Performance in a planning task: The (ir)relevance of interface style and users' cognitive style

Christof van Nimwegen
Center for Content and Knowledge Engineering
Utrecht University
Padualaan 14, 3584CH Utrecht,
The Netherlands
christof@cs.uu.nl

Daniel Burgos
Educational Technology Expertise Centre
Open University of the Netherlands
Valkenburgerweg 177, 6401 DL, Heerlen
The Netherlands
daniel.burgos@ou.nl

ABSTRACT
This research investigates whether interface style (internalization or externalization) influences performance in a problem solving task. Assistance from a user interface during problem solving is often thought to make interactions easier. Interfaces often display relevant information, making recall unnecessary and relieving working memory, called *externalization* (e.g. feedback aids such as "graying out" menu-items). By externalizing information, *display-based behavior* is provoked, which however does not necessarily instigate planning, understanding and knowledge acquisition. When certain task-information is less directly available, it needs to be *internalized,* stored in memory*,* provoking *plan-based behavior,* which may lead to better performance and knowledge. To provoke these behaviors, we manipulated the interface of a conference planning application. We also included the users' cognitive style, in this case "need for cognition" (NFC), the tendency to engage in and enjoy effortful cognitive tasks. High-NFC subjects generally plan more, leading to better performance. Results show that interface style indeed influences problem-solving behavior, but NFC does not. Internalization resulted in more planful behavior, better solution routes and less reconsidered actions. If plan-based behavior is preferred, designers should be careful in giving users too much assistance.

Keywords
Human factors, Need for Cognition, Interface, planning, problem solving, plan-based, display-based, externalization.

INTRODUCTION
Nowadays, in the software development *usability* is mostly acknowledged, at least to some extent. A recurring issue in guidelines is the importance of minimizing "user memory load" [4], also known as computational offloading; relieve users' working memory so they can devote maximum cognitive resources to the task at hand [6]. A way to implement this is to make parts of the interface context-sensitive, e.g. by hiding or disabling functions that are not applicable, such as the grayed-out menu items in Microsoft Office. This is called the *externalization* of information onto the interface. In the opposite situation, when no such features are provided a user has to *internalize* the information, to store this information in his/her memory. Although research showed that supporting the user this way can make tasks easier to perform, it may also have some negative consequences for task performance and knowledge acquisition. The knowledge gained using an externalization-based interface may be volatile and difficult to transfer to other situations. This is not always a disadvantage (e.g. with a trivial task) but it can be undesirable when learning or gaining insight *itself* is exactly the aim of the task.

Research by O'Hara and Payne [5] provide support for this point of view, stating that too strong a reliance on external information leads to negative effects regarding planning and transfer of skills. They drew a distinction between plan-based and display-based problem solving. In plan-based problem solving one uses detailed problem strategies from long-term memory. Display-based on the other hand makes little use of learned knowledge but relies on interface information. Plan-based activity leads to a shorter solution route, because steps are planned, and no unnecessary steps are taken, while a display-based strategy involves more steps because of more searching.

Van Nimwegen, Van Oostendorp and Tabachneck-Schijf [7] used an abstract version of the Missionaries and Cannibals problem called Balls & Boxes (B&B). It investigates which effect internalization vs. externalization had on task performance and knowledge, and transfer of the acquired knowledge different problems in the same domain. The experiment had two sessions with 8 months between them. In the task, externalization was realized by graying out inapplicable (momentarily unavailable) buttons (to perform certain operations). In the internalization condition this support was absent. Surprisingly, not ever did subjects in the externalization condition perform better. Moreover, it showed that subjects in the internalization condition had

better *knowledge* afterwards. After 8 months, the internalization subjects still proved to have better knowledge and now also better performance. In addition, these subjects performed better at transferring the acquired rule-knowledge to a different but similar task.

In a follow-up study, Van Nimwegen *et al.* [8] performed another experiment using B&B, but this time they also tried to *externally* provoke plan-based or display-based behavior by instructing subjects before solving. Subjects received either a low-planning instruction ("solve as fast as possible") or a high-planning instruction ("solve as economically as possible"). Again, internalization subjects performed better and had better knowledge. The instruction did not have effect on externalization subjects' behavior (no difference for high vs. low planning instructions). It seemed that subjects' attention was fully captured by the interface, and they consequently forgot the instructions. However, the instruction *did* have effect on internalization subjects. With low-planning instruction, they attempted to make twice as many illegal moves compared to those given a high-planning instruction. Also the number of unnecessary moves was higher with low-planning instruction.

