16 HOMEWARD BOUND: Ecological Design of Domestic Information Systems

David G. Wastell
Nottingham University Business School
Nottingham, UK

Juergen S. Sauer
Department of Psychology
University of Friebourg
Friebourg, Switzerland

Claudia Schmeink
Institute of Psychology
Darmstadt University of Technology
Darmstadt, Germany

Abstract *Information technology artefacts are steadily permeating everyday life, just as they have colonized the business domain. Although research in our field has largely addressed the workplace, researchers are beginning to take an interest in the home environment too. Here, we address the domestic realm, focusing on the design of complex, interactive information systems. As such, our work sits in the design science version rather than behavioral science paradigm of IS research. We argue that the home is in many ways a more challenging environment for the designer than the workplace, making good design of critical importance. Regrettably, the opposite would appear to be the norm. Two experiments are reported, both concerned with the design of the user interface for domestic heating systems. Of note is our use of a medium-fidelity laboratory simulation or "microworld" in this work. Two main substantive findings resulted. First, that ecologically designed feedback, embodying a strong mapping between task goals and system status, produced superior task performance. Second, that predictive decision aids provided clear benefits over other forms of user support, such as advisory systems. General implica-*

Please use the following format when citing this chapter:

Wastell, D., Sauer, J., and Schmeink, C., 2008, in IFIP International Federation for Information Processing, Volume 287, Open IT-Based Innovation: Moving Towards Cooperative IT Transfer and Knowledge Diffusion, eds. León, G., Bernardos, A., Casar, J., Kautz, K., and DeGross, J. (Boston: Springer), pp. 273-290.

tions for the design of domestic information systems are discussed, followed by reflections on the nature of design work in IS, and on the design science project itself. It is concluded that the microworld approach has considerable potential for developing IS design theory. The methodological challenges of design research are highlighted, especially the presence of additional validity threats posed by the need to construct artefacts in order to evaluate theory. It is argued that design theory is necessarily complex, modal, and uncertain, and that design science (like design itself) should be prosecuted in an open, heuristic spirit, drawing more on the proven methods of "good design" (e.g.,prototyping, user participation) in terms of its own praxis.

Keywords Design science, cognitive ergonomics, ecological design, domestic heating system, feedback, operator performance, goal setting theory

1 INTRODUCTION

> The same technology that simplifies life by providing more functions in each device, also complicates life by making the device harder to learn, harder to use. The paradox of technology should never be used as an excuse for poor design… the principles of good design can make complexity manageable (D. A. Norman, *The Design of Everyday Things* 1998, p. 31).

This paper is about the design and adoption of information systems artefacts in the home. Although information and communications technologies pervade more and more of contemporary life, research in the IS field (including that of IFIP WG8.6) remains preoccupied with the workplace (Brown and Venkatesh 2005; Venkatesh and Brown 2001). In neighboring disciplines, such as human–computer interaction and cognitive engineering, there is a smattering of design research addressing ICT in the domestic sphere. In our discipline, the few studies addressing the home environment typically concern themselves with contextual factors influencing technology adoption (e.g., Brown and Venkatesh 2005; Vijayasarathy 2004). While the importance of design factors (ease of use, usefulness, etc.) is acknowledged, such "adoption studies" intrinsically take the artefact as the object of passive study, not something to be actively and directly shaped. Beliefs and attitudes toward technology, and how these bear on intentions "to use or not to use," form the quintessential field of interest. Design is an element in this nexus, but only one component in a welter of other causal and contextual variables (social norms, demographics, and so on). The typical theoretical framing of these factor studies draws either directly on the diffusionism of Rogers (1995), or on derivative positions such as the ubiquitous technology acceptance model (Davis et al 1989).

The classic adoption study fall squarely within what Hevner et al. (2004) call the *behavioral science* approach to IS research, in contrast to the *design science* paradigm. The same broad dichotomy reappears in Gregor's (2006) typology of IS theory, which contrasts theory for *explanation and prediction* with theory for *design and action*. In both cases, what would seem to underlie the distinction is the degree to which the design of an artefact is a matter of immediate concern to the researcher, or whether the relationship to technology is passive and deferred. Although enjoying a well-established tradition with a lineage reaching back over many years (e.g., March and Smith 1995; Markus et al. 2002; Walls et al. 1992), IS design research remains a minority genre. This is as true for the deliberations of IFIP WG8.6 as it is for the field in general. Given the self-evident importance of design quality in decisions to adopt or reject technical innovations, this seems somewhat paradoxical. By reporting an example of design science in this paper, we hope to redress the balance and strengthen the presence of such research within our repertoire. By emphasizing the importance of design, we make common cause with cognitive ergonomists (notably Norman 1998 and Vicente 2003) who have inveighed against the proliferation of technical devices in modern society that are increasingly complex, opaque, and frustrating to use.

1.1 Designing for the Home

The design of information systems for the effective management of domestic central heating systems (CHS) provides our substantive focus. We shall describe and discuss two contiguous cycles of a design project funded by the German Research Council on ecological design. The overall goal of our research program is to explore the potential for energy savings through improving the design of domestic artefacts. CHS has by far the largest environmental impact of all technologies in the home and is, therefore, a priority concern. Its management is mediated by an information system, typically in the form of a paper-based periodic energy bill that provides crude information regarding energy consumption and costs. The potential for computerizing this information system, for exploring different modes of information presentation and decision support, is of obvious relevance, in terms not only of domestic economics but the broader ecological agenda.[1]

Designing for the home poses some unique challenges compared to the design of work-based systems (Sauer et al. 2007) . This brings us to the second sense of "bound" in the paper's title: the group of users is characterized by a high level of heterogeneity, without the possibility of selection according to technical competence; moreover, no formal training can be given, and users in effect set their own goals and tasks, with no performance supervision, standards, or systematic feedback. In general, these contextual differences demonstrate the great importance of careful design since the potential for influencing behavior is much more limited in the home than in a work context. In our investigation, we shall focus on the range of options for displaying information and

