JOURNAL OF AIRCRAFT
Vol. 48, No. 3, May–June 2011

Factors That Influence Pilot Task Demand Load During Area Navigation Approaches

Monique Heiligers,* Theo van Holten,* and Max Mulder*
Delft University of Technology, 2629 HS Delft, The Netherlands

DOI: 10.2514/1.C031188

The goal of this research is to develop a method that can predict pilot task demand load during approaches. This paper presents the results of a flight simulator experiment that aimed to give a first indication of the factors related to the approach trajectory that influence pilot task demand load. Analysis of the flight simulator data showed that the following factors influence pilot task demand load: the energy rate demand during the last parts of the approach, localizer-intercept speed, the distance available on localizer-intercept heading, the lineup distance, the required altitude at the final approach fix, localizer-intercept angle, and whether the approach is stabilized at 1000 ft. The results of the flight simulator tests are also used to validate the predictions of a Monte Carlo simulation. The Monte Carlo simulation predicts whether it is possible to meet the constraints at the waypoints and whether it is possible to achieve a stabilized approach, which are two of the factors that were shown to influence pilot task demand load. A significant correlation was found between these Monte Carlo predictions and the flight data from the flight simulator experiment.

I. Introduction

THIS research aims to develop a method that predicts the task demand load (TDL) as experienced by the pilot while flying an approach. TDL is defined as the mental workload imposed by the system to be controlled or supervised on the human operator [1]. First, this will yield insight in which aspects of an approach actually influence pilot TDL. Second, during the design of approaches, this method can be used to rapidly evaluate a potential approach and to optimize an approach with respect to pilot TDL.

The method should predict pilot TDL when flying a published approach according to standard operating procedures (SOPs) when using autopilots, autothrottle, and the flight management system (FMS), and while aiming to achieve a stabilized approach (the criteria for a stabilized approach will be explained in Sec. IV.B). The goal is to link pilot TDL, first of all, to the properties of the approach trajectory (for example, the number of waypoints, localizer-intercept speed, etc.); next, the effects of other factors such as wind conditions and aircraft weight are considered. For a definition of a stabilized approach, as well as the basic principles of the method, the assumptions and the choices that have been made as to what is and what is not incorporated in the scope of the research, see [2,3].

As explained in [2,3], the approach we have chosen to predict pilot TDL deliberately deviates from the idea behind models such as the procedure-oriented crew model [4,5] or the man–machine integrated design and analysis system [6–8] that use human operator models. The approach we have chosen is based on the principles of cognitive work analysis [9]. The main characteristic of cognitive work analysis is that it shifts the emphasis from investigating the constraints of the human operator (like memory capacity, time delay, etc.) to analyzing and describing the operator environment (like the trajectory, the aircraft dynamics, the wind conditions, etc.). The reason for this choice is that the constraints in the environment actually shape the behavior of the human working in that environment. By choosing the approach to focus on the operator environment instead of the human operator, the model will be more accessible and easier to use for approach designers, who often have little background in human operator modeling (or none).

The research is split up into five steps. The first step is to determine whether, for a specified approach, it is possible to meet the constraints at the waypoints and to achieve a stabilized approach, since it will be shown in this paper that, among others, pilot TDL during approach is influenced by these two factors. This first step was the topic of the accompanying paper [2] that, to this end, described a Monte Carlo simulation based on an aircraft model for the Boeing 747-100 (B747-100) and an initial pilot model. This pilot model consisted of a simple control model for the flight director task and a model for the pilot actions, such as selecting flaps and gear that were modeled using trigger events (for example, reaching 1200 ft) and distributions for reaction times (for instance, 2 s after reaching 1200 ft, flaps 25 are selected); see Fig. 1.

Based on assumed reaction-time distributions, the Monte Carlo simulation [2] predicted, for a given approach, the percentage of flights that can meet the altitude and airspeed constraints at the waypoints and the percentage of flights that can achieve a stabilized approach as a function of wind direction, windspeed, and aircraft mass.

The second step is to validate the predictions of the B747-100 Monte Carlo simulation for a given approach and to validate the assumptions that were made for the Monte Carlo simulation with respect to the modeling of all pilot actions. This is done by a first set of flight simulator experiments for the B747-100. These flight simulator experiments will also give a first indication of the other factors of an approach that influence pilot TDL. This second step is the topic of this paper.

In the third step, a second flight simulator experiment for the B747-100 will be performed to obtain a more detailed understanding of the factors that influence pilot TDL during approach. To check the general applicability of the factors found to influence pilot TDL, a Monte Carlo simulation, a flight simulator experiment, and real flight tests are performed for a different aircraft (a Cessna Citation) in the fourth step. The fifth and final step is to gather all relevant data and all factors that influence pilot TDL and to incorporate these in an easy-to-use tool that can be used during the design of approaches.

This paper starts with a review of literature (Sec. II); based on this review, the factors assumed to influence pilot TDL that are considered in this paper are presented in Sec. III. Section IV presents a human-in-the-loop flight simulator experiment, the results of which (both relating to the validation of the Monte Carlo simulation as well as relating to the factors assumed to influence pilot TDL) are given in Sec. V. A discussion of the results can be found in Sec. VI. In Sec. VII, a preliminary regression model is proposed that gives an indication of pilot TDL based on the numerical values of the factors that proved to influence pilot TDL. Section VIII then provides a

Received 23 July 2010; revision received 21 November 2010; accepted for publication 21 November 2010.

*Faculty of Aerospace Engineering.

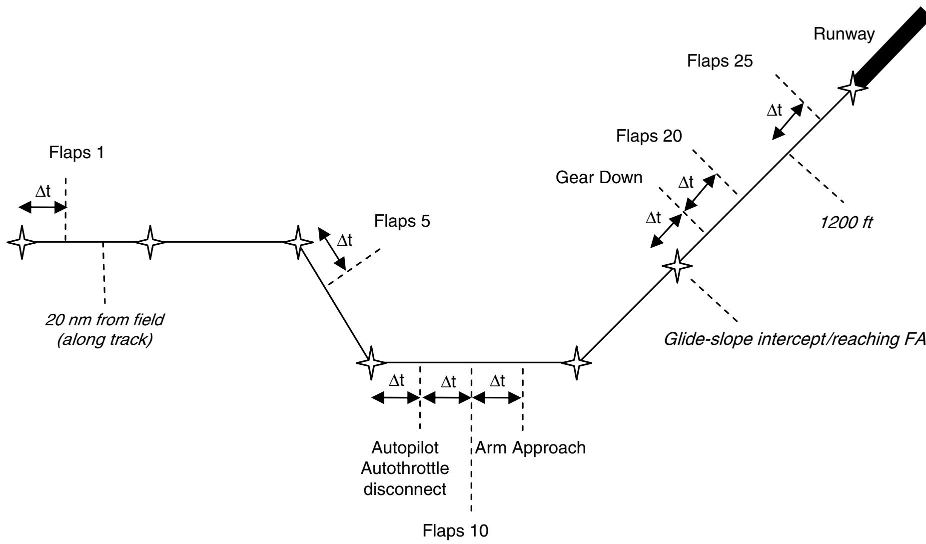


Fig. 1 Visualization of trigger events and reaction times Δt for pilot actions in the Monte Carlo computer simulation.

comparison between the human-in-the-loop experiment and the predictions of the Monte Carlo computer simulation. Finally, The conclusions are presented in Sec. IX.

II. Results of Literature Survey

The international standards for an RNAV approach using basic global navigation satellite system receivers are defined in International Civil Aviation Organization document 8168, volume 2 [(Procedures for Air Navigation Services Aircraft Operations (PANS-OPS)] [10]. These standards are given per approach segment:

1) The initial approach is the segment from the initial approach fix (IAF) until the intermediate fix (IF), and the IF is the first waypoint on a runway heading; see also Fig. 2. When used, the central initial approach segment has no maximum length [10]. The optimum length is 5.0 nm [10].

2) The intermediate approach is the segment from the IF to the FAF. If the approach is flown horizontally, the FAF is the location in the approach where the glide slope is captured. The length of the intermediate approach segment is variable but will not be less than 2.0 nm, allowing the aircraft to be stabilized before overflying the final approach fix (FAF) [10].

3) The final approach is the segment from the FAF to the runway threshold. The optimum length is 5.0 nm, but it should normally not exceed 10 nm [10].

4) The profile descent path should have an angle no greater than 3.7 deg, with an optimum descent angle of 3 deg [10].

It is not stated whether these optimum lengths and optimum descent angles are also optimal with respect to pilot TDL.

The Australian Transport Safety Board has performed an extensive survey of pilots to gain an understanding of pilot perceptions of RNAV approaches [11]. This survey was held among pilots with a Civil Aviation Safety Authority (Australian) pilot license; not all respondents were flying vertical navigation (VNAV) equipped aircraft. The findings of the Australian survey for the category of airline pilots are reported in Table 1. It should be noted that only factors that relate to the nominal trajectory are mentioned; factors concerning, for example, the naming convention of waypoints, improvement of approach charts, etc., although reported in [11], are not mentioned here. Table 1 only reports the most frequently mentioned answers.

Vormer [12] studied pilot TDL as a function of different approach trajectories and different four-dimensional (4-D) guidance displays. Twenty-two metrics, all related to the approach trajectory, were evaluated for estimating TDL during flight simulator tests. The metrics are given in Table 2. During the flight simulator tests, pilots were required to fly manually using 4-D guidance displays and were instructed to minimize the deviation from the reference flight path and reference speed. Based on task load index (TLX) workload ratings and control activity, nine metrics of the total 22 were found to influence pilot TDL (see Table 2). The energy rate demand in Table 2 is the ratio between the energy rate commanded by the trajectory and the maximum energy rate that can be achieved by the aircraft. If the value for the energy rate demand becomes larger than one, this implies, in the case of approaches, that the decrease in energy required by the trajectory cannot be met by the energy decrease of the aircraft; as a consequence, the altitude and/or airspeed constraints at the next waypoint cannot be met.

III. Task Demand Load Factors Considered

Within the present research, we specifically focus on pilot TDL during RNAV approaches, during which the part of the approach until localizer-intercept heading is flown using the autothrottle and

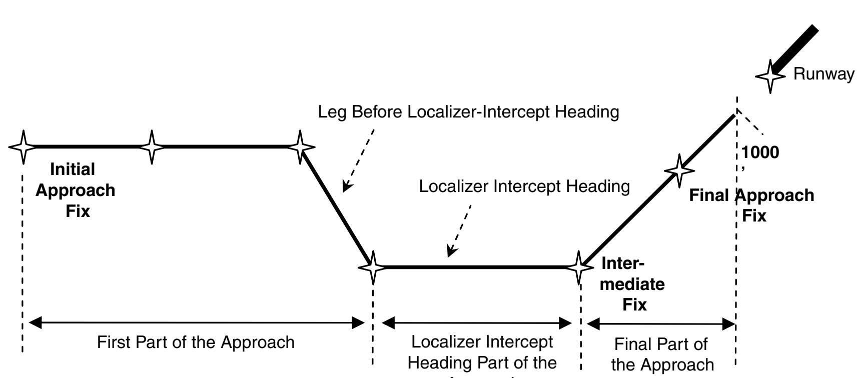


Fig. 2 Definitions relating to the approach trajectory used in this research.

Table 1 Findings of Australian survey of pilots to gain understanding of pilot perceptions of RNAV approaches [11]

Aspect nos.	Contributions to mental workload, physical workload, or time pressure	Aspects that can be improved	Most difficult circumstances
1	Programming flight management computer, setting up approach	FAF (and steps after FAF) removed from design (10 nm last segment)	Poor weather conditions
2	Varying/irregular segment lengths, many (close) steps	Reduction in number of waypoints/steps or removal of short steps	Turbulent conditions
3	Descent and position monitoring/situational awareness	Standard (PANS-OPS) distances between all waypoints/standard missed approach point position	Night
4	Aircraft configuration late or early increases or decreases workload, respectively	Reduced GPS/FMS inputs	Instrument meteorological conditions
5	Global Positional System (GPS)/FMS manipulation	3 deg slope only	Significant terrain
6	Late decision/clearance to fly RNAV approach	Runway alignment on all approaches	Speed too fast (rushed or tailwind)
7	Briefing	Overlaid approaches/waypoints matching ground-based aids	Short sectors (limited preparation time)
8	Early preparation reduces workload/time pressure		Multiple (short) limiting steps/complex approach design
9			Approach not runway-aligned
10			Short notice from ATC or limited preparation time
11			Traffic

autopilot in VNAV and lateral navigation (LNAV) modes, and at localizer-intercept heading, the pilot switches to flight director and disconnects the autothrottle (see Sec. II of [2]). The last part of the approach is flown using the instrument landing system. In this respect, the results of the studies described previously [11,12] cannot be directly applied to this research due to the different types of automation used: the study of Vormer [12] considered manual control throughout the flight, and the survey described in [11] included aircraft that were not VNAV equipped. These different levels of automation might result in different factors influencing pilot TDL. However, the two studies considered provide a good basis to start from. Therefore, we have chosen to concentrate on the following nine independent variables that might influence pilot TDL for the experiment described in this paper:

1) The first variable is the number of heading changes. Although Vormer [12] found that lateral maneuvers did not appear to increase workload, the number of heading changes might contribute to the complexity of approaches, which was found to be one of the most difficult circumstances for RNAV approaches in [11].

2) The second variable is incorporating many altitude steps in an approach compared with a continuous descent approach (CDA). Many altitude steps will increase the cumulative size of vertical path changes, the longitudinal maneuver time, and the maximum vertical acceleration for vertical path changes, which were all found to increase pilot TDL [12].

3) The third variable is the maximum energy rate demand. By applying a strong tailwind, the maximum energy rate demand is increased; this would increase pilot TDL [11] and is mentioned as one of the most difficult circumstances to conduct an RNAV approach [12].

