A Framework for Designers to Support Prospective Design

Nina Gérard, Stefanie Huber, Jens Nachtwei, Udo Schubert, Bram Satriadarma

Centre of Human-Machine Systems
Graduateschool *prometei*: Prospective Design of Human-Machine Interaction
Technische Universität Berlin
GRK1013/1, Sekr. FR 2-6, Franklinstr. 28-29, 13587, Berlin, Germany
{nge, shu, jna, usc, bsa} AT zmms.tu-berlin.de

Abstract In this paper, we analyze the term "prospective design" from various perspectives. The theoretical-, research- and empirical-based approximations are centralized. Furthermore, a working definition of "prospective design in human-computer interaction" is proposed in the first part of this paper. In practice, however, the mere definition of prospective design and awareness that user, context, technical system and task have to be known is not enough. Therefore we propose a new model of performance shaping factors in the second part of this paper. This model gives recommendations for human-computer interaction design and research in terms of factors that should be considered for a prospective design of user interfaces. The utility of this new model is shown using a case study in alarm psychology.

1 Introduction

Recently, the concept of "prospective design" has become more and more popular in human-computer research and design practice. Design projects and products are increasingly described as being prospective. Unfortunately, prospective is frequently confused with "user-centred" or "preventive". Until now a comprehensive definition of a prospective design process and its result (the prospectively designed product) has been missing. We have derived a definition from an extensive review of the literature as well as from two evaluations by experts. In practice, it is not enough merely to define prospective design and know about the user, context, technical system, and task. There is a need for a readily accessible knowledge base which contains concrete factors that should be considered if user interfaces are to be designed prospectively. We therefore bridge the goal of prospective design and an approach predominantly known from the 1980s: performance shaping factors (PSF) models. These models define influences on human performance in human-computer systems but are mostly not familiar to interface designers. Basically, a PSF model can serve as a kind of checklist for interface designers and researchers alike, and should answer the following question: Which factors have to be included in design decisions or studies on interface design in order to anticipate fundamental shortcomings and breakdowns of human-computer interaction in the future?

The importance of anticipating shortcomings in human-machine interaction as soon and as comprehensive as possible (one central aspect of prospective design) cannot be overstated. Especially accidents in aviation or process industry illustrate this vividly. In this context Leveson [19] states: "In general, safety is not a component property but an emergent property as defined by system theory. Emergent properties arise when system components operate together. Such properties are meaningless when examining the components in isolation - they are imposed by constraints on the freedom of action of the individual parts." This complexity becomes clear if we look at the results of accident analysis. For instance in 1993 an Airbus 320 crashed into an earth wall at the end of the runway due to serious braking problems shortly after landing. The wrong question would be "What was the cause?" since a whole interacting bunch of them lead to this critical incident. In essence a well-meant feature, implemented by Airbus software developers, was not adaptive to certain boundary conditions: The thrust-reverser is blocked during flight in order to avoid misuse and accidents respectively. So the system is not operating until a certain amount of force is exerted on the undercarriage. Unfortunately bad weather (especially rain and wind shears) lead to a sloped landing. Thus the force on the undercarriage was to low for some time and consequently the thrust-reverser's onset was too late. Combined with a false mental model of pilot and co-pilot as well as weather-induced mental workload this finally ended in disaster. Accordingly, we propose an integrative view on prospective design and a performance shaping factors model in order to prevent incidents and accidents of this kind.

In an effort to expand early PSF-models by more empirically verified factors and to offer a comprehensible cluster of these factors, we developed a new PSF-model. The definition of prospective design and the assembly of the PSF model will be described in the context of supervisory control, and implications for user interface design and research will be discussed.

2 Prospective Design in Theory & Practice

In recent years the term "prospective design" has become more and more prominent in human-computer interaction (HCI) research, often associated with user-centred design, adaptive or intuitive interfaces. But what is it really about?

In this section we propose a definition of "prospective design in HCI" that captures its theoretical and historical roots, but also considers current trends in practice and empirical research.

2.1 A First Theoretical Approximation

The term prospective design initially appeared in work psychology, describing one of three different strategies for designing work, or even whole work systems: corrective work design focuses on correcting given deficiencies of work design like faulty ergonomic solutions; preventive work design tries to anticipate and avoid future deficiencies that could harm the worker's physiological and psychological health; prospective work design creates opportunities for the worker to develop their own personality [40; 41]. In this context, personality development can be defined as exploiting ones potential by shaping of or searching for beneficial work conditions. Applied to Human Factors an example for the latter could be an adaptable user interface (e.g. one that allows operators in supervisory control to expand the monitoring information or vary the level of detail information, in order to meet their individual capabilities). All strategies, and in particular prospective work design, also indicate a trend towards a more anthropocentric workplace design: In former times people had to adapt to their workplace, its conditions and technical equipment - now the opposite is requested [2]. The work itself should support the worker's personality development. It should adapt or be adaptable to and consider his requirements and needs, so the user's characteristics are central for work design. These two contradictory perspectives in work design are also known as technique-centred vs. work-centred approach in work psychology [5; 41]: The operator in a work-centred job is more of a manager for different tasks, ranging from planning to control. In contrast, the human operator in a technique-centred job is solely the executor of tasks given by the automation without having any degrees of freedom.

2.2 A Second Research-Oriented Approximation

Unfortunately, in human-computer interaction research, no explicit definition of the term "prospective design" can be found. We reviewed recent research projects dealing with prospective design of human-computer interaction and identified a focus of work presented at the "7th Berlin Workshop Human-Machine Systems" on "Prospective Design of Human-Technology Interaction" in October 2007 [32]. However, the contributions there and in the broad literature did not give a clear picture of a prospective design project. Instead, we summed up all aspects we found associated with prospective design.

In some papers, authors referred directly to prospective design, including aspects such as user-oriented design, letting end-users participate during the design process, evaluating usability, or enhancing intuitive interaction [5; 11]. These aspects could also be assigned to the broader concept of user-centred design. Others refer to prospective design in terms of early phases of a product design process [21], prospective vs. retrospective design of sociological studies [14], preventive and corrective actions [3] or task allocation [24]. However, in some cases also without further explanation of the interpretation [16; 30].

2.3 A Third Empirical Approximation

To elaborate on this topic we conducted a short interdisciplinary questionnaire survey with human factors (HF) experts from the Centre of Human-Machine-Systems of the Technische Universität Berlin in April 2009. The Centre of Human-Machine-Systems consists of different parties from two universities and an institute of the Fraunhofer-Gesellschaft. In total, 26 experts (mean age: 34 years) from twelve institutions and ten faculties participated in the survey and responded to questions such as:

– How do you define "prospective design in HCI"?

