Mental Workload in Aircraft and Simulator during Basic Civil Aviation Training

Nicklas Dahlstrom
Lund University School of Aviation

Staffan Nahlinder
Swedish Defence Research Agency

Abstract

This study investigated mental workload in basic civil aviation training. Heart rate, eye movement and subjective ratings from eleven students were collected during simulator and aircraft sessions. Results show high correspondence in psychophysiological reactions between the sessions. For some flight segments heart rate was consistently lower in the simulator, suggesting higher mental workload in the aircraft. Differences in heart rate during rejected take-off and engine failure indicate that the increase of workload starts in advance of an “unexpected“ event in the simulator and seem to be of preparatory nature, while more connected to management of the situation in the aircraft.

Descriptors: psychophysiology, mental workload, aviation training, flight simulation, learning transfer

Introduction

Measurement of pilot mental workload has been recognized as an important method for understanding the dynamic and complex cognitive demands of flying (Roscoe, 1992, Wickens & Hollands, 2000; Wilson, 2001, 2002c). Due to the multifaceted nature of flight several psychophysiological methods have been developed for this purpose and used for investigation and evaluation of pilot performance, execution of missions and cockpit design (Wilson, 2001; Magnusson, 2002). However, when it comes to the foundation for the performance of civilian pilots, basic civil aviation training, little has been done to validate or use these methods to understand and improve training (Dahlstrom, 2002). Measurement of mental workload can be used to investigate cognitive demands of specific segments of flight during basic civil aviation training and the transfer between training in simulator and aircraft. In the light of recent international regulatory initiatives to significantly expand the use of flight simulation in the basic training of airline pilots, such research is essential to ensure that the training process effectively produces competent operators for the aviation industry (Dahlstrom, 2005).

Psychophysiological measurements provide objective data on the cognitive demands of flying (Wilson, 2002b). As recording equipment has become increasingly reliable and non-intrusive, this has facilitated user acceptance and increased the potential to collect in-flight data (Wilson, 2001). Heart rate has been frequently used to study the effects of flight on mental workload (Roscoe, 1992; Bonner & Wilson, 2002; Veltman, 2002; Wilson, 2002a; Wilson, 2002b) and proved a sensitive indicator; approach and landing segments have been shown to cause an increased heart rate (Wilson, 2002a; Wilson, 2002c). Other methods that have been used include heart rate variability (HRV), respiration, electrodermal activity (EDA), electroocular activity (EOG) and brain activity (EEG) (Veltman, 2002; Wilson, 2001; Wilson, 2002a; Wilson, 2002b). None of these measures have however been used and validated in aviation to the extent of heart rate, and HRV has been argued to be particularly useful in the flight environment (Wilson, 2001; Wilson, 2002a). Methods based on subjective reporting (Cooper-Harper, NASA-TLX etc.) have however been dominant (Roscoe, 1992) and psychophysiological methods have not been more than marginally used for research on basic civil aviation training or even on training in other fields. Sohn and Jo (2003) used personality tests, heart rate data and NASA-TLX to assess mental workload for studies of ideal composition of instructors and student pilots. In maritime transport Murai and Hayashi (2005) used heart rate data to study the performance of bridge staff on training ships.

Although technological advances have made recording equipment increasingly portable and reliable (Dahlstrom & Nahlinder, 2006), gaining access to collect psychophysiological data from simulated or aircraft flight remains a challenge for researchers. Due to safety and resource (cost, aircraft, pilots etc.) concerns, having aircraft deployed primarily for research purposes is uncommon and data collections normally has to be performed during flight operations planned for other purposes (Wilson, 2002a). Gaining access to simulators for research is often as difficult. According to Salas, Bower and Rhodenizer (1998): “Simulators are typically booked for training and practice continously, not leaving time for research or other experimental purposes.” (p. 201). For use of any recording equipment the primary criteria is that it is non-intrusive and in no way interferes with safety or pilot performance (Wilson, 2001). For basic training even a remote risk of recording equipment interference can render student pilots and flight instructors reluctant to participate. Beyond this safety concern, the pressure to use valuable flight time (paid for by the student pilot) effectively to complete a detailed regulatory syllabus increases reluctance to accept any risk of in-flight deviations or disturbances.

While psychophysiological data have proven valuable to detect standard segments of flight (e.g. take-off, landing) with elevated levels of mental workload (Wilson, 2002a) more short-term effects connected to individual maneuvers have normally not been detected and analyzed. These short-term effects should be easier to detect when performed during training; the mental workload of a student pilot performing advanced maneuvers or practicing emergency maneuvers should be higher than that of an experienced pilot. This would enable study of shorter sequences of flight and potentially even provide data connected to individual maneuvers. Also, comparisons between mental workload experienced during simulator sessions and flight could then further increase the knowledge of the relation between these two forms of training. Understanding of this relation has become increasingly important as the new Multi-Crew Pilot License (MPL) was adopted by the ICAO in November 2006.

While simulation has a long tradition in aviation, its use in research has primarily been focused on issues of cockpit design and avionics as well as on crew fatigue, communication and other issues related to Crew Resource Management (Lee, 2005). The recent launch of a new form of basic pilot training in the form of the MPL signals a significant increase of the use of flight simulation in early stages of basic training (Forbes, 2005; Woods, 2005). This proposal has been drafted, presented and decided upon without any openly acknowledged involvement of the research community or even any references to research (Dahlstrom, 2005). This indication of disconnect between basic civil aviation training and research has highlighted that research in this field not only is an issue of great importance but also one of great urgency. An increasing reliance on simulation already at early stages of flight training demands that renewed and increased attention is given to potential differences in how basic training is experienced by student pilots in the aircraft and simulator respectively and how this may affect their competency to function as a safe and effective operator.

