29th Congress of the International Council
of the Aeronautical Sciences
St. Petersburg, Russia; September 7–12, 2014

MENTAL EFFORT AND SAFETY IN CURVED APPROACHES

J.O. Entzinger*, T. Uemura*, S.Suzuki*
***The University of Tokyo, School of Engineering**
jorg@entzinger.nl; sky.ghaya@gmail.com; tshinji@mail.ecc.u-tokyo.ac.jp

Keywords: *Curved Approaches, Flight Safety, Workload, Psychophysics, Human Factors*

Abstract

Curved approach procedures are implemented around the world. Although typically flown by the autopilot, human pilots will need the situational awareness and skills to take over control in rare case events. We try to understand the pilot's cognitive models and differences in required (mental) effort between conventional straight-in approaches and curved approaches. We developed various methods to visualize pilots' control efforts during manual flight, show their capabilities by comparing various straight-in approach scenarios including good and bad visibility cases, and introduce preliminary results of curved approach experiments. Most experiments were carried out with a few airline pilots and student pilots who had received some elementary flight training in a fixed-base B747-400 simulator. The curved approach experiments were performed in a Dornier Do-228-200 full flight simulator. The main analyses discussed here are based on spectrograms of the pilot's elevator control, pupil diameter, and electrocardiogram (heart rate and heart rate variability). Results show increased mental effort in bad visibility scenarios and during difficult phases of flight (large deviations or flare manoeuvre) and how the control style of airline pilots is more refined than that of the trainees.

1 Introduction

We investigate which cognitive challenges human pilots face during the execution of curved approaches flown under Required Navigation Performance Authorization Required (RNP-AR) procedures. Although path design and cockpit automation for RNP-AR curved approaches have received much attention and are quite well established, descriptions of operational human factors issues in this complex environment are mostly anecdotal. This is in part due to the high reliance on automation, and the fact that curved approaches are currently mostly carried out under 'ideal' conditions (not during peak-times at airports, in good visibility conditions, etc.). Our research is trying to fill this gap and focuses on human factors issues such as situational awareness, cognitive and mental models, automation supervision, and what kind of training would be required to help pilots improve on these points for the particular application of curved approaches.

From a literature review and interviews with researchers and pilots who actually fly RNP-AR curved approaches on a regular basis, we know that various safety issues arise during actual operations [e.g., 1–3]. To mention just a few: it is practically impossible for the pilot to confirm the correctness of all the waypoints in the navigation database; the differences between normal RNAV and RNP-AR procedures are sometimes so subtle that pilots or air traffic controllers are likely to mix them up; planning and supervision by air traffic control is complex in mixed mode operation (when some aircraft make straight and other aircraft curved approaches to the same runway) and may lead to sudden changes; and when a system failure happens or decision to go-around is made in a curve, proper situational awareness and system (re-)configuration are more difficult to achieve. In our current research, we focus mainly on this last issue.

Curved approaches are generally carried out relying heavily on cockpit automation. However, the human pilot still has the final responsibility, and should at any time have full situational awareness and be able to intervene. This may become a challenge, since all pilots we spoke acknowledged that these approaches are practically always flown using the autopilot and autothrottle engaged, and pilots have little or no experience flying such approaches manually with only the Flight Director and Flight Management System. Additionally, it is much easier to cross-check tracking performance for a straight path than for a curved one, especially under strong wind, one-engine-out, or other irregular conditions.

Whereas the new RNP-AR cockpit automation is said to decrease pilot workload and increase safety in standard situations, the opposite is likely to happen for non-normal cases. For RNP-AR to be successful and to guarantee safety in the future, we will therefore have to investigate rare-event cases and particular necessities in (cognitive) pilot training.

This paper introduces new findings in our ongoing research project. Earlier analyses and findings are briefly introduced for completeness, but the reader is referred to our previous publications for details [4, 5].

2 Mental Effort and Safety

The increased capability, accuracy and reliability of aeronautical systems has left 'human error' as the largest accident cause. One way to think about this is that automation nowadays can handle all but the most extreme situation, which leaves the human pilot with only the hardest and maybe impossible problems. Another way, often noted by pilots, is that the high level of automation reduces their manual flying skills due to less frequent practice [e.g., 6, 7]. Still another reason is that pilots may overtrust automation or lose situational awareness due to its lack of transparency, and therefore suffer from complacency or plan continuation error [e.g., 8, 9].

Whereas the aeronautical systems are deterministic and can be analysed and tested intensively before receiving certification, this is different for human pilots. Their performance depends on a large number of variables, including training, experience, recent practice, and workload. Since workload has many aspects and there is no generally accepted definition of it, we will distinguish and define a few related concepts here:

- **T** *Task load*; an objective load imposed on the operator. This may be optimized by good system design (but notice the trade-off between minimizing average load and minimizing peak load).
- **R** *Available mental resources*; the capabilities the operator has to perform the task. This depends on the operator's experience and skill, and can be enhanced through training and practice.
- **E** *Mental effort*; the amount of effort the operator is investing in the task. This may depend on the operator's general physical and mental condition, as well as motivation.