– What are the advantages and disadvantages of prospective design in HCI?

– Where exactly would you locate prospective design in the product development process?

All participants deal with human-computer interaction issues in one way or another on a daily basis. Most participants had a main psychological background (16 participants), either exclusively or in combination with additional expertise, e.g. in engineering or HF. The remaining participants came from engineering, human factors, cognitive science, computer science and design. We quantified the level of expertise in human-computer interaction by three different measures: years of giving lectures in HCI (mean 3.64 years, standard deviation 7.69), years of attending HCI lectures (mean 3.57 years, standard deviation 5.24) and years of research in HCI (mean 6.77 years, standard deviation 10.18). Of all participants, five were very experienced (more than 10 years in one or more of the three scores).

2.3.1 Defining Prospective Design in HCI

The definitions given had some aspects in common. Almost half of the participants (10 participants) proposed synonyms such as *anticipatory, foresighted, goal-oriented, or future-oriented design* - terms that seem to be associated more with the concept of preventive design (see above).

A majority noted a *reference to time* (15 participants). In their opinion prospective design already tackles future user-problems or task requirements in the early stages of the product development process. However, this is generally what constitutes preventive design in work psychology [40].

Fourteen participants also regarded the *involvement of users* in the design process as a central aspect, although only a few definitions elaborated which aspects of the user it is important to consider. These included cognitive user characteristics such as user expectations and knowledge, motivational characteristics like objectives and needs, and behavioural aspects like user behaviour and skills in the design process. These aspects, too, are also considered to be central to prospective design in work psychology.

Specifying which design process was meant, it was often said that the object to be designed is either a *technical product* (11 participants) or a *human-machine interface* (14 participants). This clearly comes from the focus on HCI when asked for a definition of prospective design.

Few people also incorporated aspects such as considering context factors, including domain experts in the design process, up-to-date research knowledge out of HCI, or listed arguments that motivate the use of the concept.

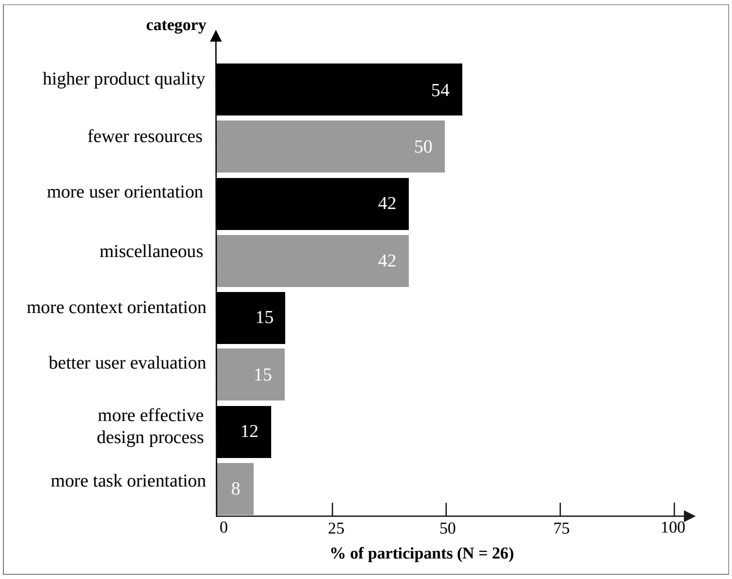

Figure 1: Advantages of prospective design processes

2.3.2 Advantages and Disadvantages of Prospective Design Processes

Most cited as advantages were (see figure 2.3.1):

- Improving product quality (14 participants), for example by avoiding design errors, by enhancing ergonomics or usability (in terms of efficacy and efficiency), or by ensuring intuitive interaction.
- Using fewer resources (13 participants) in terms of time and money - for example expensive redesigns can be avoided.
- Higher degree of user orientation (11 participants) during the design process of the technical product or interface, because cognitive, affective and physiological user characteristics and needs are considered during the entire design process.
- Furthermore, more context orientation, better user evaluation, more effective design process and more task orientation were also stated by some participants.

– We added the remaining non-classifiable contributions to the category miscellaneous.

Mentioned as disadvantages of prospective design processes were (see figure 2.3.2):

– Using more resources (14 participants), such as high efforts and costs, or having a longer time to market.

– Low predictability of future conditions (10 participants). Examples used were the impossibility to fully predict future conditions or user demands and the possibility that user needs can change over time even when involving them in the whole design process.

– Limited capability to generalize (4 participants), because engineers try to achieve the best fit between the product and a specific target user group. However, the product might not necessarily satisfy the needs of other user groups and is therefore of limited generalizability.

– In addition, many other disadvantages are classified together as ”miscellaneous”, e.g. questioning user acceptance, high efforts for the empirical evaluation of user behaviour, etc.

Note that at first glance, defining *using more resources* as a disadvantage seems contradictory, as one of the most cited advantages was *using fewer resources.* They may apply to different phases. Before market launch, more resources may be required for a prospective design process than for a usual design process, for example because user and experts are to be involved and iterative loops are to be integrated in the whole design process. After product launch, however, fewer resources should be necessary to optimize the final product resulting from a prospective design process, because potential human-computer interactions should already have been anticipated before product launch.

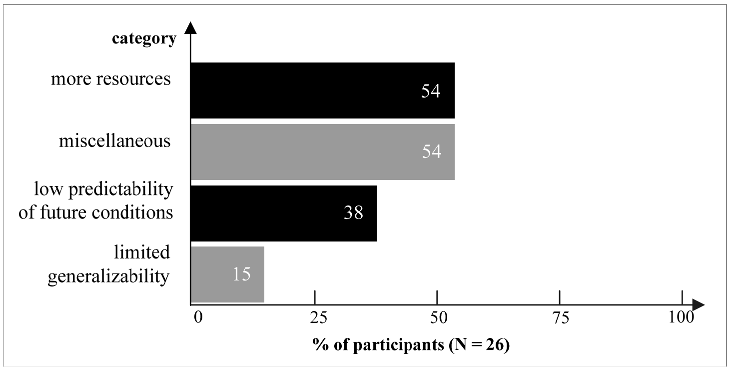

Figure 2: Disadvantages of prospective design processes

2.3.3 Locating Prospective Design in the Product Development Process

When asked how to locate prospective design in a product development process, half the participants thought that prospective design is a concern in almost all product development phases: from the first product idea to the last product evaluation test or beyond. Three participants argued that only single product phases are concerned, such as the conception phase of a product or demand analysis. Ten participants were not sure how to locate prospective design within product development. It is interesting that half the participants also mentioned an iterative approach when thinking of prospective design throughout the product development process. This finding again strengthens the close relatedness of both concepts, user-centred design and prospective design, that we have already indicated earlier in this paper, when we reviewed recent research projects in HCI.

