

AN INTEGRATED MENTAL WORKLOAD ASSESSMENT APPROACH BASED ON NASA-TLX AND SMAA-2: A CASE STUDY

Elif KILIÇ DELİCE^{1*}, Gülin Feryal CAN²

^{1*}Atatürk Üniversitesi, Mühendislik Fakültesi, ,
Endüstri Mühendisliği Bölümü, Erzurum

e-posta : elif.kdelice@atauni.edu.tr , ORCID No : orcid.org/0000-0002-3051-0496

²Başkent Üniversitesi, Mühendislik Mimarlık Fakültesi,
Endüstri Mühendisliği Bölümü, Ankara

e-posta : gfcan@baskent.edu.tr , ORCID No : orcid.org/0000-0002-7275-2012

Makale Geliş : 26.01.2018
Makale Kabul : 27.03.2018
Araştırma Makalesi

Article Received : 26.01.2018
Article Accepted: 27.03.2018
Research Article

NASA TLX ve SMAA-2 Temelinde Entegre Bir Zihinsel İş Yükü Değerlendirme Yaklaşımı: Bir Örnek Olay Çalışması

An Integrated Mental Workload Assessment Approach Based on NASA-TLX and SMAA-2: A Case Study

Öz

Zihinsel iş yükü (ZİY), özellikle karmaşık sistemlerdeki çalışan performansını etkileyen önemli bir ergonomik faktör olarak uzun zamandan beri dikkate alınmaktadır. ZİY araştırmaları genellikle, altı farklı ZİY kriterinin ağırlıklı toplamını içeren ve klasik bir ZİY değerlendirme yaklaşımı olan Ulusal Havacılık ve Uzay Dairesi İş Yükü İndeksi (NASA-TLX) yönteminin farklı uygulamalarına odaklanmıştır. Bununla birlikte, karar verici (KV) olarak çalışanlar tercihlerini, ZİY boyutlarına ait ağırlıkların ve puanların belirlenmesi için uygulanan skaladaki kesin değerleri kullanarak belirtmeyebilir veya belirtmek istemeyebilirler. Bu tür belirsizlik içeren bilgiler, etkin bir karar destek aracı olan Stokastik Çok Kriterli Kabul Edilebilirlik Analizi-2 (SMAA-2) kullanılarak stokastik bir şekilde modellenabilir. Bu kapsamda çalışmada, ZİY'nü analiz edebilmek için NASA TLX ve SMAA-2 temelinde entegre bir yaklaşım önerilmektedir. Önerilen yaklaşıma ilişkin gerçek hayat uygulaması ise kutu harf imalat prosesinde gerçekleştirilmiştir. Bu uygulamada, üç farklı problem senaryosu için NASA-TLX&SMAA-2 yöntemi kullanılarak alt prosesler ZİY açısından sıralanmıştır. Söz konusu üç problem senaryosunda da sıralama sonuçları aynı çıkmıştır. Bu çalışma sonucunda, NASA-TLX&SMAA-2 yaklaşımının NASA-TLX metoduna göre daha açık ve belirgin sonuçlar sağladığı belirlenmiştir.

Anahtar Kelimeler: Zihinsel iş yükü, NASA- TLX, SMAA-2, belirsiz bilgi

Abstract

Mental workload (MWL) has been considered as an important ergonomic factor for a long time influencing worker performance within complex systems. MWL research generally tend to focus on different applications of the National Aeronautics and Space Administration-Task Load index (NASA-TLX) method, which is a classical tool for MWL assessment based on weighted sum of ratings of six different MWL criteria. However, workers as decision makers (DMs) cannot or do not want to specify their preferences with the exact values due to the scale implemented in determining MWL dimensions' weights and their ratings. These uncertain data can be modeled in a stochastic manner by using the Stochastic Multi-Criteria Acceptability Analysis-2 (SMAA-2) which is an effective decision support tool. In this context, an integrated approach based on NASA-TLX and SMAA-2 for MWL assessment is proposed in the study. A real life application of the proposed approach is conducted in the box letter manufacturing process. In this application, three problem scenarios are discussed by using NASA-TLX&SMAA-2 to rank six sub processes respect to MWL levels. Rankings obtained from the NASA-TLX&SMAA-2 method for three scenarios are the same. This paper found that the NASA-TLX&SMAA-2 approach provides more precise and distinctive results compared to the NASA-TLX method.

Keywords: Mental workload, NASA-TLX, SMAA-2, imprecise data.

1. Introduction

Sorumlu Yazar; Tel: 0. 442.231 00 00 / 6009

There is a tendency to examine mental workload (MWL). In particular, control and audit activities of

workers have become more important in accordance with the development of new technologies (Gülkaç, 2013). Consequently, it becomes obligatory to evaluate MWL as much as physical workload in the studies. The MWL is particularly important in the research and development of human-machine interfaces in order to achieve the desired levels of satisfaction, comfort, safety and efficiency in the workplace that are the main targets of ergonomics (Rubio et al., 2004). MWL is one of the most commonly used concepts in the research and practice of ergonomics (Parasuraman, and Hancock, 2001; Flemisch, and Onken, 2002; Loft et al., 2007; Vidulich, and Tsang, 2007; Wickens, 2008; Mouzé-Amady et al., 2013; Young et al., 2015).