In a literature review Carretta & Dunlap (1998) found that only 13 out of 67 identified articles, technical reports and conference papers regarding simulator training and transfer from 1986 to 1997 were related to transfer of training from simulator to aircraft and concluded that "the literature suffers from the lack of true simulator-to-aircraft studies involving complex pilot skills" (p. 4). Regarding basic civil aviation training a positive transfer between simulation and aircraft has been shown based on subjective performance judgments rather than on psychophysiological data (Lintern, Roscoe, Koonce & Segal, 1990; Dennis & Harris, 1998; Lee, 2005). For more complex pilot skills Roessingh (2005) investigated transfer of aerobatic training between PC-based simulation and aircraft flight and found no transfer effects. Certainly, knowledge about basic flight training is to be found with flight instructors who would be able to provide performance judgments. However, due to the temporary nature of the flight instructor profession (many stay only long enough to build the number of flight hours required to apply to an airline) and small to medium enterprise character of flight training organizations (normally commercial and non-academic) this knowledge largely remains tacit and rarely becomes published or even exchanged with peers. Accordingly, input from the training industry to change basic civil aviation training has been minimal and changes have been slow and forced by the need to adapt to technological advances, such as the increasing presence of complex automation in modern transport aircraft (Dekker & Johansson, 2000; Dahlstrom, 2002).

The goal of this project has been to use psychophysiological data and subjective ratings to measure mental workload on student pilots for specific flight segments and use this for comparisons between flight training in simulator and aircraft. The aviation industry has to be able to produce safe pilots cost-effectively. Increased knowledge about the effectiveness of basic flight training has to be based on methods that produce reliable data to increase understanding of basic flight training in simulator and aircraft and the transfer between them. This knowledge can confirm or contest traditional flight instructor knowledge, be used to calibrate and improve flight training to mental workload demands and enhance effective use of simulated flight training.

Method

Participants

Eleven students (two female) and eight flight instructors at Lund University School of Aviation participated as subjects and observers in this study. The students were in the later stages of their 18-month Airline Transport Pilot (ATP) training program. Their total flight time ranged from 150 to 172 hours (mean=156, SD=6.8) and their age from 23 to 28 years (mean=24.8, SD=1.4). For the flight instructors their total flight time ranged from 1600 to 15400 hours (mean=4700, SD=4800), instructor time from 800 to 10 000 hours (mean=2960, SD=3060) and age from 36 to 63 years (mean=42, SD=19). (The large variation was due to that among the flight instructors four were of age 34 to 39 years while the other four were 55 to 63 years old.) All of the flight instructors were full-time employees at the flight school and trained and experienced as instructors on the specific flights and simulator sessions performed at this stage of the training program.

Aircraft and flights

The type of aircraft used at this stage of the training program was a twin engine propeller aircraft, the Piper Arrow 31 Navajo. The simulator sessions were performed in a Flight Navigation Procedure Trainer of category one (FNPT I) for the same aircraft type. This type of equipment is used primarily for practicing instrument flight procedures and does not provide simulation of external visual references or aircraft motion.

One of the flights and a simulator session from the training program, with corresponding content and anticipated workload profile, were chosen for this study. (Students perform simulator sessions prior to the corresponding training flights with an aircraft in the training program to let them practice certain flight procedures before the aircraft flight is performed.) The flights chosen for this study included anticipated peak workload elements for individual maneuvers (rejected take-off and engine failure), but also segments likely to represent sustained high workload (instrument approaches) as well as a segment with an anticipated low level of workload (the cruise segment of the flight).

Psychophysiological data collection

The recording equipment used in this study was a Vitaport II (Vitaport Temec Instruments BV, Gemert, the Netherlands). The Vitaport is a multi-channel recording device used for clinical recordings as well as for research purposes. In this study it was used on the student pilots to sample data for an electrocardiogram (ECG) by placing three electrodes vertically above each other on the sternum to minimize artifacts from muscle interference. The ECG signal was run through a software program that automatically detects the R-peak intervals of the QRS-complex. Heart rate was (HR) calculated by converting the time (in milliseconds) between two successive R-peaks (interbeat interval, IBI) to beats per minute. Additionally, data on electroocular activity (EOG) was collected from two electrodes placed above and next to the outer canthus (juncture of eyelids) of the right eye. The signal from these electrodes can provide data on vertical and horizontal eye movements and blinks. This data was used to calculate Eye Movement Energy (EME) a measure of the total amount of eye movement for a given time period. ECG and EOG signals were sampled at 256 Hz, with a low pass filter at 40Hz and a high pass filter at 3Hz.

Observations and questionnaires

A detailed observation form for registration of exact times of flight segments (represented by individual maneuvers) was constructed in cooperation with flight instructors. During the simulator sessions recording of times were performed by the authors. (The first author has about 200 hours of flight experience, primarily as a navigator from the Swedish Air Force, and the second author has 100 hours as licensed private pilot.) During all but one of the flights recording of times were performed by student pilots flying as passengers (and the remaining one by the second author). The students and instructors had received instruction on the procedures for recording of times and rating of mental workload. In addition, the students practiced these procedures on either two or three flights before the study commenced.

