SHS Web of Conferences **189**, 01043 (2024) https://doi.org/10.1051/shsconf/202418901043
ICESH & Ergo-Camp 2023

The use of multi-attribute task battery in mental workload studies: A scoping review

Ridwan Aji Budi Prasetyo*

Psychology Department, Faculty of Social and Political Sciences, Brawijaya University, Malang 65145, Indonesia

Abstract. Multi-Attribute Task Battery (MATB) is a software that has been arguably utilized in many ergonomics/human factors studies, including in the topic of mental workload. However, the use of this well-known program in diverse investigations has not yet been systematically tracked. Furthermore, it may be argued that a critical appraisal of MATB is urgently needed so that future researchers and users can take several crucial factors into account when planning a study or experiment using MATB. The aim of this paper is to comprehensively identify and review the use of MATB software in published studies. This aim might be accomplished by achieving two goals: (1) systematic discovery of published papers in literature databases and (2) categorization of research according to pertinent topics. In this paper, thirty-one articles were included for analysis after carefully screening for their eligibility. Our scoping review finds that MATB is a beneficial program for creating multitasking environments in general, with aviation being the area where it has been used the most. The program was also extensively used for studies on mental workload, especially by producing various stimuli that ultimately result in varying degrees of task demand or difficulty. Moreover, to successfully use MATB, researchers must be aware of a few operational issues and criticisms.

1 Introduction

Multi-Attribute Task Battery (MATB) was released in 1992 for the first time and aimed to evaluate operator performance and workload in a multitasking environment. The platform was running on DOS and early version of Windows operation system [1]. MATB was developed to simulate aircraft crew tasks as the basis to evaluate performance and workload through four main tasks: (1) system monitoring (SYSMON), which simulates system monitoring task and requires operator to detect changes in system and respond accordingly; (2) tracking (TRACK), which simulates aircraft position maintaining task that requires operator to maintain moving target within certain area; (3) resource management (RESMAN), which simulates fuel management task and requires participant to maintain fuel level and detect any failures occur in pumping system; and (4) communication (COMM), which simulates communication with air traffic controller (ATC) and requires operator to

* Corresponding author: ridwan.prasetyo@ub.ac.id

detect and respond verbal command from ATC. In addition to the four main tasks, the scheduling (SCHED) window also appears to provide information about expected workload.

As the computer system and use evolved quickly, NASA released the new version of the platform in 2011, called Multi-Attribute Task Battery II (MATB-II), to fit recent Windows operating system. MATB-II is basically the second generation of similar battery developed previously, with some noticeable changes in its interface, configurations, and operating system on which the software is running. Nothing had changed in terms of the task simulated in the platform. There are still four main tasks in MATB-II and the scheduling window as in the earlier version. Workload ratings based on NASA-TLX is also provided as the previous version [2].

Since its release in 1992, and updating in 2011, it is reported that both MATB and MATB-II has been widely used in studies involving mental workload and multitasking performance evaluation. These two concepts, multitasking and mental workload, are apparently related since performing tasks in a multitasking environment may generate higher mental demands, thus, higher mental workload. Mental workload is also a concept with strong intuitive sense, tallying with everyday lived experience, and it appears to be common in many work environments. In addition, it is apparent that modern-day jobs creating more mental/cognitive demands due to multitasking nature of the jobs [3]. A study from [2] included 41 selected publications utilizing MATB from 2003-2010 on their report, most of them are conference proceedings. However, systematic trace of the use of this popular software in various studies has not been undertaken yet. Furthermore, a critical appraisal of MATB is arguably urgent so that future researcher or users can take several important aspects into consideration when designing a study or experiment using MATB. This paper therefore focuses to systematically review and provides the scoping regarding the use of MATB software in published studies. These aims could be achieved by completing two objectives: (1) systematic finding of published articles in literature databases and (2) classification of studies according to relevant themes. Our scoping literature review will be addressing these three following research questions:

1. What is the context of the studies on which MATB has been utilized?
2. To what extent MATB has been used for mental workload studies?
3. If any, what are the operational considerations of MATB utilization?

