An Integrated Agent Model for Attention and Functional State

Tibor Bosse[1], Rianne van Lambalgen[1], Peter-Paul van Maanen[1,2], and Jan Treur[1]

[1] Vrije Universiteit Amsterdam, Department of Artificial Intelligence,
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
{tbosse, rm.van.lambalgen, treur}@few.vu.nl
[2] TNO Human Factors, P.O. Box 23, 3769 ZG, Soesterberg, The Netherlands
peter-paul.vanmaanen@tno.nl

Abstract. To provide personalized intelligent ambient support for persons performing demanding tasks, it is important to have insight in their state of attention. Existing models for attention have difficulties in distinguishing between stressed and relaxed states. To solve this problem, this paper proposes to extend an existing model for attention with a model for 'functional state'. In this integrated agent model, output of a functional state model (experienced pressure) serves as input for the attention model; the overall amount of attention is dependent on the amount of experienced pressure. An experiment was conducted to test the validity of the integrated agent model against the validity of an earlier model based on attention only. Results pointed out that the integrated model had a higher validity than the earlier model and was more successful in predicting attention.

1 Introduction

For persons performing complex and demanding tasks, it is crucial to have sufficient *attention* for the various subtasks involved. This is particularly true for tasks that involve the continuous inspection of (computer) screens. For instance, an air traffic controller inspecting the movements of aircrafts can not permit him- or herself to miss part of the events that occur. The same holds for a naval operator monitoring the movements of hostile vessels on a radar screen. In such situations, a person may be supported by an intelligent ambient agent system [1], that keeps track of where his or her attention is, and provides some personalized assistance in case the attention is not where it should be, see, e.g., [4], [13].

The current paper is part of a larger project that aims to develop such an agent-based intelligent ambient support system. The main application domain of this system will be naval missions, and one of its main goals will be to support naval operators that work in the control room of the vessels. For example, in case such an operator is directing its attention on the left part of a radar screen, but ignores an important contact that just entered the radar screen from the right, such a system may alert him or her about the arrival of that new contact. To be able to provide this kind of intelligent personalized support, the system somehow needs to maintain a model of the cognitive state of the person: in this case the human's focuses of attention*. It should have the

* Note that in this paper, a rather wide definition of the term 'attention' is used, covering not only visual attention, but also 'mental' attention for objects that have been observed some time earlier, often referred to as 'situational awareness' [5].

capability to attribute mental, and in particular attentional (e.g., [12]) states to the human, and to reason about these.

In previous work, an initial version of such a model has been developed [3], and evaluated positively [13]. This model takes two types of sensor information as input, namely information about the human's gaze (e.g., measured by an eye tracker), and characteristics of stimuli (e.g., the colour and speed of airplanes on radar screens, or of the persons on surveillance images). Based on these types of information, it estimates where the human's attention is and uses this to decide whether adaptive support is needed.

However, one important shortcoming of that model is that it assumes that the *total amount of attention* a person can spend at a particular time point (in the remainder of this paper referred to as *A(t)*) is static and known beforehand. However, it is known from literature like [8] that factors like *A(t)* usually vary over time, depending on states and characteristics of a person [6]. More specifically, it may depend on a human's *functional state*. According to [11], (an operator's) functional state refers to 'the multidimensional pattern of processes that mediate task performance under stress and high workload, in relation to task goals and their attendant physiological and psychological costs'. It is usually assumed to be based on notions like the person's *experienced pressure* and *exhaustion*.

In recent years, researchers have started to develop computational models for the concept of functional state. One of the most sophisticated models is presented in [2]. This model takes task demands, situational aspects, and some of the human's personal characteristics as input, and uses these to asses the human's functional state. Inspired by these developments, the goal of the research reported in the current article has been to develop an integrated agent model for attention and functional state. Our main hypothesis is that the integrated model (which will be called *attention+* from now on) has a higher validity (i.e., is more accurate in estimating where a person's attention is) than the original model (called *attention-*) from [3].

The structure of this paper is as follows. First, the original models for attention [3] and functional state [2] will be briefly described, as well as a proposal to integrate them. Next, an experiment is described that has been performed to compare the validity of the *attention+* model with that of the *attention-* model. The context of the experiment is a shooting task, which is representative for complex tasks that are currently performed in the naval domain. After that, the results of the experiment are analysed. The paper is concluded by a discussion.