---

[1] As one of the most complex systems within the home, CHS is of intrinsic interest. It shares problematic features with industrial process control systems (albeit in a simplified form): that is, a slow process with multiple interacting contingencies (Wickens and Hollands 2000). In particular, the lagged response and long time constants of the CHS make it difficult to manage in an optimal way. The need to improve IS design for CHS management was underscored in a preliminary survey by the authors, comprising in-depth interviews with users in their homes. This confirmed that current interfaces give generally poor feedback, providing little support for making energy-efficiency gains or understanding causal connections between system operation and costs incurred. There is also some encouraging field-based evidence that good IS design can be effective. Van Houwelingen and Van Raaij (1989) showed that goal-setting in conjunction with feedback (a novel "energy cost indicator") produced a significant reduction in energy consumption.

aiding decision-making in order to achieve the optimal balance between comfort levels and energy expenditure.

1.2 Display Design and User Support

Following Kroemer et al. (2001), we distinguish four main display categories: status displays (indicating the current system state), historical displays (information about past trends), predictive displays (projected information on future trends), and instructional displays (providing operational guidance). Research on the relative merits of different forms of display is extensive and widely scattered, clustering under various disciplinary headings: HCI, cognitive ergonomics, cognitive engineering, etc. We have not the space here to provide more than a very cursory overview of some relevant themes. As a general comment, it must be said that design absolutes are hard to find amongst the plethora of sometimes inconsistent or contradictory findings in this diverse literature.

Regarding status displays, some such feedback is essential for the operation of any device, but the choice of information to display is problematic especially where there are multiple interacting variables (Bennett et al. 2005). It might be thought that portraying historical trends is always advantageous, and such displays are indeed in widespread use. Yet the little empirical research that has been done has thrown up some negative findings (Bennett et al. 2005; Spenkelink 1990). A less equivocal picture emerges for predictive displays, with benefits reported in many domains ranging from medicine to aviation (Wickens et al 2000). Such tools are especially useful for managing lagged nonlinear systems where the anticipation of future evolution is inherently complex (Wickens et al. 2000).

Instructional displays provide qualitative feedback guiding operator action. Expert systems provide a much-hyped 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. Medicine is one such area, where clear evidence in terms of improved decision outcomes is notably lacking (Sintchenko and Coeira 2003). This resistance recalls the disinclination noted in our field to the use of decision support systems such as executive information systems (Elam and Leidner 1995; Hung 2003). Unless perfect reliability of the decision aid can be assumed, there is evidence that simple status displays may actually be preferable to instructional displays (Sarter and Schroeder 2001), which can give misleading advice.

1.3 Research Aims and Overview

Two design experiments are reported here. In the first, the potential benefits of historical displays are evaluated, together with enhanced "ecological" feedback. The second experiment assessed the relative advantages of predictive and instructional displays. The findings of the two studies will be presented and discussed, before moving on to a more general set of reflections on the design science program, motivated by recent debate over its nature and agenda (Chatterjee and Hevner 2006; Hevner et al. 2005). As an instance of design science in action, the work offers the opportunity to reflect critically on the practice of design, and indeed design science itself. There is specific interest in the investigative approach used here, which has some innovative aspects.[2] It involves an experimental methodology incorporating a dynamic, computer-based laboratory simulation. Such "microworlds" are invaluable for investigating complex, often inaccessible, settings where direct observation of key phenomena is problematic (Brehmer and Dörner 1993); their obvious potential in IS design research has been stressed by Wastell (1997). Designing for the home is methodologically challenging. While a clear understanding of user requirements is essential, the designer is not in the advantaged position of his organizational counterpart who can directly co-opt users into the design process. Alternative, less immediate methods of accruing design knowledge must be sought. The microworld is an attempt to insert the real world into the "arc of design" at a formative stage, and it will be of interest to observe how effective this method is.

2 RESEARCH METHODOLOGY

2.1 Experiments as Heuristic Devices: Introductory Reflections

Whereas experimental research is normally associated with the rigorous version of positivism, the present work was carried out in a more heuristic spirit. While empirical evaluation is an indispensable part of design science, the strong form of directional hypothesis testing would seems at odds with the inventive, creative spirit of design (Boland and Collopy 2004). It would also seem to represent an unrealistically optimistic view of the certainties of knowledge in the realm of the artificial, a world of contingent truths rather than necessary laws (as we glimpsed in the preceding section). Given the scientific uncertainties of the terrain, strong *a priori* hypotheses were eschewed. Beyond the broad expectation that more user support will, in general, lead to better performance, we ventured forward with an open pragmatic mind to appraise empirically which design concepts work and which do not, to attempt to understand some of the underlying contingencies and to abduct some tempered generalizations.

2.2 The CHESS Microworld

As noted, the experimental work was carried out using a PC-based simulation of a generic CHS, dubbed CHESS (Central HEating System Simulation). The design of CHESS itself is of methodological relevance in relation to our general design science concerns. The version for experiment 1 was largely fashioned by the experimenters themselves (i.e., with no user involvement!) supplemented by expert feedback from

---

[2] Design theory falls into two broad categories (Venable 2006; Walls et al 1992): knowledge about the design of products and about the design process itself, corresponding to the two senses of design as noun and verb. Both categories are relevant here; we are interested in the design of complex information systems and also the methodology of design.

engineering colleagues with interests in "green design." Version 2 did involve an element of user participation, in that subjects from experiment 1 were debriefed about their experience[3] and their comments led to several useful improvements in system usability. As with any microworld, verisimilitude was only carried as far as the pragmatics of the experiment required: time-scales were greatly accelerated, energy units were arbitrary, system dynamics were highly simplified, and, of course, the psychological mimicking of comfort levels could only be mediated in the crudest of ways. Within these limitations, CHESS was designed to provide a realistic user experience, in terms of the abstract nature of the operator's task, the visual appearance of the interface, and the underlying dynamic properties of the heating system (heat losses, lags, etc.). The aim was to produce a generic model that was convincing enough to generate meaningful user engagement and glean some relevant design knowledge regarding tool support and task performance.

