From experiential learning to routine execution - novices and experts

1. What is a mental model?

Payne (2003) defines users metal models of devices as one of the most tantalizing ideas in the psychology of human-computer interaction and one of the most widely discussed theoretical structures in the field. As the field of study is relatively new, the definitions of mental models are so far simple statements, indistinguishable from the concept of knowledge and usually split into different scientific domains among others cognitive science, psychology and system dynamics.

The terminology describing mental models in each of these fields varies widely. In the field of systems dynamics the terms used are cognitive map, schema, causal policy map, cognitive model, mental map, mental policy, implicit model, theories-in-use, policy theory. In cognitive science for describing mental models the terms used are mental picture, mental representation, folk theory, naive theory, naive problem representation, intuitive theory, implicit theory, knowledge map, idealized cognitive model, conceptual net, internal model, cognitive structure and knowledge structure.

The first conceptualization and definition of mental models dates back to the post war era where mental models were first described by Kenneth Craik (1943) in The Nature of Exploration (Craik, 1943), where he stated that the mind constructs “small-scale models” of reality that it uses to reason, to anticipate events and to underlie explanation. Forester continued (1971) that the mental image of the world around us that we carry in our heads is a model (Payne, 2003).

Resources:
The book ‘HCI models, theories and frameworks’ by John M. Carroll has a chapter about mental models. According to the author the concept mental model is so often used in HCI and cognitive science that the definitions vary hugely. The Author describes various perspectives and definitions of mental models. I think it is a useful starting point for this section. The book is available on the dropbox and in the library (but I have the library one this week). Another useful resource is Norman, D. A. Some observations on mental models. in: D. Gentner & A.L. Stevens (Eds). Mental models.

2. Where do mental models come from: learning from instructions

Cognitive psychology attempts to rationalise human behaviour in relation to the individual’s goals and prior knowledge (Payne, 2003), and also in the context of the ‘bounded rationality’ of the individual - in which the general limits and constraints of the human processing system (i.e. information retrieval and attention) are considered (Simon, 1955). In this context, cognitive psychology offers general theories of ‘human-information processing architecture’ (Payne, 2003), from which a structure of the human mind can be realised and applied within a more general framework. The idea of mental models emerged as an additional tool to the rationalisation of human behaviour, by offering a new perspective which places greater emphasis on the content of the human mind rather than the structure. Gent and Stevens (1983) believe that the content of people’s memories, such as their knowledge and beliefs offer greater insight into people’s behaviour than the traditional mental mechanisms.

There are several theoretical perspectives that underpin the concept of a user’s mental model. Payne (2003) offers a summary of some of them:

- It is important to try to systematically investigate what people believe to be true about particular domains (such as interactive devices).
- Mental models of machines can provide a problem space that allows more elaborate encoding of remembered methods, and in which problem solvers can search for new methods to achieve tasks.
- Mental models might in some sense share the structure of the world that they represent, i.e. a diagrammatic picture.
- Models can be derived from perception, language or imagination.
Users of a system create their own mental analogies of how that system works; however, studies have shown that user’s analogies can be fragmented, whereby one explanation for a part of the system buries no relation to the rest of the system (Payne, 1991, Norman, 1983).

Mental models can originate from the experience of the user, user assumptions created by observing or interacting with a device and making sense of feedback, or by following instruction manuals — which assist learners by helping them focus on specific information about a device or system. Much work has been conducted on examining the efficacy of instruction manuals of the following two types:

- Instructions by rote — user follows a sequential procedure for completing specific tasks.
- Instructions by model - user may be given background knowledge on how the system works, often with a diagrammatic picture.

In a study of mental models focusing on problem-space elaboration of rote methods, Halasz and Moran (1983) utilised a ‘Reverse Polish Notation’ calculator which was issued to two groups of users, both of which were issued one or the other of the two the styles of instructions listed above. It was discovered that the type of instruction made no difference to participants’ ability to solve routine arithmetic tasks, however, for more creative problems the ‘instructions by model’ group were substantially better. As these users had background knowledge of the way the calculator worked they were able to navigate the problem space through ‘mental simulation’, where the ‘instruction by rote’ group were not.

In a similar study conducted by Kieras & Bovair (1984) two groups were given the same two types of instructions for a light switch system. The model instructions were essentially a circuit diagram which showed the connections between light bulbs, switches and the power source. They discovered that participants instructed with the model were more efficient because the circuit diagram explained the contingencies in the rote-action sequences.

