



## **Optimising aircraft taxi speed: Design and evaluation of new means to present information on a head-up display**

Downloaded from: <https://research.chalmers.se>, 2025-04-20 13:32 UTC

Citation for the original published paper (version of record):

Eklund, R., Osvalder, A. (2021). Optimising aircraft taxi speed: Design and evaluation of new means to present information on a head-up display. *Journal of Navigation*, 74(6): 1305-1335.  
<http://dx.doi.org/10.1017/S0373463321000606>

N.B. When citing this work, cite the original published paper.



RESEARCH ARTICLE

# Optimising aircraft taxi speed: Design and evaluation of new means to present information on a head-up display

Rikard Eklund,\*  and Anna-Lisa Osvalder

Chalmers University of Technology, Gothenburg, Sweden.

\*Corresponding author. E-mail: [rikard.eklund@chalmers.se](mailto:rikard.eklund@chalmers.se)

**Received:** 23 December 2020; **Accepted:** 26 June 2021; **First published online:** 3 August 2021

**Keywords:** aviation, taxiing, optimisation, graphical user interfaces (GUI), head-up display

## Abstract

The objective of this study was to design and evaluate new means of complying to time constraints by presenting aircraft target taxi speeds on a head-up display (HUD). Four different HUD presentations were iteratively developed from paper sketches into digital prototypes. Each HUD presentation reflected different levels of information presentation. A subsequent evaluation included 32 pilots, with varying flight experience, in usability tests. The participants subjectively assessed which information was most useful to comply with time constraints. The assessment was based on six themes including information, workload, situational awareness, stress, support and usability. The evaluation consisted of computer-simulated taxi-runs, self-assessments and statistical analysis. Information provided by a graphical vertical tape descriptive/predictive HUD presentation, including alpha-numerical information redundancy, was rated most useful. Differences between novice and expert pilots can be resolved by incorporating combinations of graphics and alpha-numeric presentations. The findings can be applied for further studies of combining navigational and time-keeping HUD support during taxi.

## 1. Introduction

### 1.1. General

On an average day, 28,000 flights are conducted in the European airspace and the average delay for departure is 33 min (SESAR, 2015; EUROCONTROL, 2020). Delays have direct negative implications on, for example, operating costs for fuel, staff, maintenance and route planning (EUROCONTROL, 2020). Delays are absorbed at the airport departure gate, when queuing on taxiways or waiting in holding bays. The overall goal of the airport ground operations is to achieve a conflict-free and effective route for aircraft to/from the airport terminal (gates/stands) and to the runway (Ravizza et al., 2013). Airport surface congestion is responsible for increased taxi times, and increased fuel consumption and emissions (Simaiakis et al., 2014). There is a potential to resolve these issues by optimising surface operations (Lee et al., 2010).

Departures are primarily managed by issuing a calculated take-off time (CTOT), which holds the aircraft at the gate until departure time is due. CTOT is a part of the ATC (air traffic control) clearance (EUROCONTROL, 2017). Estimated off-block time, the estimated time at which an aircraft commences movement associated with departure, governs planning for the actual time of ground movement (ICAO, 2007). The reason for constraining aircraft departure is to avoid further congestion in the air en route and/or at the destination airport. An issued CTOT is communicated to the pilots by radio, data link, billboards at the parking position and/or via handling agents. During taxi, pilots use engine thrust and brakes (unlike jet aircraft, propeller aircraft may also select propellers into reverse to slow down the

aircraft) to vary ground speed to reach the assigned CTOT. Basic time-speed-distance calculations, or speed instructions from the ATC, guide the aircraft's ground speed towards the take-off runway. These calculations may be based on experience and rules-of-thumb, for pilots as well as for the ATC, for the airport infrastructure in question. Gotteland et al. (2003) suggest one of three generic aims for the ATC: (1) as early as possible an arrival to the runway, (2) reaching the runway as close to the CTOT as possible or (3) to reach the runway within a specific time window, usually minus 5 min to plus 10 min of the CTOT. Future traffic flows and four-dimensional (4D) flight trajectory planning (integration of time into the three-dimensional flight trajectory) may require precision as high as a plus/minus one-minute time window (SESAR, 2015).

The reduction of taxi time has been analysed as a target for effective airport operations (Atkin et al., 2011), often supported by means of penalties when not conforming to a CTOT (Balakrishnan and Jung, 2007). Automated systems such as the surface management systems (SMS) for ATC use (Atkins and Brinton, 2002) and the Taxiway Navigation and Situation Awareness (T-NASA) system for cockpit use, were developed to increase situational awareness (Foyle et al., 1996; Cheng et al., 2009). Airport collaborative decision making (A-CDM), where airport and aircraft operators, ground handlers and ATC collaborate to exchange accurate and timely information (ICAO, 2007; EUROCONTROL, 2017), are examples of systems introduced to improve airport efficiency and capacity planning. Complex SMSs such as the Advanced Surface Movement Guidance and Control System (A-SMGCS) perform automated surveillance, routing, guidance and control functions (Hecker et al., 2001). The essential function of A-SMGCS is the automated routing function which improves time efficiency by identifying an optimised ground route.

By optimising the 4D flight trajectory from start-up at the departure gate, via taxi, take-off, cruise, approach, landing, taxi and engine shut down at the arrival gate, fuel consumption can be reduced (Grote et al., 2014). Research on 4D flight trajectories has resulted in improved predictability of the flight segment. However, the ground segment is typically computed with an assumed ground speed, risking unrealistic planning (Marín, 2006; Couluris et al., 2008).

Furthermore, a reduction in fuel consumption, and thereby emissions, can be achieved by managing taxi speeds, primarily by reducing the number of accelerations and retardations (Green, 2011; Simaiakis et al., 2014). Nikoleris et al. (2011) suggest that the taxi segment can be divided into various stages with an estimated engine thrust setting varying between 4–9% of full engine thrust. Stettler et al. (2011) used a value of 4–7% for constant speed taxi and 7–17% for acceleration on taxiways. Nagaoka et al. (1978) showed that average taxi speed increased when the taxi distance increased. Trade-off estimations, in terms of fuel consumption, thrust settings, taxi distance and time constraints, must at the present time be performed manually.