2 The Two Submodels and their Integration

The introduced integrated agent model (i.e. the *attention+* model) is composed of two main submodels, namely (1) a basic attention model (i.e., the *attention-* model) and (2) a functional state model. Below, both of them will be briefly summarised. Next, a detailed explanation is provided about how they are combined.

2.1 Attention Submodel

The attention submodel was taken from [3]. The model uses three types of input: information about the human's *gaze direction*, about *locations* (or spaces) and about *features* of objects on the screen (see Figure 1, where the circles denotes the italicised concepts, and the arrows indicate influences between them). Based on this, at each time point t it makes an estimation of the *current attention distribution*: an assignment of attention values $AV(s,t)$ to a set of attention spaces s at that time. The attention distribution is assumed to have a certain persistency. At each point in time the new *attention level* is related to the previous attention, by:

$$AV(s,t) = \lambda \cdot AV(s,t\text{-}1) + (1 - \lambda) \cdot AV_{norm}(s,t)$$

Here, λ is the decay parameter for the decay of the attention value of space s at time point $t-1$, and $AV_{norm}(s,t)$ is determined by normalisation for the total amount of attention $A(t)$, described by:

$$AV_{\mathrm{norm}}(s,t) = \frac{AV_{\mathrm{new}}(s,t)}{\sum_{s'} AV_{\mathrm{new}}(s',t)} \cdot A(t)$$

$$AV_{\mathrm{new}}(s,t) = \frac{AV_{\mathrm{pot}}(s,t)}{1 + \alpha \cdot r(s,t)^2}$$

Here, $AVnew(s,t)$ is calculated from the potential attention value of space s at time point t and the relative distance of each space s to the gaze point (the centre). The term $r(s,t)$ is taken as the Euclidian distance between the current gaze point and s at time point t (multiplied by an importance factor α which determines the relative impact of the distance to the gaze point on the attentional state, which can be different per individual and situation):

$$r(s,t) = d_{\mathrm{eucl}}(gaze(t), s)$$

The potential attention value $AVpot(s,t)$ is a weighted sum of the features of the space (i.e., of the types of objects present) at that time (e.g., luminance, colour):

$$AV\mathrm{pot}(s,t) = \sum_{maps\ M} M(s,t) \cdot w_M(s,t)$$

For every feature there is a saliency map M, which describes its potency of drawing attention (e.g., [12]). Moreover, $M(s,t)$ is the unweighted potential attention value of s at time point t, and $w_M(s,t)$ is the weight used for saliency map M, where $1 \leq M(s,t)$ and $0 \leq wM(s,t) \leq 1$.

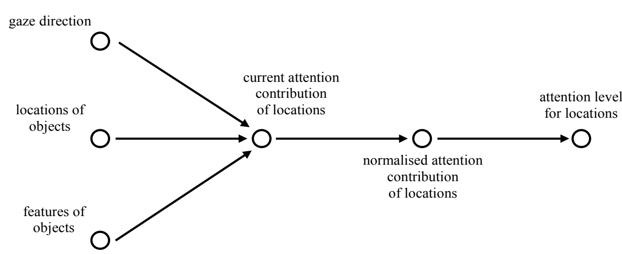

Fig. 1. Overview of the Attention model

For a more detailed description of the model and the underlying theories, see [3], [4].

2.2 Functional State Submodel

The functional state (FS) submodel was adopted from [2] and determines a person's *functional state* as a function of task properties and personal characteristics. The model is based on two different theories: (1) the cognitive energetic framework [10], which states that effort regulation is based on human recourses and determines human performance in dynamic conditions; (2) the idea, that when performing sports, a person's generated power can continue on a *critical power* level without becoming more exhausted [9]. The FS of a human represents the dynamical state of the person. In the model (see Figure 2), this is defined by a combination of exhaustion, motivation and experienced pressure, but also the amount of generated and provided effort. Due to space limitations no further details of the model are provided here. However, for a detailed description, see [2].

