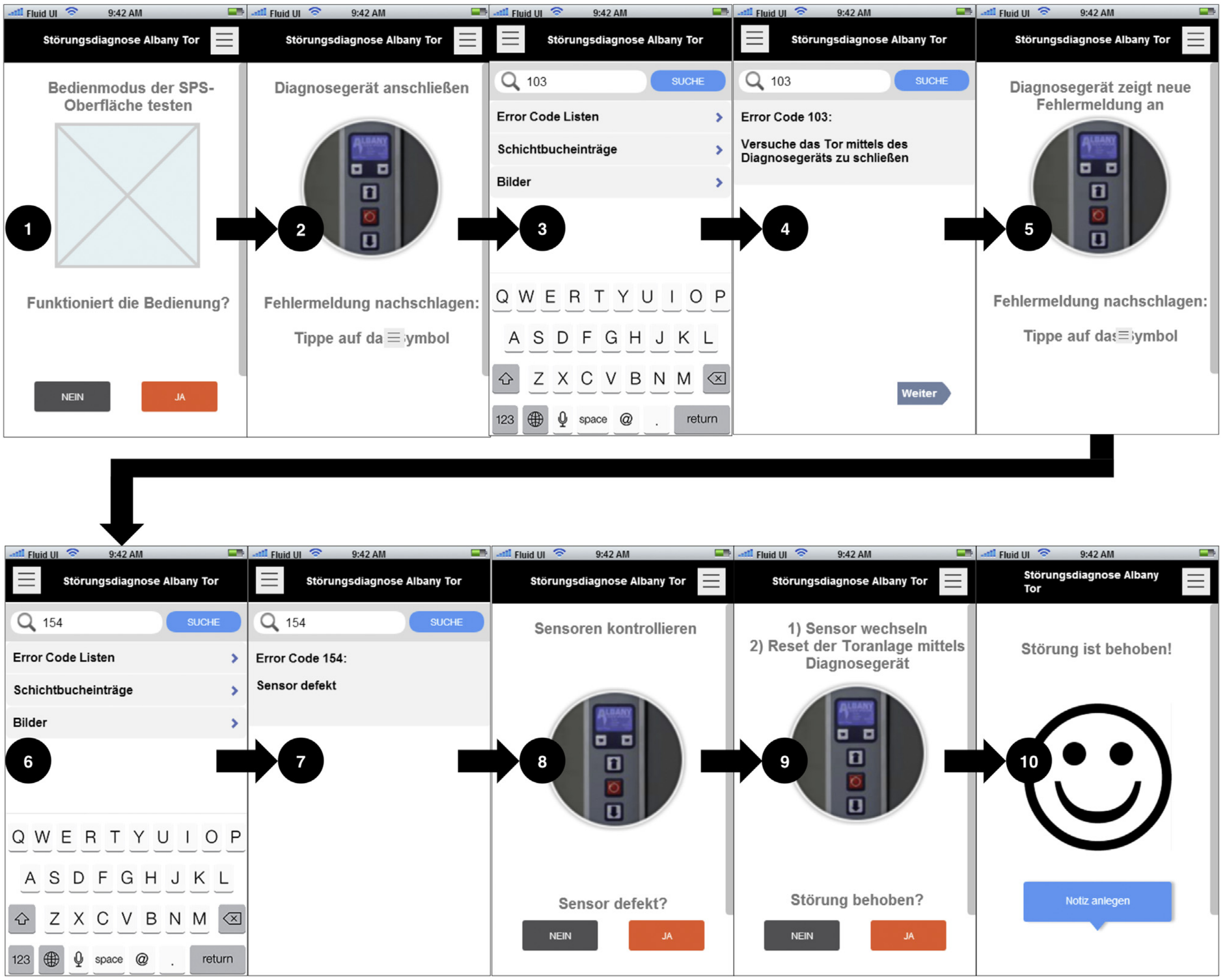

Fig. 4. Mock up development for scenario 2.

the respective error code (4). The failure persists and the diagnosis device displays a second error code (5). Again, the worker uses the dynamic search (6) and is provided with an instruction text (7). She/he follows the instructions (8) and after replacing the faulty component, resets the gate system (9). The diagnosis process ends with a positive feedback screen (10).

The mobile application contained various graphical, verbal and numerical presentations of information about the faults. Pictures and photos showing components of the production system (which were relevant for the respective fault scenario) were combined with annotations. In addition, diagnostic steps were explained in terms of a checklist. A search function was implemented to enable the participants to search for error codes in the digital technical documents of the manufacturing plant.

