ON THE CONNECTION BETWEEN PILOT'S PERFORMANCE AND INCREASING COMPLEXITY OF FLIGHT SIMULATION SYSTEMS

J.-Michael Bauschat
Institute of Flight Systems
German Aerospace Center (DLR)

Keywords: *Aircraft, Flight-Control, Man/Machine Interaction, Modeling, Model-Following Control, Simulator, Pilot-Tasks, Workload*

Abstract

In the Introduction a project is described, which deals with the comparison of ground-based simulation and real flight. The project simulators and a particular kind of real-time simulation technique, namely in-flight simulation, is explained briefly. A particular low gain / high precision pilot task is presented exemplary in greater detail. The data evaluation in the project is based on time-dependent, frequency-dependent and subjective workload data as well. The description of two subjective workload assessment approaches which are in use in the project is given. Three different kinds of simulation systems have been in use: a ground-based simulator, a full flight moving-based simulator and a nonlinear in-flight simulation using the DLR flying test-bed ATTAS. The evaluation pilots have rated their workload with the above mentioned methods. Selected results gained in the project are shown.

1 Abbreviations

ATTAS	Advanced Technologies Testing Aircraft System
DLC	Direct Lift Control
DLR	German Aerospace Center
FFT	Forcing Function Task
FTD	Flight Training Device
MCHR	Modified Cooper Harper Rating
RPT	Reproducible Pilot Task
ZFB	Zentrum für Flugsimulation Berlin

2 Introduction

Real-time flight simulation systems are used typically as pilot training devices, as engineering simulators and as research tools, if only professional applications are addressed. Independently from the application it can be stated, that three aspects are playing a central role for all mentioned systems: (1) *they reduce total costs*, (2) *they increase safety*, and (3) the *man in the loop aspect* is evident. Coming first to the monetary aspect: the initial costs as well as the running costs e.g. for a Level D Full Flight Simulator are definitively high. However, it is well-known that a training simulator has to replaces the real aircraft in an adequate way and that is the key argument. On the other hand there is the mentioned safety aspect. The training of difficult flight situations is cheaper and might be more efficient under simulation conditions than in flight. But, is the equation *Highest Simulation Standard equals Highest Education Level* always valid? What about item (3) in the above mentioned list? Needs the complex system *Human Operator* always a complex state of the art Simulation System for an efficient training?

One typical example is the application of motion cues. Simulated motion is not necessary in the case of so-called Flight Training Devices (FTDs) used generally, for initial and procedure training. These less complex simulators replicate the actual aircraft cockpit, but do not provide a visual system or motion system. On the other hand, it is well known that the pilot's behavior is influenced by the aircraft motion in the case of high precision tasks, or emergency conditions caused e.g. by bad weather such as heavy wind shear, turbulence, and gusts. High gain tasks increase the workload of the pilot significantly. Hence, a deeper understanding of these subjectively sensed influences on the pilot's reactions is necessary.

The systematical investigation of the effect of increasing simulation reality level on pilot's performance is one aspect of a research project at the Institute of Flight Systems at the German Aerospace Center DLR.

Fig. 1. DLR Flying Test-Bed VFW 614 ATTAS

One outcome of the project should be a model-based understanding of the interaction between the two highly nonlinear systems pilot and aircraft. To meet this requirement a clear project road map was defined, which is being followed consequently. Important contributions are coming from three very particular *simulation systems*. They cover all aspects and project demands from ground-based simulation to real flight. Well defined pilot tasks round off the experiment setup. After this brief overview the most important aspects as well as selected results are discussed in greater detail in the following sections.

3 Flight Simulation Facilities

3.1 ATTAS In-Flight Simulator

ATTAS (Advanced Technologies Testing Aircraft System) is based on a VFW 614, twin-turbofan, short-haul 44-passengers a/c (Figure 1). The VFW 614 is ideally suited as a general purpose test-bed due to the size, cabin space, loading capacity and flight performance. With full fuel, about 3.5 tons of test equipment can be loaded. The heart of ATTAS is the fly-by-wire/light system, which is based on a multi computer system. The mechanical control system serves as a backup in the case, that an experiment software generates a failure case. This approach makes it possible to use software in the ATTAS flight-test, which is not certified. To meet safety requirements the system has been designed as a two channel computer network consisting of four processors in each channel with one common central processor for communications and data recording.