We then define workload **W** as the ratio between the effort invested and the resources available or $W := \frac{E}{R}$. We also note that the relation between workload and performance is not a direct relation. We could say that performance is not likely to degrade as long as the $T \leq E$ AND $T < \alpha R$, with $0 << \alpha < 1$ to guarantee a sustainable level. This idea of clearly separating workload and effort has been adopted by many other researchers as well and is particularly useful because the mental effort invested by the operator can be determined by observing physiological reactions [e.g., 10–13].

A major difficulty in assessing flight safety is the extreme rarity of serious events. Under normal operations, we don't expect system trouble and workload is not likely to be a serious issue. It is in critical situation, where workload is already high, that a small difference in task load or situational awareness may have serious consequences. However, if we continuously present the pilot with such extreme scenarios in simulator experiments, the pilot will anticipate it, and may even get used to it, thereby reducing the fidelity of the experiment. The current research therefore focuses at comparing a number of cases and tries to identify which points are most likely to become critical in extreme situations.

Figure 1 shows our hypothesis tree. The boxes indicate the three locations in the closed

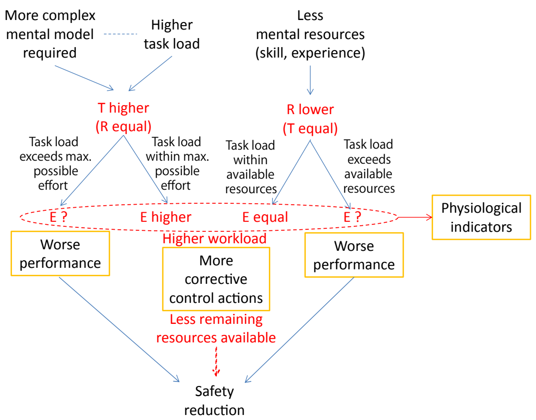


Fig. 1 Hypothesis tree

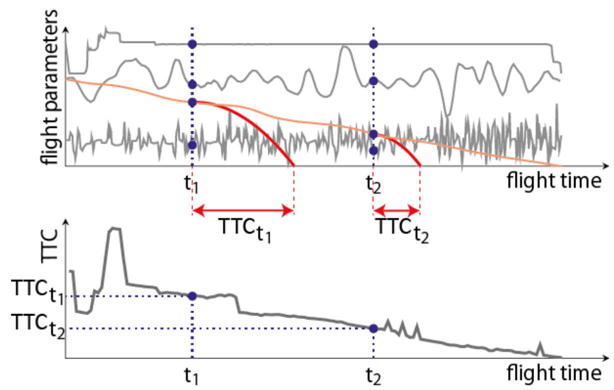


Fig. 2 Creating a time-to-crash graph (bottom) by calculating the new altitude profile (red) for the case that the flight would be resumed without any control inputs.

control loop of pilot and aircraft where we can obtain data: the pilot himself (physiological indicators), the pilot's control actions, and the resulting aircraft state (performance).

3 Analysis Methods

We gathered physiological, control input, and performance data in several experiments. In this section we will introduce the specific types of data we gathered and how we analysed that data.

3.1 Performance data analysis: Time to Crash

Gawron noted that *"The most objective measure of danger [...] is time until the aircraft is destroyed if control action is not taken." [14].* Based on this idea, we developed a 'time to crash' (TTC) analysis. Using aircraft states from subsequent points in the flight experiment as starting point, the remainder of the flight is simulated for the case where the pilot takes his hands off the controls, and the time-to-crash is calculated. Deviations of the TTC from the stabilized approach time to landing are penalized and a TTC index is calculated. This analysis process is illustrated in Fig. 2.

This method and the experiment results are detailed in an earlier publication by Nijenhuis [4] and summarized in [5] and will not be discussed in detail here. The main conclusion is that having sufficient time to stabilize the aircraft is therefore important, and this may be difficult when the final straight segment in RNP-AR curved approaches gets too short.

The TTC analysis proved to be useful as an intuitive measure of flight safety, but has the limitation that it must be possible to make fast-time simulations of the experiment approaches. This is no problem if one has access to the simulator software (source codes), but may be impossible when using off-the-shelf simulators including those used for training at airlines.

3.2 Control inputs analysis

We developed two ways to analyse the pilot's control inputs. One is by looking at the power of the elevator control input signal, which would tell something about the overall pitch control effort of the pilot. The other analysis focuses on a spectrogram of the elevator control input, which shows more details about the control strategy in different phases of the flight.

As with the TTC, the investigation of the pilot's control input power is discussed in detail in [4] and summarized in [5]. This analysis captures the pilot's control effort in a single parameter, which makes it easy to compare (parts of) approaches. Pilots showed a higher control effort in difficult situations, as we expected.

The main results of the spectrogram-based analysis of the pilot's control input will be discussed below.

