Designing for the home: A comparative study of support aids for central heating systems

J. Sauer [a,*], D.G. Wastell [b], C. Schmeink [c]

[a] Department of Psychology, Rue de Faucigny 2, University of Fribourg, 1700 Fribourg, Switzerland
[b] Nottingham University Business School, NG8 1BB, UK
[c] Institute of Psychology, Darmstadt University of Technology, 64283 Darmstadt, Germany

The study examined the influence of different types of enhanced system support on user performance during the management of a central heating system. A computer-based simulation of a central heating system, called CHESS V2.0, was used to model different interface options, providing different support facilities to the user (e.g., historical, predictive, and instructional displays). Seventy-five participants took part in the study and completed a series of operational scenarios under different support conditions. The simulation environment allowed the collection of performance measures (e.g., energy consumption), information sampling, and system control behaviour. Subjective user evaluations of various aspects of the system were also measured. The results showed performance gains for predictive displays whereas no such benefits were observed for the other display types. The data also revealed that status and predictive displays were valued most highly by users. The implications of the findings for designers of central heating systems are discussed.

1. Introduction

1.1. Designing technical systems for the domestic domain

The design of technical systems for the domestic domain has attracted increasing interest in the human factors discipline over recent years (e.g., Stanton, 1998). This may be due to the increasing proliferation of household devices around the globe, paralleled by increasing complexity (e.g., notably through increasing levels of automation).

While the goals of system design in the domestic domain are in many ways similar to those in a work context, there are also a number of important differences which make the domestic domain rather unique (Benedyk and Minister 1998; Sauer and Rüttinger, 2008). These are notably that, unlike operators in a work context, users receive no formal training, they set their own tasks, and they generally receive no performance feedback from other users. Due to these particularities, the possibilities of modifying user behaviour are much more constrained in the domestic domain than at work. This implies that designing for the domestic domain needs to focus very strongly on system design since alternative options for behaviour modification (such as training and selection) are not available. Technical systems, therefore, need to be designed such that they compensate for these shortcomings by providing system-embedded support devices for the user.

1.2. User support facilities

Operators of human–machine systems can be supported by different types of visual displays. One may distinguish between at least four types of information that can be conveyed by different display types (Kroemer et al., 2001): status information, historical information, predictive information, and instructional information. To this listing, warnings may be added as a fifth display type that directs the users' attention to changes in system states which the user may have overlooked.

Most devices in the domestic domain are limited to status displays. There are also a few (generally more sophisticated) devices that are equipped with simple warning displays (e.g., washing machine informs user that no detergent has been added). The lack of more advanced information systems in domestic technical devices is largely due to the higher costs associated with additional efforts to design more advanced systems. For some domestic technical devices, it may, however, be worth providing more advanced user support due to the potential benefits of it (e.g., energy savings for central heating, more effective cleaning of clothes in washing machines). It is, therefore, important to determine which types of user support facilities are helpful to the user.

* Corresponding author. Tel.: +41 26 3007622; fax: +41 26 3009712.
   E-mail address: juergen.sauer@unifr.ch (J. Sauer).

1.2.1. Status displays

Status displays provide information about the current state of a system. This may be in the form of check displays (i.e., indicating whether a given condition exists or not, such as the engine operates in a safe temperature range or is overheated), qualitative displays (i.e., providing an approximate value in the form of categories such as cold, warm, hot) or quantitative displays (i.e., an exact numerical value is provided such as 18.5 °C). Status displays may also be found in more advanced forms such as integrated polar displays, representing the current values, for example, of eight parameters in the form of an octagon. In the domestic domain, status displays can be found in refrigerators (e.g., indicating current temperature levels) and central heating systems (CHSs) (e.g., indicating current temperature of boiler water). Status displays provide helpful information for monitoring the current system state but are not very supportive during fault finding or for dealing with unfamiliar system situations. While these situations are clearly less frequent in the domestic domain than at work, they may still arise, in particular, during the use of more complex systems such as central heating.

1.2.2. History displays

Compared to status displays, history displays are more advanced since historical information needs to be stored in the system and displayed appropriately to the user. This is usually done in quantitative or qualitative form. The history display is very helpful in supporting the operator in fault finding activities. It may also support the user in evaluating system management behaviour in retrospect. However, it may slow down human response time compared to displays without historical information (Hansen, 1995). There is other work that also indicates that historical information does not necessarily improve human performance (Palmer et al, 1980). In contrast, in a study providing support to users operating a CHS, it emerged that historical information helped users optimise their system management strategies, resulting in reduced energy consumption (Sauer et al., 2007). The mixed findings from the literature suggest a need for empirical evaluations of the utility of history displays as their effectiveness appears to be dependent on task requirements and situational circumstances.

1.2.3. Predictive displays

While history displays only allow the evaluation of chosen system management strategies in hindsight, predictive displays support the operator in evaluating the impact of system interventions on system state. Predictive displays are of particular benefits if there are time delays (i.e., dynamic lags) in receiving feedback about the effects of one's actions, which are typical for process control systems (e.g., chemical plants, heating system) but also for some transportation systems such as large ships (Sheridan, 2002). Predictive displays may also support users of CHS to attain system management goals such as reducing energy consumption by a certain percentage. Providing a simulation of the CHS (i.e., predictive model) has the advantage over a history display that system management strategies can be optimised proactively rather than with the benefits of hindsight on the basis of historical information. An example of a simple predictive display is an indicator on the dishwasher displaying the remaining time left until termination of the cleaning process.

1.2.4. Instructional displays

There may be circumstances under which the operator does not know how to interpret the information available (e.g., due to a poor mental model of system functioning). In that case, quantitative information (be it provided by status, historical or predictive displays) is considered unhelpful because it does not offer specific recommendations about how to change behaviour (cf. Korsgaard and Diddams, 1996). Instead, the user would need explicit advice of how to improve system management strategies (e.g., shut off valve), as it is given by instructional displays. These are similar to command displays in that both give explicit instructions of how to proceed. Research has indicated a number of circumstances under which command displays have provided benefits, such as high time pressure in a decision-making situation (Sarter and Schroeder, 2001). In the context of the CHS operation, there were indications that some users did not have adequate strategies available to manage the system more energy-efficiently although they were aware of their undue energy consumption (Sauer et al., 2007). However, there may also be disadvantages associated with the instructional display. In its purest form, an instructional display does not provide any information about the system state but the system state is interpreted by an algorithm and the implications of that analysis are communicated to the user in the form of instructions. In the domestic domain, instructional displays may be considered an extension of the instruction manual in a dynamic form, providing several benefits compared to static instruction manuals (e.g., instructions are only presented if relevant to a particular situation).

