On the use of secondary tasks in addition to other objective measures of pilot workload

By Jorg O. ENTZINGER, Tsuneharu UEMURA, and Shinji SUZUKI[1)]

[1)]Department of Aeronautics and Astronautics, The University of Tokyo, Tokyo, Japan

Secondary tasks can be used to measure how much effort a given primary tasks requires from the pilot. However, there are many different types of secondary tasks, each with their own merits and demerits. In this paper we discuss a number of commonly used secondary task types and their limitations. We will specifically look at problems that may arise when combining secondary task measurements with other objective measures of pilot workload, including psychophysiological data such as heart rate and heart rate variability. We discuss our experiences with various types of secondary tasks in experiments where subjects flew our fixed-base flight simulator.

Key Words: Objective workload measurement, secondary task, human-machine interaction, human pilot

1. Introduction

Secondary task performance has been used widely in cognitive neuroscience as a measure of the cognitive effort (attention, workload) a subject puts into the primary task. The secondary task technique builds on the assumption that there is a limit to the subject's cognitive capacity, and that an increased cognitive requirement of the primary task leaves less capacity remaining for the secondary task. This assumption may not always hold true [1, 2]. In addition, the use of secondary tasks may significantly interfere with other objective measures of workload, in particular with psychophysiological measures such as heart rate, brain waves, and eye tracking data.

However, secondary tasks can provide valuable data if the type of secondary task is chosen carefully [3], considering both the primary task characteristics and interference with other measurements. Minimal secondary tasks may even be a simple and practical alternative for more elaborate psychophysiological measurements if only roughly discriminative measures are sufficient.

In this paper we introduce a number of common secondary tasks and analyze their merits and demerits. We will focus on their application in a simulated instrument flight task, and in combination with psychophysiological measurements. Finally, we share some lessons we learned from actual experiments that may benefit the design and setup of future experiments.

2. Types of Secondary Tasks

In principle, any task can be used as a secondary task. The only requirement would be that the task has a clearly measurable result, which preferably indicates the quality of task performance (score). Below we will introduce a number of commonly used task types.

2.1. Overloading

Typical secondary tasks used to evaluate a human pilot's workload (or actually spare capacity) are complex processing tasks, designed to be difficult (but doable) even as a single task. Adding such a secondary task to the primary task is expected to overload the subject, as illustrated in Fig. 1, and the performance on (score of) the secondary task is expected to degrade with increasing primary task difficulty. This heavily relies on the assumptions of single channel processing and fixed total capacity, as explained in the following section. Examples are memory tasks (such as the Sternberg task [4] or the n-back task [5]; Fig. 2), mental mathematics, or information communication & processing tasks.

2.2. Simple

Another commonly used type of secondary task requires the subject just has to press a button when he notices an external stimulus (e.g., hears a specific sound or sees a specific visual cue) presented at random time intervals [6]. This method assumes that the subject's response will be slower (longer reaction time) when his workload is higher. Such tasks generally require very little central processing effort and memory.

A variation that requires just a little bit of central processing is the choice reaction time, where one of 2 slightly different stimuli is presented at random, and the subject has to press a different button depending on which of the 2 stimuli he perceived. Examples of the simple and simple choice task procedures are shown at the bottom of Fig. 3.

