

Subjective Evaluation of Task Allocation: An Application of the Analytic Hierarchy Process

Z.G. Wei, A.R. Paauw and P.A. Wieringa

*Man-Machine Systems Group, Laboratory for Measurement and Control
Faculty of Mechanical Engineering and Marine Technology, Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands*

Abstract: This paper describes an application of the Analytic Hierarchy Process (AHP) in subjective evaluation of task allocation between human and automation for a simulated process. The task allocation was evaluated by human subjects with respect to easy operation, human control confidence, comfortable operation, and situation awareness. The experimental results reveal that the AHP can be used to assess human mental load and to give a reasonably consistent ranking of task allocation decisions based on different criteria. However, the AHP only provides a reference for the evaluation. As the number of alternatives increases, the implementation of the AHP will become more difficult.

1. INTRODUCTION

Task allocation between human operators and automation in the control room of complex systems is one of the important phases in human-machine system design (Grote *et al.*, 1995). The evaluation of decisions of task allocation is a necessary step in the task allocation phase. The evaluation involves many aspects, such as performance requirements, operational safety, cost and benefits, or more details, operational difficulty, human mental load, the level of automation and situation awareness, etc. Depending on system design requirements, the evaluation process employs different techniques, such as subjective evaluation and objective evaluation, workload analysis, part-task simulation as well as full scale simulation (Rouse and Cody, 1986). For the subjective evaluation, human subjects must be employed. After they have operated the system the subjects are asked to rate the allocation decisions based on the given criteria. The data from this rating can be processed statistically. However, if the population of subjects is small and the data does not satisfy the conditions for statistical analysis, one may have difficulty. In this paper, we present an application of a so-called Analytic Hierarchy Process (AHP) approach (Saaty, 1980) in subjective evaluation of task allocation decisions based on a simulated process.

2. METHOD

2.1 Experimental set-up

A simulation of a juice pasteurization plant (developed originally by Muir, 1989; and used by Huey, 1989, and Lee and Moray, 1992) was used as the experimental task. The plant employed in this study was adapted from the version used by Huey (1989) and some new subtasks were defined comparing with the old plant. Fig. 1 depicts the simulated pasteurization plant which was also the human-machine interface presented to the operator and was built on a Pentium PC using a graphic language,

LabVIEW. The pasteurization process was to heat the raw juice to a level where bacteria in the juice could be killed. As shown in Fig.1, the plant consisted of two loops. One was the juice circulation loop, and another one was the steam circulation loop.

In the juice circulation loop, the raw juice entered the input tank through the inflow pipe with a certain rate and temperature. From the input tank, the raw juice was pumped by the feedstock pump through the secondary heater and the primary heater to a distribution system. The secondary heater was a passive heat exchanger and was warmed by the hot pasteurized juice. The temperature, T_d , of the juice exiting the primary heater, must be maintained between 70 °C and 85 °C. If T_d was within this range, the juice was led to the distribution tanks through a 3-way valve. If T_d was below the allowed temperature, the heated juice was recycled to the input tank through the 3-way valve, unless the level of the input tank was high. If the level of the input tank was too high, the juice was diverted to the waste tank. If T_d was above the allowed temperature (the heated juice was burnt), the juice was directed to the waste tank.