1.2.5. Warning displays

The purpose of a warning is to attract the user's attention to a problem that may have been overlooked. In the domestic domain, this is most typically done in the form of static on-product warnings (e.g., Frantz and Rhoades, 1993). However, these static warnings (e.g., information label attached on product) are less effective than dynamic warning systems (Kroemer et al., 2001) because they are not adaptive to the situation (i.e., they are permanently present whether needed or not). In contrast, dynamic warning systems only issue alarms if there is a critical state. Despite these advantages, dynamic warning systems are quite rare in the domestic domain while they are much more common in work environments. Research on warnings has also revealed that compliance rates are higher for high-severity problems than for low-severity problems (Wogalter and Barlow, 1990). This raises the question of how effective warnings are that inform users about efficiency gains (lower severity) rather than safety-related issues (higher severity).

1.3. Design of CHSs

The present work addresses the user-friendly design of domestic CHS, which represents the most complex system in the domestic domain. The heating system is of great interest to ergonomics for this reason but also because of its considerable environmental impact. The importance of the latter was demonstrated by the work of Verhallen and Van Raaij (1981), which showed that user behaviour accounted for more than a quarter of the variance in energy consumption. Adequate system design including appropriate system-embedded feedback is therefore of great importance.

The effective use of CHS may be described as achieving a self-defined comfort level (i.e., indoor temperature) while simultaneously minimising energy consumption. The goal of merely minimising energy consumption is, from the present perspective, not satisfactory since it could be attained simply by turning down the heating to a minimal level (or even switching it off), which may result in severe discomfort.

There is some evidence for the positive effects of providing users with information about the effectiveness of their usage of

the CHS. Providing residents with dynamic information about their energy consumption was effective in reducing their energy bill whereas the availability of static information did not have this effect (Midden et al., 1983). Another field study demonstrated the positive effects of giving feedback to residents (Van Houweligen and Van Raaij, 1989). Energy savings amounted to about 12% when feedback on energy consumption was provided together with a reference standard. Interestingly, energy savings could not be maintained when the feedback source was withdrawn. In a recent meta-analysis 23 studies have been cited, which evaluated the effectiveness of feedback about domestic energy use (Abrahamse et al., 2005). Most of these studies showed an impact on user behaviour for the various forms of feedback examined (e.g., continuous, weekly, monthly, comparative).

A recent study compared different types of embedded system feedback using a computer-based simulation of a CHS (Sauer et al., 2007). The work showed that feedback resulted in improved environmentally friendly performance and that specific characteristics of the feedback indicator influenced the type of strategy used to improve human–machine system performance. This suggests that the feedback indicators need to be well chosen. Despite the efficiency gains that were made due to detailed quantitative feedback (e.g., energy consumed, energy wasted), this may not have been sufficient to some users because it lacked specific information of how they could improve their system management strategies (Korsgaard and Diddams, 1996). For this group of users, instructional feedback providing behavioural guidance might have been helpful (e.g. turn down bathroom radiator if you go away for the weekend).

1.4. The present study

The review of the scientific literature indicated that previous research on the use of CHSs has focussed on different aspects of feedback. Feedback refers to the provision of historical information as well as readings of the current system state, that is, information provided by history and status displays. The work confirmed the benefits of system-embedded feedback over other forms of conveying information. It also pointed to the importance of having suitable feedback indicators that are helpful in improving user performance. The present work will extend these issues to other forms of user support. These are support facilities that are commonly used in work systems but are hardly found in domestic systems. It is likely that with deceasing costs of advanced support facilities, these will increasingly enter the domestic domain. This requires research to assess the implications of the different characteristics of the domestic domain compared to the work domain (outlined at the beginning of this article) for designing support facilities.

The primary goal of this study is to empirically evaluate different forms of user support with regard to their effect on improving the efficiency of CHS operation. The most central parameter of operational efficiency is energy-efficient usage. A secondary goal of the study was to determine whether there were any associations between different person-based factors and user performance.

The methodological approach adopted in this study was based on the micro-world research paradigm (Brehmer and Dörner, 1993). This approach aims to model the complexity of the real world in a computer-based simulation environment. This reflects the difficulty of conducting ergonomics research on real systems (e.g., problems of accessibility, little control over contextual variables, slow rate of process change). The present work made therefore use of a PC-based simulation of a generic heating system. The development of the simulation

environment was guided by a general theoretical framework of micro-world design (Sauer et al., 2000). This helped make sure that the critical features of a CHS are suitably modelled in the simulation.

On the basis of the literature reviewed, it was generally predicted that increasing support quality would lead to improved ecological performance. More specifically, it was hypothesised that more advanced support facilities would lead to a linear improvements in all performance variables (i.e., user comfort, energy consumption, efficiency of energy usage) across the different experimental conditions.

2. CHS simulation

2.1. General features

The CHS simulation (CHESS V2.0) models a generic domestic heating system. It can be set-up in different modes to model different system features and task scenarios. CHESS V2.0 represents an enhancement of a previous version of the simulation environment (Sauer et al., 2007), with additional features being provided that support users during system operation. These additional system features are notably predictive display, instructional display, and warning display.

CHESS V2.0 is an operator-paced simulation, that is, the simulation process can be started and stopped at the user's discretion. The simulation can be activated by clicking upon the "run simulation" button. A complete day takes about 30 s to complete. At the end of the day, the simulation stops, allowing the operator to inspect different displays and to make changes to the heating profile. Alternatively, the simulation can be run over several days, allowing the user to interrupt the process at any time.

For the purpose of user testing, CHESS V2.0 allows the manipulation of four main classes of variables: quality of user support, weather conditions, user lifestyle, and size of accommodation.

Quality of user support. To set-up CHESS V2.0, the experimenter can choose between five different modes, representing different forms of user support: status display, history display, predictive display, instructional display, and warning display. The interfaces of the different displays are presented in Fig. 1.

Weather conditions. Outside weather conditions have an impact on temperature levels in the accommodation and determine the difficulty of heating management. For example, lower outside temperatures make the system management more difficult. Similarly, the higher the degree of variation in temperature levels is, the more difficult system operation becomes. CHESS V2.0 allows for very different temperature profiles to be set-up.

User lifestyle. Users may have different preferences with regard to temperature they find comfortable at home. They may also differ in their overall lifestyle (e.g., presence at home, desired temperature levels). If there is considerable variation in target temperatures across rooms and across days, difficulty of system operation increases. Also, central heating operation is more demanding if the user has highly irregular arrival and departure times. Time of user presence as well as target temperatures can be varied in the set-up file.

Size of accommodation. The set-up of CHESS V2.0 can be modified to model flats of different sizes, ranging from 3 to 6 rooms. Clearly, user demands rise with larger flat sizes due to increasing system complexity.