2.4 Conclusions from Different Views: Working Definition of Prospective Design in HCI

Before going on with the working definition of prospective design in HCI, we first sum up the main aspects of the different interpretations of ”prospective design” described above.

In work psychology, the prospective design approach is basically concerned with providing opportunities for workers to develop their personality. This is meant in terms of beneficial work conditions that allow the full expression of their potential. An example is to provide adaptable technology that allows users (as an HCI-equivalent to workers) to adjust task complexity to their needs.

In contrast, prospective design in HCI research is found to be closely related to user-centred design in one part of the active research, but very unspecific for the rest. It is not clear whether opportunities for personality development should be included or not, and what the main differences are to user-centred design.

Most respondents to the questionnaire survey define prospective design in HCI as an anticipatory design approach that incorporates the user (and their cognitive, motivational and behavioural characteristics) in the design process of a technical product or a human-machine interface and anticipates future problems, requirements or opportunities (e.g. intuitive interaction). It is remarkable that this definition integrates aspects of both design strategies from work psychology, i.e. preventive and prospective design. Furthermore, the most frequently mentioned advantages of prospective design processes were achieving better product quality, using fewer resources (e.g. by avoiding expensive re-design) and a higher degree of user orientation during the whole process. The main disadvantages cited were using more resources (e.g. in terms of repeating design iterations involving evaluations), lower predictability of future conditions, and limited generalization due to the involvement of a particular user group. Most participants associated prospective design with almost all phases of the product development process and considered it to be an iterative approach within the product development process.

2.4.1 Working Definition

How are these different perspectives to be integrated? We want to give credit to the historical definition in work psychology, but also want to take into account current trends in HCI research and the opinions of HF experts. Therefore, our definition includes the following aspects:

- We consider prospective design rather to be an approach to designing technical products or user interfaces than an (elaborated) strategy or a (new) methodology.
- We maintain the distinction of work psychology between corrective, preventive and prospective work design, and specify prospective design processes as going beyond traditional corrective and preventive approaches with a special focus on personality development.
- We consider the iterative involvement not only of user factors, but also of task, context and technical system factors to be very important during all phases of the design process. Therefore, we think that an interdisciplinary team of experts dedicated to the four factor classes should support the whole design process. The advantage of an interdisciplinary team lies in anticipating various influences on and forms of human-computer interaction that could arise after market launch, and creating more solutions, or more sophisticated solutions than non-interdisciplinary teams [15].
- In our opinion user involvement is a central issue if products and interfaces are to be designed to fulfil the user's needs, and shall allow for personality development.
- In a prospective design process domain-experts should also be involved whenever the complexity of the engineering task or the specificity of the product or interface requires this.
- We realize an overlap between the concept of prospective design and the concept of user-centred design described as a design philosophy based on the needs and interests of the user [27]. We furthermore see a high similarity to human-centred design as stated in ISO 13407, e.g. regarding concepts such as iteration, multidisciplinary teams and user-involvement. Nevertheless, we consider these elements (i.e. user-involvement, iteration loops during all design phases, an interdisciplinary group of designers taken from user-centred and/or human-centred design) to be of high importance also for prospective design processes.

Therefore we define *"prospective design in HCI"*:

> *Prospective design in human-computer interaction is a particular approach to designing technical products or user interfaces. It offers users the possibility to develop their personality while using the product or interface. A further objective is to enhance product quality and sustainability by anticipating influences on and forms of (un-)desired human-computer interaction as well as properties of the socio-technical system that could become important after market launch. As such, a prospective design process enhances the approach of preventive design by adding the central aspect of developing personality and decreases the need for corrective design actions. These objectives are ensured by using selected principles from user-centred design, which are aligned to serve the central purpose above: (1) constituting an interdisciplinary group of designers to satisfy user, task, technical system, context factors; (2) repeatedly involving users in various stages of the design process; (3) consulting domain experts if necessary due to high complexity or specificity of the product or interface; and (4) considering the above mentioned factors iteratively in any design phase.*

The connection of personality development and interface design could be exemplified as follows: a human-machine system for process supervision and control could offer the possibility to (a) add further information displays and to (b) choose between different degrees of information detail. The former would enable operators to utilize more of their cognitive capability which in turn elevates the chance of improving (or at least stabilizing) this capability. On the other hand the latter would satisfy their needs to consciously monitor the system's state - an operator with high consciousness would prefer detailed information instead of integrated values or overview displays. In both cases, operators could act out what is in their nature.

3 Knowing Performance Shaping Factors Faciliates Prospective Design

Two decades ago Hollnagel [13] stated that "system designers know that human capacity for perception, attention, discrimination, memory, planning, decision making, and action is limited and therefore try to make assumptions which are as reasonable and realistic as possible." (p. 511). But do they really know? As technology becomes more and more complex, it is challenging to know the user and thereby the main requirements of the human-machine system. Teams of engineers and designers try to meet this challenge by gathering data and information on user characteristics and context (like technical equipment, physical and organizational conditions). As in many other settings, laying the proper foundation is a pre-requisite for future success. Concerning the user interface design process Epstein and Beu [10] conclude that mistakes in information gathering or evaluation cause false starts (see for instance the delayed rollout of the Airbus A380). In a sense, expertise in design means avoiding such mistakes by gathering the necessary information (in terms of higher amount and relevance) in order to guide design. Cross [6] offers an extensive overview of how design novices and experts differ. Apparently one core characteristic of experts appears to be more systematic and selective gathering of information. But this does not mean that novice designers are doomed. Design guidelines and standards help in the early design phases since they incorporate the collective design related knowledge of engineering, psychology and ergonomics (among others). For instance, the ISO standard 13407 [1] offers support in software ergonomics and includes the characteristics, tasks and environment of users as important for design decisions. Standards and guidelines contain this information and directly connect it to design recommendations. Nevertheless using them does not guarantee successful design of human-computer systems, since the knowledge must be adjusted to the specific case. Some recommendations could be too specific to be useful and especially in very early design phases the use of comprehensive standards could be too time-consuming and constrain creative solutions. So how can designers (novices and experts alike) be made aware of factors that systematically influence the users' performance? Performance Shaping Factors (PSF) seem to be an answer since they comprise many conditions that influence performance without being too specific and consequently should be included in design considerations. In the following section we discuss common PSF models as well as their advantages and shortcomings. After that we will introduce an alternative model which tries to counterbalance the aforementioned disadvantages.