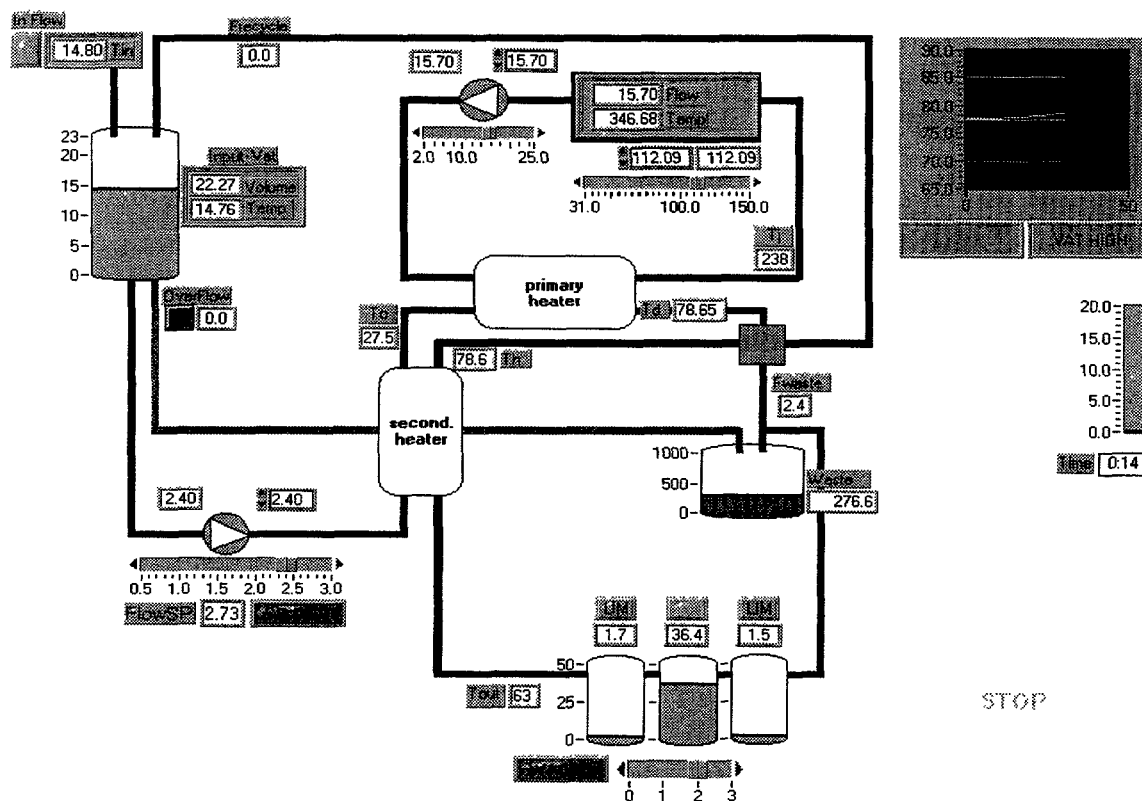


Figure 1 Pasteurisation plant interface

In the steam circulation loop, the steam was first pumped to the primary heater where the heat exchange between the juice and the steam took place, and then was sent back to the boiler. The temperature of the primary heater, T_j , could be controlled by adjusting the steam rate and the temperature of the steam heater. The heater temperature was controlled by changing the supplying amount of the energy entering the heater.

2.2 Experimental task

The *experimental task* was to pasteurize as much of the available raw juice as possible, which meant to maintain the pasteurization temperature with the allowed range, and to distribute it to an available tank. The experimental task could be broken into *subtasks*, or *sub-subtasks*, and consisted of a number of

monitoring and control subtasks. In the experimental task, the important variables to be controlled were the temperature T_d , and the juice flow rate. T_d depended mainly on the juice flow rate and the temperature of the steam heater. The juice flow was maintained to keep the input tank neither empty nor overflow. The control of the feedstock pump speed, of the steam pump, and of the steam heater could be defined as subtasks. In the present study, three control subtasks were defined and could be allocated either to the operator or to the automation.

Subtask 1: Control of juice flow rate and input tank level.

This subtask included two components: feedstock pump speed control and overflow valve control. The speed of the feed pump affected the juice flow which would further affect T_d , and the level of the input tank. The overflow valve should be opened to deliver the juice to the waste tank if the juice level of the input tank was too high. The input tank was prevented from empty by reducing the feedstock pump speed.

Subtask 2: Control of the temperature of steam entering the primary heater, T_j .

This subtask could be further divided into the heater temperature control and the steam pump control. In the control of T_j , the steam heater control was a rough adjustment, and the steam pump control was a fine adjustment.

Subtask 3: Storage distribution control.

Three storage tanks were used to store the pasteurized products. These tanks might be emptied irregularly by tanker trucks. If one of the tanks reached its maximum capacity, the surplus product was diverted to the waste tank. In order to maintain a high production, this should be avoided. Moreover, the effort should be made to let the storage tanks not empty. Otherwise, alarm would appear, and it would be difficult to maintain the maximum capacity of other tanks. It is obvious that this subtask did not affect the pasteurization process, but it did influence the amount of the pasteurized juice.

From the above definitions, it can be found that the plant provides two types of tasks. The control of the pump speed and the steam temperature is a continuous task, while the control of the overflow valve and the distribution belongs to the discrete task. Because all the subtasks could be allocated to automation, each subtask had an automatic controller. Each controller worked under its control algorithm and acted independently with each other.

2.3 Operator's task

When the operator controlled the plant, even though he performed only one subtask, he had to fulfill one of the following three objectives: (1) to maintain the temperature of the juice flowing out of the primary heater, (2) to maintain the input tank at a proper level, and (3) to distribute the pasteurized juice and to keep all storage tanks neither full nor empty. The ultimate goal was to pasteurize as much of the available raw juice as possible and to minimize the volume of the waste tank. During the experiment, the operator was asked to assess mental load associated with the operation of the overall plant. After he completed all the experimental sessions, the operator was asked to evaluate the task allocation decisions by using pair-wise comparisons.

2.4 Experimental sessions

In the experiment, allocation of task was carried out on the basis of three subtasks. Thus, in total, there were 8 experimental sessions, including the fully automated session. Table 1 presents these sessions with the configurations of task allocation.

Table 1 Experimental sessions (A: an automated subtask; M: a manually controlled subtask)

Session number	Task allocation between human and automation				
	<i>Subtask 1</i>		<i>Subtask 2</i>		<i>Subtask 3</i>
	feedstock pump control	overflow valve control	steam pump control	steam heater control	storage distribution
E ₀	A	A	A	A	A
E ₁	A	A	M	M	A
E ₂	M	M	A	A	A
E ₃	A	A	A	A	M
E ₄	A	A	M	M	M
E ₅	M	M	A	A	M
E ₆	M	M	M	M	A
E ₇	M	M	M	M	M

2.5 Subjects

Nine students (male and female) from the Faculty of Mechanical Engineering and Marine Technology of Delft University of Technology participated voluntarily as operators in the experiments. They received a fee for their participation. Before formal sessions started, the subjects performed 7 training sessions of 5 minutes each to learn to control the system and to practice to fill out mental load assessment forms. After completing the training sessions, the subject performed a 10 minute fully manually controlled session. The performance was evaluated by the experimenter based on the percentage of the pasteurized juice. If the score was too low (<80%), the subject was asked to perform more training sessions. Only when a subject was qualified to work as an operator, was he, or she, allowed to start formal sessions. Each subject performed therefore 7 formal sessions of 15 minutes each.

2.6 Performance measurement and mental load assessment

During the experiment, the following main plant variables were recorded: (1) the amount of the input raw juice; (2) the amount of the pasteurized juice; (3) the amount of the wasted juice.

The *system performance* was defined as the percentage of the properly pasteurized juice -- the ratio between the amount of the pasteurized juice and the available input raw juice during the experimental period.