## ***1.2. Presenting information in cockpit***

Presenting information such as speed, altitude, course, pitch, roll, position, terrain and traffic avoidance and systems status such as engine parameters, to the pilots is critical for their interpretation, and interaction with their environment (Shappell et al., 2007; Soo et al., 2016). Information can be presented by different interfaces: auditive by the use of speakers, visually by the use of displays, gauges and lights as well as haptically by the use of vibrations in pedals, control wheels and side sticks (Green et al., 2001; van Veen and van Erp, 2001).

Mechanical and analogue instruments have evolved over time into digital electronic systems including sensors, computers and electronic mediums of presentation such as the cathode ray tube, liquid crystal display or the head-up display (HUD) (Craig, 2012). The user interface must provide the pilot with timely and easily understood information. Nonetheless, research shows that serious incidents and accidents continue to occur due to decision and perceptual errors based on inadequate information presentation (Shappell et al., 2007). The evolution of cockpit information presentation intends to reduce pilot workload and increase situational awareness by reducing the need for the continuous use of mental models (Wiener, 1988, 1989; Sarter and Woods, 1992). A mental model is a cognitive representation



**Figure 1.** Example of HUD installation (Collins Aerospace HGS 6000). Note: Approved for publication by Collins Aerospace.

of system performance enabling the user to reason about system behaviour (Moray, 1996). When this model does not correspond to the actual system performance, a mismatch occurs which must be addressed (Palmer, 1995). Endsley (2015) points out that the requirement for creating mental models still exists since the available information afforded to the pilots has increased over time. To avoid the risk of creating overwhelming, ambiguous or confusing user information, measures have to be taken to minimise, standardise and prioritise which information to display on the cockpit instruments (Hawkins, 1987; Foyle et al., 1995; de Winter and Dodou, 2014).

Another critical factor when designing a cockpit is the crew itself. Most commercial aircraft are operated by a multi-crew (the crew of an aircraft that is equipped and required by the regulations to be operated by a crew of at least two pilots) where the pilots might have different backgrounds in terms of experience, culture and training as well as depending on (daily) cognitive prerequisites. Cockpit interfaces therefore need to match different cognitive strategies applied by both novices and experienced pilots (Stokes and Kite, 1997; Sexton et al., 2000).

### 1.3. HUD in aviation

Presentation of visual information can be achieved by head-down display (HDD), where pilots maintain their vision inside the cockpit by looking down on instrument panels, and/or by HUD (Figure 1), where pilots during specific flight phases maintain their visual attention out through the cockpit windows (Brickner and Foyle, 1990). HUD technology has a long history in the aerospace industry (Weintraub and Ensing, 1992) and is based on reflective gunsight technology from aircraft in the Second World War (Newman, 1995). HUD has also been developed and used to varying degrees in the automotive industry as well as extensively in military applications (Gish and Staplin, 1995; Holder and Pecota, 2011).

The working principle of HUD is that a projector is used to project information on a semi-transparent surface (combiner) in front of the pilot (Newman, 1987). To fuse the outside world, as seen through the HUD, information displayed must be collimated to mitigate constant changes of eye-focus. The far end for focus is infinity and the close end is the combiner itself (Newman, 1987). The pilot's view through the HUD is thus a combination of the inside world and the outside real world. HUD decreases scanning requirements (Martin-Emerson and Wickens, 1997) by presenting information at one location and in the outside world line-of-sight ensuring that information inside and outside the cockpit can be processed simultaneously. Reduced scanning requirements in turn increases situational awareness (Foyle et al., 1995; Proctor, 1999, 2000). This is accomplished by reduced dwelling times between the outside and inside world as well as reduced amounts of information to be stored in the short-term memory. Furthermore, the requirements of creating a mental model of the flight trajectory,

by synthesising information from different sources (and continuously updating that model), can be mitigated by gathering basic information at one location (Endsley, 2015).

During critical phases of the flight trajectory, such as take-off, climb, approach, flare, landing and roll-out, HUD is effective since task-specific information such as speed, pitch and vertical and lateral information can be projected in the pilot's line-of-sight (Stuart et al., 2003). HUD can, however, degrade or slow down the detection of unforeseen events (Fadden et al., 1998; Wickens and Alexander, 2009) due to pilots tending to cognitively tunnel their attention into the HUD. This occurs at the expense of outside world events by inefficient joint processing of the inside world superimposed on the HUD and the outside world as seen through it (Foyle et al., 1991).

#### **1.4. Purpose**

The airborne segment of the 4D flight trajectory has been studied to improve precision and predictability in terms of navigation and time-keeping. Studies of the ground segment (taxi) of the 4D flight trajectory indicate that improvements in time-keeping could be further studied. Based on a systematic literature review for this study, and in combination with the first author's applied experience as a commercial pilot, the purpose of the study is (1) to design a HUD presentation that could improve time-keeping during taxi and (2) to evaluate which information on such a HUD is rated the most useful for novice as well as expert pilots.

