Designing Transgenerational Usability in an Intelligent Thermostat by following an Empirical Model of Domestic Appliance Usage

Adinda Freudenthal
Delft University of Technology, Faculty of Industrial Design, Intelligence in Products group, Landbergstraat 15, 2628 CE Delft, the Netherlands, a.freudenthal@io.tudelft.nl, http://www.io.tudelft.nl/intelligentproducts

Abstract: An intelligent thermostat was designed taking into account a model of usage strategies and how people learn. The model was based on earlier observations with young and old subjects using a brand new TV/VCR combination. Evaluation in usability trials confirmed that usage strategies were according to the model and non-compliance to the model in the interface design led to usability problems. New possibilities in ICT allow designers to meet the model based design guidelines. The result is improved usability, to a level where even senior citizens in the future should be able to program complex schedules.

Keywords: human-product interaction, learning to use, elderly users, thermostat, interface design

Introduction
Present domestic appliances, such as microwave ovens and VCR’s are equipped with buttons, displays and sometimes screens with menus. Many users have problems when programming these devices. Buttons have labels, which do not indicate what to do, because in most cases several buttons are to be pressed in a certain order. Feedback is usually poor ("ERROR 4"). Feedforward about what should be done is often lacking. The correct menu needs to be opened before items can be found (or modes or screens) and therefore finding an item can be difficult. This is especially the case if the name of the required item does not match the search terms of the user. Older users have even more problems, caused by a decline of cognitive capacities (see various studies on a range of aspects needed for product usage, such as slowing of information-processing, Rabbit (1992), performance during parallel tasks (Myerson et al., 1990) and the suffering from information overload, Cann (1990), recalling the recent past and searching memory, Lovelace (1990)). Besides this elderly users of consumer products often lack relevant experience with modern interfaces. For example, they often have not learned the principle of ‘spatial organization’ (of menus) or have problems in applying it (Docampo Rama, 2001).

An empirical model of usage and learning to use, based on observations of usage of a combined TV/VCR set, was used in the development of a new intelligent thermostat. The thermostat was designed, using new technological possibilities in ICT and was evaluated in usability studies. In this paper we will investigate whether users behave and learn according to the model also with the new thermostat and whether designing according to the model can help increase usability.

The model was developed earlier on from observations of young and old subjects using a brand new combined TV/VCR set. In that study 5 subjects of 15-18 years of age, 5 subjects of 30-40 years of age and 10 subjects of over 59 years of age participated. The TV/VCR had just been introduced (1994) to the market and was equipped with menus on the screen and many options, such as 'VPT' and 'automatic Channel search'. The subjects were allowed to approach the unknown device as they normally would. This way novice use (and learning to use) was tested. The observation and analysis method, and full results are given in Freudenthal (1999, 117-149). The summarized results and model in this paper were published earlier in Freudenthal (2000).

The general approach in novice use
Users of all ages were observed to approach the apparatus in the same way. When the subjects started using the unfamiliar TV/VCR they had an open approach and seemed to apply general knowledge about apparatus. Most started either by switching on the device or reading the manual. Their actions were 'user goal' driven. They had knowledge of possible user goals, for example 'selecting a channel' or 'changing the volume'. However, often just a few of the main functions were remembered. They seemed to strive to achieve these 'real user goals' with as limited effort as possible.

This meant that they were only inclined to make functions their goal, which represented a real user goal. For example 'programming the TV channels' would only become an intermediate goal if this was necessary to achieve another real goal. Users provided with a totally empty TV would sometimes be inclined to start programming the TV; in all other cases this would not be a spontaneous action. Elderly users would rather get help than try programming.
Users were not inclined to make 'learning' their goal. They used the product and unintentionally they learned. They did not intentionally explore. One exception we found was when occasionally an interesting or surprising feature was encountered during use. The user then sometimes briefly explored this feature. However, this was not common.