Although there are several different methods used to evaluate MWL, NASA-TLX is widely employed in the field. This classical tool for MWL assessment provides an overall workload score based on a weighted sum of ratings from six subscales including mental demand, physical demand, temporal demand, performance level, effort level, and frustration level. In addition, evaluations made using this method can be biased or misleading due to the participant's psychological state, readiness level, and awareness of the overall knowledge require for the method. In the NASA-TLX application, the ratings cannot be assigned precisely by workers as decision makers (DMs) due to the structures of these subscales. However, each rating is weighted by its own coefficient obtained from the pairwise weighting technique (PWT), that is complex and time consuming. It raises some methodological and practical problems, especially in real life applications. Therefore, obtaining correct overall workloads is a problematic issue in the NASA-TLX method. On the other hand, the MWL assessment is a Multi Criteria Decision Making (MCDM) problem as it requires the consideration of multiple criteria and alternative solutions. Moreover, the presence of multiple DMs leads to variety in evaluations. Due to these uncertainties, MWL assessment is discussed as a stochastic MCDM problem in this study. An integrated approach based on NASA-TLX and Stochastic Multi Criteria Acceptability Analysis-2 (SMAA-2) method is suggested to determine the MWL. The SMAA-2 method is an effective decision support tool that allows us to solve MCDM problems with stochastic data for cases where there are uncertain, invalid or missing information (weights, parameter values e.g.).

The proposed NASA-TLX&SMAA-2 approach is used in the box letter manufacturing process in Turkey. This manufacturing field is one of the big sectors where MWL effects workers. As the main focus of the workers is related to the formation of systems and technologies, it requires more mental performance compared to physical performance. For instance, the main purpose of box letter manufacturing is to produce different kinds of letters for different famous brands such as automotive brands.

Workers in this manufacturing process must make correct decisions to situate the box letter on suitable place of product or he may cause a hole or a mark on product. On the other hand, the performance of other process of the manufacturing is equally important since it also directly or indirectly impacts product quality. If the box letter is assembled or pasted on a wrong place, this part of product cannot be used and it is called as waste. In this context, MWL assessment should be performed for this manufacturing process.

This study contributes to the related literature in two main ways. First of all, it combines NASA-TLX with SMAA-2 for the first time and discusses the MWL as a stochastic MCDM problem. Secondly, up to best knowledges, the MWL for box letter manufacturing process has not been addressed using MCDM methods in the literature so far.

The paper is organized as follows; in the second section, literature review is presented. The NASA-TLX and SMAA-2 methods are explained in the third section. The theoretical problems of NASA-TLX and the proposed NASA-TLX&SMAA-2 approach are introduced in the fourth section. In the fifth section, a case study is presented while comparative analysis is given in the sixth section. Conclusion and discussions are presented in the seventh section.

2. Literature Review

NASA TLX Method has been used to determine the MWL in various sectors. Lee and Liu (2003) measured the MWL of the pilots during the flight using physiological and multi-dimensional subjective parameters. Matthews et al. (2003) analyzed the cumulative MWL imposed by three different types of cellular phones on drivers using NASA TLX. NASA TLX has been used to determine MWL in the treatment and surgical methods in medicine sector (Stefanidis, et al., 2010; Zhen et al., 2012; Mazur et al., 2013; Durantin et al., 2014; Ruiz-Rabelo et al., 2015; Colligan et al., 2015). Relation between the MWL and task complexity degrees of fast train conductors were studied by Park et al. (2009). Filtness and Rudin-Brown (2012) used NASA TLX for the assessment of driving performance, Kuehn et al. (2013) examined MWL arising from utilization of touch screens from frontal and parallel positions, for motor control disorder people. Akyeampong et al. (2014) evaluated MWL emerged in newly proposed Human-Machine Interface (HMI) design concepts for improving the ergonomics of hydraulic excavators with NASA TLX.

Additionally, SMAA-2 was performed in the different application fields by different researchers (Hokkanen et al., 2000; Lahdelma, and Salminen, 2001; Lahdelma, and Salminen, 2002; Kangas et al., 2003; Kangas et al., 2005; Kangas et al., 2006; Lahdelma, and Salminen, 2006; Tervonen et al., 2008; Lahdelma et al., 2009; Tervonen

et al., 2009; Tervonen et al., 2011; Aertsen et al., 2011; Sporcic et al., 2011; Tervonen, 2012; Kadzinski, and Tervonen, 2013; Angilella et al., 2015). Song et al. (2017) aimed to measure national energy performance by implementin SMAA-2 employing employ the preferences among energy trilemma. Zhu et al. (2017) used SMAA-2 to select reservoir flood control operation. Zhou et al. (2017) utilized extended gray numbers, integrated with discrete gray numbers and interval gray numbers are used to express the uncertainty of stochastic MCDM problems. SMAA-2 and ELECTRE III are combined to solve stochastic MCDM problems with uncertain weight information. First, the outranking relations on interval gray numbers and interval gray numbers are defined. Then, a SMAA-ELECTRE model for dealing with gray stochastic MCDM problems is constructed. Aydoğan and Özmen (2017) proposed a new method based on SMAA-VIKOR, was proposed for stochastic multi-criteria decision making (MCDM) problems and the effectiveness of the method was shown by comparing literature data and a case study. Angilella et al. (2017) proposed SMAA-Choquet integration to compare the performances of different sailboats in regattas. Akgül et al. (2017) carried out SMAA-MOORA to determine the optimum nozzle pressure and delivery speed of Murata Vortex Spinner (MVS) machine for cotton yarn.

The aims of SMAA-2 implementation in the studies mentioned above are to identify the priorities of alternatives in terms of criteria and to determine impact levels of criteria on the selection of alternatives in different ranks and to obtain the rankings of alternatives according to the probability information (Lahdelma and Salminen, 2001). However, decision problems in real life, have a stochastic nature as; preference information obtained from DMs and criteria values are not precise, complete and correct, preference of more than one DMs lead to variety in evaluation (Lahdelma and Salminen 2001; Lahdelma and Salminen 2009). SMAA-2 method is able to implement with this kinds of information. It is able to provide the most probable preference ranking of alternatives and give the most preferred alternative, as it has a flexible structure.

As seen from the literature, NASA TLX was only used for MWL analysis of different populations and different work areas. It has not been advanced in terms of its application procedure. However, it has various shortcomings related to implementation in real life (Mouzé-Amady et al., 2013). Additionally, SMAA-2 was also used for different decision areas successfully. It has not been used for MWL analysis. However, it can provide various benefits for overcoming the shortcomings of NASA TLX. SMAA-2 can help managers to understand the differences between processes in terms of MWL and it can present the impact of each MWL sub scales on MWL. In this way, managers can identify easily sub scale/scales that should be improved based on

probabilities and impacts data.