The diagnostic process was modeled using the decision-tree process. Participants could navigate through the application by answering questions (e.g.: Connector defect?) with regard to possible causes of the fault, thus testing potential faulty components. The application displayed information in the form of instructional text and photos (which were taken and provided by the SMEs) of the faulty components of the plant. At certain steps, an *info* button was implemented; by touching the icon, participants could get additional, relevant information about the plant (see Fig. 5).

A message feature was also implemented, which enabled participants to write a brief documentation about the occurred fault. The images were annotated (see Fig. 6), thus providing a description of the faulty component. Additionally, the images were tagged (e.g. marked with circles, arrows, or labels of the components) in order to point users' attention to the relevant parts of the modules of the production unit (Fig. 6).

2.7. Dependent variables

Fault diagnosis performance was measured by

(a) the time for diagnosis, which was measured for each participant using time stamps (in the application) as well as manual time measurement. Additionally,
(b) errors were observed and listed by the experienced supervisor and expert in the respective technology, who observed (participating observation) the fault-finding process of the participants and was standing nearby (but not in the way).
(c) time savings were computed for each scenario plus overall accumulated time savings.

2.8. Subjective ratings regarding technology acceptance and task-technology fit

A questionnaire was developed based on the work of Davis

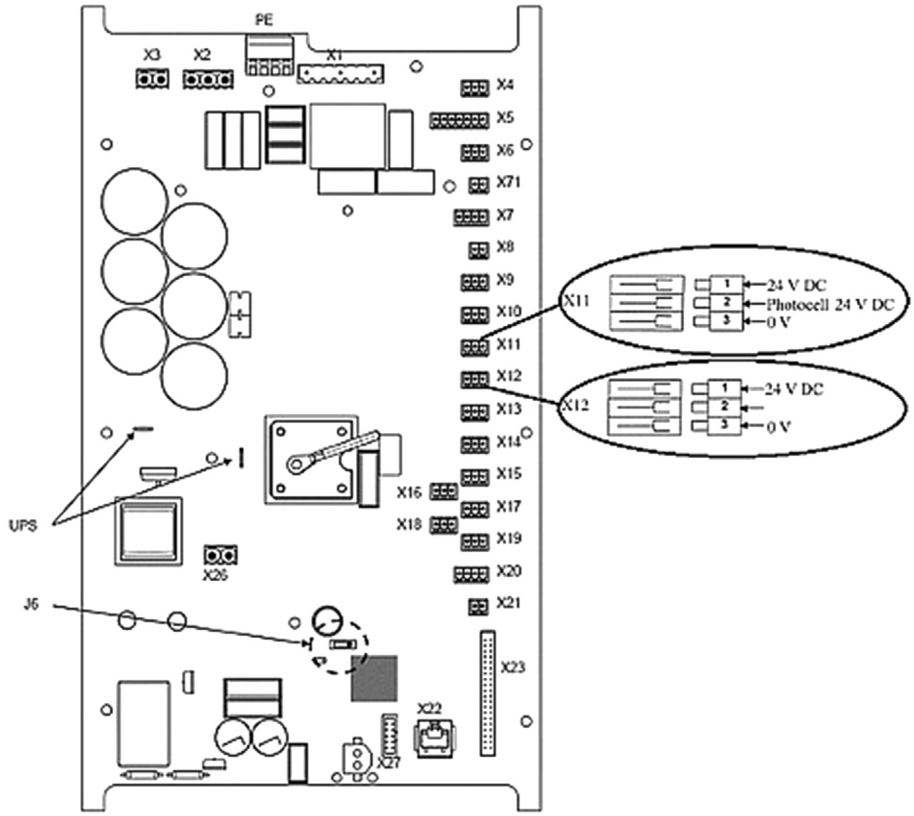

Fig. 5. Technical drawing of a faulty component.

Fig. 6. Annotations of system components (in German).

(1989) on the Technology Acceptance Model (TAM) and Goodhue and Thompson (1995) on the Task-Technology Fit Model (TTF, see Appendix 2). All items were adapted to the present context and translated by a professional translator into the language of the country where the field study was conducted.