During the experiment, the subjects were asked to assess their mental load for controlling the plant. In the experiment, a subjective rating scale, Rating Scale Mental Effort, RSME (Zijlstra, 1993), was used to assess mental load. Immediately after each session, subjects gave numbers according to the RSME that corresponded best to the mental load for controlling the whole plant.

2.7 Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) was originally developed by Saaty (1980) as a decision making aid, and is a theory of measurement for dealing with quantifiable criteria. The AHP has found its usage in a broad range of areas, including market prediction, architectural design, policy planning, project evaluation, as well as medical decision making (Vargas, 1990). It has also been used as a tool for subjective workload assessment (Vidulich and Tsang, 1987). The AHP has been demonstrated that it provides a means of evaluating multiple design options using weighted ranking scales and can be used as an experimental tool for obtaining subjective preferences in Human Factors studies (Mitta, 1993; Yang and Hansman, 1995). For task allocation between human and automation, Papantonopoulos (1990) has developed an AHP cognitive task allocation model. In comparing the AHP approach to the traditional psychophysical methods for generating measurement scales, the AHP has the following advantages: (1) it possess the ability to readily quantify consistency in human judgments, (2) it has the ability to provide useful empirical results in the event of a small sample of subjects and when the likelihood of obtaining meaningful statistical results may be restricted, and (3) the AHP requires no statistical assumptions regarding the distribution of human judgments (Mitta, 1993).

In brief, the AHP is structured to encompass the basic elements of a decision. As shown in Fig. 2, the objective of a problem is placed at the highest level, then to subobjectives affecting the objective followed by criteria in the next level and so on, from more general to the more specific. For each level of the hierarchy, the AHP breaks up multiple options into a series of paired comparisons which are then recombined to produce an overall weighted ranking. The subjects are required to compare all possible combinations. The results of the comparisons are placed in a pair-wise comparison matrix which is called *judgment matrix*. The rows and columns of the judgment matrix are headed by the options included in the comparison. The principal eigenvector for the matrix is calculated, which gives a weighted ranking scale for each option.

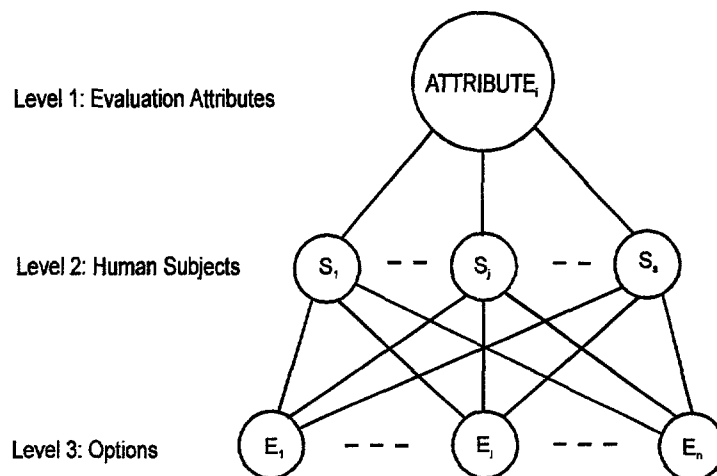


Figure 2 A hierarchic structure

For each level of Fig. 2, data in the AHP is collected using a series of the comparisons between each pair of design options at this level. Saaty (1980) and Mitta (1993) suggest a format which presents the two options to be compared on opposite ends of a 17 slot rating scale. The scale is a measure of dominance of one alternative over the other. It uses five descriptions in a pre-defined order and allows a single space between each one for comparison. The descriptions are “equal”, “weak”, “strong”, “very strong”, and “absolute”. Of course, these terms can be modified to provide easier comprehension for a particular subject group. Fig. 3 shows such a scale used in the present study to compare task allocation configurations in two experimental sessions, E_i and E_j (i

difficult to achieve. The principal eigenvalue of $\mathbf{M}$ is used to measure the consistency. A *Consistency Index* suggested by Saaty (1980) is:

$$C.I. = \frac{\lambda_{\max} - n}{n - 1} \quad (2)$$

where n is the number of options, and $\lambda_{\max}$ is the principal eigenvalue of $\mathbf{M}$. This index indicates that if C.I. is less than 0.1, the judgments are considered consistent (Mitta, 1993).

After all the pair-wise comparisons are completed, an overall ranking scale can be calculated. A frequently used approach to calculate the ranking scale is the eigenvector method. In this method, the ranking scale is produced by calculating the principal eigenvector of $\mathbf{M}$. According to the matrix theory, the principal eigenvector of a matrix, $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_n]^T$, corresponds to *the largest positive eigenvalue* of the matrix, i.e. $\lambda_{\max}$, and is determined by solving Eq. 3:

$$\mathbf{M} \cdot \mathbf{w} = \lambda_{\max} \mathbf{w} \quad (3)$$

In the AHP approach, $\mathbf{w}$ should be typically normalized such that its components sum up to 1:

$$\sum_{i=1}^n w_i = 1 \quad (4)$$

The principal eigenvector $\mathbf{w}$ can be considered as a vector in the space of the n different options where the magnitude of the components in each direction is a measure of the strength of the respective option. The degree of dominance between two options is the ratio of their $\mathbf{w}$ components.

For Subject k , a judgment matrix $\mathbf{M}^k$ can be obtained and so does a principal eigenvector $\mathbf{w}^k$. Thus, a matrix $\mathbf{W}$ containing the rankings of individual subjects can be constructed from the eigenvectors $\mathbf{w}^k$ of all subjects. Supposing that there are s subjects and n options to be compared in the experiment, $\mathbf{W}$ would be an $n \times s$ matrix:

$$\mathbf{W} = [\mathbf{w}^1, \mathbf{w}^2, \dots, \mathbf{w}^s] \quad (5)$$

Moreover, the AHP allows for results of each subject to contribute equivalently or differently depending on his skill, experience, judgment ability, and so on. The experimenter can rate all of the subjects and give a ranking for each subject in the same way as the subjects rate an option by the AHP approach. However, according to Yang and Hansman (1995), in most cases all subjects should be given equal considerations in the final analysis; otherwise, the outcome could be easily be biased toward a certain result.

