

Validation of an on-board taxi guidance system.

Marcus Biella

German Aerospace Center

DLR Institute for Flight Guidance, Braunschweig, Germany

Marcus.Biella@dlr.de

An on-board taxi guidance system with displayed ground traffic and data link to air traffic control was validated in two main studies with 57 commercial airline pilots at DLR's Institute of Flight Guidance simulation facilities. Following the validation guidelines by EUROCONTROL, results in terms of operational feasibility and operational improvements were gained. In a first step, eye gaze data were analysed to control the effect of increased head down times. In a second step (within the EU project EMMA2), hints for operational improvements were gathered in shape of questionnaires regarding workload and situation awareness.

1. Introduction

Especially on large airports under low visibility conditions, cockpit crews are still faced with difficulties regarding the clearances and the whereabouts of other traffic (Lorenz et al., 2007). To enhance safety and capacity in the future, Advanced Surface Movements Guidance and Control (A-SMGCS) activities are conducted. Different services were implemented in various test beds to show their technical feasibility in the first place. Beyond these technical tests, validation activities aim to check the operational feasibility in terms of users' acceptance and operational improvement e.g. in terms of safety, human factors.

Within the DLR internal project MOSES (2001-2005) the **operational feasibility** of the on-board taxi guidance system TARMAC-AS (Taxi And Ramp Management And Control – Airborne System [Härtl, 1997]) was tested. The system was provided to both pilots on the navigation display (ND) in after landing. All clearances were performed by a simulated data link while the traditional radio telephony was only kept as a backbone. In this study, the system was used directly via the ND touchscreen. It was crucial to test if the TARMAC-AS HMI does not provoke an attentional capture because pilot's main attention should not shift from the outside view to the display (Biella, 2009). Within the EU project EMMA2 (2006-2009) first hints on **operational improvements** were studied. The system was now equipped with a real data link functionality developed by DLR, ATN and Park Air Systems. An airworthy interface (CDTI) was modified and served here the input device. Traditional radio telephony was kept especially for time and safety critical clearances like initial calls, runway crossings, line-up and take off. It was crucial to test effects on safety and human factors issues.

2. Method

2.1 Participants

Five female and 35 male pilot students of Lufthansa Flight Training took part in the first of two **Operational Feasibility Studies** (Lorenz et al., 2007; Biella, 2009). Either they already held the commercial pilot licence or they were on the brink of the practical exam. Their average age

was 23.9 years (standard deviation=1.8) and the mean reported flight hours were 225 (standard deviation=66). The second part of the study was conducted with one female and seven male commercial airline pilots. Their average age was 42.6 years (standard deviation=8.6) and the mean reported flight hours were 5478 (standard deviation=2898). In the **Operational Improvement Studies** (Wittig et al., 2010) eight male pilots took part, seven of them were commercial airline pilots. Their average age was 42.4 years (standard deviation=15.3) and the mean reported flight hours were 7471 (standard deviation=7731). In both cases validation guidelines (MAEVA, E-OCVM) by EUROCONTROL were applied.

2.2 Equipment

Both studies took place in DLR's fixed-base flight simulator Generic Experimental Cockpit (GECO). It is designed on the Airbus A 320 architecture using the flight dynamics of the VFW614 in a fly by wire version. The ND provides the taxi guidance system TARMAC-AS (Härtl, 1997) with a Ground Traffic Display (GTD) and a Taxi-Controller Pilot Data Link functionality (TAXI-CPDLC).