2.9. Technology Acceptance Model (TAM)

Participants were asked to rate their subjective impression of working with the mobile application on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). The questionnaires for the

TAM contained two distinct scales:

1) *perceived usefulness* scale (six items), which measures the extent to which an individual believes that the use of an information technology can enhance his/her work performance. The Cronbach's alpha coefficient indicated a good internal consistency of the scale ($\alpha = 0.92$). Item example: “Using the application helps me identify faults in a better way.”
2) *perceived ease of use* scale (four items), which measures the extent to which an individual believes that the use of a technology can be carried out simply and easily (Cronbach's $\alpha = 0.90$). Item example: “It is easy for me to perform tasks using the application.”

2.10. Task-Technology Fit Model (TTF)

The TTF utilizes two task-technology fit measures,

1) *the right data and right level of detail*, which measures whether or not the technology provides the user with the right information and whether the *level of detail* meets the user's requirements (five items, Cronbach's $\alpha = 0.62$). Item example: “The information available in the application is the right information that I need to perform my work.”
2) The users' *perceived ease of software use* (Goodhue and Thompson, 1995) was measured with two items (Cronbach's $\alpha = 0.68$). Item example: “It is easy to learn how to use the application.”

Moreover, two additional items were formulated for measuring the “look and feel” (LAF) of the application. These items were: “Handling the application was intuitive”; and “I like the appearance of the application.”

Besides the closed questions, four open questions were also formulated. These asked, for example, for feedback about the application as well as for suggestions for improvement of the mobile application, and were as follows:

1. What is your general impression?
2. What did you find particularly good?
3. What did you find particularly bad?
4. What would you like to be added to the mobile application?

3. Results

For data analysis, some data sets were excluded due to missing values for some of the questions of the survey. All data sets ($n = 42$) could be included for the analysis of *time for diagnosis*. One data set was excluded from the analysis of Technology Acceptance due to one missing value. Two data sets were excluded from the analysis of the *right data and right level of detail* due to missing values. One data set was excluded from the evaluation of the *look and feel* of the application due to missing values.

3.1. Testing hypothesis 1: time for diagnosis

As the time for diagnosis was measured for each participant, it was possible to use all data sets for the analysis of the time measures. The overall mean time (independent of scenario) for the experimental group was $M = 9.31$ ($SD = 4.22$) and the mean time for the control group was $M = 14.99$ ($SD = 4.94$). The results of the ANOVA show a significant difference between the groups: $F(1,40) = 16.00$; $df = 40$; $p < 0.01$. The effect size coefficient $d = 1.23$ indicates a (very) large effect.

Participants in the experimental group finished the diagnosis 24% (scenario 1), 47% (scenario 2), and 46% (scenario 3) faster, respectively. The mean time measured for each scenario can be seen in Table 1. The results therefore support hypothesis 1.

3.2. Testing hypothesis 2: number of errors

Participants in the experimental group made fewer errors (total: 15; $M = 0.71$, $SD = 0.85$) compared to the control group (total: 36, $M = 1.71$, $SD = 1.31$). The results of the ANOVA indicate a significant difference between the groups: $F(1, 40) = 8.65$; $df = 40$; $p < 0.01$. The effect size coefficient $d = 0.91$ indicates a large effect. Thus, Hypothesis 2 is also supported.

3.3. Technology acceptance

Overall, participants perceived the mobile application as a useful tool for their work tasks. *Perceived usefulness* was rated (Likert scale from 1 to 5) at $M = 4.10$ ($SD = 0.77$) by all participants in the study, while *Perceived ease of use* was rated at $M = 4.35$ ($SD = 0.68$).

3.4. Task-technology fit

The task-technology fit, which was characterized by the *right data* in the mobile application as well as the *right level of detail* of this information, was rated (Likert scale from 1 to 5) at $M = 3.70$ ($SD = 0.68$). The *ease of use of the software* was rated at $M = 4.57$ ($SD = 0.65$).

3.5. Look and feel

The intuitive interaction with the mobile application was rated at $M = 4.34$ ($SD = 0.96$). The visual appeal of the mobile application was rated at $M = 4.46$ ($SD = 0.71$). Table 2 summarizes the scale ratings given by the participants.

3.6. Post Hoc correlation analysis of performance and acceptance variables

As shown in Table 3, in the EG, there was a medium to large correlation between the TAM (usefulness) and the TTF, and a medium to large interscale correlation between the TTF scales. The

Table 1
Mean time measures (in minutes) for experimental group (EG) and control group (CG) in each fault scenario.

