THE OVER-ASSISTING INTERFACE: THE ROLE OF SUPPORT IN KNOWLEDGE ACQUISITION

Christof van Nimwegen   Herre van Oostendorp   Hermina (Tabachneck-) Schijf
Institute of Information and Computing Sciences
Center for Content and Knowledge Engineering
Padualaan 14, 3584 TC Utrecht, The Netherlands
{christof, herre, h.schijf}@cs.uu.nl

ABSTRACT
The premise of this research is that support in an interface can make interactions easier, but does not necessarily support understanding and more importantly knowledge acquisition. Interfaces often display certain relevant information, making recall of such knowledge unnecessary, relieving working memory. An example is "greying-out" (menu-) items, indicating that an action is not possible. An experiment was conducted in which subjects solved a series of puzzles. We hypothesized that providing greyed-out items (externalization) yields better performance during initial learning. An interface without greying-out (internalization) is expected to yield better performance and knowledge in later phases. Subjects solved a puzzle in two conditions: with greyed-out items and without. Externalization had little influence on performance. All subjects learned how to solve the puzzle. With acquired knowledge it was different. Procedural knowledge afterwards was equal, but declarative knowledge, concerning the central puzzle-rule was worse for those who had greyed-out items. Also, months later the same internalization-subjects had faster problem recognition of the task and better performance on a similar task.

Keywords
Interface representation, knowledge acquisition, planning

1. INTRODUCTION
The notion that learning is more effective when people experiment themselves is far from new, and exploratory learning has been a subject of research in many domains.

More than a decade ago Carroll [1] propagated minimalism in design and instructions already. Software has seen a tremendous development and being more advanced it demands fast learning from users. Accents have been put on usability (effectiveness, efficiency & satisfaction) [2]; interfaces of today are by no means comparable with command-line interfaces of the old days. Interfaces often are complex, and applications can have hundreds of functions. A notion that came with usability is the importance of "minimizing user memory load", and a common recommendation is that users should interact on basis of recognition rather than recall [2]. In practice this means that objects, actions and options should be visible, and a user should not have to remember too much information. In interactions with software, we are nowadays quite used to see a considerable portion of the interface greyed-out (fig.1). It limits choices and "takes a user by the hand" [5]. In "wizards" with a mandatory sequence, having buttons greyed-out can indicate that one has not filled out a mandatory field. In other applications, greyed-out menu-items can indicate that an item is out of context. The interface is context-sensitive, offering only possible actions.

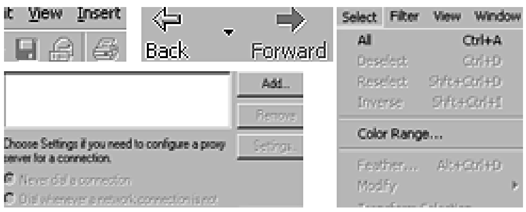

Figure 1 – Greyed out interface items

"Greying-out" is supposed to support the user, telling him/her that in the current situation applying a specific function is not allowed, and remembering this is not necessary. However, the user is left in the dark, as to the reason for the greying out. Some examples are easy to grasp. If nothing is copied, you cannot paste it, so "paste" is greyed-out, but reasons for greying-out are not always so self-evident. By externalizing certain information, working memory (WM) is relieved [6]. Several other studies showed that externalizing information can be useful for cognitive tasks; the more is externalized, the easier it is to perform a task. This was generally accepted, but as Zhang and Wang [6] pointed out, there hadn't been research on *how* external representation influences WM. They distinguished between IWM and EWM (internal vs. external representations). It showed that having *all* information in EWM always aided problem solving, but with information distributed across IWM and EWM it could go both ways, enhancing *or* hindering performance. Of influence were retrieval strategy and the way information was encoded. The results however were not clear-cut, and hard to generalize and apply to interfaces. To contribute to a more complete theory, research is needed on effects of distributing information. One could argue that with most information externalized (and little internalized) users are not triggered to look for underlying rules or reasons. In contrast, having most knowledge internalized perhaps stimulates planning and this might be important when transfer is needed, or when speed is important. O'Hara and Payne [3] and Trudel and Payne [4] provide support for this view, stating that too strong a reliance on external information leads to negative effects regarding planning and transfer of skills. They distinguished between plan-based and display-based problem solving. In plan-based problem solving one uses detailed problem strategies from long-term memory. Display-based makes little use of learned knowledge but relies on interface information. Plan-based activity leads to shorter solution routes (steps are planned) while display-based strategy involves more steps because of more searching. Our research is part of a broader research programme aiming to contribute to theories that explain and predict which type of screen representation leads to better task performance, knowledge acquisition, performance after delay, and transfer.