4) The fourth variable is applying a horizontal approach instead of a CDA. According to [12], this will increase pilot TDL by increasing the cumulative size of vertical path changes, the longitudinal maneuver time, and the maximum vertical acceleration for vertical path changes. On the other hand, pilots might be more accustomed to flying a horizontal approach and intercepting the glide slope from below, which might decrease the pilot TDL compared with flying a CDA.

These four factors were based on the findings in the literature [11,12]. Next to these factors, other factors that might contribute to pilot TDL during approaches are considered as well in this paper. These other factors are selected based on conversations with pilots and by examining the SOPs (see [2], Sec. III.C). The additional factors used as independent variables are as follows:

5) The fifth variable is the distance available on localizer-intercept heading, since according to the SOPs, many actions need to be performed on localizer-intercept heading; decreasing the distance (and thus time) available is assumed to affect pilot TDL.

6) The sixth variable is localizer-intercept speed; this is equal to the airspeed at the IF.

7) The seventh variable is aircraft mass.

8) The eighth variable is the lineup distance, this is the distance between IF and runway, and

9) The ninth variable is the heading change when turning toward localizer-intercept heading; in Fig. 2, this is the heading change when turning from the leg-before-localizer-intercept heading to localizer-intercept heading.

To test the effect of the aforementioned factors on pilot TDL during approach, 20 different approaches have been designed. These were flown in a flight simulator by nine pilots. All approaches were also simulated by the Monte Carlo simulation [2].

Table 2 Metrics evaluated for estimating pilot TDL in [12]

No.	Effect on TDL	Metric
1	a	Number of vertical path changes
2	a	Maneuver time for vertical path changes
3	b	Cumulative size of vertical path changes
4	b	Maximum vertical acceleration for vertical path changes
5	a	Number of speed changes
6	b	Maneuver time for speed changes
7	a	Maximum size of change for speed changes
8	b	Maximum acceleration for speed changes
9	a	Number of combined vertical path and speed changes
10	a	Maneuver time for path-speed changes
11	b	Cumulative size of path changes for path-speed changes
12	a	Maximum size of speed change for path-speed changes
13	b	Maximum acceleration for path-speed changes
14	b	Maximum vertical acceleration for path-speed changes
15	a	Number of track changes
16	a	Lateral maneuver time
17	a	Cumulative size of track changes
18	a	Maximum rate of turn
19	b	Maximum energy rate demand
20	a	Total number of maneuvers
21	b	Longitudinal maneuver time
22	a	Total maneuver time

[a]This factor does not influence pilot TDL, according to [12].
[b]This factor influences pilot TDL, according to [12].

IV. Human-in-the-Loop Experiment

A human-in-the-loop experiment was performed that involved pilots flying different approaches under varying conditions in a six-degree-of-freedom flight simulator.

A. Experiment Goal

The experiment was designed to test the influence of the nine independent variables (as explained in the previous section) on pilot TDL during an approach. Next to testing the influence of these independent variables, additional goals of the experiment were 1) to validate the Monte Carlo simulation (presented in the accompanying paper [2]) with respect to the simulation of the pilot's actions; 2) to validate the Monte Carlo simulation with respect to the prediction of whether or not the constraints at the waypoints are met; 3) to validate the Monte Carlo simulation with respect to the prediction of whether or not the approach is stabilized at 1000 ft; 4) to get an indication of which other factors besides the independent variables have an effect on pilot TDL; and 5) to test whether pilots are aware of the existence of an optimal RNAV approach according to the standard or optimal PANS-OPS distances, as described in Sec. II (since designing approaches according to the optimal approach criteria in [10] is one of the most common improvements for RNAV approaches mentioned in [11]).

B. Method

1. Apparatus and Boeing 747 Model

The experiment was performed in Delft University of Technology's six-degree-of-freedom Research Institute for Simulation, Motion, and Navigation (SIMONA) research simulator (SRS) with out-the-window view; see Fig. 3. The Boeing 747 (B747) aerodynamic models, as well as the autopilots, flight director, autothrottle, and yaw damper, are based on the B747 model with JT9D-3 engines, given in [13], and are identical to the models used in the Monte Carlo simulation [2]. Autopilot modes available during the experiment were VNAV, LNAV, heading hold, heading select, altitude hold, and vertical speed. The in-flight director operation additional modes available were glide-slope mode and localizer mode. In VNAV mode, a drag-equired message appeared on the navigation display when altitude and speed constraints at the next waypoint were calculated not to be met (energy rate demand too high) with the current aircraft configuration. The approaches were flown in visual flight rules conditions. During the experiment, there was no other traffic, and no emergencies (for instance, engine fire) occurred.

All approaches were preprogrammed in the FMS/control display unit (CDU). The appropriate approach was loaded in the FMS before the start of the approach, and during the experiment, pilots could switch between the progress and legs pages, but they could not use the CDU interactively or modify the approach.

There were some discrepancies between the SRS and a B747 that are of importance for the experiment. First, the cockpit layout in the SRS differed from reality (see Fig. 3). Second, the altitude setting on the mode control panel (MCP) did not overrule the VNAV mode. If the required altitude at the next waypoint was 5000 ft and the altitude set on the MCP was 6000 ft, the aircraft would continue to descend to 5000 ft instead of leveling off at 6000 ft. Third, the autothrottle did not consider the flap speed marks. For example, the aircraft was flying 230 kt with flaps up, and the up mark (the flap speed mark for flaps up) was at 220 kt. The required airspeed at the next waypoint was 200 kt. In reality, the autothrottle would not decelerate to 200 kt, but it kept the airspeed at the flap speed mark (in this case at 220 kt) until flaps 1 would be selected. This feature was not incorporated in the SRS and, as a result, the aircraft would decelerate to 200 kt, regardless of flap setting or flap speed marks. Fourth, the power levers did not move during autothrottle operation and, as a result, the power lever position at autothrottle disconnect was not necessarily the correct setting but needed adjustment by the pilot. These differences between the SRS and reality were the same for all pilots and all approaches.

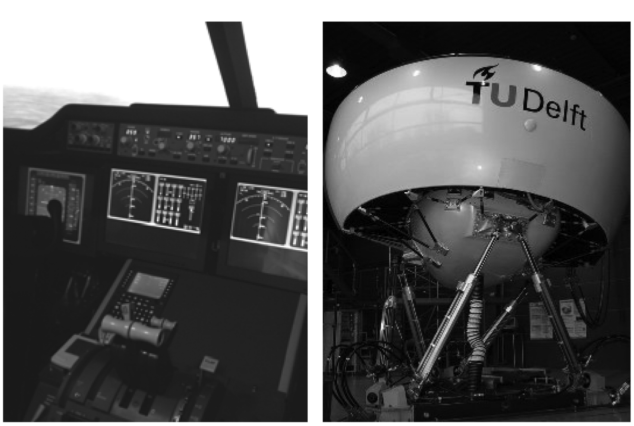


Fig. 3 SRS.

2. Subjects and Instructions

Nine B747 pilots participated in the experiment as pilots flying (PFs), with total flight hours ranging from 360 to 18,500 h ($M = 11,793$ h and $s = 6708$ h). Their flight hours on the B747 ranged from 200 to 8000 h ($M = 4228$ h, $s = 2800$ h). Two students of the Faculty of Aerospace Engineering were instructed and trained to act as pilots monitoring (PMs) during the experiment. Each crew consisted of one B747 pilot and one student. The task of the crew was to fly 18 approaches, starting at the IAF and ending at 800 ft above airport level. Some crews flew two additional approaches.

Two weeks before the experiment, the PFs received a briefing by mail. On the day of the experiment, they were briefed as well. The pilots were asked to adhere very strictly to SOPs as explained in Sec. III.C of [2], even if they could foresee that by adhering to SOPs they would not meet certain constraints at waypoints or would end up unstabilized at 1000 ft. The SOPs were briefed as follows to the pilots. The pilots were told that they were free to select flaps 1 and flaps 5 at a location in the approach they considered appropriate. They were told that they should select flaps 10, arm the approach, switch to heading select mode, and disconnect the autothrottle and autopilot at localizer-intercept heading. The pilots were additionally briefed that they should select flaps 20 and gear down at glide-slope intercept or, in the case of a CDA, that they should do so at the FAF and select that the flaps land (flaps 25) at 1200 ft altitude. The pilots were briefed that the only situation in which they were allowed to deviate from the flap settings according to SOPs was when the airspeed required by the approach was lower than the flap setting (according to SOPs) could accommodate; in that case, they were allowed to select the next flap setting. Additionally, they were asked to perform their tasks according to the principles of multiple crew coordination and to fly passenger comfort. They were briefed about the discrepancies between the SRS and the B747 (as explained in the previous paragraph), and they were informed that there would be no emergencies during the flight. They were told that they could fly the approach as published on the approach and landing charts, implying that air traffic control (ATC) would not interfere.

3. Procedure

Before starting the experiment, the pilots could familiarize themselves with the SRS and their task during three to five (depending on the pilot) practice approaches. After that, the experiment started. Before every approach, pilots could take as much time as they thought necessary to study the approach and landing charts, to brief the approach, and to prepare the SRS for the next approach. The simulation was started when the pilots indicated that they were ready.

After every approach, the PF was asked to fill in a feedback form. Each feedback form consisted of three parts: the first part required a rating of the approach on the rating scale mental effort (RSME) [14], the second part was an open-format question asking pilots to indicate which factors made the specific approach difficult or easy to fly, and the third part contained two questions asking pilot opinion on

Table 3 Independent variables and hypothesized TDL effects

Approach pair	Independent variable	Linked factor	Approaches	TDL effect
A	1) Number of heading changes	——	1 and 2	=
B	2) CDA compared with horizontal	——	3 and 4	+
C	3) Heading change toward localizer-intercept heading	——	9 and 10	=
D	4) Energy rate demand too high	Localizer groundspeed	17 and 18	+
E	5) Localizer-intercept speed (IAS)/localizer groundspeed	Energy rate demand IF-FAF	12, 13, and 14	+
F	6) Mass	——	20 and 21	+
G	7) More altitude steps compared with CDA	——	2 and 16	=
H	8) Distance available on localizer-intercept heading	Localizer groundspeed	3 and 9	+
I	9) Lineup distance (distance between IF and runway)	IF-FAF distance and energy rate demand IF-FAF	3 and 6	+

whether or not the approach was stabilized at 1000 ft and whether the pilot would have adhered to SOPs during real flight.

As stated, in the first part of the feedback form, the pilots were asked to rate each approach on the RSME. They were free to give one RSME rating for the entire approach or to divide the approach into multiple parts, giving each part a separate RSME rating. The RSME is constructed according to the magnitude estimation method [15], and the Dutch version of the scale (which was also used for this research) was used and validated in [16,17]. It is used here because of its simplicity and ease of use when compared with, for example, a NASA TLX rating procedure [18].

At the end of the day, after all approaches were flown, the PF filled in an end-of-day questionnaire. The first part of the end-of-day questionnaire regarded the realism of the SRS and the realism of the experiment as a whole. The second part contained general questions about factors that might possibly influence pilot TDL during approach.

4. *Independent Variables and Approaches*

The approaches (cases) were designed in pairs to test the independent variables; see Table 3. As explained before, the selection of the independent variables was partly based on [11,12]. All approaches were designed for Amsterdam Airport Schiphol runway 06. Within an approach pair, the independent variable was the only changing variable, except when the independent variable was inevitably linked to another factor. Between pairs, many variables were changed to explore whether pilots would comment on these variables in the feedback forms; for the same reason, approaches 5, 7, 8, and 11 were added to the experiment as additional approaches. The independent variables and corresponding approach pairs are given in Table 3, and they are explained in the Appendix. Variables that were changed between approach pairs were the required altitude at the FAF (from now on, referred to as FAF altitude), the indicated airspeed (IAS) at the FAF, the distance between the IAF and the IF, the distance between the IF and the FAF, the possibility of being stabilized at 1000 ft, and the energy rate demand. As a final note, it is mentioned that approach 6 was designed according to the guidelines of the optimal RNAV approach, as described in [10].

5. *Dependent Measures*

For every pilot, the RSME ratings for all approaches given on the feedback forms were transformed into z scores. The reason is the following: one pilot might rate all approaches during the experiment high on the RSME scale, whereas another pilot might rate all approaches low on the RSME scale; the absolute values are thus far apart and difficult to compare. However, we are only interested to find out whether both pilots rated approach x lower than approach y, irrespective of the absolute values. Therefore, the RSME scores of one pilot are converted to z scores for this one pilot, using all RSME scores given by this pilot. For each approach pair, the RSME z scores are compared to establish whether there was a significant difference between the RSME z scores for both approaches.

In addition, as explained before, pilots were free to give one rating for the entire approach or to divide the approach up in parts and give multiple RSME ratings: one for each self-assigned part. If pilot A gave one RSME rating for the entire approach (RSME_A), and pilot B decided to divide the same approach into two parts (B1 and B2), resulting in two RSME ratings (RSME_B1 and RSME_B2, one for each part), then RSME_A was assumed to be valid both for part 1 and 2, was counted as two separate (although similar) ratings, and treated as such for the calculation of the z scores per pilot. It is noted that this is an assumption, and it cannot be checked whether pilot A would indeed have given the same RSME rating to both separate parts.

Further subjective data were obtained from the pilot answers to the open-format question about why the approach was difficult or easy to fly and the pilot answers to the questions about whether the approach was stabilized and whether they would adhere to SOPs during real flight. Additionally, subjective data regarding the influence of specified factors on the difficulty of flying an approach were gained from the end-of-day questionnaire. All subjective data are compared, analyzed for inconsistencies and, whenever possible, compared with the objective flight data. These subjective data are used to determine which factors have an effect on pilot TDL.