## **2. Methods**

### **2.1. Design process**

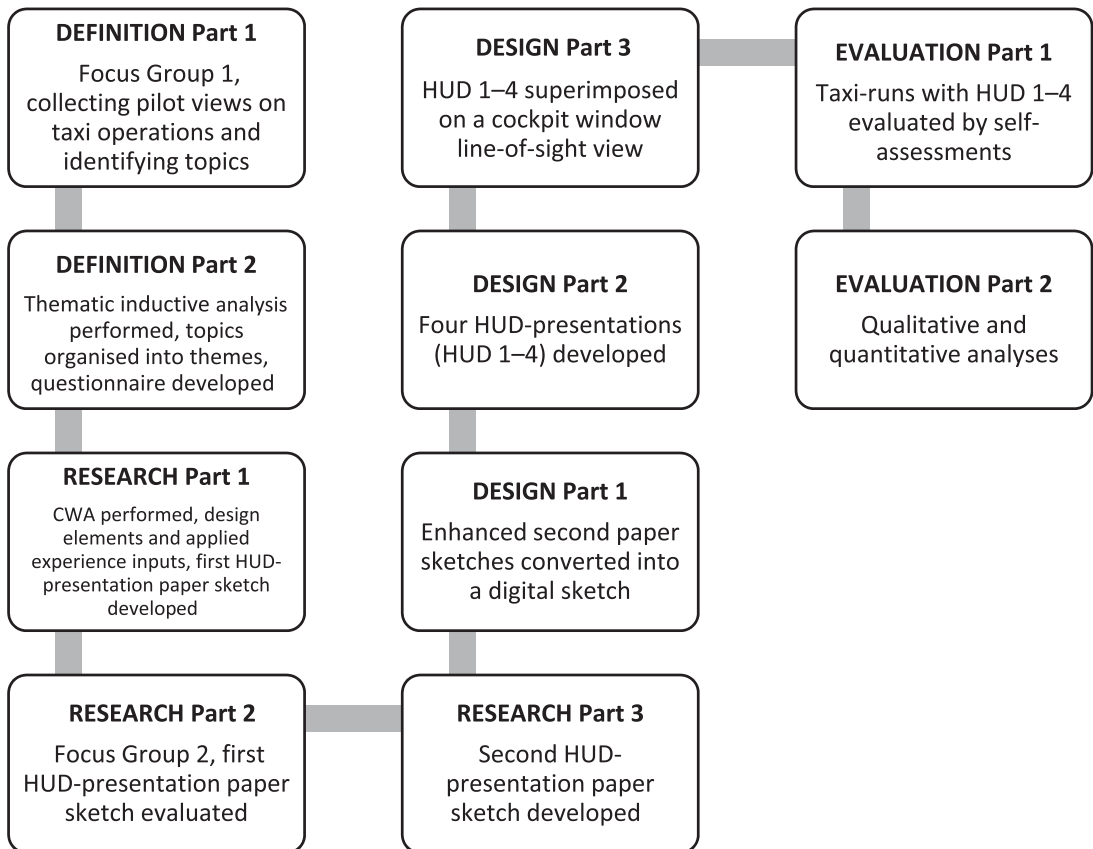
A design process is generally associated with product development where a product is conceived, designed and commercialised (Lawson, 2006). However, the present study was limited to product development and evaluation, excluding market research, manufacturing, certification and commercialisation. A process can be defined as a series of activities, changes or functions which combined leads to a result (Karlsson et al., 2009). Markus (1969) and Maver (1970) suggest a four-step linear design process of analysis, synthesis, appraisal and decision. Osvelder et al. (2009) describes a generic process including data collection, data analysis, development of products, systems and environments and presentation of results. Linear and sequential processes have been criticised by Lawson (2006), who suggests focusing to a greater extent on the tools and the components used compared with identifying specific steps. Furthermore, two basic prerequisites in a human-machine interface design are that the design process is driven by iteration and must involve the users that will handle the end product (Karlsson et al., 2009).

To match the objective of the study, a generic four-step design process adapted from Kragt (1992) was applied (Figure 2). The design process included: (1) definition, (2) research, (3), design, and (4) evaluation. Furthermore, the process was not developed or evaluated as such.

### **2.2. Definition part 1**

A first focus group two-hour session was completed and organised as a group interview with open-ended questions to increase the understanding of pilots' needs and requirements during taxi when CTOTs were in force. Organised discussions in groups can be effective to increase the understanding of individuals and their ideas, feelings and attitudes (Krueger and Casey, 2000). Furthermore, Virzi (1992) suggests that 80% of potential usability problems can be detected using focus groups. The ideal number of participants is approximately five participants (Rubenstein et al., 1984; Virzi, 1992).

Focus Group 1 consisted of five male pilots, all holding an Airline Transport Pilot Licence (ATPL) with a mean age of 48 years and a mean flight time of 10,228 flight hours. All five pilots were full-time employees at major airlines, with four working as captains and one as a co-pilot. The focus group session



**Figure 2.** Overview of the design process used in this study. Adapted from Kragt (1992).

was recorded on a voice-recording device and transcribed on level III (Linell, 1994). Additionally, notes were also taken as a complement to the recording.

### 2.3. Definition part 2

A thematic inductive analysis (Braun and Clarke, 2006) was used with emerging topics transcribed on level III (Linell, 1994), examined, categorised and clustered into six themes: information, workload, situational awareness, stress, support and usability. These six themes were subsequently calibrated on information as described by Rasmussen (1986) and Bisantz et al. (2009); on workload as described by Megaw (2005) and Sexton and Helmreich (2003); on situational awareness as described by Endsley (1996) and Smith and Hancock (1995); on stress as described by Lazarus and Folkman (1984) and Merritt and Helmreich (1996); on support as described by Morrison et al. (1998); and on usability as described by Shackel (1981) and Williams and Ball (2003).

Based on these descriptions, the following definitions were created to be used during design and evaluation. *Information*: the amount, presentation position, colour, layout, and practicality of presented alpha-numerical and graphical information on the four displays to be used to reduce uncertainty. *Workload*: the amount of cognitive effort in terms of work and/or time that was experienced when using the displays to make calculations and estimations and as a base for decision making. *Situational awareness*: the perception, understanding and assessment of environmental elements and/or events in terms of time or space, and the projection of future events to come. *Stress*: the physiological or psychological reaction on pressure to handle internal as well as external demands. *Support*: the overall support gained from the displays in terms of surveillance and decision making. *Usability*: the capability

of a tool, that can be used by specified users, to achieve a specified goal with effectiveness within a specified context.

Furthermore, the six themes were also operationalised into a psychometric six-point Likert scale questionnaire as variables to be used during the evaluation part.