2 Methods

Studies analysed in this paper were selected through online scientific databases, including Web of Science, ErgoAbs, ScienceDirect, and Scopus. The databases were accessed via university walk-in user (WIU) resources. The reason of choosing those databases was mainly because they cover most published articles from various disciplines, as the MATB might be utilized loosely for different purposes. The exception was ErgoAbs since studies using MATB might have been done mostly in human factors or ergonomics area. The term used for articles finding was simply “multi-attribute task battery” and the phrase was intended to act as a whole phrase, by applying Boolean operator “AND” (i.e., multi-attribute AND task AND battery) or simply putting quotation marks on the phrase (as in ScienceDirect), meaning that the search result must include all words with the specified order. The target field for the searching was set to either “topic” or “title, abstract, keywords”. Since MATB/MATB-II is a name of the software that is very specific, thus the search result for the term (“multi-attribute task battery”) is likely to appear either in article titles or abstracts for studies that mainly utilized the software, including if the author was using the abbreviation.

The total of 111 articles were found from all databases (Web of Science: 32, Scopus: 39, ScienceDirect: 11, ErgoAbs: 29). Subsequently, title and abstract scan were conducted manually to select articles to be included in the analysis by applying inclusion criteria consisting of (1) journal article and (2) research articles. The reason for this filter is that journal article is mostly peer-reviewed, making them more reliable, and some of them were also coming from proceedings. Exclusion criteria were then implemented by excluding any type of documents other than journal article, such as proceedings, report, or book chapter, removing duplicates, and excluding articles that were not written in English. The total of 35 articles were considered relevant. The final phase was to find the full article either using "find full text" feature on EndNote reference manager software, downloading directly from the databases, or even searching in Google for the articles that could not be accessed due to subscription issue. Figure 1 describes the flowchart of articles inclusion process.

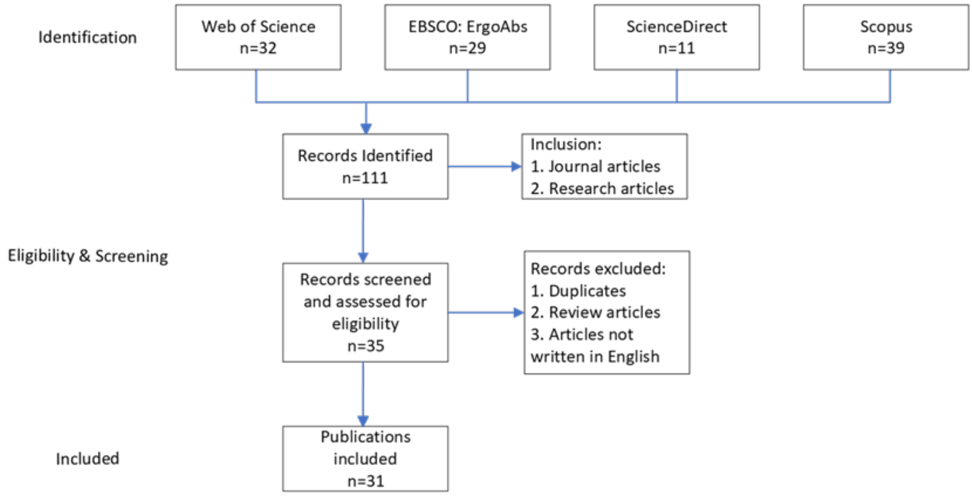

Fig. 1. Flowchart of records inclusion process.

3 Results and discussion

The very final of 31 articles were considered for the analysis, since the full version of four articles were still unable to be obtained. These articles were then classified according to three relevant topics in terms of the use of MATB in studies: (1) context of the studies using MATB, (2) the use of MATB for mental workload studies, and (3) MATB operational considerations. Table 1 shows articles included in the analysis in chronological order and the overview of the MATB use in respective study.

Table 1. Articles included in the analysis.

No.	Source	Aims of the study	Summary of the Utilization of MATB/MATB-II
1	[4]	Examine the effect of reaction modality on reactions to alarm	Serving as the primary task by employing tracking (TRACK) and monitoring (SYSMON) subtasks to test the reactions to alarm using different modalities
2	[5]	Examine the effect of task duration on sensitivity of sleep deprivation to performance	Providing task (all subtasks) on which sleep deprivation sensitivity to performance is assessed