The form identified 15 distinct flight segments. These are presented below along with the average times for them as recorded during simulator sessions and flights in the study.

Table 1. The fifteen selected flight segments, including average times recorded and descriptions of respective segment. (For times marked with * see below.)

Number	Name	Time	Description
1	Line up	00:00:00	Aircraft is lined up and getting ready for take-off
2	T/O Brake release	00:01:20	Aircraft starts accelerating down the runway
3	Rejected T/O	00:01:40	Take-off cancelled and aircraft brakes on runway
4	T/O Brake release	00:04:00*	*See number 2 above*
5	Gear up	00:05:45	Gear is brought in
6	Engine failure	00:13:10*	Flight on one engine and execution of emergency checklist
7	Back on two engines	00:14:40	Both engines running normally again
8	Cruise	00:22:00*	Straight and level flight
9	Flap 7	00:25:00	Extend 7 degrees of flap, prepare for landing
10	Gear down	00:28:00	Gear is lowered
11	Go around	00:32:00	Landing is aborted, go-around procedure performed
12	Flap 7	00:59:00	*See nr 9*
13	Gear down	01:02:00	*See nr 10*
14	Full flaps	01:04:30	Flaps fully extended
15	Touch down	01:05:30	Touch down on runway

Directly after respective flight a questionnaire was used to collect student pilot ratings of their own mental workload. The questionnaire was constructed using a Likert scale from 1 to 9, with anchors selected for the end points of the scale (Nahlinder, Berggren & Persson, 2005).

Results

Data from three participants were excluded, one set of data from a session in the aircraft and two from simulator sessions, due to poor electrode connection. Data from eight individuals were analyzed and for each participant data from the simulator sessions were compared to data from aircraft sessions.

Regarding the recorded times of flight segments in table 1 above, three have been marked with an asterisk (*). For these, there were significant differences in the recorded times between the simulated and the aircraft flight. The second T/O Brake release occurred on average one minute later in the simulator than in the aircraft due to conversation between the student and the instructor after the Rejected take-off. The engine failure occurred earlier in the simulator (on average 20 seconds after “Gear up”) while it in the aircraft occurred at safe altitude during the Cruise (on average 8 minutes after “Gear up”). Also, the Cruise segment was longer in the simulator than in the aircraft.

Heart rate

Heart rates were calculated as average values of two-minute segments for the flight segments. The flight segments were represented by the recorded time of individual maneuvers. These times were used as centre points for the calculation of average heart rate for the segments. The Cruise segment was the only segment of flight not represented by an individual maneuver and the point in time chosen as centre point for it was at three quarters of the total cruise time.

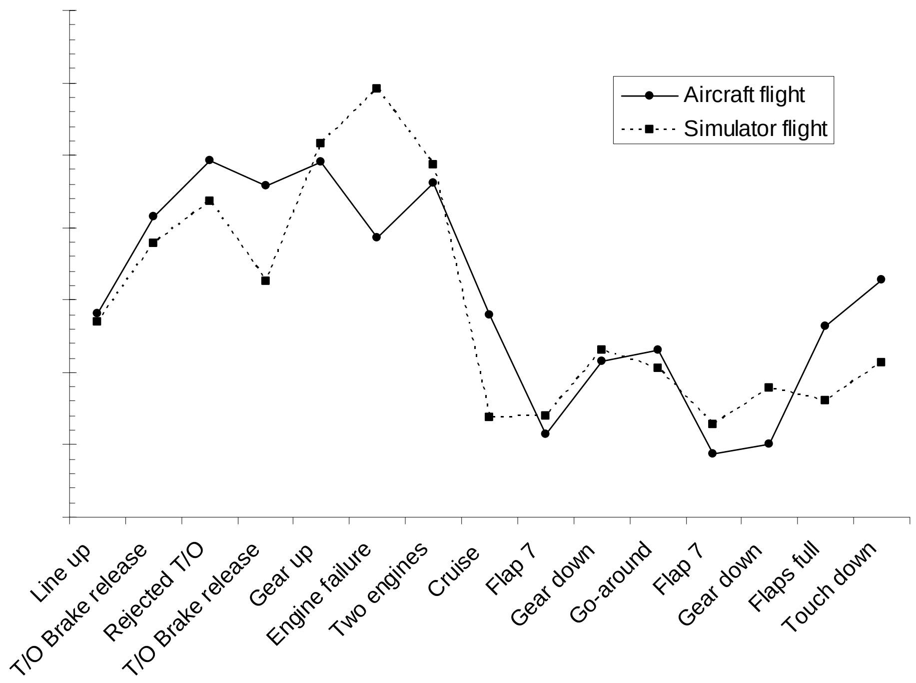

Figure 1. Average heart rate for 8 student pilots for the 15 selected flight segments.

The standardized and normalized heart rates for eight participants are presented in Figure 1. These were analyzed as a dependent variable in a repeated measures analysis of fifteen points in time (the selected flight segments) and two different conditions (aircraft and simulator). This analysis shows no significant main effect of type of flight. There is however a strong main effect of flight segment ($F[14, 84] = 13.457$; $p<.0005$). There is no effect of the interaction between type of flight and flight segment. This indicates that the two types of flight (in aircraft and simulator) can not be shown to be statistically different from each other.

