TU/e technische universiteit eindhoven

Intuitive Interfaces:

A literature review of the Natural Mapping principle and Stimulus Response compatibility

Author: Arnout R. H. Fischer

Supervisor: Prof Dr Matthias Rauterberg

JFS
J.F. Schouten School for
User-System Interaction Research

1999

Table of contents

Abstract

In this paper the possibilities for research into the development of a design method for user centred design based upon seven design principles formulated by Donald Norman (Norman, 1988) are reviewed. Special attention is focused on the principle of natural mapping and the possibilities for grouping and configuration of interfaces therein.

Although it is concluded that Norman's principles could be a good way of thinking for a designer, due to the lack of formalization the principles are (at the moment) not fit for use as a method. It also appears that a significant part of the theoretical background; especially on mental models, a unified compatibility theory, and the possibilities of usability gain by grouping and configuring, are lacking.

Introduction

Many problems in operation are based upon a misunderstanding between how the designer interprets and facilitates the intention of the user and how the actual intentions of the user. These problems occur when people have to work with machines. This is often the case with simple products such as door handles. It should not be necessary to need a manual when operating a door but this is exactly what the often-used signs 'Push' and 'Pull' are. The majority of simple applications however can be used easily and intuitively. It seems that the more complicated the application gets the harder it is for the user to understand what function each control element actually operates. Infamous examples are the programming of VCR devices and the operation of the current PC systems. Although in most cases a usable solution is not found yet, there is an increasing amount of realization among designers that the technology should serve human and that humans should not be expected to perform inhuman tricks in a machine environment. In the mind of the designer the user should come first instead of the technology; this attitude towards design is called user-centered design.

When looking at the user-interface of either a computer application or a complex electronic consumer product, a number of control elements, labels and displays are designed and configured on the interface. When a user operates a product or application he will form an idea how the product and its control elements work. This idea will be formed by the use of analogies and former experience into an interpretation of the underlying structure of the operated machine. The developed interpretation of reality is called 'mental model' and problems in operating the device will be solved using this model. The more success is obtained using this model, the more firmly it is set in the mind of the user as a 'true' interpretation of the operation of the device. All kind of 'tricks' and tools are used to explain interaction effects (Payne, 1991). These tricks are applied to fit the perceived effects with the mental model. When looking at this it is of importance that the designer facilitates the user with clues that allow him or her to form the correct mental model or at least a mental model that allows the user to perform the required actions. It is not necessary to form an exactly correct model as long as the formed model is consistent with the behavior of the system in all situations. Who has a perfect internal model of for example a car? (Norman, 1998) The designer has certain possibilities to ensure that the user understands the intention of the offered control elements. In a chapter of his book 'The Psychology of Everyday Things' (Norman, 1988) concerning his view on user-centered-design Norman formulates seven principles to make complex tasks easier. Although these seven principles can help designers in their way of thinking the principles it will be shown that these principles are not formulated in such a way they can be implemented as a design method.

In the next chapter each of Norman's principles will be explained shortly. After that one of these principles: the principle of 'Natural mapping' will be explained more elaborately and its aspects will be explained. A short review of current literature on compatibility effects, which is the research field in which natural mapping is researched, will be given as well as a review of grouping and configuration effects, which are other aspects of the natural mapping principle. Finally will be concluded what work has to be done to transform Norman's seven principles into a usable design method.

1 Seven principles for User-Centered-Design

Donald Norman concludes his book 'The Psychology of Everyday Things' (Norman, 1988) with seven principles for user-centered-design to transform difficult tasks into easy ones. These principles suggest a way of thinking about user centered design to designers of applications.
His seven principles are:

1. Use both knowledge in the world and knowledge in the head
2. Simplify the structure of the task
3. Make things visible: bridge the gulfs of execution and evaluation
4. Get the mappings right
5. Exploit the power of constraints both natural and artificial
6. Design for error
7. When all else fails, standardize

The following paragraphs will give some explanations on all these principles.

1.1 Use both knowledge in the world and knowledge in the head

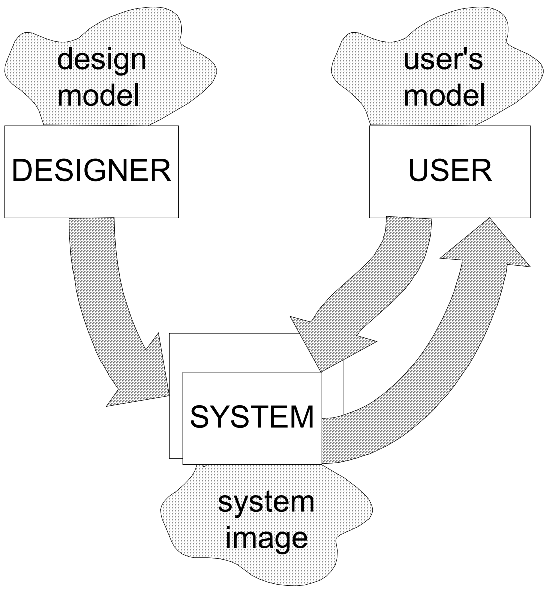

Figure 1: Three aspects of mental models: Designer - User models - System Image

Norman argues that a task is more easily learned when the needed information is available in the world. This information is only useful when it is presented in a 'natural easily' interpreted way. When a user becomes experienced less knowledge has to be extracted from external sources and more will be in the head. The offered information in the world may never interfere with the use of experienced users.
This principle depends heavily on the theory of mental models. The designer formulates a certain conceptual model of the application. During use the user also forms a certain model explaining the operation of the system. Ideally, the user's model and the design model are equivalent (see fig. 1). However, the user and the designer only communicate through the system itself. The system image of the design model is critical, this is the only way the designer can convey information of the design model to the user.

1.2 Simplify the structure of tasks

Tasks should be simple in nature, minimizing the amount of planning and problem solving actions they require.
This is where the designer should pay attention to psychological aspects of the user.
The offered stimuli should not result in an overload situation of the workload of the user. It is however also important that a minimum level of challenge is guaranteed because the user might easily make errors based on inattention and boredom (Neerincx, 1995).

This implies that the limits of short term memory (STM) and long term memory (LTM) are taken into account. This means that only five to nine unrelated items can be kept in the memory for a short time. (Miller, 1956) It appears that information stored in the long time memory is better stored when it makes sense and can be integrated into a conceptual model framework. Information retrieval from the LTM is slow and will probably contain errors (Norman, 1988).
Recommendations made by Norman are:

- Keep the task much the same but provide mental aids.
- Use technology to make visible what would otherwise be invisible, thus improving feedback and the ability to keep control
- Automate but keep the task much the same
- Change the nature of the task

This principle; simplifying the structure of tasks; urges the designer to reconsider the task he is asking the user to perform more than giving the designer ideas how to improve the interface he will be offering the user.