To give ATTAS a 5-DOF simulation capability, five independent control surfaces must be available. Therefore, ATTAS was equipped with a specifically developed 'Direct Lift Control' System (DLC) for pitch/ heave motion decoupling and gust/load control. The rear parts of the landing flaps have been divided into six (three on each wing) fast moving flaps having about 85 deg/sec flap rate and +35 degrees flap deflection capability for high frequency direct lift modulation.

3.1.1. ATTAS Nonlinear In-Flight Simulation

The aim of in-flight simulation is to imprint the characteristics of a vehicle to be simulated on a

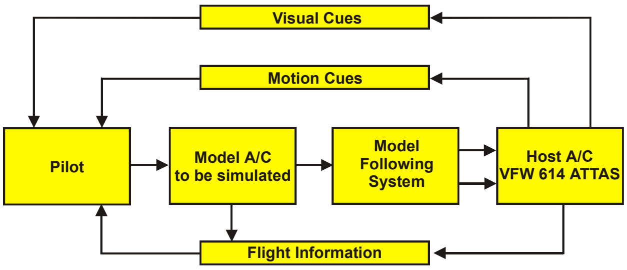

Fig. 2. Rough Structure of ATTAS In-Flight Simulation

host a/c. Figure 2 gives a survey of this particular technique.

The figured elements have the following tasks and qualities:

- Pilot: He has to perform prepared simulation tasks, which depend, for instance, on the scientific work to be done. Simulating the automatised flight of an a/c, the pilot would be replaced by a digital flight control system. In this case he would have the task to observe the simulated flight.
- Model a/c to be simulated: In the discussed system this part contains the algorithms which are needed to simulate a nonlinear model a/c under real-time conditions.
- Model following system: It consists of a combination of feedforward and feedback control laws.
- Host a/c: The essential part of the in-flight simulation is an adequate flying test-bed. It must be equipped with a fly-by-wire control system. A data acquisition system must be available.

The principle is as follows: The pilot has direct control of the computed nonlinear model with his inputs. This technique is identical to that found in ground-based simulation. In an explicit model following system the idea is to make selected model states and their changes due to time available to the model following system. The computed outputs from the model following system are inputs to the actuators of the host a/c. The actuators affect the required deflections of the control surfaces (elevator, direct lift control (DLC), flaps, aileron, rudder) and the required thrust. Using this simulation technique has the advantage, that the pilot has the real visual cue and the correct motion stimuli of the model a/c (Fig. 2). Information needed by the pilot are attained from an electronic flight information system (EFIS).

The in-flight simulation technique makes it possible to simulate two different transport aircraft in this project. One is the VFW 614 itself (weight of about 20 tons). The second one is a typical widebody transport aircraft with a weight of about 115 tons, which is simulated nonlinear in flight.

3.2 ATTAS Ground-Based Simulator

The real-time simulation of ATTAS on ground is an important tool. It simulates the a/c as precisely as it is possible without motion cue, but an with a visual system. The aboard data-processing system consists, as in the real a/c, of MIL-specified computers. An original ATTAS-cockpit belongs to the simulator. Every experiment flown on the in-flight simulator can be tested on the ground. The scientist is able to validate his software and obtain first results. Another important fact is to train the pilots and find out their opinions before being airborne. The ground simulation reduces the costs and also a lot of development risks. The standard of the ATTAS ground-based real-time simulator allows the realisation of typical experiments concerning simulation technique.

3.3 Level D Full Flight Training Simulator

Part of the framework of this scientific project is the particular full flight training simulator of the *Zentrum für Flugsimulation Berlin (ZFB)*. The company operates a moving based transport aircraft real-time simulator. This simulation system has two faces: one is the training simulator and the second is the research simulation option. As a research facility the simulator provides the scientist with full access to nearly all software and hardware systems. The simulator can be used as it is with a good data acquisition system, but can also be

Fig. 3. Full Flight Simulator of ZFB

modified concerning the model dynamic, Displays, etc.

4 Reproducible Pilot Tasks

Part of the discussed project framework are particular reproducible pilot tasks (RPTs), which the evaluation pilots have to perform. The specially designed RPTs are precision/low-gain and precision/high-gain tasks. The first category is covered by the use of an artificial instrumental landing system, which is computed in real-time. The pilot tasks of the second category are mainly based on different kinds of tracking tasks, where the pilot has to follow a given cockpit indication and to compensate for offsets. The layout of the tracking tasks has been developed using sophisticated algorithms considering the aircraft dynamics as well as the pilot´s reaction.