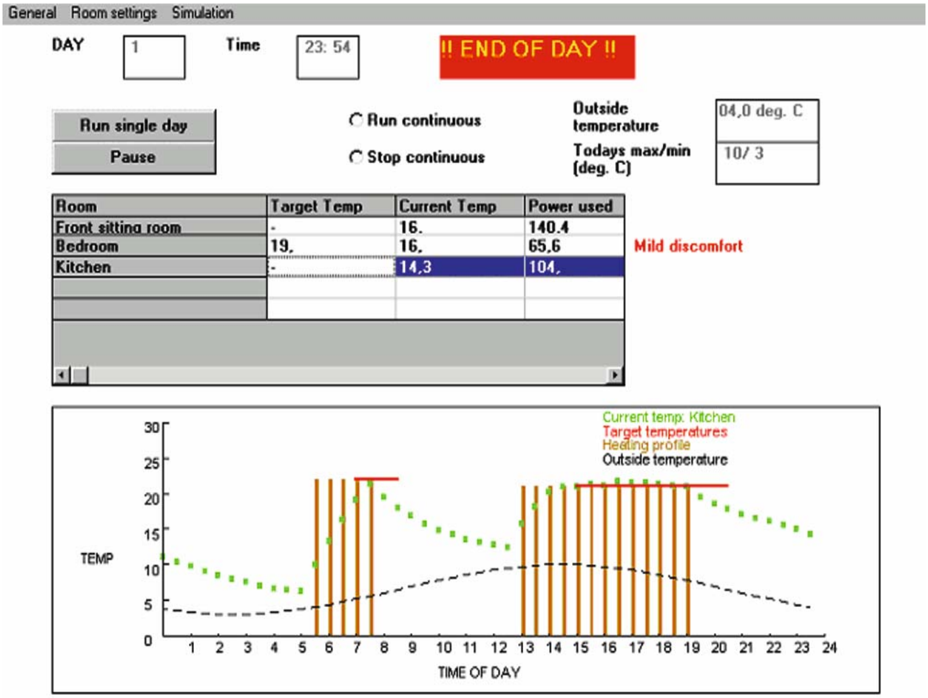
CHESS can be configured to represent a range of types of accommodation and heating arrangements. Specific temperature and weather profiles can also be created by the experimenter and stored prior to experimental sessions. CHESS also creates a results file in which all key performance parameters (energy consumption, comfort levels, etc.) are logged for each "day" the simulation runs. All settings made and sources of information sampled by the operator are also recorded in the results file.

A small "one person" apartment with three rooms was used for both experiments. The operator's main task was to define a heating profile for each of the rooms (i.e., sitting room, bedroom, and kitchen) according to a target temperature profile. There is a "set up" screen for each room, and the heating profile can be typed in via a dialogue box specifying the start, end, and thermostatic temperature level for one or more "heating blocks"; alternatively, the mouse may be used, drawing the block directly on a temperature by time-of-day graphic. Once the user is satisfied with the room settings s/he has made, the main simulation screen is selected (Figure 1). Clicking the "run simulation" button, fast-forwards the simulation for a complete day before pausing again; this takes around 30 seconds unless the operator decides to interrupt and intervene. Half-hourly status information on temperatures, cumulative energy use and comfort levels are provided on this screen, as well as a graphic showing diurnal trends for individual rooms.

Satisfactory comfort was defined as the attainment of a room temperature within 1 degree of the target. If the room temperature falls more than 1degree below target but within 3 degrees, "mild discomfort" is indicated. A more severe discrepancy is signaled by the "serious discomfort" indicator. Figure 1 shows the situation for one participant at 23:54, when the day is nearly at a close. Only the bedroom has a target value. Clearly the heating has been switched off slightly prematurely, reflected in the disparity between target and actual temperatures marked by the appearance of the "mild discomfort" warning.

---

[3] Interestingly, their comments related largely to usability issues and no radical proposals were made; their feedback led to a number of minor enhancements in version 2 (e.g., the provision of press-button controls for editing heating blocks). Further discussion of general issues in relation to microworld methodology may be found in Sauer et al. (2000), and a longitudinal review of one such application (CAMS) in the context of industrial process control may be found in Wastell et al. (2003)



DAY 1

Time 23: 54

Outside temperature 04,0 deg. C

Todays max/min (deg. C) 10/ 3

Room	Target Temp	Current Temp	Power used
Front sitting room	-	16.	140.4
Bedroom	19.	16.	65.6
Kitchen	-	14.3	104.

Figure 1. The Main Simulation Display Is Shown at the End of a Simulated Day. The graphic at the bottom of the display shows the outside and room temperatures (dashed and dotted curves) superimposed on the heating profile (vertical bars) and the target temperatures (horizontal lines).

A number of sources of management information were potentially available in both experiments at any point, depending on the experimental condition: a weather forecast provided a prediction of daily peak and low temperatures over the next four days. There was also a monthly report summarizing daily energy consumption and comfort levels for the house as a whole. Detailed daily reports were also potentially available, showing the hourly temperature profile for each room (akin to Figure 1) for any previous day, together with a summary of energy consumption and comfort.

3 EXPERIMENT 1

The first experiment compared the following three levels of operator support. The standard mode (STAND) was intended to represent the normal situation in most households; cumulative information only was provided in the form of an overall indication of total energy used at the current point (equivalent to reading the gas or electricity meter).