## **2.4. Research part I**

### *2.4.1. Cognitive work analysis*

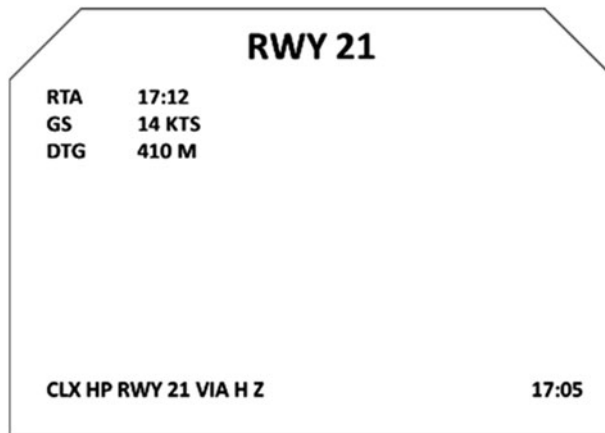
Work in complex socio-technical systems can be assessed by different methods. A common method is the cognitive work analysis (CWA) framework originally developed by Rasmussen et al. (1994). A variety of domains such as health care (Effken et al., 2011; Ashoori et al., 2014), military (Jenkins et al., 2008; Brady et al., 2013) and aerospace have utilised CWA. CWA identifies constraints enabling the creation of a model to understand how work is performed. CWA focuses on system design where human, social, technological and organisational prerequisites are reflected and converted into a design (Rasmussen et al., 1994). Furthermore, CWA is formative, focusing on what is required to perform the task regardless of the agent, the event or system environment (Roth and Bisantz, 2013). CWA can however elicit massive amounts of data, which makes it less useful in smaller research groups. CWA, in its original form, also analyses a system already in place followed by recommendations on (re)design (SESAR, 2021). The formative nature of CWA contrasts with the normative methods found in hierarchical transaction analysis (HTA) (Annett and Duncan, 1967), which focuses on understanding the task based on goals and sub-goals and the potential for error (Stanton, 2006; Salmon et al., 2010), or GOMS (goals, operators, methods, and selection rules), which produces a set of instructions on how work is to be completed (Benyon et al., 2005). HTA and GOMS have limited value in dynamic and unanticipated situations. Descriptive methods, focusing on observing what is done, also have a limited value in the design of new products but offer less limitation compared with normative analysis (Vicente, 1999).

In its basic form, CWA consists of five phases: (1) work domain analysis, where the ecology of the work is studied (Vicente, 1999) and where the product is a model of the work domain, (2) control task analysis, where work functions and specific tasks are studied (Vicente, 1999) and where the product is which information is required to execute specific tasks, actions and decisions, (3) strategies analysis on how the work can be implemented (Vicente, 1999) and where the product is a repository of different strategies for accomplishing work tasks, (4) social organisation and cooperation analysis of how work is distributed, communicated and coordinated among different agents in the system and where the product is a repository on who is generating information and who is receiving and acting on it, and (5) worker competence analysis, where the skill levels to perform a task are analysed and where a product is the identification of which competencies a worker needs to accomplish a specific task (Vicente, 1999).

Based on CWA (Rasmussen et al., 1994), in this study an adapted CWA was applied in order to identify, analyse and describe the tasks, constraints, information and actions within the cockpit work-environment during taxi (Appendix Figure A1).

### *2.4.2. Design elements*

The development of the HUD presentation was based on the following design elements (Appendix Table A1): consistency and relative location (U.S. Department of Transportation, 2014), control movement stereotypes (van Cott and Kinkade, 1972; U.S. Department of Defense, 2012), control response compatibility (Helander, 2006; Proctor and Vu, 2016), HUD screen layout (Wertheimer, 1923/1938; Palmer, 1992), emergent features (Dashevsky, 1964), contrast (Wood and Howells, 2001; U.S. Department of Defense, 2012; ESSE, 2021), characters, numbers and symbols attributes (Wood and Howells, 2001; Hirsch, 2004; U.S. Department of Defence, 2012; Academio, 2021; Eclipse, 2021), HUD screen dimensions (Collins Aerospace, 2021), display-eye distance, design eye position, head motion box and field of view (Wood and Howells, 2001; U.S. Department of Defense, 2012).



*Figure 3. HUD 1.*

Based on definition parts 1–2, the CWA, design elements, and the first author's applied experience as a commercial pilot, a first HUD presentation paper sketch prototype was developed.

### **2.5. Research part 2**

A second focus group session was created and organised as a group interview, including four male and one female pilot with a mean age of 51 years and a mean flight time of 10,898 flight hours, all five holding an ATPL. Four of the pilots were full-time employees at major airlines and one was a full-time pilot at a regional airline. All five were working as aircraft captains. These pilots assessed the first HUD presentation paper sketch in terms of usability. The focus group session was recorded on a voice-recording device and transcribed on level III (Linell, 1994). Additionally, notes were also taken as a complement to the recording.

### **2.6. Research part 3**

Based on the inputs from Focus Group 2, the first HUD presentation paper sketch was improved accordingly. The resulting second HUD presentation paper sketch was subsequently used during the first part of the design process.

### **2.7. Design part 1**

The second HUD presentation paper sketch was converted into a digital format by the use of vector graphics editing software (Adobe Illustrator 12.01).

### **2.8. Design part 2**

Four different HUD presentations (hereafter HUD 1–4) based on the second HUD presentation digital paper sketch and reflecting different levels of information were developed (Figures 3–6).

HUD 1 (Figure 3) reflected basic support, displaying descriptive alpha-numerical information. Alpha-numeric were used to display required time of arrival (RTA) in hour/minute format, ground speed (GS) in knots, and distance to go (DTG) in metres. In addition, actual time in hour/minute format, runway in use and a taxi clearance were also displayed.