This study was taken further by Bibby and Payne (1993, 1996), this time utilising a laser system. Similarly to the previous study, participants were given either ‘instructions by model’ (a table showing the conditions under which each indicator light be illuminated), or ‘instructions by rote’. They were also asked to complete two different types of tasks, a straightforward switch task and a fault task (in which the participant had to solve a problem). They discovered that the fault task was easier for the users given the model instructions, but the switch task was easier for the users given the procedures.

3. Where do mental models come from: learning from exploration

Different conditions bring people to learn directly from experience, rather than through formal learning:

- people choose to start using a device without reading the manual (e.g. assembling, connecting and operating a new BlueRay device at home);
- it is acceptable for the type of task to remain novices in an interaction, hence no pressure is placed on learning a device in advance (e.g. experienced employee onboarding in a new working may not be required to undertake a mailing system training);
- people are not expected to have to learn a device from a manual due to the ecology of the environment where they are performing the task (e.g. train automatic ticket machine or vending machines).

In all these cases, the users of the devices have to learn through exploration, examining the the device for cues on its operating rules.

In other cases people go through an instruction based training, however their learning continues when they start to use the device.

Learning from pure exploration

Cognitive scientists attempt to understand what knowledge of a device is built through exploration and how.

Shrager (1985) observed people learning to operate a programmable toy vehicle (BigTrak) and divided the learning phases in two:

- an initial orientation: users mainly observe the device and try to operate it, until the device key purpose is discovered (or achieved), i.e. the capability to move;
- a successive systematic investigation: users build through hypothesis, experiment and evaluation, the syntactic, semantic and model knowledge, i.e. the knowledge of how the device works and how to operate it.

Most relevant observations come from the second phase: users formulate hypothesis with different degrees of refinement and scope, sometimes leaving uncertain the expected outcome of the experiment. Their hypothesis reveal a novice scientist attitude, embedding biases and confounding, reflecting the inclination to proof an hypothesis rather than its contrary. Contradictory hypothesis were held simultaneously, users were assessing results falling into the trap of confirmation biases and memory failed to recall BigTrak unforecasted behaviours. Overall users were able to master most of the controls, but the more sophisticated ones remained uncovered.

Another interesting finding comes from Trudel and Payne (1995), replicated by Cox and Young (2011): users in an exploratory learning experiment of a device spent most of their time building device-oriented knowledge (how-the-device-works) rather than task-oriented knowledge (how-to-do-the-task). As observed in Cox and Young, knowledge on how-to is built only if the users assign to themselves realistic task goals. But this is not observed to spontaneously happen in an exploratory learning experiment.

Comparing the experiments result with real environment user behaviour, it emerges that in the latter it is exactly the opposite: users explore further devices when they have a clear task problem to solve (Rieman, 1996). They would give priority to the ‘try’ strategy, following most likely by the ‘read the documentation’, and eventually the ‘ask someone’, often combining one or more of these strategies.

Learning from exploration after learning from instruction
Formal learning through instruction is a modality of learning. Even in Kieras and Bovair (1984) and Bibby and Payne (1993, 1996) experiments, learning from instructions was followed by the use of the system in a ‘real’ context. Bibby and Payne experiment in 1996 highlighted that use of the device allowed the users to adapt their strategies, however knowledge compilation (building new declarative knowledge specific to the task) can be more or less successful according to the starting learning material (i.e. procedures versus table). The users would learn by experience and apply heuristics (Neche 1987) to build this new declarative knowledge, becoming more efficient and performing.

With time, expert users would show that the starting learning material is not anymore influencing task performances: knowledge gaps were filled through the experience and knowledge compilation. However when presented with a new task, the initial instruction advantage or disadvantage according to the combination task type/instruction type emerges again.

4. Mental models and conceptual models

// Outline

1. Mental model vs. Conceptual model
   a. Mental model - user view of how system works
   b. Conceptual model - designer's vision
2. How mental models vary based on instruction (studies from lecture) - I think this is content of Section 2 here (Jiri)
3. Progression through learning and it's effects on the mental model - ditto Section 3? (Jiri)

Jiri & Philip: You want to structure some other way? change anything?

Terri & Philip: Isn't content of 2. and 3. supposed to be in sections 2. and (surprisingly) 3. here on this section of Wiki?

//

Mental Models and Conceptual Models

How people approach using systems is greatly influences by their mental model. They use these models to help plan and execute their interactions with objects and systems, and to understand the effects of their actions. In an interaction the users mental model encounters the system representation, their actions and the feedback they receive can mismatch the conceptual model of the system. These models evolve over time, a representation of this is demonstrated below.