Participants were also informed of how effective they had been in heating the house (i.e., the average comfort level achieved thus far); this information was provided in aggregate form on the monthly report (i.e., without any daily breakdown). In the historical mode (HIST), participants were given full access to detailed historical information for each day, via the monthly and daily report screens. They were able to inspect the temperature and heating profiles for any room on any selected day, together with the total energy and comfort achieved for that day.

In the ecological (ECOL) mode, participants were provided with enhanced feedback in the form of a "waste estimator," which used a crude algorithm to estimate how much energy had been unnecessarily consumed by switching the heating on too soon, or switching it off too late. This indicator was intended to overcome obvious weaknesses in the basic status information available in the STAND and HIST modes. Although the two relevant status variables were presented (energy and comfort), it is not readily clear how they trade off against one another, nor how they are impacted by operator behavior. The waste indicator provides a simple example of feedback that is ecological (Rasmussen et al. 1994) in the sense that it enables the operator "to directly perceive the state of affairs in the environment…the deep structure of the work domain" and provides some guidance for adaptive action (i.e., reviewing switching on/off times).

3.1 Participants, Design, and Procedure

In all, 45 participants (30.0 percent female) were recruited from the Darmstadt University student population. They were aged 19 to 38 years (mean = 24.3 yrs) and received a payment of €25 for their participation. A one-factorial between-subjects experimental design was employed, with operator support as the primary independent variable, varied at three levels, as above. The experiment took place in a laboratory setting, and participants were tested in groups of three or four working on individual PCs, separated by screens. The testing session comprised two main phases. In the first, they were introduced to the CHESS software and initial training was given. In the second "treatment" phase, participants completed two task scenarios in their assigned feedback condition, each scenario involving the operation of the heating system for a whole month. Sessions on average lasted around 3 hours. Reflecting the home context, participants were not assigned specific goals; they were instructed to "do their best" to achieve the comfort levels prescribed by the set target temperatures while minimizing energy consumption.

User performance was measured on three dimensions: energy consumption, comfort (as defined above), and energy efficiency. Energy consumption was measured as the amount of energy (expressed as the average cost per day) consumed during task completion. The energy waste index was used to assess efficiency, although it was only fed back to participants in the ECOL condition. Patterns of information sampling and intervention were also assessed throughout. The frequency of inspection of the monthly and daily reports, the room settings, and the weather forecast were measured. System control activity was assessed by measuring the frequency of changes to the heating profiles.

Table 1. Effects of Feedback Level on Key Performance Indicators (Experiment 1)

Feedback Mode	Energy Cost per Day	Comfort (Max 100%)	Estimated Waste
Ecological (ECOL)	5.4	93.8	5.3
Historical (HIST)	5.7	92.6	13.4
Standard (STAN)	5.6	88.5	12.8
F ratio	11.8	26.7	53.2

3.2 Results

Given the multiplicity of dependent variables (nine in all, including the five measures of user interaction), a two factor MANOVA was carried out as the first step in the formal statistical analysis, with operator support (feedback mode) and day-of-the-month as the two independent variables. Only the second treatment month was analyzed, as participants were at their most experienced and proficient at this point.

Wilks lambda was significant for both *support* ($F = 16.9$, $p < 0.000$) and *day* ($F = 27.7$, $p < 0.000$). Table 1 shows the impact of operator support, in terms of group average performance levels across the month. Univariate F ratios were significant for all parameters. Tukey *post hoc* tests indicated that participants achieved significantly higher levels of comfort in both the HIST and ECOL conditions, and that their energy use and waste levels were significantly reduced in the ECOL condition. All other differences were not significant. It is apparent that while the HIST group achieved significantly higher comfort levels than the standard group, they used more energy to realize this. Their waste index was also the same as the standard group. Participants in the ECOL condition, on the other hand, achieved improved comfort, but with no significant increase in energy use; in fact, a small reduction was noted. Their waste index was less than half the other two groups!

Table 2 summarizes patterns of user interaction with the simulation. First we note that there is a low level of interaction with the system overall, generally less than one interaction per day per option. Looking at the monthly trend, it was found that most interaction occurred at the outset of the month, as would be expected, thereafter rapidly falling away and achieving steady state at around day 6. Room screens, for instance, were accessed 2.7 times on average on day 1 (to set/adjust heating profiles); by day 8, this had dropped off to less than once per day (0.7). Tukey HSD *post hoc* tests revealed

Table 2. Information Sampling and System Intervention as a Function of Feedback, Experiment 1 (All F ratios are based on df = 2:1,318 (pooled error term))

Feedback Mode	STAN	HIST	ECOL	F ratio
Sampling of monthly report (No./day)	0.28	0.25	0.40	11.6
Sampling of historical daily reports	N/A	0.16	0.26	12.5
Accessing room screens (No./day)	0.90	0.74	1.31	14.8
Changes to profile settings (No./day)	0.82	0.60	1.07	14.9
Sampling of weather forecast	0.87	0.55	1.31	23.5

two themes. First, that for all forms of interaction, user activity was at its most prolific in the ECOL condition. Second, that for temperature forecasts and changes to heating profiles, there was consistently higher interaction in the STAND than the HIST condition, whereas for monthly reports and room screens, there were no differences between these two feedback conditions.

4 EXPERIMENT 2

The second experiment differed from the first in two substantive respects. First, in the provision of more sophisticated support tools: a predictive display and an instructional adviser, both optional features. The second difference was the posing of a more demanding heating management task, including an external temperature profile that fluctuated significantly more markedly.