3. Methods

3.1. NASA TLX

NASA-TLX is one of the most commonly used instruments for assessing MWL (Hart and Staveland, 1988). Since its presentation in 1988, it has been tested many times and frequently used in human performance studies and also considered to be a reliable measure of MWL (Jorgensen and Garde; 1999; Moroney et al., 1995). It has been translated into many languages, administered verbally, and modified in various ways. It has also been evaluated many times in terms of reliability, sensitivity, utility, and assessed and compared to other methods of measuring workload. Additionally, NASA-TLX is stated to be more reliable than two other well-known workload evaluation methods which are Subjective Workload Assessment Technique (SWAT) and Workload Profile (WP) method (Rubio et al., 2004).

The NASA-TLX is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of rating on six subscales as shown in Table 1 (NASA Ames Research Center, 2018.) First three subscales (mental demand, physical demand, and temporal demand) show the expectations of workers from the work. Other subscales (performance, effort, frustration) are self-evaluations of workers. These six subscales are formed criteria in SMAA-2. NASA-TLX evaluates the general workload at three steps. First step is the rating step where participants are asked to rate the effect of six sub-criteria on a scale from "very low" to "very high". After this, values based on points between 0-100 are obtained. The second step is the weighting step, in which participants are asked 15 pairwise comparisons to identify which criterion among the six is the most significant with regard to the work's nature and load. After the responses are obtained, responses are subjected to PWT, which reveals the importance weight of each criterion. In the third step, values found from the first two steps combined into a single Task Load Index (TLX) using Eq. (1). (Rubio et al., 2004; Mouzé-Amady et al., 2013):

$$TLX = \sum_{i=1}^N S_i W_i \quad (1)$$

where,

S_i indicates rating of i th criterion, ($i = MD, PD, TD, PL, EL, FL$), W_i denotes the weight of the i th criterion. W_i is computed as in Eq.(2).

$$W_i = \frac{f_i}{\sum_{i=1}^N f_i} \quad (2)$$

Where,

f_i presents the frequency of i th criterion providing that $\sum_{i=1}^N f_i = 15$.

Table 1. Rating Scale Definitions and Endpoints of The NASA-TLX

Factors	Endpoints	Definitions
Mental Demand (MD)	Low/High	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical Demand (PD)	Low/High	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand (TD)	Low/High	How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance (PL)	Good/Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort (EL)	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration (FL)	Low/High	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

3.2. Stochastic multi-criteria acceptability analysis SMAA-2

SMAA-2 is suggested by Lahdelma and Salminen (Lahdelma and Salminen, 2001). SMAA-2 is a multi-criteria decision support technique for problems including imprecise, uncertain information. Hence, it is suitable for the situations where criteria measurements and weights are not precisely known which is common especially in ranking and choosing problems. The inverse weight space analysis is applied in SMAA-2 to evaluate the ranking possibility of the alternatives.

The decision problem in SMAA-2 consists of “ m ” alternative sets $\{x_1, x_2, x_3, \dots, x_m\}$ evaluated in terms of “ n ”

criteria as $\{g_1, g_2, g_3, \dots, g_n\}$. $g_j(x_i)$ represents the evaluation of x_i alternative according to g_j criterion. The preference of DM is stated by real-valued utility function “ $u(x_i, w)$ ”. Uncertain or imprecise criteria are represented by the combined probability distributed intensity function “ $f(\xi)$ ” and the density function “ ξ_{ij} ” in the space $X \subseteq R^{m \times n}$. Similarly, the DM’s unknown or partially known preferences are expressed by a weight distribution with joint density function $f(w)$ in the feasible weight space W . Total lack of preference information is shown by a uniform weight distribution in W , i.e. $f(w) = 1/vol(W)$. The feasible weight space is defined in Eq.(3). Note that the weights are normalized and have non-negative values.

$$W = \{w \in R^n \mid w \geq 0 \text{ and } \sum_{j=1}^n w_j = 1\} \tag{3}$$

The utility functions are used to map stochastic criteria and weight distributions $u(\xi_i, w)$ into utilities. Based on value distributions, the ranking of each alternative is stated as an integer from the best (1) to the worst (m) by using the ranking function given in Eq. (4).

$$rank(i, \xi, w) = 1 + \sum_{k=1}^m \rho(u(\xi_k, w) > u(\xi_i, w)) \tag{4}$$

Here, $\rho(true) = 1$ and $\rho(false) = 0$. Then, SMAA-2 is based on analyzing the stochastic sets of appropriate rank weights as shown in Eq.(5) (Lahdelma and Salminen, 2001).

$$W_i^r(\xi) = \{w \in W : rank(i, \xi, w) = r\} \tag{5}$$

Weight $w \in W_i^r(\xi)$ represents the values of alternative x_i having rank r .

Additionally, the SMAA-2 introduces three measures for each alternative i and rank r as the rank acceptability index (RankAcc), the central weight vector (CW) and the confidence factor (CF). The RankAcc b_i^r shows the possibility of alternative x_i to be in rank r . Alternatives which have the highest acceptability for the best ranks are considered as the best alternatives. RankAcc takes values in [0,1]. The “0” value shows that the alternative will never provide the given rank while value “1” shows that the given rank will always be provided by the alternative whatever its weight is. With the multi-dimensional integral operations made on the distribution of criteria and appropriate rank weights, RankAcc index is calculated as follows:

$$b_i^r = \int_{w \in W_i^r(\xi)} f_w(w) dw d\xi_r \quad (6)$$

The second measure is CW provided by SMAA-2. It is defined for each alternative and it shows the rank preferences probability of each alternative. The CW represents the DM's typical preferences for the alternative. CW factors also determine the probability of the alternative to be chosen by the decision makers. These probabilities are calculated by the help of a multi-dimensional integral of the criteria and weight distributions as illustrated in Eq.(7).