3	[6]	Examine the accuracy of artificial neural networks (ANN) in real time using psychophysiological variables to determine participant's level of mental workload while performing a complex task	Providing tasks (all subtasks) that generate different level of mental workload to be measured by using physiological methods and input to ANN classifier
4	[7]	Determine the effect of desloratadine (a long-acting, nonsedating antihistamine) on healthy individuals placed under condition of simulated cabin pressure	Providing multitasking (all subtasks) and performance measuring platform to see the effect of desloratadine
5	[8]	Investigate if the sensitivity and diagnosticity of psychophysiological variables to task demands was influenced by learning/skill acquisition	Providing different level of task demands (low and high) to be measured by using psychophysiological variables (three subtasks, excluding COMM)
6	[9]	Investigate whether psychophysiological measures could predict changes in subjective states	Generating multitasking environment (excluding COMM subtask) to be measured by using psychophysiological variables and correlated with individual subjective states
7	[10]	Develop a quantitative model of the human operators (HO) and machine interaction during multiple task performance	Providing tasks (excluding COMM subtask) that were quantified into baud rate (bits/s) to model different level of task demands
8	[11]	Determine the effect of fatigue on the range of task performance and psychophysiological consequences, specifically in the context of UAV piloting	Providing aviation/UAV related tasks (all subtasks) as a basis to measure the effect of fatigue on performance and psychophysiological and subjective parameters
9	[12]	Investigate cognitive mechanism of procedural knowledge degradation	Providing resource management task (only RESMAN included) to trigger procedural knowledge degradation through interference task
10	[13]	Examine timesharing ability among inexperienced cadets of USAF Reserve Officer Training Corps (US AFROTC)	Providing single (TRACK), dual (TRACK and SYSMON), and tri-tasks (TRACK, SYSMON, and RESMAN) as a measuring basis for timesharing ability
11	[14]	Examine the test/retest consistency of physiological responses induced by mental tasks	Providing tasks with different level of difficulties (excluding COMM subtask) to be measured by using subjective and physiological methods, as well as performance indices

12	[15]	Determine the effects of a single dose of loratadine/montelukast (L/M) versus placebo and diphenhydramine on daytime somnolence and psychomotor performance in individuals in simulated cabin pressure	Providing tasks (all subtasks) as a measure of psychomotor performance during administration of the before mentioned drugs
13	[16]	Examine the relationship between personality variables (state and trait) and cognitive performance after sleep deprivation	Providing single task (SYSMON) with which participant's response time was measured as indicator of cognitive performance
14	[17]	Evaluate the effects of alcohol intoxication on multi-task performance	Generating tasks (excluding COMM subtask) as a platform to measure performance and perceived workload after alcohol intoxication
15	[18]	Examine the effects of 35 hours of continuous sleep deprivation on performance in a variety of cognitive tasks as well as simulated flight	Providing cognitive tasks (excluding RESMAN subtask) as a measuring index of performance and predictor of flight performance in simulator
16	[19]	Test the accuracy of cross-subject workload classifier based on hierarchical Bayes model	Providing three levels of task difficulty (all subtasks) to be measured by physiological methods (EEG) and then sampled for classifier model training
17	[20]	Examine the effects of action videogames training in improving multitasking performance in high workload environments	Providing high workload multitasking environments (all subtasks) to assess the effects of videogames training
18	[21]	Identify physiological indices that accurately monitor mental workload while participants perform multitasks	Providing multitasks (excluding TRACK subtasks) with different level of difficulties (low, medium, high level) as a platform to measure MWL using both physiological and subjective methods
19	[22]	Find out the possibility of obtaining quantitative information about the degree of the learning process throughout a training period by analysing the variations of some physiological signals	Providing multitasks (all subtasks) with different level of difficulties, from hyper-easy to hard, as a platform to capture some physiological signals to assess training improvement
20	[23]	Examine the effect of inspired partial pressures of CO_2, N_2, and O_2 to cognitive performance in immersed, exercising divers	Providing multitasking environment (all subtasks) to assess diver's cognitive performance
21	[24]	Examine the effects of time-on-task to the resistance of task switching	Providing different level of task difficulties (particularly on TRACK subtask) and a platform to measure task switching from ongoing task (OT) to alternative task (AT)

