Adaptive Attention Allocation Support: Effects of System Conservativeness and Human Competence

Peter-Paul van Maanen[1,2], Teun Lucassen[3], and Kees van Dongen[1]

[1] TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands
{peter-paul.vanmaanen,kees.vandongen}@tno.nl
[2] Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
[3] Department of Cognitive Psychology and Ergonomics, University of Twente
P.O. Box 215, 7500 AE Enschede, The Netherlands
t.lucassen@gw.utwente.nl

Abstract. Naval tactical picture compilation is a task for which allocation of attention to the right information at the right time is crucial. Performance on this task can be improved if a support system assists the human operator. However, there is evidence that benefits of support systems are highly dependent upon the systems' tendency to support. This paper presents a study into the effects of different levels of support conservativeness (i.e., tendency to support) and human competence on performance and on the human's trust in the support system. Three types of support are distinguished: fixed, liberal and conservative support. In fixed support, the system calculates an estimated optimal decision and suggests this to the human. In the liberal and conservative support types, the system estimated the important information in the problem space in order to make a correct decision and directs the human's attention to this information. In liberal support, the system attempts to direct the human's attention using only the assessed task requirements, whereas in conservative support, the this attempt is done provided that it has been estimated that the human is not already paying attention (more conservative). Overall results do not confirm our hypothesis that adaptive conservative support leads to the best performances. Furthermore, especially high-competent humans showed more trust in a system when delivered support was adapted to their specific needs.

1 Introduction

In the domain of naval warfare, information volumes for navigation, system monitoring and tactical tasks will increase while the complexity of the internal and external environment also increases [6]. For tactical picture compilation tasks also the dynamics in behavior and ambiguity of threat is expected to increase. Furthermore, the trend of reduced manning is expected to continue as a result of economic pressures and humans will be responsible for more and more demanding tasks. Although attention can be divided between tasks, problems with attention allocation and task performance are expected since attentional resources are limited [13,8]. Experience, training and better interfaces can lift these limitations, but only to a certain level. Even with experienced users, attentional problems are still likely to occur [11].

D.D. Schmorrow and C.M. Fidopiastis (Eds.): FAC 2011, HCII 2011, LNAI 6780, pp. 647–656, 2011.

Automation can assist humans by directing attention to critical events [14]. It can heuristically identify and prioritize objects of interest by highlighting high priority objects and dimming low priority objects. This helps humans to focus on the right subset of objects and thereby effectively reduces the number of objects that must be monitored. The downside is that this form of cuing can impede the detection of important objects that are mistakenly left unhighlighted when the automation is imperfect or when the situation is uncertain [12]. In these situations problems with inappropriate trust in the automation, resulting in over-reliance on the automation, are expected. This paper suggests to use different types of adaptive automation that adapts the support to the human need of support: only support when really needed, and in this way limit the above mentioned downside of automation as much as possible.

In [7] adaptive automation is defined as follows: "adaptive automation refers to a system capable of dynamic, workload-triggered reallocations of task responsibility between human and machine". There are multiple reasons for applying adaptive automation. How well people perform a certain task is affected by the allocation of their attention. People that are more experienced will be better at dividing attention between different sources of information. Research on the effects of playing video games has shown that visual attention abilities often improve with training. Experienced players of video games required less attentional resources for a given target [5]. In the case of a tactical picture compilation task, experts will be able to track more contacts. Experts will also be able to determine more quickly whether a contact is a possible threat. As opposed to poor performers, good performers will apply the rules correctly. Adaptive automation can help by assisting (the less well performing) humans in their allocation of attention through estimating their current allocation of attention and intervene when the human should reallocate his attention.

In [10] it is proposed to use adaptive automation to prevent 'complacency' and to increase the changes of detecting errors. By shifting task responsibility between humans and machines, humans will be more involved in tasks, which causes more errors to be detected and therefore the performance with respect to monitoring automation will increase.

In this paper the effect is investigated of different types of adaptive automation (decision support) with respect to system conservativeness (high and low) and human competence (good and poor) in terms of task performance, trust, understandability and responsibility. The paper is composed of the following sections. In Section 2 the proposed support types are described and several hypotheses are given and motivated about these support types. Then, in Section 3 a experiment is described in which the support types are evaluated. The results of this experiment are given in Section 4 and the paper ends with a discussion and conclusions in Section 5.

2 Attention Allocation Support

2.1 Generic Support Model

One way of implementing adaptive automation is to use computational cognitive models of attention as a basis for triggering change in automation. A cognitive model of attention is a model which estimates a human's focus of attention at each moment in time for a given task (see e.g., [2,1]). Together with a normative model, which estimates where attention should be optimally allocated for that same moment in time and task, a decision support system can aid the user in distributing limited attentional resources when there is a large difference between the two. In [2,1] it is shown, for instance, that in this way it is possible to support humans in their allocation of attention.