3.1 Theory of Everything vs. Useful Support for Designers: Discussion of PSF-Models

Especially in safety critical domains, researchers and practitioners are concerned about the pre-assessment of failures and accidents in human-machine systems. The technical perspective of such evaluation is covered by a Probabilistic Risk Assessment (PRA), which in essence is a selection of techniques to quantify and assess the potential for failure, reducing the probability of accident risks (on the distinction between this and resilience engineering see [35]). The human factors perspective in turn is represented by a so called Human Reliability Analysis (HRA). This approach stems from the experience gained in the nuclear industry in the late 1970s (see Three Mile Island, 1979) and models actions of human operators in order to estimate the probability of human error [12]. There are several HRA methods which will not be discussed in this article since only their commonality is of interest: all methods include PSF (since these are necessary for the probability evaluation of human error). We think that especially in early design phases the mere awareness of such factors that influence performance can be more valuable than the concrete and complex calculation of error probabilities as in HRA.

In accordance with the development of HRA methods in the 1980s and 1990s many structured collections of PSF evolved. One of the most prominent examples is the model by Swain and Guttmann [38], which however reveals one of the severest stumbling blocks: the proper stop criterion. Since behaviour is determined by multiple causes one could (in theory) trace these all the way back to the "Big Bang". The authors for instance mention hunger and thirst as factors influencing human actions. In our opinion the benefit of such factors for design (as well as for its original purpose, the human error rate prediction) is not apparent. One could reply that in some settings these variables are relevant (e.g. in military combat missions where rations are limited). But this leads us to another problem of PSF models - they are too general and therefore difficult to apply in concrete settings. A theory of everything is useless if designers want to benefit from the knowledge of factors that influence user performance (like completion time or errors). Swain and Guttmann [38] collected the PSFs more than 20 years ago which may explain why factors linked to automation and function allocation (such as compliance and reliance) were not included. In times of omnipresent automation the problems concerning user interfaces have to be further differentiated and not be met by just including "man-machine factors".

Roughly ten years later Bubb [4] introduced another, more process oriented view on PSFs, where direct and indirect influences are responsible for variability in human performance. Related to accidents, this view is based on the notion of latent conditions by Reason [29]. Furthermore the influences on performance are subdivided into external and internal PSFs which are further clustered: external into organizational and technical pre-requisites and internal into capacity and readiness of the human interacting with the technical system. Although this subdivision approach seems clear and usable, one important aspect is missing: an intermediate cluster which is used as an interface between external and internal PSF. Situation awareness (SA) for example is commonly defined as an influence on human performance [9]. But where does SA belong to? Is it an external vs. internal factor? Salmon et al. [33] reviewed several SA models and picked up this question by referring to the perceptual cycle model of SA by Smith and Hancock [36]. This model proposes that SA resides through an interaction of person and environment. These kinds of interaction (which are understandably common in human-computer *interaction* research) have to be modelled by PSF approaches. Another shortcoming of Bubb's model: process (or problem) complexity and the operator's personality are missing. The former could be connected to technical pre-requisites whereas the latter could be linked to readiness (or better be an independent aspect within the cluster of internal PSF).

To sum up, we define three main rules for constructing a proper PSF model to support designers of human-computer systems. Rule 1: Define an appropriate stop criterion. One criterion could be chance for manipulation. If a factor cannot be manipulated at all it has nothing to do with design and should not be included in the model. Rule 2: Be domain specific. If PSFs are too general they are useless for design. For example one could concentrate on factors that influence performance in supervisory control. Rule 3: Define appropriate clusters. The clusters should be linked to the knowledge of designers and engineers: for instance sophisticated information processing models do not guarantee the awareness needed to include PSF in design decisions. Sub-division of clusters must not be exaggerated in order to offer hands-on support for designers. Especially for interaction between human and computer/ machine intermediate clusters are needed. We tried to take these rules to heart in order to support prospective design by introducing a new PSF model which is described below.

3.2 The Warnemuende Model of Performance Shaping Factors

The aforementioned PSF model is named after a German town on the Baltic Sea, where it was discussed with human factors experts during a conference. This "Warnemuende Model" represents an integrative framework of clustered factors influencing operator performance. Our goal was to offer developers a comprehensive framework of operator performance in the context of supervisory control. Sheridan [34] describes a supervisory control environment as follows: "One or more human operators are intermittingly programming and continually receiving information from a computer [...] The human may remain as a supervisor, or may from time to time assume direct control, or may act as supervisor with respect to control of some variables and direct controller with respect to other variables." A three-fold model of influences is proposed, in which personal factors (for instance traits, states of the operator), system factors (characteristics of the system to be controlled), and environmental factors (external influences like noise, light, etc.) are used to classify influences found in a broad literature review. Designers as well as researchers can use this model as a basic underlying framework whenever operator performance shall be enhanced by a prospective approach, or if the design and discussion of experiments are questioned.

The initial list of factors from the literature was evaluated by a number of experts from various fields during a workshop in April 2008 working in interdisciplinary groups. Factors were classified and the relations to operator performance were discussed together with inner- and inter-class influences among single factors. The collection of a reasonably large number of factors for an acceptable determination of operator performance and the initial classification was of highest priority in the first part of this work.

Of course, this framework is not exhaustive as the number of factors influencing operator performance is practically unlimited. The model is restricted to assistance systems that leave the operator concerned with supervisory control tasks (in the phases: activation, preparation, supervision, decision-making, control and learning). These restrictions show that we stuck to the rules for constructing a proper PSF model (see section 3.1): we defined a stop criterion, namely the manipulability of the variables, we restricted our framework to variables influencing supervisory control and we clustered the different variables into three clusters that are plausible and easy to access (see figure 3.2). The value of our model compared to other