THE CURRENT EXPERIMENT & NEED FOR COGNITION

In the current experiment we use a more realistic task, still investigating the influence of interface style. In the previous experiment [8] we tried to induce a mental state on the subjects. Here, we decided to see whether more stable, personal qualities of people influence behavior. We focused on a trait regarding attitude towards problem solving: "Need for Cognition" (NFC). NFC is a well proven construct measuring the tendency of individuals to engage in and enjoy effortful cognitive tasks [2]. Humans have to make sense of their world, but tend to do this in different ways. High-NFC persons love to seek, reflect on and reason about information, whereas low-NFC persons only think as hard as they have to, and are inclined to rely on others. The fact that high-NFC persons recall more information than low-NFC persons could have implications for HCI. For example in research on adaptation emphasizes the non-existence of a 'one-size-fits-all'-solution and the importance of adapting the external representation to the emotional state of learners. Other literature on learning systems proposes adapting to various constant and volatile user characteristics [1]. Could NFC be used to adapt interfaces to users? Using an information-seeking task, it was shown [3] that NFC can indeed influence perceived system usability and users' responses, or as they put it: "*If, in a particular HCI context, individuals high in NFC exhibit significantly different behavior than individuals low in NFC, usability could likely be improved by providing interfaces optimized for each group.*"

In context of the research by Van Nimwegen *et al.*, this could mean that high-NFC persons display different behavior than low-NFC persons, and even interactions with interface style might exist. We formulated two hypotheses:

H1: Internalization leads to a more plan-based strategy and better performance than externalization. Having information externalized tempts users *not* to form plans and rely on the interface. The internalized condition lacks this guidance and encourages to plan and think before acting.

H2: High NFC leads to a more plan-based strategy than low NFC. People with high NFC have high intrinsic motivation to think and engage in effortful cognitive tasks, will show more plan-based behavior, and perform better

METHOD

Design

Subjects: There were 43 subjects (17 male, 26 female, age 19-32) following or having followed higher education. The experiment had two independent variables: interface style (internalization vs. externalization) and cognitive style (low NFC vs. high NFC, based on median).

Material

NFC-questionnaire

We used the 18-item NFC scale [2]. Statements were rated on a Likert-scale ranging from 1-7. Score "1" meant 'strongly disagree' and "7" meant 'strongly agree'.

The Conference Planner Application

The application simulated the planning of a conference. Speakers, each of them having different demands, had to be scheduled. They had to be scheduled in slots in one of three rooms (each with its facilities and availability). Without planning, the scheduling will not be optimal and extra moves besides the optimal path must be made. If one does not examine the *entire* situation, one will get stuck in a later phase because there will remain speakers that do not fit in the slots that are left, and one must make superfluous moves. The difference between externalization (fig. 1) and internalization (fig. 2) was implemented by showing, in the externalization condition where a speaker can be placed. When clicking on a speaker, the possible slots in the timetable turn green ("possible" means "satisfying constraints and available", says nothing about "smartness").

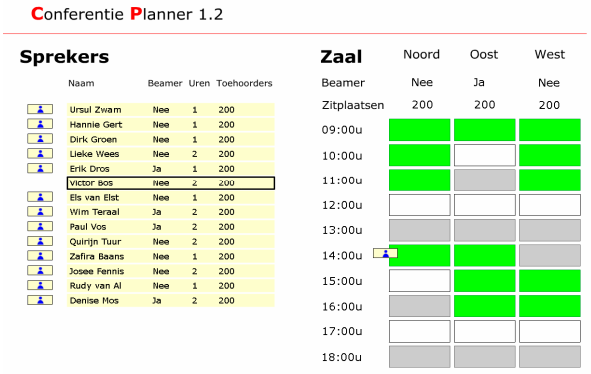

Figure 1 - Conference Planner in externalized condition

In the internalization condition, the interface did not provide this information (fig. 2).

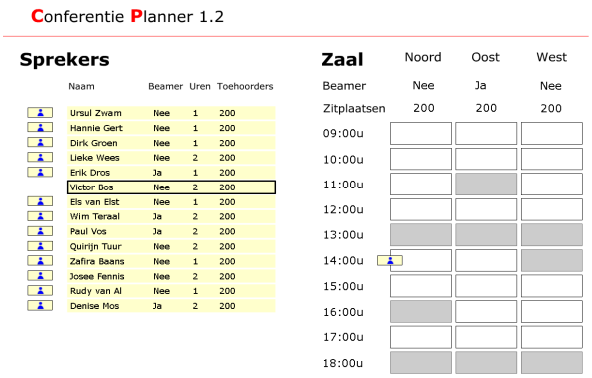

Figure 2 - Conference Planner in internalized condition

On the left are the to be scheduled speakers. Their constraints could vary on a maximum of three variables:

- Amount of speaking hours (1 or 2, in Dutch: *uren*)
- Beamer needed (yes or no, in Dutch: *beamer ja/nee*)
- Number of expected listeners (in Dutch: *toehoorders*)

The speakers had to be assigned to slots in the timetable on the right of the screen in one of the three rooms. The rooms can hold a maximum number of people and have either a beamer or not. Not all timeslots were available (indicated with gray, e.g. lunch/dinnertime and others). The speakers had to be placed, while taking the different constraints into account. The subjects had to perform 5 tasks, in which constraints of the speakers and the rooms varied.

Measures

To estimate planning and performance during the 5 trials we collected:

Time measures

- Time before their first move (how long subjects analyzed the problem before they started solving it)
- Time between moves (how long subjects analyzed the problem state after a move was made, and continued)

Move measures

- Superfluous moves (unnecessary actions or moves, deviations from most economical solution path)
- Correction moves (moves necessary to fix situations in which the subject was stuck).
- Reconsidered moves (how often subjects picked up speakers, and while dragging changed their minds and put it back.

Procedure

First NFC was measured, followed by an explanation of the application. Then the conference planning task started, containing 5 trials. The experiment took about 30 minutes.

RESULTS

The minimum average NFC score was 2.39, the maximum 6.50 (M=4.89, SD=0.83), Median = 5.06. Cronbach's alpha of the 18 statements was 0.89. We statistically analyzed the effects of interface style and cognitive style using ANOVA. All 5 solutions were eventually found across conditions, but "the path" via which subjects reached solutions differed.

Time

There were no significant (p>.05) main effects of interaction style or cognitive style on the *total time* needed to complete the 5 trials.

	Internalization				Externalization			
	Low NFC		High NFC		Low NFC		High NFC	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Time before first move	22.7	7.3	16.9	6.6	15.1	7.2	15.4	7.6
Time between moves	5.3	1.7	4.3	1.0	4.2	1.4	3.7	1.1

Table 1 – Average times (s) in task across conditions

There was a significant main effect of interface style on the time passed before subjects made their *first* move across the 5 trials, F(1,39)=4.32, p<0.05. Internalization subjects took longer than externalization subjects, M=19.8, SD=7.4 vs. M=15.3, SD=7.2. There was also a significant main effect of interface style on the average time *between* moves, F(1,39)=4.82, p<0.05. Internalization subjects took more time between moves, M=4.8, SD=1.4 vs. M=3.9, SD=1.3. No significant main or interaction effects for cognitive style were found.

Moves

Again no significant effects of cognitive style were found.

	Internalization				Externalization			
	Low NFC		High NFC		Low NFC		High NFC	
Number of moves	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Superfluous moves	2.9	2.8	2.1	2.3	4.1	2.1	4.4	3.9
Correction moves	1.1	0.3	1.0	0.3	1.8	0.4	1.6	0.3
Reconsidered moves	0.5	0.5	0.4	0.5	1.7	0.5	2.4	0.5

Table 2 – Average moves in task across conditions

There was a significant main effect of interface style on superfluous moves F(1,39)=4.17, p<0.05. Internalization subjects made fewer superfluous moves than externalization subjects, M= 2.46, SD=0.61 vs. M=4.27, SD=0.63 (fig. 3).

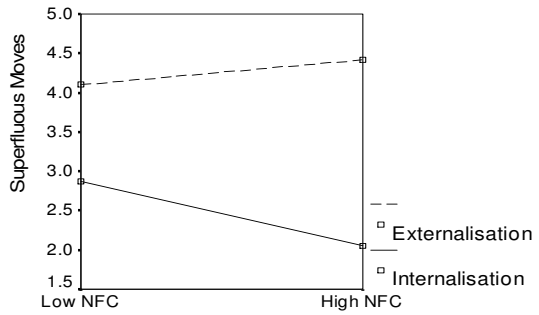

Figure 3–Average superfluous moves across conditions

The effect of interface style on the number of correction moves was nearly significant, $F(1,39)=3.80$, $p=0.06$. Externalization subjects did this more than internalization subjects, M=1.70, SD 0.25 vs. M=1.1, SD=0.24 (fig. 4).