1.3 Make things visible: bridge the gulfs of execution and evaluation

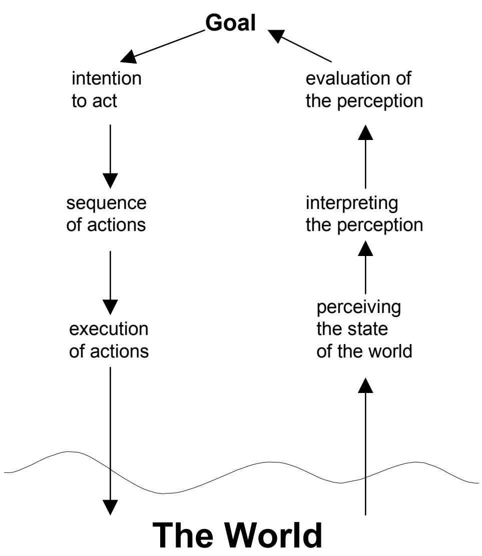

Figure 2: Norman's 7-stages of action

To explain the principle of making things visible, Norman's stages model of action has to be considered (fig. 2). The left part of the figure; made up out of the stages: intention to act, sequence of actions, and execution of actions; is the execution side. The user-interface should provide the user with information, feed-forward, to decide which actions he should undertake. The right side of the schema with the stages: perceiving the state of the world, interpreting the perception, and evaluation of the perception; is the evaluation side. The user-interface should give feedback that can be understood. Thus the user can see what can be done and what the results are.
But another aspect should also be considered. The product should supply information to all stages. Thus it should give the user support all the way from goals which contain rather vaguely specified and highly semantic information to formulate the users intention to act; down to specialized information on the execution of the actions.
Another, more formal, way of describing the abstraction levels in the operation of applications is formulated by M.M. Taylor (Taylor, 1988) in his layered protocol model. This model separates actions from (sub) goal state down to physical operation. This is however not a strict, theoretical model but an applicable protocol. The protocol of Taylor has been tested in a study in which an audio device was designed based on the implications of Taylor's protocol (Haakma, 1998). In this study it was found that when designing an audio device taking Taylor's protocol into account it was tested as more usable than a comparable device that was designed without taking Taylor's protocol into account.

1.4 Get the mappings right

This is the area of natural mapping[1]. Natural mappings are the base of the research into 'compatibility' effects in the field of human factors and ergonomics. This field has been studied for some decades now and a number of interesting outcomes were produced. Chapter 3 of this paper is completely devoted to the compatibility research, so it will not be explained any further here.

The four relationships as summed up by Norman can be seen as being related to two types of compatibility. The designer should make sure the user can easily determine the following relationships:

- Between intentions and possible actions (Stimulus-Response (S-R) compatibility (Fitts, 1953))
- Between actions and their effects on the system (S-R compatibility)
- Between actual system state and what is perceivable by sight, sound or feel (Stimulus-Cognition (S-C) compatibility (Andre, 1992))
- Between the perceived system state and the needs, intentions and expectations of the user (S-C compatibility)

1.5 The last three principles of Norman

The last three principles as summed by Norman have little relationship with the further scope of this report and will therefore be mentioned only shortly.

1.5.1 Exploit the power of constraints both natural and artificial

Design the product in such a way that only one action is possible or logical in any given situation.

1.5.2 Design for error

Assume that any error will be made. A user will make errors so the system should be designed to anticipate all possible errors and allow the user to correct them.

1.5.3 When all else fails, standardize

When something can not be explained in any way completely logical or culturally determined, make sure a universal standard is followed. When standards are followed rigorously they may well become part of a cultural stereotype and change into a natural mapping. Things like that have happened with the analogue clock, type writer keyboards, and the side of the road people drive on.

1.6 Some remarks on Norman's seven principles

When reviewing Norman's seven principles some remarks can be made.

First: Most of Norman's principles overlap each other in some way.

Especially the principles: "1: Use both knowledge in the world and knowledge in the head", "3: Make things visible: bridge the gulfs of execution and evaluation", "4: Get the mappings right", and "5: Exploit the power of constraints both natural and artificial", have much in common. All of them depend heavily on the forming of mental models, the cultural and perceptual aspects and aspects found in compatibility research.

The use of standardization (principle 7) will, eventually, also result in a cultural stereotype thus becoming more than mere standardization but part of a cultural mapping effect.

---

[1] Although the terms 'mapping' and 'compatibility' are widely spread in this type of research the term 'natural mapping' as an integrated concept is developed and used by Donald Norman.

Second: Norman's principles are the backbone of a philosophy that designers could follow.

Some problems for designers to implement these principles seem apparent:

1. The principles give no operational support in the current form to be used by designers as guidelines.
2. Although the principles look simple and straight forward the construction of operational guidelines requires the results of a lot of ongoing studies in the field of cognitive psychology. This means that a lot of the knowledge necessary to construct an overall theory is still missing.

Both factors have to be solved before it will be possible to transform Norman's principles into an effective design method.

2 Natural Mapping

One of the design principles that Donald Norman (Norman, 1988) gives is the principle that the designer should get the mappings right. The mappings meant by Norman indicate the relationships of the controls and their movements on the one hand and the results of these actions in the real world on the other hand. To get the mappings 'right' a 'natural' mapping should be used. A natural mapping is a mapping that leads to immediate understanding because the representation of the functionality in the controls is made considering physical analogies and cultural standards. Although this seems a straightforward approach it includes at least two open questions. One of these questions is what parts of the interface can be described using natural mapping and which effects, generating a natural mapping can be used. Second a more fundamental question is what exactly are physical analogies and cultural standards. Research programs in experimental psychology and human-factors sciences are working on this latter question for almost fifty years. This research area is called the area of S-R compatibility, (stimulus-response compatibility).

Possible aspects that can be of aid to the designer are (according to Norman, 1988):

- spatial analogy (the control space is an immediate mapping of the response space, e.g., lighting lights using the same pattern for the controls as for the lights (Fitts, 1953))
- cultural standards (the control-response mapping follows a culturally determined convention, e.g., clockwise is more (Brebner, 1976))
- biological effects (the biological functions of the user have effect, e.g., in dual task performance the stimulus offered in the same hemisphere of the user as the hemisphere that directs the response offers faster response time and lower error rate (Wickens, 1984))
- perception effects that allow for natural grouping and patterning of controls and feedback (similar controls are grouped in a way to reveal their relation, e.g., placing certain displays close together as a group to show their functional equality (Wickens, 1990), proximity compatibility)