3.3 Physiological data: Eye data

We looked in detail at the analysis of pupil diameter as a measure for mental effort, which is an established method for baseline psychological experiments with discrete task [e.g., 15, 16]. We also looked at the number of blinks per time unit. Although we recorded the gaze direction, we did not analyse it in detail, but only used it as a reference when interpreting the other data.

One analysis method suggested in literature is the 'index of cognitive activity' as developed by Marshall [17]. However, this method did not provide the robust and useful results we had hoped for. We therefore now focus on the pupil diameter itself.

3.4 Physiological data: ECG data

We recorded electrocardiograms (ECG) and analysed the data using the open source ecgBag software [18].We then calculated the instantaneous heart rate (HR) and the heart rate variability (HRV). These two parameters are related to stress and effort. In particular the HRV power spectrum band from 0.06 to 0.14 Hz is said to be suppressed in cases of high mental effort [10, 11].

4 Experiments

4.1 Baseline Experiments

We did a number of baseline experiments to get feeling for the background noise and signal amplitude we could expect from our eye-mark and ECG measurements.

A problem using the pupil diameter as a measure for mental effort, is that it also changes depending on the illumination level. We therefore did some eye-mark measurements where subjects in a darkened room looked at a large projection screen, covering almost the whole field of view, and white boxes of various sizes were projected for 30 s each. In another test, subjects looked at various cockpit instruments for 30 s each and at the outside view from 2000 ft and 20 ft height.

Several ECG measurements were taken during relaxation. Additionally, ECG measurement continued for some time after the landing in each trial, although talking or preparation for the next trial may have influenced these final data.

Fig. 3 Briefing and training of a student pilot in the fixed-base B747-400 simulator

We also had subjects observe automated landings, while recording eye-mark and ECG data. Although subjects were told to just look, especially experienced pilots may still have been 'supervising' the process, and thus invested some mental effort as if they were in control.

4.2 Straight-in Experiments

We carried out a large number of informal experiments to develop and choose useful analysis methods using a fixed base B747-400 simulator at The University of Tokyo. Participants were the researchers themselves and a few interested members of the laboratory ('student pilots'). The student pilots received basic flight training and specific approach and landing training from a retired airline pilot twice a week for several weeks before the main initial experiment (Fig. 3). For a few of the experiments, other (current or retired) airline pilots volunteered to take part.

Since the simulator is not RNP-AR capable, we started out with straight-in approaches under different conditions, in particular Visual Meteorological Conditions (VMC, or 'good visibility') where the runway is visible all the time and Instrument Meteorological Conditions (IMC R800, or 'bad visibility') where the pilot has to fly using the cockpit instruments and can only see 800 m ahead, meaning the runway approach lights become visible at a height of ca. 500 ft. The hypothesis here is that IMC approaches are more difficult (require higher mental effort) than VMC approaches, in analogy with our original hypothesis comparing curved and straight approaches.

We measured the aircraft states and pilot control inputs at 20 Hz, recorded eye-data (gaze di-

Fig. 4 The ParamaTech EP-301 portable ECG recorder

Fig. 5 The NAC EMR-8 eye-mark recorder

rection, blinks, and pupil diameters) at 60 Hz, and electrocardiogram (ECG) data at 250 Hz. The ECGs were recorded using the ParamaTech EP-301 (Fig. 4), and for the eye data recordings, the NAC EMR-8 was mostly used (Fig. 5), although for a few flights the similar NAC EMR-9 was used.

4.3 Curved Approach Experiment

Recently we carried out a first experiment with curved approaches in the full flight simulator owned by the Japan Aerospace eXploration Agency (JAXA). Although not truly RNP-AR capable, this Dornier Do-228-200 turboprop simulator can be used to make curved approaches, and was programmed with the RJTT/TOKYO INTL RNAV(RNP) RWY23 (curved) and ILS Z RWY23 (straight) approaches to Tokyo Haneda airport. The approaches started around 1500 ft height above ground level, and for the IMC approaches the cloud ceiling was set to coincide with the end of the last curve in the approach path, that is at 585 ft, which is about 4 times as high as in fixed-base IMC experiments.

One airline pilot and one student pilot flew 2 straight VMC approaches, 4 straight IMC approaches, and 3 curved IMC approaches each. We used the same measurement equipment as mentioned above for the physiological data, and simulator data was recorded at 25 Hz.

5 Results

5.1 Baseline Experiments

5.1.1 Physiological data: Eye data

Figure 6 shows an example of the baseline experiment measuring the pupil diameter. We can clearly distinguish the different illumination levels by looking at the time history of the pupil diameter. It also seems like higher frequency fluctuations are larger when focusing on an area with larger contrast (smaller boxes or lines), although this cannot explain the initial 30 s.

The other experiment, looking at different cockpit displays, did not show such large and distinct changes in pupil diameter. This leads us to believe the illumination conditions around our simulator are sufficiently controlled. There is however quite a lot of ‘noise’ in the signal, probably due to the fact that scene is very contrast-rich (in particular the cockpit instrument displays).