Two factors that will be shown to influence pilot TDL are whether the approach is stabilized at 1000 ft and whether the constraints at the waypoints are met. To establish whether the approach was stabilized at 1000 ft, the following criteria were used (based on [19]):

1) Heading change and pitch change are within 5 deg/s.
2) The IAS is not more than $V_{\mathrm{REF}} + 20$ kt.
3) Flaps 25 are selected, and landing gear is down.
4) Sink rate is not larger than 1000 ft/min.
5) Localizer and glide slope are within one dot.

An average value is calculated for each of the criteria: for the time slot starting 5 s before reaching 1000 ft and ending at 1000 ft. A larger time slot is used to calculate the average sink rate: it starts 1 min before reaching 1000 ft and ends at 1000 ft.

The constraints at the waypoints are considered to be met when the airspeed is within 10 kt of the required airspeed and the altitude is within 100 ft of the required altitude.

To validate the Monte Carlo simulation (presented in the accompanying paper [2]) with respect to the simulation of the pilot actions, the following pilot actions were logged during the experiment: selection of flaps, autothrottle off, autopilot off, arm approach, heading select, and gear down, as well as the use of the speedbrakes.

C. Hypotheses

Regarding the influence of the independent variables, the following factors were hypothesized to not influence pilot TDL (see also the last column of Table 3 and the explanation in Sec. III): 1) the number of heading changes (independent variable 1), 2) the heading change toward localizer-intercept heading (independent variable 3), and 3) whether the approach is a stepped approach (approach with several altitude steps) or a CDA (independent variable 7).

Factors that were hypothesized to increase pilot TDL are 1) a CDA compared with a horizontal approach (independent variable 2), 2) an energy rate demand that is too high (independent variable 4), 3) a higher localizer-intercept speed (independent variable 5), 4) a lower aircraft mass combined with the speed restrictions for this particular approach pair (pair F) (independent variable 6), 5) a decrease of the distance available on localizer-intercept heading (independent variable 8), and 6) a shorter lineup distance (independent variable 9).

Table 4 Approaches and number of pilots that flew each approach (total $N = 172$)

No. of pilots	Approach No.
1	9
2	9
3	9
4	9
5	6
6	9
7	9
8	9
9	9
10	9
11	5
12	9
13	9
14	9
15	0
16	9
17	9
18	9
19	9
20	9
21	8

V. Results

The experiment results consist of subjective data (collected using feedback forms after each approach and an end-of-day questionnaire) and of objective flight data; the results are presented next.

A. Feedback Forms

The feedback forms were filled in after each approach. The number of pilots that flew each approach is given in Table 4, and a total of 172 feedback forms were available for analysis (approach 15 could not be simulated in the SRS, approach 11 was flown by only five pilots due to time considerations, and approaches 5 and 21 were flown by all pilots but only yielded usable results for seven and eight pilots, respectively). As explained before, each feedback form consisted of three parts (a rating of the approach on the RSME scale, an open format question asking pilots to indicate which factors made the specific approach difficult or easy to fly, and two questions asking pilot opinion on whether or not the approach was stabilized at 1000 ft and whether the pilot would have adhered to SOPs during real flight), and the results for each of these parts are discussed next.

1. Rating Scale Mental Effort Scores

As explained earlier, the pilots were free to give one RSME rating for the entire approach or to divide the approach into multiple parts, giving each part a separate RSME rating. It is interesting to note that, when pilots decided to divide the approach into multiple parts, they almost always split the approach at the IF and/or at the first waypoint on localizer-intercept heading. A distinction can therefore be made between three approach parts (see also Fig. 2): 1) the first part of the approach, which is the segment from the IAF until the first waypoint on localizer-intercept heading; 2) localizer-intercept heading part, which is the segment from the first waypoint on the localizer-intercept heading until the IF (this part includes localizer capture); and 3) the final part of the approach, which is the segment from the IF until 1000 ft.

The results of the average RSME z scores for all approaches flown by nine pilots during the experiment (but excluding approaches 5, 11, 15, and 21, since these were flown by less than nine pilots) are given in Fig. 4. Approach 21 was flown by only eight pilots; therefore, in order to be able to compare approaches 20 and 21 (approach pair F), all RSME z scores were also calculated for only eight pilots and all approaches except approaches 5, 11, and 15 (see Fig. 4). Given the small sample size of average RSME z scores for each approach, it is noted that the following statistical analysis of these scores should be considered merely an indication of possible effects.

The RSME z scores for all approaches (see Fig. 4) were tested for normality using Kolmogorov–Smirnov tests. Only the RSME z score for approach 4, $D(9) = 0.29$, $p < 0.05$, was significantly nonnormal. Levene's test showed that the variances could be considered equal, $F(16, 136) = 1.18$ ns. The RSME z scores for all independent variables were compared using t tests for the normally distributed RSME z scores and Wilcoxon signed-rank tests for the nonnormally distributed RSME z scores. The RSME z scores for approach pair E were analyzed using a one-way repeated-measures analysis of variance (ANOVA). The results are shown in Table 5.

The pilots' average RSME z scores indicate that the following independent variables have a significant effect on pilot TDL: 1) a smaller heading change toward localizer-intercept heading, a surprising result that will be further elaborated upon in Sec. VI.B (independent variable 3); 2) an energy rate demand that is too high (independent variable 4); 3) a smaller distance available on localizer-intercept heading (independent variable 8); and 4) a shorter lineup distance (independent variable 9).

The other independent variables did not cause a significant difference in pilot RSME z scores, and thus in pilot TDL.

Additionally, it can be noted that the RSME scores for approach 6, which was designed according to the optimum design criteria for an RNAV approach in the PANS-OPS, are among the lower ratings on the RSME scale. Approach 6 is thus one of the approaches with lower TDL compared with the other approaches, but it cannot be regarded the approach with the lowest TDL.

2. Factors Influencing Difficulty of Approach

After each approach, the pilots were asked to explain in writing why, in their opinion, the approach was difficult or easy to fly. The

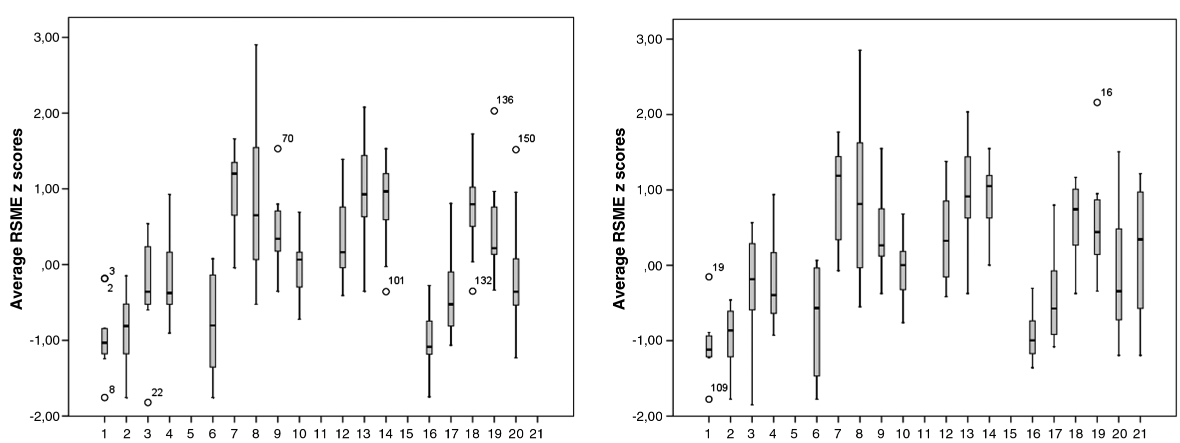


Fig. 4 Pilots' average RSME z scores for all approaches ($N = 9$), except approaches 5, 11, 15, and 21 (left), and for all approaches ($N = 8$), except approaches 5, 11, and 15 (right).

Table 5 Independent variables and results of comparison of RSME z scores

Approach pair	Independent variable	Statistic	Effect size	Significant effect?
A	1) Number of heading changes	$t(8) = -0.851$	$r = 0.29$	No
B	2) CDA compared with horizontal	$T = 15$	$r = 0.10$	No
C	3) Heading change toward localizer-intercept heading	$t(8) = 2.57^{a}$	$r = 0.67$	Yes
D	4) Energy rate demand too high	$t(8) = -5.03^{b}$	$r = 0.87$	Yes
E	5) Localizer-intercept speed (IAS)/localizer groundspeed	$F(2, 16) = 2.06$	$\omega^2 = 0.08$	No
F	6) Mass	$t(7) = -1.14$	$r = 0.4$	No
G	7) More altitude steps compared with CDA	$t(8) = -0.58$	$r = 0.2$	No
H	8) Distance available on localizer-intercept heading	$t(8) = -2.28^{a}$	$r = 0.63$	Yes
I	9) Lineup distance (distance between IF and runway)	$t(8) = 1.89^{a}$	$r = 0.55$	Yes

[a] $p < 0.05$
[b] $p < 0.001$

most frequently mentioned answers for all pilots and all approaches are summarized in Fig. 5. It is important to note that this was an open-format question, and none of the factors in Fig. 5 were preprinted on the feedback form. Pilots could mention more than one factor for each approach; in total, there were 238 comments relating to 24 factors. A factor was included in Fig. 5 when it was mentioned by at least three different pilots; consequently, Fig. 5 shows a total of 208 comments relating to 11 factors.

It is interesting to note that only few comments, 13 out of 208 total (6%) in Fig. 5, are related to the first part of the approach (the factor time available during first part of approach). The factors of wind and energy rate demand apply to the entire approach and comprise 71

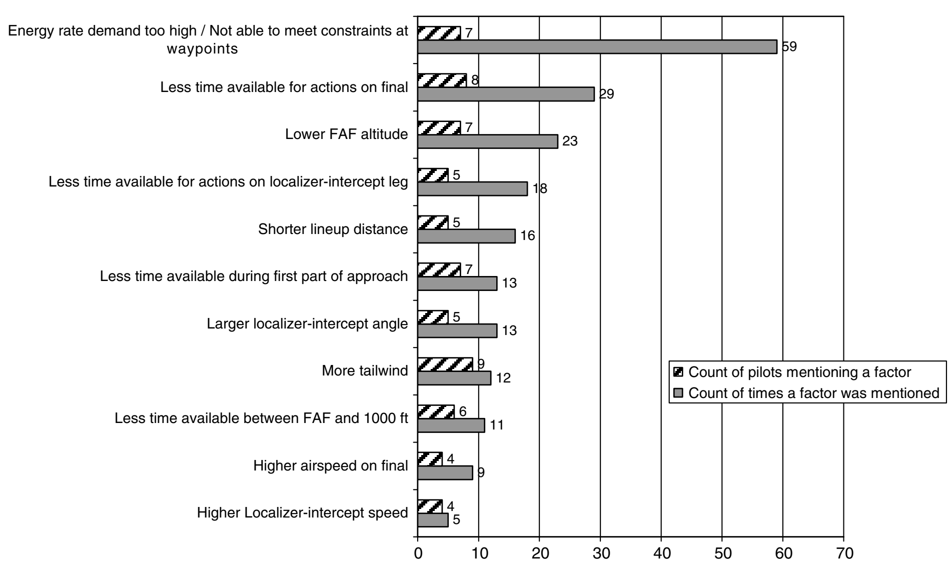


Fig. 5 Factors increasing pilot TDL during approach, as mentioned by at least three pilots on the feedback forms.

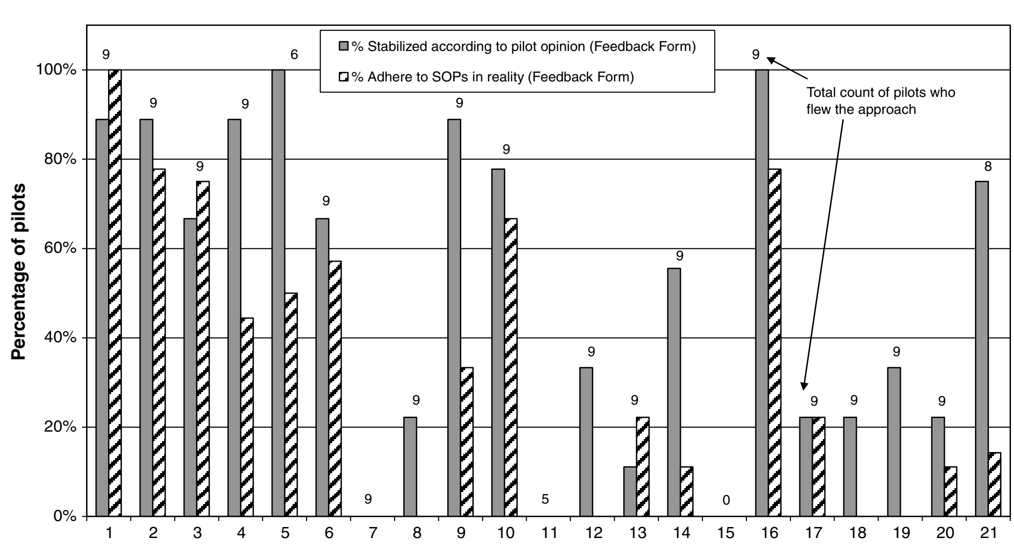


Fig. 6 Percentage of pilots indicating (on feedback forms) that approach was stabilized at 1000 ft and percentage of pilots indicating they would fly the approach according to SOPs in reality.