The support evaluated in this study has three variants, namely the *fixed*, *liberal* and *conservative* support type. The fixed support is defined as support that advices a human user what decision to make, without taking into account whether it is needed to support the human at that moment. The outcome of the task is shown to the human who can then decide whether to comply with the advice or to rely on his own judgment. As stated earlier, a potential risk of fixed support is inappropriate reliance. The fixed support system always gives its advice. So the easiest way to perform the task is to follow the advice as given by the system, which can lead to problems with complacency. This means that if the fixed system occasionally gives incorrect advice, it is more likely to be (incorrectly) taken over by the human, compared to an adaptive support system.

The alternative for the fixed support system is to direct the attention of the human to areas that are estimated to need human attention, instead of suggesting a specific decision. This way, the human is supported during an earlier stage of information processing, namely information acquisition, and hence leaving information interpretation and decision making to the human. The result is that the human can no longer completely rely on the support with respect to deciding what to do. Errors in the support are thus likely to be less influential on the decisions of the human. Wrong advice of the support system is also expected to be detected more easily by humans, because the advice is checked more thoroughly due to the fact that it needs to be processed more before a decision is made. This basic idea of bringing the human back 'in the loop' is also the underlying property of the last two support variants.

The liberal and conservative support type are different with respect to system conservativeness. *System conservativeness* is defined as the inverse tendency of the system to provide support to the human. It can be varied through adaptation to the behavior of the human. Examples of this behavior are mouse clicks, reaction times to events, and point of gaze. The models used for liberal adaptive support will use less behavioral data than those for conservative adaptive support. For instance, in the context of the tactical picture compilation task described in [2,1], liberal can be defined as support that adapts only to the current selection of threatening contacts. In this case, it is estimated (by a mathematical model) whether support is needed through adaptation to the clicking behavior of a human operator. For the conservative support, next to adaptation to the clicking behavior, an estimation of the current human attention allocation is also incorporated. Overall this means that the liberal support is less adapted to the user than the conservative support.

2.2 Hypotheses

For the above mentioned three support types, the effects on 1) task performance, 2) trust, 3) understandability and 4) responsibility are studied, which are discussed in sections 2.2, 2.2, 2.2 and 2.2, respectively. In sections 2.2 and 2.2 overall effects of system conservativeness and human competence are discussed. The above mentioned discussions lead to a total of 6 hypotheses.

Task Performance. When support is fixed, humans are expected to be more prone to over- and under-rely on the automation, whereas with adaptive support less problems with inappropriate reliance or complacency are expected. This can be explained by the fact that fixed support allows humans to rely entirely on the support (i.e., just take over the computer's advice). Adaptivity can stimulate the human's involvement in the task by automatically applying support, only where and when needed. It is expected that higher levels of such adaptivity to the human also results in higher task performance. This basically boils down to the following hypothesis:

Hypothesis 1. *The proposed adaptive support results in higher task performance than fixed (non-adaptive) support.*

Trust. Another important factor is trust in the support: Do participants trust adaptive decision support more than fixed decision support? Errors will inevitably occur in the support. However, these errors are likely to be much more salient in when applying fixed support. This boils down to the following hypothesis:

Hypothesis 2. *The proposed adaptive support results in more trust in the support system compared to fixed support.*

Understandability. A potential problem in adaptive support is understandability. Adaptive support systems are likely to be more complex than fixed support systems (or in any case, no support). This leads to the following hypothesis:

Hypothesis 3. *The proposed adaptive support results in a poorer understanding of the support compared to fixed support.*

Responsibility. Since we expect the human to be more involved in the task when applying the proposed adaptive support, we also expect that the responsibility for a good result as felt by the human is higher.

Hypothesis 4. *The proposed adaptive support results in a greater feeling of responsibility for the eventual outcome compared to fixed support.*

System Conservativeness. As has been mentioned, conservativeness can be varied through adaptation to the behavior of the human. This adaptivity can come in various degrees: A more conservative adaptive support system depends to a higher degree on the behavior of the human. When the system is uncertain about information, conservative support will withhold information longer than liberal support. This is expected to result in a stronger effect with respect to task performance, trust and understandability. This boils down to the following hypothesis:

Hypothesis 5. *The claimed effects in Hypotheses 1 to 4 are stronger for conservative adaptive support than for liberal adaptive support.*

Human Competence. We expect that, since adaptive support takes actions of the human into account, the task performance of the human (with or without support) contributes to the performance of the support. When the actions of the human are in line with the task model of the support, the support will be more appropriate. The hypothesis:

Hypothesis 6. *The claimed effects in Hypotheses 1 to 4 are stronger for good performers than for poor performers.*

3 Method

3.1 Participants

Forty college students (17 male, 23 female) with an average age of 23.9 years ($SD = 2.6$) participated in the experiment as paid volunteers.