Non-standardized and non-normalized heart rates are also presented below in detail for some of the flight segments. All of them are presented as group averages for two minutes of flight segments connected to specific events during the flight.

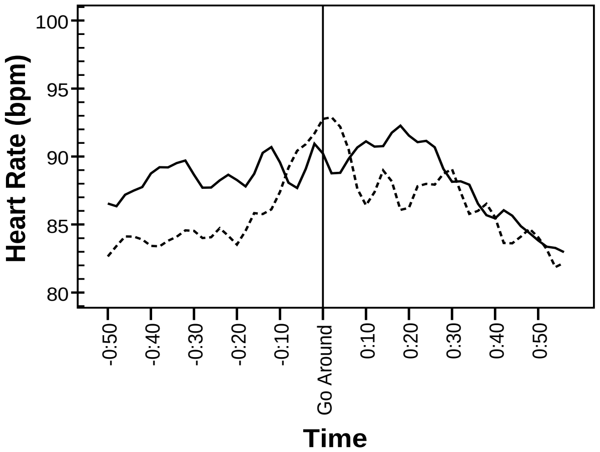

Figure 2. Average heart rate for 8 student pilots during the last two minutes of the first instrument approach. Solid line is aircraft flight, dashed line is simulator flight

For the last two minutes of the first instrument approach (Figure 2) heart rate was generally higher for aircraft flight. However, in the aircraft Go-around was planned for already ahead of or the approach. In the simulator it was called out by the flight instructor when deemed suitable and this “surprise effect” seems to have generated a peak in heart rate which caused average heart rate for the segments to become similar.

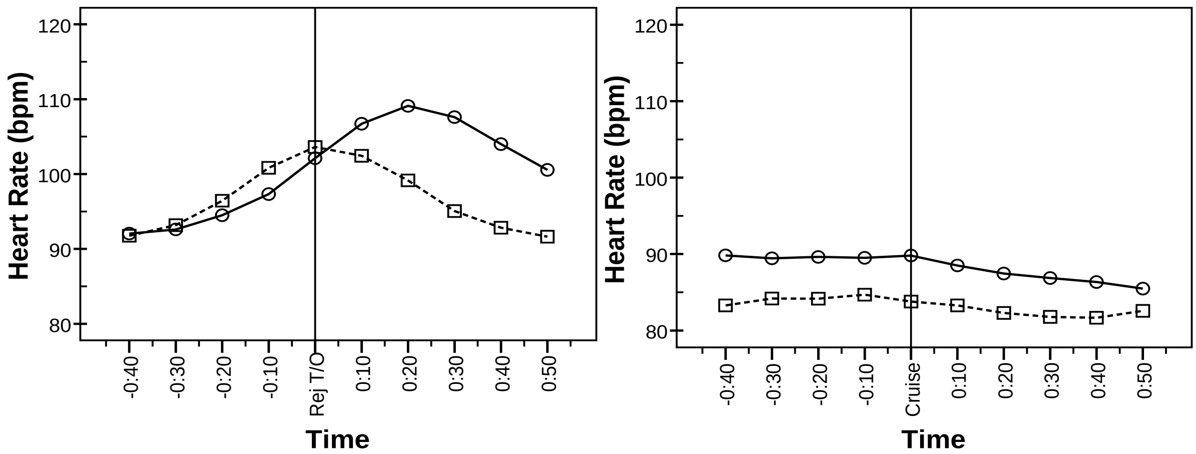

Figure 3. Average heart rate for 8 participants during the flight segment (a) Rejected take-off and (b) Cruise. Solid lines are aircraft flight, dashed lines are simulator flight

For the Rejected take-off segment (Figure 3a) there was no main effect of type of flight, but there was a main effect of time (F[9, 63]=7.645; p<.0005) and interaction between type of flight and time (F[9, 63]=5.413; p<.0005). Peak heart rate in the simulator was registered at the occurrence of the rejected take-off. In the aircraft, peak heart rate was registered after the occurrence as the student pilots tried to manage braking to a full stop.

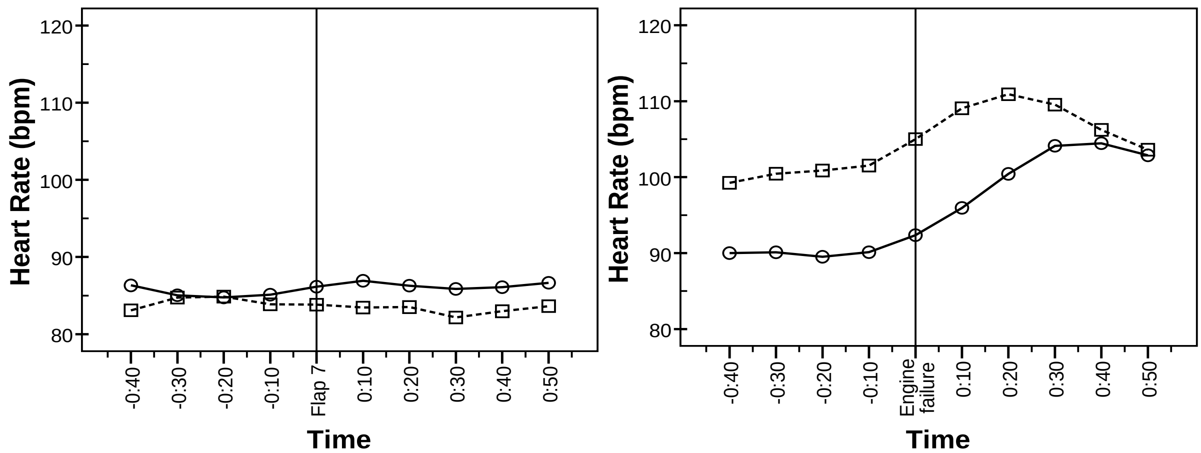

Figure 4. Average heart rate for 8 student pilots during the flight segments (a) Flap 7 and (b) Engine failure. Solid line is aircraft flight, dashed line is simulator flight