4.1 The Localizer Intercept Task

The in-flight simulation software package includes an artificial instrumental landing system (ILS). This tool provides the evaluation pilot with an approach to a runway at high altitude for training or scientific purposes. All navigation aids are simulated and their signals are indicated on the screens in the ATTAS cockpit. At higher altitude test-pilots can perform ILS approaches without disturbing the traffic or Air Traffic Control (ATC) of an actual airport.

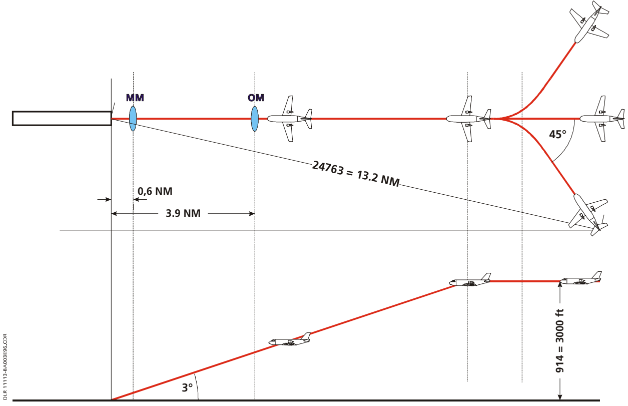

Fig. 4. Setup of the Localizer Intercept Task

The Localizer Intercept Task is based on the artificial ILS. The pilot has the task to intercept the localizer from one side (intercept angle 45 deg). The intercept itself is a heading change to the selected runway heading. This part is followed by a descent, which starts at the indication of the incoming glideslope. The task ends at an altitude of 200 feet above the artificial runway. The evaluation pilot then has to perform a go-around. Simulated abnormal flight conditions (e.g., engine failure), and demanding environmental conditions (e.g., turbulence) are introduced.

4.2 The ILS-Tracking Task

The task is again based on the above mentioned artificial ILS. The ILS-tracking task in general causes the pilot to compensate for generated offsets in localizer and glidepath. Programmable CRTs allow it to generate a symbol, which indicates to the pilot an offset from a given fixed reference.

For this tracking task a glidepath-, a localizer transmitter and a DME are necessary. The cone effect is also simulated. The experiment starts usually between FL 160 and FL 200. The pilot has to stabilize the a/c in a given configuration and a given speed with a climb angle of -1.5. The flight-test engineer starts the tracking with a switch. The data on the test-pilot‘s CRTs , like heading, altitude, speed and distance are related to the simulated ILS.

After a given time the localizer symbol moves e.g. to the right, indicating that the a/c is left to the localizer beam. The pilot now has a given time to compensate this offset and so on. While approaching the transmitters the degree of difficulty of the task increases because of the cone effect. The indication becomes more and more sensitive.

4.3 The Forcing Function Task (FFT)

A special tracking task has been designed by v. Gool et al. [9] for the longitudinal motion of an aircraft, which can be adapted to the lateral motion. The command sequences seem to be random, but they are a based on a predefined set of several sine functions. The main difference between the available FFTs concerns the amplitude and not the frequency.

The task is indicated to the evaluation pilot on his primary flight display. In the center a fixed square is displayed, which allows the reading of the actual pitch angle. The pitch angle is indicated on a moving scale. The aim is to hold a fixed square in the circle of a birdy symbol, which would be the desired performance.

5 Database, Data Evaluation and Performance Assessment

The data evaluation is based here on different kinds of data representations and data sources. All described simulators have their own acquisition systems to store all interesting data. The data evaluation following the experiments is done in the time domain as well as in the frequency domain (e.g. power spectral density of the stick command). It was found out, that objective data cannot answer all questions concerning the man/machine system. The database had to be supplemented by subjective workload data. Two subjective methods to evaluate workload have been applied in this project.

5.1 NASA Task Load Index (NASA-TLX)

Hart et al. [7] describe the results of a study, which ended in the proposal of a multi-dimensional rating scale. NASA-TLX is based on six subscales representing possible sources of workload. These subscales are:

Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Own Performance (OP), Effort (EF), Frustration Level (FL).