The first category, that of spatial analogy, has been researched in the studies looking into stimulus-response compatibility (Fitts, 1953) that have supplied an enormous amount of data in a wide diversity. One underlying theory of the compatibility effects has not been formulated yet but the understanding of this phenomena has

increased to a point where predictions can be made about compatibility effects among seemingly unrelated situations (Proctor, 1990).
The second category concerns cultural analogies. This category has a lot of problems since certain perceptual values are far from the same in different cultures. An example of cultural difference is the case of the names allocated to colors and the interpretation of colors. Not in all cultures have the same colors the same emotional power. Even the number of colors with distinct names differs from culture to culture in a range from two to fourteen (Smets, 1986).
An important aspect in this line of research is the segregation of cultural analogies between generations. This issue is in focus in Europe because an increasingly large part of the population is elderly. Some interesting results on differences in the construction of mental models due to experience and diminished mental abilities were found (Walsh, 1993) but the research field seems relatively unexplored as yet.
A number of methods to describe cultural differences are reviewed by Nancy Hoft (1996). One of the problems in describing cultural factors is that they are susceptible to change.
The third category of biological aspects is also fairly complex. It involves biological factors such as those studied in the research projects of Christopher Wickens (1984) where auditory stimuli were offered to one ear and visual stimuli offered to the opposite eye. This study was originally meant to investigate parallel processing of multiple stimulus-response pairs for aviators. When a tactile response was required from the hand at the same side of the user where the visual stimulus was offered, meaning that the brain processes were confined to one hemisphere both response time and error rate showed a better performance
The fourth category of grouping and configuration received little attention. Most of this attention was primarily directed at perception and was aimed at display groups and hardly for controls. One exception is perhaps the case where Norman comes up with an example of the seat operating system of a Mercedes car, in which case the components of the operating system form visual rendering of the actual seat (Norman, 1988). In the research into display configuration and grouping, it seems that some theories are available that might be used to form ideas about configuring and placing control elements due to their function. The available theories should be mainly looked for in: perception theories such as Gestalt theories, in graphical presentation and in function analysis.

2.1 Problems with the use of Natural Mapping as a design aid

Before natural mapping can become a useful design aid a number of problems has to be solved:

1. The concept of natural mapping holds no specific guidelines or strict methodology and is therefore not operational for designers. To overcome these problems the following aspects should be considered:
   a) An overall framework of the boundaries and possibilities of natural mapping should be formulated.
   b) All current available data and models should be translated into a way that can be fit into the framework for natural mapping.
2. Some of the necessary data to construct a methodology based on natural mapping is missing or not organized. Progress has to be made to a certain point in the following field to allow the construction of a design methodology based on natural mapping:
   a) A single underlying theoretical description of the data from compatibility research is missing

b) Methods to describe the construction of mental models is incomplete
c) The available data on cultural aspects and differentiation is scarce and hardly organized into formal theories
d) Data on grouping and configuration is scarce and hardly organized

3 S-R compatibility

The concept of stimulus-response compatibility has become a fundamental concept in designing an interface for use. This principle states that when a stimulus and a response share common factors the reaction of the user will be both more efficient and more effective. In the first study using the name of compatibility a stimulus set consisting of lights in a certain pattern and a response set of controls in another pattern were tested together. It was shown that for every layout in this experiment both the response time and the error rate scored significantly better when the stimulus set and the response control set share the same pattern (Fitts, 1953). This congruency in shape is also called spatial compatibility. Soon after this first study several programs were set up to determine whether other types of compatibility could be found.

3.1 Research projects

In the time since S-R compatibility has become an important research field. Several related research projects have been examined in a large number of papers. Some of the more important effects will be shortly discussed in the following paragraphs.

3.1.1 The Stroop effect

One of the earliest studies in which the effects of incongruency between display aspects was shown, was the study in 1935 by J. R. Stroop (MacLeod, 1991).
Stroop discovered a significant interference between the speed of reading the name of a color and the ink color of the word.
When subjects were asked to read out loud a color name that was printed in almost any hue of colored ink the color name was read as fast as the same color name printed in black ink. Only when the color of the ink matched the color name printed in text a small difference in reaction time was found.
When subjects were asked to give the color of the ink of the message it took significantly longer to interpret a color when the printed word, a color name, did not match the ink color. This was compared with asking the color of the ink of control squares and neutral words, no color names, printed in colored ink.
This effect is called the 'Stroop-effect'. In the years since Stroops article was published over 700 studies into the Stroop effect were published many of which were highly theoretical, based on many of these articles Colin MacLeod has written an extensive review (MacLeod, 1991).

3.1.2 A review of four burner stove experiments

One example, in which was tried to formulate a spatial analogy between controls and operated functions, is the four-stove burner plate. In the first paper on this

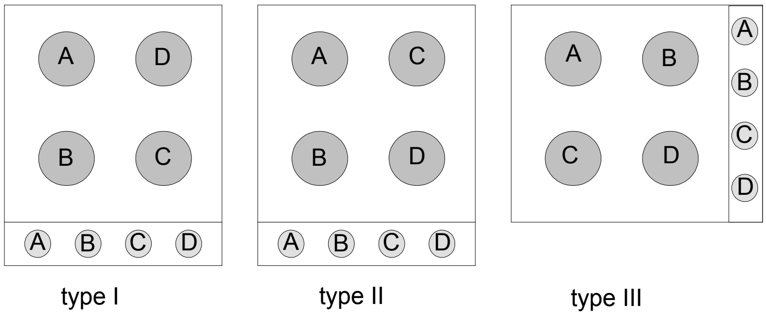

Figure 3: Four burner stoves (adapted from Chapanis 1959 and Chapanis 1967)

issue it was shown that a configuration in which the linear control set was 'wrapped' around the square response set (fig 3; type I) had faster response time and less errors than a configuration in which the controls are placed otherwise (fig 3; type II) (Chapanis, 1959). Later research using a questionnaire (Shinar, 1978) came up with different results than Chapanis. It appeared that in a paper and pencil situation the preference was for a layout consistent with western writing (A top left, B top right, C bottom left, D bottom right). Chapanis results were confirmed in other experiments where the performance of users was monitored (Ray, 1979). Other questionnaires showed consistency with Shinar's experiment (Smith, 1981). In 1995 it was found that naïve subjects, meaning no experts in design or human-factors fields, believed Shinars solution more often favorable than Chapanis solution (Payne, 1995) showing that compatibility effects are sometimes far from intuitive. It appears that the layout most liked is not always the layout that is most favorable for actual use (Sanders, 1987).

In another research study, in which an actual experiment was performed, it was shown that when a burner stove with the controls on top of the stove to right were placed in a line (fig 3; type III) the western writing style held preference in use over the 'wrapping' as it did in the other layout examples (Chapanis, 1967). None of the layouts tested in this project with a configuration in which the topmost two plates were not operated by the topmost two controls had acceptable performance. So it is clear that spatial compatibility where possible (top – bottom) has a high priority (Chapanis, 1967). According to this finding the more important factor in the earlier research appeared to be the spatial relation between the left-right groups (columns) which did not allow the writing style.

In other larger control panel it was examined whether sensor lines, meaning visual lines connecting the control operators to the related response operators, might improve the operation (Chapanis, 1965). It is shown that the effect of sensor lines can be especially helpful for non-compatible lay-outs. For small displays the positive effect is high during the first trials but the effect decreases to insignificant level as the user gains more experience. When applied on large control-response panels the effect decreases a little but stays well above significance levels. Note that non-compatible lay-outs, even with sensor lines, score in every case worse than compatible lay-outs (Chapanis, 1965). However even in small interfaces the use of sensor lines reduces the chance for errors in critical situations, even after a learning period (Osborne, 1987). This shows that sensor lines could help solve learning problems for non-compatible layouts and after that give support in critical situations.