2.3 Simulation scenarios and procedure

In the **Operational Feasibility Studies** (Lorenz et al., 2007; Biella, 2009) each pilot flew a total of eight scenarios containing approach, landing and taxiing on Zurich Kloten airport under different visibility conditions. The investigator acted as Pilot Non Flying without actively intervening in the events. In these sessions, the ATC side was simulated by a pseudo controller within the GECO environment. Two different test conditions were applied: baseline taxiing with paper chart vs. use taxi guidance systems via navigation display. In the **Operational Improvement Studies** (Wittig et al., 2010) each pilot flew a total of eight experimental scenarios, four of them as Pilot Flying resp. as Pilot Non Flying. Scenarios consisted of an inbound and outbound leg including taxiing on Prague Ruzyne airport under CAT I visibility conditions. Four different test conditions were applied: baseline, Electronic Moving Map (EMM), EMM plus GTD, EMM plus GTD and standard TAXI-CPDLC, EMM plus GTD and advanced TAXI-CPDLC including a clearance update by the Ground Executive Controller. In these sessions, a controller team from the Prague (Air Navigation Service Provider, Czech Republic) was responsible for ATC using DLR's Apron and Tower Simulator which was connected to the GECO together with other traffic by pseudo pilots. The complete procedures are described in Wittig et al. (2010). The A-SMGCS provides expected route information by TAXI-CPDLC on pilot's request, after the departure clearance has been issued in outbound legs resp. before final approach in inbound legs.

2.4 Assessment of Operational Feasibility and Operational Improvement

2.4.1 Operational Feasibility via pilots' gaze

Pilots' eye movements during taxiing were measured with the system iViewX by SMI (resolution: <0.1 degrees [pupil]; <0.5 degrees [compensated for movements]; frequency: 50 Hz). Raw data were condensed into cumulated dwell times for the apron segment on predefined areas of interest. Dwell times reveal if head down times increase significantly while an on-board taxi guidance system is in use (Biella, 2009).

2.4.2 Operational Feasibility via tailor-made questionnaire

A tailor-made questionnaire (Wittig et al., 2010) was created to check the fulfilment of the requirements basing on the description in the Systems, Procedures and Operational Requirements document of EMMA2. Each of the 72 items has a six point Likert scale ranging from from 1 "strongly disagree" to 6 "strongly agree". After finishing off all of test runs, the questionnaire was given to the pilots. Their response reveals if basic requirements are fulfilled if an on-board taxi guidance system is in use.

2.4.3 Operational Feasibility via Standard Usability Scale (SUS)

The SUS (Brooke, 1996) was given to the pilots after the very last trial. Answers range from 1 “strongly disagree” to 5 “strongly agree” on a five point Likert scale. After finishing off all of test runs, the SUS was given to the pilots. Their response reveals if usability in general is fulfilled if an on-board taxi guidance system is in use (Wittig et al., 2010).

2.4.4 Operational Improvements via tailor-made questionnaire

An additional tailor-made questionnaire (Wittig et al., 2010) was created to check the operational improvements e.g. in terms of safety basing on the “expected benefits” that were identified in EMMA2’s “Systems, Procedures and Operational Requirements” document (www.dlr.de/emma2/). Each item has a six point Likert scale ranging from from 1 “strongly disagree” to 6 “strongly agree”. After finishing off all of test runs, the questionnaire was given to the pilots. Their response reveals if safety increases is perceived by the pilots while using a taxi guidance system (Wittig et al., 2010).

2.4.5 Operational Improvements via Workload and Situation Awareness questionnaires

The Instantaneous Self Assessment Scale (ISA) was originally conceived at NATS. It was applied within EMMA2 trials twice during the inbound and the outbound segment of each test run. The scale ranges from “low” (1) to “high” (5) for workload and from “low” (1) to “high” (10) for situation awareness. Pilots were questioned approx. in the middle of the segment (e.g. after a runway cross) and at the end of the segment (before reaching the gate resp. before line-up). Pilots’ response reveals if an increase of workload resp. situation awareness is perceived by the pilots while using a taxi guidance system.

3. Results

3.1 Change of pilots’ gaze depending on use of taxi guidance system

A two-way analysis of variance with repeated measures was used to assess the effect of the taxi guidance system on pilot’s information gathering. Factor 1 represents taxi guidance support (standard paper chart vs. taxi guidance display), factor 2 pilots’ experience (flight students vs. experienced pilots). Dependant variable in each analysis was the cumulated dwell time on an area of interest for the taxi phase on apron, averaged over four scenarios with resp. without taxi guidance support. Both in the condition with and without taxi guidance support the main source of information remains the outside view. Nevertheless, the difference between 53.4% in the baseline condition and 46.0% in the experimental condition is highly statistically significant ($p < 0.01$). The second most important source of information is the taxi chart, with 8.7% for the paper chart in the baseline condition resp. 13.7% for the navigation display in the experimental condition. Independent from the taxi guidance display, experienced pilots use the outside view more than flight students do ($p < 0.01$) (Biella, 2009).