Procedures to reach user goals

Users expected that their user goals could be reached through the execution of a 'procedure', a sequence of actions in a certain order. However, they did not expect procedures to necessarily equate to those on their own equipment. In fact, they seemed to find it more logical that the new apparatus would function somewhat differently. This could be the result of previous experiences with new devices.

Users did not tend to use a 'trial and error' approach in finding out the procedure to operate functions. This was only observed with extremely simple functions, such as 'how do I switch on the TV?'. Probably the procedures of operation were too complicated to remember the results of trial and error and be able to deduce a strategy from this. This might explain the 'step by step' approach towards their final goal: seldom more than one step at a time seemed to be planned. In deciding what to do in the next step the users expected the device to guide them by providing relevant information.

Although users expected the product and the manual to guide them in product use they were rather casual about using information supplied by the product and the manual. If they thought they knew what needed be done, they would hardly study the provided information, but would immediately act. This was even the case if they were, for instance, in the middle of carrying out a procedure in the manual. They would forget about that and carry on by themselves. This was already happening early in the learning process.

During first product use subjects used general knowledge such as 'press a button to activate a function' or 'one should first program the channels'. Later on more knowledge would be available. For example, knowledge of required sequences of main actions was used, but only if the user was aware of the fact that a correct order was needed, and this was not always the case.

The users expected the available product functions to be organized according to how they mentally organize their user goals. For example, the subjects seemed to expect that 'programming the VCR for today' would not differ fundamentally from 'programming the VCR for tomorrow'. Problems occurred frequently when other procedures and/or other buttons were needed. One of them was that the wrong section in the manual was consulted and followed.

Mental storage of product rules

During novice use storage of information in memory was required. However, the available capacity of working memory, during the operation of the home appliance, did not seem to be sufficient to store all separate actions for the various procedures. This might be the reason why users remembered only general rules and forgot details almost immediately. They seemed to search for recurring patterns, translate them into rules and use them for predictions of future product reactions to user actions.

They seemed to deduce these 'laws', probably in the same way they have learned to explain and predict the behavior of their natural surroundings. The subjects all appeared to expect that the product would function according to constant 'laws' and would react consistently. It seemed that, consciously or unconsciously, users expected that reactions of the product to users' actions would reflect its 'laws'.

Executing a procedure that required actions, which did not conform to the general rule, would increase the number of items to be remembered in working memory. It seemed that the available capacity of working memory was usually insufficient for this. Younger users would then start to forget crucial information. Older users would easily reach a level of complete overload after which sensible actions would be rare. General product rules appeared to be rather easily stored, while exceptions were forgotten over and over again.

If a product reacts unexpectedly this can seriously disrupt the learning process. The worst cases were observed when users made mistakes while learning. Making mistakes is generally recognized by a user from the fact that feedback from the product is not according to expectations. If the user has encountered exceptions to the general product rule earlier on he may start to expect more exceptions, (which is in line with the assumption of consistency). We observed users assuming that another exception had been encountered, instead of deducing that a mistake had been made. They had to correct their developing 'mental model' again later on. For younger users this seemed to make learning difficult. For older users such problems were observed to completely frustrate further use.

A model of usage strategies and how people learn
From our observations we derived a model of novice product use of a TV/VCR combination, see figure 1 (Freudenthal, 2000). In the model we organized the way in which subjects appeared to use their internal knowledge and signs from the device.

Well-known elements, as can be found in almost every article or book with design guidelines, were recognized as being important (e.g. on feedforward and feedback, consistency, and guidelines to anticipate learned aspects in language, operation of devices and icons). Their roles in the process of novice product usage are indicated in the model. The figure indicates the relationships by arrows, which indicate the flow in time. During the usage steps the arrows are followed. Meanwhile the internal knowledge grows. After interaction has started the product rules will start to build up.

The described process of building up knowledge seems to take place only partly consciously. The user for the most part does not seem to be aware of the application of knowledge or of the laws and does not form the laws intentionally or consciously.