3.1.3 Experiments on the effect of scale side

In a research project the effects of operating a linear scale with a rotary control was examined (Brebner, 1976). Brebner expected three effects for this type of operation: (1) the culturally determined rule that clockwise turning means an increase of value (see fig. 4a), (2) the seemingly 'physical' links between the control and the display.

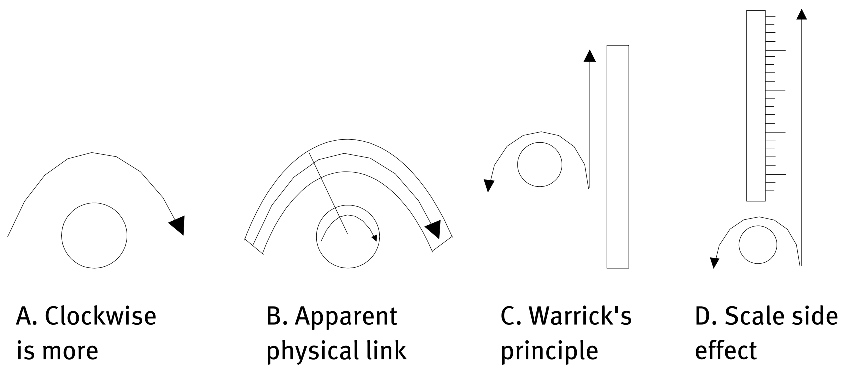

Figure 4: Rotary control effects

When using a rotary control this applies to a semi-circular display (see fig. 4b). The final effect was (3) Warrick's principle (see fig. 4c). Warrick's principle constitutes that the pointer is expected to move in the direction of the point on the control element closest to the indicator. Another effect was studied in this experiment, the scale side effect (see fig 4d). The scale side effect states that the pointer will move in the direction of the movement of the side of the control at the same side as the scale of the display. In Brebners test (Brebner, 1976) it was shown that the scale side effect had significant influence. Furthermore it was shown that effects could interfere with each other, enhancing or diminishing mutual effects. In a re-examination of Brebners results the additional information was found that of the effects of influence Warrick's effect was dominant (Petropoulos, 1981). Some criticism on Brebners results was given in another re-examination when Courtney (Courtney, 1992) found that 19% of his subjects were not influenced by the compatibility effects that Brebner described. Furthermore Courtney found that the scale side effect also applies when no scale is used but only the triangular arrow shaped pointer as used in the experiment of Brebner. Courtney concluded that the effect of the shape of the pointer had a greater influence than the scale side.

3.1.4 The Simon effect

An effect that makes predictions on the interference between the source of the offered stimulus and the desired response is the Simon effect (Simon, 1990). In an experiment in 1967 Simon discovered this effect as a side effect of an experiment on the response of people with different preferred hands (test subjects were either strongly left or strongly right handed). In this experiment subjects were offered a stimulus consisting of the words 'right' or 'left' in either the right or the left ear and were asked to operate a switch with either the right or the left index finger consistent with the stimulus. It appeared that the stimulus to operate the right switch offered in the right ear resulted in a faster response time than the stimulus word 'right' offered in the left ear. A similar effect applied for the 'left' stimulus. Whether the subjects were right or left-handed had far less influence. This led to the discovery of the Simon effect implicating that a stimulus offered on the same side as the desired response resulted in a faster response time. A series of experiments showed that this effect is based on perceptual factors and has effects on visual stimulus sets as well as auditory. The auditory effect appeared when pure tones not prior associated with left-right response were used. Shifting the stimuli, from right to left, resulted in a gradual shifting of the preference. By applying noise the effect could be weakened and even reversed. By applying irrelevant directional clues a distortion effect on the processing can be found. Based on this it is suggested to design displays in such a way as to avoid such clues.

3.1.5 Effect of response selection rules

An experiment was performed in which four numbered stimuli (from 1 to 4) corresponded to four response keys placed in a line parallel to the stimuli (A, B, C, D) (Duncan, 1977). The subjects were asked to respond on the illumination of the

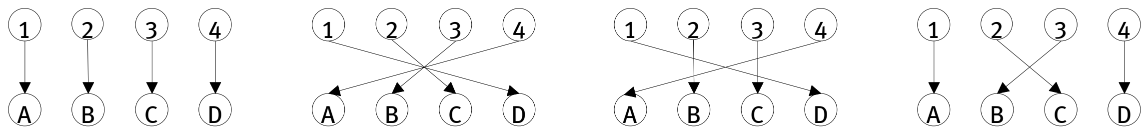

Figure 5: Duncan's four conditions

stimuli lights. Four different conditions were tested (see fig.5). The first condition was a condition where a spatial link was made between the numbers (1 to A, 2 to B, 3 to C, 4 to D). In the second condition the controls were reversed 180° meaning that 1 corresponded to D (2 to C, 3 to B, 4 to A). Both other conditions were mixed conditions. In the third condition the outermost mappings were reversed (1 to D, 4 to A) while the middle were spatially linked (2 to B, 3 to C) while in the fourth condition the outer links were spatially connected and the inner links were crossed. It appeared that the first condition resulted in the fastest reaction times with the least errors. The second condition scored second best on both factors and far better than condition three and four. According to Duncan this could be explained by the construction of a single rule that reversed the order of the stimuli and responses according to a spatial compatibility. The conditions three and four however do not answer to a single conversion rule, which complicates the operation. It might be stated that the mental model for the operation is more complex in condition three and four, resulting in a longer learning time and for novice users less accurate operation. Note however that the term mental model was no common practice to describe this type of interaction at the time of this experiment.

3.1.6 Proximity compatibility

In Wickens' research program it is suggested that proximity either in color or in spatial distance can provide a powerful clue to relation between controls and displays (Wickens, 1990). A number of possible manipulations to change the proximity of display elements are given. The object integration can be altered in spatial proximity aspects such as: connections, source similarity and code homogeneity and by object integration aspects like: contiguity, contour and spatial integration can be altered. Many of these manipulations follow Gestalt laws of perceptual organization (Wickens, 1995). Two mechanisms seem to underlie object display advantages for integration. One of these mechanisms is the production of task-relevant emergent features that provide visual shortcuts for the mental integration of data values. The second mechanism is not explained. It appears that even in heterogeneous sets a proximity compatibility effect can have an influence.

This might be explained by the theory that it is easier to divide attention within an object than between objects or by an alternative theory that the spatial compactness allowed attentional resources to be focussed more narrowly (Carswell, 1996).

3.1.7 Biological factors

It was shown by Wickens et. al. (Wickens, 1984) that stimuli offered to the right hemisphere resulted in lower response times and lower error rates when response from the right hand was required. Wickens used these findings for research in dual task missions where auditory stimuli were offered in one ear and visual stimuli were offered in the opposite eye. It appeared that when the stimulus and corresponding response sets were kept separated for processing between the two hemispheres both response time and error rate were superior than in the case where the stimulus and response requirements were crossed. This experiment was one of the main indications for Wickens' S-C-R (Stimuls-Cognition-Response) compatibility theory (this theory will be further discussed in 3.2.2).
In a number of experiments considering the compatibility of vertically oriented stimuli and horizontally oriented response Weeks et. al. found (Weeks, 1995) that the stimulus-response mapping up-left/down-right is more compatible than the mapping up-right/down-left for responses executed by the left hand in the left hemisphere, but this relation is reversed for responses executed by the right hand in the right hemisphere.