5.1.2 Physiological data: ECG data

Figure 7(a) shows typical data from a baseline ECG measurement while relaxing. We see that the inter-beat-intervals (the reciprocal of which would be the heart rate per second) are varying quite a bit, and the mental effort as expressed by the suppression of the power spectrum band from 0.06 to 0.14 Hz of the heart rate variability is also fluctuating, but generally quite low (high values). The fluctuations may be because thoughts are wandering wile relaxing.

If we then take a look at Fig. 7(b), we clearly see smaller inter-beat-intervals (i.e., higher HR), and higher mental effort. After the touchdown, however, we see a very steep reduction of the mental effort (the final increase is probably be-

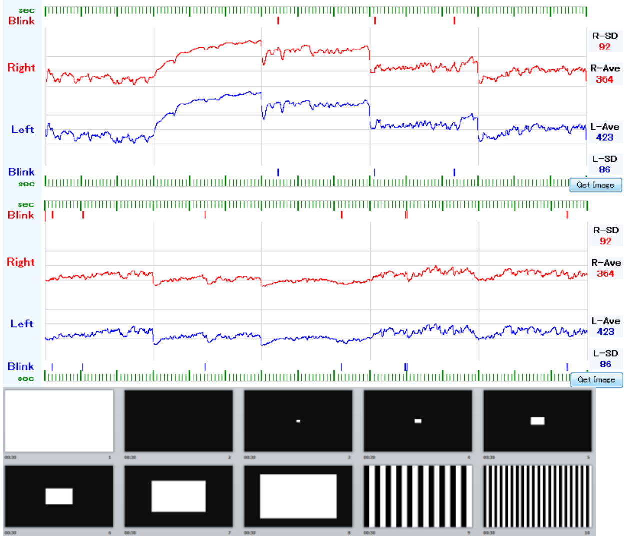
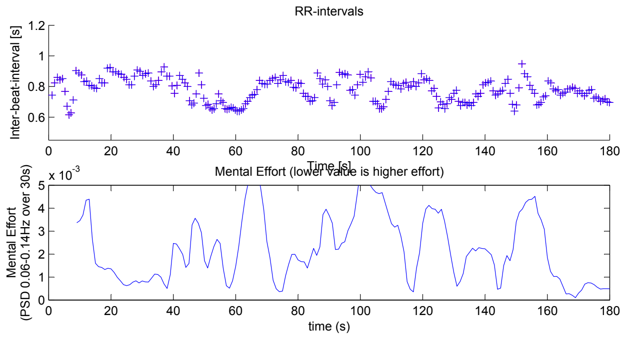
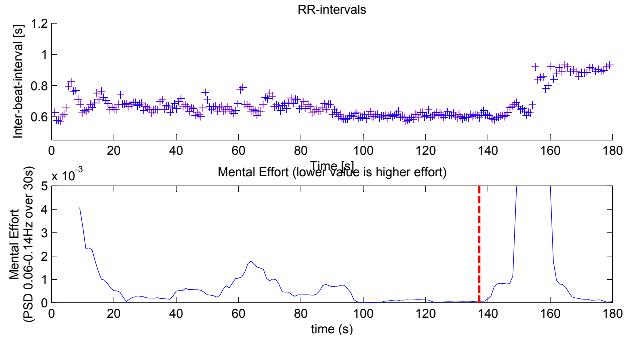


Fig. 6 Example of a baseline experiment. The upper graphs show the pupil diameter of the right and left eye (average and standard deviation indicated with the horizontal gray lines and values given on the right). The lower image shows the respective slides that were projected on a large screen for 30s each, and filling almost the whole field of view.



(a) Relaxing



(b) Flying an IMC R800 approach

Fig. 7 Example of inter-beat-interval (IBI = 60/HR) and mental effort data. (a) while relaxing (b) while flying a bad visibility IMC R800 straight-in approach. The vertical red line indicates the moment of touchdown.

cause of self-evaluation or spoken feedback).

Table 1 compares the HR averaged over the trial (in case of flights up to the moment of touchdown). It is clear that the HR is increased during flight trials, as compared to the relaxation periods. There also seems to be a small difference between the good and bad visibility flights, but this difference was not statistically significant.

Student Pilot H, Day 1		Student Pilot J, Day 2	
RELAX	65	RELAX	68
VMC	85	VMC	75
VMC	84	VMC	78
VMC	79	VMC	76
IMC R800	81	IMC R800	78
IMC R800	83	IMC R800	82
IMC R800	78	IMC R800	77
Student Pilot H, Day 2		**Student Pilot J, Day 2**	
RELAX	77	RELAX	65
IMC R800	93	VMC	77
IMC R800	88	VMC	71
VMC	83	IMC R800	75
VMC	82	IMC R800	73
RELAX	75	RELAX	69
Average		**Average**	
RELAX	72.3	RELAX	67.7
VMC	82.7	VMC	75.2
IMC R800	84.5	IMC R800	77.1

Table 1 Comparison of average heart rates when relaxing, when flying 'good visibility' VMC straight-in approaches and 'bad visibility' IMC R800 straight-in approaches.