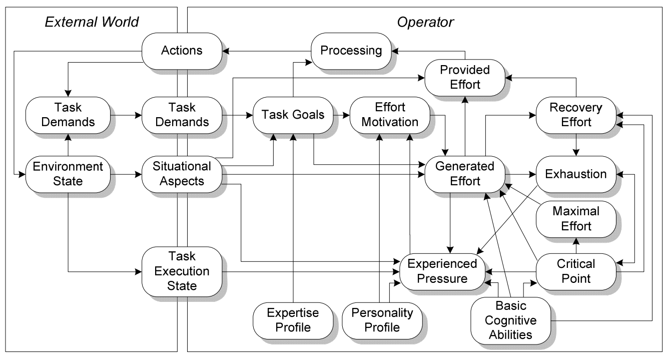

Fig. 2. Overview of the Functional State model

The most important variable from the FS model that is used in this paper is the *experienced pressure*. Here, this variable is used to determine the amount of available attention (the precise relation is explained in the next section). In the FS model, experienced pressure is related to a number of factors, such as the amount of exhaustion, the amount of effort related to the critical point and the performance quality. The strength of these relations is dependent on personality characteristics like exhaustion sensitivity, performance norm and performance sensitivity.

2.3 Integrating the Attention and Functional State Model

One of the drawbacks of the *attention-* model is that it does not take into account that the amount of attention may vary over time. However, in reality, this amount of attention is influenced by different aspects of the functional state, in particular by the experienced pressure. Experienced pressure results in variances in concentration and motivation, which are directly related to attention. This is the idea behind the integration of the previously explained submodels: the output variable experienced pressure of the Functional State submodel is used as an input variable for the total amount of attention $A(t)$ in the Attention submodel. This is done in the following way:

$$A(t) = a + b \cdot (1 - EP(t))$$

where $EP(t)$ is the experienced pressure at time point t, and $0 \le a \le 1$ and $-1 \le b \le 1$ are parameters that can be tuned. If the used Functional State submodel is valid, this means that also the *attention+* model will have an improved validity, due to the FS's capability to alter $EP(t)$ at the appropriate times and therefore dropping or taking into account the part of the estimation where *attention-* was the least certain of.

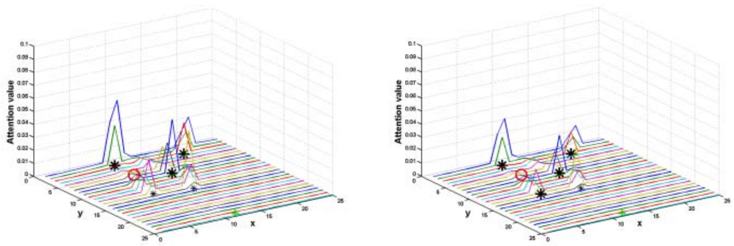

Fig. 3a. Output *attention-* (left) and *attention+* (right) with a low situational demand

In Figure 3a and b an example is given of this capability. A visualisation of the outcomes of the models *attention-* and *attention+* is shown for the situations in which there is a low situational demand (a) and a high situational demand (b). Here, the area determined by the x- and y-axis represents a radar screen, the circles denote contacts, and the z-axis indicates the estimated level of attention. The situational demand is related to the amount of contacts to be handled with a certain time interval.

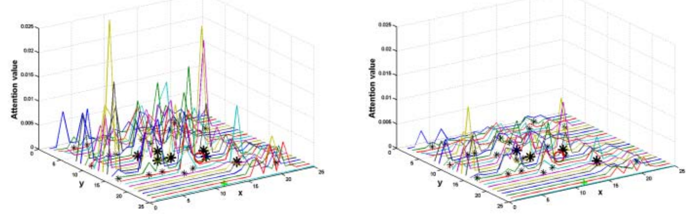

Fig. 3b. Output *attention-* (left) and *attention+* (right) with a high situational demand

For fixed decision criteria *attention+* is able to adapt to the expected change of the functional state of the user, whereas *attention-* is not. In Figure 3 this means that more objects are estimated to be attended to in the low situational demand (4 opposed to 3) and less in the high situational demand condition (4 opposed to 7).

3 Experiment

The goal of the experiment was to investigate the difference in validity between the attention model connected to the FS model (*attention+*) and the original attention model (*attention-*). The hypothesis is that the validity of the *attention+* model is higher than the validity of the *attention-* model.

3.1 Participants

Three female and two male participants with a mean age of 24.67 took part in this study. All participants already had some experience with the task environment.