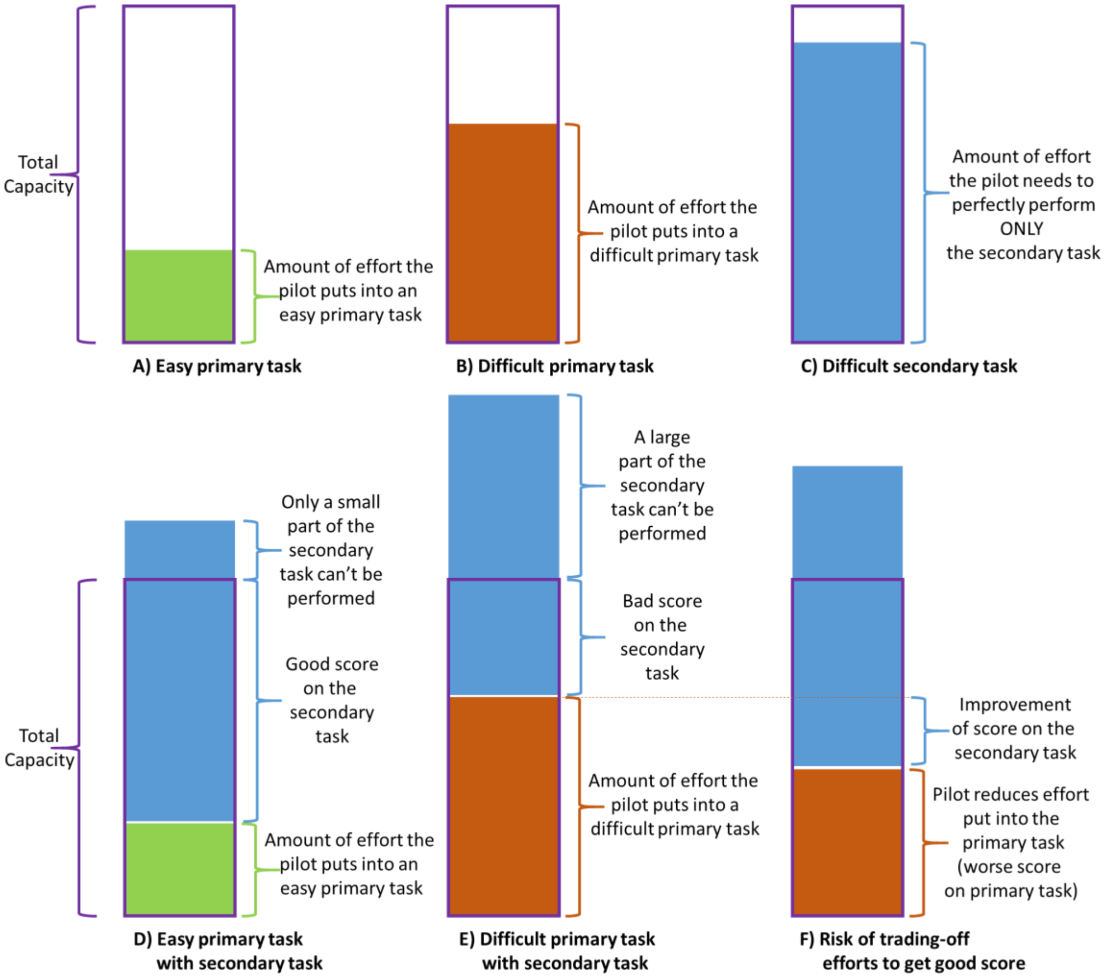

Fig. 1. Illustration of how a difficult secondary task can be used to measure the amount of effort a subject puts into the primary task, assuming the single channel capacity model.

2.3 Continuous

Tracking or monitoring tasks require continuous attention. In tracking task the subject has to operate a device to follow a guidance signal. Tracking tasks may require varying amounts of central processing effort. Monitoring tasks focus on the perception, rather than on control. The subject has to filter perceived information to find specific cues obscured by "noise". These tasks assume that tracking performance or detection rate of the secondary task degrades when the effort required for the primary task increases.

2.4 Time Estimation

Time estimation tasks assume that subjective perception of time changes with workload [7]. For instance, when waiting for 1 minute, it feels like 5 minutes, whereas "time flies when you're having fun". The subject can be ask to create a regular beat (which is likely to become irregular due to workload changes), or the subject can be asked how much time he thinks has passed during a specific interval.

3. Limitations of Secondary Tasks

3.1. Single Channel Processing and Fixed Total Capacity Assumptions

Using secondary tasks for workload measurement only works under the assumptions of single channel processing and fixed total capacity, meaning that additional effort put into one task will lead to a reduction of effort put into another task (especially when the subject is using his full processing capacity). As mentioned in the introduction, this may not always be the case [1, 2].

For instance, remembering a shopping list requires a certain amount of mental capacity, and each additional item will reduce the amount of spare capacity. If we need to remember more than 7 items, each additional item will increase the probability that we forget some of the original items (reduced main task performance). However, if the additional (secondary) task is not remembering additional shopping list items, but walking to the supermarket, this task doesn't influence remembering the shopping list items (primary task performance), because walking requires a different type of effort/capacity ("channel").