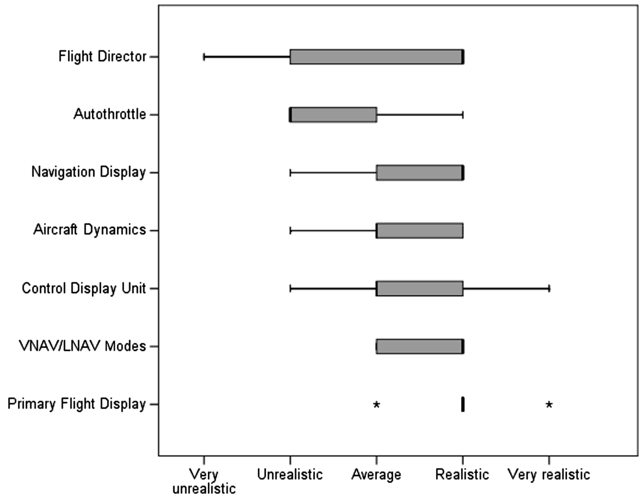


Fig. 7 Pilot answers ($N = 9$) regarding reality of SRS.

comments (34%) of the total amount of 208. The majority of the factors (8 out of 11) and the majority of the comments (124 out of 208, which is 60%) relate uniquely to the last part of the approach and regard the phase from localizer-intercept heading to 1000 ft.

3. Stabilized at 1000 Feet and Standard Operating Procedures

To complete the feedback form, the pilots were asked whether the approach (in their opinion) was stabilized at 1000 ft, with response options of yes or no, and whether they would have flown the approach according to SOPs in reality (again, with response options of yes or no). Figure 6 shows the results. There was a positive relationship between the pilot indicating the approach was stabilized and the same pilot indicating that he would adhere to SOPs in reality when flying that approach; $\tau = 0.45$ (Kendall's tau), $p(\text{one tailed}) < 0.01$.

B. End-of-Day Questionnaire

1. Reality of Simulation, Motion, and Navigation Research Simulator and Reality of Experiment

Pilots were asked their opinion regarding the reality of the SRS by answering closed-format questions with categorical response options for certain aspects of the simulator. For example, for the question, "In your opinion, how realistic was the autothrottle in SIMONA?" the corresponding response options were as follows: very unrealistic, unrealistic, average, realistic, and very realistic. Figure 7 shows that the flight director and the autothrottle were regarded to be especially unrealistic by some of the pilots. When asked, two pilots (out of nine) answered that pilot TDL increased due to the fact that some of the aspects in Fig. 7 were unrealistic.

All pilots indicated that ATC provided them with sufficient information regarding the approach and that the communication between ATC and their flight was average to very realistic. Nevertheless, two pilots (out of nine) reported that the TDL was influenced by the fact that the contact with ATC differed from reality.

The pilots answered that the communication between the PF and the PM was average to realistic and that the way the PM performed his tasks was average to very realistic. However, one pilot indicated that pilot TDL increased due to the fact that the PM was not a real pilot.

All pilots answered that they had sufficient time to prepare for the approach. Of all nine pilots, one pilot reported that the approach and landing charts provided insufficient information, while eight pilots reported that the charts provided sufficient or more than sufficient information. According to two pilots, the TDL during the approach was influenced by these factors.

To conclude this part of the end-of-day questionnaire, the pilots were asked two more questions. The first question asked whether the fact that the cockpit layout of the SRS did not resemble the actual layout of a B747 influenced the difficulty of flying the approaches. One pilot replied, "Yes," and all other pilots replied, "No." The second question was whether the fact that the altitude and speed settings on the MCP did not overrule the VNAV mode had an influence on the difficulty of flying the approaches: three pilots replied, "Yes," and six pilots replied, "No."

It should be noted that, although some of the aspects mentioned previously deviated from reality, and therefore might have had an influence on the difficulty of flying the approach, all these aspects were the same for all pilots and remained constant during all approaches.

2. Factors Influencing Difficulty of Approach

In the second part of the end-of-day questionnaire, the pilots were asked to indicate the factors that influenced pilot TDL while flying an approach by answering closed-format questions. An example of such a question was, "Considering an RNAV approach: when localizer-intercept angle becomes larger, the approach becomes:". Provided response options were as follows: a lot easier, easier, no influence,

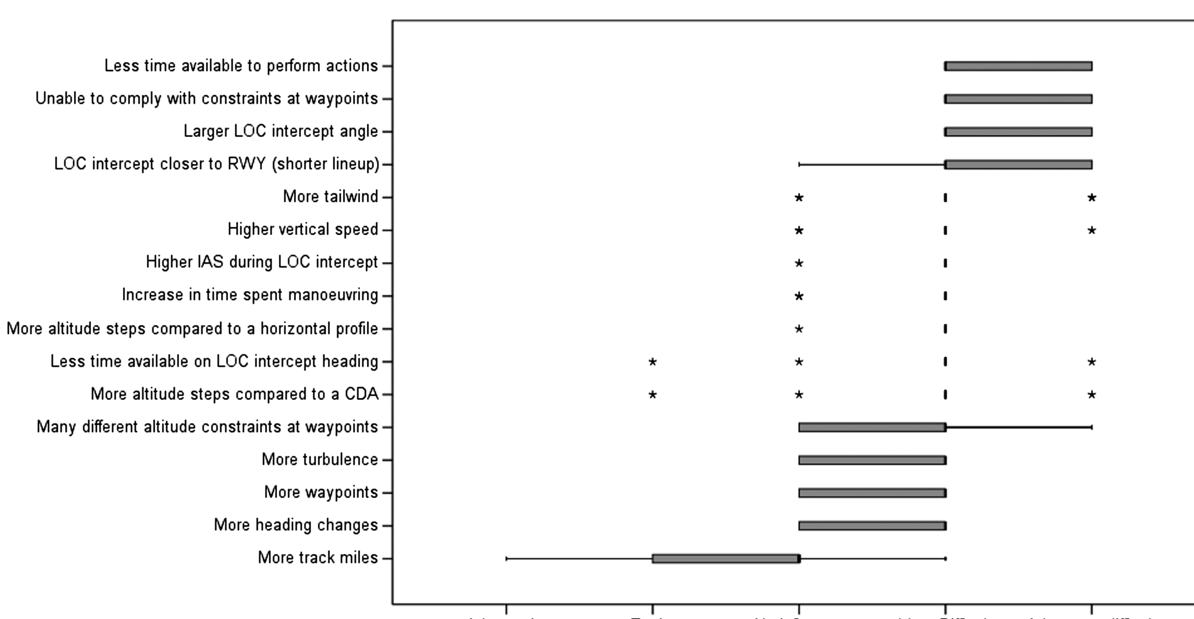


Fig. 8 Pilots' answers ($N = 9$) regarding influence of factors on pilot TDL while flying an approach.

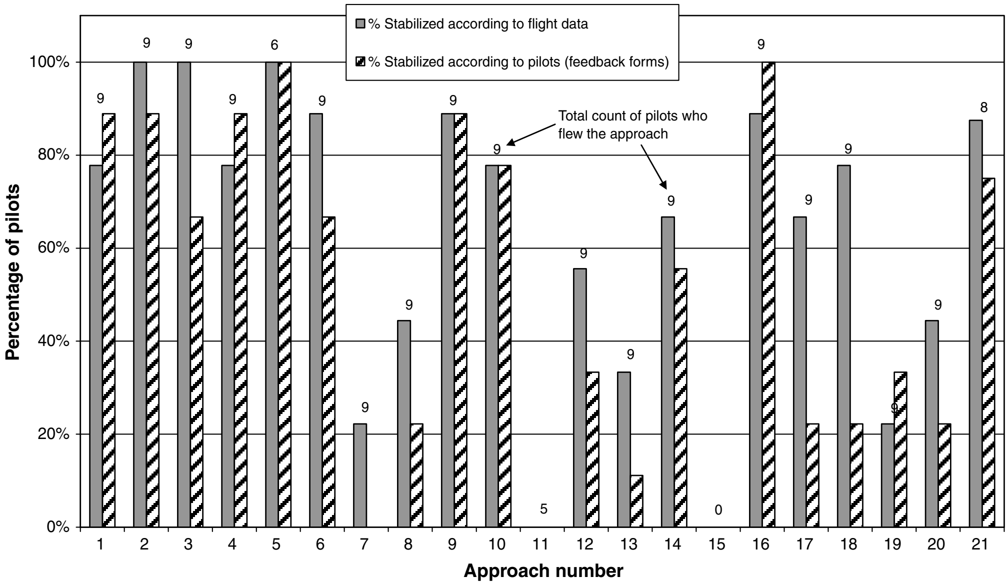


Fig. 9 Percentage of pilots who were stabilized at 1000 ft according to flight data.

more difficult, a lot more difficult. The pilot responses are summarized in Fig. 8.

The final two closed-format questions regarding factors that influenced an RNAV approach were formulated in a different way. The first question was, "If you do not meet the altitude and speed constraints at the waypoints during an RNAV approach, but you do meet all the requirements for a stabilized approach at 1000 ft, would you classify this approach as difficult?". The response options were yes or no. Of all the pilots, seven pilots answered, "Yes," and two replied, "No." The second question, "If you do meet the altitude and speed constraints at the waypoints during an RNAV approach, but you do not meet the requirements for a stabilized approach at 1000 ft, would you classify this approach as difficult?" was answered, "Yes," by eight of the pilots, and one pilot answered, "No."

Following the closed-format questions, the pilots were asked whether there were any other factors that influenced the difficulty of

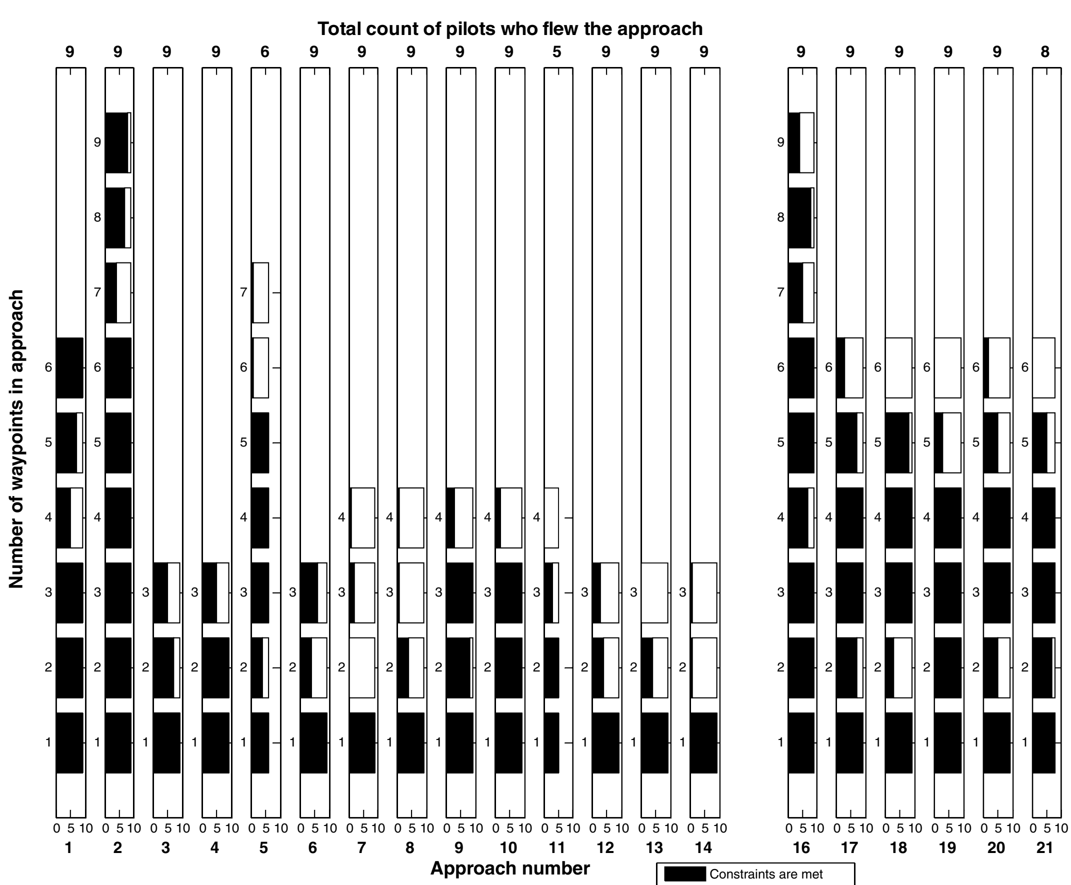


Fig. 10 Count of pilots who met constraints at waypoints and did not meet constraints at waypoints according to flight data for all approaches and all waypoints.

flying an RNAV approach (open format). The answers given were as follows: mountainous terrain, high landing weight, experience on aircraft, FMS programming, and fatigue. All these factors were mentioned only once.

3. *Standard Area Navigation Approach*

In the third part of the end-of-day questionnaire, pilots were asked whether, to their knowledge, a standard or optimal RNAV approach existed. Five out of nine pilots replied, "Yes," and four pilots replied, "No." Further explanations about what this standard or optimal approach would look like mainly focused on the fact that an optimal approach does not have too-strict speed and altitude constraints at waypoints but rather, it has at- or above-altitude constraints or at- or below-speed constraints. None of the pilots mentioned the standard RNAV approach as described in the PANS-OPS [10]. This differs from the findings in the Australian survey. This might be due to the fact that pilots participating in the Australian survey were used to flying approaches according to the optimum design criteria in the PANS-OPS, since these design criteria are applied to many approaches in Australia (only 21.5% of the RNAV approaches published in late 2006 varied from the optimum 5 nm configuration [11]), whereas the pilots participating in our experiment were all Dutch pilots, and they were not used to flying approaches according to the optimum design criteria.

C. Flight Data

During the experiment, flight data were recorded. These data were analyzed to determine how many flights were stabilized at 1000 ft, to determine how often the constraints were met at the waypoints, to calculate the amount of time speedbrakes were used, and to gain insight into all pilot actions. The results are discussed next.