Time (in min.)	EG M (SD)	CG M (SD)	Savings
Scenario 1	13:47 (4:16)	18:14 (3:08)	4:27
Scenario 2	8:11 (1:11)	15:29 (4:12)	7:18
Scenario 3	6:00 (1:23)	11:15 (4:58)	5:15

Table 2
Mean scale ratings (Likert scale 1–5): Overall, experimental group and control group.

Scales/Items	Overall M (SD)	EG M (SD)	CG M (SD)
Perceived Usefulness (TAM)	4.10 (0.77)	4.02 (0.75)	4.19 (0.80)
Perceived Ease of Use (TAM)	4.35 (0.68)	4.36 (0.63)	4.34 (0.75)
The Right Data/Level of Detail (TTF)	3.70 (0.68)	3.62 (0.60)	3.79 (0.76)
Ease of Software Use (TTF)	4.57 (0.65)	4.67 (0.53)	4.48 (0.75)
Intuitive Interaction (LAF)	4.34 (0.96)	4.19 (1.08)	4.50 (0.83)
Visual appeal (LAF)	4.46 (0.71)	4.29 (0.85)	4.65 (0.49)

Note: TAM = Technology Acceptance Model, TTF = Task Technology Fit, LAF = Look and Feel.

Table 3
Correlation between dependent variables in the experimental group ($n = 21$).

EG	No.	1	2	3	4	5	6	7
Errors	1	–						
Diagnosis time in min.	2	0.540*						
Usefulness (TAM)	3	0.073	−0.208					
Ease of Use (TTF)	4	−0.270	−0.387	0.486*				
Right Data (TTF)	5	−0.285	−0.157	0.356	0.508*			
Ease of Software Use (TTF)	6	−0.389	−0.354	0.275	0.638**	0.318		
Intuitive Handling (LAF)	7	−0.212	−0.236	0.192	0.450*	0.443*	0.465*	
Appreciation of Appearance (LAF)	8	0.050	−0.115	0.677**	0.412	0.581**	0.278	0.102

Note: *. $p < 0.05$, **$p < 0.01$, TAM = Technology Acceptance Model, TTF = Task Technology Fit, LAF = Look and Feel.

intuitive handling is mainly affected by ease of use, the right data, and the ease of software use.

For the CG, it can be shown that the maintenance workers in the control group who made many diagnostic errors before experiencing the application rated the intuitive handling of the application less positively. Nevertheless, usefulness, right data, and ease of software use are also highly correlated in the control group. The best predictors for the appreciation of appearance are the perceived usefulness, ease of use, right data and intuitive handling (Table 4).

In summary, for both groups, there is a strong relationship between the usefulness and ease of use. In particular, the right data seems to be a central concept for predicting the ease of use. Task-Technology Fit and Technology Acceptance are very closely connected, and the right data needs to be carefully considered in HCD through the use of the cognitive strategies of the SME and the information used by the maintenance workers.

3.7. Additional analysis of answers to open questions

Participants provided their feedback on their performance with and use of the technology in four open questions. They were asked to state their feedback in the form of a general impression of the technology, which features were good, which were poor, and whether there is anything they would like to be integrated into the mobile application.

Overall, participants provided 126 answers. 73 answers were unambiguously positive, while 16 answers were unambiguously negative.

3.8. The workers' general impression

Sixteen participants indicated having a good overall impression ("good", "good idea"). Seven participants described a positive overall impression using adjectives such as "interesting", "super", "great" "pleasant" or "excellent". Three participants stated that this technology is "useful", while five participants stated that it is particularly "useful for inexperienced workers". Four participants noted that the mobile application enables a "faster diagnosis" and "saves time". Three participants stated that they desired "more details", whereas three participants described the mobile application as "impractical" because sometimes you need "both hands" in order to perform diagnosis and that the tablet could be damaged while performing certain tasks (e.g. gluing of components).

What did the maintenance workers find particularly good?

Nineteen participants liked the "ease of use" and "understandability" of the mobile application. Fifteen participants enjoyed the "visualization" and "graphical presentation" of the diagnostic task via images and explanations of the faults. Four participants outlined the "clarity" of the content, while five participants praised the "intuitive interaction". Three participants mentioned that using the mobile application "saved time" and "accelerated the diagnosis".

What did the maintenance workers find particularly bad?