In a first step the subject has to weigh the subscales by performing fifteen pair-wise comparisons. Fifteen comparisons are required to decide which of each pair of the six factors is the most significant in creating the level of workload experienced in performing a particular task. In a second step the subject has to give a workload rating for each of the six subscales. E.g. in the case of PD it is a scale between Low and High, which corresponds with an internal value between 0 and 100. The values gained in the pair-wise comparisons are used to weigh the magnitude ratings obtained for the six scales after each experimental condition. Hart et al. [7] describe the algorithm for obtaining a weighted workload score, the so-called TLX-value, from the subscale scores and the weighting values. The TLX-value is a number between 0 (no workload) and 100 (overload).

5.2 Modified Cooper/Harper Rating Scale

The way to determine the handling qualities of an aircraft or rotorcraft is typically based on the Cooper/Harper rating scale described in [5]. A test-pilot can give a handling quality rating between 1 (excellent) and 10 (severe deficiencies). Wierwille et al. ([11]) found a way to use the Cooper/Harper scale to determine mental workload by modifying it. In their scale 1 stands for an easily solvable instructed task, operators mental effort is minimal and desired performance is easily attainable. Accordingly, 10 means impossible - instructed task cannot be accomplished reliably.

5.3 Assessment of pilot's performance

The performance of an evaluation pilot can be assessed first of all on the basis of the objective data. It is possible for all mentioned Reproducible Pilot Tasks to specify ranges of desired, adequate and not adequate performance. Taking the Localizer Intercept Task as an example, the localizer offset, the glide path offset and the error in speed as well can be assessed that way.

5.3.1 The Cost Function

An adequate approach to assess the performance is given by the cost function $I(t)$ in Equ. (1), which includes an error trend. Based on the given definitions for desired and adequate performance a relative error can be determined and summed up. In the following the principle of a cost function will be described, which was designed for the Reproducible Pilot Tasks in particular.

$$I(t) = \int_{0}^{\infty} f\left[e(t)\right] dt \qquad (1)$$

The cost function for the whole task is a function of the error, which occurs in the main task and of the error in a given subtask. In a pitch task the pilot has to follow e.g. a commanded pitch angle, which is the main task. The subtask is in this case is to hold the wings in level (no bank angle).

The result of Equ. (1) increases with every deviation from the commanded attitude and is not easy to interpret. Fokken [6] developed a particular performance index (Equ. (2)), which supports the data analysis better than Equ. (1) and the result of Equ. (2) can be displayed in real-time in the cockpit. It provides the pilot with information about his actual performance and should be motivating.

5.3.2 The Optimum Profile Indicator (OPI)

The OPI represents pilots performance in each time step on a scale between 0 and 100%. A value of 100% indicates that the pilot solved the task at any time with *desired performance*. Zero percent is a synonym for *not adequate performance*. The Optimum Profile Indicator is calculated as follows:

$$OPI(t) = \left(1 - \frac{I(t)}{I_{max}(t)}\right) 100 \qquad (2)$$

6 Selected Results – Localizer Intercept Task

Some of the results gained in this project are shown and discussed in this section. Due to the fact, that not all Reproducible Pilot Tasks can be discussed in this article, only Localizer Intercept results are discussed here. Fig. 5 gives an impression concerning the control activity and command amplitude of one of the evaluation pilots. The evaluation pilot solved the task in the fixed-base simulation and in flight with ATTAS. The flight-tests included flights with the VFW 614 ATTAS itself and a widebody transport aircraft which was simulated in flight (see Fig. 2). In the simulator large stick deflections can be observed, while the deflections in the flight-test are significantly smaller for both aircraft. That indicates differences in the control behavior depending on the presence of motion and the level of realism.

An idea about differences in the subjective workload ratings illustrates Fig. 6, where the TLX subscales are plotted. The exemplary ratings were given by one evaluation pilot. He performed Localizer Intercept Tasks in the