5.2 Straight-in Experiments

5.2.1 Control inputs: Spectrogram analysis

The spectrogram of the elevator control input contains a lot of information about flying style (Fig. 8). This style changes remarkably with training and experience. Before training (Fig. 8(a)) the student pilot does not know well what to do, and spends a lot of effort correcting his own mistakes. These very strong and low frequency (long period) control inputs gradually become more subtle and faster throughout the training (Fig. 8(b)).

For the airline pilots we see something similar (Figs. 8(c) and (d)). It should be mentioned that the younger airline pilot has not only less flight hours (experience), but also considerably less experience operating our particular simulator, compared to the veteran. This might somewhat enhance the difference between both.

It is particularly interesting to see that the control frequency seems to increase in the final phase before touchdown. This can be explained by the increased salience of visual cues as well as the increasing sensitivity of the glide slope indicator and PAPI[1] throughout the approach.

For both airline pilots we show a good visibility (VMC) approach on top, and a bad visibility (IMC) approach at the bottom, and it is clear that control input frequency is generally higher in good visibility. We assume that this is because the pilot can obtain more information quicker from the outside visual scene, than through scanning his various cockpit instruments.

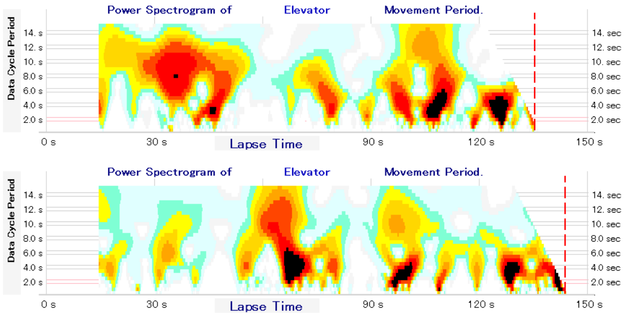
5.2.2 Physiological data: Eye data

In many of the trials we saw an increase of the pilot's pupil diameter around the flare phase, and in particular for 'good' landings. The flare is a pitch-up manoeuvre just seconds before touchdown, and is generally considered the most difficult part of the landing control. Although the observed pupil dilatation would perfectly match our hypothesis of high mental effort, we thought it might be just because the rapidly approaching runway is relatively dark. We therefore did a control experiment where the pilot is just watching the landing, without controlling it.

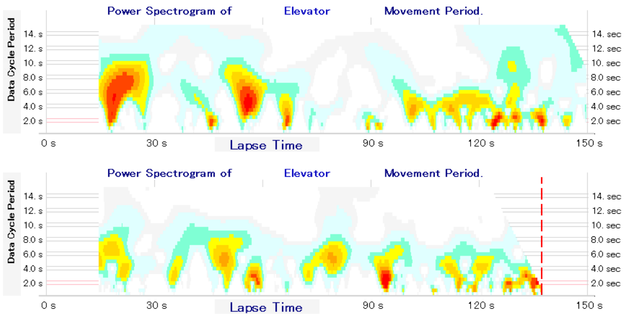
An example of the results is shown in Fig. 9. There are a few important differences between the 4 cases shown in the figure:

- The 'controlling' cases show clearly fewer blinks in general, and in the 20 s or so before touchdown, implying more visual attention.
- The overall standard deviation (R-SD value at the right of the graphs) of the pupil diameter is larger for the 'controlling' cases.
- The increase of the pupil diameter in the 20 s before landing is larger for the 'controlling'

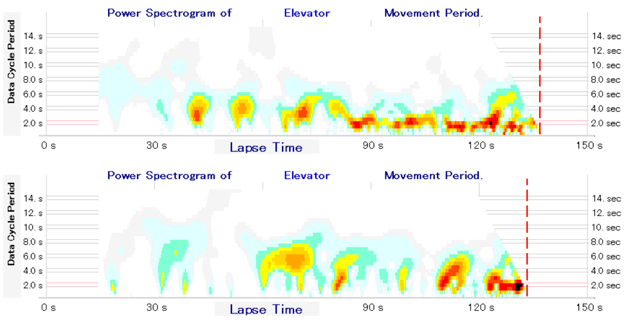
[1]PAPI: Precision Approach Path Indicator



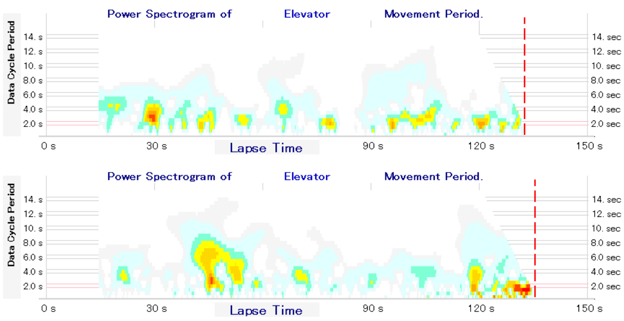
(a) Student pilot before training (2x)



(b) Student pilot after training (2x)



(c) Young airline pilot (above: VMC, below: IMC)



(d) Veteran airline pilot (above: VMC, below: IMC)

Fig. 8 Example spectrogram analysis of the elevator control input. Darker colours represent stronger control inputs. The vertical red line is the moment of touchdown.