3.2 Apparatus

Participants had to perform a (simplified) naval tactical picture compilation task as performed in naval warfare. The goal of this task is to build up awareness of possible threats surrounding the own ship (contacts). A screenshot of the task environment is shown in Figure 1.

Participants had to mark the five most threatening contacts by clicking on them. To determine if a contact is a possible threat the following criteria were used: speed, heading, distance to own ship and position in or out a sea-lane. All criteria were equally important. The five contacts scoring highest on these criteria had to be selected as most threatening. The behavior of the contacts was such that the threat varied over time. For instance, a contact could get out of a sea lane, speedup, or change its heading toward the own ship. Contacts that were mistakenly identified as threats (false alarms) or contacts that were mistakenly not identified as threats (misses) resulted in a lower task performance.

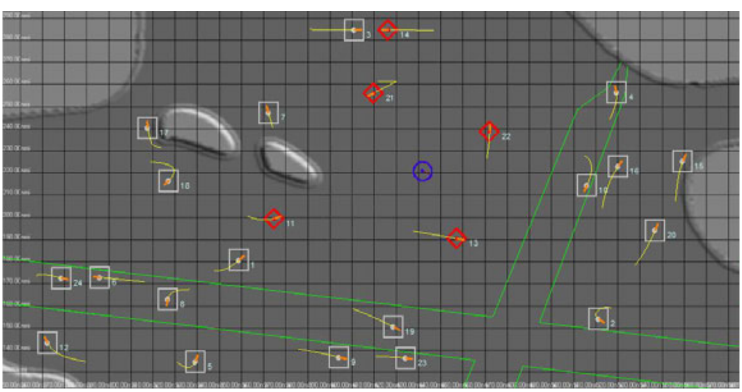

Fig. 1. Screenshot of the task environment

All participants were exposed to the same task complexity. The complexity was determined by the ambiguity and the dynamics of the behavior of the contacts. Concerning ambiguity, small differences in the threat level of contacts made it more difficult to identify the five most threatening contacts. Dynamics were determined by the varying number of threat level changes of contacts over time. Changes in threat levels were such that the number of times that each contact needed to be re-evaluated was high.

The task, including the different developed adaptive support conditions (see Section 3.4), was implemented using the game implementation software Gamemaker.[1]

3.3 Design

A 4 (system conservativeness) $\times$ 2 (human competence) mixed design was used. System conservativeness was a within-subjects independent variable and the order was balanced between the participants. Human competence was a between-subjects quasi-independent variable.

3.4 Independent Variables

Two independent variables were used: system conservativeness and human competence.

System Conservativeness. There were four levels of conservativeness for the used support system: no support, fixed support, liberal adaptive support and conservative adaptive support.

In the no support (NS) condition participants performed the tactical picture compilation task without any form of support.

The fixed support (FS) used a task model to determine the five most threatening contacts. Threat levels were determined on the same criteria as used by the participants. The five most threatening contacts were highlighted.

The liberal adaptive support (LAS) also takes the current selection of the user into account. Firstly, the two contacts with the lowest threat levels which are already selected by the user were highlighted. This was done because they are candidates for deselection and should therefore receive attention. Secondly, the three contacts with the highest threat levels which are not selected by the user were highlighted. This was done because they are candidates for selection and should receive attention for this reason.

The conservative adaptive support (CAS) basically highlighted the same contacts as the liberal support, but now only when the user paid little attention to these contacts. Attention values for all contacts were calculated using the cognitive model of attention proposed in [2,1].

The reliability of the task model in each support condition was manipulated by adding errors to the actual threatlevels. This manipulation was done in order to simulate an imperfect task model. If the task model was perfect, there was no use of comparing FS with the other support types, since it would always be a good decision to follow the advice of FS, resulting in a maximum task performance.

[1] For more information on Gamemaker, see `http://www.yoyogames.com/gamemaker`

Human Competence. After the experiment, a median split on the task performance in the NS condition was performed to distinguish a good and poor human competence group.

3.5 Dependent Variables

Four dependent variables were measured: task performance, trust, understandability and responsibility.