Suppose that the principal eigenvector which reflects the ranking of subjects is $\mathbf{s} = [s_1 \ s_2 \ \dots \ s_s]^T$ and that all subjects are given the equal consideration, after normalizing, there is

$$\mathbf{s} = \begin{bmatrix} \frac{1}{s} & \frac{1}{s} & \dots & \frac{1}{s} \end{bmatrix}^T \quad (6)$$

where $\mathbf{s}$ is an s vector and s is the number of subjects who participate the evaluation.

The final overall ranking in which the subject contributions are taken into account is:

$$\mathbf{r} = \mathbf{W} \cdot \mathbf{s} \quad (7)$$

$\mathbf{r}$ should be normalized if it does not sum up to 1.0. Thus, the entries of $\mathbf{r} = [r_1, r_2, \dots, r_n]^T$ provide weights for the n options as well as the relative differences between two options. Fig. 4 summaries the AHP procedure.

The degree of dominance between two is represented by the ratio of their weights, r_i and r_j . Based on this ratio, the dominance of one over another one can be converted into a qualitative description in the terms presented in Fig. 3. Table 2 shows the conversion.

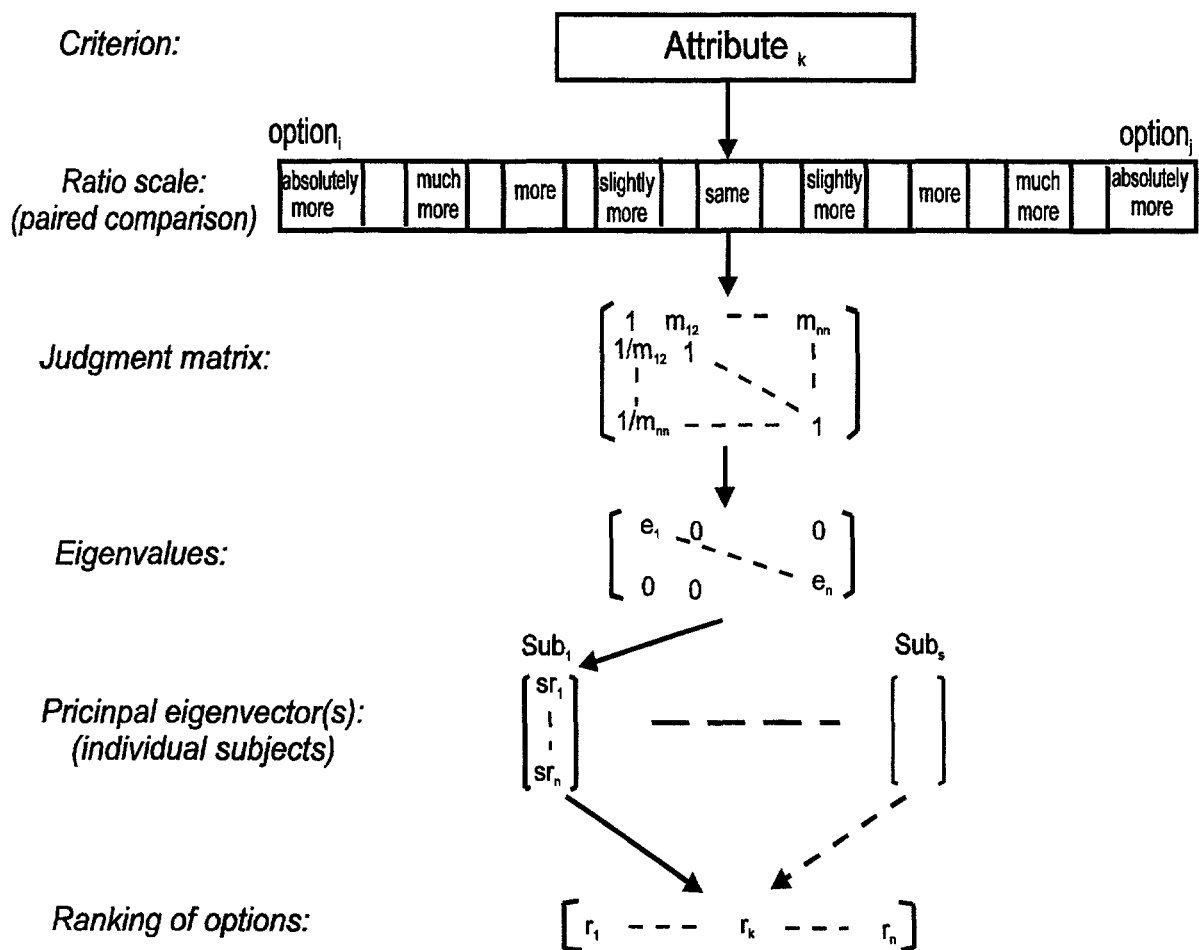


Figure 4 Scheme of the AHP evaluation process

2.8 Evaluation criteria and procedure

The attributes, or criteria, used in the evaluation are listed in Table 3.