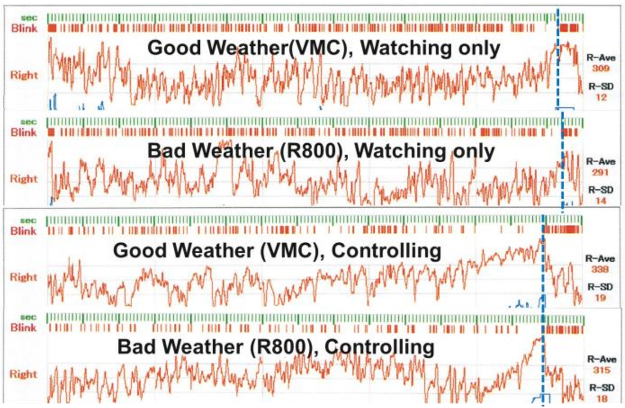


Fig. 9 Comparison of pupil diameters of a captain pilot. The vertical blue line indicates the moment of touch down. Each graph shows 150s recording time.

cases.

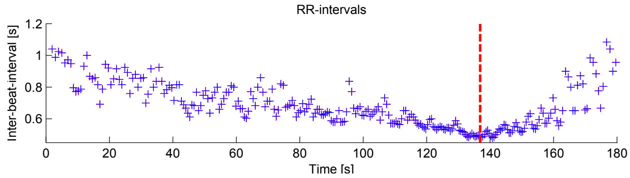
- The increase of the pupil diameter in the 20 s before landing is largest and most distinct for the 'bad visibility, controlling' case.

These results indicate that it is not just because of the approaching dark runway that the pupil diameter changes, and confirm the hypothesis that mental effort is higher in more challenging cases such as when actually controlling the aircraft and in particular when doing so in bad visibility conditions and when flaring.

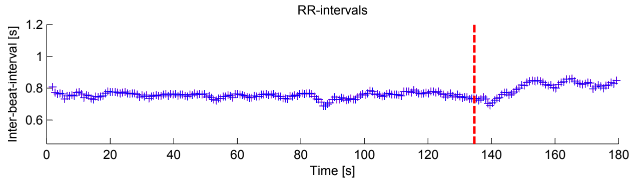
5.2.3 *Physiological data: ECG data*

Heart rate is well known to be related to stress. A simple comparison between a trainee and a veteran, as shown in Fig. 10, tells us that the trainee has a clear increase in HR (decreasing IBI) as the moment of touchdown comes closer, whereas this kind of routine landing does not arouse the experienced airline pilot at all. Although we do not see this pattern with all trainees, and even for the same trainee not in all trials, it clearly indicates a different level of readiness.

Another point one immediately notices is the general variation in the student pilot's HR, whereas the airline pilot's is extremely constant. This could indicate the airline pilot is investing more effort (more mental processing), but there are many other possible explanations based on interpersonal differences, including age and ethnicity.



(a) Student pilot



(b) Experienced airline pilot

Fig. 10 Example of how inter-beat-intervals may change toward the touchdown for an inexperienced operator, compared to a professional airline pilot.

As already mentioned in §5.1.2, Table 1 suggests that the difference in difficulty between approach types may also be reflected in the (average) heart rate. The same pattern of a slightly higher average HR in IMC than in VMC and a clearly lower heart rate during relaxation, was observed for the professional airline pilots. However, more trials will be needed to further investigate this, since this difference is not statistically significant. The variation between people and trials, but especially also between days is so large, that the average heart rate does not seem very useful for the evaluation of a single trial or pilot.

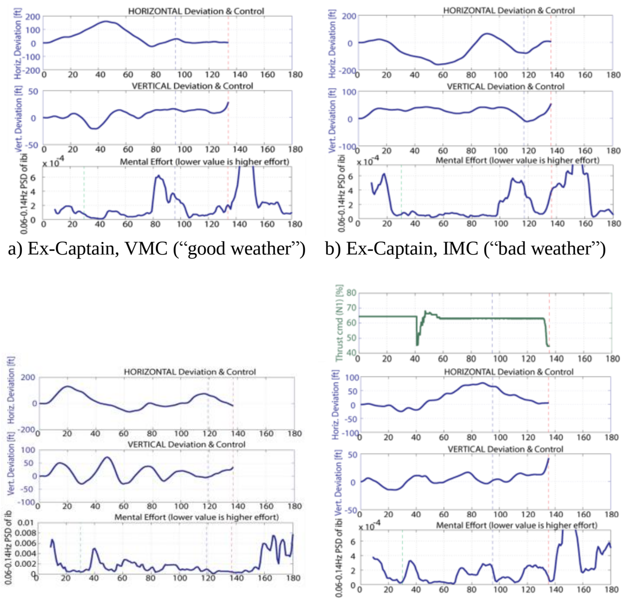
Rather that trying to compare people, it seems more meaningful to see how a single person's HRV changes over time during the approach. We therefore calculated the value of the 0.06–0.14Hz of the PSD of the HRV with a moving window over the duration of the experiment.