There was no main effect of type of flight or time for the Cruise or Flap 7 (Figure 3b and 4a) segments as well as no interaction. However, higher levels of heart rate in the aircraft for non-normalized and non-standardized data were seen not only for relatively uneventful segments as these, but also for more demanding segments with peak heart rate levels (e.g. Go-around, Flaps full, Touch down).

The Engine failure segment (Figure 4b) did display overall higher heart rate for the simulated flight. There was a significant main effect of type of flight (F[1, 7]=12.24; p<.01) and also a main effect of time (F[9, 63]=385.4; p<.0005) as well as a significant interaction (F[9, 63]=2.91; p<.01). The increase of heart rate is higher and more rapid in the aircraft than in the simulator. However, the heart rate just before the event is considerably higher in the simulator, resulting in an over-all higher average than in the aircraft.

Eye Movement Energy (EME)

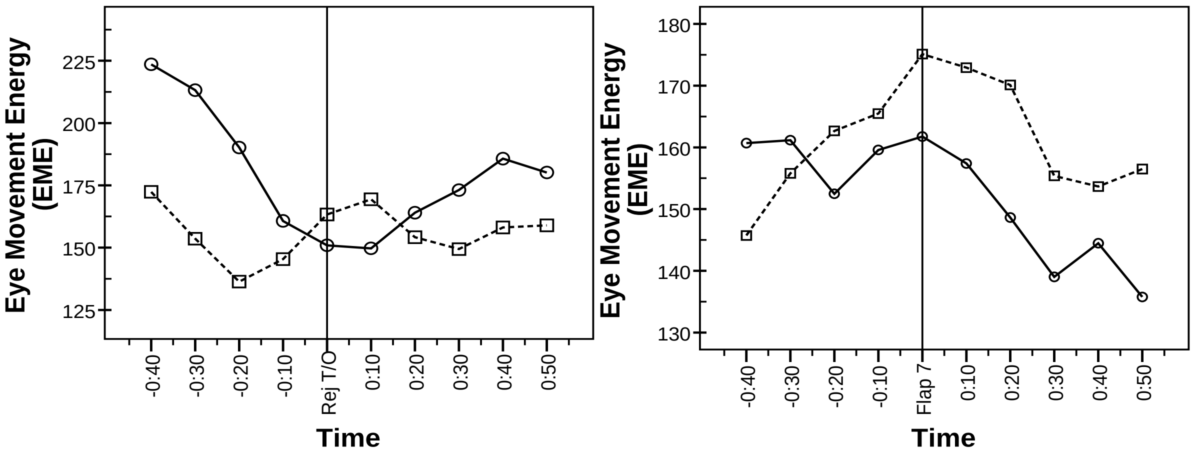

Figure 5 a and b. Average Eye Movement Energy during the flight segments Rejected take-off and Flap 7. Solid lines are aircraft flight, dashed lines are simulator flight

The absence of a visual system in the simulator meant that differences in EME were expected. In the aircraft the student pilots were able to shift references between inside and outside of the aircraft, something that leads to an increased amount of eye movement (Wilson, 2001). For the Rejected take-off (Figure 5a) there was a main effect of time (F[9, 63]=2.38; p<.05) as well as an interaction effect (F[9, 63]=1.83; p<.10). There was a higher level of EME in the aircraft, except for when the rejected take-off occurred and the student pilots had to keep their eyes on the runway to stay on it. The Flap 7 segment (Figure 5b) was a part of the instrument approaches. This meant that the student pilots either flew in weather conditions allowing only instrument flight or were screened off from being able to look outside. For this segment there was no main effect of type of flight or time and no interaction effect. The pattern of EME in this case was similar for the two different types of flight.

Ratings of mental workload

Student ratings of their mental workload were not performed for all flight segments and the ratings of the incompletely rated segments were excluded from further analyses. Ratings of the landings for instance, where not given by all participants for the simulator sessions.