Table 2 Conversion of ratio value to qualitative description
(adapted from Yang and Hansman, 1995)

r/r_i	<i>Dominance of option i over option j</i>
1	equal
3	weak dominance
5	strong dominance
7	very strong dominance
≥ 9	absolute dominance

Table 3 Subjective evaluation attributes

<i>Number</i>	<i>Attributes</i>	<i>Descriptions</i>	<i>Questionnaire</i>
1	Easy operation	The overall system is easy to control	Which system is easier to control?
2	Human control confidence	Human is actually in the control of the system, rather than automation	In which system do you feel more that you are in control instead of automation?
3	Comfortable operation	Operators feel uncomfortable because too much automation has been used	With which system do you feel more comfortable to control?
4	Situation awareness	Whether the operator is aware of basic system state variables and understands them	For which system do you know the most about the system status (all necessary variables)?

Attribute 1 was chosen as an indicator for assessing the workload level. We expected that the results should be comparable with the mental load rating based on the RSME. Since the operator often complained, in a highly automated system, that he was not in the control of the system, but automation. The operator lost the confidence in controlling the plant. Attribute 2 was used to obtain the subjective opinion on this issue. Too much automation could make the human operator uncomfortable in operating a system. It was our interest, by using Attribute 3, to have a quantitative estimation on how a high level of automation affected the operator's feeling in the operation. Situation awareness has become an important aspect in the design of a human-machine system (Endsley, 1995). In this experiment, by using Attribute 4, we tried to perform a subjective evaluation on the situation awareness associated with different levels of automation.

The hierarchic structure for the evaluation in this study was similar to Fig. 2, where $n = 7$ was the number of the experimental sessions. Because two of the nine subjects did not have time to complete the evaluation, the number of subjects, s , was seven. Each subject was asked to carry out pair-wise comparisons after all of the sessions were performed.

3. RESULTS

3.1 System performance and mental load

The average percentage of pasteurized juice and the average mental load for each experimental session are presented in Table 4, from which one can see that there is not significant difference between the system performance for different task allocation configurations, but the mental load perceived by the operators appears very differently.

Table 4 System performance and operator mental load

Session number	Manual operation	System performance		Mental load (RSME)	
		Mean	SD	Mean	SD
E ₁	Steam loop	0.92	0.03	35.9	19.1
E ₂	Flow loop	0.88	0.04	44.0	18.9
E ₃	Distribution	0.90	0.00	15.3	7.6
E ₄	Steam + Distribution	0.93	0.03	66.3	11.3
E ₅	Flow + Distribution	0.85	0.05	70.3	9.3
E ₆	Flow + Steam	0.85	0.09	81.0	16.1
E ₇	Fully manual	0.79	0.14	92.3	15.2

3.2 AHP Subjective evaluation of task allocation decisions

For *easy operation*, i.e. *Attribute 1*, the easy level was compared. Matrix W_{A1} shows each subject's evaluation on how *easy* to operate the plant based on a task allocation configuration. A large value means an easier operation. Each column of W_{A1} represents a subject, and each row represents an option, in an order: E₁, E₂, E₃, E₄, E₅, E₆, and E₇.

$$W_{A1} = \begin{bmatrix} 0.3944 & 0.1742 & 0.1434 & 0.3631 & 0.2150 & 0.2335 & 0.1422 \\ 0.2434 & 0.1576 & 0.2207 & 0.2747 & 0.1134 & 0.1312 & 0.1252 \\ 0.1796 & 0.4916 & 0.5084 & 0.2123 & 0.5116 & 0.4650 & 0.5441 \\ 0.0937 & 0.0614 & 0.0443 & 0.0390 & 0.0538 & 0.0521 & 0.0523 \\ 0.0480 & 0.0678 & 0.0443 & 0.0665 & 0.0306 & 0.0766 & 0.0900 \\ 0.0236 & 0.0305 & 0.0262 & 0.0285 & 0.0594 & 0.0268 & 0.0231 \\ 0.0173 & 0.0169 & 0.0126 & 0.0159 & 0.0161 & 0.0148 & 0.0231 \end{bmatrix}$$

When the ranking of subjects is taken into account, we substitute W_{A1} and Eq. 6 with $s = 7$ into Eq. 7 and yields the vector specifying the operational difficulty ranking for task allocation configurations which were represented by the experimental sessions as shown in Table 1:

$$r_{A1} = w_{A1} \cdot s = [0.238 \quad 0.181 \quad 0.416 \quad 0.057 \quad 0.061 \quad 0.031 \quad 0.017]^T$$

For *human control confidence*, i.e. *Attribute 2*, the subjects were asked to compare for which configuration they felt *more* in the control rather than automation. W_{A_2} shows subjects' rating on how *much* they felt that they were in the control. A large value indicates that the operator felt more control.

$$W_{A_2} = \begin{bmatrix} 0.0487 & 0.0519 & 0.1429 & 0.0348 & 0.0452 & 0.0281 & 0.0551 \\ 0.1119 & 0.0663 & 0.1429 & 0.0941 & 0.0481 & 0.0539 & 0.0551 \\ 0.0388 & 0.0241 & 0.1429 & 0.0157 & 0.0222 & 0.0189 & 0.4958 \\ 0.0867 & 0.1157 & 0.1429 & 0.0396 & 0.0734 & 0.1269 & 0.0985 \\ 0.1606 & 0.1105 & 0.1429 & 0.1077 & 0.0967 & 0.1190 & 0.0985 \\ 0.2736 & 0.2440 & 0.1429 & 0.3489 & 0.3398 & 0.2677 & 0.0985 \\ 0.2798 & 0.3875 & 0.1429 & 0.3593 & 0.3747 & 0.3856 & 0.0985 \end{bmatrix}$$