Task Performance. Task performance was determined by measuring the accuracy of the identification of the five most threatening contacts during the task by means of a penalty function. The severity of errors was also taken into account by the task performance measure.

Trust. After each trial participants estimated the reliability of the support system on a scale between 0 and 100% correct. Since trust is for an important part determined by perceptions of reliability [9,4,3], this was considered as a good measure of trust.

Understandability. Participants rated after each trial the degree to which they thought the decision making process of that of the support system was understandable on a 7-point Likert scale between -3 (not understandable) and 3 (understandable).

Responsibility. Participants rated after each trial the degree to which they felt responsible for the outcome of the task on a 7-point Likert scale between -3 (not responsible) and 3 (responsible).

4 Results

Two out of the 40 retrieved data sets have been removed due to unintended errors during the experiment. All participants passed the test on paper and were therefore expected to be able to correctly apply the classification criteria.

Task Performance. Lilliefors tests have shown that task performance in the NS, FS, LAS and CAS conditions were all normally distributed (i.e., null hypothesis that they are normally distributed could not be rejected).

To check whether the design of the fixed support system was a fair competitor for the adaptive variants, at least it should hold that FS results in higher performances than NS. This was indeed the case: Participants in condition FS ($M = 89.24, SD = 2.20$) performed significantly better compared to NS ($M = 87.62, SD = 3.27$), $t(74) = 2.53$, $p < .001$.

A repeated measures analysis of variance (ANOVA) showed a significant main effect ($F(2, 72) = 27.40, p < .001$) of system conservativeness on task performance.

A post-hoc Bonferroni test showed that there is a significant difference between conditions FS and LAS ($p < .001$), FS and CAS ($p < .001$), but not between LAS and CAS ($p = 1$). Hence participants performed worse in the LAS ($M = 87.34$, $SD = 1.73$) and CAS condition ($M = 87.16$, $SD = 1.87$) than in the FS condition ($M = 89.24$, $SD = 2.20$). Hypotheses 1 and 5 (for task performance) are therefore not accepted.

No interaction effect was found ($F(2, 72) = 0.22$, $p = .80$) between system conservativeness and human competence on task performance. Hence Hypothesis 6 (for task performance) is not accepted.

Trust. A repeated measures analysis of variance (ANOVA) did not show a significant main effect ($F(2, 72) = 0.47$, $p = .63$) of system conservativeness on trust. Hypotheses 2 and 5 (for trust) are therefore not accepted.

A significant interaction effect was found ($F(2, 72) = 3.17$, $p = .048$) between system conservativeness and human competence for trust.

A post-hoc Bonferroni test showed that there is a significant difference in trust between the good ($M = 1.11$, $SD = 0.29$) and poor ($M = -0.32$, $SD = 0.29$) competence group in the CAS condition ($p = .02$), but not in the LAS ($p = 1$) and FS ($p = 1$) condition. Hence the claimed effect in Hypothesis 2 is stronger for good performers than for poor performers in the case of CAS. For CAS, Hypothesis 6 (for trust) is therefore accepted, but not for LAS.

Understandability. A repeated measures analysis of variance (ANOVA) did not show a significant main effect ($F(2, 72) = 0.42$, $p = .66$) of system conservativeness on understandability. Hypotheses 3 and 5 (for understandability) are therefore not accepted.

No significant interaction effect was found ($F(2, 72) = 0.92$, $p = .40$) between system conservativeness and human competence for understandability. Hypothesis 6 (for understandability) is therefore not accepted.

Responsibility. A repeated measures analysis of variance (ANOVA) did not show a significant main effect ($F(2, 72) = 0.37$, $p = .69$) of system conservativeness on responsibility. Hypotheses 4 and 5 (for responsibility) are therefore not accepted.

No significant interaction effect was found ($F(2, 72) = 1.39$, $p = .26$) between system conservativeness and human competence for responsibility. Hypothesis 6 (for responsibility) is therefore not accepted.

5 Conclusion and Discussion

In this study we investigated the benefits of adaptive attention allocation support over fixed (non-adaptive) support in a tactical picture compilation task. We expected task performance using adaptive support to be higher than in fixed support. However, this first hypothesis was not accepted. Trust in adaptive and fixed support did not differ significantly, also rejecting the second hypothesis. In contrary to our third hypothesis, our participants did not report to have a poorer understanding of the more complicated adaptive support than of the fixed support. Also the fourth hypothesis, stating that the feeling of responsibility would be higher in the adaptive condition, could not be accepted based on the results in this study.

The influence of system conservativeness and human competence was also investigated on the first four hypotheses on task performance, trust, understandability and responsibility. The results did not show a significant effect of system conservativeness on any of these variables, so the fifth hypothesis could not be accepted.