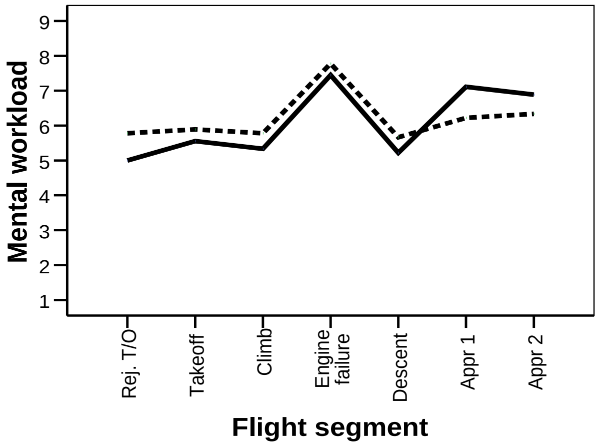

Figure 6. Student subjective ratings of mental workload during seven flight segments. Solid line is aircraft flight, dashed line is simulator flight

For the subjective ratings of mental workload (Figure 6), there was a main effect of flight segment (F[6, 48] = 9.22; p<.0005). There was no main effect of type of flight and no interaction effect. Accordingly, this indicates that there was no difference in how the two types of flights were rated. The Engine failure produced the highest ratings of mental workload for both types of flight. The engine failures did however occur at different times and conditions in the simulator (after take-off) and in the aircraft (after reaching straight and level flight at safe altitude). The lack of availability of corresponding heart rate data and subjective ratings for flight segments implied that statistical analysis of the relation between them would be unproductive.

Discussion

The overall lack of statistically significant differences in heart rate data for simulator and aircraft flight indicates that the mental workload is similar for the two type of flight; flight segments that produce either higher or lower levels of heart rate or do so in both types of flight. This supports previous findings (Roscoe, 1992; Wilson, 2001) of increased heart rate as an effective indicator of in-flight mental workload.

Certain flight segments (T/O Brake release, Engine failure, Cruise, Flaps full and Touch down) did however produce significant differences in heart rate that may be explained by the differences in the conditions for the segments:

- Rejected take-off in the aircraft was followed directly by another brake release and take-off. In the simulator it was followed by a longer period, providing the student with time to relax before the second take-off.
- In the simulator engine failure occurred immediately after the Take-off and Gear up segments and this was known to the student pilots. This means that they were aware of going from one demanding segment to another and could prepare for it mentally. In the aircraft the engine failure occurred at straight and level flight at higher altitude and at a less predictable point in time. Also, an engine failure at low altitude is a more serious emergency (and more cognitively demanding) than one at higher altitude.
- The cruise segment was considerably longer in the simulator, with potentially more time for the student to relax before the next flight segment.

- The touch down was not an essential part of the training in the simulator due to that id did not have a visual system. The student pilots descended decision height and then simply continued to descend until zero feet of altitude. In the aircraft the landing is normally the most cognitive demanding flight segments (Wilson, 2001, 2002a).

Overall the results confirm the results of Magnusson (2002) that there is a high degree of correspondence not only in psychophysiological reaction but for mental workload between simulator and aircraft flight. The similar patterns of heart rate and the subjective ratings support this. The student pilots reacts similarly to the events of a certain flight segment regardless of they are experiencing the event in the simulator or in the aircraft, i.e. the simulated flight seem to replicate the workload demands from aircraft flight.

However, when comparing heart rate data that has not been normalized and standardized there are differences, as seen in Figure 3b and 4a for the Cruise and Flap 7 segments. Although not verified by statistics, for these segments heart rate is consistently lower in the simulator. This difference was observed for other segments (e.g. Go-around, Flaps full) and has been observed in previous research (Magnusson, 2002). That it is not statistically significant may be due to a floor effect; the level of heart rate is not sensitive to decreases in workload demands below the level experienced in uneventful segments in the simulator. To separate flight segments statistically is also difficult since “surprise” effects produce higher heart rates in the simulator and smoothes out averages (see Figure 2). What this difference in heart rate could mean is open to interpretation. It was considered and discussed during this study that student pilots perceived simulated flights more as training and the aircraft flights more as opportunities to prove their ability. This was however directly opposed by student pilots and flight instructors. Another suggestion was that an emotional component of mental workload in regards to the potential consequences of aircraft flight (risk of an accident) may have caused the higher levels of heart rate. Although interesting and rarely brought up in the literature, this suggestion would represent a step away from the well established use (particularly in aviation research) of heart rate as a valid and reliable measure of mental workload. If heart rate is such a measure, the difference in levels of heart rate indicates that there is a difference in mental effort invested in the task, i.e. that mental workload is higher in the aircraft. Since the simulator in this study was equipped with neither a visual or motion based system the difference in heart rate could be reflecting the level of simulation. Although a visual system could have had an impact on the mental workload in the simulator the flights in this study were primarily focused on instrument flight (and a majority of aircraft flights were flown in instrument conditions). The absence of a motion-based system can be considered of no importance for this study. Both Roscoe (1991) and Longridge, Bürke-Cohen, Tiauw & Andrew (2001) concluded that research has failed to find any operationally significant effect of the use of motion-based systems on training. Further investigation of mental workload in simulator and aircraft flight should however look at the impact of visual and motion-based systems on differences in heart rate between the two types of flight.

Heart rate during the rejected take-off (Figure 3a) is interesting for consideration of the differences between simulator and aircraft flight. The peak heart rate in the simulator is reached as the rejected take-off occurs while it is delayed until after the occurrence in the aircraft. In both cases the event is expected by the student pilots as they accelerate to the target speed. However, in the simulator the student pilots can focus all attention on the speed indication and prepare mentally for the event before it occurs. In the aircraft the student pilots need to keep their eyes on the runway to stay on it and look out for other traffic, something that decreases the focus on speed and increase the “surprise effect” when the rejected take-off occurs. The difference in peak heart rate seem to indicate that in the simulator the increase of workload starts in advance of the occurrence of the event and the occurrence itself represents peak mental workload. In the aircraft the occurrence itself is the starting point for handling the situation and the peak mental workload seems to be connected to the handling rather than to the occurrence itself. Since simulators often are used to drill emergency or unusual maneuvers, more than to provide “free play” where unexpected events occur, this difference is interesting. Even though it may be argued that scenarios with unexpected events are often used for simulator training these events are often not unexpected at all, but well known to the pilots exposed to them. (A common complaint among airline pilots is about being first in the proficiency check period, when the scenarios are not known to them, something that is changed soon after the period has started and “the word is out”.) If similar differences in heart rate of “preparatory” nature could be shown to occur consistently for “unexpected” events in simulator training compared to in aircraft flight the relevance and transfer for simulator training of such events may need further consideration.