Three levels of operator support were provided. The first condition (HIST2) was a direct replication of the HIST condition of experiment one. The other two conditions involved additional support tools for setting room heating profiles. The first of these (PRED) provided a predictive display showing the estimated temperature profile for that room, given the weather forecast and the heating blocks defined by the operator. Building on the results of experiment 1, an indicator of energy efficiency was also provided. This gave a comparison of the predicted performance against the optimal that could be achieved for a given set of target and forecast external temperatures.[4] In the third condition (INST), an instructional display was also available, providing explicit feedback on efficiency and comfort, as well as advice on how to improve performance (e.g., to switch on sooner and/or to set lower target temperatures).

4.1 Participants, Design, and Procedure

Forty-five new participants (48.0 percent female) took part in the study, again from the Darmstadt University student population. Their ages ranged from 19 to 47 years (mean = 23.7 yrs). They were paid €15 each for their participation, and general procedures were the same as experiment 1. Sessions lasted approximately 2 hours. After a brief overview of the experiment, participants were introduced to the CHESS software, and received practical instruction in its use. They then completed one test session involving the operation of the heating system for 30 days, which took approximately 60 minutes. This involved a complex sequence of daily scenarios modeled on typical lifestyle motifs: a normal working day, an extended working day, a weekend day including a late party, a day at home. Each scenario had a specific heating profile to be implemented.

---

[4] This assessment was more sophisticated than the waste estimator of experiment 1, and involved the use of a simple "hill-climbing" heuristic that progressively adjusted the set heating profile until the optimal level of efficiency was obtained.

4.2 Results

Again, MANOVA was performed with experimental condition (three groups) and day-of-the-month as the independent variables. This found that both factors were significant, yielding a Wilks lambda of 34.5 ($p < 0.000$) for the main effect of user support. Table 3 shows that whereas the participants with access to the predictive display were able to cope adequately with the more exacting challenge of this experiment, the performance of participants with only historical information was significantly impaired in terms of their ability to attain satisfactory comfort levels. Univariate F tests were significant for all three parameters in Table 3, with Tukey HSD *post hoc* comparisons indicating that the participants in the HIST2 mode, although more parsimonious in their energy use, were less effective in attaining comfort levels and showed a generally lower level of energy efficiency. No differences in performance were found between the PRED and INST groups for any parameter.

Table 4 summarizes patterns of operator interaction. The need for more intensive monitoring and intervention is borne out by these figures, which are notably elevated compared to their counterparts in experiment 1 (they show similar intensity to the levels evidenced at the outset of the month in the earlier experiment). The general pattern across conditions is striking. There are two main themes. A more active profile of interaction is consistently seen for the HIST2 condition; this is especially remarkable for the temperature forecast and the historical reports, which are hardly accessed at all in the PRED and INST conditions. The second theme is the very similar profile of interaction for the two conditions where the predictive display is available (PRED and INST). Tukey *post hoc* tests revealed this general pattern to be statistically valid for all parameters except the monthly reports, which were sampled more often in the PRED condition.

Table 3. Energy Consumption, Comfort, and Energy Efficiency Levels for the Second Experiment (Note that energy efficiency is simply comfort divided by daily energy cost)

User Support Mode	Daily Energy Cost	Comfort	Energy-Efficiency
History (HIST2)	6.0	74.4	12.4
Predictive (PRED)	6.3	90.3	14.3
Instructional (INST)	6.3	91.6	14.5
F ratio	13.8	86.5	97.6

Table 4. Information Sampling and System Intervention for Experiment 2 (All F ratios are based on df = 2:1,318 (pooled error term))

	HIST2	PRED	INST	F ratio
Sampling of monthly report (no./day)	0.34	0.45	0.13	43.2
Sampling of historical daily reports	2.8	0.7	0.3	57.7
Accessing room screens (no./day)	4.6	3.3	3.1	15.1
Changes to profile settings (no./day)	3.3	2.6	2.8	28.4
Sampling of weather forecast (no./day)	1.8	0.3	0.1	117.2

Regarding the use of the advanced support tools, there was striking evidence of a strong preference for the predictive display. This was accessed on average 10.3 times per day by participants in the INST condition, and 11.5 times per day in the PRED group (this difference was not significant, $t = 0.18$). In contrast, the energy adviser was rather scantily consulted, with the average rate of use being less than once per experimental day (mean = 0.81).

To triangulate these behavioral measures, participants were asked to rate their subjective opinions of the various information displays on a seven point scale, running from 1 (low utility) to 7 (very useful). The HIST2 group gave a significantly higher assessment of the value of the historical display (6.1) than either of the two other groups (4.2 and 3.7 on average for PRED and INST respectively, $F = 4.9$, $p < 0.01$). This group also gave a notably higher rating (4.3) to the weather forecast than subjects in either of the other conditions, who gave an average rating of 2.5 ($F = 12.3$, $p < .001$). Both groups with access to the predictive display rated it very highly (6.3 and 6.4 on average), markedly higher than the general evaluation of the instructional display (5.2 on average).

5 DISCUSSION

In contrast to the mainstream "behavioral" approach to research on the adoption of IT/IS, the present work takes a more active tack. The aim is to influence directly the take-up of technology by improving its design, centering attention on the delivery of relevant functionality embodied in a supportive user interface. We have also set foot in an application domain that is relatively neglected in our discipline, building on the meager discourse addressing the home environment. We have noted that the home is, in many ways, a more challenging environment for the designer than the workplace, where supervision and training can help to optimize the utilization of IT. Without such support, it becomes of paramount importance that technology is well-designed if it is to be used effectively. But as others have lamented, all too often this is not the case; bad design would seem to be the norm for everyday objects (Norman 1998). Our discussion will be organized under two broad headings, following the twin aims prefigured in the "Introduction." In this penultimate section, we will reflect directly on the specific knowledge furnished regarding IS design for the home. In the final section, we will stand back and reflect more generally on design work and design science, based on our practical experiences reported here.