Mental models are incomplete and people are limited in applying these models. Thus to reduce the dissonance between the conceptual model and mental model, the system should be learnable, functional and usable (Norman, 1983).

5. Bridging the gap between the user and designer model

Many users form mental models that are often incorrect, incomplete or more simple than reality. If a user does not feel that the system reflects their understanding of it they are likely to get confused, irritated or they may be put off from using the system altogether. The designer on the
Discovering and Using Models in Design

Techniques to gain insight into the user's mental model can be surveys, interviews and focus groups, contextual inquiry or participatory design. These can help inform design and then after an iteration is made, user testing can help to validate them. Fundamental differences in mental models can also be recorded into varying personas.

Nielsen suggests conducting think-aloud user testing, as this can gain a deeper insight into users' mental models as they move through a website or application, and they can verbalise their predictions. He then suggests that the designer:

- Makes the system conform, to the mental model that users already have assuming that most people have a similar mis-match. For example, if users repeatedly look for something in the wrong place then it may be beneficial to move something to where users expect to find it. In the case of navigation, Card Sorting is a good tool for relabelling taxonomy.

- Alternatively, designers can try to improve the users' mental models so that they more accurately conform to the system in question. This can be done by perhaps explaining the concept better or making labelling clearer, in order to present a more intuitive design. Or by giving explicit training e.g. video tutorial.

Potential Issues

There can be some issues with the conceptual models that designers present. Too much information may interfere with a user's model. Goodwin gives an example where it may be beneficial to just provide the information needed to allow a user's mental model to work:

"A clothing store database might note three separate product numbers for a navy blue sweater, a black sweater, and a red sweater of the same style. This structure is useful from the standpoint of inventory management, but it shouldn't be surfaced to shoppers, who will be frustrated if your site has three separate listings for what they think of as a single sweater available in three colours."

Norman states that many things can be designed into systems to help users understand how they work, they include, affordances, constraints, good visibility (e.g. buttons, labels), good feedback and appropriate mapping. Affordances and constraints are covered below.

However, it is not possible to create a conceptual model that acts as a silver bullet. Users' models are built on extensive years of experience (and also imagination), and although many experiences are shared by a lot of people, they are still interpreted on a personal level so it still may not be possible for all users to understand intuitively the first time they use something. It's also important to remember that users who have used a system for a long time may be accustomed to how it behaves but that does not mean it is a natural structure, or easy to learn.

6. Knowledge in the world: affordances & constraints

Outline:
- Introduction
- The origin of the theory of affordances
- Evolution of the theory of affordances
- Constraints
- Implications for design (what happens with novice and expert users?)
- Conclusion

Introduction

In “The Psychology of Everyday Things”, Norman (1988) argues that users can deal with a novel object either with their past and transferable knowledge, which is stored in the head or with the knowledge, which is presented in the world. Therefore, there are three major aspects that need to be considered in the design process: conceptual models, constraints, and affordances. All these concepts contribute in a different way to the design. In this chapter, the focus is placed on the affordances and the constraints.
The origin of the theory of Affordances

The term «affordance» was coined by Gibson (1979) to highlight the correlation between humans and the environment. Without an explicit definition provided, “the affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill”. They are a fundamental element of the ecological approach to visual perception, which tries to explain how human perceive the environment, with main reference to Gestalt theory (see Chapter 2). In brief, the main argument of the ecological approach is that humans perceive the environment directly, without any internal processes occuring. This argument contradicts the information-processing model (see Chapter 1). Moreover, the affordances exist independently from a human’s action, needs or goals, e.g. a wall made of concrete is solid despite human’s previous experience. However, the perception of the affordances depends on human’s experience, value and culture, e.g. someone knows that punching a concrete wall will cause pain to his fingers. An affordance may or may not be perceived, which means that has a binary nature.

Norman (1988) introduced the term in the HCI community with the aim to provoke designers’ thinking of how to design a product or an interface so as to be easily perceived. He acknowledges that humans perceive the environment with the information provided in the world but this perception also depends on their processing mechanisms, past knowledge and experience. In contrast to the independent nature of Gibson’s affordance, Norman believes that if the affordance is not perceived, it does not exist. Thus, the designer has to make the affordances visible. For that reason, Norman (1999) clarified that he talks about two different kind of affordances:

- The real affordances which are about the physical characteristics that allow a product to function, e.g. the mouse affords pointing and clicking. This type is similar to Gibson’s definition of affordance.
- The perceived affordances which provide cues about the possible functions of a product. The perceived affordances determine the usability of the product (McGrenere and Ho, 2000).