HUD 2 (Figure 4) displayed descriptive alpha-numerical information as in HUD 1. However, it was also augmented by predictive information on estimated time of arrival (ETA) in hour/minute format and



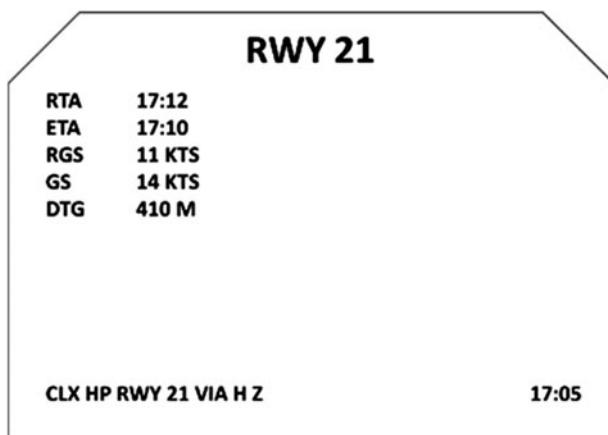


Figure 4. HUD 2.

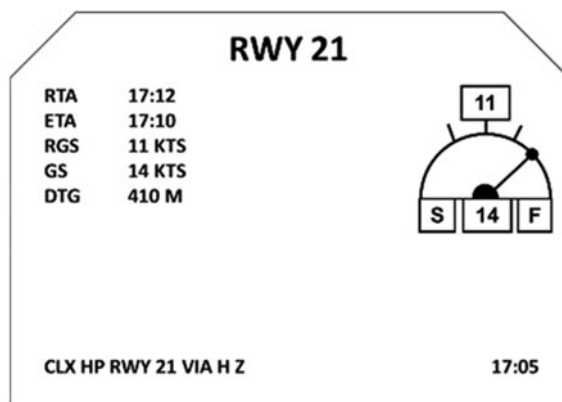


Figure 5. HUD 3.

required ground speed (RGS) in knots to support CTOT compliance. The combination of descriptive and predictive information can reduce cognitive workload (Meister, 1985), provide redundancy and increase vigilance (Lewis et al., 1990), as well as increase the level of situational awareness (Foyle et al., 1995; Proctor, 1999, 2000).

HUD 3 (Figure 5) provided redundancy where alpha-numerical and graphical information was displayed simultaneously. Presenting redundant information can improve situational awareness and lower workload result (Howarth and Bullimore, 2005). However, HUD 3 displayed not only descriptive/predictive information but also graphical information using a semi-circular, fixed index, moving pointer dial GS presentation. The utilisation of scales, dials and tapes with fixed and/or moving pointers can improve flight crew interpretation of numeric data (U.S. Department of Transportation, 2014). A circular dial is one of two basic ways to convey quantitative information, the other is by using a tape, where changes are easily detected (Harris, 2011). HUD 3 was also based on emergent features (Dashevsky, 1964) where alignment of the GS moving point dial with the RGS index vertically allows for quick detection of divergences.

HUD 4 (Figure 6) displayed the same redundant information as HUD 3. However, HUD 4 had a different graphical layout with a fixed index graphical scale, a moving GS ball indication, and a vertical tape presentation. Presenting information using a tape in combination with alpha-numerical characters is effective to determine variables such as speed and altitude (Zhang et al., 2002). The main disadvantage with the vertical tape presentation is that it is more difficult to detect trends (Zhang et al., 2002). As

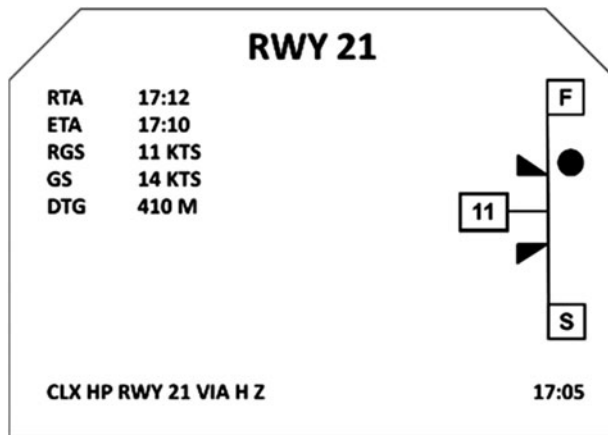


Figure 6. HUD 4.

with HUD 3, HUD 4 was also based on emergent features. Aligning the GS moving ball with the RGS index horizontally allows for quick detection of divergences.

### 2.9. Design part 3

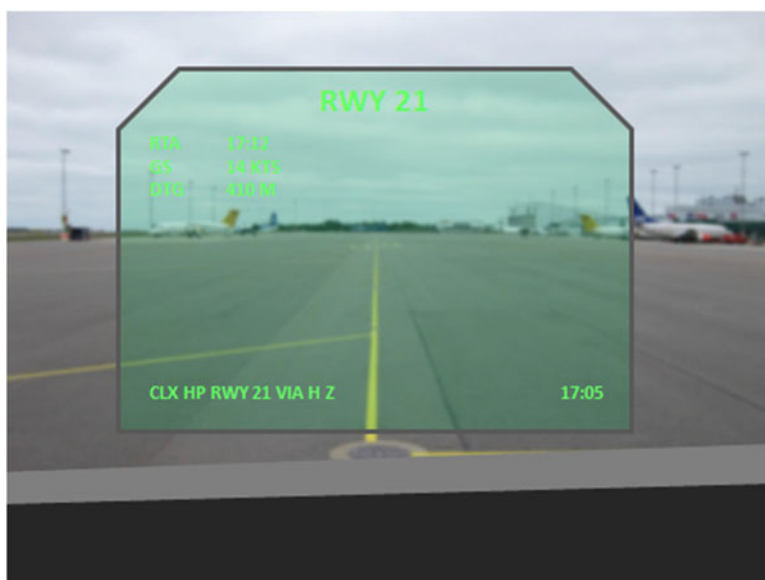
All four HUDs were individually superimposed on an actual airport infrastructure digital photograph to mimic a view out of the cockpit window.