2.2. Operator tasks

The task of the operator was to manage the heating most efficiently to achieve the comfort levels prescribed by the set predefined target temperature and to minimise energy consumption at the same time. In order to achieve these goals a heating profile for each room needs to be set up. The heating profile contains the times when the burner is switched on and off and the maximum temperature setting of the thermostat. This is done by clicking on a graphical surface and drawing out an appropriate rectangle or, alternatively, by clicking on a dialogue box.

2.3. Operator support facilities

CHESS V2.0 offers several support facilities to the operator. However, their availability depends on the set-up mode chosen by the experimenter. The weather forecast and the status display are available to operators in all set-up modes.

Weather forecast. This facility provides an accurate overview of the outside temperatures on a daily basis over a period of up to 4 days.

Status display. This facility provides current readings of primary system parameters, such as temperature levels, comfort levels, and energy usage (see Fig. 1a). This information is presented for individual rooms.

History display. There are two displays that provide historical information. Both can be called upon by the user. While one of them provides a numerical and graphical representation of the several feedback parameters for the previous 24 h (see Fig. 1b top section) the other one offers a summary of the main parameters over a full month in the form of a table (see Fig. 1b bottom section).

Predictive display. On the basis of environmental conditions (e.g., outside temperature), this facility allows the simulation of different heating set-ups to assess their impact on different parameters. The following parameters are calculated by the system: energy consumption, efficiency of energy use, comfort levels and optimal consumption levels as a benchmark (see Fig. 1c). Energy efficiency was assessed by comparing predicted performance with the optimal one that could be achieved for a given set of target and forecast external temperatures. The latter was computed by using a simple 'hill-climbing' heuristic which
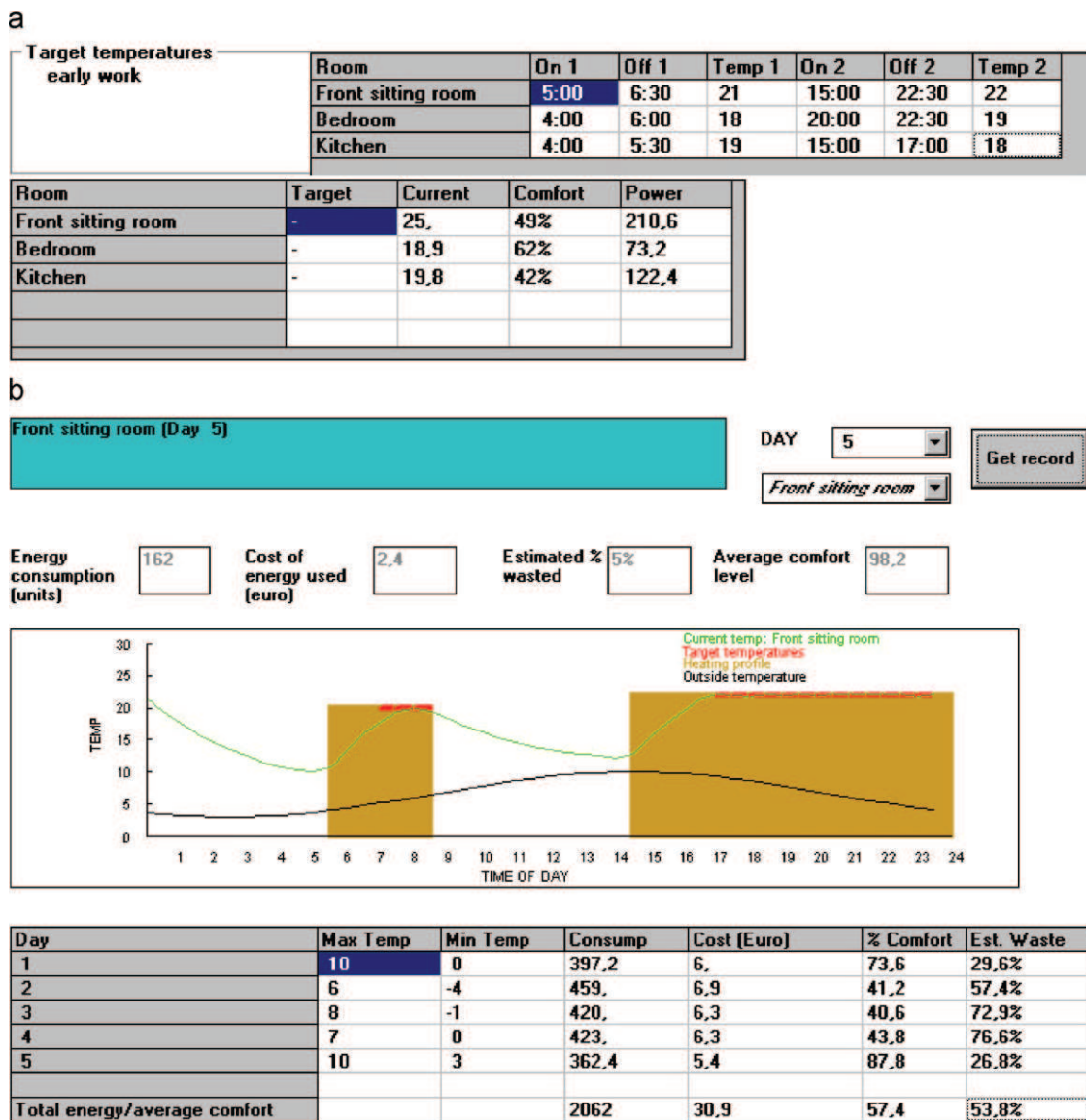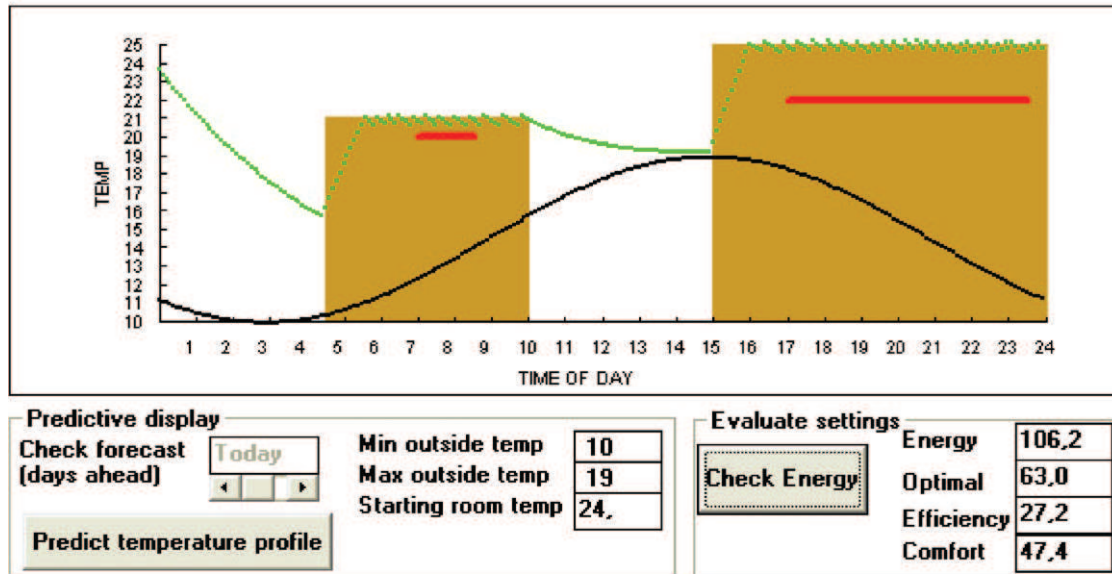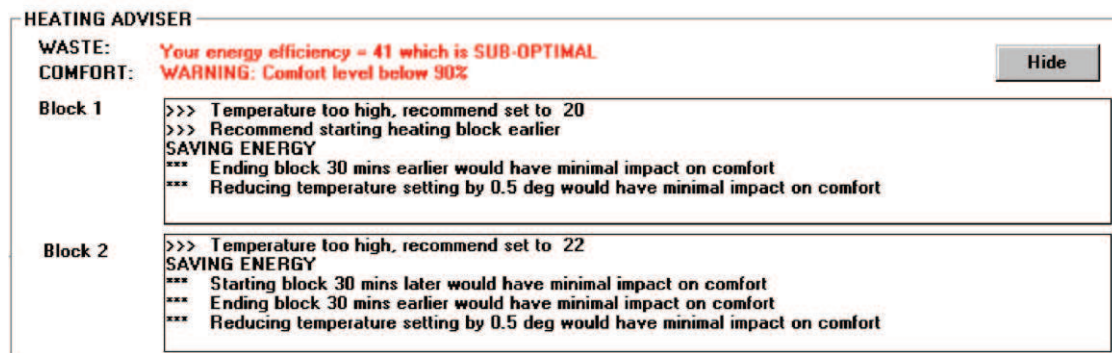