2. HYPOTHESES

We investigated our questions using 3 hypotheses:

1. Externalization will initially yield better task performance than internalization. Internal knowledge-elements are not yet acquired and externalization is helpful.

2. Internalization yields better task performance in later phases. Reliance on better-internalized knowledge leads to better performance in later phases.

3. Internalization yields better knowledge. Not relying on externalized information provokes planning steps oneself, consequently knowledge of the task will be better.

3. METHOD

3.1 Material

The problem-solving task we used is called "Balls & Boxes" (B&B). It is an informationally equivalent isomorph of "Missionaries and Cannibals" (M&C). 5 missionaries and 5 cannibals stand on a riverbank, and all have to reach the other bank by boat. The boat only fits 3 people, the minimum to sail is 1. If cannibals ever outnumber missionaries at any place, the latter will be eaten (fig.5, with 5 M&C instead of 3). The B&B problem (fig.3) uses the same problem space, but is more abstract (in M&C rules are quite intuitive; cannibals eat people, boats cross rivers). With boxes, coloured balls and a dish instead, we avoided too easily learned rules. In B&B the rules translate to the following:

1. Balls should be transported using the dish
2. You can transport at most 3 balls at a time
3. To move, the dish must contain at least 1 ball
4. The dish has to reach the boxes in turn
5. No more blue than yellow balls in the dish
6. No more blue than yellow balls left in the boxes

Below is the formal problem space of the puzzle. Solving it from start to end involves a sequence of 11 steps, some variations are possible, but there are also several moves that take you nowhere. These "dead-end-states" (circled in fig.2) leave you with only one choice; go back and find the right track.

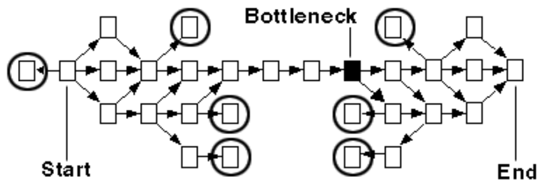

Figure 2 - The balls & boxes puzzle's problem space

It is a puzzle with a "trick" to it, after an aha-moment the solution is easy. At one point a contra-intuitive move has to be made, this bottleneck is passed after 7 moves (fig. 2, black square). The controls were simple: to get balls into the dish, blue or yellow up-arrows had to be clicked and to move the dish horizontally, one had to click a pink arrow (left or right). The rules could be consulted at all times by clicking on a tab. There were 2 puzzle versions: <u>With greying-out (Externalization):</u> Arrows are only coloured (clickable) when an action is legal, and greyed out (unclickable) if illegal. E.g. moving the dish empty in fig.3 is illegal because it violates rule 3. In this situation, it is externalized by greying out all the pink arrows. <u>Without greying-out (Internalization):</u> All arrows are always coloured providing no information about the legality of moves. One can push buttons at all times. If one wants to move the dish empty (illegal) and clicks a pink arrow, the dish moves, but an error notification pops up saying "not possible". Subjects had to click "ok" to undo the move.

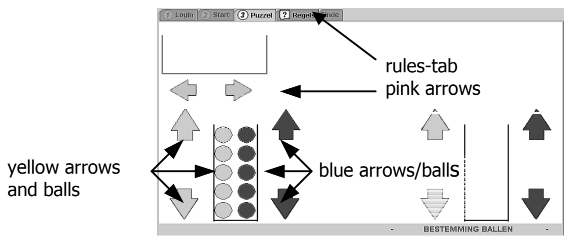

Figure 3 - The balls & boxes puzzle

3.2 Knowledge test

To estimate the acquired knowledge of subjects afterwards, we developed a knowledge test of 8 questions. The puzzle has more than 30 legal states. There were 7 procedural knowledge questions about 7 of those states (open and multiple choice) in which subjects were visually presented with puzzle situations (fig.4). They had to judge whether or not certain moves were leading to the solution, and why.