22	[25]	Evaluate the efficacy of a non-invasive brain stimulation technique (transcranial direct current stimulation/tDCS) to improve information processing capabilities during a multitasking environment	Providing a multitasking platform (all subtasks) with a baud input increments to evaluate the effect of tDCS; This version of MATB also employed information throughput (IT) mode and baud output as a measure of performance
23	[26]	Evaluate the efficiency of a workload estimation based on the ERPs elicited by infrequent, task-independent and ignored auditory probes	Providing multitasking environment (excluding COMM) that generates low and high workload, parallel with auditory stimuli, as a platform to record EEG/ERP activity whose data was used to classify workload
24	[27]	Assess the effect of a single dose of bilastine 20 mg on alertness and complex task performance of healthy individuals in a hypobaric condition/simulated cabin pressure	Providing simulated aircrew tasks (all subtasks) as a basis to measure performance after drug administration
25	[28]	Validate the strategic task overload management (STOM) model that predicts task switching when concurrence is impossible	Providing different level of task difficulties (particularly on TRACK subtask) to assess task switching strategies
26	[29]	Demonstrate improved cross-day workload classification accuracy by combining statistical parameters and variety of RNN architectures	Generating low and high workload condition (all subtasks) to obtain EEG dataset that will be used for deep RNN models
27	[30]	Examine the use of fractal dimension of pupil size changes to measure the level of task complexity in multitasking environment	Serving as the secondary task (excluding COMM) for Time Window-based Human-In-The-Loop (TWHITL) as a measure of task complexity level
28	[31]	Validate the use of new wearable device that captures electro-physiological data in response to a variety of aviation tasks	Generating aviation tasks with two level of task load (low/high; all subtasks) while physiological data from the wearable device was captured
29	[32]	Validate a coloured petri nets (CPN) based workload evaluation model	Generating three levels of task difficulty (all subtasks) as a basis to measure workload using a subjective method (NASA-TLX), whose data will be compared with CPN model values
30	[33]	Examine cross-participant estimation of operator workload in a non-stimulus-locked, multi-task environment by developing and evaluating seven neural network architectures	Generating low and high workload condition (all subtasks) to obtain EEG dataset that will be used for neural network training

31	[34]	Validate a quantitative model for the study of user performance improvement in a multitasking environment	Generating three levels of task difficulty (all subtask) based on baud rate and to be used as a basis for subjective workload rating and for performance measurement

3.1 Context of the studies

Since the development of MATB aimed to simulate aircraft crew tasks, it is not surprising that some of the studies were conducted in the context of aviation (26% of our included papers, see Figure 2), with various group of dependent variables such as mental workload (12.5%), performance (62.5%), reaction times (12.5%), and time-sharing ability (12.5%). The main aims of the studies in this context vary, e.g., to examine the effect of reaction modality [4], task duration and sleep deprivation [5], [18], fatigue [11], and certain drugs [7], [27] to variety of performance. The examination of timesharing ability among inexperienced flying cadets [13] and new wearable device to measure aircrew task loads [31] were also conducted using this software. For these studies, MATB provided a platform to measure multitasking performance, which represents flying performance, to see the effect of the independent variables. Several studies even employed pilots (including flying cadets) as participants [4], [5], [13], [18].

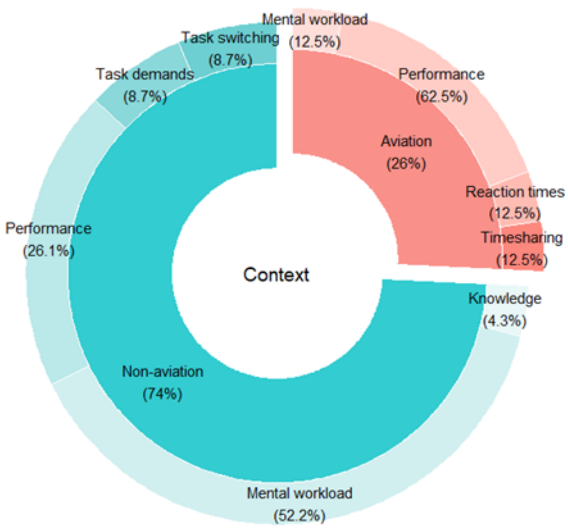

Fig. 2. Context of studies on which MATB has been used.

Other than aviation context, studies using MATB has been applied in another specific context, such as automotive [17] and diving [23]. For the study in automotive, MATB provided multitasking environment which was considered equal to driving tasks that act as a measurement basis to see the effect of alcohol intoxication. Meanwhile, for study in diving, MATB was employed directly in the water tank (using special equipment) to check whether divers' cognitive performance was affected by partial pressure of certain gases. Regarding environmental condition of the study, this extreme condition (hyperbaric) on [23] was similar to other studies [7], [15], [27], which simulated hypobaric condition (as in high altitude) by using special chamber on which participants performed MATB mission.

The rest of the enlisted studies (around 74%) were conducted in non-aviation, more general context. Those studies took benefits from MATB ability to generate multitasking environment as a basis to assess drug effect [15], workload/task demands/task complexity [8], [9], [14], [21], [26], [30], multitasking strategies [20], [24], [25], learning improvement [22], and cognitive performance [12], [16]. Another context that is possibly one of the prominent topics in this context is modelling [6], [10], [10], [19], [24], [29], [32]–[34]. Two studies mostly performed quantitative modelling of multitasking performance in the context of human-machine interaction using baud rate stimulus (bits per second) and baud rate response [10] and of task switching strategies in overload condition [28]. Meanwhile, the rest of the studies about modelling were related to mental workload evaluation using neural network training [6], [29], [33], hierarchical Bayes model [19], coloured petri nets [32], and baud rate [34].