Fig. 1. (a) Status display. (b) History display: daily reports in graphical format separately for each room (top section) and monthly report in numerical format (bottom section). (c) Predictive display. (d) Instructional display. (e) Warning display.
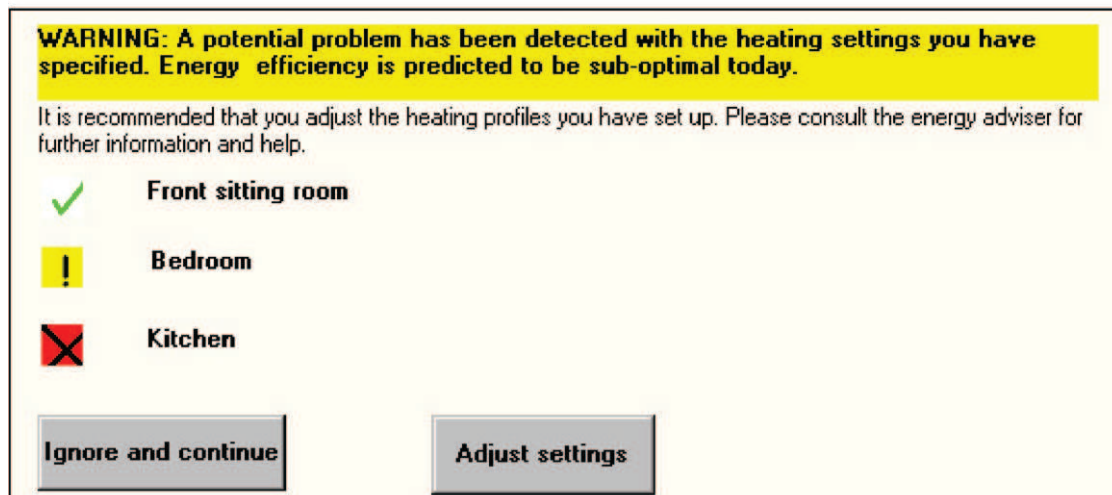
c

Predictive display

Check forecast (days ahead) | Today

◄ | ►

Predict temperature profile

Min outside temp | 10
Max outside temp | 19
Starting room temp | 24.

Evaluate settings

Check Energy

Energy | 106.2
Optimal | 63.0
Efficiency | 27.2
Comfort | 47.4

d

HEATING ADVISER

WASTE: Your energy efficiency = 41 which is SUB-OPTIMAL
COMFORT: WARNING: Comfort level below 90%

Hide

Block 1
>>> Temperature too high, recommend set to 20
>>> Recommend starting heating block earlier
SAVING ENERGY
*** Ending block 30 mins earlier would have minimal impact on comfort
*** Reducing temperature setting by 0.5 deg would have minimal impact on comfort

Block 2
>>> Temperature too high, recommend set to 22
SAVING ENERGY
*** Starting block 30 mins later would have minimal impact on comfort
*** Ending block 30 mins earlier would have minimal impact on comfort
*** Reducing temperature setting by 0.5 deg would have minimal impact on comfort

e

WARNING: A potential problem has been detected with the heating settings you have specified. Energy efficiency is predicted to be sub-optimal today.

It is recommended that you adjust the heating profiles you have set up. Please consult the energy adviser for further information and help.

✓ Front sitting room

! Bedroom

✗ Kitchen

Ignore and continue | Adjust settings

Fig. 1. (*Continued*)

involved progressively adjusting the heating profile until the optimal level of efficiency was obtained.

Instructional display. This display provides explicit advice of how to improve the settings chosen for running the CHSs (see Fig. 1d). This involves verbal instructions of what to do but also includes a qualitative evaluation of important system parameters (e.g., energy efficiency is non-optimal).

Warning display. This display, presented in Fig. 1e, provides a simple tool for checking overall energy efficiency. In one of the conditions (WARN, see below) it was triggered automatically if the system appraised that the chosen heating profile would be less than 90% 'energy efficient' for any of the three rooms. The warning display provides a summary evaluation of the settings chosen for each room separately, using a three-category "traffic light"

system: a green tick indicated that the settings for the room were satisfactory (energy efficiency above 97%), an amber exclamation mark indicated that the settings were acceptable though less than optimal (better than 90% but less than 97%), and a red cross indicated that the settings were seriously out of the acceptable range (less than 90% energy efficiency).

2.4. Experimenter set-up screen

There is a special set-up screen available to the experimenter, permitting the variation of variables such as target temperatures, type of accessible displays and duration of experimental trials.