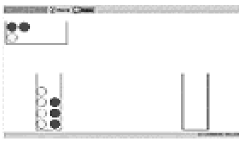

Figure 4 – Puzzle situation from knowledge test

To test declarative knowledge, there are limited possibilities. The puzzle had only 6 rules. Rules 1 to 4 were easy to grasp and remember. Rules 5 and 6 are more crucial and define the difficulty of the puzzle. They can be merged to just 1 rule: "blue balls can never be in the majority at any place except when there are only blue balls". Subjects were asked about this rule with a multiple-choice question.

3.3 Material for delayed re-testing

To see what happens over time we re-invited the same subjects after 8 months. Again they received the B&B puzzle, but no rules were consultable, since we purely wanted to check performance after a long delay. In addition, subjects solved a "real" Missionaries & Cannibals puzzle that was semantically richer (fig. 5). The solution (11 steps) is similar as in B&B. However, the fact that the playing direction was reversed and that there were 3 instead of 5 of each objects will force subjects to stop, think and apply (transfer) learned knowledge to this similar problem.

Figure 5 - The missionaries & cannibals puzzle

3.4 Procedure

Thirty subjects (15 per version) aged 19-28, experienced with PC's had to solve the B&B puzzle 9 times. The starting situations always differed, to avoid subjects relying too much on "having learned the trick" and repeat actions. For a delayed re-test 8 months later, we again invited 14 of the same subjects, 7 from each version. They had to solve the same B&B puzzle again (5 times) without further instructions. After this they had to solve the realistic M&C.

4. RESULTS

Some measures we analysed were the number of solved trials and time and steps needed to solve the puzzle. Mostly the average score of internalization-subjects was higher, but not significantly so. In session 1, in both groups, already after 4 trials on average, a ceiling was reached, and all subjects regardless of the version they worked with, were perfectly able to perform fast and efficiently (the minimum of 11 steps). We confirmed that there was no difference in rule checking; in both versions they were consulted equally often, and as expected especially in the beginning.

The scores on the procedural knowledge questions of both externalization and internalization-subjects were high, 5.8 and 6.2 respectively (maximum 7). This small difference was not significant; both groups answered these questions equally well. For declarative knowledge however (fig.6), we found that subjects in the internalization condition scored far better. All the internalization-subjects correctly answered this question, whereas of the externalization-subjects only 60% answered it correctly. This was significant (Phi =-.50, $p < 0.01$).

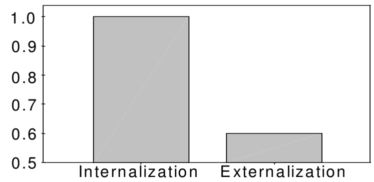

Figure 6 – Proportion correct declarative knowledge p.version

In the delayed-retest more encouraging results were found (fig.7). Internalization-subjects needed significantly less time to solve the same B&B puzzle again for the first time (mean **432** sec, Sd 314 and mean **778** sec, Sd 397 resp; T (df=12)=-1.81, $p < .05$). After the first success, all subjects solved the puzzle equally well, just as 8 months ago.

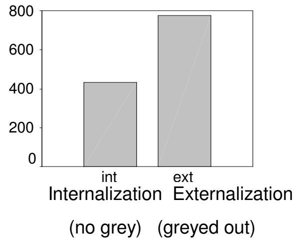

Figure 7 – Time (sec) until success B&B per version

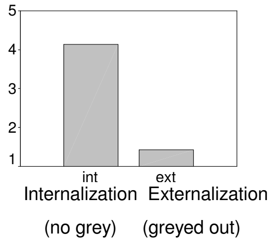

Figure 8 – Number of M&C solved per version

During the second (transfer) task internalization-subjects managed to solve the M&C (fig.8) puzzle significantly more often than externalization-subjects (mean **4.14**, Sd 2.54 and mean **1.43**, Sd 1.81; T (df=12)= 2.30, $p < .05$).