### 2.10. Evaluation part 1

#### 2.10.1. Participants

A sample of 32 pilots of fixed wing aircraft ( $n=32$ ) participated in the usability tests. The sample consisted of 24 males and eight females with a mean age of 37 years ( $SD=11$ ). The number of flight hours varied between 241 and 17,278 with a mean of 6,610 flight hours ( $SD=5,485$ ). Of the participating pilots, 21 had flying as their main occupation while 11 were unemployed or had non-flying employments outside the aviation industry. The 21 employed pilots operated in multi-crew environment flying jet transport aircrafts (Airbus 320 or Boeing 737). Amongst them, 12 pilots were captains holding an ATPL, while the other nine pilots were co-pilots holding a Commercial Pilot Licence. Among the 11 unemployed pilots, seven had successfully completed a type rating for either Airbus 320 or Boeing 737. Furthermore, two of these seven pilots had prior employment experience. The remaining four pilots had no flying experience other than their flight training plus time-building flight hours. All pilots had ratings for instrument flying as well as for multi-engine operations. The nine co-pilots had also written ATPL exams. Of the 21 pilots with an employment, seven were also employed as flight instructors. All pilots held a valid Class 1 Medical Certificate as well as a valid Proficiency Check (PC).

Based on the amount of flight hours, the participants were divided into three groups. Pilot Group 1 (novice pilots) consisted of 10 pilots (six male and four female). They were less experienced pilots with a mean age of 24 years ( $SD=3$ ) and with mean of 541 flight hours ( $SD=554$ ). In this group, two pilots were employed and eight were unemployed, two pilots were co-pilots and the remainder held no pilot rank since they had not been employed. Pilot Group 2 (semi-experienced pilots) consisted of 10 pilots (eight male and two female). They were experienced pilots with a mean age of 35 years ( $SD=3$ ) and with a mean of 5,835 flight hours ( $SD=2,393$ ). In this group, eight pilots were employed and two were unemployed, six pilots were captains and two were co-pilots. The two unemployed pilots had previously been employed as co-pilots. Pilot Group 3 (expert pilots) consisted of 12 pilots (10 male and two female) with a mean age of 48 years ( $SD=5.0$ ) and with a mean of 12,315 flight hours ( $SD=3,384$ ). In this



*Figure 7. Example: HUD 1, Position 1.*

group, 11 pilots were employed and one was unemployed; nine were captains and one was co-pilot. The unemployed pilot had previously been employed as a co-pilot.

The flight hour limits, which defines the three pilot groups, were based on aerospace industry common practice and on interviewing flight instructors, pilot recruiting officers and members of approved training organisations. Flight hours prerequisites were subsequently defined. For Pilot Group 1: flight hours for first employment, typically ranging between 250 and 1,500 flight hours. For Pilot Group 2: flight hours for the upgrade from co-pilot to captain, typically around 5,000 flight hours. For Pilot Group 3: extensive flight hour experience as captain, typically above 10,000 flight hours.

### *2.10.2. Apparatus*

At the test session each participant was assigned identical test equipment: a portable personal computer (HP Pavilion Entertainment PC TX 1320 [Windows Vista] with Microsoft PowerPoint 2010, Microsoft Word 2010 and Adobe Acrobat X Pro 2010). They received a pencil, paper and written test instructions. HUD user descriptions, a compilation of acronyms and explanations, and an airport navigation chart were also distributed. The tests were conducted in an office type environment including a desk, a computer, an office chair and a side table.

### *2.10.3. Experiment design and procedure*

At each test session, 60 min were assigned to an introduction. The participants were first asked to read the test instructions for 15 min followed by the test leader reading them aloud for 5 min. The participants were then asked to fill in a form regarding basic data such as age, sex and flying background for 10 min. User descriptions (Appendix Figures A2–A5) and acronyms/explanations (Appendix Table A2) for the HUDs were then studied for 30 min followed by a 5 min break.

The evaluation task was to assess simulated aircraft taxi supported by HUD 1–4, one at a time. A four-picture slide show (Figures 7–10) termed Taxi-run 1–4 showed four fixed positions (Position 1–4) of the apron and taxiways on Gothenburg Landvetter airport (ESGG/GOT), Sweden, for each HUD (Appendix Figure A6). Simulated outside conditions were daytime with good visibility. Each of the four HUDs was superimposed on a synthetic cockpit forward view of the airport infrastructure. The forward view was a combination of the respective HUD, including alpha-numeric characters/symbols (depicted in green monochromatic colour), and actual colour photographs from the airport infrastructure. A 12·1



Figure 8. Example: HUD 2, Position 2.



Figure 9. Example: HUD 3, Position 3.

in. (307 mm) PC colour screen placed on a desk with a display–eye distance of 500 mm and with a participant posture compliant with design eye position (DEP), head motion box and field of view (FOV) were used during the test sessions. In order to achieve visual clarity, the HUD’s elements were not collimated with the background; instead, the background was blurred at a 25% level.

The test was initiated by positioning the participant outside airport Gate 17 on the apron followed by the initiation of one of four taxi-runs. Each taxi-run had an identical route performed with HUD 1–4 randomly swapped to minimise the risk of learning effects. Each taxi-run had a duration of 4 min with the respective HUD displayed. Between each position, the PC screen was dark for 1 min as well as before and after the trial. The slideshow was automatically set to keep these time limits. Helton et al.