3.1.8 Destination compatibility

In a re-evaluation of results presented in a paper by Claire Michaels (1988) Proctor and his colleagues (1993) state that the phenomena observed by Michaels should not be assigned to the perception of affordances as Michaels had done. Proctor concluded that an explanation along the lines of a spatial coding of relative direction shows more promise. Both the original experiment by Michaels and Proctor's experiments show that a stimulus, which seems to move in a certain direction results in a shorter response time in the case when the stimulus appears to move in the same direction as the required response. In his article Proctor calls this phenomenon 'destination' compatibility.

3.2 Integrated theories

A number of attempts was made to unify the compatibility research in one single theory. Many of the studies mentioned below were done in the late 1980's and early 1990's. Although all researchers emphasize that follow-up studies are highly wanted hardly any further studies were found that follow the paths of these studies.

3.2.1 The overview by Proctor

In 1990 Robert Proctor and Gilmour Reeve published a collection of papers that covers a wide range of research topics of the field of S-R compatibility (Proctor, 1990). In their (editorial) closing chapter they state that '*Although a theoretical explanation of compatibility phenomena has been slow in coming, the understanding of the phenomena has increased substantially*'. Thus although no comprehensive theory is formulated yet several general points among the outcomes of the experiments can be pointed out, which are presented here:

1. S-R compatibility effects occur in a wide variety of situations, from basic perceptual motor tasks to highly complex programming tasks.
2. A conceptual correspondence, rather than physical correspondence of the mappings as perceived in the real world is the source of S-R compatibility effects.

3. For spatial-location stimuli assigned to responses at different locations, compatibility effects occur regardless of whether the stimulus location is relevant or irrelevant for determining the correct response.
4. S-R compatibility effects arise primarily from a stage of information processing that is referred to as the translation stage or the response-selection stage. These effects are independent from effects of Response-Response compatibility.
5. The codings of stimulus and response sets, and how these codings relate, play an important role in most S-R compatibility effects.
6. The coding system is hierarchical but flexible. Meaning that effects have a certain hierarchical order but that a situation can alter the importance of an effect.

Several models of S-R compatibility are developed in a range of research programs. Although the terminology varies Proctor distinguishes three models of description of S-R compatibility

1. Attentional models. The models emphasize the direction of attention to location.
2. Coding models. These models are favored as being a complete explanation of compatibility effects where attentional factors have been rules out. Based on the coding hypothesis in performing a task at a certain moment a response code has to be formed that facilitates the user to perform the task and while at the same time a stimulus code is offered to the users mind. Congruency between the formulated response code and the perceived stimulus codes facilitates easy operation.
3. General information processing models. These models have been developed in Human-Factors engineering to enable the consideration of S-R compatibility in the design process.

Most models rely heavily on mental coding and the description of mental models.

3.2.2 S-C-R compatibility

Christopher Wickens (Wickens, 1983) suggests to explain the concept of stimulus-response compatibility with the aid of a third factor, a mediating central processing or cognitive (C) compatibility factor. His reasons to propose this solution are twofold: First it seems necessary in extremely complex systems that an operator does not act on a stimulus directly but integrates the stimulus in his mental model of the current state of the system. Only then the operator constructs consciously with the aid of cognitive power a response. Second cognitive psychology has suggested that there are two fundamentally different codes underlying the central processing operations; approximately labeled spatial and verbal coding. Support for the theory of S-C-R compatibility can be found in differences between auditory and visual reactions and certain C-R (central processor-response) compatibility effects. The argument for S-C-R compatibility lies in the possibility and problems in transforming an auditory / vocally based process to a visually / spatially based process. In certain cases such as reading the altimeter of an aircraft the S-C combination explains the effects while the perception of the altimeter while changing the height relies on S-R compatibility (Andre, 1992). A further use for this theory can be found when applied to dual task loads in which one of the tasks is operated through a spatial-visual system and the other through a verbal-auditory system (Wickens, 1984). This application holds some promises in cockpit design although attention has to be paid to more resource competition and violation of cockpit tradition.

Some criticism on the S-C-R approach was given by Eberts and Posey (Eberts, 1990). Although they see the potential to explain certain aspects of the observed effects, the reviewers lack a method to extract experimental data with high validity. They also have some criticism on the role of the mental model in Wickens' work were it is only the binding factor for the central processing factor. The reviewers state that the mental model is more important and already shaped considerably by training and prior experience. Another study looking into the differences between the S-C-R compatibility model and the theory of dimensional overlap as suggested by Kornblum (see 3.2.3 of this paper) (Guadagnoli, 1994) suggests that the S-C-R theory is more in accordance with their experimental results than the theory of dimensional overlap. Guadagnoli maintains some reservations for his own results mainly because the task the experiment was conducted upon was a very simple task.

3.2.3 Cognitive model with the use of dimensional overlap

To construct the theory of dimensional overlap a number of experiments in the S-R compatibility research field were examined. A general cognitive processing diagram was proposed to model the processes involving compatibility (Kornblum, 1990) and is represented in block diagrams. In these block diagrams clues on the mapping are placed on one path of the diagram and the knowledge of the user (e.g., experience) are placed on another path.

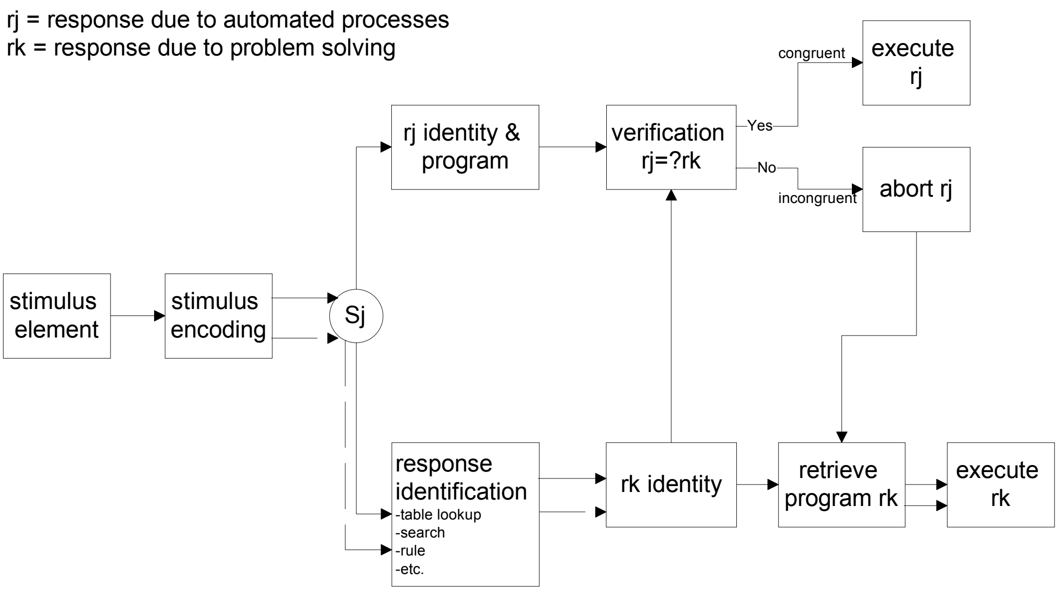

Figure 6: Dimensional overlap as a cognitive explanation for S-R compatibility

When these paths overlap a situation of either compatibility or incompatibility is in effect (see fig 6.). When these paths do not overlap no compatibility effect, neither compatibility nor incompatibility, but an indifferent process appears.
When the indifferent path is taken problem solving processes are applied. When an incompatible situation appears the user will have to cancel the automatic action triggered by the seemingly compatible mapping and has to fall back to resolving the problem using problem solving processes.