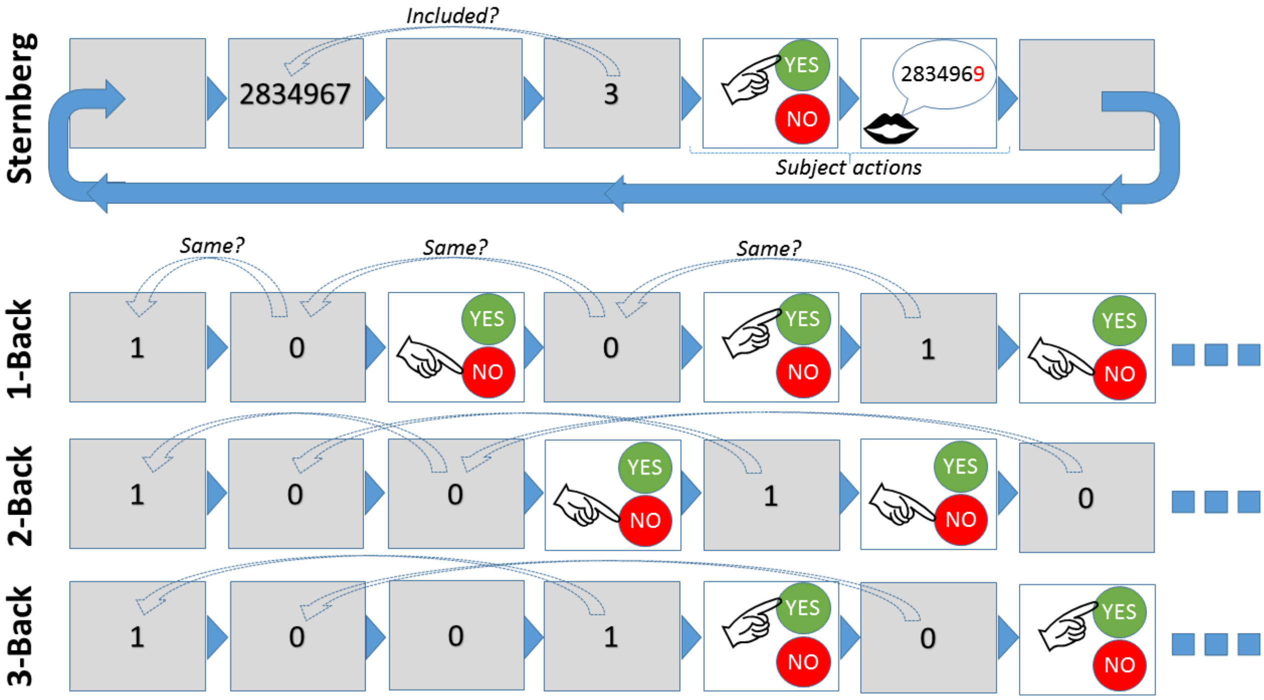

Fig. 2. Illustration of the displays and subject's required actions for some common "overloading type" secondary tasks. For all tasks the response time from the last display change until the subject's button press is recorded. In case of multiple input options, correct/incorrect score is also recorded. For the Sternberg task, the number of mistakes in reading back the original digit series is also recorded.