After the rankings of all subjects being considered, the vector specifying the ranking of the human control confidence ranking for task allocation configurations is:

$$r_{A_2} = w_{A_2} \cdot s = [0.058 \quad 0.082 \quad 0.108 \quad 0.098 \quad 0.119 \quad 0.245 \quad 0.290]^T$$

For *comfortable operation*, the subjects were asked to compare for which configuration they felt *more comfortable* in the operation of the plant. W_{A_3} shows subjects' rating on how *comfortable* they felt. A large value indicates that the operator felt more comfortable.

$$W_{A_3} = \begin{bmatrix} 0.2885 & 0.3417 & 0.1748 & 0.0536 & 0.1988 & 0.4103 & 0.3568 \\ 0.1184 & 0.2138 & 0.1748 & 0.1724 & 0.1224 & 0.2689 & 0.2387 \\ 0.0217 & 0.0809 & 0.5041 & 0.0166 & 0.4427 & 0.0479 & 0.0190 \\ 0.3044 & 0.1426 & 0.0546 & 0.0543 & 0.1217 & 0.1502 & 0.1123 \\ 0.1251 & 0.1232 & 0.0546 & 0.1427 & 0.0356 & 0.0675 & 0.0871 \\ 0.0729 & 0.0571 & 0.0247 & 0.2906 & 0.0644 & 0.0375 & 0.1098 \\ 0.0691 & 0.0408 & 0.0124 & 0.2698 & 0.0145 & 0.0177 & 0.0762 \end{bmatrix}$$

The vector specifying the ranking of the effect of automation on human by combining all subjects' judgments is as follows:

$$r_{A_3} = w_{A_3} \cdot s = [0.261 \quad 0.187 \quad 0.162 \quad 0.134 \quad 0.091 \quad 0.094 \quad 0.072]^T$$

For *situation awareness* rating, the subjects were asked to compare for which configuration they knew *more about* the plant status. W_{A_4} shows each subject's rating. A large value indicates that the operator knew more variables and had more situation awareness.

$$W_{A_4} = \begin{bmatrix} 0.1405 & 0.0753 & 0.0643 & 0.1040 & 0.0651 & 0.0988 & 0.0435 \\ 0.0842 & 0.0683 & 0.0432 & 0.1336 & 0.0559 & 0.1179 & 0.0435 \\ 0.0182 & 0.0245 & 0.0141 & 0.0186 & 0.0216 & 0.0235 & 0.0435 \\ 0.0949 & 0.0953 & 0.0559 & 0.0562 & 0.0858 & 0.1621 & 0.0435 \\ 0.1065 & 0.0769 & 0.0954 & 0.2084 & 0.0984 & 0.1212 & 0.0435 \\ 0.2352 & 0.3248 & 0.3635 & 0.2396 & 0.2804 & 0.2009 & 0.3913 \\ 0.3205 & 0.3349 & 0.3635 & 0.2396 & 0.3929 & 0.2756 & 0.3913 \end{bmatrix}$$

The vector specifying the ranking of the situation awareness from a combination of all subjects' judgments is:

$$\mathbf{r}_{A_4} = \mathbf{w}_{A_4} \cdot \mathbf{s} = [0.085 \quad 0.078 \quad 0.023 \quad 0.085 \quad 0.107 \quad 0.291 \quad 0.331]^T$$

4. DISCUSSION

4.1 AHP evaluation and mental load

Easy operation and mental load. The relationship of the easy level in the operation (AHP) and the measured mental load (RSME) appears linear as shown in Fig. 5. The correlation coefficient between two was -0.963, $\rho < 0.001$.

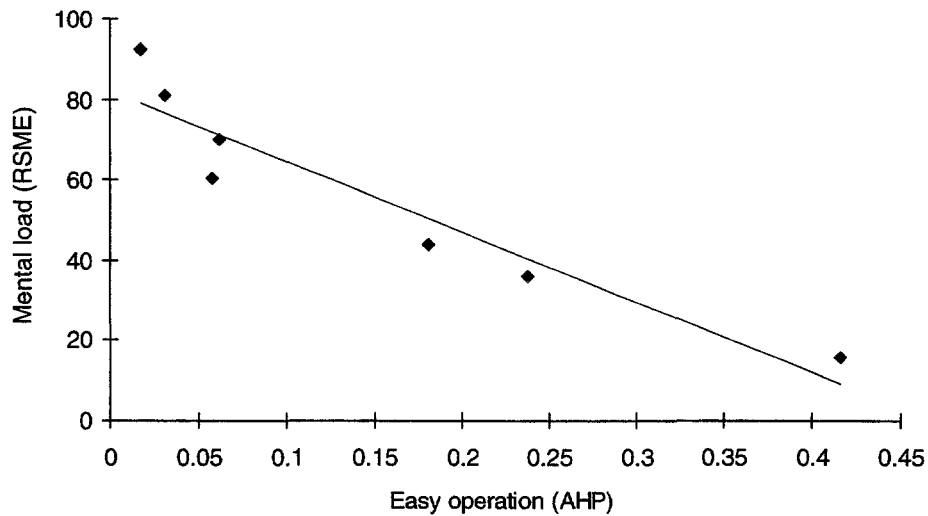


Figure 5 Relationship between easy operation and mental load

Fig. 5 indicates that the easy operation level based on the AHP evaluation agreed quite well with the subjective assessment of mental load. The easiest controlled system had a lowest mental load. As the easy level increased, the mental load decreased accordingly. The only exception in the order of the two measures was in the cases of steam loop plus distribution control and feed loop plus distribution control. Both the mental load and the AHP values in these two configurations were very close. This means that the two configurations were at the same level. The results presented Fig. 5 have further demonstrated that the AHP approach could be used as a method for the mental load assessment as suggested by Vidulich and Tsang (1987).