$$w_i^c = \int_{\xi \in X} f_x(\xi) \int_{w \in W_i^1(\xi)} f_w(w) dw d\xi / a_i \quad (7)$$

The last descriptive measure of SMAA-2 is the CF. This factor is defined as the probability of an alternative to achieve the first rank provided that the CW is chosen. CF states the reliability of an alternative while being available in its rank. It is described as in Eq. (8).

$$p_i^c = \int_{\forall k=1, \dots, m} \xi \in X: u(\xi_i, w_i^c) \geq u(\xi_k, w_k^c) f_x(\xi) d\xi \quad (8)$$

Multi-dimensional integrals presented in SMAA-2 formulations can be calculated by numerical techniques such as Monte Carlo Simulation. In order to facilitate the modeling of decision problems with SMAA-2, JSMAA program, which is available in the open source as JSMAA 1.0.3 is used in this study (Tervonen, 2012).

4. Methodological and Practical Problems in NASA-TLX and The Proposed NASA-TLX&SMAA-2 Approach

NASA-TLX depends on the participant's perception, which may be affected by external factors. For example, the analyst might not explain the methods properly to participants. Similarly, the instant psychological states of the participants might mislead the results since they are overloaded or their level of stress is high at the time of NASA-TLX application. Factors such as a bad social life or low levels of harmony between the company and participants may lead to inaccurate results in the study. Moreover, two problems are encountered in the NASA-TLX method. One of them occurs in the process of obtaining a score value between 1 and 100 after the participants mark the effect of six criteria on a scale from "very low" to "very high" in the first part. Difficulty arises from the scale having a very wide range between 'very low' and 'very high' and not having intervals in between them. They may not be sure which exact point corresponds their mark. Weighting of the six criteria poses another problem. As they indicate the importance

rates of criteria with regard to nature and load of work, the participants compare the six criteria to each other. PWT is used that can easily be calculated but this raises some methodological and practical problems, especially in real work environments (Mouzé-Amady et al., 2013). Consequently, it is quite difficult to identify the importance of all criteria between each other. This study aims to eliminate the above mentioned problems related to the traditional NASA-TLX implementation by means of the SMAA-2 method, which allows definition of criteria values and weights in the form of interval, distribution, or data loss, and to lower the negative effects of external factors. In MCDM problems, in some situations, the DMs may not prefer to express their preferences explicitly or implicitly. In addition, the performance values of alternatives for criteria may be imprecise, uncertain. For these reasons, exact parameter values and weights cannot be obtained. SMAA-2 can be applied to decision making in these situations.

5. An Application for Box Letter Manufacturing Process

The proposed MWL assessment approach is utilized in box letter manufacturing process in an advertisement company. The box letter manufacturing process involves the installation of emblems, especially with the use of materials determined by the companies on vehicles of major automotive companies. The emblems are brand logos and the box letters are attached or glued to different locations of vehicles such as the steering wheel, the bonnet or the luggage compartment. Accordingly, the presence of any fault in the attachment or glue processes leads to deformation and a bad appearance in the relevant part of the vehicle. For this reason, it is necessary to fit the logo exactly and to fit it without any positional deviation. All this process requires high attention. In this context, this process has a characteristic that MWL can affect workers.

This manufacturing process is completed with six different sub processes. Each sub process is implemented by two workers and there are six sub processes in the box letter manufacturing process. MWL of the 12 workers in these six sub processes are evaluated and compared in the context of the study. All workers are graduated from vocational high school. Workers working in the first sub-process are 38 and 42 years old, workers in the second sub-process are 27 and 42 years old, workers in the third sub-process are 33 and 35 years old, workers in the fourth sub-process are 41 and 45 years old, workers in the sixth sub-process are 46 and 40 years old.

At first, a decision maker team (DMT) is created from all workers in six sub processes as Plywood Cutting (PC), Chrome Lettering (CL), Plywood Varnish (PV), Plywood

Chamfering (PCh), Acrylic Cutting (AC) and Plywood Emery (PE). These sub processes are formed alternatives evaluated in terms of MWL in this study and six sub scales of NASA TLX are formed criteria. Later on, DMT determines the weights of NASA TLX's six sub scales as criteria weights by using Eq.(2). At the same time, workload level of each process varies between [0-100] is determined by workers considered each sub scale's rating.

In this study, the NASA TLX scale was applied to the all workers at end of the working day. The survey was given as a written document to the participants after the important points were explained to them in a face to face manner. The survey was distributed to the participants at the same time around 05.00 pm at the end of the working day. The participants were asked to leave their forms in the boxes located in predetermined areas in order to maintain anonymity. Criteria weights are shown in Table 2 and MWL levels for six sub processes are presented in Table 3.

Table 2. Criteria Weights (w_i)

Normalize	W_{MD}	W_{PD}	W_{TD}	W_{PL}	W_{EL}	W_{FL}
Weights	0.222	0.233	0.200	0.156	0.089	0.100

As shown in Table 2, the most important criterion is obtained as physical demand (0.233) of processes.

Table 3. Ratings of subscales for Each Sub Process (S_j) Between [0-100]

Process	MD	PD	TD	PL	EL	FL
PC	95	50	85	90	95	70
CL	95	65	90	80	90	70
PV	70	60	90	80	90	85
PCh	65	90	65	85	75	85
AC	65	85	75	90	80	80
PE	60	85	60	70	80	70

According to Table 3, in terms of work done in PC process, workers were having most difficulty in terms of the MD and effort level that the work required. Similarly, workers were having most difficulty in terms of the MD in CL process. In PV process, workers perceived the most strain in terms of TD and required effort level of process. Workers in the PCh process have the most difficulty in terms of PD criterion. In AC process, workers feel strain about PL criterion. In PE process, it is determined that the work done in this process has difficulty for PD criterion.