With the use of this view on compatibility effects some useful connections between seemingly unrelated effects could be made.
Experiments conducted to compare results as predicted by the theory of dimensional overlap and as predicted by the theory of S-C-R compatibility as suggested by Wickens (see 3.2.2 of this paper) give strong indications that for simple situations this theory holds little predictive power (Guadagnoli, 1994).

When applying and enhancing the dimensional overlap (DO) model Kornblum is able to code the Stroop and Simon effects in the same terminology as traditional S-R compatibility effects (Kornblum, 1995). In this way Kornblum finds eight different S-R ensemble types when he differentiates for multiple stimuli (such as the Stroop effect). These possibilities are compared on the presence of relevant and irrelevant stimuli and on S-S (Stimulus-Stimulus) interaction.
Based on these eight taxonomies a model is presented in which the reaction time for different tasks is predicted (Kornblum, 1999). This model was implemented for four types of interaction in a first article, a second article treating the other taxonomies is forthcoming (Kornblum, in preparation). A computer application is presented to simulate the different possible reaction times (the DO '98 model[2]).

3.2.4 GOMS based hierarchic description of compatibility

Based on the GOMS method (Card, 1983), three leading compatibility experiments (Fitts and Seeger, 1953; Duncan, 1977; Morin and Forrin, 1962) were analysed (Rosenbloom, 1988).
A description of these experiments was made based on goal hierarchy. The assumption was taken that the total reaction time was constructed from a certain processing time for each processing cycle with an added initial base time.
For the separate experiments a high fit was found between the proposed model based on the hierarchic method and the experimental results (Duncan, r^2 = 0.981, Fitts, r^2 = 0.999 and Morin r^2 = 0.900). A slight difference in cycle processing time was found between the experiments. There was however a large difference between the base times.
When the results were combined with a single cycle processing time a high fit was found (r^2=0.933). To obtain this fit the values had to be corrected for base times typical for each of the experiments.
This last leads to the main problem associated with the model of Rosenbloom and Newell. The model can simulate (afterwards) effects when the base time is added for an existing test, but because this value is not known for new situations, it cannot predict results in a new test
Another problem or important omission in this model is the fact that it cannot simulate the error rate, another important indicator of compatibility.

4 Grouping and Configuration

The use of grouping and configuring could help in translating the internal structure of the application toward a psychological understandable structure.
Only a few papers were published in which configuration of control elements as a powerful tool is described.
Some early recommendations on the configuration of control panels of large industrial plants were made (mid 1960's end of 1970's) but those recommendations were mainly concerned with the anthropometric aspects of the control panel (Sargent, 1997).
One of the few recent papers takes on the design of a control panel of a nuclear power plant. Sargent and her colleagues order the controls and displays according to a cognitive ordering based upon the analytic hierarchy process of Saaty (Saaty, 1977). Although the method suggested by Sargent is highly formalized and has

---

[2] The Kornblum Lab Homepage, *http://www-personal.umich.edu/~kornblum/*, University of Michigan, Michigan: USA

certain values it can be doubted whether it will be of major use in the design of simpler interfaces. Sargent mentions the lack of research on this field in the paper of 1997.

For the use of alphanumeric computer screens Tullis performed a series of experiments to show that clustering and grouping of certain information can give better response times and less errors. These experiments were summarized in the development of a computer program (Tullis, 1985) and a chapter in the Handbook of Human Computer interaction (Tullis, 1988).

Basic research into cluster analysis for computer graphics as an aid for relational judgement based upon theories of perceptual grouping was done by Liu and Wickens (Liu, 1992). Wickens performed other research programs in which relations between displays according to Gestalt (Wickens, 1992) laws and the proximity compatibility which is derived from the Gestalt law of proximity (Wickens, 1995) are shown. Most of these research programs are directed at display configuration and design or even multi-sensory stimuli and are of a high abstract and fundamental level, which means that it cannot be easily applied in a wide range of applications.

Gestalt laws themselves were formulated early in this century to explain holistic effects that could not be explained by traditional, compositional perception theories. One of the forerunners of the Gestalt theories, Koffka, formulated (amongst others) Gestalt laws of proximity, (Koffka, 1935). These theories had a high impact and influenced later research into the perception of visual stimuli. The view of Gibson on the perception of vision (Gibson, 1950) has certain aspects in common with the Gestalt psychologists, but his more holistic and ecological approach can explain more actual facts. It is not in the scope of this report to embellish any further on this widespread field of research and theory, although these lines are far to few to even give a glimpse of the ideas of either Gestalt or Gibson.

4.1 Cluster analysis in relational judgement

Based on work of Garner and Pomeranz; Liu and Wickens (Liu, 1992) performed tests to validate the proximity compatibility of Wickens (Wickens, 1990). Subjects were asked to detect the number of clusters for four representations of a data set in a matrix form. These four representations are a digit form, clustered by size, clustered by color or clustered by 3-D line length[3]. For all representations, except for 3-D line length, observations from clustered displays recorded significant smaller error scores and faster response times than observations from the data set put in a random order. This experiment suggests that color, size and digit representation of data in clusters enhances the awareness of people that data points are related.

4.2 Alphanumeric computer displays

For the use of alphanumeric displays experiments have been performed that show that an effect on human performance can be found for overall density, local density, and grouping of variables (Tullis, 1983). The performance seems to be related to measure described by a '∩' shape function (inverted 'U' shape as Tullis calls it). Thus, at low levels of the measure increases have a beneficial effect while at high levels of the measure increases have a detrimental effect (Tullis, 1983). No empirical evidence was found on the effects of layout on human performance. Structured information was faster read and understood than narrative information. When related information is grouped the search time in general improves even if the

[3] 3-D line length indicates a line in a 3 dimensional orthogonal space with its base in the x-y plane and a certain length in the z-direction.

overall density is high. When the average size of groups remains below a visual angle of 5 degrees the search time is most closely related to the number of groups. With visual angles above this angle of 5° the search time is most closely related to the size of the groups (Tullis, 1986). Based on these findings Tullis has developed a computer program that analyses display layout for alphanumeric displays and suggest grouping and layout changes (Tullis, 1985). In a chapter on display design issues Tullis recommends a number of factors in the layout to take into account. Among those are: sequence of use, conventional usage (standardization), importance, frequency of use, generality versus specificity, alphabetical or chronological (Tullis, 1988).