A few aspects will be explained a bit more: (1) We found that subjects expect consistency always and throughout a device, even if the device has a rather inconsistent design, therefore this is indicated as a given fact. The assumption of 'consistency' is used in the process of deducing the product rules and needs not to be deduced to be a product rule. (2) We found that feedback feeds the process of developing the product rules in the mind and the growing general knowledge of devices - and therefore of the interpretation of next feedforward. (3) The environment can effect the process. Think of, for example, low lighting or other persons changing settings.

A key finding is the step-by-step manner, in which internal knowledge is used together with the feedforward from the device and the manual. Seldom more than one step at a time seemed to be planned. (The steps indicated in the model are the actual steps taken by the user, and not the steps required by the product).

The behavior of subjects observed during the use of the TV/VCR did not meet expectations based on the three levels of cognitive control according to Rasmussen (1987). If problems of use were encountered, according to these theories, an approach on the highest level - 'knowledge-based' - is to be expected. Users are supposed to plan their actions in advance and carry them out. It seemed that short-term memory was simply insufficient for the complex problems users encountered in the apparatus.

There is a mismatch between what users are willing to invest and the required effort to operate and/or operating the brand new domestic device was above available human capacities. These capacities might be relatively low in this situation, due to a low motivation to (learn to) use. The observed low motivation is probably caused by usage goals, which are rarely of major importance and a life-threatening situation does not occur if things go wrong. Therefore goals could be adjusted if other goals were satisfactory as well.

Finally we must mention that, as all models are simplifications of what really happens, so is this model. For example, sometimes a user goal is not reached or the goal is not equal to the aim and mistakes are made. We decided to not explicitly indicate intermediate goals in the model, but these can be seen as reflected in the sequence of usage steps to be made.
Research target
In Freudenthal (1999, 201-231) guidelines were presented to optimally support young and old users. Many of these guidelines were composed to meet the model in a design. The set of guidelines was tested in new appliance designs in industry and proved to be helpful to improve usability. The degree of improvement is substantial but has its limits, because several guidelines conflict or are ‘impossible’ to meet in currently applied technologies.

An example is the requirement to substantially enlarge typefaces on static products. This requires space. However, unknown abbreviations (technospeak) should not be used, nor parts of sentences (which often mentally are completed incorrectly). Using icons as labels is not an option, because understandability is low for older users. Nevertheless users should be completely guided in carrying out the relevant procedure, including error corrections. It should be ‘shown’ at all times what should be done and how. Enough guidance should be given for all users, including senior citizen.

If a designer actually aims at meeting all these requirements it does not suffice to adapt existing products. Whole new interaction principles are needed. We expected that new technologies such as speech recognition and embedded help would provide possibilities to do this.

In this investigation we would like to find out whether:
the model applies for other domestic devices (besides the TV/VCR);
to what extent the model can be anticipated in a design using new technological possibilities in ICT;
whether applying such solutions actually increases usability for young and old users.

Method
The Delft Intelligence in Products group designed an intelligent thermostat (Keyson et al., 2000, TUDelft patent) based on literature and earlier research (see Freudenthal, 1999, Keyson et al., 2000). Also practical design experience from earlier work by the team members was used. A substantial part of the interface was based on the described model of usage and learning.

Once a simulation was available user tests were carried out to find out about the quality of usability (Freudenthal et al., 2001, Freudenthal and Mook, in press). The subjects used the simulation of the thermostat, one subject at a time. (The version tested was one version before the one described below; differences will be mentioned in the next section). In the trial 7 subjects with ages between 25 and 45 and 7 subjects with ages between 60 and 73 were observed. The subject was given tasks, but they were as much as possible given as contexts for use. This was done to, as much as possible, not give away clues on how to carry out the task. No ‘thinking aloud’ during the tasks was asked for. There was no paper manual available. Afterwards the subjects were interviewed mainly to find out about their opinions about the thermostat. Results from the usability test were selected for this paper, which concerned design concepts according to the model and their assessed effect on usability. Also problems in interaction were analyzed to see whether it is likely that they are caused by applying the model or by not applying the model. General behavior was analyzed to detect approaches according or in contrast with the model.