3.2 Simulation-Based Training Environment

The main task that was used in this study consists of identifying incoming contacts and, based on the outcome of identification, deciding to eliminate the contact (by shooting) or allowing it to land (by not shooting). A screenshot of this simulation-based training environment is displayed in Figure 4. The object at the bottom of the screen represents the participant's (stationary) weapon. In addition, contacts (allies and enemies in the shape of a dot with a radius of 5 pixels) appear at a random location on the top and fall down to random locations at the bottom of the screen.

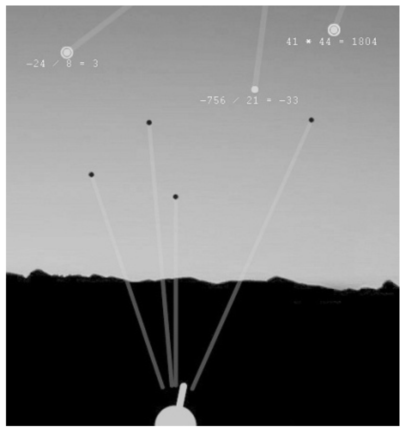

Fig. 4. Screenshot of the Task Environment

Before a contact can be identified, it has to be perceived. This is done by a mouse click at the contact, which reveals a mathematical equation underneath the contact. The identification task is to check the correctness of the mathematical equation (which is less difficult in less demanding situations). A correct equation means that the contact is an ally; an incorrect equation indicates that the contact is an enemy. Identification is done by pressing either the left or right arrow for respectively an ally or enemy. When a contact is identified a green (for an ally) or a red (for an enemy) circle appears around the contact.

The contacts that are identified as an enemy have to be shot before they land. A missile is shot by executing a mouse click at a specific location; the missile will move from the weapon to that location and explode exactly at the location of the mouse click. Any contact within a radius of 50 pixels of the exploding missile is destroyed.

3.3 Procedure

The experiment consisted of 4 blocks of 20 minutes of the simulation-based task environment. In the first 10 minutes of one block, task demands were low (contacts appear every 10 to 20 seconds) and in the second 10 minutes of one block, task demands were high (contacts appear every 2.25 to 4.5 seconds). In the first and third block, 'freezes' were made after each 2.5 minute, in the second and fourth block no freezes were made. When a freeze was made, the experiment was put on hold and the following sentence was shown: "Gameplay frozen. Select contacts, press space when done.” At this moment, participants had to select all contacts which they thought to have recently paid attention to. After selection, a computer version of the NASA-TLX was shown, where participants had to indicate their performance and mental effort.

First, a Tobii x50 eye tracker (http://www.tobii.se) was connected to measure eye movements as input for the attention model. After calibration of the eye tracker, the experiment was started and onscreen instructions were given on the task environment and freezes. The instructions were followed by a practice block of 1 minute low task demands and 1 minute high task demands to get familiar with the environment. After practice, participants started with the first block. Before each block the eye tracker status was checked and after each block, the participant was given a three minute break before continuing with the next block.

When the participant finished the experiment, the data on the task difficulty (situational demands) and performance quality were used for tuning of the parameters in the FS model. This was done using a Simulated Annealing technique. This method initially selects a random parameter setting as the best available parameter setting, then introduces a small change in these settings to generate a neighbour of the current parameter settings in the search space. If this neighbour is found to be a more appropriate representation of the observed human behaviour then it is marked as the best known parameter setting, otherwise a new neighbour is selected to evaluate its appropriateness. For more details about how this technique was applied, see [2].

Next, the obtained personal parameters were used to calculate the experienced pressure and the exhaustion in the FS model, which served as input for predicting attention in the $attention_{+}$ model. Furthermore, both eye movements and features of contacts (luminance, colour, ...) served as input for the attention model. At this point, also the remaining parameters of the attention model (i.e., λ, α, and the different w_M, see Section 2.1) and of the connection between both models (i.e., a and b, see Section 2.3) are tuned using Simulated Annealing, to obtain an optimal performance.

3.4 Data Analysis

The output of the *attention*- and *attention*+ models have been compared with subjective data retrieved during freezes in the experiment. In 40 minutes, after each 2.5 minutes of the task execution time a freeze was initiated, where the participant was asked to point out to what objects she was paying attention to. At the same time, *attention*- and *attention*+ also pointed out what they thought was the case. Each freeze in an easy condition was coupled with one in a hard condition in order to be able to evaluate the performance of the models given that the task demand changes over time (see Table 1).