3.2 Further Results to Operational Feasibility (RTS)

Since the sample size is only eight different pilots per item, the binominal test as a non-parametric statistic was used to prove the results of the questionnaires for their statistical significance. Finally, 56 operational requirements or procedures could be regarded as verified, twenty of them statistically significant. The remaining 16 operational requirements were not answered positively which was caused by the use of the modified airworthy interface (CDTI) which was unfamiliar for the pilots.

For example, pilots agree that the new cockpit services were well integrated into the existing systems. This result is highly statistically significant for the GTD ($M=5.38$; $sd=0.74$; $p < 0.01$) but only significant by trend for TAXI-CPDLC ($M=3.88$; $sd=0.83$ and $p=0.07$). According to

the pilots the GTD is capable of being used appropriately when operating within the movement area ($M=5.25$; $sd=1.04$; $p<0.01$); results for the TAXI-CPDLC are positive as well ($M=4.25$; $sd=1.28$; $p=0.29$) but not significant (Biella, 2009).

3.3 SUS Questionnaire

The pilots answered all ten items in favour of the new system. Three of them became even highly statistically significant: According to the pilots there was not too much inconsistency in the system ($M=2.50$; $sd=0.53$; $p<0.01$). They find the system not very difficult to use ($M=2.38$; $sd=0.92$). Finally, they agree that they do not need to learn a lot of things before start to work with the system ($M=1.88$; $sd=0.83$; $p<0.01$) (Wittig et al., 2010).

3.4 Operational Improvements Questionnaire, Sub-Scale Safety

Highly statistically significant results include that the graphical taxi clearances on the display would enable pilots to follow an assigned taxi route more safely ($M=5.63$; $SD=0.52$; $p<0.01$). Furthermore the indication of the surrounding traffic is an additional information source to navigate and manage the aircraft speed more safely ($M=5.75$; $sd=0.46$; $p<0.01$) (Wittig et al., 2010).

3.5 Workload and Situation Awareness

The ISA means of each test run were analysed in a 4 x 2 (A-SMGCS treatment x pilot role) two-way repeated measures analysis of variance (ANOVA).

Regarding workload the ANOVA revealed a highly significant main effect of the A-SMGCS treatment ($F_{(3,21)} = 5.418$; $p<0.01$) with a mean of $M = 2.20$ for the (EMM) baseline, respectively $M = 2.12$ for GTD and $M = 2.48$ for standard TAXI-CPDLC (incl. GTD) and $M = 2.36$ for advanced TAXI-CPDLC (incl. GTD) on a scale reaching from “under-utilised” (1) to “excessive” (5). No significant main effect could be shown for the pilot role ($F_{(1,7)} = 1.932$; $p = .207$), both pilot flying and pilot non flying seem to have a similar level of workload. Yet there is a significance for the interaction between treatment and pilot role ($F_{(3,21)} = 3.743$; $p<0.05$), with the highest mean $M = 2.52$ in the condition pilot non flying using advanced TAXI-CPDLC.

Regarding situation awareness the ANOVA revealed no significant main effect of the A-SMGCS treatment ($F_{(3,21)} = .146$; $p = .931$) with a mean of $M = 8.09$ for the (EMM) baseline, respectively $M = 9.03$ for GTD and $M = 8.70$ for standard and $M = 8.74$ for advanced TAXI-CPDLC (incl. GTD) on a scale ranging from “low” (1) to “high” (10). There is no significant main effect for the pilot role ($F_{(1,7)} = 1.337$; $p = .286$), both pilot flying and pilot non flying seem to have a similar level of situation awareness. There is no significance for the interaction between treatment and pilot role ($F_{(3,21)} = 1.391$; $p = .273$) (Wittig et al., 2010).

4. Discussion

Regarding **operational feasibility** it could be shown that the use of the taxi guidance system results in a different pattern of dwell times without resulting in an attentional capture on the navigation display as no incidences occurred while using the display. Single case studies indicate that attention changes in form of a “trade off” only from the taxi charts to the navigation display (Lorenz et al., 2004). Main source of information by far remains the outside view. Further single case studies reveal that pilots use the display especially only during holding times and therefore not in safety critical situations (Biella, 2009). Compared with simulator data it can be shown that pilots do not taxi faster while using TARMAC-AS compared to scenarios with traditional taxi charts (Lorenz et al., 2004).