For a complete record of the thermostat design, the usability test method and results we refer to Freudenthal et al. (2001) and Freudenthal and Mook (in press).

Design suppositions
The main properties of an interface, if it is to serve the user in a way, which matches the observed ‘natural’ way of interacting with domestic appliances, are:
• It should help the user define his (main) usage goals (it may be expected that the user can remember his goal during operation, but hardly more than that).
• It should fully guide the user in setting his first and every next step to reach his goal taking into account the level of knowledge already available. It should therefore provide sufficient guidance for users with different levels of knowledge to begin with and different learning curves.
• Information from the product during usage must be completely consistent to build up the knowledge of product rules to be used in next actions; only general rules will be remembered.

In the thermostat the richness of the various modalities was used in combination to support a range of types of information. The dialog between user and thermostat is a regular conversation in spoken language (the mother language of the user). A GUI (Graphical User Interface) on touchscreen was used to present settings and allow direct manipulations for changes with instant feedback. Sound was used to support clicking and dragging on the touch screen.
Expected foreknowledge: Crucial is the level of knowledge of users. When we started we knew that knowledge of devices, menus and icons is extremely poor with older subjects. To be able to exclude as few users as possible we assumed that users would have knowledge of their mother language (not technospeak, but everyday spoken sentences) and of known devices, which have been in use for such a long period that all users would know them. We chose an 'agenda', a 'thermometer' and a 'map of the home' (in floors) and turned these into screens for direct manipulation, feedback and information overview.

The current design can be divided into two modes. The default mode is meant to be used by anybody, also an (elderly) person without any previous knowledge. This mode can overrule all programs in the thermostats 'agenda', which allows inexperienced users to not be hindered by programs set by others. Very basic instructions will be given if a user just stands in front of the device and does nothing. The thermostat will explain that the screen should be touched or that the users should give an answer from the 'things to say' list. (See figure 2.)

The advanced mode (figure 3) can be activated by a regular user, e.g. a person who lives in the home. First time authorized users already have learned some rules about the thermostat: e.g. they know that they can speak to it and that they can set temperatures by dragging on the arrows. Later on they will learn more rules. Depending on the user's age and usage experience, assessed by the thermostat (or asked) the embedded help will adapt and be present when it might be needed. For example, a novice user might ask the thermostat 'help me switch floors' (listed under 'things to say', figure 4). The thermostat would then say: “Tap on the floor you want”.

Guidance to set goals and sub goals: The possible (main) usage goals are presented to the authorized user at the beginning of (new) interaction (figure 3), so not in the middle of carrying out a procedure. They are presented as complete messages in the list of options. The user tells the thermostat his main goal and this directs the thermostat’s next step or question. New sub goals will be given through a verbal dialog and direct next steps.

Figure 2 - Non-authorized users can use only this screen. They can program a temporary setting in separate rooms. Current (black, 20) and desired (gray, 19) temperatures in the living room are displayed. To view the figure in color, please consult http://www.dcs.gla.ac.uk/~johnson/sem2002/ or ----/fria2002/
Figure 3 - The activities in the agenda can be scheduled in any desired recurrence. The user can program through direct manipulation in the agenda or main user goals can be said or touched (right part of the screen). Once a goal has been chosen next ‘things to say’ will be shown, depending on the situation and the next question posed by the thermostat.

To view the figure in color, please consult http://www.dcs.gla.ac.uk/~johnson/eam2002/ or —

Guidance to make settings: Some users will have specific wishes to set their thermostat and others will not want to bother; they just want the thermostat to see it that they do not waist energy. For all users, but in particular for this last group, it is important that the thermostat system will recognize recurring living patterns. It can present suggestions, based on these patterns, to the user. The user needs to discuss the suggestion through a normal verbal dialog. If the user agrees the thermostat can program the suggestion. If temperatures are needed from the user he is asked to set these in the relevant rooms in the map of the home (figure 4). Also the thermostat might need a name for the new activity in the agenda and asks the user to type that in on a keyboard presented at that time. The user will find the new activity in the agenda with a self-chosen name and can adjust it later if he wants to.