Table 1. Freeze couples

Freeze couple nr	Freeze nr (easy)	Freeze nr (hard)
1	1	5
2	2	6
3	3	7
4	4	8
5	9	13
6	10	14
7	11	15
8	12	16

The procedure used to compare the models with the subjective data retrieved during the freezes is described in the next section.

4 Results

This section presents the results of the experiment. An example of the experienced pressure as predicted from the functional state model is shown in Figure 5, for participant 1 and 2. These two participants were quite extreme cases, in the sense that participant 1 experienced much pressure, and participant 2 experienced little pressure. As can be seen, both participants experience less pressure during the blocks with low task demands (time point 0-1500, 3000-4500, and 6000-7500) than during the blocks with high task demands. Recall that these fluctuations of experience pressure were used to determine the values of the total amount of attention *A(t)* in the *attention+* model. Unfortunately, results of participant 5 could not be used, as the functional state model provided unreliable data.

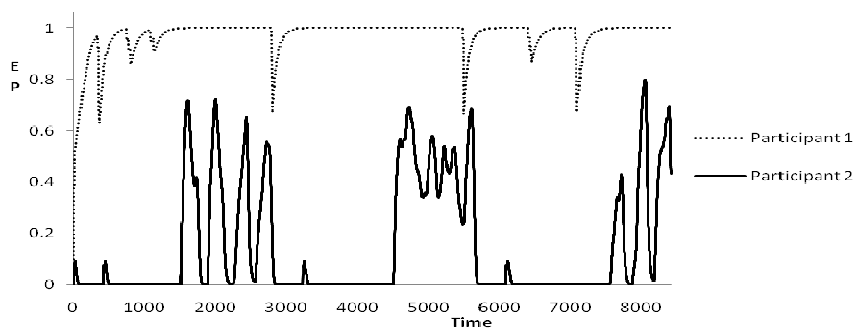

Fig. 5. Estimated Experienced Pressure over time of participant 1 and 2 during three blocks

To evaluate the models, a performance measure was chosen based on the calculation of true positives (hit rate) and false positives (false alarm rate), which can be extracted from confusion matrices, as is shown in Table 2.

Table 2. Confusion matrix

		Participant		
		t	f	*total*
Model	t'	Hits	False Alarms	T'
	f'	Misses	Correct Rejections	F'
	total	T	F	

In this table, *t* and *f* represent whether the participant indicated that he allocated attention to an object or not, respectively, and *t'* and *f'* indicate that one of the models indicated it or not, respectively. Moreover, *Hits/T* results in the hit rate and *False Alarms/F* results in the false alarm rate. Based on these notions, the *sensitivity score d'*, which is a measure for the performance of the model, is determined as follows:

where *zscore(X)* is a function that represents *X* in terms of standard deviations from the average.

The results of the experiment show that the average performance of the *attention+* model ($M+ = 0.736$, $SD+ = 0.118$) was significantly higher than for the *attention-* model ($M- = 0.661$, $SD- = 0.128$), with $t(31) = 4.709$, $p < 0.001$ (paired t-test). Average d-primes per participants are displayed in Figure 6.

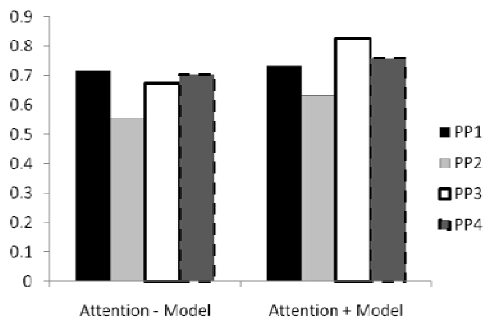

Fig. 6. Mean d-prime per model averaged over participants.

5 Discussion

For personalized ambient agents supporting persons performing demanding tasks, it is important that they have insight in various aspects of the person's mental state [5], [14]. One important aspect is the distribution of the person's attention over the objects (s)he observes. However, existing models for human attention (e.g., [3]) assume that the total amount of attention a person can spend is static. As a result, such models cannot distinguish situations in which a person experiences a lot of pressure from situations in which (s)he is completely relaxed, whereas in reality these situations result in very different behaviours.