4.2.1 Other automated layout construction

A number of other publications in which computer generated grouping is formalized have been published. Computer programs were derived from these formalizations to perform the task of configuring control panels and even control rooms (Hendy, 1989).
Sargent (Sargent, 1997) found only limited literature on the field of configuration but in this research program not only anthropometric but also cognitive aspects, albeit in limited amount, were considered. A mathematical model to construct interfaces of large layouts such as power plants was built.

4.3 Anthropometric view

A number of publications on the configuring of control elements from an anthropometric view were published (Sanders, 1987). These configurations were examined in particular for control rooms of large industrial plants. A major point of view in the design of these layouts was an anthropometric view.
The configuration has to be optimal for human sensory, anthropometric and biomechanical characteristics. Hardly any attention was paid to the understanding of the provided layout by the user although Sanders gives an example were a control panel is enhanced by emphasizing control relations by drawing boxes around related controls. Most connections have to be made intuitively though.

4.3.1 Principles of arrangement

Sanders (Sanders, 1987) recommends several principles of arrangement

- Sequence of use principle – Sequences of actions should be followed in the design of the layout. This principle lowers the time of operation by a large amount.
- Functional principle – Related functions should be placed together. This reduces the mean time of operation considerably.
- Importance principle – Vital functions should be placed in easily accessible places. Although it does not improve operation time the principle is still recommended because less errors especially under stressful circumstances are to be expected.
- Frequency of use principle – Functions that are used often should be located in convenient positions.

When a layout is designed according to the sequence of use principle the highest operating speed is achieved, followed by an interface designed on the functional principle.

5 Conclusions

The use of natural mapping and results of S-R compatibility studies is limited by a number of aspects.

Concerning the seven principles for changing difficult tasks into simple ones some remarks can be made. It can be noted that except the fact that Norman's principles are formulated in a narrative rather than a formal way the current state of the research also does not allow making the seven principles of Norman useable as a set of operational design guidelines.
Necessary developments are:

- The development of a methodological framework in which all principles are incorporated.
- The formalization, unification and complementation of all relevant knowledge.

To render the concept of natural mapping operational as an independent design aid certain developments are necessary:

- A formal framework describing all factors involved in natural mappings should be formulated.
- All factors that are placed into this framework should be described in similar formal ways.
    - Taking together all the numerous research projects on this topic, a 'general' compatibility theory has to be formulated. In this theory result from several types of compatibility must be united.
    - A general formalization of the involved cognitive work and the construction of mental models must be developed.
    - A description of a way to organize different cultural user groups and how to describe the involved factors must be developed.
    - An overview of the effects and real influence of biological factors must be made and formalized.
- Additional research into some of the aspects of natural mapping has to be done:
    - The power and possibilities of grouping and configuration in natural mapping should be investigated.
    - A more exact description of mental models should be generated
    - More knowledge on the effects of culturally (and otherwise) determined differences between users should be found.

When in the future all these aspects are fulfilled Norman's seven principles could form the backbone of a design methodology. With the current state of research natural mapping alone could be a powerful tool in developing user products and applications. It might be possible to build a full methodology based on this first start. This is partly so because other principles are based on similar factors which play an important role in natural mapping. However some research into fundamental research areas has to be performed and the concept of natural mapping should be formalized before this could be applied.

Bibliography

Andre, Anthony D. & Wickens, Christopher D. (1992) Compatibility and consistency in display-control systems: implications for aircraft decision aid design, *Human Factors, 34(6),* 639-653

Brebner, John & Sandow, Bevin (1976) The effect of scale side on population stereotype, *Ergonomics, 19(5),* 571-580

Card, Stuart K., Moran, Thomas P. & Newell, Allen (1980) Computer text-editing: An information-processing analysis of a routine cognitive skill, *Cognitive Psychology, 12,* 32-74

Card, Stuart K., Moran, Thomas P. & Newell, Allen (1983*) The psychology of human-computer interaction,* Hillsdale, NJ, USA: Lawrence Erlbaum

Carswell, C. Melody & Wickens, Christopher D. (1996) Mixing and matching lower-level codes for object displays: evidence for two sources of proximity compatibility, *Human Factors, 38(1),* 1-22

Chapanis, Alphonse & Lindenbaum, Lionel E. (1959) A reaction time study of four control-display linkages, *Human Factors, 1(nov-1959),* 1-7

Chapanis, Alphonse & Lockhead, Gregory R. (1965) A test of effectiveness of sensor lines showing linkages between displays and controls, *Human Factors, 7,* 219-229

Chapanis, Alphonse & Mankin, Donald A. (1967) Test of ten control-display linkages, *Human Factors, 9,* 119-126

Courtney, Alan J. (1992) Control-display stereotypes for multicultural user systems, *IEEE transactions on systems, man and cybernetics, 22(4),* 681-687

Duncan, J. (1977) Response selection rules in spatial choice reaction tasks, in: Dornic, S. (Ed.) *Attention and Performance VI,* Hillsdale, NJ: Erlbaum, 49-61

Eberts, Ray E. & Posey, Jack W. (1990) The mental model in Stimulus-Response Compatibility, in: Proctor, Robert W. & Reeve, T. Gilmour (Eds.), *Stimulus-Response Compatibility,* Amsterdam, NL: Elsevier Science, 389-426

Fitts, Paul M. & Seeger, Charles M. (1953) S-R Compatibility: Spatial characteristics of stimulus response codes, *Journal of Experimental Psychology, 46,* 199-210

Gibson, James J. (1950) *The Perception of the Visual World,* London, UK: Allen and Unwin

Guadagnoli, Mark A. & Reeve, T. Gilmour (1994) Stimulus-Response compatibility and motor-programming effects: A test of theoretical accounts of compatibility, *Human Performance, 7,* 291-304

Haakma, Reinder (1998) *Layered feedback in user-system interaction (Ph.D.-Thesis),* Eindhoven, NL: Philips Electronics / Technical University of Eindhoven

Hendy, Keith C. (1989) A model for Human-Machine-Human interaction in workspace layout problems, *Human Factors, 31(5),* 593-610

Hoft, Nancy L. (1996) Developing a cultural model, in: Nielsen, Jakob & del Galdo, Elisa (Eds.) *International interface design,* 41-73, Chichester, UK: Wiley

Koffka, K. (1935) *Principles of Gestalt,* New York, USA: Harcourt Brace

Kornblum, S., Hasbroucq, T. & Osman, A. (1990) Dimensional overlap: cognitive basis for stimulus-response compatibility – A model and taxonomy, *Psychological Review, 97,* 253-270

Kornblum, Sylvan, Stevens, Gregory T. & Requin, Jean (1999) The effects of irrelevant stimuli: 1 The time course of Stimulus-Stimulus and Stimulus-Repsonse consistency effects with Stroop-like stimuli, Simon-like tasks and their factorial combinations, *Journal of Experimental Psychology: Human Perception and Performance, 25,* 688-714