The Engine failure segment (Figure 4b) indicates the potential of a similar interpretation of heart rate data. In the simulator student pilots knew that the engine failure would occur immediately after take-off and this seem to have increased heart rate in advance of the actual occurrence of the event. In the aircraft the engine failure occurred during straight and level flight with a longer time window for it to occur, thus making the occurrence less predictable. This could explain the lower initial heart rate and more rapid increase when it was experienced in the aircraft. Again, if the mental workload for “unexpected” events could be shown to be partly of preparatory nature in simulators and more connected to handling of situations in aircraft flight such a difference would be of importance for the design of pilot training.

Subjective ratings of mental workload (Figure 6) were high for Engine failure, moderately high for Approach and low for Rejected take-off, Climb and Descent segments. This seems to be in accordance with heart rate data, except for the Rejected take-off. When heart rate was examined in detail (Figure 3a) this segment showed a distinct rise and peak of heart rate around the time for the occurrence of the rejected take-off. This shows that psychophysiological data can provide a higher resolution of data than subjective ratings. (Interfering with a pilot repeatedly during an individual maneuver is impossible, particularly if it is an emergency maneuver, and asking the pilot to rate his mental workload afterwards for different segments of the maneuver should be equally impossible.) Also, Wilson (2002a) showed that mental workload in landings are consistently rated lower in subjective ratings compared to the indications of heart rate and other psychophysiological data. This was explained with landings being a routine maneuver performed so many times that the effort invested was underrated by the pilots. Wilson (2002a) concluded: “The heart rate results may be more sensitive to the actual demands placed on the pilots by each segment“ (p. 16). A similar explanation could account for the difference of subjective ratings and heart rate data (and EME) for the rejected take-off. Since this was a ground maneuver it would not be perceived as difficult and consequently rated low, even though psychophysiological data indicates that experiencing it created a significant increase in mental workload. This shows a weakness in using subjective ratings only for investigation of mental workload and support the conclusion that psychophysiological data is of great for importance when research on in-flight mental workload.

Bell & Waag (1998) concluded that "Although a fair amount of opinion data exists that suggests there is training potential in using simulation, actual transfer data are extremely limited." In spite of an increase in transfer studies (such as by Taylor, Emanuel & Rantanen, 2005) this situation seems to prevail. Increased face validity has certainly been helpful in gaining acceptance for simulator training in the pilot community (Roscoe, 1991) but the persistent attention to further increase of face validity may have little impact on learning and transfer to aircraft flight. To improve simulator training there is a need for increased knowledge about the transfer effects between simulator and aircraft. This in turn means that there is a need for increased knowledge about the mental workload experienced in respective environment. To rely only on subjective ratings by pilots and instructors for such investigations is according to Hays, Jacobs, Prince & Salas (1998) "major problem with research in this area" (p. 153) and may mislead the attention of research to pursue ever increasing demands of higher face validity (resolution of visual system, sophisticated motion systems, simulated air traffic control environment etc.).

Conclusion

This study has shown that measurement of mental workload in basic civil aviation training with psychophysiological methods provides interesting results on mental workload demands during specific flight segments and for comparisons between simulator and aircraft flight . The results identify flight segments and individual maneuvers with high mental workload demands (Take-off, Rejected take-off, Engine Failure, Flap 7/Approach, Go-around, Touch down/Landing) of particular importance to focus on during basic flight training. Also, overall the results confirm the training value of using simulators. They do however also question some well accepted assumptions on simulator training and transfer to aircraft flight, particularly regarding unexpected events. Differences in heart rate between simulator and aircraft flight, indicating that mental workload is higher in the aircraft, as well as the preparatory nature of mental workload in the simulator questions if replacing aircraft training with simulator training is completely unproblematic. At very least the results highlight important research needs to ensure the success of new Multi-Crew Pilot License. Further investigation mental workload of student pilots in simulator and aircraft flight training can reveal more about the mental workload demands of flight segments and individual maneuvers as well as more on matches and mismatches between the two types of training. This can provide valuable information for improvement of basic civil aviation training as well as for development of flight simulators, confirming or disconfirming beyond traditional face-validity the relevance and transfer of simulated flight training.

References

Bell, H. H., & Waag, W. L. (1998). Evaluating the effectiveness of flight simulators for training combat skills: A review. *The International Journal of Aviation Psychology, 8*(3), 223-242.

Bonner, M. A., & Wilson, G. F. (2002). Heart Rate Measures of Flight Test and Evaluation. *The International Journal of Aviation Psychology, 12*(1), 63-77.

Carretta, T. R., & Dunlap, R. D. (1998). *Transfer of training effectiveness in flight simulation: 1986 to 1997* (No. AFRL-AZ-TR-1998-0078). Mesa, AZ: Air Force Research Laboratory.