1. *Stabilized at 1000 Feet*

Figure 9 shows, for each approach, the percentage of pilots that were stabilized at 1000 ft according to the flight data. For comparison, the data from Fig. 6 are included as well (representing the percentage of pilots indicating on the feedback forms that, in their opinion, the approach was stabilized). It should be noted that the pilots also took the criterion "all checklists conducted" into account, whereas this was not incorporated in the criteria to analyze the flight data. This causes a difference between the results from the flight data and the pilot answers. [This difference becomes larger when all other criteria for a stabilized approach (such as correct airspeed, correct flap setting, etc.) are met, but the only reason for not being stabilized is the fact that there was not enough time to complete the checklist. This is the case for approaches with a very low FAF (because the final part of the checklist can only be performed after the FAF is reached) and for which the velocity constraints were such that the airspeed at 1000 ft could be lower than VREF + 20. This was the case for, for instance, approaches 17 and 18, which indeed showed a larger deviation between pilot perception and flight data.]

2. *Meet Constraints at Waypoints*

Figure 10 shows how many pilots did, and did not, meet the constraints at each waypoint during the experiment; this is depicted for every approach. For each approach, the first waypoint (waypoint 1) in this figure is the IAF, and the last waypoint (waypoint with the highest number) is the FAF. For example, approach 7 consisted of four waypoints, where waypoint 4 is the FAF. In total, nine pilots flew approach 7. At the first waypoint, all nine pilots met the constraints, which is logical, since this was the starting position of the simulation. At waypoint 2, none of the pilots met the constraints; at waypoint 3, only two pilots met the constraints; and at waypoint 4, only one pilot met the constraints.

3. *Speedbrakes*

Figure 11 shows the average percentage of time that pilots used the speedbrakes during the approach (percentage calculated with respect to the total flight time during the approach). It can be seen that, during some approaches, the pilots used the speedbrakes for a substantial amount of time.

4. *Pilot Actions and Reaction Times*

During the experiment, all pilot actions (selecting flaps, gear, etc.) were logged. These data were analyzed to verify whether the modeling of the pilot actions in the Monte Carlo computer simulation of [2] was realistic. In the Monte Carlo simulation, the pilot actions were modeled using trigger events (for instance, reaching 1200 ft) and reaction times (for instance, 2 s after reaching 1200 ft, flaps 25 are selected). See Table 6 and the accompanying paper [2] for assumed trigger events and reaction times.

Analysis of the flight simulator results showed that the trigger events for the pilot actions in the Monte Carlo simulation and Table 6 were chosen correctly, except those for the selection of flaps 1 and flaps 5. Experiment results showed that it was better to model the selection of flaps 1 and flaps 5 using the IAS instead of a predefined location in the approach; see Figs. 12 and 13. Better trigger events for selecting flaps 1 and flaps 5 were reaching the flap

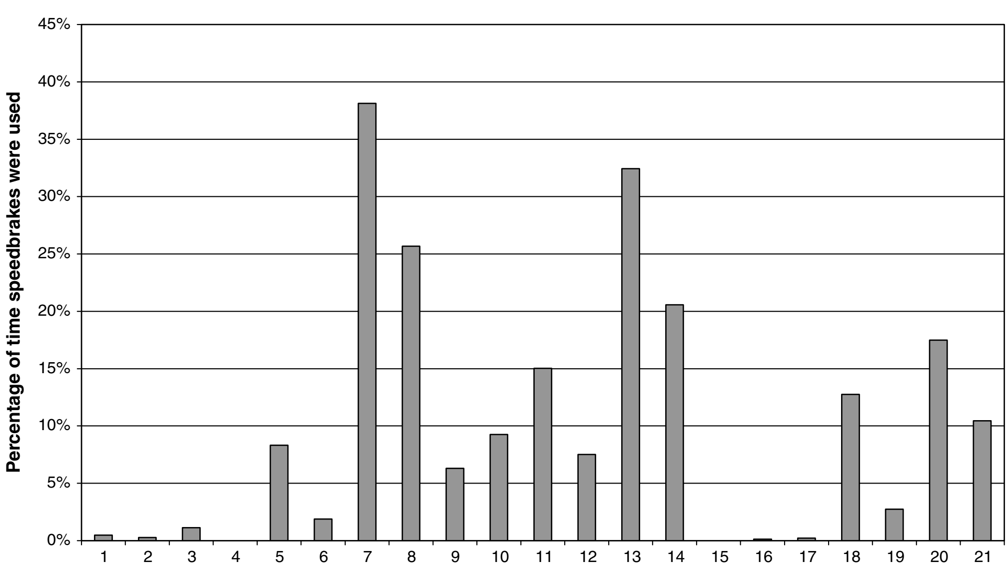


Fig. 11 Average percentage of time (total approach time) that speedbrakes were used.

Table 6 Trigger events and distributions on which reaction times are based for all pilot actions in Monte Carlo simulation

Pilot action	Trigger event	Mean	Standard deviation
Flaps 1	Reaching first waypoint of leg on which aircraft is 20 nm from field	$T/3$[a]	$T/3$[a]
Flaps 5	Reaching first waypoint of leg-before-localizer-intercept heading	$2 \cdot T/3$[b]	$T/3$[b]
Autothrottle/autopilot off	Turn to localizer-intercept heading	$T/2$[c]	$T/2$[c]
Flaps 10	Autothrottle/autopilot off	2 s	0.5 s
Arm approach	Selecting flaps 10	4 s	1 s
Gear down	Glide-slope intercept/reaching FAF	2 s	0.5 s
Flaps 20	Gear down	2 s	0.5 s
Flaps 25	Reaching 1200 ft	2 s	0.5 s

[a]T = amount of time spent on leg on which the aircraft is 20 nm from field.
[b]T = amount of time spent on leg-before-localizer-intercept heading.
[c]T = amount of time spent on localizer-intercept heading.

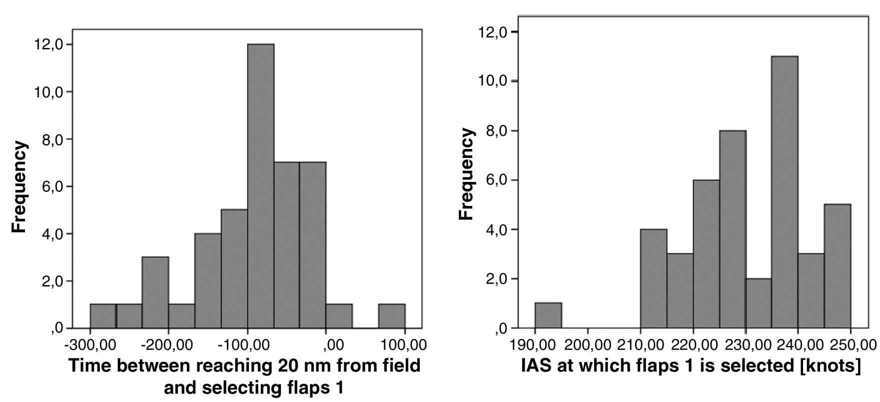
speed mark for flaps up (up mark) at 223 kt (below this speed, flaps 1 should be selected) and reaching the flap speed mark for flaps 1 (1 mark) at 203 kt, respectively (below this speed, flaps 5 should be selected).

Additionally, the trigger events for the actions on localizer-intercept heading (see Table 6) implied that these actions would always be performed in the fixed order given in Table 6. The flight data showed that this is not the case. A more realistic trigger event for each of these actions is the turn-to-localizer-intercept heading, and each of the actions has its own distribution of reaction times with respect to this trigger event; see Fig. 14.

The reaction times for all pilot actions that were assumed in the Monte Carlo simulation were considerably shorter (see Table 6) than the reaction times observed during the experiment; see Figs. 12–16. Additionally, the flight simulator results showed that the pilot sometimes performed the actions before the trigger event was reached; in the Monte Carlo simulation, actions are always modeled to take place after the trigger event.

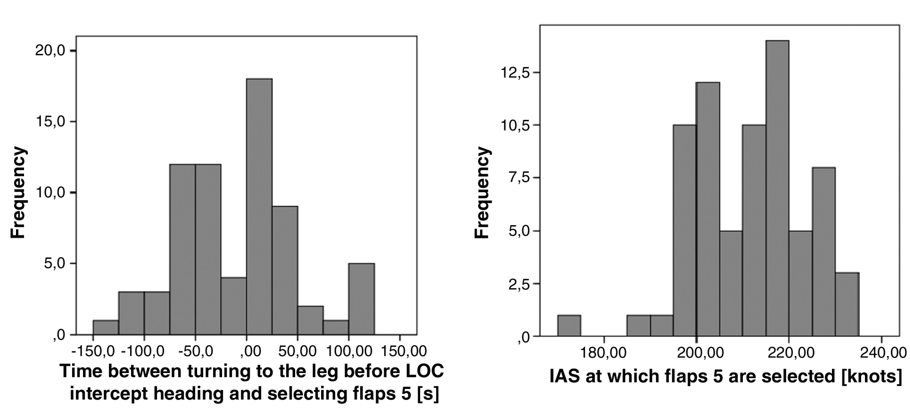
Regarding the actions performed on localizer-intercept heading, it can be stated that there was no significant effect of the time available on localizer-intercept heading on the reaction times for the actions that were performed on localizer-intercept heading. The correlations between the time available on localizer-intercept heading and all pilot actions are $\tau = 0.008$, $p = 0.9$ for the arm approach, $\tau = 0.089$, $p = 0.155$ for flaps 10, $\tau = 0.015$, $p = 0.811$ for autopilot off, $\tau = 0.011$, $p = 0.865$ for autothrottle off, and $\tau = 0.026$, $p = 0.374$ for heading select. The flight data thus showed that, irrespective of the time that was available, pilots tended to start performing these actions the moment they started to turn toward localizer-intercept heading, to get it over with as soon as possible.

Additionally, a very high correlation existed between the reaction times for autothrottle off [significantly nonnormal, $D(127) = 0.099$,



a) b)

Fig. 12 Histograms for selecting flaps 1; approaches with an initial condition equal to flaps 1 or flaps 5 are not incorporated in the histograms.



a) b)

Fig. 13 Histograms for selecting flaps 5; approaches with an initial condition equal to flaps 5 are not incorporated in the histograms.

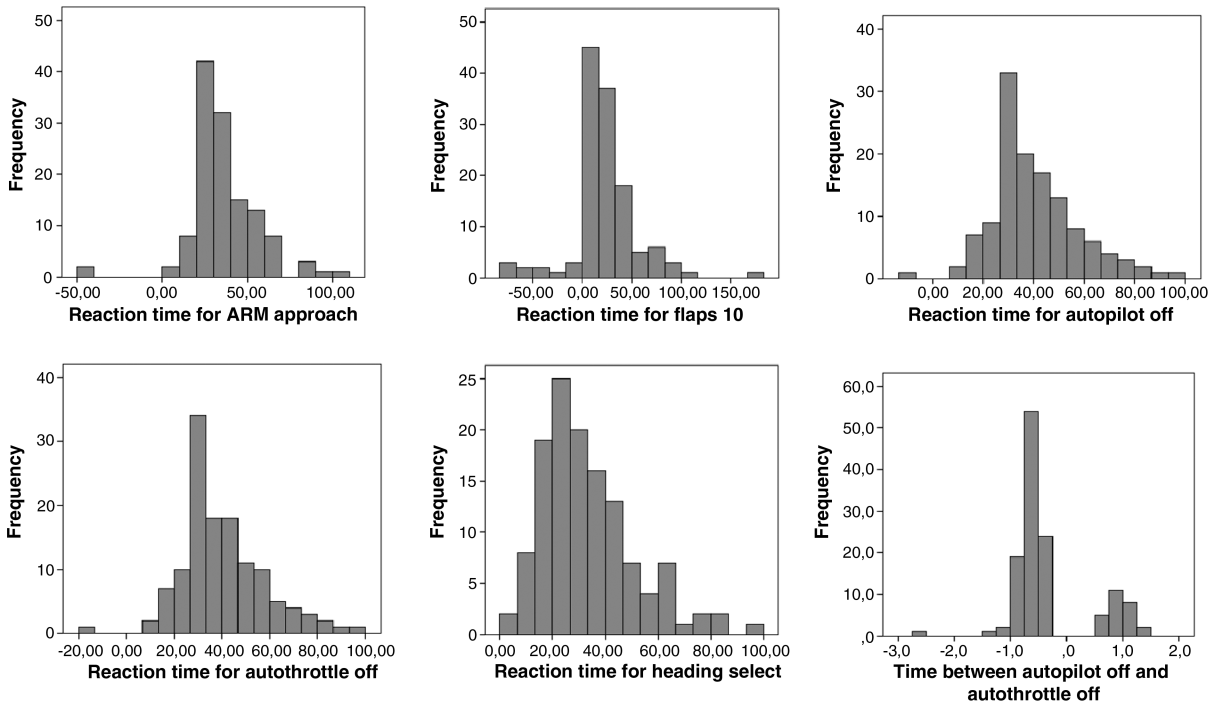


Fig. 14 Histograms for actions on localizer-intercept heading.

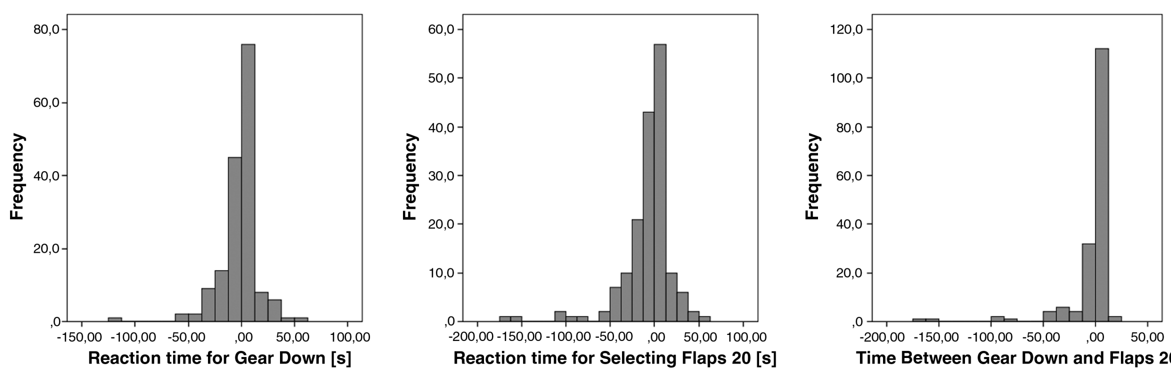


Fig. 15 Histograms for actions at FAF.