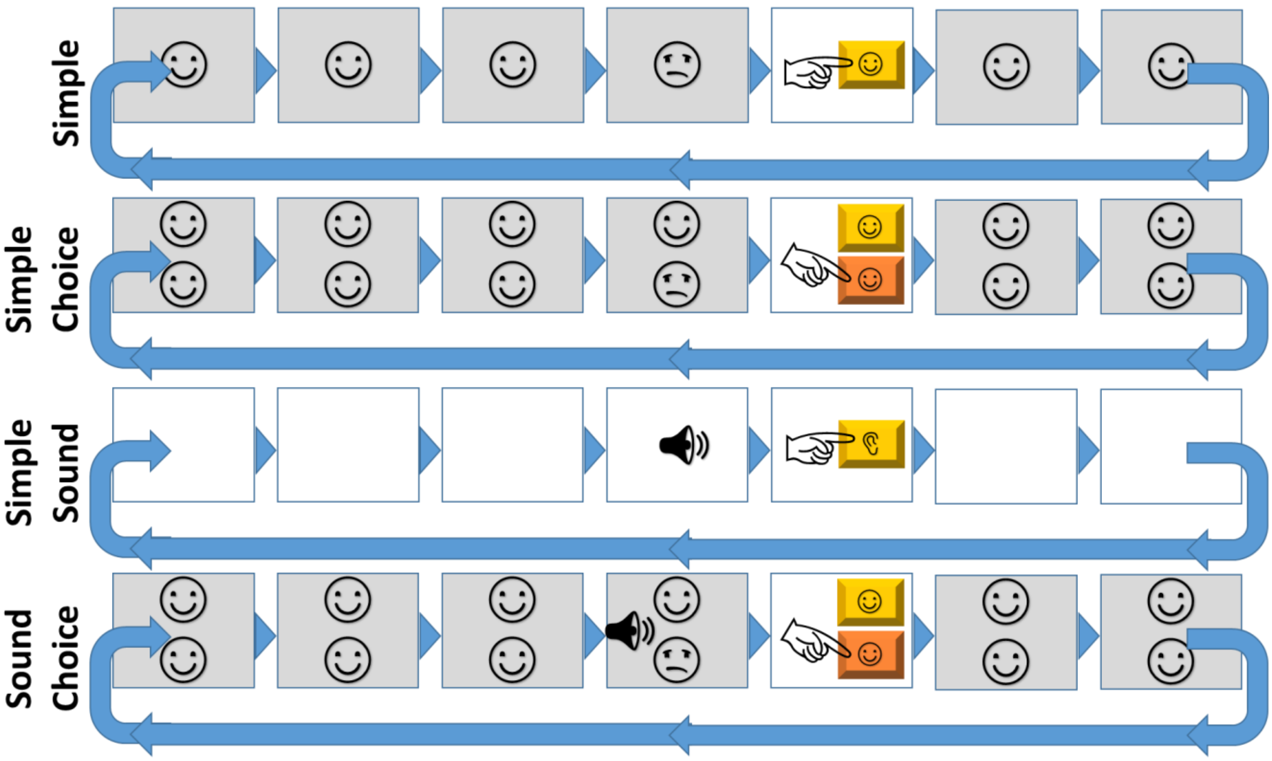

Fig. 3. Illustration of different variations of "simple" response time secondary tasks. Task stimuli may be presented in a different modality (e.g., auditory or tactile instead of visually). For all tasks the response time from the last display change until the subject's button press is recorded. In case of multiple input options, correct/incorrect score is also recorded. For the Sternberg task, the number of mistakes in reading back the original digit series is also recorded. The "Sound choice" task displayed here adds the a sound just to warn the subject that a change has happened. The visual "simple choice" task could also be implemented with sounds only, where the subject has to differentiate between high and low pitch sounds.

The total capacity (information processing capacity, memory capacity, physical movement capacity, etc.) may also not be constant. Specifically, it may change with motivation, stress, and tiredness. We constantly manage our workload internally, and if we feel a specific need, motivation, opportunity for quick gain, we may temporarily increase our capacity to a non-sustainable working level. Such fluctuations will affect the secondary task results, but are typically ignored.

3.2. Interference with Primary Task

A large risk is that the secondary task may interfere with the primary task. This is illustrated in Fig 1F. The basic idea is that the primary task should always be prioritized, and the secondary task should only be executed to the degree of spare capacity the subject has at that time. However, the secondary task may be distracting, or the subject may manage his workload differently when he has to perform a secondary task. Specifically, the subject may feel that a slight decrease of performance on the main task can get him a large performance gain on the secondary task, and therefore shift his priorities. Although undesirable, and against the clear instructions (to prioritize the primary task) that should always be given, it is only human to manage our own workload and show our best (overall) performance.

3.3. Artificiality

Secondary tasks are typically artificial tasks, which the subject is not used to. The secondary task may therefore require some training or practice to get used to it.

In some cases, it is possible to use an "embedded" task, for example, when the primary task is to control an aircraft in a landing approach, carrying out a checklist or communicating with the control tower could be used as an embedded secondary task. However, the risk of using embedded tasks is that they could easily be considered as an inherent part (sub-task) of the primary task, and therefore receive a higher priority than just using up "spare capacity".

4. Lessons Learned from Workload Evaluation Experiments

Recently, there has been a significant increase in the use of psychophysiological measures such as heart rate, brain waves, and eye tracking data for the objective evaluation of workload, mainly due to the fact that the necessary equipment has become much cheaper. In this section we discuss our experiences combining various types of secondary tasks with psychophysiological data for the evaluation of workload.

4.1. Experiment Overview

We carried out various experiments with secondary tasks in the fixed based training device (hereafter "simulator") owned by the Suzuki-Tsuchiya laboratory at The University of Tokyo (Fig. 4). Various subjects took part in these experiments, including students of the Aeronautics & Astronautics Department, researchers, and experienced pilots. During the experiments we recorded the aircraft state, pilot's control inputs, secondary task performance, and psychophysiological data [8, 9]. We used and the Takei TalkEyeLite eye-mark camera (to record the gaze direction, pupil diameter, and blinks), the ParamaTech EP-301 portable ECG recorder (for heart rate and heart rate variability), and eMotiv EPOC+ EEG headset (for brainwaves).