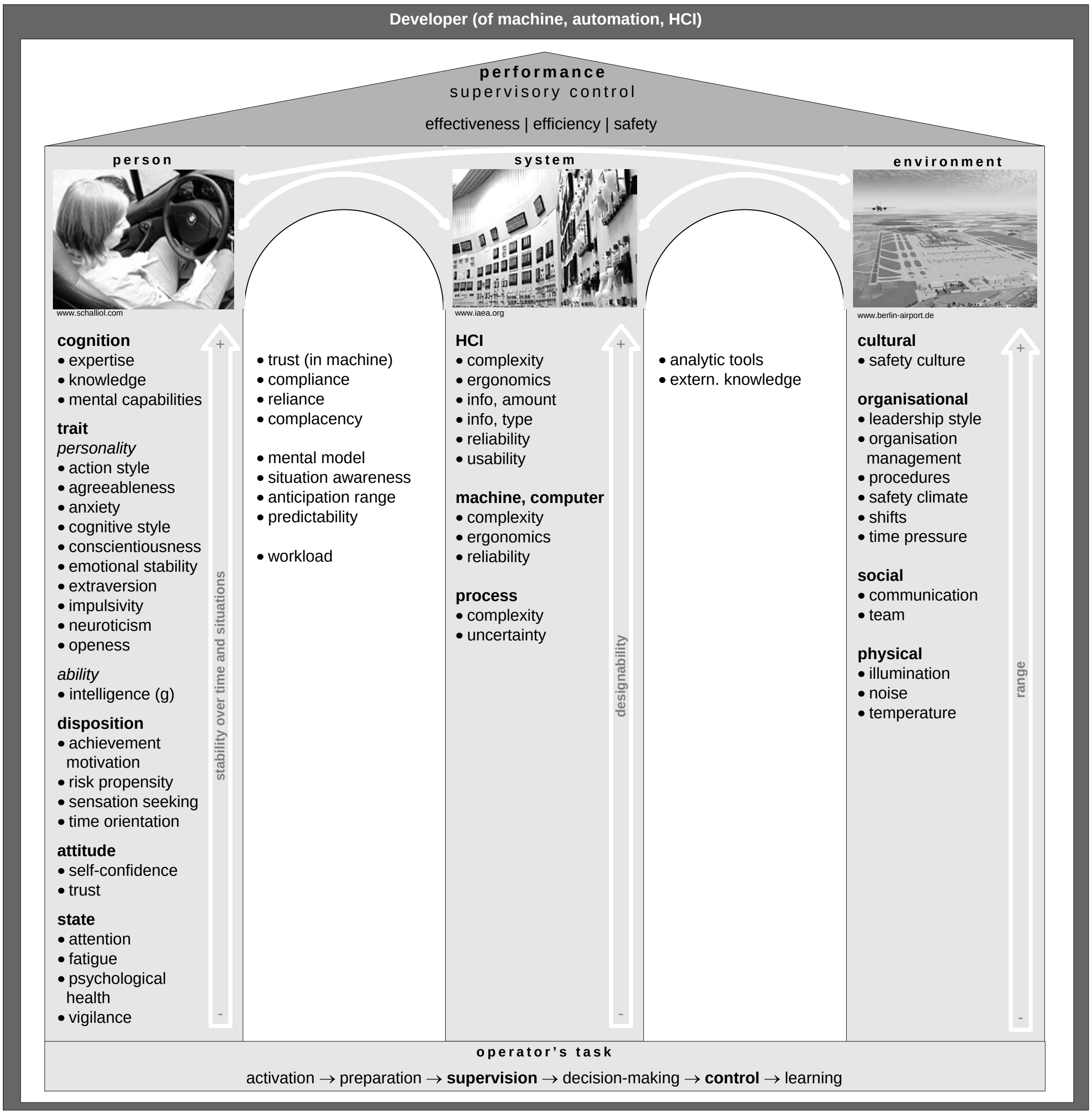

Figure 3: The Warnemuende Model of Performance Shaping Factors

models is that rather than only listing numerous performance shaping factors, we tried to sort these into useful clusters, including a rationale for sorting within each cluster. In the following, the structure of our model and of performance shaping factors is described in more detail.

Personal factors The first cluster contains personal variables influencing performance in supervisory control settings. The variables in this cluster are arranged along the dimension "stability over time and situations". Reviewing the literature, personal factors have always been distinguished as being states vs. traits. States represent personal variables that vary across situations and time whereas traits refer to stable characteristics that are relatively independent of contextual factors. The sub-cluster "state" is at the bottom, followed by the slightly more stable sub-clusters "attitude", "disposition", "trait", with the stable "cognition" at the top. (Expertise and knowledge are only representatives of the latter if long-term and highly trained users in supervisory control settings are addressed).

To give an example, one of the variables in the sub-cluster "trait" is cognitive style. Cognitive styles, representing the constancy of a person to process cognitive stimuli in a particular way [39; 18], can themselves be classified by dimensions like serial vs. holistic style, concerning the type of information processing [31]. Whereas holists scan a large amount of information for characteristic pattern, serialists assess rather less information to test a specific hypothesis. Torenvliet et al. [39] found that operators with a holistic style were better at detecting faults in a process control micro world setting than their serial colleagues. The authors show, that the effect of cognitive style clearly depends upon the interface at hand. Furthermore different results are expected in relation to different operators' tasks. Both (interface and task) should carefully be examined before cognitive style is included as a covariate (as proposed in our performance shaping factors model).

System variables System variables represent the second cluster of factors affecting performance. These are arranged along the dimension "designability", ranging from the sub-cluster "process" that may be more or less complex (an aspect that can hardly be influenced by the designer), to the sub-cluster "machine and computer" up to the sub-cluster "human-computer interaction" which represents the main level of interest to the developer because of its high manipulability.

One important variable in the sub-cluster human-computer interaction is complexity. A classic example of a complex, dynamic system is a nuclear power plant, which has a number of elements with high interaction and change rates. Woods [43] describes four dimensions of complexity: the measure of dynamics, the number of interacting elements, the measure of uncertainty, and the measure of risk. In an automated system these dimensions may be characterized by the following aspects: the rate of change of parameters, the number of tasks to be completed concurrently (single, dual or multiple task) and their degree of interaction, the expected outcome of a decision made by the operator, and the reliability and consistency of the system. For instance, complexity is high when the operator is confronted with simultaneous tasks. Especially when the parameters of interdependent tasks are displayed successively one at a time, then the complexity for the operator increases. The number of tasks to be accomplished, the time pressure caused by high change rates and the diverse parameters to be controlled lead to a high workload and to a deterioration of the operator's performance [7; 20].

Environmental variables The third cluster consists of environmental variables that vary according to the actual context. These variables are organised in terms of the "range" of influence from the direct ("physical") to the social environment, the organisational environment, and finally to more abstract factors like cultural aspects, for example safety culture. So the influences on performance can be imagined from physical conditions such as noise and lighting levels to organisational culture, although they cannot all be adjusted to the same extent.

Bridges We outlined the internal structures of personal variables, system variables and environmental variables, but what are the relations between these clusters? And what about variables, such as "situation awareness" or "trust" that are known to play a crucial role in performance in human-computer interaction but that have not been mentioned so far? These variables can neither be assigned solely to the personal cluster nor to the system cluster. There have to be some "bridges" between the clusters that contain performance shaping factors that in turn depend critically on two underlying factors.