2.5. Data logging facility

The simulation software produces a results file in which all important information is stored, that is, user interventions (e.g., sampling of instructional display), and system responses (e.g., change in comfort levels). As the precise occurrence of each event is registered by the system clock, CHESS V2.0 allows the recreation of an experimental session for detailed analyses.

3. Method

3.1. Participants

Seventy-five participants (48.0% female) took part in the study. Their ages ranged from 19 to 47 years ($M = 23.7$ yrs). The vast majority of participants were members of the Darmstadt University student population, with most of them reading engineering and pure sciences (55.1%) and social sciences (43.5%). Participants were paid € 15 each for their participation in the study.

3.2. Experimental design

A two-factorial mixed design (5×2) was employed. The independent variable *display type* was varied between participants at five levels: status display (STAT), history display (HIST), predictive display (PRED), instructional display (INST), and warning display (WARN). Table 1 shows the support features that were available under each condition.

The second independent variable *session phase* was a repeated-measures variable being manipulated at two levels: first half vs. second half. Participants were asked to complete a 30-day experimental trial, divided into two phases of 15 days each.

3.3. Measures

3.3.1. User performance

Three measures of user performance were taken: energy consumption, comfort level and efficiency of energy use. Energy consumption measures the amount of energy (in arbitrary units) consumed during task completion. Comfort level is a measure of the degree to which target temperatures are reached (expressed as % of time). Efficiency of energy use refers to the amount of energy (units/comfort level) consumed as a proportion of the comfort level achieved. This index is equivalent to "gallons per mile", which is a commonly used indicator of a car's energy efficiency.

3.3.2. Information sampling and system intervention

Besides these performance variables (i.e., how well the heating is managed), a number of measures of information sampling and system control behaviour were taken (i.e., in what way the system is operated by the user). *Information sampling* referred to the frequency (number of interrogations/day) with which information sources and support facilities were accessed by users (e.g., weather forecast, history display, and instructional display). Accordingly, *system control action* measured the frequency (interventions/day) with which the user made changes to the profile settings.

3.3.3. Subjective evaluation of support facilities

To be able to assess the perceived utility of the support facilities, users were asked to provide subjective ratings on a seven-point Likert scale (1 = low utility, 7 = high utility) for each of them. The ratings were made after the working session had been completed.

3.4. Procedure

The experiment was conducted in a university-based laboratory that was equipped with five personal computers (which were screened off from each other). After a general introduction to the experiment, the participants were given an introduction to the CHESS software (ca. 10 min). They then began the working session with the CHESS task, involving the operation of a heating system for 30 days (ca. 60 min).

The 30-day simulation scenario that participants were asked to complete comprised of a sequence of daily task scenarios that were modelled on typical user lifestyle scenarios. These were, for example, normal working day, long working day, party, day at home, and holiday. Each task scenario had a specific heating profile that the participant had to implement. For example, for one day at home the temperature in the sitting room had to be maintained at 21 °C from 9 am to 5pm and at 22 °C from 5 pm until midnight. For the next day, the temperature profile might be 21 °C from 9 am to 7 pm and 20 °C from 7 pm until 10.30 pm.

4. Results

Given the multiplicity of dependent variables, MANOVA was first performed, with experimental condition (five groups) and session phase as the independent variables. This found that both main effects and their interaction were significant, yielding a Wilks lambda of 32.8 ($p < .001$) for the experimental manipulation. Univariate ANOVA and post hoc LSD-tests were used to locate the specific source of the significant effects.

4.1. Performance

Comfort levels. This measure presents the extent to which target temperatures were achieved (as a percentage), representing an indicator of user comfort. The general picture indicated that

Table 1
Independent variable "display type" and the support features available in each condition

Experimental condition	Status display	History display	Predictive display	Instructional display	Warning display
STAT	✔	–	–	–	–
HIST	✔	✔	–	–	–
PRED	✔	✔	✔	–	–
INST	✔	✔	✔	✔	–
WARN	✔	✔	✔	✔	✔

better user support resulted in higher comfort levels (see Table 2), being confirmed by a significant main effect ($F = 3.79$; df $= 4,70$; $p < .01$). Post-hoc LSD-tests indicated that the conditions WARN, INST, PRED were significantly different from HIST and STAT ($p < .05$). Analysis of variance also showed a significant effect of session phase, with comfort levels being higher in the second half than in the first ($F = 32.5$; df $= 1,70$; $p < .001$). A significant interaction between the two independent variables showed that this increase in comfort levels in the second half was only observed for HIST and STAT but not for the three other conditions ($F = 2.96$; df $= 4,70$; $p < .05$).

Energy consumption. This parameter refers to the amount of energy consumed during task completion. The data are presented in Table 2. Visual inspection of the means indicated that all users had similar energy consumption levels, independently of the level of support provided the different display types ($F < 1$). There were no changes in consumption levels with increasing time-on-task ($F < 1$). No interaction occurred ($F < 1$).

Energy efficiency management. This refers to a composite index, combining the parameters comfort level and energy consumption. Comfort levels are divided by energy consumption with higher values of the index indicating better performance (since less energy is consumed to reach a given comfort level). The data in Table 2 show higher efficiency levels for WARN, INST and PRED than for HIST and STAT. This was confirmed by analysis of variance ($F = 4.13$; df $= 4,70$; $p < .005$; LSD-test: $p < .01$). The analysis also revealed a second main effect, with the efficiency index increasing in the second half ($F = 91.1$; df $= 1,70$; $p < .001$). A significant interaction was also observed. This was due to performance improvements from the first to the second phase being more pronounced for HIST than for STAT and PRED ($F = 3.08$; df $= 4,70$; $p < .05$).

4.2. Information sampling and system intervention

Sampling of daily report (history display). This information source was not available in the STAT condition. The data in Table 3 show that in the HIST condition the daily report was much more frequently sampled than under the other three conditions ($F = 9.94$; df $= 3,56$; $p < .001$). The data also revealed a main effect of session phase, with sampling frequency decreasing over time ($F = 12.21$; df $= 1,56$; $p < .001$). No interaction was observed ($F < 1$).

Sampling of monthly report (history display). As the parameter above, this information source was not available in the STAT condition. As the data in Table 3 show, sampling frequency of this history display was higher for HIST and PRED than for INST and WARN ($F = 4.35$; df $= 3,56$; $p < .01$; LSD-test: $p < .05$). No effect of session phase ($F < 1$) and no interaction was found ($F = 1.70$; df $= 1,56$; $p > .05$).