$p < 0.01$] and autopilot off [significantly nonnormal, $D(127) = 0.091$, $p < 0.05$]; $r_s = 0.998$, $p < 0.001$. This correlation is due to the fact that both actions are asked for in a single call, autothrottle off and autopilot off; therefore, these two actions were, most of the time, performed very closely together in time (see Fig. 14). The same is true for the actions gear down and flaps 20, which are performed at the FAF; see Fig. 15.

Based on these findings, the trigger events and reaction times in Table 6 are adjusted, as given in Table 7. The normal distributions of the reaction times per pilot action given in Table 7 are based on the histograms shown in Figs. 11–15. One exception is made: for the time between selecting gear down and flaps 20, the mean and standard deviation of the flight data do not represent the flight data at all when these are applied to a normal distribution. Therefore, based

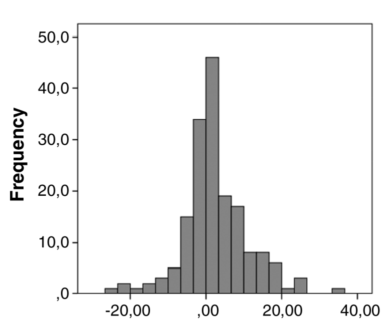


Fig. 16 Histogram for selecting flaps 25 at 1200 ft.

Table 7 Adjusted trigger events and reaction times based on flight data

Pilot action	Trigger event	Mean	Standard deviation
Flaps 1	IAS	230 kt	11.8 kt
Flaps 5	IAS	212 kt	11.8 kt
Flaps 10	Turn to localizer-intercept heading	22.9 s	33.1 s
Heading select	Turn to localizer-intercept heading	33.2 s	17.9 s
Arm approach	Turn to localizer-intercept heading	35.9 s	20.3 s
Autopilot off	Turn to localizer-intercept heading	40.7 s	17.4 s
Autothrottle off	Autopilot off	−0.35 s	0.69 s
Gear down	Reaching FAF	−1.7 s	15.9 s
Flaps 20	Gear down	2 s	2 s
Flaps 25	Reaching 1200 ft	2.5 s	8.6 s

on the recorded reaction times, a different mean and standard deviation are chosen for the approximating normal distribution ($M = 2$ s, $s = 2$ s). The values in Table 7 will be used to validate the simulation of the pilot actions in the Monte Carlo simulation.

VI. Discussion

The experiment results that were derived from the feedback forms, end-of-day questionnaire, and flight data are compared in this section to verify whether these results are consistent.

A. Stabilized at 1000 Feet

There was a significant relationship (see Fig. 9) between the percentage of approaches that were stabilized according to the objective flight data and the percentage of approaches that were stabilized according to the subjective answers of the pilots on the feedback forms; $\tau = 0.58$, $p(\text{one tailed}) < 0.01$ (two times, pilots forgot to complete the relevant question in the feedback form, hence $N = 171$). This implies that the judgment of the pilots about whether the approach was stabilized at 1000 ft correlates well with the objective flight data. An important fact is that, in most cases, pilot judgment is more conservative than the flight data; that is, although the approach was stabilized at 1000 ft according to the flight data, pilots classified the approach as not stabilized. In this respect, it should also be noted again that the criterion checklists conducted for a stabilized approach were taken into account by the pilots, but they were not taken into account when analyzing the flight data.

B. Factors Influencing Pilot Task Demand Load During Approach

To check the consistency between the pilot RSME ratings, the factors mentioned on the feedback forms, and the answers given in the end-of-day questionnaire, all results are summarized in Table 8. For example, the number of heading changes was an independent variable and was also changed between approach pairs. The RSME ratings for the approach pair that considered the number of heading changes yielded no significant difference, and the pilots did not mention the number of heading changes on any of the feedback forms as a factor that contributed to the difficulty of flying an approach. However, in the end-of-day questionnaire, the pilots answered that an approach becomes more difficult when the number of heading changes increases.

The second column thus indicates whether the factor has been changed at all during the experiment (between approach pairs). This is important, for example, when considering the factor turbulence. The turbulence intensity was not varied during the experiment and was kept at a low setting. Therefore, the pilots might not have been triggered to write down any comments regarding turbulence on the feedback forms. The fact that it is not mentioned on the feedback forms does not, however, imply that it does not have an influence on the difficulty.

The factor distance available on localizer-intercept heading and the time available for actions on localizer-intercept heading are directly linked. They are mentioned separately to indicate that pilots always commented on time instead of on distance in the feedback forms. The factors lineup distance and time available for actions on the final are also directly linked; for these factors, pilots did comment on both factors, and the same holds for the factors of FAF altitude and time between FAF and 1000 ft.

Looking at Table 8, there are two factors that require further explanation. The first factor is heading change toward localizer-intercept heading. It was hypothesized that this factor would not influence the difficulty of an approach. The RSME scores, however, indicated that pilot TDL increases when the heading change toward localizer-intercept heading decreases.

On the other hand, none of the pilots mentioned this factor on the feedback forms. The authors were somewhat puzzled by this outcome, especially since, if there would be an effect, it would be expected that pilot TDL would increase when the heading change increases as well. Therefore, a couple of months after the experiment, four pilots were asked to fly these two approaches again, and they were asked to explain which approach they found easier or more difficult to fly. All four pilots commented that there was no difference between the two approaches. This clearly illustrates the limited reliability of the RSME scores and, as stated before, the RSME scores should be considered only an indication of possible effects.

The second factor that needs further explanation is localizer-intercept speed/groundspeed, which was linked to a higher energy rate demand between IF and FAF. Analysis of the RSME scores showed no difference in average RSME scores with increasing localizer-intercept speed and increasing energy rate demand between IF and FAF. This seems strange when regarding the findings for the factor energy rate demand too high (linked to a higher localizer groundspeed), which did show a significantly higher RSME score with increasing energy rate demand and higher localizer-intercept groundspeed. Also, when looking at the pilot comments on the feedback forms and the results of the end-of-day questionnaires, one would expect that localizer-intercept speed would have an influence on the RSME scores. This will need to be tested again in the next experiment.

Using Table 8, it is now possible to discuss which factors have an influence on pilot TDL. Ideally, factors that influence the difficulty of flying an approach would yield a significant difference in RSME scores, would be mentioned on the feedback forms, and would be rated more difficult or a lot more difficult in the end-of-day questionnaire. The only factors that fulfill this description are factor 4 (energy rate demand too high), factor 8 [distance (or time) available on localizer-intercept heading], and factor 9 (lineup distance).

On the other hand, factors that do not influence pilot TDL would not yield a significant difference in RSME scores, would not be mentioned on the feedback forms, and would not be rated more difficult or a lot more difficult in the end-of-day questionnaire. This applies to none of the factors in Table 8. However, factors 1 (number of heading changes), 2 (CDA compared with horizontal), and 7 (more altitude steps compared with CDA) were tested three times (by t tests or Wilcoxon signed-rank tests for the RSME z scores, by analyzing whether they were mentioned on the feedback forms and by asking about these factors in the end-of-day questionnaire). Only in the end-of-day questionnaire did the pilots indicate that these factors had an effect on TDL; the other two tests did not show an effect on TDL. Therefore, it is assumed that factors 1, 2, and 7 do not have an effect on pilot TDL.

Factor 3 (heading change toward localizer-intercept heading) was tested again, as mentioned before, by flying two approaches in the flight simulator, and pilots commented that there was no difference in TDL due to a smaller heading change toward localizer-intercept heading. For that reason, factor 3 is assumed not to influence pilot TDL.

Factor 12 (tailwind) is, by definition, directly correlated with the value of the energy rate demand (energy rate demand, factor 4), and therefore not regarded as an additional factor that influences pilot TDL. Factor 20 (increase in time spent maneuvering) is already covered by factors 1 (number of heading changes), 7 (more altitude steps compared with CDA), and 17 (more altitude steps compared with horizontal), and it is not regarded a separate factor.

Given the preceding, some factors indicated in Table 8 can be removed from the list of factors that influence pilot TDL. The remaining factors in Table 8 are indicated by [d] in the first column.

VII. Regression Analysis for Pilot Task Demand Load

A regression analysis was performed on the pilots' average RSME z scores for the approaches. The goal of the regression analysis was to identify the factors that were dominant in affecting the pilot RSME ratings, and thereby pilot estimates of TDL during an approach. (The RSME ratings are an indication of the mental load experienced by the pilots. However, by choosing pilots with different levels of experience, different ages, and different levels of fatigue, and by testing the approaches in random order, it is assumed that through the RSME ratings of the pilots, a good indication of pilot TDL can be obtained.) The predictors used for the regression analysis are based on the

Table 8 Factors that might influence pilot TDL, with corresponding results for RSME ratings, comments on feedback forms, and answers given in end-of-day questionnaire[a]

Factor no.	Factors	Independent variable	Variable between pairs	RSME scores	Number of times factor was mentioned on feedback forms	Factor indicated to have an effect in the end-of-day questionnaire?	Hypothesis for TDL effect
1	Number of heading changes	Yes	Yes	NS[b]	Not mentioned	Yes	No effect
2	CDA compared with horizontal/many different altitude constraints	Yes	Yes	NS[c]	Not mentioned	Yes	Effect expected
3	Heading change toward localizer-intercept heading	Yes	Yes	$p < 0.05$[b]	Not mentioned	Not asked	No effect
4[d]	Energy rate demand too high/unable to comply with constraints at waypoints (linked factor: localizer groundspeed)	Yes	Yes	$p < 0.01$[b]	>10 times	Yes	Effect expected
5[d]	Localizer-intercept speed (IAS)/Localizer groundspeed (linked factor: energy rate demand IF-FAF)	Yes	Yes	NS[e]	<10 times	Yes	Effect expected
6[d]	Mass	Yes	No	NS[b]	Not mentioned	Not asked	Effect expected
7	More altitude steps compared with CDA	Yes	Yes	NS[b]	Not mentioned	Yes	No effect
8[d]	Distance available on localizer intercept . heading (linked factor: localizer groundspeed)	Yes	Yes	$p < 0.05$[b]	Not mentioned	Not asked	Effect expected
	Time available for actions on localizer-intercept leg (linked factor: localizer groundspeed)				>10 times	Yes	
9[d]	Lineup distance (linked factor: IF-FAF distance)	Yes	Yes	$p < 0.05$[b]	>10 times	Yes	Effect expected
	Time available for actions on final (linked factor: IF-FAF distance)				>10 times	Yes	
10[d]	FAF altitude	No	Yes	Not tested	>10 times	Not asked	
	Time available between FAF and 1000 ft				>10 times		
11[d]	Localizer intercept angle	No	Yes	Not tested	>10 times	Yes	
12	Tailwind	No	Yes	Not tested	>10 times	Yes	
13[d]	Vertical speed	No	Yes	Not tested	Not mentioned	Yes	
14[d]	Number of waypoints	No	Yes	Not tested	Not mentioned	Yes	
15[d]	Turbulence	No	No	Not tested	Not mentioned	Yes	
16[d]	Track miles	No	Yes	Not tested	Not mentioned	Yes	
17[d]	More altitude steps compared with horizontal	No	No	Not tested	Not mentioned	Yes	
18[d]	Airspeed on final	No	Yes	Not tested	>10 times	Not asked	
19[d]	Time available during first part of approach	No	Yes	Not tested	>10 times	Not asked	
20	Increase in time spent maneuvering	No	Yes	Not tested	Not mentioned	Yes	
21[d]	Stabilized at 1000 ft	No	Yes	Not tested	Not mentioned	Yes	

[a] NS denotes no significance.
[b] t test.
[c] Wilcoxon signed-rank test.
[d] Predictors used for regression analysis.
[e] ANOVA.

factors that are indicated by [d] in Table 8. For different reasons, not all these factors are incorporated as predictors.

Factors 15 (turbulence) and 17 (more altitude steps compared with horizontal) were not varied among the different approaches and can, therefore, for this experiment, not be used as predictors.

Factor 13 (vertical speed) is analytically defined by the IAS, the wind conditions, and the inertial flight-path angle. When pilots commented on the vertical speed, they always referred to the vertical speed on final. On final, the inertial flight-path angle is equal to 3 deg for all approaches, implying that the vertical speed on final only depends on the IAS on final and the wind conditions on final. Compared with the influence of the IAS on the vertical speed, the influence of the wind is very small for the approaches considered. Therefore, the vertical speed can also be represented by the IAS on final, which already is a factor (factor 18). Therefore, for this experiment, vertical speed is not regarded as a separate factor for the approaches considered.

Factor 19 (time available during first part of approach) could only be calculated for a part of the approaches, since some approaches did not contain a first part of the approach. It is therefore decided to not use factor 19 as a predictor in the regression analysis for this experiment.

Factor 18 (airspeed on final) is calculated as the average of localizer-intercept speed (which is factor 5) and IAS at the FAF. Since factor 5 is already a predictor, the choice is made to include the IAS at the FAF as an additional predictor instead of the average airspeed on final.