In this study, three problem scenarios (S_1, S_2, S_3) are discussed by using NASA-TLX&SMAA-2 to rank six sub

processes respect to MWL levels.

In the first scenario (S_1), criteria importance weights are not taken into consideration. In the present study, it was assumed that NASA-TLX criteria weights cannot be determined by DMT, so, in the SMAA-2 method, criteria weights were obtained by taking equal values to produce solutions. Six sub processes were ranked using only alternative values depicted in Table 3. These alternative values are crisp values for each criteria and for each sub process obtained from workers. RankAccs values of rankings for six sub processes are shown in Table 4.

Table 4. Ranking of Processes For S_1

Process	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
PC	0.21	0.36	0.11	0.09	0.16	0.06
CL	0.44	0.23	0.15	0.18	0.00	0.00
PV	0.00	0.01	0.09	0.13	0.29	0.48
PCh	0.11	0.13	0.10	0.22	0.35	0.08
AC	0.19	0.22	0.35	0.24	0.00	0.00
PE	0.05	0.05	0.20	0.14	0.19	0.37

According to RankAcc values given in Table 4, it was identified that CL is at the first rank with 44% probabilities for MWL. Likewise, it is identified that PC process is at the first rank with 21% probabilities in terms of MWL. PC process is at the second rank with 36% probabilities whereas AC unit at the third with 35%, PE at the fourth with %14, PCh at the fifth with 35%, and finally the PV at the sixth with 48%. According to RankAccs, the PV process has 0% probability to be at first rank. This means that, this process does not have higher MWL compared to other sub processes. For this reason, the improvements in terms of MWL are never implemented in this sub process firstly. CF and CW values for each sub process are presented in Table 5.

Table 5. CW and CF Values for Ranking of Sub Processes in S_1

Process	CF	MD	PD	TD	PF	EF	FL
PC	1.0	0.20	0.07	0.14	0.31	0.09	0.19
CL	1.0	0.21	0.14	0.23	0.08	0.14	0.19
PV	1.0	NA	NA	NA	NA	NA	NA
PCh	1.0	0.11	0.28	0.08	0.10	0.36	0.07
AC	1.0	0.07	0.24	0.14	0.27	0.18	0.10
PE	1.0	0.06	0.25	0.06	0.06	0.24	0.33

It can be seen from Table 5 that CF values are the same. These values had taken value as 1.00 for all sub processes. For example, the confidence level of CL

process to be preferred by DM at the first rank with 44% probabilities is 100%. This shows that the reliability of the CL process being at the first rank is 100% in terms of MWL.

For example, criteria that have the highest and lowest impact levels for CL in terms of being at first rank are TD (23%) and PF (8%) respectively. Additionally, PV was not selected as the first in terms of MWL. For this reason, CW vectors for this process were not calculated. NA (Not applicable) implies this situation.

In the second scenario, points of each MWL criterion is assigned to alternatives as interval values. In the determination of interval values, the values shown in Table 3 were used considering NASA TLX traditional scale. This scale was created between very high and very low with five point increments. For this reason, it is predicted that the values given in Table 3 may change at ± 5 when the interval values are generated. This gives flexibility for determining MWL. RankAccs values of sub processes are presented in Table 6.

Table 6. Ranking of Sub Processes for S_2

Process	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
PC	0.07	0.63	0.28	0.02	0.00	0.00
CL	0.91	0.08	0.01	0.00	0.00	0.00
PV	0.00	0.00	0.00	0.09	0.75	0.16
PCh	0.00	0.01	0.11	0.78	0.09	0.00
AC	0.02	0.28	0.60	0.10	0.00	0.00
PE	0.00	0.00	0.00	0.01	0.16	0.83

According to RankAccs values given in Table 6, it is identified that CL is the first candidate to be chosen to reduce the workload with 91% probabilities. However, it was determined that the PV, PCh, PE processes which had the RankAcc values as 0.00 % have never selected for the first rank. On the other hand, PC is at the second rank with 63% probabilities. AC is at the third rank with 60% probabilities. PCh is at the fourth order with 78% probabilities. PV is at the fifth with 75% probabilities, and finally the PE is at the sixth order with 83% probabilities. After the processes are ranked according to their workloads, their ranks were interpreted with CW and CF values presented in Table 7.

Table 7. CW and CF Values for Ranking of Sub Processes in S_2

Process	CF	MD	PD	TD	PF	EF	FL
PC	0.07	0.22	0.23	0.20	0.16	0.09	0.10
CL	0.91	0.22	0.23	0.20	0.16	0.09	0.10
PV	1.00	NA	NA	NA	NA	NA	NA
PCh	1.00	NA	NA	NA	NA	NA	NA
AC	0.02	0.22	0.23	0.20	0.16	0.09	0.10
PE	1.00	NA	NA	NA	NA	NA	NA

As can be seen from the CF values in Table 7, the CL is preferred at the first rank with 91% confidence level, while PC is preferred at the second rank with 7% confidence level and AC is preferred at the third rank with 2% confidence level. PV, PCh and PE sub processes have 0.00 probability for preference at the first rank with 100 % confidence level. Besides, as there is 0.00 probability of these processes for being at the first rank, the CW values of them were not calculated and shown as NA. Finally, as can be seen in Table 7, the most important criteria in the ranking of process in terms of MWL are PD and MD because of having the highest CW values. This is an expected situation which reflects the truth.

In the third scenario, six sub processes are ranked according to MWL levels combining S_1 and S_2 . Here, criteria importance weights are not taken into consideration. In this scenario, equal importance weights and interval values of alternatives are used. RankAccs values for S_3 are shown in Table 8.