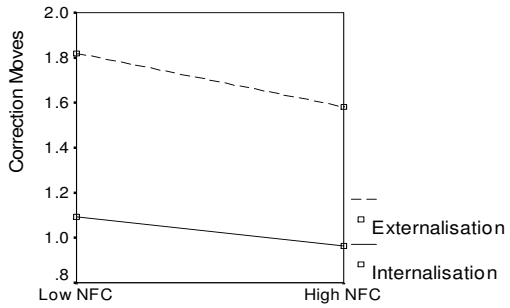

Figure 4–Average correction moves across conditions

Interface style also significantly influenced the number of reconsidered moves, $F(1,39)=9.59$, $p<0.01$. Internalization subjects reconsidered moves less often than externalization subjects, M=0.46, SD=0.36 vs. M=2.05, SD=0.37 (fig. 5).

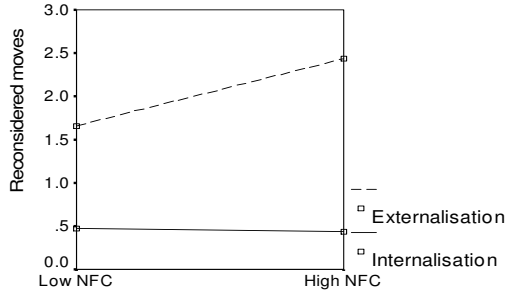

Figure 5–Average reconsidered moves across conditions

DISCUSSION AND CONCLUSION

We investigated the influence of interface style and cognitive style on performance. Our first hypothesis stating that internalization leads to more planning and better performance than externalization is accepted. As in earlier experiments [7] [8], it once again showed that also in this more realistic task, user behavior was different. An interface requiring more internalization resulted in longer thinking time before subjects started working on the problem, and also more time between moves. This indicates that when information has to be internalized, more contemplation from the users is provoked. Still more interesting are the conclusions that can be drawn from the move-based measures. Internalization caused subjects to solve the tasks in a more straightforward way, with less deviation, thus with greater economy, conforming [5]. The issue here was no so much "can they solve it" but "how smart or economical do they solve it", since in the end each problem had a solution. There was also the fact that externalization subjects significantly reconsidered moves more while making them (like a chess player realizing his mistake the moment he picks up a piece. This also suggests that those subjects do less "thinking before you act" than internalization subjects. The second hypothesis has to be rejected. Cognitive style along the dimensions of high versus low NFC had no influence. Pre-existing attitudes towards problem solving did not have effect on the displayed behavior and performance of subjects. Interface style again [8] predicted problem solving behavior quite strongly, and could overrule possible effects of pre-existing individual differences. Also, our subjects were volunteering university students, where the NFC average was very high.

There are still many challenges in HCI. Understanding reactions to interface information (based on cognitive findings) is important in tuning software, thereby allowing it to achieve its goal. With multimedia being present in all corners of society, our findings can be valuable in the development of applications in the realm of education, and multimedia learning. We will continue the research, and broaden the types of problem solving activities, and include other variables related to attitudes to problem solving.

ACKNOWLEDGEMENTS

We thank Imke de Jong, Marijn Plomp & Jasper Roques for their invaluable assistance, and Herre van Oostendorp and Hermi Schijf for their advice during the project.

REFERENCES

1. Brusilovsky, P. Adaptive hypermedia. *User Modeling and User-Adapted Interaction, 11 (2001),* 87-110.
2. Cacioppo, J.T., & Petty, R.E. The need for cognition. *Journal of Personality and Social Psychology*, 42 (1982), 116-131.
3. Crystal, A. & Kalyanaraman, S. Personality variables and information processing on the web: Methodological issues and empirical evidence. *International Communication Association Annual Meeting* (2005), New York, NY.
4. Nielsen, J. *Usability Engineering*. San Francisco: Morgan Kaufmann, 1995.
5. O'Hara, K.P. & Payne, S.J. Planning and the user interface: The effects of lockout time and error recovery cost. *International Journal of Human-Computer Studies, 50,* 1 (1999), 41-59.
6. Scaife, M. & Rogers, Y. External cognition: How do graphical representations work? *International Journal of Human-Computer Studies, 45*(2) (1996), 185-213.
7. Van Nimwegen, C., Van Oostendorp, H. & Tabachneck-Schijf, H.J.M. The over-assisting interface: The role of support in knowledge acquisition, in *Proceedings of HCI2004: Design for Life* (Leeds, UK, 2004), 133-136.
8. Van Nimwegen, C., Van Oostendorp, H. & Tabachneck-Schijf, H.J.M. The role of interface style in planning during problem solving, in *Proceedings of the 27th Annual Cognitive Science Conference* (Stresa, Italy, 2005). Mahwah, NJ: Lawrence Erlbaum, 2271-2276.