The remaining 14 factors are used as predictors for the regression analysis; these factors and their factor numbers are given in the first column of Table 9. Because of the small amount of different

Table 9 Factors used as predictors in the regression analysis[a]

		Energy rate demand localizer -intercept heading	Energy rate demand final	Localizer-intercept speed (IAS)	Aircraft mass	Distance available on localizer-intercept heading	Lineup distance	FAF altitude	Localizer-intercept angle	No. of waypoints	Track miles	IAS at FAF	Stabilized at 1000 ft (Monte Carlo original simulation)
4	Energy rate demand localizer-interceptor heading	——	0.78	0.91	——	——	——	——	——	——	——	0.84	—— 0.80
4	Energy rate demand final	0.78	——	0.92	——	——	——	——	——	——	——	0.87	—— 0.71
5	Localizer-intercept speed (IAS)	0.91	0.92	——	——	——	——	——	——	——	——	0.93	—— 0.80
6	Aircraft mass	——	——	——	——	——	——	——	——	——	——	——	——
8	Distance available on localizer-intercept heading	——	——	——	——	——	——	——	——	——	——	——	——
9	Lineup distance	——	——	——	——	——	——	0.91	0.80	——	——	——	0.71
10	FAF altitude	——	——	——	——	——	0.91	——	0.92	——	——	——	——
11	Localizer-intercept angle	——	——	——	——	——	0.80	0.92	——	——	——	——	——
14	No. of waypoints	——	——	——	——	——	——	——	——	——	0.92	——	——
16	Track miles	——	——	——	——	——	——	——	——	0.92	——	——	——
18	IAS at FAF	0.84	0.87	0.93	——	——	——	——	——	——	——	——	—— 0.80
21	Stabilized at 1000 ft (Monte Carlo original simulation)	—— 0.80	—— 0.71	—— 0.80	——	——	0.71	——	——	——	——	—— 0.80	——

[a]Correlation coefficients larger than 0.7 are denoted.

approaches that were tested, some of these predictors are highly correlated, which will most probably pose a problem for the regression analysis; correlations higher than 0.7 are given in Table 9. It should be noted that, for localizer-intercept speed (factor 5), the IAS during the localizer intercept predicted by the original Monte Carlo simulation is used, and for the airspeed at FAF (factor 22), the IAS at the FAF predicted by the original Monte Carlo simulation is used.

A stepwise regression was performed on these predictors (probability of F to enter is ≤ 0.05, and probability of F to remove is ≥ 0.01) for the dependent measure average RSME z score; the resulting model is shown in Table 10. The average variance inflation factor (VIF) for these three predictors is 1.44, which indicates that for these predictors, collinearity is not an issue. However, considering the large amount of high correlation coefficients between these three predictors and other factors not included in the regression model in Table 10, other possible regression models are considered by performing a best subsets regression. The best subsets regression revealed, as expected, that several different combinations of the list of predictors in Table 9 provide comparable R-squared values to the regression model resulting from the stepwise regression in Table 10.

The authors infer that the following factors should be used as predictors in a regression model for the average RSME z score:

1) Factor 4, which is the factor energy rate demand too high, should be used as a predictor. This factor is calculated separately for the localizer-intercept part and the final part of the approach.

2) Factor 5, which is the localizer-intercept speed, should be used as a predictor. For this factor, the IAS during localizer intercept predicted by the original Monte Carlo simulation is used.

3) Factor 8, which is the distance available on localizer-intercept heading, should be used as a predictor.

4) Factor 9, which is the lineup distance, should be used as a predictor.

5) Factor 10, which is the FAF altitude, should be used as a predictor.

6) Factor 11, which is the localizer-intercept angle, should be used as a predictor.

7) Factor 21, which represents whether the aircraft is stabilized at 1000 ft, should be used as a predictor. For this factor, the prediction of the original Monte Carlo simulation is used.

This means that of all factors in Table 9, factor 14 (number of waypoints), and factor 16 (total track miles) are no longer considered predictors. This choice is made based on observations during the experiment: none of the pilots mentioned the number of waypoints or total track miles, either in written form or orally, although these factors were varied significantly among approaches. Factor 6 (aircraft mass) is also left out, since from a flight mechanical perspective, this only has an influence through a change in energy rate demand, which is already covered by factor 4. Finally, factor 22 (airspeed at FAF) is no longer considered as a predictor; we think that the speed at the FAF in itself is not a factor for pilot TDL, but that the possibility to reduce from the speed at the FAF to a speed less than $V_{\mathrm{REF}} + 20$ at 1000 ft is a factor for pilot TDL, and this is again covered by the energy rate demand for the final part of the approach (factor 4).

The resulting regression model for the remaining predictors given in the preceding list is shown in Table 11. The largest VIF factor for these predictors equals 25, and thus is cause for concern. Cook's distance is smaller than one for all cases, and the centered leverage value is smaller than twice the average leverage for all cases.

The way forward is to use the predictors of the regression model in Table 11 as a starting point for the design of the approaches that will be tested during the second flight simulator experiment (earlier indicated as step 3 in this research). Special care should be taken to ensure a low correlation between these factors in the next experiment. Additionally, pilots should be asked specifically about these factors in the feedback forms and end-of-day questionnaire.

VIII. Comparison of Experiment Results and Monte Carlo Simulation

Some of the predictors for pilot TDL presented in the previous section are directly evident when looking at the approach charts (for instance, the FAF altitude or localizer-intercept angle); the flight mechanical factors (these are the energy rate demand and the whether

Table 10 The stepwise linear regression model for the dependent variable RSME z score

	B	Standard error B	β	R^2	Adjusted R^2
Step 1					
Constant	2.06	0.234			
Lineup distance, nm	−0.311	0.035	−0.597[a]	0.356	0.351
Step 2					
Constant	−2.16	0.894			
Lineup distance, nm	−0.226	0.037	−0.434[a]		
Localizer-intercept speed, kt	0.019	0.004	−0.345[a]	0.449	0.441
Step 3					
Constant	−1.61	0.899			
Lineup distance, nm	−0.185	0.039	−0.355[a]		
Localizer-intercept speed, kt	0.017	0.004	−0.301[a]		
Total track miles, nm	−0.018	0.007	0.194[a]	0.475	0.464

[a] $p < 0.001$.

Table 11 Linear regression model for the dependent variable RSME z score chosen as best model from a best subsets regression

	B	SE B	β	p	R^2	Adjusted R^2
Constant	−5.07	2.489		0.044		
Energy rate demand localizer-intercept part	−0.091	0.138	−0.109	0.512		
Energy rate demand final part	−0.076	0.047	−0.374	0.112		
Localizer-intercept speed, kt	0.045	0.015	0.800	0.003		
Distance available on localizer-intercept heading, nm	−0.128	0.068	−0.156	0.060		
Lineup distance, nm	−0.108	0.094	−0.208	0.251		
FAF altitude, ft	−0.003	0.001	−0.822	0.008		
Localizer-intercept angle, deg	0.053	0.022	0.548	0.018		
Stabilized at 1000 ft, %	0.004	0.003	0.198	0.218	0.497	0.467

the approach is stabilized at 1000 ft), however, are not directly available and need to be calculated. The Monte Carlo computer simulation aims to do exactly that. As stated in the experiment goals, this experiment aimed to validate the Monte Carlo simulation with respect to the simulation of the pilot's actions, with respect to the prediction of whether or not the constraints at the waypoints are met, and with respect to the prediction of whether or not the approach is stabilized at 1000 ft. For each of these three items, the flight simulator results are compared with the Monte Carlo simulation in the following sections.

A. Simulation Pilot Actions

The simulation of the pilot actions in the Monte Carlo simulation is adjusted according to the findings presented in Sec. V.C and based on the trigger events and reaction times in Table 7. As a first approximation, normal distributions are used based on the means and standard deviations given in Table 7.

Additionally, the use of speedbrakes is incorporated in the Monte Carlo simulation. In the original Monte Carlo simulation [2], speedbrakes were assumed not to be used; the experiment results, however, showed that pilots did use the speedbrakes (see Fig. 11). Speedbrakes are modeled to be selected when a drag-required message would appear on the pilot displays, indicating that the constraints at the next waypoint would not be met. There is no time delay in applying speedbrakes. Speedbrakes are selected the moment the drag-required message appears, and they are selected consistently every time the message appears. Because of simulation technical reasons, the speedbrakes could only be used in the Monte Carlo simulation until the localizer was captured. In reality, pilots could use the speedbrakes for a slightly larger part of the approach until they reached the FAF, since they are instructed not to use the speedbrakes for flap settings greater than flaps 10 (flaps 20 is selected at the FAF).

The results of this adjusted Monte Carlo simulation are compared with the flight data in the following sections.

B. Prediction Stabilized

Figure 17 shows, for each approach, the percentage of pilots who were stabilized according to the flight data, the percentage of pilots who would be stabilized according to the adjusted Monte Carlo simulation when speedbrakes were applied, and the percentage of pilots who would be stabilized according to the adjusted Monte Carlo simulation when speedbrakes were not applied. It can be seen that whether or not speedbrakes were applied (in the Monte Carlo simulation) can have a large influence on the prediction for the percentage of stabilized approaches (for instance, approach 7: 56% is stabilized when using speedbrakes, and 0% is stabilized without speedbrakes).

There is a significant correlation between the percentage of pilots who were stabilized according to the flight data and according to the adjusted Monte Carlo prediction, $\tau = 0.53$, $p(\text{one tailed}) < 0.001$, for the adjusted Monte Carlo prediction when speedbrakes are used, and $\tau = 0.64$, $p(\text{one tailed}) < 0.001$ for the adjusted Monte Carlo prediction when speedbrakes are not used.

The discrepancies between the flight data and the adjusted Monte Carlo simulation can be due to the fact that, during the experiment, the pilots have used the speedbrakes for a longer amount of time (e.g., also after capturing the localizer) than the amount of time modeled in the adjusted Monte Carlo simulation, or they have used the speedbrakes for a shorter amount of time. During the experiment, pilots sometimes forgot to retract the speedbrakes due to the absence of buffet cues (while speedbrakes were deployed) in the flight simulator. Another reason for discrepancies could be the fact that, as explained before, the throttle operation differed from reality. Pilots, therefore, sometimes found it more difficult to regulate the airspeed than they would normally have found it during real flight. This sometimes resulted in an airspeed that was too low or too high. This was not incorporated in the Monte Carlo simulation. Additionally, during the experiment, pilots sometimes maintained an IAS at the FAF that was lower than defined on the approach charts; as a result, they could sometimes achieve a stabilized approach at 1000 ft, while the Monte Carlo simulation (maintaining the correct higher IAS at the FAF) predicted an unstabilized approach.

Figure 17 illustrates the large influence that the use of speedbrakes can have on the possibility of achieving a stabilized approach, and hence on the correlation with the predictions of the Monte Carlo simulation. It is therefore better to not allow the use of speedbrakes during the next experiment and to compare these results to the Monte Carlo simulation without the use of speedbrakes. This is also a better choice regarding the intended use of the method developed during this research: to provide an additional tool for the design of approaches. Approaches should be designed such that they can be flown without speedbrakes.

C. Prediction Meet Constraints

The percentage of pilots that met the constraints according to the flight data (significantly nonnormal, $D(76) = 0.23$, $p < 0.05$) was significantly related to the percentage of pilots that met the constraints according to the Monte Carlo simulation. This is true for the adjusted Monte Carlo simulation during which speedbrakes were

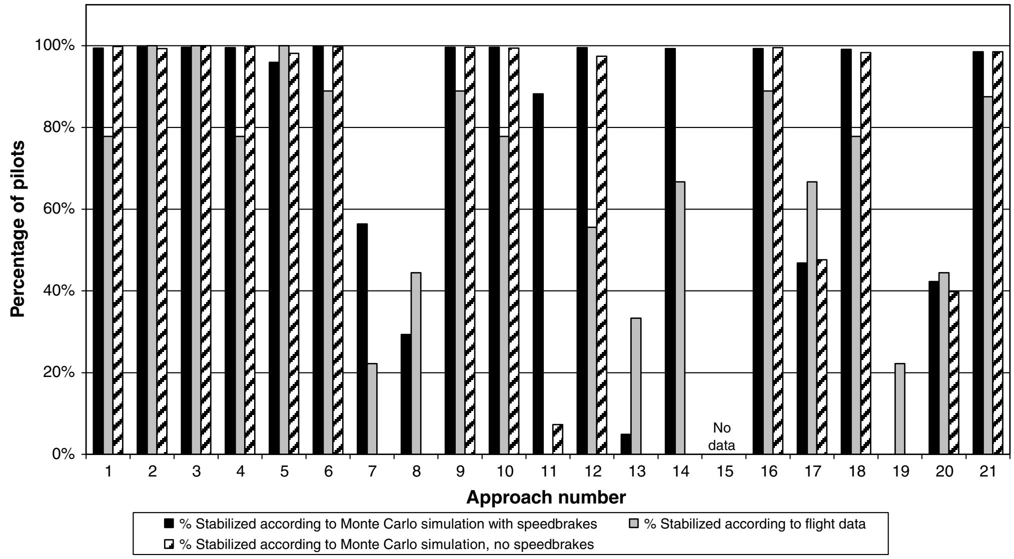


Fig. 17 Percentage of pilots stabilized according to flight data and stabilized according to the adjusted Monte Carlo simulation (with and without the use of speedbrakes).

used, $\tau(\text{one tailed}) = 0.407$, $p < 0.001$, and for the adjusted Monte Carlo simulation during which speedbrakes were not used, $\tau(\text{one tailed}) = 0.500$, $p < 0.001$. However, in order to be able to use the Monte Carlo simulation to accurately predict whether the constraints at the waypoints will be met, a larger correlation coefficient is desirable and might be achieved by not allowing the use of speedbrakes during the next flight simulator experiment. Since the Monte Carlo simulation will ultimately be used during the design of approaches, and approaches should be designed such that they can be flown without the use of speedbrakes, not allowing the use of speedbrakes corresponds better to the goal of this research.

IX. Conclusions

In this paper, a relation has been found between pilot TDL during approaches (measured in terms of the RSME rating scale) on the one hand and flight mechanical factors and factors relating to the approach trajectory on the other hand. Based on the results of flight simulator tests for a B747-100, it has been inferred that the following factors influence pilot TDL: the energy rate demand during the last parts of the approach, localizer-intercept speed, the distance available on localizer-intercept heading, the lineup distance, the FAF altitude, localizer-intercept angle, and whether the approach is stabilized at 1000 ft.