The GTD was appreciated and highlighted by all pilots. The added value is seen especially under low visibility conditions. TAXI-CPDLC is regarded as potential means which will ease pilots work considerably. It must be distinguished clearly between the positively rated graphical

information on the EMM display and the intermingled rating of the usability of the CDTI, which served for selecting CPDLC messages to be sent and for the textual displaying of received clearances. The graphical presentation of the cleared taxi route on the EMM (compared to textual clearances on the CDTI) will enable pilots to follow an assigned taxi route in an intuitive way. Debriefing revealed that TAXI-CPDLC operation requires too much head down times while using the CDTI as interface. This means either the aircraft needs to be stopped (time is lost) or redundancy of the pilot non flying who operates the CDTI is lost. Pilots responded that they would like to use the TAXI-CPDLC in the future and that they agree in general with its concept given and realised in EMMA2. It should be stressed that use of the CDTI was necessary in the studies because this certified interface was necessary for preparing and conducting the subsequent flight trials in DLR's test aircraft with pilots and additional safety pilots which were conducted successfully (Wittig et al., 2010).

Methodically the results encourage to follow E-OCVM in general and to develop tailor-made questionnaire according to the operational requirements of the project and to conduct thoroughly debriefings as they deliver more important results than standard questionnaires. For example, pilots say in the SUS that they don't need to learn a lot of things before start to work with the system. Only comments in the tailor-made questionnaire and in the debriefing reveal the novelty and problems of the input device. Pilots made additional remarks in the debriefing how to solve that problems, e.g. by integrating so called Left and Right Line Select Keys and by ensuring accessibility of the main menu to have easier access to new and old messages in the system (Wittig et al., 2010).

Regarding **operational improvement** it could be shown that pilots experience an increment of safety. Yet further tests with an improved CDTI are necessary to check if a decrement of workload and an enhancement of situation awareness will take place. So far, this could be shown for the TARMAC-AS display but not for the input device (Wittig et al., 2010).

Literature

- Biella, M. (2009). Pilot gaze performance in critical flight phases and during taxiing: Results from DLR-project MOSES. In W. Kallus et al. (Ed.), *Aviation Psychology in Austria. Human Factors and Resources*. (Pp. 61-70). Wien: Facultas.
- Brooke, J. (1996). "SUS: a "quick and dirty" usability scale". in P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland. *Usability Evaluation in Industry*. London: Taylor and Francis. <http://www.usabilitynet.org/trump/documents/Suschapt.doc>.
- E-OCVM http://www.eurocontrol.int/valfor/gallery/content/public/E-OCVM_v2_Small.pdf
- Härtl, D. (1997). *Investigations with an Aircraft Taxi Assistance and Guidance System in a Future Airport Traffic Scenario (DLR Forschungsbericht)*. Köln: DLR.
- Lorenz, B., Biella, M., Schmerwitz, S., & Többen, H. (2004). *Verlässlichkeit der Interaktion zwischen Pilot und Assistenzsystem*. In M. Grandt (Ed.), *Verlässlichkeit der Mensch-Maschine-Interaktion* (Pp. 271-294). Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt e.V.
- Lorenz, B., Biella, M., Teegen, U., Stelling, D., Wenzel, J., Jakobi, J., Ludwig, T., & Korn, B. (2007). Performance, situation awareness, and visual scanning of pilots receiving onboard taxi navigation support during simulated airport surface operation. *Human Factors and Aerospace Safety*, 6, 135-154.
- MAEVA http://www.eurocontrol.int/valfor/gallery/content/public/MAEVA_VGH_Part1.pdf
- Wittig, T., Biella, M., Jakobi, J., Ludwig, T., Wehrstedt, C., Drege, C., Urvoy, C., & Friebel, R. (2010). *Airborne Validation Results. Part A. Project report. EMMA2 2-D6.6.1a*. http://www.dlr.de/emma2/maindoc/2-D661a_VO-TR_V1.0.pdf