As we can see from the graphs in Fig. 11, mental effort quickly decreases (higher value) immediately after the touchdown (the later increase is probably due to the fact we start discussion, self-evaluation, etc.). We can also see an increased effort in the period just before touchdown.

For Fig. 11a), we can assume the higher effort just before 40 s is because the pilot recognizes his vertical deviation, and the subsequent one around 60 s because of horizontal deviation. Figure 11b) shows an IMC approach, which requires integrating information from various cockpit instruments. We can therefore see a quite constant high mental effort throughout most of the approach, but a slight relaxation around the time where the pilot gets the runway in sight. The student pilot's data shown in Fig. 11c) also shows a quite constant high mental effort throughout the IMC approach, although a little relaxation after recovering from the high horizontal deviation, only to find out that his vertical deviation has become unacceptable.

Figure 11d) shows characteristics similar to the ones discussed before, although the increased mental effort between 40–60 s cannot be explained by high horizontal or vertical deviations. After some searching, however, we found something peculiar happened with the thrust (throttle, power) setting that probably caught the pilot's attention.

This example shows the HRV analysis can be a powerful way to investigate mental effort in various phases throughout the approach.



a) Ex-Captain, VMC ("good weather") b) Ex-Captain, IMC ("bad weather")

c) Student, IMC ("bad weather") d) Ex-Captain, VMC ("good weather")

Fig. 11 Analysis of Heart Rate Variability (IBI = inter beat interval = 60/Heart Rate). The red vertical line is the moment of touchdown. The blue vertical line is the moment where the pilot starts looking mostly outside, rather than watching the cockpit instruments.

5.3 Curved Approach Experiments

5.3.1 Control inputs: Spectrogram analysis

We are still analysing the data, but at first glance the airline pilot's control style seems very similar for the straight-in cases, including the higher frequency control input in good visibility and short before touchdown. For the curved approaches, there seems to be slightly more and lower frequency control just after coming out of the last curve, and possibly also just before landing. We also see more and lower frequency aileron (lateral) control.

5.3.2 Physiological data: Eye data

The pupil diameter data for the good and bad visibility straight-in approaches is in line with the expectations from the data obtained in the fixed-base simulator. This means that in the last 10 or 20 s before touchdown, we see a slight increase in pupil diameter for the good visibility approaches, and a more distinct increase for the bad visibility approaches (Fig. 12).

In the curved approaches, we saw a slow upward trend in the pupil diameter (which could indicate increasing mental effort) until ca. 110 s, which happens to be the moment the aircraft comes out of the last curve and the pilot can confirm his position with the outside view. Then, during the final 10-20 s before landing, the pupil diameter increases rapidly, even more distinctly than in the straight-in bad visibility cases.

These observations hold for both the airline and the student pilot.

5.3.3 Physiological data: ECG data

There were no meaningful differences in the average heart rates for the different types of trials. The mental effort as obtained from the heart rate variation seems to be slightly higher for the first two curved approach trials. However, the mental

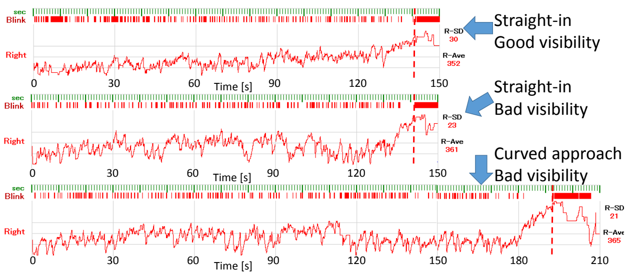


Fig. 12 Comparison of pupil diameter data obtained in the full flight simulator experiment.

effort in the last trial is fluctuating and comparatively low for both pilots. This might be a sign of fatigue.

5.4 Discussion

We showed how pupil diameter, heart rate, heart rate variability, and control style can reveal a pilot's mental effort and task load. This will be important now aircraft systems and operations become more and more complex, such as with RNP-AR curved approaches. Particularly in rare events, where the human pilot suddenly has to take over the control authority from the autopilot, workload may become a serious issue.

We investigated how mental effort can be measured during (simulated) flight, and have shown differences between good and bad visibility cases,for different training levels, and in different flight phases. Most experiment trials were straight-in approaches in a fixed-base simulator to confirm the appropriateness of the suggested analysis methods. An initial analysis of curved approach a few trials extended these findings.