For users who have (gained) some more knowledge about the thermostat and devices in general there is the possibility to interact in a more active way. They can program intentional scheduling through a dialog.

They can even manipulate in their agenda without waiting for instructions or they can overrule parts of longer procedures by acting directly, without a verbal dialog.

Consistency: Consistency has been one of the main focus points in the design. For example: the user gets an overview of what he can say, so he cannot say anything else. Speech recognition allows for more commands, but we choose to always provide explicit feedforward.

Figure 4 - The rooms of the home are presented in a map of the house. The temperatures in rooms belonging to an activity can be programmed here. The left-hand side can be operated by direct
manipulation. The options at the right can be said or touched. An example of adaptive embedded help is shown (‘help me switch floors’).

To view the figure in color, please consult http://www.dcs.gla.ac.uk/~johnson/eam2002/ or —/iria2002/

Learnings from the usability trial
We found that some users indeed need instructions on how to operate through the verbal dialog, the first time they use it. Once they received instructions interaction through a spoken dialog worked very well for all subjects. Even the oldest subjects could rather effortlessly follow the suggestions. They listened to the questions and gave the appropriate answers. They appreciated the suggestions as they found the usage goals to be sensible (especially the saving of energy) and they liked the interaction style. Although the suggestions were easy to operate several younger subjects expressed their dislike of the rigidity of the wizard.

The provided guidance was sufficient, except for one family of subgoals. ‘Checking the settings made’ was not supported the way other subtasks were. For example, the provided final feedback was a flashing instance of the last changed setting. The user should scroll his agenda or the map of the home to check settings. This, however, was not the natural thing to do. Completely according to our own theory young and old users were not inclined to take sufficient action, even though they in many cases were clearly in need of such information. Because checking final settings is crucial to avoid errors in programming, we added spontaneous and complete final feedback in the present redesign. The thermostat’s voice now tells the user what settings were made. If more freedom for the younger users will be developed the thermostat will have to know when it should give the final feedback.

All subjects understood the large metaphors to present the programming data (the thermometer, agenda and map) without any additional explaining. The feedforward to indicate how the metaphors should be operated through direct manipulation were clear, such as arrows to indicate where to drag. As a comparison we had also implemented direct manipulation without explicit feedforward (such as the stairs in the home, to be tapped to go up or down). This was unclear to all subjects.

Subjects could find an activity (birthday) in their agenda; they could all scroll their agenda and set temperatures. They were very well supported by the immediate feedback during manipulations to guide the user in adjusting the agenda or a temperature, such as a little clock, which appeared next to the pulling finger when changing the time of an activity. Activation of items was indicated by sound. Changes were immediately visible in the agenda, map, etc.

In one of the procedures - making a new activity - users were required to do more than just answer questions and carry out dictated small tasks. This type of interaction worked well for the young subjects, but it was too difficult for some elderly subjects. It was also not clear whether all elderly subjects could relate to the user goal. At least some of them did not seem to feel a need to program a schedule for temperature control.

Icon understanding proved again to be extremely low, especially with older subjects. Icons, which were supported by optional embedded instructions, were easy to use. For example ‘deleting an activity’ required an icon, the trashcan, see figure 3, and was successfully used.

Error recovery was not sufficiently supported yet. Elderly subjects used the item ‘one step back’ to navigate (several times for more steps). They did not notice that it changed their settings back. Specific support anticipating the lack of knowledge about ‘spatial organization’ (of menus) and error recovery principles needs to be developed for older users.