**Figure 10.** Example: HUD 4, Position 4.

**Table 1.** Summary of the self-assessment questions in the post-taxi-run questionnaire.

- 
- Q: I felt I had sufficient information when solving the task (1 Strongly Disagree – 6 Strongly Agree)
- Q: I did not feel that my mental workload was high when solving the task (1 Strongly Disagree – 6 Strongly Agree)
- Q: I felt that my situational awareness was good when solving the task (1 Strongly Disagree – 6 Strongly Agree)
- Q: I did not feel stress when I solved the task (1 Strongly Disagree – 6 Strongly Agree)
- Q: I felt that I had the support I needed when solving the task (1 Strongly Disagree – 6 Strongly Agree)
- Q: I felt that the display was useful to solve the task (1 Strongly Disagree – 6 Strongly Agree)
- 

(2007) and Teichner (1974) suggest that reduced vigilance may occur between the first 5 and 15 min in test situation. To reduce the risk of vigilance challenges during the taxi-runs, a 5 min break was included between each taxi-run. After each taxi-run a two-page questionnaire was completed. The questionnaire was designed to gather self-assessments on the six themes (information, workload, situational awareness, stress, support and usability) derived from the definition part of the study. A six-point ordinal Likert scale with self-assessments (strongly disagree, disagree, slightly disagree, slightly agree, agree and strongly agree) on each theme, was used (Table 1).

## 2.11. Evaluation part 2

### 2.11.1. Data analysis

A two-tailed Wilcoxon paired-sample test (Matlab R2015b, MathWorks BV) was therefore used to enable pair-wise comparisons on 6(Variables), 4(HUDs), and 3(Pilot Groups) for each of the four HUDs. Since several pair-wise comparisons were performed simultaneously (108 in total) on the single data set, a Bonferroni correction test was performed to reduce the risk of results incorrectly appearing

as statistically significant. The test showed that a significance level of 5% required a critical probability value of  $p < 0.00046$ . Due to the explorative design of the evaluation, this conservative  $p$ -value was omitted. The results are reported as differences and not as evidences. For this critical  $p$ -value a sample size of  $n=32$ , with the power set to 0.8, was considered adequate (Sandelowski, 1995). Stress and workload were reversed, to keep the questionnaire consistent, since a high self-assessment median equals an unfavourable outcome and a low self-assessment median equals a favourable outcome.

### 3. Results

The results are presented as differences in median values (M) between all HUDs for all pilots, as significant  $p$ -values, and as critical  $W$ -values (Table 2) as well as differences in median values between all pilots for each HUD (Figure 11). Section 3.1–3.6 below expands on the results.

#### 3.1. Information

Differences were found between HUD 1–3 for all pilots but when comparing HUD 3 with HUD 4 this was not the case. All pilots ranked the two graphical displays (HUD 3 and HUD 4) higher compared with the non-graphical HUD 1 and HUD 2. The graphical design as such did not make any difference. Displays providing predictive/descriptive information (HUD 2–4) were preferred over the descriptive only display (HUD 1). Displaying descriptive/predictive information in combination with graphical information was ranked high by all the pilots.

#### 3.2. Workload

Differences were found between HUD 1 and HUD 4 only for novice and semi-experienced pilots. A difference was also found between HUD 2 and HUD 4 but only for novice and expert pilots. All pilots ranked the two graphical displays (HUD 3 and HUD 4) higher compared with the non-graphical HUD 1 and HUD 2. Differences were also found between HUD 3 and HUD 4 but only for novice pilots. Expert pilots ranked the semi-circular graphical display lower than the vertical tape display. Displaying descriptive/predictive information in combination with graphical information was ranked high by all pilots.

#### 3.3. Situational awareness

Differences were found for novice and semi-experienced pilots between HUD 1 and HUD 3, for all pilots between HUD 1 and HUD 4, between HUD 2 and HUD 4, and between HUD 3 and HUD 4. No difference was found between HUD 1 and HUD 2 or between HUD 2 and HUD 3. All pilots ranked graphical displays higher compared with the non-graphical displays. Furthermore, expert pilots ranked HUD 1 and HUD 2 equally low. Displaying descriptive/predictive information in combination with graphical information was ranked high by all pilots.

#### 3.4. Stress

Differences were found for all pilots between HUD 1 and HUD 3, between HUD 1 and HUD 4, between HUD 2 and HUD 3, and between HUD 2 and 4. No difference was found between HUD 1 and HUD 2 and between HUD 3 and HUD 4. Again, the two graphical displays were ranked higher by all pilots but neither of the two displays was preferred when compared with each other. All pilots ranked HUD 1 and HUD 2 equally low. Displaying descriptive/predictive information in combination with graphical information was ranked high by all pilots.

**Table 2.** Summary of *W*-values, probabilities, and medians for a two-tailed Wilcoxon paired-sample test on self-assessment variables for HUD 1–4 and Pilot Group 1–3.