Our main substantive findings can be expressed in the following two generalizations. First, that historical displays, while replete with detailed management information, do not in themselves assist users to manage the heating system more effectively, whereas the provision of feedback directly indicating energy efficiency strongly motivated more interaction, which in turn enhanced performance. Secondly, that decision aids in the form of predictive displays are well-liked by users, extensively used, and engender enhanced outcomes, whereas instructional support was relatively disregarded.

Concerning the first finding, it is notable that the HIST group in experiment 1 actually intervened less often than the standard group, despite the radical improvement in their management information. By contrast, the addition of a single database field (the waste index) in the ECOL condition evoked a gestalt change in operator behavior: all information sources were sampled more frequently, including the weather forecast (strongly indicative of a more proactive orientation). Increased levels of system intervention were also observed. These changes in engagement directly translated into more energy-efficient system management, with comfort levels and energy consumption being jointly optimized. How can this dramatic transformation be explained?

For an explanation, we first return to the concept of ecological information systems design (EISD) briefly alluded to above (Rasmussen et al. 1994). The essence of EISD is the imperative to support human judgment in the management of complex problems through the clear representation of task objectives together with an information system ("measuring functions" in Rasmussen et al.'s terminology), enabling the *direct perception* of critical environmental variables and relationships. Although the *a priori* motivation for the waste indicator had been merely to provide additional feedback, it was realized retrospectively that its real potency lay elsewhere, in that it provided an implicit performance goal. The powerful influence of goals over performance is well known. Goal-setting theory (Locke and Latham 2002) argues that specific performance goals are the primary regulators of task performance, and that the mere provision of feedback will have limited potency (McCalley and Midden 2002). A clear mapping between goals and performance feedback is thus vital; its all-too-common lack is memorably described by Norman (1998) as the "gulf of evaluation."

Setting specific external goals is, however, unrealistic for the domestic environment; goals are voluntary and self-imposed in this context. Although not designed as a goal-setting mechanism (and no explicit waste goals were set), the waste indicator nonetheless carried a potent implicit goal in its moral labeling (all waste is bad and the target should, therefore, be to reduce it to zero). Not only does it set a goal but it directly conveys performance feedback against that goal, combining information from two oppositional parameters (comfort and cost) into a single unequivocal index. No mental computation is required to gauge performance; the error signal is directly given. Combining goal and feedback, it thus provides a simple example of ecological design that neatly bridges the gulf of evaluation. We thus regard the superior performance of the ECOL group as a direct reflection of the efficacy of sound ecological information system design. It is important to reemphasize, however, that the success of the waste-indicator was largely serendipitous; its teleological properties were only appreciated *ex post facto*.

Looking back at the first experiment, we can read the pattern of results as a general commentary on the contingencies of design and the limitations of design agency (Richardson 1993). Designers sometimes get it right, but sometimes they do not. Both forms of enhanced feedback were expected to enhance system management, but only one afforded genuine benefit. The same generic pattern comes through in the second study: one aid worked whereas another failed to deliver the expected gains.

The benefits we found for predictive displays[5] are consistent with the broad swathe of research that demonstrates the clear advantages of such displays in managing systems

---

[5] The poorer performance of the HIST2 group did not reflect any general lack of motivation. Indeed, this group interacted more intensively with the system, making more use of the information resources available to them, rating these resources more highly. The enhanced performance of the other two groups can thus be attributed to the superior technical tools at their disposal.

that are complex, lagged, and dynamic (Wickens et al. 2000). Regarding the benefits of instructional support, the research literature is markedly less certain, as we noted in the "Introduction." Users in our simulated home environment were apparently just as disinclined to utilize such advisory systems as are professional users in real work settings (Sintchenko and Coeira 2003). Although consistent with literature that is equivocal on the benefits of prescriptive aids, definitive conclusions should be drawn carefully. The failure could, of course, simply reflect a poorly designed implementation rather than a fundamental flaw in the concept of instructional support. Nonetheless, we may safely infer that the design of such support is more challenging than of predictive aids, and that there is a greater range of contingencies to be addressed. We may conclude that designers of complex artefacts for the home, and indeed for other environments that involve the management of dynamic processes, should, in the first instance, devote their efforts to the provision of effective predictive displays, and exercise caution in investing their design labor on forms of support that involve advice and guidance.

6 CODA: REFLECTIONS ON DESIGN AND DESIGN SCIENCE

As well as an experiment *for* design, the CHESS studies constitute an experiment on the process of designing itself. Let us begin these final reflections by considering the insights into design work (and its potential dysfunctions) revealed by the studies before considering more strategic implications for design science. First, we note that design is not a top-down, linear process moving deductively from a body of *a priori* knowledge (Boland and Collopy 2004; Weick 1990). Its open, emergent character is clearly manifest in the first experiment, where the full properties of an important design feature (the waste indicator) only came to light adventitiously, and the relevance of a significant body of "kernel theory"[6] was only fully recognized once the experimental work had been completed (i.e., goal-setting theory).

The tendency in design work to incorporate novel features that please the intellectual curiosity of the designer, rather than the pragmatic needs of users, has been dubbed "creeping featurism" (Norman 1998). This phenomenon is present in both experiments,

---

[6] The role of kernel theories in design science is highlighted by Walls et al. (1992). Kernel theories are theories emanating from the natural and social sciences that fundamentally bear on design processes and products. Goal-setting theory would seem an important kernel theory for the design of complex interactive systems, its core proposition being that specific challenging goals, broadly speaking, engender superior task performance in comparison to vague goals (such as "do-your-best") or personality attributes (e.g., achievement motivation). Important nuances should, however, be noted; for example, for complex problems, where learning and exploration are critical, setting "learning goals" (rather than performance goals) is generally more effective. Although goal-setting is inherently problematic in the domestic environment, it is not intractable and further design ideas could have been developed following the general direction suggested by the present results. It is possible, for instance, to envisage setting explicit *relative goals*, for example, setting a target for an efficiency improvement relative to past performance in similar conditions, in the manner of van Houwelingen and van Raaij (1989).

and would seem to be an endemic feature of design. Significant effort was devoted in the first experiment to the development of the historical database, which provided no real benefit for decision making (not unlike many real-life MIS!), and, in the second experiment, the expert system was relatively underutilized. We have noted that there was little user involvement in our design process. The salutary thought occurs that had we consulted our users more, it is possible that such redundant features would have been abandoned, or perhaps refined to be more useful.