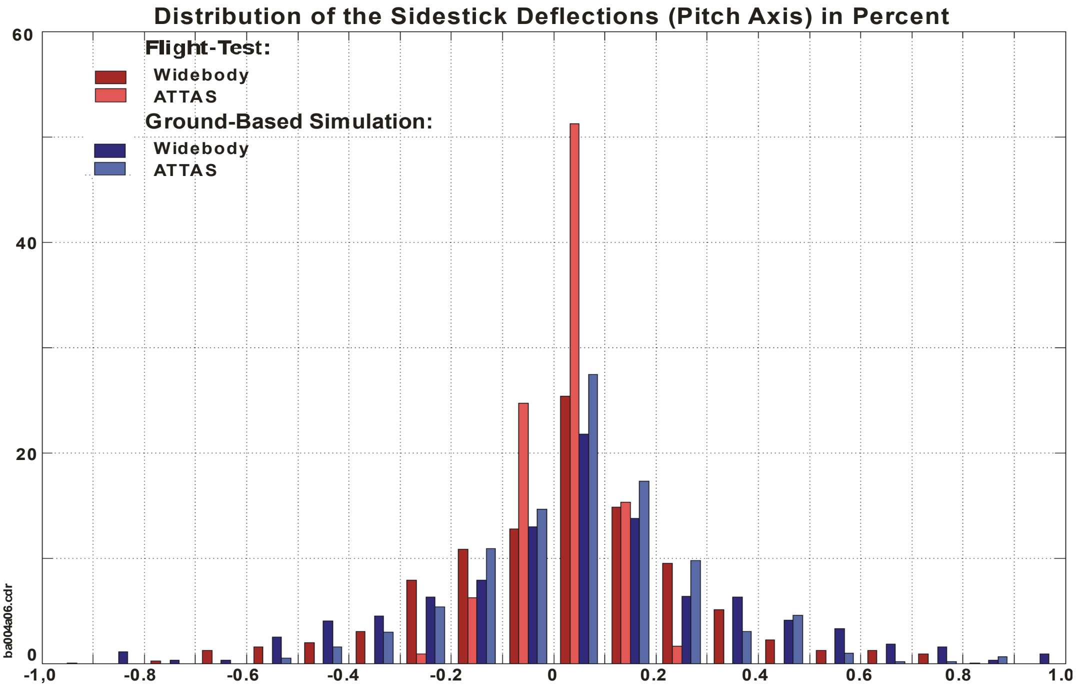

Fig. 5. Distribution of Sidestick Pitch Position Values

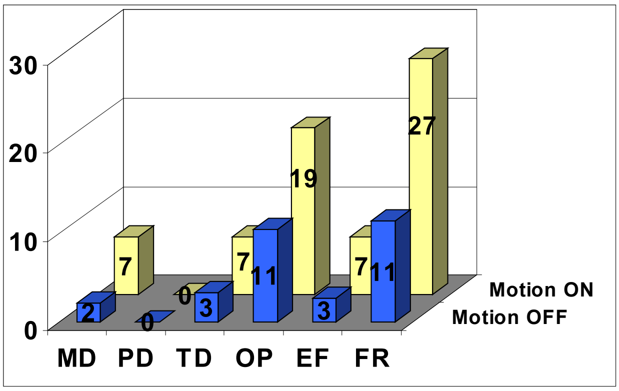

Fig. 6. TLX Subscales of LocIC Flights

motion base simulator which was described in section 3. One of the tasks was performed with active motion system, in the other case the motion was switched off. The task performed under the influence of motion got a TLX workload rating of 65%. This value is comparable to a MCHR of 6,5: *major difficulty – maximum operator mental effort required to bring errors to moderate level*. The task performed without active motion system was rated TLX=30% (*MCHR=3: fair, mild difficulty – acceptable operator mental effort is required to maintain adequate system performance*). This result shows the typical finding: concerning the Localizer Intercept Task it can be stated, that the influence of motion leads in a simulator to significant higher workload levels.

Fig. 7 shows the influence of the motion cue on the task solution and on the pilot (see also Heine [8]). Again particular findings concerning the Localizer Intercept Task performed in the Full Flight Simulator with and without motion are illustrated. The figure shows the average improvement for lateral deviation, horizontal deviation, roll command activity, pitch command activity and workload, when the motion system was switched off. A clear performance improvement can be observed (decreasing localizer deviation: 8% and decreasing glideslope deviation: 15%). The project pilots indicated a decreasing subjective workload of 25% during the task, without motion cue.