Given vector $\mathbf{r}_{A1}$ in Section 3.2, one can find that E_3 , where only distribution was manually controlled, was ranked as the easiest configuration to operate, and E_7 , where all subtasks were manually controlled, was ranked the most difficult to control by a factor of at least (with respect to E_6) 1.82. $\mathbf{r}_{A1}$ also tells us that subjects perceived E_3 as being easier to control than E_1 , E_2 , E_4 , E_5 , E_6 , and E_7 by factors 1.75, 2.23, 6.82, 7.30, 13.42, and 24.47, respectively. The degree of the dominance of E_i over E_j and qualitative interpretations can be found in Table 2.

Human control confidence and comfortable operation. These two criteria aimed at evaluating how the level of automation affected human operator's feeling during operation. Vector $\mathbf{r}_{A2}$ indicates that E_7 was ranked as the most human control configuration, and E_1 , where the steam loop was manually controlled, was ranked the least to control by the operators by a factor of at least (with respect to E_2) 1.41. Vector $\mathbf{r}_{A2}$ also indicates that subjects perceived E_7 as having more human control than E_1 , E_2 , E_3 , E_4 , E_5 , and E_6 by factors 5.0, 3.54, 2.69, 2.96, 2.44, and 1.18, respectively.

Vector r_{A3} indicates that E_1 was ranked as the most comfortable configuration to control because of automation, and E_7 was ranked the least comfortable to control by a factor of at least (with respect to E_5) 1.26. Vector r_{A3} also indicates that subjects perceived E_1 as being more comfortable than E_2 , E_3 , E_4 , E_5 , E_6 and E_7 by factors 1.40, 1.61, 1.95, 2.87, 2.78, and 10.07, respectively.

Comparing vector r_{A2} with vector r_{A3} , one can find that the ranking of the human control confidence perceived by the subjects was opposite to the ranking for the operational comfort. E_1 was ranked as the most comfortable configuration, but was the least in the human control confidence. E_7 was ranked as the most in the human control confidence, but the least comfortable configuration. This implies that a compromise has to be made between these two aspects. An interesting point is that E_3 was the easiest to control, but was not the most comfortable configuration. This means that a configuration with a low workload does not necessarily make the operator feel comfortable in the operation.

Situation awareness. Task allocation between human and automation will directly affect operator's awareness of the system situation. For the subjective evaluation of the situation awareness in this study, vector r_{A4} indicates that E_7 had the highest situation awareness, and E_3 had the lowest situation awareness by a factor of at least (with respect to E_2) 3.39. Vector r_{A4} also indicates that subjects perceived that for E_7 , the situation awareness was higher than for E_1 , E_2 , E_3 , E_4 , E_5 , and E_6 by factors 3.89, 4.24, 14.39, 3.89, 3.09, and 1.14, respectively. This is understandable because in E_7 the operators performed all subtasks and monitored all necessary system variables.

4.2 Synthesis of the subjective evaluation

All criteria, or attributes ($A_1, \dots, A_4$), used in the evaluation could also be ranked based on the AHP approach so that a final ranking of all task allocation decisions may be obtained. For four attributes used in the study, a judgment matrix given by the experimenter may be as follows:

$$M_C = \begin{matrix} & \begin{matrix} A_1 & A_2 & A_3 & A_4 \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{matrix} & \begin{bmatrix} 1 & 2 & 3 & 1 \\ 1/2 & 1 & 2 & 1/2 \\ 1/3 & 1/2 & 1 & 1/3 \\ 1 & 2 & 3 & 1 \end{bmatrix} \end{matrix}$$

The principal eigenvector of M_C is:

$$r_c = [0.351 \ 0.189 \ 0.109 \ 0.351]^T$$

The consistent index of the comparison is C.I. = 0.0035 which indicates the paired comparison is consistent.

The rankings of task allocation configurations in four attributes can construct a new matrix with the following form:

$$R_{TA} = \begin{bmatrix} r_{A1} & r_{A2} & r_{A3} & r_{A4} \end{bmatrix}$$

Thus, the final rank specifying all task allocation decisions ($E_1, \dots, E_7$) by combining four criteria is:

$$r_F = R_{TA} \cdot r_c = [0.153 \ 0.127 \ 0.192 \ 0.083 \ 0.106 \ 0.170 \ 0.185]^T$$

From vector $\mathbf{r}_F$, one can find that E_3 with the distribution subtask being manually controlled, and E_7 with all subtasks being allocated to the operator have a high ranking. However, we must take into account all aspects in the evaluation before making a decision. For example, E_3 was ranked an extremely high score at Attribute 1, which effects a high score in the final ranking. However, at Attribute 4, E_3 got the lowest ranking. E_7 was the most difficult configuration to operate, i.e. it demanded a very high mental load, although it got the highest ranking at Attribute 2 and 4. Both E_3 and E_7 had a ranking at two extreme ends of Attributes 1 and 4 in an opposite order, and they could not be considered as an optimum task allocation. In the final ranking E_1 and E_6 had a high ranking, but they did not have an extremely high or low ranking at any attributes. Thus, they could be considered as an optimum task allocation in the control of the pasteurization plant. If the system performance as well as the mental load as shown in Table 4 were taken into account, E_1 might be the best task allocation decision.

4.3 Consistency of the judgments

One of the advantages of the AHP is its capability to calculate consistency in human judgments. The consistency indices for all the subjects in the evaluation were calculated by Eq. 2, and are listed in Table 5.