3.2 Mental workload studies using MATB

The ability to generate multitasking mission makes MATB is widely used in studies related to mental workload measurement or evaluation. From our included studies, mental workload has been applied as the dependent variable as many as 41.93%. Task scenario within the software can be easily configured as necessary, giving full control to experimenter for manipulating level of task difficulty or complexity to induce mental workload to participants. There are 13 studies from the list on Table 1 that involve mental workload measure, evaluation, or estimation (i.e., [6], [8], [14], [17], [19], [21], [26], [29]–[34]). Some studies overlap with previous topic as the authors applied certain modelling technique to evaluate mental workload [6], [19], [29], [32]–[34].

Regarding methods to study mental workload, most of the studies were utilizing psychophysiological measures such as EEG [8], [21], [26], ECG [9], [14], [21], [26], [31], respiration rate [8], [14], [31], tissue blood volume in nose and skin potential level [14], and eye-tracking [30]. Subjective measures through one of the widely used tools in the field, NASA-TLX, were also undertaken either as the main tool for mental workload measurement [17], as a measure to determine level of mental workload in a pilot study [8], [21], or as a comparison for psychophysiological techniques [14], [31]. The high practicality to operate the software gives benefits for experimenter who employs a wide range of psychophysiological devices, as well as subjective measures through questionnaires (NASA-TLX, RSME, etc.).

Since mental workload may be affected by certain psychological factors and demographics, we attempted to identify studies considering individual differences when using MATB. A study from Carlozzi et al. [16] attempted to investigate the influence of personality traits and states on reaction time after sleep deprivation. Cloninger’s tri-dimensional personality variables of novelty seeking, harm avoidance, and reward dependence, and state personality variables examined by McNair’s Profile of Mood States (POMS) including tension/anxiety, depression/dejection, anger/hostility, vigour/activity, fatigue/inertia, and confusion/bewilderment were employed to achieve the goal. Meanwhile, another study tried to investigate the role of individual difference in working memory capacity (WMC) examined by the operation span task (OSPAN) on multitasking performance through a quantitative modelling [34].

There are also other studies which involved participants’ psychological properties, such as subjective states i.e., task engagement, distress, and worry as examined by Dundee Stress State Questionnaire/DSSQ, to see whether psychophysiological measures could predict changes in those subjective states [9]. POMS and Visual Analog Scales from Penetar et al. were also employed in a study [11] to determine fatigue effects on participants’ subjective states in the context of UAV piloting task. Study from [14] also included mood states and feelings i.e., stress and arousal (Stress Arousal Checklist), state anxiety (State and Trait Anxiety Inventory), and fatigue and vigour mood scores from POMS as an additional analysis for interpreting physiological response. Individual differences might also be related to the population used for the studies. As partially mentioned earlier, some studies employed special population as participants, i.e., rated pilots [4], [5], [18] or flying cadets [13] and participants that were strictly selected to meet Navy fitness criteria [23]. These sampling strategies apparently aimed to correspond to the study purposes (ecological validity issues), resource availability, and safety issues.

3.3 MATB operational considerations

Most of the studies utilized earlier version of the software (MATB). If it is chronologically traced, studies from 1997 to 2016 used this version. Meanwhile, studies from 2016 onwards tends to employ the newer version of this software (MATB-II), with exception of two studies, which still uses the earlier version [27] or uses modified version of earlier MATB called Air Force MATB (AF-MATB) [25]. The AF-MATB implements the information throughput (IT) mode for more independent adjustment of the stimulus (for details about AF-MATB, see [35]). With regards to the subtasks provided by the software (TRACK, SYSMON, COMM, and RESMAN), most studies (59.26%), as seen in Figure 2, used all these subtasks whereas the remaining studies applied them partially with two or three subtasks. The subtasks used in respective studies vary yet the most common one was the combination of TRACK, SYSMON, and RESMAN subtasks (25.93%). In general, MATB was used as a platform to generate multitasking environment (87%) even though it is operationally possible to implement MATB as a single tasking platform, for instance, by applying only TRACK, RESMAN, or SYSMON individually. Table 1, as shown earlier in this paper, included notes about subtasks that were implemented in every study.