Situation awareness and anticipation are examples of concepts linking personal variables and system variables. These should be taken into account as important factors affecting performance, especially in complex environments where the knowledge of the current and future state of a system is essential. Situation awareness involves "the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" [9], and its link to operator performance seems to be inconclusive. Anticipation, the ability to predict future developments, represents the third level of Endsley's situation awareness model. Anticipation in contrast to situation awareness is supposed to have an undisputable high impact on performance [8] and seems to explain a major part of variance in accidents [42].

The clusters containing system variables and environmental variables can also be linked by variables like analytic tools or external knowledge. One could certainly link the environmental cluster and the personal cluster as well, but we do not consider this bridge as a central part of the presented framework dealing with Human Factor issues. We would categorize interactions between humans and the organisation rather in terms of organisational psychology and interactions between humans and physical conditions in the broader fields of labour science or medicine.

To sum up, the Warnemuende Model illustrates how many aspects the designer should at least be aware of (and if possible take into account) when trying to conceptualize and develop a human-computer system. Our model tends to give a structured overview over variables that are important to the designer and that are accessible and can be manipulated in the specific context. As such it should serve as a kind of checklist for developers and researchers into human factors.

3.3 Application / Use Case of the Warnemuende Model

The Warnemuende Model serves as a checklist to scientists and designers to remind them of possibly relevant factors they should take into account in their study or design process. However, this does not mean that all factors of the model have to be included, but that they are offered at least as covariates (instead of independent variables) in order to explain variance in the data. For instance, research results in alarm psychology are a good illustration of the crucial effect the interaction of personal, system and environmental variables can have on behaviour and consequently on performance of operators interacting with an automated system. Nowadays, operator tasks have metamorphosed from active controlling and intervening to supervisory control tasks [34] which require the operator to monitor an automated system and to intervene whenever a system fault occurs. This operator task is usually supported by alarm systems which are used to alert an operator in the case of critical events. However, the effectiveness of alarm systems directly depends on operators' responding to them in a rational and responsible way. This responsible manner of handling alarms is most challenging due to the complexity of the system's underlying components and unreliable alarm systems.

The Warnemuende Model serves as a checklist to scientists and designers to remind them of possibly relevant factors they should take into account in their study or design process. However, this does not mean that all factors of the model have to be included, but that they are offered at least as covariates (instead of independent variables) in order to explain variance in the data. For instance, research results in alarm psychology are a good illustration of the crucial effect the interaction of personal, system and environmental variables can have on behaviour and consequently on performance of operators interacting with an automated system. Nowadays, operator tasks have metamorphosed from active controlling and intervening to supervisory control tasks (Sheridan, 1992) which require the operator to monitor an automated system and to intervene whenever a system fault occurs. This operator task is usually supported by alarm systems which are used to alert an operator in the case of critical events. However, the effectiveness of alarm systems directly depends on operators' responding to them in a rational and responsible way. This responsible manner of handling alarms is most challenging due to the complexity of the system's underlying components and unreliable alarm systems.

In one of our studies we simulated an operator workplace in a chemical plant. Seventy-four subjects had to complete two tasks simultaneously, i.e. a monitoring task and an ordering task (a simple arithmetic task). In the monitoring task, participants had to supervise a chemical process with different reaction chambers (containers). For this purpose, fuzzy images representing the container content were shown with a mean display rate of 7.5 containers per minute. This task was supported by an automated alarm system. Together with the fuzzy container image, an alarm signal was displayed below the monitoring screen indicating whether the chemical product in the container was in specification (green light) or if the molecular weight in the container was too high (red light). In addition, an alarm state monitor supplied detailed diagnosis with respect to the current container label (e.g. "Molecular weight in container rt-22 too high" or "Container ll-41 OK"). Participants then had a 5 seconds reaction period to respond to a given alarm ("red light") by choosing one of three response alternatives: (1) immediately comply with the alarm and request appropriate treatment by clicking on the "rework" button (compliance), (2) ignore the alarm and wait for the next container ("cry wolf effect), or (3) validate the alarm by inspecting the raw process data using the "check" button, before making a decision whether or not to comply with it. In order to use this option, they had to open a menu using the "check"-button and select the container they wanted to check from a list. Thereby, the fuzzy container image got augmented with process data that is necessary to exactly assess the actual state in the actual container. If the operator did not respond within the 5 seconds period, the container passed the inspection unchanged, and the fuzzy image for the next container appeared. In response to a "green light", however, also three alternatives were possible in general: (1) rely on the validity of the "green light" (reliance), (2) discard the green light and request treatment of the container (i.e. respond as if the green light would represent a miss), or (3) validate the alarm by inspecting raw process data.

The secondary task was a simple ordering task that required participants to subtract the amount of stocked chemicals ("available amount") and the amount of chemicals required for process operation ("required amount"). To preserve sufficient amount of chemicals in stock, used material had to be refilled individually by ordering supply.

The main research question in the study was how participants' tendency to react directly to, ignore or check the alarm was affected by different levels of alarm reliability. Four levels of alarm reliability for each type of system diagnosis (alarm / no alarm) had been investigated.

So how could the Warnemuende model of be of use in this scneario? We used it to check for relevant performance shaping factors in our setting. For instance it names "reliability" as one important aspect of the machine or computer. Alarm systems differ in their level of alarm reliability, represented as the probability that, given an alarm, there is really a critical incident in the system (the so-called "a posteriori probability"). The reliability of alarm systems is mainly determined by technical settings but the designer can set the response criterion, i.e. determine if the alarm is liberal, resulting in higher hit-rates at the cost of higher false alarm-rates, or if it is conservative, meaning fewer hits, but fewer false alarms [37]. This resulting alarm reliability in turn seems to influence operator trust in the system (stressing the importance of the bridge between the clusters "system" and "person"). The level of trust finally affects operator's performance depending on its appropriateness (mistrust or too much trust lead to a deterioration of performance, whereas appropriate trust can enhance performance; [17]).

Another glance at the Warnemuende model calls attention to the organisational factor ”time pressure” in the environmental cluster. In the context of a control room of a chemical plant, operators must make quick decisions and reactions to avoid severe system failures (e.g. an uncontrolled and critical rise in temperature). Thus, it seemed appropriate to include time pressure in the experimental design to increase the external validity of the research results.