Fig. 4. Simulator of the Suzuki-Tsuchiya laboratory at The University of Tokyo. Left: overview. Middle: Sternberg 7 digits secondary task and response switches (Yellow & Orange) attached to the control column. Part of the primary flight display is covered for experiment's primary task purposes. Right: Eye mark camera, Portable ECG device, and brainwave measurement headset.

4.2. Overload Tasks (e.g., Fig. 2)

Combining psychophysiological measurements with a secondary task poses some significant problems, because the effort a pilot puts into the secondary task will influence the psychophysiological data. This means that the commonly used "overload" types of task will be particularly unsuitable, since they will always use up all the pilot's capacity. Therefore, we won't be able to measure the psychophysiological effects of (only) the primary task anymore (unless we would run separate experiments). In our experiments, we noticed clear differences between trials with and without overload type secondary tasks in heart rate variability and some brainwave indices.

We therefore propose the use of minimal secondary tasks, such as a "simple" reaction time task or a "simple choice" reaction time task (fig. 3). Although not as comprehensive as the complex processing tasks, these minimal tasks are easier to combine with the psychophysiological measurements, which in turn will provide the more detailed information about the degree to which the pilot is using his internal processing resources.

4.3. Visual-only simple tasks (e.g., "Simple" and "Simple Choice" in Fig. 3)

Visual secondary tasks will require the pilot to change his scan-pattern (i.e., order & timing of looking at various cockpit displays & outside visual scene) to include the secondary task display. This may be minimized by positioning the display close to the primary flight display and making changes very salient, so that they may be perceived in peripheral vision, rather than requiring foveal attention.

The fact that pilot will need to look at the display to perform the task may actually be beneficial. A slower response of the pilot may indicate that he has high workload, or at least less spare visual scanning capacity. Remembering the flaws of the single channel processing assumption mentioned in section 3.1, it may be well possible that a pilot has spare memory capacity, but no spare visual perception capacity. Therefore, if we are particularly interested in visual scan patterns or the way visual information is acquired (e.g., when evaluating the design of cockpit displays), a visual secondary task may provide useful data.

A drawback is that, when using an eye mark camera to record gaze direction and fixations, the secondary task will undoubtedly influence the reliability or usefulness of such data.

4.4. Sound-only tasks (e.g., "Simple Sound" in Fig. 3)

Sound-only secondary tasks have a very minor influence on the primary task. The human brain can process sound much faster than visual information, and unlike visual information, we don't have to pay specific attention to perceive sound. This also means that reaction times are typically much shorter and show less variation, so they should be measured more accurately.

One interesting result from our experiments is that we actually become unable to perceive sounds in extremely high workload conditions. For example, when making landing approaches, the flare maneuver, which is critical and carried out just seconds before touchdown requires a pilot's full attention. At this time, our subjects didn't notice the sounds, and generally for all sounds presented from a few seconds before until a few seconds after the touchdown, the responses timed out. This indicates that rather than measuring response time to evaluate spare capacity on a continuous scale, measuring timeouts in a simple sound tasks could be useful as a binary evaluation of whether or not there is any remaining capacity at all.

4.5. Visual & Sound Combined Tasks (e.g., "Sound Choice" in Fig. 3)

In an attempt to get the best of both, we developed a combined visual & sound task, indicated as "Sound Choice" in Fig. 3. The task is identical to the visual-only simple choice task, but at the moment a new choice is presented, a sound will be played as well. This should remove the necessity to continuously scan the secondary task screen for changes. At the same time, the choice element will still make it necessary to check the screen, rather than immediately responding to the perceived sound (as would be the case in the sound-only task). We assume that the pilot postpones the checking of the secondary task screen for a longer time when his remaining (visual perception) capacity is low. We could therefore use the pilot's response time as a continuous measure of remaining (visual perception) capacity, while of course any timeouts would still be useful to know no capacity remained at all.

4.6 General experiences

Even though we always instruct subjects that the primary task is absolutely first priority, and the secondary task should only be carried out if and to the extent at which they have spare capacity, we found that different subjects treat each task differently.

First of all, there was a difference between inexperienced pilots (students), trained students, and experienced (private and professional) pilots. We would expect experienced pilots would have more spare capacity, and therefore shorter response times or less timeouts. However, we found that the least experienced students generally scored better on the secondary task. We think this can be explained by the task interference (mentioned in section 3.2). The more experienced the subject, the easier it is for them to ignore the secondary task in high-workload situations, and focus on the primary task. Subjects who are not comfortable yet with the primary task probably find it difficult to focus when needed, and give up on it in favor of the secondary task (Fig. 1F).