Table 8. Rankings of sub processes for S_3

Alt	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
PC	0.20	0.32	0.11	0.11	0.16	0.10
CL	0.43	0.24	0.15	0.16	0.02	0.00
PV	0.01	0.03	0.11	0.15	0.30	0.41
PCh	0.13	0.14	0.14	0.21	0.29	0.09
AC	0.19	0.22	0.32	0.22	0.05	0.00
PE	0.04	0.06	0.17	0.15	0.19	0.39

According to RankAcc values given in Table 8, it can be seen that CL process should be considered at the first rank with 43% probabilities whereas PC process at the first rank with 20% probabilities. On the other hand, PC is at the second rank with 32% probabilities, AC is at the third rank with 32% probabilities, PCh is at the fourth rank with 21% probabilities, PE is at the fifth rank with 19% probabilities, and finally the PV department is at the sixth rank with 41% probabilities. CW and CF values are given in Table 9.

Table 9. CF and CW Values For Ranking of Sub Processes in S_3

Alt	CF	MD	PD	TD	PF	EF	FL
PC	0.67	0.21	0.08	0.15	0.27	0.09	0.20
CL	0.92	0.21	0.14	0.22	0.11	0.14	0.19
PV	0.14	0.04	0.09	0.54	0.12	0.16	0.06
PCh	0.59	0.11	0.27	0.09	0.15	0.30	0.08
AC	0.75	0.09	0.23	0.13	0.24	0.20	0.11
PE	0.64	0.07	0.23	0.07	0.06	0.23	0.34

According to CF values in Table 9, it is seen that confidence level of being at the first rank of CL is 92% and CL of being at the first rank of PV is 14%. On the other hand, respect to CW values, MD and TD criteria, with their 21% and 22% impact values, are the most significant criteria for CL process in terms of being at the first rank for MWL.

6. Comparative Analysis

In this section, the rankings obtained from NASA-TLX and NASA-TLX&SMAA-2 methods are compared. The MWL was determined according to Eq.(1) by using Table 2 and 3. Rankings for six sub processes obtained from NASA-TLX method are depicted in Table 10 by using NASA TLX 1.0.0.1 software.

Table 10. NASA-TLX Results

	MD	PD	TD	PF	EF	FL	TLX
PC	38	6.6	17	12	6.3	4.6	84.6
CL	31.6	8.6	18	16	6	4.6	85
PV	8.6	30	8.6	17	10	5.6	79.3
PCh	8	28.3	8	9.2	10.6	9.3	80
AC	13	22.6	15	12	5.3	10.6	78.6
PE	9.3	12	30	10.6	6	11.3	73.6

As shown in Table 10, according to NASA TLX results, although the MWL of the sub processes are close to each other, it is possible to determine which process has more workload. NASA-TLX&SMAA-2 can be used effectively to measure workload as it provides DMs with information as to which alternatives are selectable and to what extent, rather than presenting them with the best alternative among others.

Comparison of rankings obtained by NASA-TLX&SMAA-2 method for three different scenarios and NASA TLX are shown in the Table 11.

Table 11. NASA TLX&SMAA Ranking Results

Rankings	NASA-TLX&SMAA-2			NASA-TLX
	S_1	S_2	S_3	
1	CL	CL	CL	CL
2	PC	PC	PC	PC
3	AC	AC	AC	PCh
4	PE	PCh	PCh	PV
5	PCh	PV	PE	AC
6	PV	PE	PV	PE

As seen in Table 11, rankings obtained from the NASA-TLX&SMAA-2 method for three scenarios are the same. The CL process is found to be the first department required to reduce workload levels in all rankings. This is because the CL process purports to be the primary station for box letter manufacturing process. In addition, the PC process is ranked as second in all scenarios. It is an expected result according to DMT. In CL process, letter is covered by chrome and this covering should be performed carefully. Smoothness of the material surface is ensured by proper coating. Additionally, in PC process, it is necessary to cut the letters in the appropriate shape and size. This sub process also needs to pay attention. Actually, workers are struggling much more in terms of MWL in these two sub processes.

Because of using RankAccs, CW and CF arguments in NASA TLX&SMAA-2 approach, the obtained ranking results are more realistic and accurate. The reliability of rankings can be debated in NASA TLX&SMAA-2. The impact of each criterion for sub process rankings can be seen. Finally, NASA TLX&SMAA-2 provides flexibility for manager to establish different action plans to reduce MWL. The proposed approach provides flexibility because it determines which sub-process needs to be improved primarily in terms of MWL, taking into account the probability values. Managers structured their decisions based on probability values. Different scenarios can be employed by using NASA TLX&SMAA-2. However, NASA TLX produces only one result as the sub process that has the highest MWL. It can only use the PWT argument. It can only work with crisp values. It cannot give flexibility for DMs to rate sub scales with interval values. However, in real life cases, people can make decisions by using approximate values more easily.

In S_1 , it is assumed that weight of each sub scale of NASA TLX has the same importance weight and ratings of each sub scale between [0-100] are assigned in a crisp manner, only one value. In these conditions, it was determined that CL process is at the first rank and PC process is at the second rank. It means that, the mental workload that occurs in the CL process is the highest

when the workers ignore which criterion is compulsory so, MWL occurs in this process should be reduced firstly. In terms of contribution to this highest MWL, MD criterion has the highest impact for this sub process. Managers should improve problems that cause to feel MD highly.

In S_2 , workers have more flexibility to score each sub scale for each sub process because of using interval valued points. Additionally, in this scenario it is also considered that the effects of the criteria on the MWL may be different. According to results, CL process is at the first rank and PC process is at the second rank again. This means, even if they feel difficulty in different dimensions of the MWL, this process is the most difficult process. PD of work done in this process has the highest effect this MWL.

In S_3 , workers have more flexibility to score each sub scale for each sub process like S_2 . They could give interval valued scores to each sub scale. However, it was assumed that all MWL related criterion have the same weight. CL process is at the first rank and PC process is at the second rank again for this scenario. CL process is at the first rank that should be improved in terms of MWL, when assumed that each sub scale of MWL has the same importance.

According to three scenarios, S_2 analyzes MWL more detailed. Because, this scenario, considers different effect of six sub scales on MWL as criteria weights and this scenario gives flexibility to score each sub scale for workers at the same time. It is a well-known fact that different criteria have the different effect on decisions. As MWL assessment is a MCDM problem, different sub scales should have different effect on MWL.