Further, the proposed framework involves personal variables in the context of decision-making. Depending on individual dispositions, subjects might vary in their tendency to prioritize speed over accuracy (resulting in less checking behaviour and accepting more inappropriate reactions in the monitoring task but at the same time accomplishing more in the secondary task) or to prioritize accuracy over speed (resulting in more appropriate behaviour in the monitoring task, but accomplishing less in the secondary task). Consequently, the study used a short version of the Big Five-Inventory (BFI-10) [28] and a questionnaire about action strategies and expertise (Fragebogen zur Erfassung von Strategie und Expertise in Experimenten-Revision, FESE-R; English: Questionnaire for the Assessment of Strategy and Expertise in Experiments-Revised) [25; 26] as relevant personality factors as a covariate. The BFI-10 contains two items to each of the five personality traits: extraversion, neuroticism, agreeableness, conscientiousness and openness to experience. The FESE-R investigates action strategies, such as an individual trade-off between speed and accuracy. Some people may vote for a speeding strategy, meaning that they will not check raw data very frequently but will try to compensate the lack in accuracy by saving time and treating more items. Others will take their time to check every item accurately so as to avoid making mistakes.

Putting it all together, the results indicate that behavioural tendencies in the context of checking behaviour show no evidence for a cry-wolf effect, i.e. subjects kept on checking the raw data behind the alarm even with very low levels of alarm reliability. It seems that the alarm is only completely ignored as an extreme reaction when the operator is forced to choose between two mutually exclusive options to act (to ignore the alarm or to comply to it). On the other hand, in alarm-free trials, checking behaviour was even more strongly influenced by even subtle changes of misses of the alarm system. These results show that misses have a far greater impact on behavioural effects than do false alarms. Moreover the experimental set-up is an approach to support the distinction between the concepts of reliance and compliance proposed by Meyer [23]. In addition, a moderating effect of personality traits (see above) was not found, so the results are robust in this regard [22]. This illustrates that the effect of performance shaping factors clearly depends on setting (lab vs. field) and task (e.g. supervision vs. control) but offers a guideline for planning and interpreting a study. This guideline may not be restricted to HCI designers but should of use for experimenters (in a sense another form of designers when planning a study) as well.

4 Conclusions & Discussion

In this paper we present an innovative framework for designers to support prospective design of HCI, as well as for researcher to get an overview of important performance shaping factors that may influence their empirical results. The Warnemuende Model comprises three clusters influencing operator performance. As there is no common understanding of what is meant by prospective design, we firstly approached the term ”prospective” based on three different perspectives: a theoretical one, based on the findings of early work psychology; a research-oriented one, where we reviewed current projects dealing with prospective design; and an empirical one, gathering experts’ opinions on the issue.

Analyzing the three perspectives we found no consistent definition of prospective design. However, ”user-centred design” was mentioned more often than other aspects. For our definition we therefore adopted the basic principles of user-centred design, such as iteration, multidisciplinary teams and user-involvement. However, the distinction between user-centred design and prospective design remains, to some extent, vague. Some argue that the distinction lies particularly in the assignment to different stages in the product development cycle. In this context, prospective design is only associated with early design phases, whereas user-centred design spans the whole development process. Others refer more to work psychology where the emphasis of ”prospective” is rather on developing the own personality at the workplace. Of course, the original term encompasses the entire work environment. The area of human-computer interaction only represents a small section in this field, so that the transfer is negotiable.

As we mentioned in the definition, a prospectively designed interface facilitates personality development. However, the potential for users to develop their personality depends on the tasks and the specific interfaces used in HCI (which are mutually influencing each other). Consequently, a change of the interface already results in a different task to be accomplished. According to Hacker (1986), hierarchically organized criteria already exist for task analysis and job evaluation (e.g. harmlessness, avoidance of interferences or distraction). We emphasize that fulfilment of these criteria has to be taken as a prerequisite if personality development is to be achieved. Therefore, if these conditions are met, the objective of prospective design can be considered the most challenging in HCI design.

After clarification of the meaning of prospective design we proposed a new framework for designers and researchers. In contrast to other PSF-models we assigned all factors to three main clusters (personal, system, environmental factors) and then ordered them again within the clusters according to design-relevant aspects, i.e. stability over time, designability, and range. This helps users of our model (designers and researchers alike) to specify for their task, which factors should be considered, and which are accessible and can be manipulated.

We demonstrated a case where the Warnemuende Model constitutes the foundation for identifying the relevant factors (i.e. system variables, environmental variables and personality traits) to be considered and controlled for planning and analysis of an experimental study. Researchers should always carefully choose those factors proposed by the Warnemuende Model that are closely related to the actual research subject and that are supposed to explain variance in their application (at least as covariates). With this work, we hope to offer designers and researchers in interdisciplinary work settings a shared knowledge base in order include relevant factors into the design process and to expand their horizon for the better anticipation and avoidance of potential problems.

Finally, we encourage researchers and designers to apply the proposed framework. We appreciate any form of feedback on missing factors or relations. Arguments that promote the discussion of prospective design in HCI are also welcome. Please refer to the url http://www.operator-performance.com for updates of the Warnemuende Model, a broad literature reference to all factors and general discussion.

References

[1] ISO 13407. *Human-centred design processes for interactive systems.* International Organization for Standardization, 1999.

[2] B. Bergmann. *Kompetenzentwicklung '96: Strukturwandel und Trends in der betrieblichen Weiterbildung.* Waxmann Verlag, Münster, 1996.

[3] G. Boy and Y. Barnard. Knowledge management in the design of safety-critical systems. In D. Schwartz, editor, *Encyclopedia of Knowledge Management*, pages 389–397. Idea Group Inc., Hershey, 2005.

[4] H. Bubb. Human reliability in system design. In IEC/TC 56/WG11 Committee, editor, *Application Guide to Human Reliability - Part 1*. 1994.

[5] J. M. Corbett. Prospective work design of a human-centred cnc lathe. *Behaviour & Information Technology*, 4(3):201–214, 1985.

[6] N. Cross. Expertise in design: an overview. *Design Studies*, 25(5):427–441, 2004.

[7] S. R. Dixon, C. D. Wickens, and D. Chang. Mission control of multiple unmanned aerial vehicles: A workload analysis. *Human Factors*, 47(3):479–487, 2005.

[8] S. M. Doane, Y. W. Sohn, and M. T. Jodlowski. Pilot ability to anticipate the consequences of flight actions as a function of expertise. *Human Factors*, 46(1):92–103, 2004.

[9] Mica R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37:32–64, 1995.

[10] A. Epstein and A. Beu. Design of a graphical user interface for process control based on the example of a paper recycling plant. *International Journal of Human-Computer Interaction*, 12(3):387 – 400, 2000.

[11] F. Flemisch, J. Kelsch, A. Schieben, J. Schindler, C. Löper, and J. Schomerus. Prospective engineering of vehicle automation with design metaphors: Intermediate report from the h-mode projects. In M Rötting, G. Wozny, A. Klostermann, and J. Huss, editors, *7th Berlin Workshop Human-Machine Systems*, Berlin, 2007. VDI Verlag GmbH.