Further analysis and more data is needed to confirm the current findings, especially considering the curved approaches. If successful, a next step would be to do experiments in fully RNP-AR capable simulators with pilots who are authorized to fly such approaches.

Acknowledgements

This research is financially supported by the Japan Aerospace eXploration Agency (JAXA) under the next generation air-traffic management system DREAMS program, and we like to express our gratitude to Dr. Kohei Funabiki of JAXA for his guidance.

The authors would like to thank Mr. Marijn Nijenhuis for his contribution during his internship and to Mr. Daiichiro Matsunaga for his assistance with operating the simulator and other equipment. Additional thanks go out to the various pilots and researchers who we met and interviewed for this research, the students and pilots who took part in the simulator experiments.

References

[1] B. Lyall, R. England, G. Harron, B. Hoffa, E. Jones, M. Lyall, S. Taylor, and J. Wilson. "Flight Crew Training for Nextgen Automation". Final report, Federal Aviation Administration, Sep. 2011.

[2] R. Barhydt and C.A. Adams. "Human Factors Considerations for Performance-Based Navigation". Technical Memorandum NASA/TM-2006-214531, NASA Langley Research Center, Dec. 2006. URL `http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20070003553_2007003019.pdf`.

[3] D. Mitchell, R. Eklund, and H. Ekstrand. "A Human Factors Assessment of Mixed-Mode Air Traffic Arrival and Approach Procedures". In *The Swedish Human Factors Network (HFN) Conference, Linköping, Sweden, November 24-25, 2011*, Nov. 2011.

[4] M. Nijenhuis. "Measuring Pilot-Affected Safety of Flight". Master internship report, University of Twente/University of Tokyo, Oct. 2013.

[5] J.O. Entzinger, M. Nijenhuis, T. Uemura, and S. Suzuki. "Assessment of Human Pilot Mental Workload in Curved Approaches". In *Proceedings of the 2013 Asia-Pacific International Symposium on Aerospace Technology (APISAT), Nov.20-22, 2013, Takamatsu, Japan*. Japan Society for Aeronautical and Space Science, Nov. 2013. (Peer Reviewed).

[6] H. Bhana. "Trust but Verify". *AeroSafety world*, vol. 5, 2010.

[7] J. Drappier. "Air Carrier Training Insights: Maintaining Manual Flying Skills". In *World Aviation Training Conference and Tradeshow*, Rosen Shingle Creek Resort, Orlando, Florida, USA, 2012.

[8] S.M. Casner. "Perceived vs. Measured Effects of

Advanced Cockpit Systems on Pilot Workload and Error: Are Pilots' Beliefs Misaligned with Reality?". *Applied Ergonomics*, vol. 40, no. 3, pp. 448–456, 2009.

[9] M. Causse, B. Baracat, J. Pastor, and F. Dehais. "Reward and Uncertainty Favor Risky Decision-Making in Pilots:Evidence from Cardiovascular and Oculometric Measurements". *Applied Psychophysiol Biofeedback*, vol. 36, pp. 231–242, 2011. doi: 10.1007/s10484-011-9163-0.

[10] J. Aasman, G. Mulder, and L.J.M. Mulder. "Operator Effort and the Measurement of Heart-Rate Variability". *Human Factors*, vol. 29, no. 2, pp. 161–170, 1987.

[11] K.J. Vicente, D.C. Thornton, and N. Moray. "Spectral Analysis of Sinus Arrhythmia: A Measure of Mental Effort". *Human Factors*, vol. 29, no. 2, pp. 171–182, 1987.

[12] A.F. Kramer. "Physiological Metrics of Mental Workload: A Review of Recent Progress". *Multiple-task performance*, pp. 279–328, 1991.

[13] J. Veltman and A. Gaillard. "Physiological Indices of Workload in a Simulated Flight Task". *Biological psychology*, vol. 42, no. 3, pp. 323–342, 1996.

[14] V.J. Gawron. *Human Performance, Workload, and Situational Awareness Measures Handbook, Second Edition*. Taylor & Francis, 2008. ISBN 9781420064506.

[15] J. Beatty. "Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources". *Psychological bulletin*, vol. 91, no. 2, pp. 276–292, 1982.

[16] E. Van Der Meer, R. Beyer, J. Horn, M. Foth, B. Bornemann, J. Ries, J. Kramer, E. Warmuth, H.R. Heekeren, and I. Wartenburger. "Resource Allocation and Fluid Intelligence: Insights from Pupillometry". *Psychophysiology*, vol. 47, no. 1, pp. 158–169, 2010.

[17] S.P. Marshall. *Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity*. Google Patents, Jul. 18 2000. US Patent 6,090,051.

[18] G.D. Clifford. "ecgbag Software". Online, 2010. URL `www.mit.edu/~gari/CODE/ECGtools/ecgBag/`.

Copyright Statement