Some of these factors are directly evident when looking at the approach charts (for instance, the FAF altitude or localizer-intercept angle); the flight mechanical factors (these are the energy rate demand and whether the approach is stabilized at 1000 ft), however, are not directly available and need to be calculated. This paper also presented the validation of a Monte Carlo computer simulation that aims to do exactly that (the Monte Carlo computer simulation itself was presented in the accompanying paper [2]). By modeling the pilot actions according to the SOPs and applying a distribution in time for all these actions, the Monte Carlo computer simulation predicts the percentage of pilots who will meet the altitude and velocity constraints at the waypoints and the percentage of pilots who will achieve a stabilized approach at 1000 ft. A significant correlation was found between these predictions and the flight data from the flight simulator tests. However, a larger correlation coefficient than achieved in this paper is desirable and might be realized by not allowing the use of speedbrakes during the next experiment.

The factors that were found to influence pilot TDL in this paper are regarded as an initial indication, and further research based on these findings is required. The predictors of the regression model will be used as a starting point for the design of the approaches that will be tested during the second flight simulator experiment (indicated as step 3 in this research). Special care should be taken to ensure a low correlation between these factors in the next experiment. Additionally, pilots should be asked specifically about these factors in the feedback forms and the end-of-day questionnaire.

The next flight simulator experiment will also be used to further validate the Monte Carlo simulation. Pilots should not be allowed to use the speedbrakes in the next experiment in order to cancel out this additional variable and to obtain solid validation data for the Monte Carlo simulation. This is also a better choice regarding the intended use of the method developed during this research: to provide an additional tool for the design of approaches. Approaches should be designed such that they can be flown without the use of speedbrakes.

Appendix: Design of Approaches

The table in Fig. A1 shows all relevant data for the approaches that were tested. Most variables in Fig. A1 are self-explanatory, except the energy demand ratio. The maximum value of the energy demand ratio is calculated for each approach part separately. The values in Fig. A1 are determined using the Monte Carlo simulation: by calculating the average value of the energy demand ratio during a 10 s time interval before a waypoint is reached. The value in Fig. A1 thus differs from the value that is predicted by the PMM, as explained in the accompanying paper [2], which gives an energy demand ratio that is an average value for the entire leg.

If the value for the energy rate demand is larger than one, this indicates that the constraints at the next waypoint will not be met; the larger the value for the energy rate demand, the more the actual IAS will exceed the required IAS at the waypoint. It is important to note that pilots received a drag-required message when the energy rate demand became higher than one during the experiment. No message would show up for energy rate demand values smaller than one, and pilots would thus not be aware of whether the energy rate demand equaled, for instance, 0.1, 0.6, or 0.9, since, in all these cases, the constraints would be met. All approach pairs are now briefly described next:

Approach pair A considers the independent variable "number of heading changes". The difference between approaches 1 and 2 was the amount of heading changes within the same amount of track miles in the first part of the approach. A difference in energy rate demand can be observed for the first part of the approach (Fig. A1); however, the energy rate demands do no exceed one.

Approach pair B considers the independent variable "CDA compared with horizontal". Approaches 3 and 4 both start on localizer-intercept heading. The final part of the approach is identical for approaches 3 and 4; localizer-intercept heading part is a horizontal segment for approach 4 and a CDA segment for approach 3. However, due to keeping the IAS constraints on localizer-intercept heading the same, the energy rate demand will increase for the CDA segment (approach 3) compared with the horizontal segment (approach 4). However, the energy rate demand does not exceed one for either of the approaches.

Approach pair C considers the independent variable "heading change toward localizer-intercept heading". Approaches 9 and 10 are identical, except for the heading change that is required when turning from the leg-before-localizer-intercept heading to localizer-intercept heading. A very small difference in energy rate demands can be observed, but all are lower than one.

Approach pair D considers the independent variable "energy rate demand is too high". Approaches 17 and 18 are exactly identical. The difference in energy rate demand is established by applying a wind of 40 kt in approach 18, which resulted in a strong tailwind during the first part of the approach and localizer-intercept heading part. As a result, the energy rate demand during approach 18 was larger than one. In approach 17, the wind was only 10 kt and all constraints could be met (energy rate demand smaller than one). Since the IAS during localizer intercept was identical for both approaches, the groundspeed during localizer intercept differed for both approaches due to the wind.

Approach pair E considers the independent variable "localizer-intercept speed". The IAS during localizer intercept is different for approaches 12, 13, and 14. This difference in IAS is amplified by adding wind to the experiment, resulting in an even larger difference in groundspeed during localizer intercept. Since the remainder of the approach is exactly identical for approaches 12, 13, and 14, a higher localizer-intercept speed (IAS) inevitably means that the energy rate demand will increase between IF and FAF.

Approach pair F considers the independent variable "aircraft mass". Approaches 20 and 21 are exactly identical, but they are flown with a different aircraft mass. According to the predictions of the PMM and the Monte Carlo simulation, it would be easier meet the constraints at the waypoints with the higher aircraft mass than with the lower aircraft mass (represented by different energy rate demands in Fig. A1).

Approach pair G considers the independent variable "CDA compared with altitude steps". The first part of the approach for approach 2 was a CDA, whereas the first part of the approach for approach 16 consisted of three altitude steps (stepped approach). The location of waypoints and the IAS restrictions were the same for both approaches. A difference in energy rate demands can be observed (Fig. A1), but all remain below one.

Approach pair H considers the independent variable "distance available on localizer-intercept heading". The total amount of track miles is equal for both approaches. For approach 3, localizer-

Approach number	No. of heading changes	No. of altitude steps	Max energy rate demand: first part	Energy rate demand LOC intercept heading	Max energy rate demand: final part	Heading change towards LOC intercept heading, deg	Mass, 10^3 lb	Lineup distance, nm	Distance on ocalizer-intercept heading, nm	FAF altitude, ft	FAF speed, kt	Localizer-intercept angle, deg	Localizer-intercept speed (IAS), kt	Localizer-intercept speed (groundspped), kt	Distance IAF-IF, nm	Distance IF-FAF, nm	Stabilized accordingg to Monte Carlo simulation, %
1	3	CDA	0,2	0,1	0,1	20	237	9	5	2000	180	60	180	178	30	3	100
2	6	CDA	0,8	0,1	-0,1	20	237	9	5	2000	180	60	180	177	30	3	100
3	1	CDA	·	0,6	0,9	·	237	7	5.1	1590	160	45	180	170	5	2	100
4	1	1	·	-0,2	-0,1	·	237	7	5.1	1600	160	45	180	176	5	2	100
5	2	CDA	1,6	0,9	1,2	90	237	7	4.1	1590	160	90	180	180	20	2	100
6	1	CDA	·	0,8	0,3	·	237	10	5	1600	160	45	180	170	5	5	100
7	2	CDA	1,8	3,7	14,9	45	237	5	3	1270	160	35	180	181	8	1	0
8	2	CDA	0,6	2,8	5,4	45	237	5	3	1270	160	35	180	181	8	1	0
9	2	CDA	0,6	0,0	-0,2	30	237	7	1.6	1590	160	45	180	179	5	2	100
10	2	CDA	0,5	-0,3	-0,1	50	237	7	1.6	1590	160	45	180	175	5	2	100
11	2	CDA	0,8	0,4	3,7	30	237	7	1.6	1590	160	45	240	240	5	2	0
12	1	CDA	·	1,3	3,0	·	237	5	4	1350	160	45	170	150	4	1	0
13	1	CDA	·	2,1	13,0	·	237	5	4	1350	160	45	190	205	4	1	0
14	1	CDA	·	1,8	9,5	·	237	5	4	1350	160	45	180	176	4	1	0
16	6	3	0,2	-0,3	-0,3	20	237	9	5	2000	180	60	180	180	30	3	100
17	4	CDA	1,0	0,1	0,3	30	237	5	3	1270	160	30	170	161	18	1	100
18	4	CDA	1,4	1,7	2,3	30	237	5	3	1270	160	30	170	205	18	1	0
19	4	CDA	0,6	0,4	2,9	30	237	5	3	1270	160	30	200	202	18	1	0
20	4	CDA	1,4	1,2	1,5	30	237	5	3	1270	160	30	170	166	18	1	3
21	4	CDA	0,7	0,1	0,1	30	295	5	3	1270	160	30	170	166	18	1	100

Fig. A1 Relevant data for all approaches.

intercept part is 5.1 nm, and there is no first part of the approach. For approach 9, localizer-intercept part is 1.6 nm long, and the remaining 3.5 nm constitutes the first part of the approach. Unfortunately, due to an incorrect wind setting during the experiment, there is also a difference in groundspeed during localizer intercept within this approach pair: 170 kt for approach 3, and 179 kt for approach 9. The energy rate demands are all below one.

Approach pair I considers the independent variable "lineup distance". Approach 3 has a lineup distance of 7 nm, and approach 6 has a lineup distance of 10 nm. Since the altitude of the FAF is the same for both approaches, this inevitably leads to the fact that the distance between IF and FAF will differ for both approaches, as well as the energy rate demand between IF and FAF (but remains below one).

Acknowledgments

This research was financially supported by the Dutch Technology Foundation STW (DLR7071). We would like to thank all pilots who participated in the flight simulator experiment.

References

[1] Stassen, H. G., Johannsen, G., and Moray, N., "Internal Representation, Internal Model, Human Performance Model and Mental Workload," *Automatica*, Vol. 26, No. 4, 1990, pp. 811–820. doi:10.1016/0005-1098(90)90057-O

[2] Heiligers, M. M., Van Holten, T., and Mulder, M., "Flight Mechanical Evaluation of Approaches," *Journal of Aircraft*, Vol. 48, No. 3, 2011, pp. 963–974.

[3] Heiligers, M. M., Van Holten, T., and Mulder, M., "Predicting Pilot Task Demand Load During Final Approach," *International Journal of Aviation Psychology*, Vol. 19, No. 4, Oct.–Dec. 2009, pp. 391–416. doi:10.1080/10508410902983987

[4] Baron, S., Zacharias, G., Muralidharan, R., and Lancraft, R., "PROCRU: A Model for Analyzing Flight Crew Procedures in Approach to Landing," *NASA-MIT 16th Annual Manual 1980*, 1980, pp. 488–520.

[5] Baron, S., Kruser, D. S., and Huey, M. B., *Quantitative Modeling of Human Performance in Complex, Dynamic Systems*, National Academy Press, Washington, D. C., 1990, pp. 1–108.

[6] Corker, K. M., "Human Performance Simulation in the Analysis of Advanced Air Traffic Management," *Proceedings of the 1999 Winter Simulation Conference*, Vol. 1, edited by P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, 1999, pp. 821–828.

[7] Gore, B. F., and Corker, K. M., "A Systems-Engineering Approach to Behavioral Prediction of an Advanced Air Traffic Management Concept," *19th Digital Avionics Systems Conference, Entering the Second Generation of Powered Flight, 1*, Philadelphia, 2000, pp. 1–8.

[8] Smith, B. R., and Tyler, S. W., "The Design and Application of MIDAS: A Constructive Simulation for Human-System Analysis," *2nd Simulation Technology and Training Conference*, Canberra, Australia, 1997.

[9] Vicente, K. J., *Cognitive Work Analysis, Towards Safe, Productive, and Healthy Computer-Based Word*, Lawrence Erlbaum Assoc., Mahwah, NJ, 1999, pp. 1–416.

[10] "Constructions of Visual and Instrument Flight Procedures," *Procedures for Air Navigation Services Aircraft Operations*, Vol. II, 5th ed., International Civil Aviation Organization, Montréal, 2006, pp. I-1-1-1–IV-1-17.

[11] Godley, S. T., "Perceived Pilot Workload and Perceived Safety of RNAV (GNSS) Approaches," Australian Transport Safety Bureau Rept. 20050342, Canberra, Australia, 2006.

[12] Vormer, F. J. P., *Theoretical and Operational Aspects of Optimal Airport Arrival Trajectories*, Delft Univ. of Technology, Delft, The Netherlands, 8 Dec. 2005, pp. 1–240.

[13] Hanke, R. C., and Nordwall, D. R., "The Simulation of a Jumbo Jet Transport Aircraft Volume II: Modeling Data," NASA Rept. D6-30643, Sept. 1970.

[14] Zijlstra, F. R. H., "Efficiency in Work Behavior: A Design Approach for Modern Tools," PhD Thesis, Delft Univ. of Technology, Delft, The Netherlands 1993.

[15] Zijlstra, F., and Meijman, Th., "Het Meten van Mentale Inspanning met Behulp van een Subjectieve Methode," *Mentale Belasting en Werkstress, een Arbeidspsychologische Benadering*, edited by T. Meijman, Van Gorcum, Assen/Maastricht, The Netherlands, 1989, pp. 42–61.

[16] Lodge, M., *Magnitude Scaling, Quantitative Measurement of Opinions*, Sage University Paper Series on Quantitative Applications

in the Social Sciences, Sage, Newbury Park, CA, pp. 7–25.
[17] Meijman, T., Zijlstra, F., Kompier, M., Mulders, H., and Broersen, J., *The Measurement of Perceived Effort, Contemporary Ergonomics 1986*, edited by D. J. Oborne, Taylor and Frances, London, 1986, pp. 242–247.
[18] Hart, S. G., and Staveland, L. E., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in psychology*, Vol. 52, 1988, pp. 139–183. doi:10.1016/S0166-4115(08)62386-9
[19] "Killers in Aviation: FSF Task Force Presents Facts About Approach-and-Landing and Controlled-Flight-Into-Terrain Accidents," *Flight Safety Digest*, Nov. 1998–Feb. 1999, pp. 1–278.