5. DISCUSSION & CONCLUSION

Our first hypothesis stating that having greyed-out items (externalization) yields better performance initially, was hardly supported (unlike in other research). There were minor differences, but not significantly so. The second hypothesis stating that *not* having greyed-out items (internalization) would yield better performance in later stages was not supported in session 1, but in session 2 after a long delay, some interesting differences emerged which supported the hypotheses, see further in the conclusion. We first want to focus on the results concerning the third hypothesis. We expected subjects that had no greyed-out items (internalization) to have better knowledge after the experiment, because they had to build a stronger, more elaborated plan. For procedural knowledge, we were surprised to find no differences. Being confronted with screenshots of puzzle states and asked to judge them, both groups did equally well. Was it because the questions were asked with a visual example, and was this sufficient to trigger the needed procedural knowledge, or was the puzzle just so easy? There seems to be more to it. Looking at declarative knowledge it came as a surprise that even though the mentioned performance and procedural knowledge were almost equal in both groups, the externalization-subjects that had support were much worse at correctly recalling the actual rule the puzzle was about. Just 60% of the subjects that used greyed-out items knew it correctly, opposed to 100% in the internalization version. They "did it right", but not based on correct declarative knowledge. Perhaps (procedural) knowledge they possessed here was comparable with the way people sometimes remember safe combinations or phone numbers; just when they do it. We interpret these findings as indicators of better knowledge instigated by internalization. More confirmation came after the delayed re-test. In context of hypothesis 2, after 8 months, the same subjects had to solve B&B again. The first success took subjects that had worked with greyed-out items (externalization) twice as long as the internalization subjects. This indicates that after this time-lapse there was a difference in procedural knowledge as well. After this first success, performance in both groups was again at top-level soon, so we assume that procedural knowledge by then was equal again. Next, subjects solved a more realistic version of the same problem (M&C). Subjects that had no greyed-out items 8 months before had an advantage. They solved the puzzle almost 3 times more often in the same time. We interpret this in a transfer-context; when facing "new" situations internalization-subjects understand better that it was actually the same problem.

Some remarks remain. Perhaps in session 1, internalization-subjects might have done better in the B&B-puzzle if the application hadn't forced unintentional delays on them. They had to click away messages to undo "wrong" moves, and may have lost time in recovering, described by O'Hara and Payne [3] as the effect of "lockout time and error recovery cost".

The results were encouraging and we will continue the research with more difficult tasks (see Future Research) and more realistic planning-related tasks, e.g. educational, spreadsheet or drawing applications that are less repetitive and more complex. With interfaces that facilitate learning a task (such as "learning Photoshop"), making mistakes has no severe consequences. In other situations however, it is critical for an operator to know exactly what to do. Imagine when a system is down or fails and processes (e.g. shutting down a part of a factory) have to be performed without the system's support. Here solid underlying task knowledge is necessary. For situations where knowledge acquisition itself is the aim (language CD-ROMs, school situations), future research can results in specific guidelines for specific tasks.

6. FUTURE RESEARCH

The fact that we did not find convincing performance differences in the first session led us to assume that the material itself was fairly easy. Furthermore the repetitive nature enabled fast automated solving (after 3-4 trials all subjects performed very well). To give the effect of our manipulations more chance we are currently preparing another experiment with B&B in the same conditions again, but now more difficult without consultable rules available. Next, we plan to have subjects solve this puzzle with incrementing difficulty by increasing the amount of balls gradually. Finally, since we did find interesting results regarding a far transfer task (after time though) we will also investigate a more near transfer with another abstract version instead of the very realistic M&C. We designed another puzzle called "Squares and Triangles" (fig. 9) with slightly different controls.

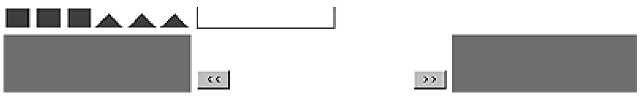

Figure 9 – Squares and Triangles puzzle

Regarding acquired knowledge we found that after the first session the distinguishing factor seemed to be declarative knowledge. We will design a more elaborate questionnaire with items more focused on distinguishing between procedural and declarative knowledge.

7. REFERENCES

1. Carroll, J.M. (1990). The Nurnberg Funnel: Designing Minimalist Instruction for Practical Computer Skill Cambridge, Mass: MIT Press.
2. Nielsen, J. (1994). Usability Engineering. Morgan Kaufmann, San Francisco.
3. O'Hara, K. and Payne, S.J. (1999). "Planning and the user interface: the effects of lockout time and error recovery cost", Int. Journal of Human-Computer Studies, 50, 41-49.
4. Trudel, C.I. and Payne, S.J. (1996). "Self-monitoring during exploration of an interactive device", Int. Journal of Human-Computer Studies, 45, 723-747.
5. Van Oostendorp, H. and De Mul, S. (1999). "Learning by exploration: Thinking aloud while exploring an information system", Instructional Science, 27, 269-284.
6. Zhang, J. and Wang, H. (1998). "An Exploration of the Relations Between External Representations and Working Memory" (in review).