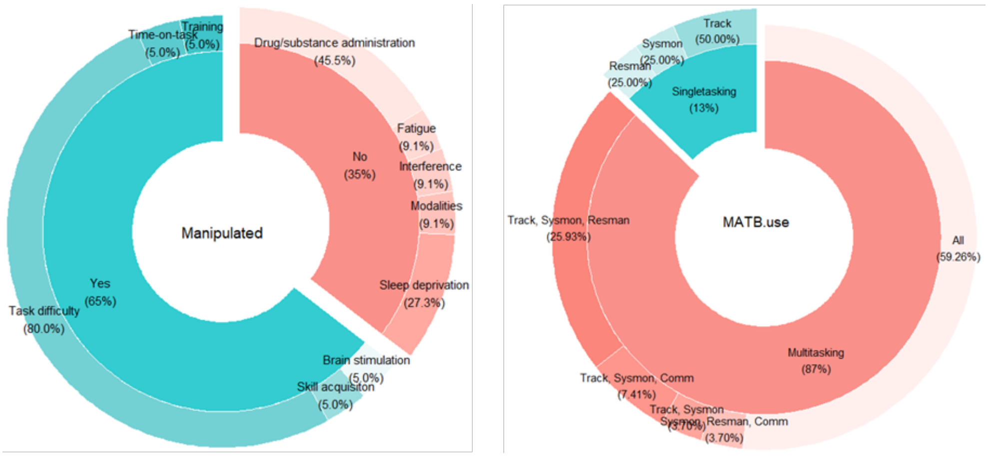

Fig. 3. Percentage of MATB utilization by (a) whether researcher manipulate the MATB stimuli and (b) whether researcher use MATB as single- or multi-tasking platform.

Justification to employ all or partial subtasks seems to be loose. Most authors did not include any argument or description regarding subtask implementation. Several authors attempted to communicate the relevance of using subtask(s) to the purpose or context of their studies [12], [13], [21], even with some evidence [16], [18]. From the papers we have analysed, most studies have used MATB as a platform to manipulate task difficulty (80%) and other relevant independent variables. MATB stimuli on each subtask were varied by researcher to create different levels of difficulty for demonstrating changes in mental workload [6], cognitive or psychomotor performance [21], [22], or reaction times [4]. Nevertheless, MATB may be also used without varying subtask stimuli. Researcher may create a single set of stimuli to be tested to participants after being exposed to a particular treatment, for example, drug or substance administration. In this kind of study, MATB task with an identical set of subtask stimuli was usually employed to both treatment and control group. Figure 3 shows the detail percentage of MATB utilization for stimuli manipulation or as a platform for multitasking and single task environment.

3.4 General discussion

Aviation was possibly the most popular context in which studies using MATB software are conducted, based on the articles included in this paper. There is none to very limited studies for the specific area other than aviation, such as maritime, railway, special vehicles, construction, or general automotive. However, the number of studies using MATB for more general context is also considerable, with the specific emphasize on the use of modelling techniques to quantify tasks or consequences of the tasks such as workload or multitasking performance. Therefore, the opportunity to explore the use of this software in more various, peculiar context is wide-open.

Mental workload is also a popular topic in terms of the construct to be measured by utilizing the software. Studies related to mental workload measurement utilize MATB for generating multitasks with different levels of difficulty or complexity. The tasks act as a "source" of mental workload to be induced to participants. Mental workload is then attempted to be measured by using either psychophysiological (e.g., EEG, ECG, respiration rate, eye-tracking) or subjective methods (mostly NASA-TLX as it is attached to the software), usually along with performance indices generated by the software such as reaction time or tracking deviation. In terms of psychophysiological methods, some popular and emerging devices such as functional near-infrared spectroscopy (fNIRS) or transcranial Doppler ultrasound (TCD) did not appear in our search results, thus matching those devices in a study using MATB seems to be prospective. In addition, it seems to be potential to apply qualitative analysis of workload measurement through, for example, deep interview or even ethnographic studies, though that might cost considerable amount of time.

Considering individual differences in a study using MATB seems to be methodologically reasonable and practically wide-open. There was a small number of studies that involved individual differences, particularly as an independent variable or as a mediating role of the variables such as personality states and traits [16] working memory capacity [34]. Another psychological aspect e.g., self-control, intelligence, or coping strategies or demographic aspect e.g., cultural background, age, or sex might have influence on cognitive performance in multitasking environment.

Even though the software could be easily used by anyone, the use of special population, however, may become essential to increase ecological validity in certain studies, depending on the purpose of the study. Studies mental workload in the context of aviation might be more appropriate if employing aviators since the generalization of this kind of study is obvious. Nevertheless, hiring common people as participants [7], [11], [27], [31] is sometimes more practical due to, for instance, resource availability.