Four participants were "missing details" that would have been useful for the respective fault. Three participants stated that while the mobile application was "good", it might impair "critical thinking and action". One participant criticized the size of the device, another reported "hand pain", while a third stated that "working with one hand" is not always possible, as both hands are required for executing certain tasks. One further participant noted a "problem with the touch function", and another criticized the "design" of the mobile application.

What would the maintenance workers like to be added to the mobile application?

Four participants asked for further "details" and "additional information" in general, while three requested further "instructions" and "technical documentation" and one would have liked to have a "tree structure of the diagnostic steps", and an extended level of detail (images, schemes) in order to enhance "learning" with the mobile application. Three participants wanted an account of the time taken and desired a "time display". Two participants asked for a linkage with other information-system modules of the production unit. One participant wished for a real-time request of the stock via the mobile application (e.g. information about available spare parts). Another participant asked for a "message system".

Table 4
Correlation between dependent variables in the control group ($n = 21$).

CG	No.	1	2	3	4	5	6	7
Errors	1	–						
Diagnosis time in min.	2	0.427						
Usefulness (TAM)	3	0.332	0.016					
Ease of Use (TTF)	4	0.267	−0.042	0.710**				
Right Data (TTF)	5	0.349	0.041	0.676**	0.710**			
Ease of Software Use (TTF)	6	−0.033	−0.002	0.410	0.547*	0.552**		
Intuitive Handling (LAF)	7	−0.479*	−0.150	0.361	0.170	0.252	0.108	
Appreciation of Appearance (LAF)	8	0.134	−0.032	0.561**	0.520*	0.518*	0.245	0.455*

Note: *. $p < 0.05$, **$p < 0.01$, TAM = Technology Acceptance Model, TTF = Task Technology Fit, LAF = Look and Feel.

4. Discussion

The goals of the field study were to explore the potential of a human-centered fault diagnosis tool for mobile devices in order to enhance the overall quality of error diagnosis, to reduce time in troubleshooting, and to show the usefulness of applying HCD principles in order to derive the layout and design concept for maximizing user acceptance. The results show that the first goal concerning the reduction of time, was achieved successfully. Fault diagnosis in manufacturing can be enhanced through the introduction of mobile information technology that is aligned with the cognitive strategies used by maintenance workers. We showed that the usage of a mobile application for fault finding can, in fact, enhance job performance. Participants in the experimental group completed troubleshooting faster, proceeded more systematically, and made fewer errors compared to the control group. By integrating mobile information technology into the work processes, manufacturing enterprises could make use of an array of positive effects with regard to reduced costs, improved job performance of the work staff and higher levels of job satisfaction due to cognitive relief (Wendemuth et al., 2014).

The methodology of a human-centered design (HCD) process proved to be very useful in order to maximize technology acceptance by the maintenance workers and their perceived task-technology fit. The interviews at the beginning of the process made us aware of the questions:

- who will use the product, e.g. maintenance workers with different years of fault finding experience but especially those without experience,
- what they will use it for, e.g. fast und correct fault diagnosis with the help of user manuals, error codes and written instructions that are available “at hand”, and
- under what conditions they will use it, e.g. under time pressure and in order to reduce the time for moving between different locations in the manufacturing hall.

By creating design solutions which were iteratively discussed with the maintenance workers in order to best meet their expectations concerning usefulness, it was possible to demonstrate the usability in the quasi-experimental field study with the actual users. This shows that the time needed for directly speaking to and discussing with the maintenance workers “paid off”: Solutions were created, as throughout the whole process of design and implementation, the maintenance workers were very supportive and full of ideas regarding how to improve their efforts in a fast fault-finding process by means of a mobile fault-finding application. As the mobile fault-finding application was mainly created based on the ideas of the future users and how they already use their mobile phones (e.g. downloading the supplier's App on their private phones), the “costs” for the implementation were also reduced, as the workers were really interested in using the App created for them.

4.1. Limitations

It should be noted that this was a preliminary study which was designed to assess the possible benefits of further work. Through the quasi-experimental design of the field study, it was not possible to fully replicate the real-life conditions of the diagnostic process. Nevertheless, the conditions of the diagnostic process were well imitated. Since the purpose of the study was to evaluate the effects and benefits of a mobile diagnostics application, to assess how troubleshooters can be supported during troubleshooting, and to gauge the user acceptance of the technology, it is a step in the right direction. In future research, it will be necessary to evaluate the findings under real-life conditions of fault-finding tasks.