Kornblum, Sylvan & Lee, Ju-Whei (1995) Stimulus-Response compatibility with relevant and irrelveant stimulus dimensions that do and do not overlap with the response, *Journal of Experimental Psychology: Human Perception and Performance, 21,* 855-875

Kornblum, Sylvan (in preparation) *The effects of irrelevant stimuli: 2 Congruence, Stimulus-Stimulus consistency and Stroop tasks.* Manuscript in preparation

Liu, Yili & Wickens, Christopher D. (1992) Use of computer graphics and cluster analysis in aiding relational judgement, *Human Factors, 34(2),* 165-178

MacLeod, Colin M. (1991) Half a century of research on the Stroop effect: An integrative review, *Psychological Bulletin, 109(2)*, 163-203

Michaels, Claire F. (1988) S-R compatibility between response position and destination of apparent motion: evidence of the detection of affordances, *Journal of Experimental Psychology: Human Perception and Performance, 14(2)*, 231-240

Miller, George A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information, *Psychological Review, 63(2)*, 81-97

Morin, Robert E. & Grant, David A. (1955) Learning and performance of a key pressing task as a function of the degree of spatial stimulus-response correspondence, *Journal of Experimental Psychology, 49*, 39-47

Neerincx, Mark A. (1995) *Harmonizing tasks to human knowledge and capacities (Ph.D.-Thesis)*, Groningen, NL: Rijskuniversiteit Groningen

Norman, Donald A. (1988) *The Psychology of Everyday Things* (reprinted *as 'The Design of Everyday Things'*, 1998), New York, USA: Bantam Books

Norman, Donald A. (1998) *The invisible computer; Why good products can fail, the personal computer is so complex and information appliances are the solution*, Cambridge Massachusetts, USA: MIT press

Osborne, David W. & Ellingstad, Vernon S. (1987) Using sensor lines to show control-display linkages on a four burner stove, in: *Proceedings of the human factors society – 31st annual meeting – 1987*, 581-584

Payne, Stephen J. (1991) A descriptive study of mental models, *Behaviour & Information Technology, 10(1)*, 5-21

Payne, Stephen J. (1995) Naïve judgements of Stimulus-Response compatibility, *Human Factors, 37(3)*, 495-506

Petropoulos, Harry & Brebner, John (1981) Stereotypes for direction-of-movement of rotary controls associated with linear displays: the effects of scale presence and position, of pointer direction, and distances between the control and the display, *Ergonomics, 24(2)*, 143-151

Proctor, Robert W. & Reeve, T. Gilmour (1990) Research on Stimulus-Response compatibility: toward a comprehensive account, in: Proctor, Robert W. & Reeve, T. Gilmour (Eds.), *Stimulus-Response Compatibility*, Amsterdam, NL: Elsevier Science, 483-494

Proctor, Robert W., Zandt, Trisha van, Lu, Chen-Hui & Weeks, Daniel J. (1993) Stimulus-Response compatibility for moving stimuli: Perception of affordances or directional coding, Journal of Experimental Psychology: Human Perception and Performance, 19(1), 81-91

Ray, R. D. & Ray, W. D. (1979) An analysis of domestic cooker design, *Ergonomics, 11*, 1243-1248

Rosenbloom, Paul S. & Newell, Allen (1988) An integrated computational model of stimulus-response compatibility and practice, in: Bower, G. H. (Ed.) *The Psychology of Learning and Motivation 21*, 3-52 San Diego, USA: Academic Press

Saaty, Thomas L. (1977) A scaling method for priorities in hierarchical structures, *Journal of Mathematical Psychology, 15*, 234-281

Sanders, Mark S. & McCormick, Ernest J. (1987) *Human factors in engineering and design, 6th edition*, New York, USA: McGraw-Hill

Sargent, Tiffany, A, Kay, Michael G. & Sargent, Robert G. (1997) A methodology for optimally designing console panels for use by a single operator, *Human Factors, 39(3)*, 389-409

Shinar, David & Acton, Marie B. (1978) Control display relationships on the four-burner range: population stereotypes versus standards, *Human Factors, 20*, 13-17

Simon, J. Richard (1990) The effects of an irrelevant directional cue on human information processing, in: Proctor, Robert W. & Reeve, T. Gilmour (Eds.), *Stimulus-Response Compatibility*, Amsterdam, NL: Elsevier Science, 31-86

Smets, Gerda (1986) *Vormleer de paradox van de vorm* (in Dutch), Amsterdam, NL: Bert Bakker

Smith, Sidney L. (1981) Exploring compatibility with words and pictures, *Human Factors, 23(3),* 305-315

Taylor, M. M. (1988) Layered protocols for computer-human dialogue. I: Principles, *International Journal of Man-Machine Studies 28,* 175-218

Tullis, Thomas S. (1983) The formatting of alphanumeric displays: A review and analysis, *Human Factors, 25(6),* 657-682

Tullis, Thomas S. (1985) A computer-based tool for evaluating alphanumeric displays, in: Shackel, B. (Ed.), *Human-Computer interaction – Interact '84,* Amsterdam, NL: Elsevier Science Publishers B.V., 719-723

Tullis, Thomas S. (1986) Optimizing the usability of computer-generated displays in: Harrison, M. D. & Monk, A. F. (Eds.), *People and computers: Designing for usability; proceedings of the second conference of the British computer society* Cambridge, UK: Cambridge University Press

Tullis, Thomas S. (1988) Screen Design, in: Helander, M. (Ed.), *Handbook of Human Computer Interaction,* Amsterdam, NL: Elsevier Science Publishers B.V. (North Holland), 377-411

Walsh, David A. & Hershey, Douglas A. (1993) Mental models and the maintenance of complex problems solving skills in old age, in: Cerell, John, Rybash, John, Hoyer, William, & Commens, Michael L. (Eds.), *Adult information processing: Limits on loss,* San Diego, USA: Academic Press, 553-584

Weeks, Daniel J., Proctor, Robert W. & Beyak, Brad (1995) Stimulus-response compatibility for vertically oriented stimuli and horizontally oriented responses: Evidence for spatial coding, *The Quarterly Journal of Experimental Psychology, 48A(2),* 367-383

Wickens, Christopher D. & Andre, Anthony D. (1990) Proximity compatibility and information display: effects of color, space, and objectness on information integration, *Human Factors, 32(1),* 61-77

Wickens, Christopher D. & Carswell C. Melody (1995) The proximity compatibility principle: its psychological foundation and relevance to display design, *Human Factors, 37(2),* 473-494

Wickens, Christopher D. (1992) *Engineering Psychology and human performance, 2nd edition,* New York, USA: Harper Collins

Wickens, Christopher D., Sandry, Diane L. & Vidulich, Michael (1983) Compatibility and resource competition between modalities of input, central processing, and output, *Human Factors, 25(2),* 227-248

Wickens, Christopher D., Vidulich, Michael & Sandry-Garza, Diane (1984) Principles of S-C-R compatibility with spatial and verbal tasks: The role of display-control location and voice-interactive display-control interfacing, *Human Factors, 26(5),* 533-543