The evolution of the theory of Affordances

Gaver (1991) highlights the value of affordances when it comes to design technological artefacts people can interact with. He adopts Gibson’s notion about the invariant nature of affordances and he introduces the perceptible and the hidden affordances. If there is correct information provided about a possible action, then the affordance is perceptible, otherwise without any information the affordance is hidden. Furthermore, if the information provided is wrong, the affordance is false. Therefore, the goal of the design is to make affordances perceptible.

Hartson (2003) revises the aforementioned theories and proposes a framework to guide design decisions in terms of affordances. For this aim, he renames some of the existing types of affordances and introduces two other types: the sensory and the functional affordance.

According to his framework, the designer must start by considering the artefact’s functionality and purpose or in other words, the functional affordance. The functional affordance allows the user to accomplish a task and achieve his goal. In the next phase, the design of the cognitive affordance will ensure that the artefact is easily perceived. The cognitive affordance guides the design in terms of usability, similarly to Norman’s perceived affordance. A text label informing the function of the artefact is an example of cognitive affordance.

Likewise Norman’s real affordance, the physical affordance indicates the physical characteristics of an artefact which users can manipulate. Physical affordances can maximise the performance of experienced users and help novice users to get familiar with the artefact. Size, texture and location are factors that determine the artefact’s physical affordance.

Users must be able to sense the cognitive and physical affordances in order to understand and act upon them respectively. For that reason, Hartson (2003) introduces the concept of the sensory affordance which supports the crucial role of the cognitive and physical affordances. A sensory affordance is a design feature in visual, auditory or haptic form.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Types of affordances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibson</td>
<td>Affordance</td>
</tr>
<tr>
<td>Norman</td>
<td>Real</td>
</tr>
<tr>
<td></td>
<td>Perceived</td>
</tr>
<tr>
<td>Gaver</td>
<td>Perceptible</td>
</tr>
<tr>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Hartson</td>
<td>Physical</td>
</tr>
<tr>
<td></td>
<td>Cognitive</td>
</tr>
<tr>
<td></td>
<td>Sensory</td>
</tr>
<tr>
<td></td>
<td>Functional</td>
</tr>
</tbody>
</table>

Table: Different types of affordances

![Diagram of affordance types](image)
7. Designing for novices

Rieman (1996) states that learning a new system should be intrinsically motivating; features should be revealed incrementally; and the system should be at least minimally useful with no formal training. According to Carroll (1985), in the absence of manuals or instructions, most people prefer to use an explorative behaviour when learning new systems. However, there are reasons to believe that unsupported instructionless learning, without having access to manuals or other training aids, can be problematic. In this sense, "walk-up-and-use" systems that allow for a trial and error approach should be balanced with making manuals and the possibility of asking for help. Rieman also notes that "on-line help is widely disliked", but it is sometimes used in conjunction with trial and error. As a consequence of this, most users have evolved skills to identify what type of "help" is required in each moment and acquire that assistance.

The minimalist approach to instructional design is based on decreasing negative impact of instructional material and allowing for self-directed learning through meaningful tasks, "The goals is to let the learner get more out of the training experience by providing less overt training structure." Carroll (1985).

Kearsley offers 4 principles for the application of a minimalist approach to instructional design:

1. Allow learners to start immediately on meaningful tasks.
2. Minimize the amount of reading and other passive forms of training by allowing users to fill in the gaps themselves
3. Include error recognition and recovery activities in the instruction
4. Make all learning activities self-contained and independent of sequence.

Further practical recommendations for a minimalist instructional design are given by Robert E. Horn, which are summarised in the following nine principles:

1. Use real tasks for the training exercises and let users select their own tasks. It enables people to use their prerequisite competence and engages a "powerful source of motivation."
2. Get the learner started on real tasks fast by eliminating almost all front-end orientational material. Extensive preambles can "obstruct meaningful activity."
3. Guide learners' reasoning, exploring and improvising with questions and other hints. This includes incomplete training materials, so that learners have to explore. He also suggests presenting summaries in place of complete texts.
4. Design the materials so that they can be read in any order in so far as possible. This principle permits learners to "support their own goal-directed activities"
5. Help learners to coordinate training materials and software by providing landmarks for normal or error situations, e.g. illustrations which show what the screen should look like if everything is OK.
6. Focus early attention in the training materials on enabling the learner to recognize and recover from errors. Learners make many kinds of errors in learning computer systems. "Training materials must therefore explicitly support the recognition of and recovery from error both to make the materials robust with respect to user error and to train error recovery skills."
7. Engage the learner's prior knowledge in introducing novel concepts. Use familiar office tasks, language and metaphors. Highlight differences in operation of the system from what might be expected from the learner's background.
8. Consider using the learning situation, as opposed to practical on-the-job examples, for learning examples, exercises and explorations. Help the learner understand the "fine detail of the actual situations in order to create practical solutions."
9. Aim for optimizing learning designs by repeated testing and avoiding the temptation to systematize approaches into checklists. Carroll says, "There is no deductive theory of minimalist instruction; that is, given a set of minimalist principles, we cannot just crank out a training manual. Design never works this way."