Sampling of predictive display. This information source was not available in the STAT and HIST conditions. Overall, the data in Table 3 show that this support facility was very extensively used, with sampling rates of around 11 per day (accumulated across all three rooms). The analysis revealed stable sampling pattern across display types ($F < 1$), indicating that the additional availability of instructional and warning displays did not change the sampling frequency. The sampling rate somewhat decreased from the first half to the second ($F = 53.7$; df $= 1,44$; $p < .001$) No interaction was observed ($F < 1$). The rather intensive use of this facility was consistently observed across participants, ranging from 5.0 to up to 22.1 checks per day.

Sampling of instructional display. This support facility was much less frequently used than the predictive display, with 2.1 checks per day on average for the whole flat. There were only small differences in means across the different conditions, with none of them being statistically significant. It is of some interest that usage frequency was very variable across participants ranging from 0 to 11.3 checks per day, with 4 very intensive users inflating the mean (more than 6 checks per day).

Issuing of warning display. Overall, few warnings were issued in the WARN condition, on average once every 5.5 days (i.e., less than one in 15 of the individual room profiles set). The comparatively small number of alarms reflected the attainment of a generally high standard of system management; a warning only appeared if

Table 3
Frequency of information sampling and system intervention as a function of display type and session phase

	STAT	HIST	PRED	INST	WARN	Overall
History display: daily report (no/day)	N/A	2.75	.69	.25	.65	
Day 1– 15	N/A	3.0	.98	.41	.83	1.31
Day 16– 30	N/A	2.51	.39	.08	.47	.86
History display: monthly report (no/day)	N/A	.34	.44	.13	.12	
Day 1– 15	N/A	.30	.43	.12	.17	.26
Day 16– 30	N/A	.38	.45	.13	.08	.26
Predictive display (no/day)	N/A	N/A	11.5	10.3	10.7	
Day 1– 15	N/A	N/A	13.0	12.0	13.2	12.7
Day 16– 30	N/A	N/A	10.0	8.6	8.2	8.9
Weather forecast (no/day)	1.30	1.62	.37	.09	.09	
Day 1– 15	1.60	1.96	.34	.18	.18	.85
Day 16– 30	1.01	1.29	.39	.01	.01	.54
Changes to profile settings (no/day)	2.97	3.21	2.83	2.79	2.85	
Day 1– 15	3.00	3.29	2.46	2.64	2.73	2.83
Day 16– 30	2.93	3.12	3.20	2.94	2.98	3.03

N/A: not available.

Table 2
Performance parameters as a function of display type and session phase

	STAT	HIST	PRED	INST	WARN	Overall
Comfort levels (%)	81.1	77.7	91.6	92.4	92.5	
Day 1– 15	78.7	72.9	90.2	91.2	91.1	84.8
Day 16– 30	83.6	82.5	92.9	93.6	93.9	89.3
Energy consumption (units)	367.0	355.3	372.8	366.5	372.2	
Day 1– 15	363.4	347.6	370.9	370.5	375.2	365.5
Day 16– 30	370.7	362.9	374.7	362.5	369.1	368.0
Energy efficiency management (comfort/consumption)	.222	.216	.245	.251	.249	
Day 1– 15	.217	.206	.241	.244	.243	.230
Day 16– 30	.226	.225	.248	.257	.255	.242

the settings chosen by the user were highly inappropriate (an estimated energy efficiency of less than 90%, see above). No significant differences between experimental conditions were observed.

Sampling of weather forecast. This display provided information about outside temperatures for up to 4 days in advance. The general picture was that sampling frequency decreased with increasing user support facilities (see Table 3). The overall main effect was significant ($F = 20.5$; df = 4,70; $p < .001$). Post-hoc LSD-tests showed that HIST and STAT had higher sampling rates than PRED, followed by INST and WARN with the lowest sampling rates. For the latter two, sampling rates nearly dropped to zero in the second session half. The data also showed a main effect of session phase, with sampling frequency strongly decreasing over time ($F = 12.13$; df = 4,70; $p < .001$). No interaction occurred ($F = 2.31$; df = 4,70; $p > .05$).

Changes to profile settings. This parameter referred to the frequency with which users made changes to heating profiles (see Table 3). Although the analysis revealed no main effect of display type ($F = 1.19$; df = 4,70; $p > .05$), there was a significant interaction between display type and session phase. As the data in Table 3 show, this interaction was due to increases in intervention frequency from the first to the second half for PRED, INST and WARN whereas reductions were observed for HIST and STAT ($F = 2.86$; df = 4,70; $p < .05$). Despite this differential pattern, intervention rates were overall on the increase over the session, leading to a significant main effect of session phase ($F = 4.87$; df = 1,70; $p < .05$).

4.3. Subjective evaluation of support facilities

A one-factorial analysis of variance was carried out on the subjective utility ratings of the different support facilities.

Primary support facilities. The data and the results of the statistical analysis are presented in Table 4. It shows that the *status display* was very highly valued by users across all conditions (i.e., no significant difference between conditions), with an average rating of $M = 6.2$ out of a maximum rating of 7. Even when advanced support facilities were available, this did not reduce its perceived utility. The pattern was clearly different for the *history display*, which showed very high ratings only for the HIST condition but its perceived utility decreased significantly with the availability of more advanced support facilities such as the predictive display. The *predictive display* attained ratings in excess of six points in all three conditions in which it was available. In contrast, the instructional display was considered to be much more useful in the INST condition whereas the additional presence of the warning display reduced its perceived utility. The warning display was only evaluated in one experimental condition, with its ratings being the lowest of all display types.

Examining the display evaluations in the condition WARN (i.e., the only condition in which all displays were available) revealed that the predictive display enjoyed the highest value together with the status display while all other displays had considerably lower ratings.

Weather forecast. In addition to these primary support facilities, a support system providing a weather forecast was also available. The data in Table 4 show that the perceived utility of the weather forecast significantly decreased with the availability of more advanced user support facilities. This perceived utility was in line with the objective sampling rate (see Table 3), which was confirmed by a strong positive correlation between the two parameters ($r = .50$; $p < .001$).

5. Discussion

The general pattern suggested that the predictive display led to a significant improvement of ecological performance while no such improvements were observed for any of the other display types. The benefits of the predictive display were not only indicated by the performance data but also by user ratings, which showed that the predictive display was highly valued by users, independently of the availability of other support facilities. While the status display was equally highly valued by users under all conditions, there was a general trend for the other display types to lose some degree of importance as soon as a more advanced display was made available.

For all display types, performance improvements were observed from phase 1 (i.e., days 1–15) to phase 2 (i.e., days 16–30). Interestingly, these changes were limited to comfort levels and energy efficiency while consumption levels remained stable. While practice effects are a common occurrence over the task completion period (e.g., Patrick, 1992), the interesting point is that they were not observed in the same magnitude in each condition. Performance gains for comfort were more pronounced under less advanced support facilities (i.e., status and history display), suggesting that practice effects become smaller under enhanced support facilities. To some extent, this differential pattern represents a natural "ceiling effect". Participants with the advanced aids were able to attain a high level of competence in system management right from the start. The scope for further improvements (indeed the need to improve their performance at all) was thus much reduced compared to the less well supported groups who began from a baseline that was below the standard for satisfactory performance. Nonetheless, the results imply that users being less familiar with a system will particularly benefit from advanced support facilities.