To solve this problem, the contribution of the current paper was to extend the original attention model from [3] with a component to keep track of a person's functional state. A first experiment provided evidence that the validity of the *attention+* model was slightly higher than the validity of the *attention-* model.

Despite this encouraging result, the limitations of the approach should not be ignored. First, the amount of participants in the experiment (only 5) was too low to be able to draw strict conclusions. Second, the results were difficult to evaluate, due to a number of complicating factors. For example, the presence of the 'freezes' used for the subjective evaluation may have interfered with the task. Third, the approach assumes that participants are sufficiently capable of estimating where their own attention is. Although some evidence exists that this is indeed the case [13], this assumption can be tested more precisely. Finally, it is an open question to what extent the results can be generalised to other scenarios and circumstances.

In future work, it is planned to address these concerns. For example, experiments with higher numbers of participants are planned, both with the current setup and within a different experimental context. In addition, more work will be spent on fine-tuning of the parameters involved in the model (using standard techniques such as simulated annealing and gradient-based parameter estimation). Finally, on the long term, it is planned to actually implement an intelligent ambient agent system supporting humans in demanding circumstances, and test this in more realistic scenarios.

Acknowledgement

The authors wish to thank Rogier Oorburg and Michael de Vos for their contribution to the training environment, and Alexei Sharpanskykh and S. Waqar Jaffry for their contribution to the tuning of the parameters of the model.

References

1. Aarts, E., Harwig, R., and Schuurmans, M.: Ambient Intelligence. In: P. Denning (ed.), The Invisible Future, McGraw Hill, New York, pp. 235-250 (2001)
2. Bosse, T., Both, F., Hoogendoorn, M., Jaffry, S.W., Lambalgen, R. van, Oorburg, R., Sharpanskykh, R., Treur, J., and Vos, M. De: Design and Validation of a Model for a Human's Functional State and Performance. International Journal of Modeling, Simulation, and Scientific Computing 2, 1-31 (2011)
3. Bosse, T., Maanen, P.-P. van, Treur, J.: Simulation and Formal Analysis of Visual Attention. Web Intelligence and Agent Systems Journal 7, 89-105 (2009)
4. Bosse, T., Lambalgen, R. van, Maanen, P.P. van, and Treur, J.: A System to Support Attention Allocation: Development and Application. Web Intelligence and Agent Systems Journal, in press. (2012)
5. Endsley, M.: Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors 37, 32-64. (1995)
6. Gonzalez, C.: Task workload and cognitive abilities in dynamic decision making. Human Factors 47, 92-101 (2005)
7. Green, D.M., Swets J.A.: Signal Detection Theory and Psychophysics. New York: Wiley (1966)
8. Hancock, P.A., Williams, G., Manning, C.P., and Miyake, S.: Influence of task demand characteristics on workload and performance. The International Journal of Aviation Psychology 5, 63-86 (1995)
9. Hill, D.W.: The critical power concept. Sports Medicine 16, 237-254. (1993)
10. Hockey, G.R.J.: Compensatory control in the regulation of human perfromance under stress and high workload: a cognitive-energetical framework. Biological Psychology 45, 73-93 (1997)
11. Hockey, G.R.J.: Operator Functional State as a framework for the assessment of performance degradation. In: Hockey, G.R.J., Gaillard, A.W.K., and Burov, O., Operator Functional State, pp. 8-23. Amsterdam, IOS press (2003)
12. Itti, L. and Koch, C.: Computational Modeling of Visual Attention, Nature Reviews Neuroscience 2, 194-203 (2001)
13. Maanen, P.-P. van, Koning, L. de, and Dongen, K. van: Design and Validation of HABTA: Human Attention-Based Task Allocator. In: Proc. of the First International Workshop on Human Aspects in Ambient Intelligence. Published in: M. Mühlhäuser, A. Ferscha, and E. Aitenbichler (eds.), Constructing Ambient Intelligence: AmI-07 Workshops Proceedings, CCIS, vol. 11, pp. 286-300. Springer Verlag, Berlin-Heidelberg (2008)
14. Wilson, G.F. and Russell, C.A.: Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. Human Factors 49, 1005-1018 (2007)