7. Conclusion and Discussions

In this study, the MWL level determining problem is discussed as a stochastic MCDM problem, which is well suited to cases where criteria values are not precisely known by different DMs. Therefore, to solve this problem, the NASA-TLX method that is commonly used in MWL measurement and the SMAA-2 method, which allows to record of alternative performance values and importance weights of criteria as intervals, distribution, and missing data, were used together. This paper found that the NASA-TLX&SMAA-2 approach provides more precise and distinctive results compared to the NASA-TLX method.

SMAA-2 is a practical tool for MWL analysis to enable the sequential ordering of the MWL related criteria, to determine the effectiveness or impact of these criteria. NASA-TLX&SMAA-2 supports managers about which criteria should be improved to reduce MWL in sub processes.

The MWL levels of workers in a box letter

manufacturing process were analyzed. It was observed that MWL levels were high in processes CL, PC and AC for three scenarios.

When NASA-TLX&SMAA-2 is employed, it is observed that workload levels of some sub processes are low. In order to provide for the health, safety, comfort and efficiency of workers in the long term, tasks should be regulated in such a way that the individual is neither under loaded nor overloaded. Dangers in overloading have been known for a long time. At this point, however, we are concerned about the effects of under load and boredom stress, particularly related to the transactions that are operated in an automated fashion. Therefore, the workload levels of processes having low MWL should also be increased to higher levels by rearranging the scope and responsibilities of the work.

In further studies, the proposed NASA-TLX&SMAA-2 method can be employed in different sectors in the field of MWL as a Decision Support System by developing a user friendly interface. This allows easy data entry with a user-friendly interface and results can be listed on the computer screen so that the administrator can understand them. NASA-TLX&SMAA-2 approach, proposed in the present study, can be extended by using different methods that consider the relations between sub scales.

Conflict of Interest

No conflict of interest was declared by the authors.

References

- Aertsen, W., Kint, V., Orshoven, J.V. and Muys. B. (2011). Evaluation of modelling techniques for forest site productivity prediction in contrasting ecoregions using stochastic multicriteria acceptability analysis (SMAA). *Environmental Modelling & Software*, 26(7), 929-937.
- Akgül, E., Aydoğan, E. K., Özmen, M., and Türksoy, H. G. (2017). Optimization of the Murata Vortex Spinning machine parameters by the SMAA-MOORA approach. *Industria Textila*, 68(5), 323-331.
- Akyeampong, J., Udoka, S., Caruso, G. and Bordegoni, M. (2014). Evaluation of hydraulic excavator Human-Machine Interface concepts using NASA TLX. *International Journal of Industrial Ergonomics*, 44(3), 374-382.
- Angilella, S., Corrente, S. and Greco, S. (2015). Stochastic multiobjective acceptability analysis for the Choquet integral preference model and the scale construction

- problem. *European Journal of Operational Research*, 240(1), 172-182.
- Angilella, S., Arcidiacono, S. G., Corrente, S., Greco, S., and Matarazzo, B. (2017). An application of the SMAA-Choquet method to evaluate the performance of sailboats in offshore regattas. *Operational Research*, 1-23.
- Aydogan, E. K., and Ozmen, M. (2017). The stochastic vikor method and its use in reverse logistic option selection problem. *RAIRO-Operations Research*, 51(2), 375-389.
- Colligan, L., Potts, H. W., Finn, C. T. and Sinkin, R. A. (2015). Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record. *International Journal of Medical Informatics*, 84(7), 469-476.
- Durantín, G., Gagnon, J. F., Tremblay, S. and Dehais, F. (2014). Using near infrared spectroscopy and heart rate variability to detect mental overload. *Behavioural Brain Research*, 259, 16-23.
- Filtness, A. J., and Rudin-Brown, C. M. (2012). Drinking and riding: is subjective workload related to performance? Australian Road Safety Research, *Policing and Education Conference*, Wellington, New Zealand, 4-6 October.
- Flemisch, F. O., and Onken, R. (2002). Open a Window to the Cognitive Work Process! Pointillist Analysis of Man-Machine Interaction. *Cognition, Technology & Work*, 4, 160-170.
- Gülkaç, H. (2013). Pilotların zihinsel işyüklerinin NASA-TLX yöntemiyle ölçülmesi, Yüksek Lisans Tezi, Kocaeli Üniversitesi, Kocaeli.
- Hart, S. G., and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139-183.
- Hokkanen, J., Lahdelma, R. and Salminen, P. (2000). Multicriteria decision support in a technology competition for cleaning polluted soil in Helsinki. *Journal of Environment Management*, 60(4), 339-348.
- NASA Ames Research Center, NASA-TLX. <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000021487.pdf>. Erişim tarihi:14 Mart 2018.
- Jorgensen, A.H., Garde, A.H. and Laursen, B., Jensen, B.R. (1999). Applying the concept of workload to IT work. *In Virtual Proceedings of CybErg*.
- Kadzinski, M., and Tervonen, T. (2013). Stochastic ordinal regression for multiple criteria sorting problems. *Decision Support Systems*, 55(1), 55-66.
- Kangas, J., Hokkanen, J., Kangas, A. R. and Lahdelma, P. S., (2003). Applying stochastic multicriteria acceptability analysis to forest ecosystem management with both cardinal and ordinal criteria. *Forest Science*, 49(6), 928-937.
- Kangas, J., Store, R. and Kangas, A. (2005). Socioecological landscape planning approach and multicriteria acceptability analysis in multiple-purpose forest management. *Forest Policy and Economics*, 7(4), 603-614.
- Kangas, A., Kangas, J., Lahdelma, R. and Salminen, P. (2006). Using SMAA-2 method with dependent uncertainties for strategic forest planning. *Forest Policy and Economics*, 9(2), 113-125.
- Kuehn, K. A., Chourasia, A. O., Wiegmann, D. A. and Sesto, M. E. (2013). Effects of Orientation on Workload during Touch Screen Operation among Individuals With and Without Disabilities. *In Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), 1580-1584.
- Lahdelma, R., and Salminen, P. (2001). SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. *Operations Research*, 49(3), 444-454.
- Lahdelma, R., and Salminen, P. (2002). Pseudo-criteria versus linear utility function in stochastic multicriteria acceptability analysis. *European Journal of Operational Research*, 141(2), 454-469.
- Lahdelma, R., and Salminen, P. (2006). Classifying efficient alternatives in SMAA using cross confidence factors. *European Journal of Operational Research*, 170(1), 228-240.
- Lahdelma, R., Makkonen, S. and Salminen, P. (2009). Two ways to handle dependent uncertainties in