Table 5 Consistency indices

Attributes	Consistent Index (C.I.)							
	Subj. 1	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	MEAN
Easy operation	0.318	0.228	0.451	0.233	0.266	0.297	0.257	0.293
Human control confidence	0.135	0.101	1.167	0.226	0.201	0.206	0.525	0.366
Comfortable operation	0.148	0.059	0.475	0.123	0.412	0.223	0.046	0.212
Situation awareness	0.171	0.070	0.293	0.008	0.075	0.241	0.000	0.123

From the consistency indices, we can conclude that the most consistent comparison took place in the evaluation of the situation awareness (an average C.I. approaching 0.1). In other evaluations, the average C.I. was larger than 0.2 which implies that the consistency in the paired comparisons was lower. The possible reasons are: (1) the comparison capabilities of some subjects were poor, (2) the questionnaire were not completely understood by the subjects, and (3) there were too many options, i.e. the number of experimental sessions. After all of the sessions completed, the subjects could not accurately remember what they had experienced.

The consistency indices of each subject are very helpful in evaluating each subject's ability to judge. So, a rank, s , for all subjects could be obtained. The judgment of the subjects that were least consistent could be eliminated by the individual preference. Thus, although the data from all subjects are included in the final ranking, the judgments from the subjects with the least judgment consistency will have less effect than the subjects with the high consistency. However, in the

present study, each subject was given an equal weight to avoid biased judgments although some of the subjects had demonstrated less capabilities of providing a sound judgment.

4.4 Problems in the AHP application

The advantages and limitations of the AHP approach have been addressed by Mitta (1993). Here, we only address the problems met in the application of the AHP for subjective task allocation evaluation. In our application, we met with three problems. The first one was that the number of pair-wise comparisons increased greatly as the number of options and the number of the attributes. This was a very heavy burden to the subjects who felt very boring during paired comparisons. This made it difficult for the subjects to have a reasonable consistency in their judgments.

The second problem deals with the combination of the individual rankings. When a final ranking for all options was obtained by combining the ranking from each attribute, an extremely high ranking of an option in one of the attributes might play a significant role, even though the weight for this attribute was not so large. For example, E_3 was ranked as the easiest configuration to operate, and, according to Table 2, it dominated strongly over most of the other configurations. Therefore, at the final ranking, E_3 ranked very high even though other attributes had been considered in a weight approach. When we took into account other attributes, E_3 could not be an optimum task allocation configuration. From this example, we could conclude that the AHP ranking can only be taken as a reference, and other aspects have to be taken into account when making a final decision.

The third problem was that for task allocation evaluation, the consistency of the subjective judgments were difficult to reach. When the number of the options is large, or the attributes are not simply defined and not clearly expressed, it will be more difficult to have a reasonable consistency. Thus, to carry out the task allocation evaluation by using the AHP, it is better to reduce the options as many as possible, and to make the evaluation criteria be easily understood.

5. CONCLUSION

The paper presents an application of the Analytic Hierarchy Process in the subjective evaluation of task allocation decisions between human and automation. From the experimental results, the following conclusions can be drawn:

- A subjective ranking of different task allocation decisions can be obtained using the AHP approach instead of a statistical approach; but,
- The AHP ranking can only be taken as a reference in the evaluation of task allocation;
- A configuration which is easy to control does not necessarily make the human operator comfortable;
- The AHP can be used to measure the relative level of human workload.

REFERENCES:

Grote, G., S. Weik, T. Wafler and M. Zolch (1995). Criteria for the complementary allocation of functions in automated work systems and their use in simultaneous engineering projects. *International Journal of Industrial Ergonomics*, 16, pp.367-382.

- Endsley, M.R. (1995). Towards a new paradigm for automation: Designing for situation awareness. In: *Preprints of the 6th IFAC/IFIP /IFORS /IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems*, MIT, MA, USA, pp.421-425.
- Huey, B. M. (1989). *The effect of function allocation schemes on operator performance in supervisory control systems*. Unpublished Doctoral Dissertation, George Mason University, USA.
- Lee, J. and N. Moray (1992). Trust and the Allocation of Function in the Control of Automatic Systems. *Ergonomics*, **35**, pp.1243-1270.
- Mitta, D.A. (1993). An application of the analytic hierarchy process: a rank-ordering of computer interfaces. *Human Factors*, **35**(1), pp. 141-157.
- Muir, B.M. (1989). *Operator's trust in and use of automatic controllers in supervisory process control task*. Unpublished Doctoral Dissertation, University of Toronto, Canada.
- Papantonopoulos, S.A. (1990). *A design model for cognitive task allocation*. Unpublished Doctoral Dissertation, Purdue University, IN, USA.
- Rouse, W.B. and W.J. Cody (1986). Function allocation in manned system design. In: *Proc. of the IEEE International Conference on System, Man, and Cybernetics*, pp.1600-1606. IEEE, New York.
- Saaty, T.L. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill Book Company.
- Vargas, L.G. (1990). An overview of the analytic hierarchy process and its application. *European Journal of Operational Research*, **48**, pp.2-8.
- Vidulich, M.A. and P.S. Tsang (1987). Absolute magnitude estimation and relative judgment approaches to subjective workload assessment. In: *Proc. of the Human Factors Society 31st Annual Meeting*, pp.1057-1061.
- Yang, L.C. and R. J. Hansman, Jr. (1995). Application of the analytic hierarchy process for making subjective comparisons between multiple automation/display options. In: *Preprints of the 6th IFAC/IFIP /IFORS /IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems*, MIT, MA, USA, pp.555-559.
- Zijlstra, F.R.H. (1993). *Efficiency in Work Behavior*. Ph.D. thesis, Delft University of Technology, Delft University Press, Delft, The Netherlands.