Although it seems there is no agreed rule of thumb on which subtasks of the software should be used for a study, communicating the rationale behind utilization certain or all subtasks is essential. Providing evidence or practice from similar previous studies, if any, is possibly the best way to increase validity of the study. The current version of the software (MATB-II) would be seemingly the only choice for future studies as only this version is available in NASA's MATB website (https://matb.larc.nasa.gov/), unless the older version is still retained privately. Still, apparently there is no strong reasons to apply the older one while the newer version with better user-interface and configuration, which might affect study validity and reliability, is publicly available.

Some criticism regarding the software inevitably exists. A study from [4] revealed that when applied to pilot population as a primary task, MATB is considered unrealistic in several way compared to traditional flight task generated by flight simulator. Besides, TRACK and SYSMON subtasks are considered less interesting and engaging compared to two other subtasks, making performance indices from those subtasks could be falsely impressed [11]. Using MATB to distinguish workload level by varying the number of tasks to perform in parallel gives the impression that only participant's divided attention does matter, whereas working memory load and executive functions seem to be essential in workload measurement [26]. Meanwhile, a study from [9] assumed that high demand and sustained performance in MATB tasks are also considered less powerful to provoke expected distress.

4 Conclusion

Our scoping review reveals that MATB is a useful software for generating mainly a multitasking environment, with aviation is a field that has utilized it at most. The software was also heavily employed for mental workload studies primarily by creating different set of stimuli that eventually generate different levels of task demand or difficulty. Several operational considerations and criticism exist and thus must be taken care by researcher in using MATB.

References

1. Comstock, J. R. & Arnegard, R. J. *The Multi-Attribute Task Battery for Human Operator Workload and Strategic Behavior Research*. https://matb.larc.nasa.gov/files/2014/03/Comstock-Arnegard-Original-TM.pdf (1992).
2. Santiago-Espada, Y., Myer, R. R., Latorella, K. A. & Comstock, J. R. *The Multi-Attribute Task Battery II (MATB-II) Software for Human Performance and Workload Research: A User's Guide*. https://matb.larc.nasa.gov/files/2014/03/NASA-TM-2011-217164.pdf (2011).
3. Sharples, S. & Megaw, T. Definition and measurement of mental workload. in *Evaluation of Human Work Fourth Edition* (eds. Wilson, J. R. & Sharples, S.) 515–548 (CRC Press, Boca Raton, FL, 2015).
4. Bliss, J. P. Alarm reaction patterns by pilots as a function of reaction modality. *International Journal of Aviation Psychology* **7**, 1–14 (1997).
5. Caldwell, J. A. & Ramspott, S. Effects of task duration on sensitivity to sleep deprivation using the multi-attribute task battery. *Behavior Research Methods Instruments & Computers* **30**, 651–660 (1998).
6. Wilson, G. F. & Russell, C. A. Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors* **45**, 635–643 (2003).
7. Valk, P. J. L., Van Roon, D. B., Simons, R. M. & Rikken, G. Desloratadine shows no effect on performance during 6 h at 8,000 ft simulated cabin altitude. *Aviation Space and Environmental Medicine* **75**, 433–438 (2004).
8. Fairclough, S. H., Venables, L. & Tattersall, A. The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology* **56**, 171–184 (2005).
9. Fairclough, S. H. & Venables, L. Prediction of subjective states from psychophysiology: A multivariate approach. *Biological Psychology* **71**, 100–110 (2006).