To be added:

- Explorability, device knowledge and task knowledge - (Cox, A. L., Young, R. M., 2000)
- Learning and Retention - Ritter et al (2011)

8. Designing for experts

Outline (details/prose to be added):
Model of human memory: Declarative vs procedural
Stage 3: procedural (fine tuning to the rules, small adjustments to make things as fast as possible)
Novices: need info in the real world
Experts: hold knowledge in their head

- efficient
- allows better aesthetics
- requires lots of learning
- may not be retrievable (so must design for re-learning)

What is expert? What do people learn?

- faster
- more accurate
- new strategies for a more efficient journey
• mental model - how the device works (hypothetical construct stored in LTM)
• quickly reading current state of system/task
• looking ahead/taking a longer-term view
• pre-emptively choosing appropriate routes and strategies
• has holistic perception of activity (e.g. "shorten this doc" or "clean up hard drive", rather than "cut text" or "move file")
• understands "terrain" and rapid judgement of next steps - so needs to immediately see implications of actions
• observation of others
• personal reflection
• has mental "framework" making it easier to add detail in context (so it's not confusing, as it may be for a novice)

Considerations for design:
• short cuts
• hot keys
• ability to use both hands
• design to facilitate creation and refinement of accurate mental model (close to designer's model)
• surface future implications of their actions (take a longer-term view)
• design so least effort option is most effective (shortest time, fewest errors)
• "Chaining": amalgamating single movements together
• personalisation (e.g. macros) of interaction
• immediate, detailed feedback
• multiple undos - ability to review ones own activity and reflect

Examples: Photoshop (Good); Word Tracking Changes (not so good)

Need to add references

9. Application to systems

We look at effective methods for learning through empirical studies.

Kieras and Bovair - Learning from instructions

Kieras and Bovair(1984) had an experiment where there were 2 groups that learnt how to use a control panel using a set of procedures.

The **Device Model Group** had additional information with the panel described as a phaser bank, shown a digram of the internal structure and trained in procedures.

the **Rote Group** was just trained in procedures.

The effects from the Device Model Group was that they

• remembered procedures more often
• used efficient procedure more often
• executed procedures faster

This had shown that knowledge of an internal workings of a system helped users infer how to operate it, hence has aligned the mental model of the user to the designer's concept.

It had also shown that traditional instruction manuals are not efficient to train people.

Bibby and Payne - Laser Fire

Bibby and Payne's (1993, 1996) experiment was to train users to make a laser fire.

There were two groups with different sets of instructions

Group 1 - Table of conditions to make the laser fire
Group 2 - Set of Procedures to make each laser fire

Both groups were asked to perform tasks, key the final switch and name the broken component.

Due to the nature of the group's instructions it was found that

Group 1 was faster at naming the broken component.

Group 2 was faster at keying the final switch.
It was noticed that through test phases <80 that participants were faster at the task for which their instructions had given them the advantage. After continuous practice 80+ test phases the advantage had disappeared. However when going to new tasks the advantage had reappeared, as their orginal knowledge persisted in the same form as they had learnt it, the repetitive test phasaes had not changed the facts that they had learnt.

**Implications of Bibby and Payne**

Bibby and Payne states that different forms of instructions tend to support different types of tasks. The form of instructions can have a long lasting effect.

**Carroll - Text Editing System**

Carroll 1990 had two groups to learn using a text editing system.

Group 1 - traditional system manual

Group 2 - set of cards related to realistic tasks

The nature of instruction for Group 2 guided users to explore the system.

Group 2 were

- quicker to learn
- quicker at completing tasks
- more successful at completing tasks

Carroll found that people find it difficult to learn from manuals, different styles of training produce different results.

The importance of realistic tasks when learning to use a system.

**10. References**


Kearsley, Greg Minimalism (J. Carroll)


11. Exam questions

Previous exam questions

Other exam questions that we’ve thought of