The usefulness of each display type for CHS design is now considered in more detail. The status display (together with the predictive display) was considered by users to be the most critical information source across all experimental conditions. Its rating was hardly influenced by the presence of other (more advanced) display types. This suggest that status displays are considered to be an essential basic support facility, indispensable for effective system operation since they provides up-to-date information about the current system state on a permanent basis (Norman, 2002; Ulich, 1986). Although some form of status display is a *sine qua non* for the operation of any device, the choice of information to display is problematic where there are multiple interacting variables (Bennett et al., 2005). Here the situation may seem to be straightforward, with only two key parameters (cost and comfort). Nonetheless, the relationship between the two variables (i.e., how they trade off against each other) is far from transparent. When this relationship is made more tractable, for example, by providing some form of waste indicator (as in an earlier

Table 4
User ratings of utility of various system support facilities as a function of experimental condition (1 = low utility; 7 = high utility; N/A: not available; [**]$p < .01$; [***]$p < .001$)

	STAT	HIST	PRED	INST	WARN	Result of analysis of variance
Status display	6.7	6.2	5.7	6.4	5.9	$F(4,70) = 1.54$
History display	N/A	6.1	4.2	3.7	4.0	$F(3,55) = 4.98$[**]
Predictive display	N/A	N/A	6.3	6.4	6.3	$F < 1$
Instructional display	N/A	N/A	N/A	5.2	4.0	$F(1,28) = 2.89$
Warning display	N/A	N/A	N/A	N/A	3.3	N/A
Weather forecast	5.2	4.3	3.1	2.0	1.7	$F(4,70) = 12.28$[***]

experiment), significant improvements in system management were obtained (Sauer et al., 2007).

Although the history display provided more detailed data (such as energy consumption and energy efficiency use) than the status display, it is interesting that this additional information did not result in better performance. This was unexpected and, at first sight, may appear to contradict the findings of a previous study using the CHESS environment which found some benefits of the history display over the status display for performance (Sauer et al., 2007). A closer look at the operational scenarios faced by users suggests that scenario stability may represent a major factor that influenced the utility of history displays. In the previous study, the operational scenario was much more stable than in the current one (i.e., less pronounced day-to-day changes in outside temperature, a steadier user lifestyle with less changing target temperature levels, and more regular presence/absence cycles at home). More stable operational scenarios allow a more accurate extrapolation of required heating profiles on the basis of historical information due to the strong similarities between the current and the future situation. Overall, this would suggest that the effectiveness of the history display decreases as operational conditions become increasingly unstable. While there is considerable empirical evidence in the literature that feedback on historical energy consumption is highly useful for energy conservation (e.g., Abrahamse et al., 2005), this may mainly result in simple conservation strategies such as "keeping-the-home-cool". For more complex energy management strategies (e.g., minimising consumption while maintaining comfort levels), different user support facilities may be needed. When examining the usage of history displays in an industrial context, it seems that their biggest advantage is in the context of fault diagnosis (i.e., backward reasoning) rather than, as required here, in system control with a forward flow of events (Wickens and Hollands, 2000). Although there was a diagnostic component to the users' task in the present study, in the sense that they were expected to reflect on how well they were managing the system, there were no faults as such. Had the users been required to investigate possible system malfunctioning, especially faults that waxed and waned through time, we may well have seen more intense scrutiny of historical information.

The predictive display clearly provided a very valuable tool, being confirmed by superior ecological performance and consistently positive subjective evaluations across all experimental conditions. It was regularly consulted by all users because it seemed to provide all the information needed (witnessed by a reduced inclination to access other sources of information). The worse performance of the groups without this display did not reflect any lack of motivation or disaffection. Indeed, these groups interacted just as intensively with the system and if anything made more use of the informational resources available to them, and rating these resources relatively more highly. The benefits of the predictive display thus reflect its specific technical advantages, rather than any indirect effects on general levels of intrinsic motivation. A clear benefit of the predictive display is that it lowers working memory load, relieving users of the cognitive effort of mentally calculating the effects of their actions; these can be anticipated with the tool and do not need to be based on one's own deliberations. The predictive display is not only effective in near-transfer situations (e.g., extrapolating from a familiar situation) but also in far-transfer situations (e.g., extrapolating from a novel situation). More varied operational conditions appeared to have increased the benefits of predictive displays. Overall, the results are in line with the broad swathe of research which demonstrates the clear benefits of predictive displays in managing systems that are complex, lagged and dynamic (Wickens and Hollands, 2000). Domains ranging from medicine (Kennedy et al., 2004) to aviation (Wickens and Hollands, 2000) consistently show the advantages of such decision aids.

While instructional displays have demonstrated some utility in a number of work-related application areas (e.g., Sarter and Schroeder, 2001), evidence on the benefits of support aids providing qualitative feedback is generally rather more equivocal than for the benefits of predictive aids. Experts systems provide a good example of the genre. Although such systems have the potential to aid decision-making in constrained, highly structured settings (Wickens and Hollands, 2000), there is often considerable resistance to their use in professional domains, such as medicine, where clear evidence in terms of improved decision outcomes is also wanting (Sintchenko and Coeira, 2003). In the present study there was no evidence for further improvements in ecological performance when this display type was added. The relatively low use of the instructional display compared to the users' enthusiasm for the predictive aid echoes the resistance to advisory tools alluded to above, although users certainly found the qualitative feedback to be of some interest. Their use of the instructional tool however failed to translate into improved performance, which indicates that all the key information enabling enhanced system management was furnished by the predictive aid. It is possible that they were simply using the instructional aid as a check that all was well with their settings. It is notable that users in the present study constituted a relatively homogeneous, highly motivated and computer-literate group. It is conceivable that less computer-literate users (e.g. older people) might have benefited more from instructional displays since the predictive display may have been more difficult for them to use. It would thus be premature to abandon the idea of employing instructional displays in central heating design; in particular, they may be considered an important means to compensate for the lack of training opportunities in the domestic domain. The results do suggest though that the most fruitful line of development should focus on the provision of predictive displays rather than qualitative aids, although clearly more research is needed on the specific requirements of different segments of the user population.