An important subset of design theory concerns the efficacy and aptness of design methods and tools, and the performance of the microworld was of obvious interest. The general value of such microworlds in the context of IS research has been argued by Wastell (1997). By inserting something of the complexity of the real world into the controllable space of the laboratory, it is possible to engineer and explore realistic and important scenarios that would be difficult to examine *in vivo*. It also becomes feasible to evaluate formally the influence of a range of environmental, task, and design variables. Although artificial, the medium fidelity realism of the microworld manifestly has the dramaturgical power to engage users and to elicit realistic behavior. Certainly our users took their work seriously; this is evident from the pattern of their interaction with the system and their task performance. Of course, the simulation is not the real thing, and external validity is inevitably problematic, as for any laboratory experiment. Such limitations are certainly recognized here. Nonetheless, we believe (albeit with due circumspection) that some valid and useful design knowledge has been generated by the microworld, and that the trends seen here have a degree of generalizability. The fact that outcomes ran against some of our broad expectations in itself provides a cogent argument for the use of such empirical methods. Design science differs from other realms of science (March and Smith 1995; Walls et al. 1992) in that its knowledge-base must be judged against utilitarian criteria as well as predictive and explanatory validity. This utility test (Venable 2006) has been passed; we have confirmed the heuristic potential of microworld methodology for generating realistic behavioral data, testing ideas, and developing design theory.

Design science differs from everyday design work in its aspiration to produce *theory*, not to build specific working systems. The artefacts constructed here are simply the means to this loftier end, vehicles for testing ideas and hypotheses in the endeavor to construct a body of generalizable knowledge regarding the design of a certain class of artefact. Methodologically, we would certainly affirm design science to be challenging. The experimental psychologist deals with one primary source of validity threats in her theory-building labors, those arising from the quality of the experimental design. She has the luxury of working with highly simplified experimental models (decocting real-world phenomena into some putative canonical form) and the theoretical knowledge sought is relatively simple and well-circumscribed. In contrast, empirical realism is vital in design science, enabling knowledge to generalize across a wide range of contexts of use, tempered by relevant singularities. We should, therefore, expect to see the accumulation of a relatively unstable body of contingent knowledge punctuated by exceptions and caveats (rather than necessary, timeless truths), resisting easy codification. The design researcher must always be alert for emergent phenomena and unintended consequences, dexterously ready to switch direction, to change approach, to explore different theoretical perspectives. In this disorderly world, induction and abduction represent more relevant modes of reasoning than strict deduction, as design researchers try to make sense of the complex and inconsistent findings furnished by even relatively simple experiments such as the present ones.

Doing design science necessarily involves design work, regarding both the design of the "experimentation" (field or laboratory) and of the artefact itself. Like design, we have argued that it is best prosecuted heuristically and opportunistically, but perhaps this argument should be pushed further. There is much that reflects the *habitus* of conventional behavioral science in the present work. The experimenters were very much in charge, and the empirics had a largely linear trajectory with little adaptation to user feedback. Critical elements of good design practice were lacking, in particular prototyping and user participation. We will end on the chastening thought that, had we worked more collaboratively and iteratively with our users (using prototypes, for instance, to explore design ideas), we may well have produced not only a better artefact but more robust theory as well. Reflecting this, users in our future work will be more fully engaged as partners and coproducers of design knowledge, rather than passive guinea pigs!

Acknowledgments

We gratefully acknowledge the financial support of the German Research Foundation (DGF) for carrying out this work (Research Grant: SFB392/TFB55).

References

Bennett, K. B., Payne, M., and Walyers, B. 2005. "An Evaluation of a 'Time Tunnel' Display Format for the Presentation of Temporal Information," *Human Factors* (47:2), pp. 342-359.

Boland, R., and Collopy, F. 2004. *Managing as Designing*, Stanford, CA: Stanford University Press.

Brehmer B., and Dörner, D. 1993. "Experiments with Computer-Simulated Microworlds: Escaping Both the Narrow Straits of the Laboratory and the Deep Blue Sea of the Field Study," *Computers in Human Behavior* (9:2/3), pp. 171-184.

Brown, S. A, and Venkatesh, V. 2005. "Model of Adoption of Technology in Households: A Baseline Model Test and Extension Incorporating Household Lifecycle," *MIS Quarterly* (29:3), pp. 399-426.

Chatterjee, S., and Hevner, A. 2006. "Program Co-Chairs Message," First International Conference on Design Science Research in Information Systems and Technology, Claremont Graduate University, February 24-25 (http://ncl.cgu.edu/designconference/Cochairmsg.htm).

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. 1989. "User Acceptance of Technology: A Comparison of Two Theoretical Models," *Management Science* (35:8), pp. 982-1003.

Elam, J. J., and Leidner, D. G. 1995. "EIS Adoption, Use, and Impact: The Executive Perspective," *Decision Support Systems* (14:21), pp. 89-103.

Gregor, S.. 2006. "The Nature of Theory in Information Systems," *MIS Quarterly* (30:3), pp. 611-642.

Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75-105.

Hung, S. 2003. "Expert vs. Novice Use of Executive Support Systems: An Empirical Study," *Information and Management* (40:3), pp. 177-189.