However, the results of the study can be seen as having a high internal validity. At the same time, it was not possible to control for some of the variables in the field study. First, although participants were briefed about the goals of the study and were explicitly told that the purpose was not to evaluate their work performance, it cannot be ruled out that the test situation per se led to higher levels of stress and might consequently have biased their decision making. Second, despite using different staffrooms for the maintenance staff, and advising the workers not to reveal information about the tasks, it was not possible to control whether or not participants exchanged information about the fault scenarios.

4.2. Future research

Since the procedure was based on fault scenarios, future studies should validate the shown effects under real conditions in manufacturing. Moreover, future work should take into account the challenges inherent in designing information technology of mobile services in order to achieve user acceptance. In order for a new technology to be adopted, an HCD approach is crucial for its development. It is important to specify the intended context of use, the requirements for use, and the organizational requirements, to produce design solutions and to evaluate the solutions against the requirements (ISO, 2010). By extending research into various fields in manufacturing, it would then be possible to provide a deeper understanding of the potential of mobile information technology as well as the requirements in their respective application context.

5. Conclusions

Results of the field study confirm the broad potential for use of IT in a number of areas based on an HCD process. The results of the in-depth interviews confirm the demand for a mobile technology as daily work support for inexperienced as well as more experienced employees in the maintenance staff. The diagnostic process was accelerated, while simultaneously, the overall diagnostic quality improved. Notably, inexperienced employees made fewer observed errors. The empirical results validate the significant impact of information technology on job performance. Just for the three scenarios, the time for diagnosis was reduced by a total of 119 min. Additionally, participant observation showed a reduction in errors made during the diagnostic process. Furthermore, the importance of an HCD approach during the development process is a formative criterion of particular importance. Since experts have been shown to exhibit superior performance on memory tasks (Lindgaard, 1995), by conceptualizing a performance-supporting system based on the mapping of expert knowledge, an optimized diagnostic strategy can be achieved. Thus, it is important to attain a deeper understanding of the domain within which problem solving needs to be supported, and additionally of how the relationships between relevant variables are perceived by experts (Lindgaard, 1995). Moreover, the mobile context of use should be more focused, as users reported some concerns regarding the non-functional features, such as the weight of the device on which the mobile application is running.

Nevertheless, software development is a complex and often difficult process. The acceptance of technology is a key issue in software design, especially for maintenance goals. Only if a certain level of user acceptance is achieved and if the technology leads to higher job satisfaction and is viewed as enrichment for one's work as well as a relief of cognitive load is there an added value of using new technology.

Appendix 1

Main questions from the in-depth interview guide	
1.	What does the typical procedure look like?
2.	How are failures perceived and classified?
3.	Which information is used to plan further steps of troubleshooting?
4.	Which considerations lead to a decision?
5.	Are decisions made on your own?
6.	How is the error diagnosis time comprised?
7.	Are error code lists being used?
8.	How could one work more efficiently?
9.	Which aids could be helpful in order to facilitate the diagnosis process?

Appendix 2

TAM and TTF scales

Item	Scale
The application allows me to get my work done faster.	Perceived Usefulness Scale (TAM)
I think the application helps me to work more systematically.	Perceived Usefulness Scale (TAM)
Using the application enables me to accomplish tasks more quickly.	Perceived Usefulness Scale (TAM)
Using the application makes it easier to do my job.	Perceived Usefulness Scale (TAM)
Using the application helps me identify faults in a better way.	Perceived Usefulness Scale (TAM)
I find the application useful in my job.	Perceived Usefulness Scale (TAM)
The use of the application is easy for me to understand.	Perceived Ease of Use Scale (TAM)
The use of the application is flexible.	Perceived Ease of Use Scale (TAM)
It is easy for me to perform tasks using the application.	Perceived Ease of Use Scale (TAM)
I find the application easy to use.	Perceived Ease of Use Scale (TAM)
The information available in the application is the right information that I need to perform my work.	The Right Data & Right Level of Detail (TTF)
It is more difficult to do my job effectively because some of the information I need is not available.	The Right Data & Right Level of Detail (TTF)
The application maintains sufficient information.	The Right Data & Right Level of Detail (TTF)
The diagnosis application maintains information at an appropriate level of detail for my purposes.	The Right Data & Right Level of Detail (TTF)
The application enables me to get information simply and quickly.	The Right Data & Right Level of Detail (TTF)
It is easy to learn how to use the application.	Ease of Use of Software (TTF)
The diagnosis application is convenient and easy to use.	Ease of Use of Software (TTF)
Using the application is intuitive.	Look and Feel (LAF, Ad Hoc Item)
I like the appearance of the application.	Look and Feel (LAF, Ad Hoc Item)
What is your general impression?	Open question
What did you think was particularly good in the application?	Open question
What did you think was particularly poor in the application?	Open question
Is there something you would like to have added into the application?	Open question

References

Adams, D.A., Nelson, R.R., Todd, P.A., 1992. Perceived usefulness, ease of use, and usage of information technology: a replication. MIS Q. 16, 227–247.