Dahlstrom N. (2002). Current aspects on aviation training and its relevance for safety. *Human Factors and Safety in Aviation: Proceedings from a conference*, 47-54. Lund, Sweden: Swedish Centre for Aviation Research & Development.

Dahlstrom N. (2005, November 7-8). *Pilot selection and training – the academic connection*. Presentation at the European Aviation Training Symposium, Amsterdam, Netherlands. Farnborough, UK: Halldale Media Ltd.

Dahlström, N., & Nählinder, S. (2006). A Comparison of Two Recorders for Obtaining In-flight Heart Rate Data. *Applied Psychophysiology and Biofeedback, 31*(3), 273-279.

Dekker, S. W. A., & Johansson, B. (2000). *JAR-FCL and pilot knowledge needs*. Linköping, Sweden: Swedish Centre for Human Factors in Aviation, Linköping University.

Dennis, K. A., & Harris, D. (1998). Computer-based simulation as an adjunct to ab initio flight training. *The International Journal of Aviation Psychology, 8*(3), 261-276.

Forbes G. (2005, November 7-8). *An update on the Multi-Crew Pilot License (MPL)*. Presentation at the European Aviation Training Symposium (EATS), Amsterdam, Netherlands. Farnborough, UK: Halldale Media Ltd.

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Requirements for Future Research in Flight Simulation Training: Guidance Based on a Meta-Analytic Review. *The International Journal of Aviation Psychology, 2*(2), 143-158.

Lee, A. T. (2005). *Flight Simulation: Virtual Environments in Aviation*. Aldershot, UK: Ashgate Publishing Ltd.

Lintern, G., Roscoe, S. N., Koonce, J. M., & Segal, L. D. (1990). Transfer of landing skills in beginning flight training. *Human Factors, 32*(2), 319-327.

Longridge, T., Bürki-Cohen, J., Go, T. H., & Kendra, A. J. (2001, March 5-8). *Simulator Fidelity Considerations for Training and Evaluation of Today's Airline Pilots.* Paper presented at the International Symposium on Aviation Psychology, Columbus, OH.

Magnusson, S. (2002). Similarities and differences in psychophysiological reactions between simulated and real air-to-ground missions. *The International Journal of Aviation Psychology, 12*(1), 49-61.

Nählinder, S., Berggren, P., & Persson, B. (2005). *Increasing training efficiency using embedded pedagogical tools in a combat flight simulator.* Paper presented at the Human Factors and Ergonomics Society (HFES), Orlando, FL.

Roessingh, J. J. M. (2005). Transfer of manual flying skills from PC-based simulation to actual flight - Comparison of in-flight measured data and instructor ratings. *The International Journal of Aviation Psychology, 15*(1), 67-90.

Roscoe, A. H. (1992). Assessing Pilot Workload: Why measure Heart Rate, HRV and Respiration? *Biological Psychology, 34*(2-3), 259-288.

Roscoe, S. N. (1991). Simulator Qualification: Just as Phony as it Can Be. *The International Journal of Aviation Psychology, 1*(4), 335-339.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It Is Not How Much You Have but How You Use It: Toward a Rational Use of Simulation to Support Aviation Training. *The International Journal of Aviation Psychology, 8*(3), 197-208.

Sohn, S. Y., & Jo, Y. K. (2003). A Study on the Student Pilot's Mental Workload due to Personality Types of Both Instructor and Student. *Ergonomics, 46*(15), 1566-1577.

Taylor, H. L., Talleur, D. A., Emanuel, T. W., & Rantanen, E. M. (2005). *Transfer of Training Effectiveness of a Flight Training Device (FTD).* Paper presented at the International Symposium on Aviation Psychology, Columbus, OH.

Telfer, R. A. (1993). Introduction. In R. A. Telfer (Ed.), *Aviation Instruction and Training* (pp. 1-7). Aldershot, UK: Ashgate Publishing Ltd.

Veltman, J. A. (2002). A Comparative Study of Psychophysiological Reactions During Simulator and Real Flight. *The International Journal of Aviation Psychology, 12*(1), 33-48.

Wickens, C. D., & Hollands, J. G. (2000). Assessing Mental Workload. In C. D. Wickens (Ed.), *Engineering Psychology and Human Performance* (3rd ed., pp. 459-471). Upper Saddle River, NJ: Prentice-Hall.

Wilson, G. F. (2001). Psychophysiological Inflight Monitoring. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in Ambulatory Assessment* (pp. 435-454). Seattle, WA: Hogrefe & Huber Publishers.

Wilson, G. F. (2002a). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *The International Journal of Aviation Psychology, 12*(1), 3-18.

Wilson, G. F. (2002b). Psychophysiological test methods and procedures. In S. G. Charlton & T. G. O'Brien (Eds.), *Handbook of human factors testing and evaluation* (2nd ed., pp. 127-156). Mahwah, NJ: Lawrence Erlbaum Associates.

Wilson, G. F. (2002c). A comparison of three cardiac ambulatory recorders using flight data. *The International Journal of Aviation Psychology, 12*(1), 111-119.

Woods F. (2005, November 7-8). *A Flight Training Regulatory Update – The latest on JAA, EASA, JAR-FCL and ICAO MPL*. Presentation at the European Aviation Training Symposium, Amsterdam, Netherlands. Farnborough, UK: Halldale Media Ltd.