Kroemer, K., Kroemer, H., and Kroemer-Elbert, K. 2001. *Ergonomics: How to Design for Ease and Efficiency*, Upper Saddle River, NJ: Prentice-Hall.

Locke, E. A., and Latham, G. P. 2002. "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35 Year Odyssey," *American Psychologist* (57:9), pp. 705-717.

March, S. T., and Smith, G. 1995. "Design and Natural Science Research on Information Technology," *Decision Support Systems* (15:4), pp. 251-266.

Markus, L.. Majchrzak, A., and Gasser, L. 2002. "A Design Theory for Systems That Support Emergent Knowledge Processes," *MIS Quarterly* (26:3), pp. 179-212.

McCalley, L. T., and Midden, C. J. H. 2002. "Energy Conservation Through Product Integrated Feedback: The Roles of Goal-Setting and Social Orientation," *Journal of Economic Psychology* (22:5), pp. 589-603.

Norman, D. A. 1988. *The Design of Everyday Things*, London: MIT Press.

Rasmussen, J., Pejtersen, A. M., and Goodstein, L. P. 1994. *Cognitive Systems Engineering*, New York: Wiley.

Richardson, A. 1993. "The Death of the Designer," *Design Issues* (9:2), pp. 34-43.

Rogers, E. M. 1995. *Diffusion of Innovations*, New York: The Free Press.

Sarter, N. B., and Schroeder, B. 2001. "Supporting Decision-Making and Action Selection Under Time Pressure and Uncertainty: The Case of In-Flight Icing," *Human Factors* (43:4), pp. 573-583.

Sauer, J., Wastell, D., and Hockey, G. R. J. 2000. "Using Micro-Worlds to Simulate Highly-Automated Work Environments: The Case of the Cabin Air Management System," *Computers in Human Behavior* (165:1), pp. 45-58.

Sauer, J., Schmeink, C., Wastell, D. 2007. "Feedback Quality and Environmentally Friendly Use of Domestic Central Heating Systems," *Ergonomics* (50:6), pp. 795-813.

Sintchenko, V., and Coeira, E. 2003. "Which Clinical Decisions Benefit from Automation: A Task Complexity Approach," *International Journal of Medical Informatics* (70:3), pp. 309-316.

Spenkelink, G. P. J. 1990. "Aiding the Operator's Anticipatory Behavior: The Design of Process State Information," *Applied Ergonomics* (21:3), pp. 199-206.

Van Houwelingen, J., and Van Raaij, W. 1989. "The Effect of Goal-Setting and Daily Electronic Feedback on In-Home Energy Use," *Journal of Consumer Research* (16:1), pp. 98-105.

Venable, J. R. 2006. "The Role of Theory and Theorizing in Design Science Research," in *Proceedings of the First International Conference on Design Science Research in Information Systems and Technology*, S. Chatterjee and A. Hevner (eds.), Claremount Graduate University, February 24-25 (http://ncl.cgu.edu/designconference/DESRIST%202006%20Proceedings/2A_1.pdf).

Venkatesh, V., and Brown, S. 2001. "A Longitudinal Investigation of Personal Computers in Homes: Adoption Determinants and Emerging Challenges," *MIS Quarterly* (25:1), pp. 71-102.

Vicente, K. 2003. *The Human Factor*, New York: Routledge.

Vijayasarathy, L. R. 2004. "Predicting Consumer Intentions to Use On-Line Shopping: The Case for an Augmented Technology Acceptance Model," *Information and Management* (41:6), pp. 747-762.

Walls, J. G., Widmeyer, G. R., and El Sawy, O. A. 1992. "Building an Information Systems Design Theory for Vigilant EIS,". *Information Systems Research* (3:1), pp. 36-59.

Wastell, D. G. 1997. "Human-Machine Dynamics in Complex Information Systems: The "Microworld" Paradigm as a Heuristic Tool for Developing Theory and Exploring Design Issues*," Information Systems Journal* (6:4), pp. 245-260

Wastell D. G., Sauer, J. S., and Hockey, G. R. J. 2993, "Using Micro-Worlds in Research on Distributed Cognition in Complex Dynamic Worlds: A Ten Year Retrospective on the Cabin Air Management System," in *Cognitive Science Approaches to Process Control,* G. van der Veer and J. F. Hoorn (eds.), Le Chesnay, France: EACE, pp. 145-150.

Weick, C. 1990. "Organizational Redesign as Improvisation," in *Organizational Change and Redesign*, G. Huber and W. Glick (eds.), New York: Oxford University Press, pp. 346-382.

Wickens, C. D., Gordon, S., and Lui, Y. 2000. *Introduction to Human Factors Engineering*, Upper Saddle River, NJ: Pearson-Prentice Hall.

Wickens, C. D., and Hollands, J. G. 2000. *Engineering Psychology and Human Performance*, Upper Saddle River, NJ: Prentice Hall.

About the Authors

David Wastell is a professor of Information Systems at Nottingham University Business School. He began his research career as a psycho-physiologist before moving into information systems. His research interests are in public sector reform, innovation and design, management epistemology, and cognitive ergonomics. He has co-organized two previous IFIP WG8.6 conferences (1997 and 2007) and was research co-chair for the IFIP WG8.2 conference in Manchester in 2004. David may be contacted at dave_wastell@hotmail.com or david.wastell@nottingham.ac.uk.

Juergen Sauer is Professor of Cognitive Ergonomics at the University of Freibourg in Switzerland, a position he took up following his previous appointment at the Darmstadt University of Technology (Germany), where the present work was conducted. His interests are in the human factors design of complex systems and the ecological design of domestic artefacts. Juergen may be contacted at juergen.sauer@unifr.ch.

Claudia Schmeink works as a research assistant at the Institute of Psychology, Darmstadt University of Technology, Germany. She has a background in educational psychology and counseling. Claudia may be contacted at schmeink@psychologie.tu-darmstadt.de.