Even between professional pilots, there are personal differences in how they deal with the secondary task. Comparing the visual simple choice task with and without additional sound, some find it easier to "not scan the visual only task" than to "not immediately check it when they hear a sound" if they have no spare capacity, for others it is exactly the opposite. This might be a difference in general between people who prefer finishing simple tasks first, then focus on difficult things, or people who prefer to do the difficult things first, then the easy ones.

5. Conclusion

While secondary tasks may provide objective data that can be useful for workload evaluation, it is important to carefully consider the type of task to use. We introduced various task types and discussed their merits and demerits, as well as the general assumptions on which the secondary task method is based.

The main factors that play a role when choosing the type of secondary task in an experiment are:

- The type of primary task (mainly visual, mainly information processing, mainly memory, …)
- Risk of interference with the primary task (single channel assumption, experience level with primary task, personality type)
- Risk of interference with other measurements (psychophysiological measurements)
- Type & quality of data needed (continuous scale remaining capacity, or only timeouts)
- Amount of preparation/post-processing required (hardware, software implementation, carrying out, data analysis)

We found that subjects may trade off the efforts spent on the primary and secondary tasks to some extent, even when they are specifically instructed to fully prioritize the primary task. Therefore, in addition to choosing a task carefully, it is important to check for deterioration of primary task performance when a secondary task is present.

An interesting finding is that a minimal sound-only task can at least indicate moments when zero capacity is left, and can therefore constitute a baseline for the interpretation of the psychophysiological data.

We are currently gathering more data with various subjects to quantify the experiences shared in this paper.

Acknowledgments

The authors like to thank the subjects for their participation in the experiments, and their input in the discussions about the validity of secondary tasks.

References

1) R.S. Owen, "Clarifying the Simple Assumption of the Secondary Task Technique", *NA - Advances in Consumer Research, Volume 18*, eds. Rebecca H. Holman and Michael R. Solomon, Provo, UT : Association for Consumer Research, 1991, Pages: 552-557. http://acrwebsite.org/volumes/7216/volumes/v18/NA-18
2) J.B. Isreal, G.L. Chesney, C.D. Wicken, E. Donchi, "P300 and tracking difficulty: evidence for multiple resources in dual-task performance", *Psychophysiology*, Vol. 17, Nr. 3, (May 1980). Pages: 259-273. http://dx.doi.org/10.1111/j.1469-8986.1980.tb00146.x
3) K.R. Boff & J.E. Lincoln (eds.), *Engineering Data Compendium Human Perception and Performance, Volume II*, Harry G. Armstrong Aerospace Medical Research Laboratory Wright-Patterson Air Force Base, Ohio, 1988. (Section 7.719~).
4) S. Sternberg, "High-Speed Scanning in Human Memory", *Science, New Series*, Vol. 153, No. 3736 (Aug. 5, 1966). Pages: 652-654. http://www.jstor.org/stable/1719418
5) A.F. Monk, D. Jackson, D. Nielsen, E. Jefferies & P. Olivier "N-backer: An auditory n-back task with automatic scoring of spoken responses", *Behavior Research Methods*, Vol. 43, Nr. 3, 2011. Pages: 888-896. http://dx.doi.org/10.3758/s13428-011-0074-z
6) Robert J. Kosinski, *A literature review on reaction time*, working paper, Clemson University, 2012. (Available from: http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/biae.clemson.edu/bpc/bp/Lab/110/reaction.htm or http://homepage.univie.ac.at/andreas.franz.reichelt/intro2cogsci2/data/literature_review_reaction_time.pdf).
7) Scott W. Brown, "Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks", *Perception & Psychophysics*, Vol. 59, Nr. 7, 1997. Pages: 1118-1140
8) J.O. Entzinger, T. Uemura, S. Suzuki, "Mental Effort and Safety in Curved Approaches", 29th Congress of the International Council of the Aeronautical Sciences (ICAS2014), Sept. 2014, St. Petersburg, Russia. http://repository.dl.itc.u-tokyo.ac.jp/dspace/handle/2261/56357
9) J.O. Entzinger, T. Uemura, T. Iijima, N. Matayoshi, and S. Suzuki, "Objective Evaluation of Pilot Operations in Wind Shear Using Psychophysiological Measurements", *30th congress of the International Council of the Aeronautical Sciences (ICAS2016)*, Daejeon, South Korea, Sept 25-30, 2016.