Chavaillaz, A., Wastell, D., Sauer, J., 2016. Effects of extended lay-off periods on performance and operator trust under adaptable automation. Appl. Ergon. 53, 241–251.

Crandall, B., Klein, G., Hoffman, R., 2006. Working Minds – a Practitioners Guide to Cognitive Task Analysis. Massachusetts Institute of Technology, Cambridge, MA, USA.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 13 (3), 319–339.

Davis, F.D., Bagozzi, R.P., Warshaw, P.R., 1989. User acceptance of computer technology: a comparison of two theoretical models. Manag. Sci. 35, 982.

Duncan, K.D., 1985. Representation of fault-finding problems and development of fault-finding strategies. Innov. Educ. Train. Int. 22 (2), 125–131.

Fitter, M.J., Sime, M.E., 1980. Responsibility and shared decision making. In: Smith, T.H., Green, T.R.G. (Eds.), Human Interaction with Computers. Academic Press, London.

Gebauer, J., Shaw, M.J., 2004. Success factors and impacts of mobile business applications: results from a mobile e-procurement study. Int. J. Electron. Commer. 8 (3), 19–41.

Goodhue, D.L., Thompson, R.L., 1995. Task-technology fit and individual performance. MIS Q. 19 (2), 213–236.

Gramopadhye, A.K., Drury, C.G., Prbhu, P.V., 1997. Training strategies for visual inspection. Hum. Factors Ergon. Manuf. Serv. Ind. 7 (3), 171–196.

Ham, D.H., Yoon, W.C., 2007. The training effects of principle knowledge on fault diagnosis performance. Hum. Factors Ergon. Manuf. Serv. Ind. 17 (3), 263–282.

Henneman, R.L., Rouse, W.B., 1984. Measures of human problem solving performance in fault diagnosis tasks. IEEE Trans. Syst. Man Cybern. 14, 99–112.

Hermawati, S., Lawson, G., D'Cruz, M., Arlt, F., Apold, J., Andersson, L., Lövgren, M.G., Malmsköld, L., 2015. Understanding the complex needs of automotive training at final assembly lines. Appl. Ergon. 46, 144–157.

Hochholdinger, S., Schaper, N., 2013. Training troubleshooting skills with an anchored instruction module in an authentic computer based environment. J. Tech. Educ. 1, 7–22.

ISO/IEC, 2010. ISO 9241-210:2010 Ergonomics of Human-system Interaction – Part 210: Human-centred Design for Interactive Systems. ISO/IEC, Geneva, Switzerland.

Jasperneite, J., Niggemann, O., 2012. Intelligent assistance systems to control the complexity of industrial automation systems. Atp-Edition 54 (9), 36–45.

Jonassen, D.H., Hung, W., 2006. Learning to troubleshoot: a new theory-based design architecture. Educ. Psychol. Rev. 18 (1), 77–114.

Kluge, A., 2014. The Acquisition of Knowledge and Skills for Taskwork and Teamwork to Control Complex Technical Systems. A Cognitive and Macroergonomics Perspective. Springer, Dortrecht.

Kluge, A., Grauel, B., Burkolter, D., 2013. Job aids: how does the quality of a procedural aid alone and combined with a decision aid affect motivation and performance in process control? Appl. Ergon. 44, 285–296.

Kluge, A., Greve, J., Borisov, N., Weyers, B., 2014. Exploring the usefulness of two variants of gaze-guiding-based dynamic job aids for performing a fixed-sequence start-up procedure after longer periods of non-use. Int. J. Hum. Factors Ergon. 3 (2), 148–169.