[12] B. Hallbert, D. Gertman, E. Lois, J. Marble, H. Blackman, and J. Byers. The use of empirical data sources in hra. *Reliability Engineering and System Safety*, 83(2):139–143, 2004.

[13] E. Hollnagel. The design of error tolerant interfaces. In *The First International Symposium on Ground Data Systems*, pages 511–514, Darmstadt, Germany, 1990.

[14] C. G. Janson. Retrospective and prospective design and data. *European Child & Adolescent Psychiatry*, 5(0):11–17, 1996.

[15] C. Krinner and S. Henkel. Entwicklung von Assistenzkonzepten unter verschiedenen ressourcenreichen Bedingungen. In M Rötting, G. Wozny, A. Klostermann, and J. Huss, editors, *7th Berlin Workshop Human-Machine Systems*, Berlin, 2007. VDI Verlag GmbH.

[16] W. Lauer and K. Radermacher. Simulation - a valuable tool for the prospective usability-assessment of complex or-components. In M Rötting, G. Wozny, A. Klostermann, and J. Huss, editors, *7th Berlin Workshop Human-Machine Systems*, Berlin, 2007. VDI Verlag GmbH.

[17] J. D. Lee and K. A. See. Trust in automation: Designing for appropriate reliance. *Human Factors*, 46 (1):50–80, 2004.

[18] D. Lev, E. Hershkovitz, and E. Yechiam. Decision making and personality in traffic offenders: A study of israeli drivers. *Accident Analysis & Prevention*, 40(1):8, 2008.

[19] G. Leveson, N. Safety as a system property. *Communications on the ACM*, 38(11):146, 1995.

[20] B. R. Levinthal and C. D. Wickens. Supervising two versus four uavs with imperfect automation: A simulation experiment. Technical report, University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation Human Factors Division, 2005.

[21] W. Liebert, W. Bender, and J. C. Schmidt. Prospektive Gestaltung von Wissenschaft und Technik - Umgang mit Fragilität und Ambivalenz . In *Technik in einer fragilen Welt. Die Rolle der Technikfolgenabschätzung*, Berlin, 2004.

[22] C. Meyer, J. Nachtwei, and S. Kain. Der moderierende Effekt von Persönlichkeitsmerkmalen auf die Leistung von Operateuren bei der Prozessführung und -überwachung in komplexen Mensch-Maschine-Systemen. In *8th Berlin Workshop Human-Machine Systems*, Berlin, 2009.

[23] J. Meyer. Conceptual issues in the study of dynamic hazard warnings. *Human Factors*, 46(2):196–204, 2004.

[24] J. G. Morrison, D. Cohen, and J. P. Bluckman. Prospective principles and guidelines for the design of adaptively automated crewstations. Technical Report NAWCADWAR-93031-60, Human Factors Technology Development Branch, 1993.

[25] J. Nachtwei. *Ein experimenteller Ansatz in der Kognitiven Ergonomie: Personenmerkmale als Kodeterminanten der Komplexität einer belebten Mikrowelt.* Vdm Verlag, Saarbrücken, 2008.

[26] J. Nachtwei. Testhandbuch zum Fragebogen zur Erfassung von Strategie und Expertise in Experimenten, revidierte Fassung. Technical report, Humboldt-Universität, 2008.

[27] D. Norman. *Psychology of Everyday Things.* Basic Books, New York, 1988.

[28] B. Rammstedt and O. P. John. Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of Research in Personality*, 41(1):203–212, 2007.

[29] J. Reason. *Human Error.* Cambridge University Press, 1990.

[30] M. Rentzsch and D. Seliger. Prospektive Gestaltung der Mensch-Maschine-Interaktion am Beispiel schienengeführter Triebfahrzeuge im grenzüberschreitenden Verkehr. In M Rötting, G. Wozny, A. Klostermann, and J. Huss, editors, *7th Berlin Workshop Human-Machine Systems*, Berlin, 2007. VDI Verlag GmbH.

[31] R. Riding and I. Cheema. Cognitive styles - an overview and integration. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 11(3):193 – 215, 1991.

[32] M Rötting, G. Wozny, A. Klostermann, and J. Huss. *Prospektive Gestaltung von Mensch-Technik-Interaktion 7. Berliner Werkstatt Mensch-Maschine-Systeme 10. bis 12. Oktober 2007, Berlin*, volume 25 of *Fortschritt-Berichte VDI-Reihe 22.* VDI Verlag, Düsseldorf, 2007.

[33] P. Salmon, N. Stanton, G. Walker, C. Baber, D. Jenkins, R. McMaster, and M. Young. What really is going on? review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, 9:297–323, 2008.

[34] T. B. Sheridan. *Telerobotics, automation, and human supervisory control.* MIT Press, Cambridge, 1992.

[35] Thomas B. Sheridan. Risk, human error, and system resilience: Fundamental ideas. *Human Factors*, 50 (3):418–426, 2008.

[36] K. Smith and P. A. Hancock. Situation awareness is adaptive, externally directed consciousness. *Human Factors*, 37:137–148, 1995.

[37] R. D. Sorkin and D. D. Woods. Systems with human monitors: A signal detection analysis. *Human-Computer Interaction*, 1(1):49–75, 1985.

[38] A. D. Swain and H. E. Guttmann. *Handbook of human reliability analysis with emphasis on nuclear power plant operations*. Sandia National Labs, US Nuclear Regulatory Commission, Washington D.C., 1983.

[39] G. L. Torenvliet, G. A. Jamieson, and K. J. Vicente. Making the most of ecological interface design: the role of individual differences. *Applied Ergonomics*, 31(4):395–408, 2000.

[40] E. Ulich. Gestaltung von arbeitstätigkeiten. In H. Schuler, editor, *Lehrbuch Organisationspsychologie*. Hans Huber, Berlin, 2 edition, 1995.

[41] E. Ulich. *Arbeitspsychologie*. Schäffer-Poeschel, Stuttgart, 6th edition, 2005.

[42] C. D. Wickens and S. R. Dixon. The benefits of imperfect diagnostic automation: a synthesis of the literature. *Theoretical Issues in Ergonomics Science*, 8(3):201–212, 2007.

[43] D.D. Woods. Coping with complexity: The psychology of human behavior in complex systems. In L.P. Goodstein, H.B. Andersen, and S.E. Olsen, editors, *Mental Models, Tasks and Errors*, pages 128–148. Taylor & Francis, London, 1988.