For human competence, the effect was significant for trust, but only in the conservative adaptive support condition. This means that well performing participants had more trust in conservative adaptive support than poorly performing participants. This confirms the sixth hypothesis (for trust), but only for the conservative condition. The sixth hypothesis could not be accepted for task performance, understandability, and responsibility. The increase of trust in the conservative adaptive support for good performers can be explained by the effect that good performers are more likely to understand the task and the effect support systems have on task performance. In [9], for instance, it is shown that the use of automation decreases when the effect of automation to performance is not properly perceived.

There are several possible explanations for why an increase of task performance was not found in our experiment. Our implementation of adaptive support aimed at reducing inappropriate reliance on fixed support. However, this comes with a cost in the form of added complexity. Although participants did report a clear understanding of how both adaptive support systems worked, it is still possible that the disadvantage of added complexity is larger than the advantages of such a system. Working with complex (support) systems can raise the cognitive load on the human, leaving less capacity to focus on the actual monitoring of contacts. In this case, this resulted in a significantly higher task performance in the fixed support condition than in both adaptive support conditions. Future design of adaptive support systems should aim at keeping the system as simple as possible, though preserving the expected advantages of adaptivity.

For the adaptive support investigated in this study, it was not possible for the human to simply follow suggestions of the support system. This was because, instead of suggesting a possible answer to a problem, only areas of interest were indicated by the system. This meant that, in any case, the proposed adaptive support must have eliminated inappropriate reliance on the support. The found results in this study are not a reason for rejecting this principle and therefore more research on adaptive attention allocation support is suggested, focusing on the *requirements* in which such a system can help to gain task performance.

Acknowledgments

This research has partly been supported by the programme Cognitive Modeling" (V524), funded by the Dutch Ministry of Defense. The authors would also like to thank Karel van den Bosch, Tibor Bosse, Anja Langefeld and Jan-Willem Streefkerk for their helpful comments.

References

1. Bosse, T., van Lambalgen, R., van Maanen, P.P., Treur, J.: Attention manipulation for naval tactical picture compilation. In: Proceedings of the 2009 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2009 (2009)
2. Bosse, T., van Lambalgen, R., van Maanen, P.P., Treur, J.: Automated visual attention manipulation. In: Paletta, L., Tsotsos, J.K. (eds.) WAPCV 2008. LNCS, vol. 5395, pp. 257–272. Springer, Heidelberg (2009)
3. Dzindolet, M.T., Beck, H.P., Pierce, L.G., Dawe, L.A.: A framework of automation use. Technical Report ARL-TR-2412, Army Research Laboratory, Aberdeen Proving Ground, MD (2001)
4. Gao, J., Lee, J.D.: Extending decision field theory to model operator's reliance on automation in supervisory control situations. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans 36(5), 943–959 (2006)
5. Green, C.S., Bavelier, D.: Action video game modifies visual selective attention. Nature 423, 534–537 (2003)
6. Grootjen, M., Neerincx, M.A., van Weert, J.C.M.: Task-based interpretation of operator state information for adaptive support. In: Foundations of Augmented Cognition: Strategic Analysis, 2nd edn. LNCS, pp. 236–242. Springer, Arlington (2006)
7. Hilburn, B., Jorna, P., Byrne, E., Parasuraman, R.: The effect of adaptive air traffic control (atc) decision aiding on controller mental workload. In: Human-Automation Interaction: Research and Practice, pp. 84–91 (1997)
8. Kahneman, D.: Attention and effort. Prentice Hall, Englewoods Cliffs (1973)
9. Lee, J.D., See, K.A.: Trust in automation: Designing for appropriate reliance. Human Factors 46(1), 50–80 (2004)
10. Parasuraman, R., Riley, V.A.: Humans and automation: Use, misuse, disuse, abuse. Human Factors 39, 230–253 (1997)
11. Pavel, M., Wang, G., Li, K., Li, K.: Augmented cognition: Allocation of attention. In: Proceedings of 36th Hawaii International Conference on System Sciences, pp. 286–300. IEEE Computer Society, Los Alamitos (2003)
12. John, M.S., Smallman, H.S., Manes, D.I., Feher, B.A., Morrison, J.G.: Heuristic automation for decluttering tactical displays. Human Factors 47, 509–525 (2005)
13. Wickens, C.D.: Processing resources in attention. In: Parasuraman, R., Davies, D.R. (eds.) Variaties of attention, pp. 63–101. Academic Press, Orland (1984)
14. Wickens, C.D., McCarley, J.S.: Applied attention theory. CRC Press, Boca Raton (2007)