Conclusions and discussion
A model of domestic product usage was applied as design guidance for an intelligent thermostat. The model was derived from observations of users operating a brand new TV/VCR. Based on our findings with the thermostat we expect that the model also applies for the first encounter with other domestic appliances, provided that the level of motivation to achieve usage goals and perceived risks during operation is similar.

If we compare the performance observed with the thermostat with usual performance levels in usability studies with home electronic devices, the difference is remarkable. With many current devices, which provide a huge amount of possibilities, we tend to see subjects struggle (see for example Vermeeren, 1999). Elderly subjects often get completely lost when they try to program. Our thermostat provided even more functions than the most advanced thermostats on the market. Nevertheless we observed even elderly subjects actually program complex patterns into a schedule, with relative little effort.

We found a difference between young and older users. Younger users require less support to carry out tasks, because they have more knowledge available. We even expect that they might want to work in the thermostat GUI without verbal dialogue after a while, for certain subtasks. In this context maybe even
'knowledge based' behavior could occur later on (because their level of knowledge could (become) sufficient as well as their youthful working memory).

The impression is very strong that following the model in our design was for a large part responsible for the observed ease of use. It is unlikely that it is responsible for all increase in usability, because besides the aspects related to the model; many more guidelines for design were taken into account. To mention just one: we included verbal explanations about the way the smart home system (of which the thermostat is a part) works to improve user understanding and thereby acceptance of the new principles.

Nevertheless, the way users behaved was according to what we should expect, based on the model. The subjects were supported so that they should be able to operate with little effort and did so. In those situations where the design asked for some more effort, problems occurred or subtasks were not even carried out. Both by the design and by the trial set up the subjects were encouraged to learn by doing. They complied with this effortlessly. The subjects used the device from the very first encounter and did not take the time to explore or intentionally learn. Therefore they learned by using.

This way of learning by doing – mainly of how to carry out procedures to accomplish programming goals, is not common in the usage of safety critical devices. With such devices first usage tends to take place in a separate training phase, without real risks, and usually with help from others or from simulations or training-documentation. Motivation to learn is usually higher and intentional learning is strived for. Such a learning phase will require another model of human behavior (which possibly could be an extension of this model).

Application of this other model of human behavior is not likely to mean that our model based design principles will not apply for the learning process of operation of devices for safety critical tasks. Our model based design principles comply with most well known existing design guidelines, relevant for ICT based product (interface) design. (An endless list of sources with relevant ergonomic guidelines could be mentioned here. Examples come from different disciplines, such as Pirkl and Babic (1988), for consumer electronics and appliances, http://www.useit.com, for updates on website design by Jakob Nielsen, and Schneiderman (1998) and The Mitre Corporation (1986), on desktop applications and Preece et al. (2002), who combine the various disciplines for professional and consumer products.) In many cases our guidelines are stricter or more specified for specific age groups (Freudenthal, 1999).

If young and old users can learn to program a complex thermostat by the described design aspects, it is likely that they will benefit in other situations as well. Users of professional devices have many problems to operate as well, see for example Bogner (1994). Even if the model is only partly valid for safety critical devices, consequences for design requirements of such devices could be great. Present devices which are used in safety critical situations tend not to meet the requirements depicted, e.g. no complete guidance is provided, products are inconsistently designed and the burden on working memory could become high with functions not used recently, resulting in forgotten exact steps in procedures. Therefore users of safety critical devices could especially be supported by these principles with seldom-used procedures of operation.

We have demonstrated that there are ways to apply ICT in domestic appliances so that ease of use can be substantially increased for young and old users. Older residents need not be excluded from the advantages of modern devices, just because the interfaces are too complex to operate. It is even possible to give them back the control over their own devices without help from others, or have them overrule programs in devices they have never seen before - a gerontologist's dream.

Acknowledgement
The author expresses her thanks to all team members and others who have contributed to the design and research work: David Keyson, Marc de Hoogh, Elyon DeKoven, Riejanne Mook, Liek Voorbij and Henk Arisz.

References