In contrast, little benefit was found for the warning display. As noted above, this reflected the generally high standard of proficiency of users in all conditions involving the predictive display, resulting in very few operational scenarios in which a warning needed to be issued. It is possible that greater gains may have been found had such a warning system been available in the STAT and HIST conditions. We should also note the inherent dangers of such warning displays. It is possible that the present one served a default validation function; if no warning was given, then users assumed that all was well. This could mean that opportunities for fine tuning the system and obtaining further efficiency gains were lost. The use of warning systems as a validity check could be particularly consequential should the warning mechanism itself fail, leading operators to assume that all is well when it is not. Operator complacency is a major concern in system design for work environments (e.g., Bailey and Scerbo, 2007). With the introduction of more advanced automatic systems, complacency can be expected to become more pertinent in the domestic domain, too.

In this study, students were employed as participants, which may have increased the benefits of support aids because of the higher computer-literacy of this group compared to a more representative user population. We also acknowledge that the study was conducted in the artificial context of a research laboratory instead of a field setting. However, the aim was to draw some broad conclusions regarding the relative value and utility of different forms of support, rather than to estimate the

exact gains that might be accrued in everyday operational use. Having shown the potential benefits that can be achieved; the challenge for designers is to incorporate such support features into the design of operational systems paying particular attention to their usability (e.g., increasing intuitiveness and efficiency of use). Optimising usability was not our primary goal here, and we would assume that any reduction in effect sizes had we used a more representative user population could have been off-set by additional gains made through improved usability.

6. Conclusion

The study demonstrated that advanced support facilities like the predictive display can provide benefits for the ecological management of CHS operation. Several recommendations for further research can be given on the basis of the present work. An analysis of typical task scenarios and environmental factors should be conducted since the effectiveness of support devices may be influenced by factors such as the steadiness of user lifestyles and the range of local climatic variability. For example, history displays may well prove sufficient if the degree of external temperature variation is moderate (cf. Sauer et al., 2007) whereas predictive displays appear to be particularly promising for more challenging situations characterised, for instance, by a greater degree of day-to-day external temperature variability. Further research should also examine how the support devices could be improved with regard to usability, and it would also be appropriate to test future version of CHESS on a more representative sample of users. In particular, efforts need to be made to enhance usability in order to accommodate the needs of special user groups such as the elderly who represent an important subgroup of CHS users. Overall, the findings of the study suggest that in heating system design, we need to think beyond the development of good feedback facilities by also examining more advanced forms of user support. This could entail examining other forms of decision support from other application domains to determine whether they are transferable to the domestic realm.

Acknowledgements

References

Abrahamse, W., Steg, L., Vlek, C., Rothengatter, J.A., 2005. A review of intervention studies aimed at household energy conservation. J. Environ. Psychol. 25, 273–291.

Bailey, N.R., Scerbo, M.W., 2007. Automation-induced complacency for monitoring highly reliable systems: the role of task complexity, system experience, and operator trust. Theor. Issues Ergon. Sci. 8, 321–348.

Benedyk, R., Minister, S., 1998. Applying the BeSafe method to product safety evaluation. Appl. Ergon. 29, 5–13.

Bennett, K.B., Payne, M., Walters, B., 2005. An evaluation of a "Time Tunnel" display format for the presentation of temporal information. Hum. Factors 47, 342–359.

Brehmer, B., Dörner, D., 1993. Experiments with computersimulated microworlds: escaping both the narrow straits of the laboratory and the deep blue sea of the field study. Comput. Hum. Behav. 9, 171–184.

Frantz, J.P., Rhoades, T.P., 1993. A task-analytic approach to the temporal and spatial placement of product warnings. Hum. Factors 35, 719–730.

Hansen, J.P., 1995. An experimental investigation of configural, digital, and temporal information on process displays. Hum. Factors 37, 539–552.

Kennedy, R.R., French, R.A., Gilles, S., 2004. The effect of a model-based predictive display on the control of end-tidal sevoflurane concentrations during low-flow anaesthesia. Anesth. Analg. 99, 1159–1163.

Korsgaard, M.A., Diddams, M., 1996. The effect of process feedback and task complexity on personal goals, information searching, and performance improvement. J. Appl. Soc. Psychol. 26, 1889–1911.

Kroemer, K., Kroemer, H., Kroemer-Elbert, K., 2001. Ergonomics: how to design for ease and efficiency. Prentice-Hall, NJ.

Midden, C.J., Meter, J.E., Weenig, M.H., Zievering, H.J., 1983. Using feedback, reinforcement and information to reduce energy consumption in households: a field experiment. J. Econ. Psychol. 3, 65–86.

Norman, D.A., 2002. The design of everyday things. Basic Books, New York.

Palmer, E.A., Jago, S., Baty, D.L., O'Connor, S., 1980. Perception of horizontal aircraft separation on a cockpit display of traffic information. Hum. Factors 22, 605–620.

Patrick, J., 1992. Training: Research and Practice. Academic Press, San Diego.

Sarter, N.B., Schroeder, B., 2001. Supporting decision making and action selection under time pressure and uncertainty: the case of in-flight icing. Hum. Factors 43, 573–583.

Sauer, J., Rüttinger, B., 2008. Designing technology for work and home applications. In: Chmiel, N. (Ed.), Work and Organizational Psychology: A European Perspective, second ed. Blackwell, Oxford, pp. 471–476.

Sauer, J., Wastell, D., Hockey, G.R.J., 2000. A conceptual framework for designing micro-worlds for complex work domains: a case study of the Cabin Air Management System. Comput. Hum. Behav. 16, 45–58.

Sauer, J., Schmeink, C., Wastell, D., 2007. Feedback quality and environmentally friendly use of domestic central heating systems. Ergonomics 50, 795–813.

Sheridan, T.B., 2002. Humans and automation: system design and research issues. John Wiley & Sons, Santa Monica, CA.

Sintchenko, V., Coeira, E., 2003. Which clinical decisions benefit from automation: a task complexity approach. Int. J. Med. Inform. 70, 309–316.

Stanton, N.A., 1998. Human Factors in Consumer Products. Taylor & Francis, London.

Ulich, E., 1986. Einige Anmerkungen zur Software—Psychologie. Hum. Prod. Hum. Arbeitsplätze 8, 10–15.

Van Houweligen, J.H., Van Raaij, W.F., 1989. The effect of goal-setting and daily electronic feedback on in-home energy use. J. Consum. Res. 16, 98–105.

Verhallen, T.M.M., Van Raaij, W.F., 1981. Household behavior and the use of natural gas for home heating. J. Consum. Res. 8, 253–257.

Wickens, C.D., Hollands, J.G., 2000. Engineering Psychology and Human Performance. Prentice-Hall, NJ.

Wogalter, M.S., Barlow, T., 1990. Injury severity and likelihood in warnings. In: Proceedings of the Human Factors Society 34th Annual Meeting, pp. 580–583.