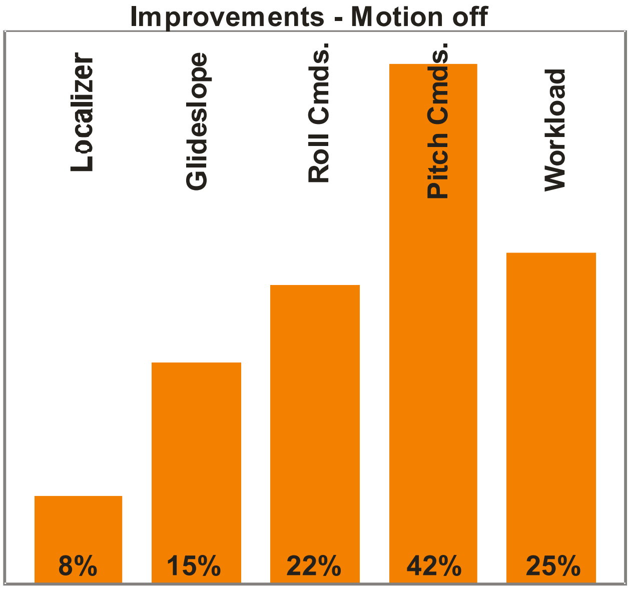

Fig. 7. Influence of the Motion on the Task

One possibility to build a bridge between real-time simulation and flight-test shows Fig. 8. Results gained in the ground-base simulator (no motion system) and the flight-test are illustrated. The Localizer Intercept Task was subdivided into 4 phases to allow a clearer analysis:

Phase 1: approach on intercept course
Phase 2: localizer intercept and stabilizing runway heading
Phase 3: glideslope intercept, stabilizing the descent
Phase 4: final

Fig. 8 shows the averaged duration of the sidestick inputs in the pitch and the roll axes in relation to the mentioned phases. The figure makes it obvious, that the pilots have a clear tendency to give much longer control commands in the flight-test. So, in this case it can be shown that they have obviously different control strategies to solve the task adequately in the fixed-base simulator and in real flight.

7 Conclusions

An experiment has been described where a fixed-based, a moving-based and an in-flight simulator are used to improve the understanding of the existing dependence between pilot task, pilot performance and pilot workload. The different pilot tasks, which were designed in the project, have been explained briefly. The evaluation pilots have to perform both so-called precision low-gain and precision high-gain tracking tasks. The results from one particular

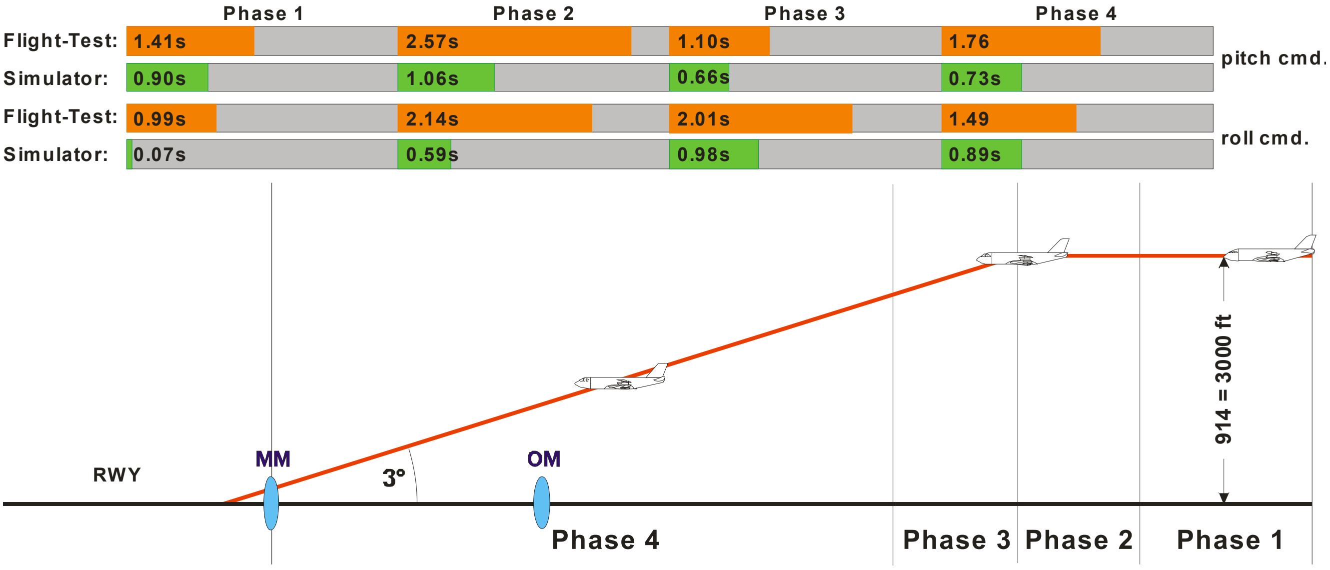

Fig. 8. Averaged Duration of Sidestick Command Inputs (Schmerwitz [10])