- multicriteria decision problems. *Omega*, 37(1), 79-92.
- Lee, Y. H., and Liu, B. S. (2003). Inflight workload assessment: Comparison of subjective and physiological measurements. *Aviation, Space, and Environmental Medicine*, 74(10), 1078-1084.
- Loft, S. P. Sanderson, A. N. and Mooij, M. (2007). Modeling and predicting mental workload in en route air traffic control: critical review and broader implications. *Human Factors*, 49(3), 376-399.
- Matthews, R., Legg, S. and Charlton, S. (2003). The effect of cell phone type on drivers subjective workload during concurrent driving and conversing. *Accident Analysis & Prevention*, 35(4), 451-457.
- Mazur, L. M., Mosaly, P. R., Hoyle, L. M., Jones, E. L. and Marks, L. B. (2013). Subjective and objective quantification of physician's workload and performance during radiation therapy planning tasks. *Practical Radiation Oncology*, 3(4), 171-177.
- Moroney, W.F., Biers, D.W. and Eggemeier, F.T. (1995). Some measurement and methodological considerations in the application of subjective workload and measurement techniques. *International Journal of Aviation Psychology*, 5, 87-106.
- Mouzé-Amady, M., Raufaste, E., Prade, H. and Meyer, J.P. (2013). Fuzzy-TLX: using fuzzy integrals for evaluating human mental workload with NASA-Task Load index in laboratory and field studies. *Ergonomics*, 56(5), 752-763.
- Parasuraman, R., and Hancock, P. A. (2001). *Adaptive Control of Mental Workload. In Stress, Workload and Fatigue*, Lawrence Erlbaum, 305-320.
- Park, J. I. N. K. Y. U. N., Jung, W. and Ko, J. O. N. G. H. Y. U. N. (2009). Investigating the appropriateness of the TACOM measure-application to the complexity of proceduralized tasks for high speed train drivers. *Nuclear Engineering and Technology*, 42(1), 115-124.
- Rubio, S. E., Díaz, J., Martín. and Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review*, 53(1), 61-86.
- Ruiz-Rabelo, J. F., Navarro-Rodriguez, E., Di-Stasi, L. L., Diaz-Jimenez, N., Cabrera-Bermon, J., Diaz-Iglesias, C.,... and Briceño-Delgado, J. (2015). Validation of the NASA-TLX Score in Ongoing Assessment of Mental Workload During a Laparoscopic Learning Curve in Bariatric Surgery. *Obesity Surgery*, 25(12), 2451-2456.
- Song, L., Fu, Y., Zhou, P., and Lai, K. K. (2017). Measuring national energy performance via Energy Trilemma Index: A Stochastic Multicriteria Acceptability Analysis. *Energy Economics*, 66, 313-319.
- Sporcic, M., Landekic, M., Lovric, M. and Martinic, I. (2011). Planning and Decision Making Models in Forestry. *Croatian Journal of Forest Engineering*, 32(1), 443-456.
- Stefanidis, D., Wang, F., Korndorffer Jr, J. R., Dunne, J. B. and Scott, D. J. (2010). Robotic assistance improves intracorporeal suturing performance and safety in the operating room while decreasing operator workload. *Surgical Endoscopy*, 24(2), 377-382.
- Tervonen, T., Hokkanen, J. and Lahdelma, R. (2008). Elevator planning with stochastic multi criteria acceptability analysis. *Omega*, 36(3), 352-362.
- Tervonen, T., Figueira, J.R., Lahdelma, R., Dias, J.A. and Salminen, P. (2009). A stochastic method for robustness analysis in sorting problems. *European Journal of Operational Research*, 192(1), 236-242.
- Tervonen, T., Valkenhoef, G.V., Buskens, E., Hillege, H.L. and Postmus, D. (2011). A stochastic multicriteria model for evidence based decision making in drug benefit-risk analysis. *Statistics in Medicine*, 30(12), 1419-1428.
- Tervonen, T. (2012). JSMAA: open source software for SMAA computations. *International Journal of Systems Science*, 45(1), 69-81.
- Vidulich, M.A, and Tsang. P.S. (2007). Methodological and theoretical concerns in multitask performance: A critique of Boles, Bursk, Phillips, and Perdelwitz. *Human Factors*, 49(1), 46-49.
- Wickens, C. D. (2008). Multiple Resources and Mental Workload. *Human Factors*, 50(3), 449-455.

- Young, M. S., Brookhuis, K.A. Wickens, C. D. and Hancock, P.A. (2015). State of science:mental workload in ergonomics. *Ergonomics*, 58(1), 1-17.
- Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Pantou, O. N. M. and Atkins, M. S. (2012). Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surgical endoscopy*, 26(10), 2746-2750.
- Zhou, H., Wang, J. Q., and Zhang, H. Y. (2017). Stochastic multicriteria decision-making approach based on SMAA-ELECTRE with extended gray numbers. *International Transactions in Operational Research*.
- Zhu, F., Zhong, P. A., Wu, Y. N., Sun, Y., Chen, J., and Jia, B. (2017). SMAA-based stochastic multi-criteria decision making for reservoir flood control operation. *Stochastic Environmental Research and Risk Assessment*, 31(6), 1485-1497.