10. Phillips, C. A., Repperger, D. W., Kinsler, R., Bharwani, G. & Kender, D. A quantitative model of the human-machine interaction and multi-task performance: A strategy function and the unity model paradigm. *Computers in Biology and Medicine* **37**, 1259–1271 (2007).
11. Wilson, G. F., Caldwell, J. A. & Russell, C. A. Performance and Psychophysiological Measures of Fatigue Effects on Aviation Related Tasks of Varying Difficulty. *The International Journal of Aviation Psychology* **17**, 219–247 (2007).
12. Brannon, N. G., Koubek, R. J. & Voss, D. Mechanisms of knowledge degradation in a resource management task. *Theoretical Issues in Ergonomics Science* **9**, 25–44 (2008).
13. Hardy, D. J. & Mitrovich, A. K. Preliminary Examination of Timesharing in United States Air Force ROTC Cadets. *Perceptual and Motor Skills* **107**, 21–28 (2008).
14. Miyake, S. *et al.* Physiological responses to workload change: A test/retest examination. *Applied Ergonomics* **40**, 987–996 (2009).
15. Valk, P. J. L. & Simons, M. Effects of loratadine/montelukast on vigilance and alertness task Performance in a simulated cabin environment. *Advances in Therapy* **26**, 89–98 (2009).
16. Carlozzi, N. E. *et al.* Personality and Reaction Time after Sleep Deprivation. *Current Psychology* **29**, 24–33 (2010).
17. Takae, Y. *et al.* Effects of Alcohol Intoxication on a Multi-task Simulator Operation Performance. *Review of Automotive Engineering* **31**, 150–155 (2010).
18. Lopez, N., Previc, F. H., Fischer, J., Heitz, R. P. & Engle, R. W. Effects of sleep deprivation on cognitive performance by United States Air Force pilots. *Journal of Applied Research in Memory and Cognition* **1**, 27–33 (2012).
19. Wang, Z., Hope, R. M., Wang, Z., Ji, Q. & Gray, W. D. Cross-subject workload classification with a hierarchical Bayes model. *Neuroimage* **59**, 64–69 (2012).
20. Chiappe, D., Conger, M., Liao, J., Caldwell, J. L. & Vu, K.-P. L. Improving multi-tasking ability through action videogames. *Applied Ergonomics* **44**, 278–284 (2013).
21. Hsu, B.-W., Wang, M.-J. J., Chen, C.-Y. & Chen, F. Effective Indices for Monitoring Mental Workload While Performing Multiple Tasks. *Perceptual and Motor Skills* **121**, 94–117 (2015).
22. Borghini, G. *et al.* Quantitative Assessment of the Training Improvement in a Motor-Cognitive Task by Using EEG, ECG and EOG Signals. *Brain Topography* **29**, 149–161 (2016).
23. Freiberger, J. J. *et al.* Assessment of the interaction of hyperbaric N-2, CO2, and O-2 on psychomotor performance in divers. *Journal of Applied Physiology* **121**, 953–964 (2016).
24. Gutzwiller, R. S., Wickens, C. D. & Clegg, B. A. The Role of Time on Task in Multi-task Management. *Journal of Applied Research in Memory and Cognition* **5**, 176–184 (2016).
25. Nelson, J. *et al.* The Effects of Transcranial Direct Current Stimulation (tDCS) on Multitasking Throughput Capacity. *Frontiers in Human Neuroscience* **10**, 1–13 (2016).
26. Roy, R. N., Bonnet, S., Charbonnier, S. & Campagne, A. Efficient Workload Classification based on Ignored Auditory Probes: A Proof of Concept. *Frontiers in Human Neuroscience* **10**, 1–12 (2016).
27. Valk, P. J. L., Simons, R., Jetten, A. M., Valiente, R. & Labeaga, L. Cognitive Performance Effects of Bilastine 20 mg During 6 Hours at 8000 ft Cabin Altitude. *Aerospace Medicine and Human Performance* **87**, 622–627 (2016).

28. Wickens, C. D. *et al.* Time Sharing Between Robotics and Process Control: Validating a Model of Attention Switching. *Human Factors* **58**, 322–343 (2016).
29. Hefron, R. G., Borghetti, B. J., Christensen, J. C. & Kabban, C. M. S. Deep long short-term memory structures model temporal dependencies improving cognitive workload estimation. *Pattern Recognition Letters* **94**, 96–104 (2017).
30. Kim, J. H. & Yang, X. Applying fractal analysis to pupil dilation for measuring complexity in a process monitoring task. *Applied Ergonomics* **65**, 61–69 (2017).
31. Nixon, J. & Charles, R. Understanding the human performance envelope using electrophysiological measures from wearable technology. *Cogn Tech Work* **19**, 655–666 (2017).
32. Wang, P., Fang, W. & Guo, B. A colored petri nets based workload evaluation model and its validation through Multi-Attribute Task Battery-II. *Applied Ergonomics* **60**, 260–274 (2017).
33. Hefron, R., Borghetti, B., Kabban, C. S., Christensen, J. & Estepp, J. Cross-participant EEG-based assessment of cognitive workload using multi-path convolutional recurrent neural networks. *Sensors* **18**, 1–27 (2018).
34. Liu, S. & Nam, C. S. Quantitative modeling of user performance in multitasking environments. *Computers in Human Behavior* **84**, 130–140 (2018).
35. Miller, J. *et al. An Updated Version of the U.S. Air Force Multi-Attribute Task Battery (AF-MATB)*. http://www.dtic.mil/docs/citations/ADA611870 (2014).