Kluwe, R.H., 1997. Informationsverarbeitung, Wissen und mentale Modelle beim Umgang mit komplexen Systemen [Information processing, knowledge and mental models]. In: Sonntag, K., Schaper, N. (Eds.), Störungsmanagement und Diagnosekompetenz [Failure management and diagnostic competence]. Vdf, Zürich, pp. 13–37.
Krems, J.F., 1997. Expertise und diagnostische Urteilsbildung [Expertise and diagnostic judgment]. In: Sonntag, K., Schaper, N. (Eds.), Störungsmanagement und Diagnosekompetenz [Failure management and diagnostic competence]. Vdf, Zürich, pp. 77–94.
Langlotz, C.P., Shortliffe, E.H., 1983. Adapting a consultation system to critique user plans. Int. J. Man-Machine Stud. 19, 479–496.
Leung, G.T.C., Yucel, G., Duffy, V.G., 2010. The effects of virtual industrial training on mental workload during task performance. Hum. Factors Manuf. Serv. Ind. 20 (6), 567–578.
Lindgaard, G., 1995. Human performance in fault diagnosis: can expert systems help? Interact. Comput. 7 (3), 254–272.
Linou, N., Kontogiannis, T., 2004. The effect of training systemic information on the retention of fault-finding skills in manufacturing industries. Hum. Factors Ergon. Manuf. Serv. Ind. 14 (2), 197–217.
Mao, J.-Y., Vredenburg, K., Smith, P.W., Carey, T., 2005. The state of user-centered design practice. Commun. ACM 48 (3), 105–109.
Morris, N.M., Rouse, W.B., 1985. Review and evaluation of empirical research in troubleshooting. Hum. Factors 27 (5), 503–530.
Ortegon, K., Nies, L.F., Sutherlnad, J.W., 2014. The impact of maintenance and technology change on remanufacturing as a recovery alternative for used wind turbines. In: 21st CIRP Conference on Life Cycle Engineering, pp. 182–188.
Patrick, J., 1993. Cognitive aspects of fault-finding. Le. Trav. Hum. 56 (2/3), 187–209.
Patrick, J., Haines, B., Munley, G., Wallace, A., 1989. Transfer of fault-finding between simulated chemical plants. Hum. Factors 31, 503–518.
Patrick, J., James, N., Friend, C., 1996. A field study of training fault-finding. Le. Trav. Hum. 59 (1), 23–44.
Rouse, W.B., 1978. Human problem solving performance in a fault diagnosis task. IEEE Trans. Syst. Man Cybern. 8 (4), 258–271.
Rouse, W.B., 1979. Problem solving performance of maintenance trainees in a fault diagnosis task. Hum. Factors 21 (2), 195–203.
Rouse, W.B., Rouse, R.M., 1979. Measures of complexity of fault diagnosis tasks. IEEE Trans. Syst. Man Cybern. 9, 720–727.
Sauer, J., Nickel, P., Wastell, D., 2013. Designing automation for complex work environments under different levels of stress. Appl. Ergon. 44 (1), 119–127.
Teague, R.C., Allen, J.A., 1997. The reduction of uncertainty and troubleshooting performance. Hum. Factors 39 (2), 254–267.
Usher, J.M., Kaber, D.B., 2000. Establishing information requirements for supervisory controllers in a flexible manufacturing system using GTA. Hum. Factors Ergon. Manuf. Serv. Ind. 10 (4), 431–452.
Venkatesh, V., Davis, F.D., 2000. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag. Sci. 46, 186–204.
Venkatesh, V., Morris, M.G., 2000. Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. MIS Q. 24, 115–139.
Wendemuth, A., Biundo, S., 2012. A companion technology for cognitive technical systems. In: Esposito, A., Esposito, A.M., Vinciarelli, A., Hoffmann, R., Müller, V.C. (Eds.), Cognitive Behavioural Systems, pp. 89–103.
Wendemuth, A., Diedrich, C., Böck, R., 2014. Dialogues in industrial diagnosis. In: Proceedings 3 Interdisciplinary Cognitive Systems Workshop: Human, Teams, Systems and Automats (page s.p).
Wilson, J.R., 2000. Fundamentals of ergonomics in theory and practice. Appl. Ergon. 31 (6), 557–567.
Woods, D.D., 1985. Cognitive technologies – the design of joint human-machine cognitive systems. AI Mag. 6 (4), 86–92.
Woods, D.D., Hollnagel, E., 2006. Joint Cognitive Systems. Patterns in Cognitive Systems Engineering. Taylor & Francis Group, Boca Raton, FL.