task, the Localizer Intercept Task (LocIC), has been discussed in this paper in greater detail. In this case the workload is determined using subjective determination methods: NASA-TLX and the Modified Cooper/Harper Rating (MCHR) Scale. NASA TLX as well as the Modified Cooper/Harper Rating has a good acceptance by the evaluation pilots. It is easily possible to get subjective workload data in the simulator and in the flight-test after each task using, for example, a laptop computer.

The preliminary results concerning the comparison of different real-time simulation techniques, which have been achieved to date in the project already show some very interesting aspects leading to the following conclusions:

- The performance of a pilot during demanding tasks in a real-time environment is clearly influenced by the workload generated by the task. In all cases the workload ratings given by the evaluation pilots have values which indicate that they can not be neglected.
- The problem to get physiological workload data can be avoided in the case, that an experiment requires a more general impression about pilot's workload. In such a case subjective workload information are sufficient.
- The occurrence of a conspicuous task performance (one example of which has been shown in the paper) can be better understood if typical performance data is supplemented by workload information.
- The more demanding a pilot task is, the more significant the difference between ground-based real-time simulation and flight-test.
- The result from the high precision / low gain task Localizer Intercept illustrate, that also a state of the art motion system may have negative influence on the task performance, the control behavior and the training effect of a pilot.

References

[1] Bauschat, J.-M. On the Application of a Nonlinear In-Flight Simulation Technique, *Proceedings of the First European Control Conference (ECC 91)*, Grenoble (France), Vol. 3, pp. 2415-2422, (1991)

[2] Bauschat, J.-M. The Role of In-Flight Simulation for the Definition of Simulation Fidelity Criteria, *ICAS/AIAA CP Vol. 1*, Anaheim (USA), Paper ICAS-94-8.1.2, pp. 261-271, (1994)

[3] Bauschat, J.-M., Gestwa, M., Leißling, D.; A Score Monitoring System to Support Evaluation Pilots in Flight, *Conference Proceedings of the AIAA Modeling and Simulation Technologies Conference*, paper AIAA-2004-5444, Providence, Rhode Island, 16-19 August (2004)

[4] Bussolari, S.R., Lee, P.D.A. The Effects of Flight Simulator Motion on Pilot Performance and Simulator Acceptability in Transport Category Aircraft, *Proc. 2éme Colloque International, La Securité Arienne*, pp. 361-371, (1986)

[5] Cooper, G.E. & Harper, R.P. The use of Pilot Rating in the Evaluation of Aircraft Handling Qualities“, *NASA TN D-5153*, (1969)

[6] Fokken, M. Ein Cockpit-System zur Anzeige der Lösungsqualität von Pilotenaufgaben in einem Forschungsflugzeug, *Diplomarbeit am Institut für Luft- und Raumfahrt der Technischen Universität Berlin*, (2000)

[7] Hart, S.G. & Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Human Mental Workload*, North-Holland, Amsterdam, pp. 139-183, (1988)

[8] Heine, U. Ein kriteriengestützter Vergleich zur Untersuchung der Lösungsqualität einer niederfrequenten Präzisionspilotenaufgabe auf der Basis verschiedener Echtzeitsimulationssysteme, *Diplomarbeit am Institut für Luft- und Raumfahrt, Technische Universität Berlin*, (2002)

[9] v. Gool, M.F.C, & Hanke, D. & Lange, H.-H. Flight Path Angle Tracking Experiments in the DFVLR HFB320 Equipped with Direct Lift Control, *Interner Bericht IB 154-75/32*, Institut für Flugmechanik, DFVLR Brauschweig, (1975)

[10] Schmerwitz, S. Untersuchung einer Präzisionsflugaufgabe bezüglich der Pilotenlösungsstrategie im Simulator und im Flugversuch, *Studienarbeit am DLR-Institut für Flugsystemtechnik in Braunschweig und Technische Universität Braunschweig*, (2003)

[11] Wierwille, W.W. & Casali, J.G. A Validated Rating Scale for Global Mental Workload Measurement Applications, *Proceedings 27th Annual Meeting, Human Factors*, (1983)