Variable	Pilot group	HUD1 versus HUD2	HUD1 versus HUD3	HUD1 versus HUD4	HUD2 versus HUD3	HUD2 versus HUD4	HUD3 versus HUD4
Information	1	$p = 0.008$	$p = 0.002$	$p = 0.002$	$p = 0.031$	$p = 0.016$	$p = 1.0$
		$W = 36$	$W = 55$	$W = 55$	$W = 34$	$W = 28$	$W = 16$
		$M = (2, 3)$	$M = (2, 4.5)$	$M = (2, 4.5)$	$M = (3, 4.5)$	$M = (3, 4.5)$	$M = (4.5, 4.5)$
	2	$p = 0.004$	$p = 0.002$	$p = 0.002$	$p = 0.023$	$p = 0.008$	$p = 1.0$
		$W = 45$	$W = 55$	$W = 55$	$W = 42$	$W = 36$	$W = 6$
		$M = (2, 3)$	$M = (2, 4.5)$	$M = (2, 4.5)$	$M = (3, 4.5)$	$M = (3, 4.5)$	$M = (4.5, 4.5)$
	3	$p = 0.014$	$p = 0.001$	$p = 0.004$	$p = 0.002$	$p = 0.014$	$p = 1.0$
		$W = 51, 5$	$W = 66$	$W = 45$	$W = 55$	$W = 51, 5$	$W = 20$
		$M = (2, 3)$	$M = (2, 4)$	$M = (2, 4)$	$M = (3, 4)$	$M = (3, 4)$	$M = (4, 4)$
Workload	1	$p = 0.20$	$p = 0.062$	$p = 0.004$	$p = 0.48$	$p = 0.040$	$p = 0.016$
		$W = 41.5$	$W = 15$	$W = 45$	$W = 19$	$W = 33.5$	$W = 28$
		$M = (3, 3)$	$M = (3, 4)$	$M = (3, 5)$	$M = (3, 4)$	$M = (3, 5)$	$M = (4, 5)$
	2	$p = 0.12$	$p = 0.16$	$p = 0.008$	$p = 0.33$	$p = 0.06$	$p = 0.53$
		$W = 30$	$W = 28$	$W = 36$	$W = 21$	$W = 32.5$	$W = 29$
		$M = (3, 3)$	$M = (3, 4)$	$M = (3, 5)$	$M = (3, 4)$	$M = (3, 5)$	$M = (4, 5)$
	3	$p = 0.38$	$p = 0.78$	$p = 0.14$	$p = 0.33$	$p = 0.008$	$p = 0.46$
		$W = 20, 5$	$W = 45, 5$	$W = 51$	$W = 21$	$W = 36$	$W = 35, 5$
		$M = (3, 3.5)$	$M = (3, 4)$	$M = (3, 5)$	$M = (3.5, 4)$	$M = (3.5, 5)$	$M = (4, 5)$
Situational awareness	1	$p = 0.17$	$p = 0.008$	$p = 0.002$	$p = 0.09$	$p = 0.002$	$p = 0.008$
		$W = 23$	$W = 36$	$W = 55$	$W = 25$	$W = 55$	$W = 36$
		$M = (2, 3)$	$M = (2, 4)$	$M = (2, 5)$	$M = (3, 4)$	$M = (3, 5)$	$M = (4, 5)$
	2	$p = 0.06$	$p = 0.02$	$p = 0.002$	$p = 0.16$	$p = 0.004$	$p = 0.05$
		$W = 26$	$W = 28$	$W = 55$	$W = 18, 5$	$W = 45$	$W = 26, 5$
		$M = (2, 3.5)$	$M = (2, 4)$	$M = (2, 5)$	$M = (3.5, 4)$	$M = (3.5, 5)$	$M = (4, 5)$
	3	$p = 1.0$	$p = 0.41$	$p = 0.0001$	$p = 0.19$	$p = 0.001$	$p = 0.008$
		$W = 19$	$W = 30.5$	$W = 78$	$W = 22$	$W = 66$	$W = 36$
		$M = (3.5, 3.5)$	$M = (3.5, 4)$	$M = (3.5, 5.5)$	$M = (3.5, 4)$	$M = (3.5, 5.5)$	$M = (4, 5.5)$

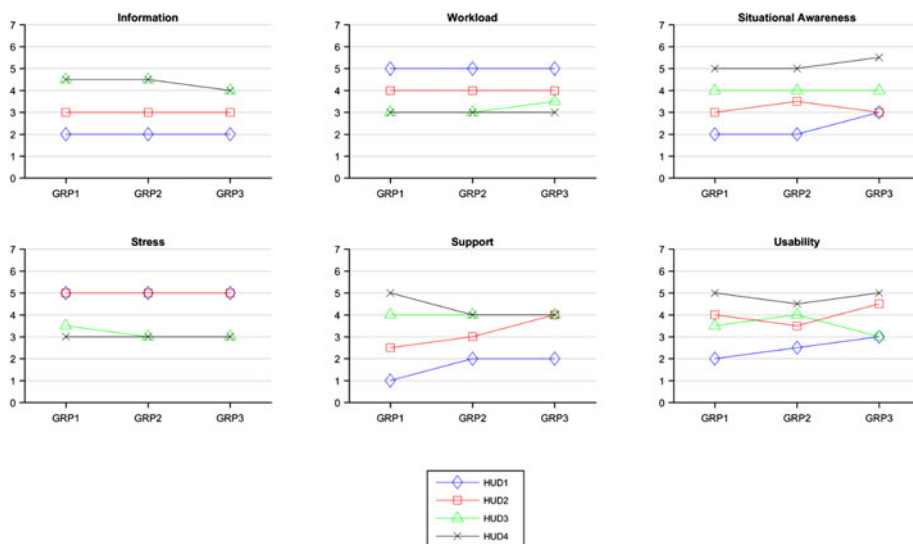
*Continued.*

**Table 2.** Continued.

Variable	Pilot group	HUD1 versus HUD2	HUD1 versus HUD3	HUD1 versus HUD4	HUD2 versus HUD3	HUD2 versus HUD4	HUD3 versus HUD4
Stress	1	$p = 1.0$	$p = 0.004$	$p = 0.008$	$p = 0.004$	$p = 0.002$	$p = 1.0$
		$W = 10, 5$	$W = 45$	$W = 53$	$W = 45$	$W = 55$	$W = 20$
		$M = (3, 3.5)$	$M = (3, 5)$	$M = (3, 5)$	$M = (3.5, 5)$	$M = (3.5, 5)$	$M = (5, 5)$
	2	$p = 0.78$	$p = 0.002$	$p = 0.002$	$p = 0.012$	$p = 0.008$	$p = 1.0$
		$W = 27$	$W = 55$	$W = 55$	$W = 43, 5$	$W = 53$	$W = 12$
		$M = (3, 3)$	$M = (3, 5)$	$M = (3, 5)$	$M = (3, 5)$	$M = (3, 5)$	$M = (5, 5)$
	3	$p = 0.79$	$p = 0.0001$	$p = 0.001$	$p = 0.005$	$p = 0.0001$	$p = 0.62$
		$W = 21, 5$	$W = 78$	$W = 66$	$W = 63, 5$	$W = 78$	$W = 10, 5$
		$M = (3, 3)$	$M = (3, 5)$	$M = (3, 5)$	$M = (3, 5)$	$M = (3, 5)$	$M = (5, 5)$
Support	1	$p = 0.004$	$p = 0.002$	$p = 0.002$	$p = 0.004$	$p = 0.002$	$p = 0.19$
		$W = 45$	$W = 55$	$W = 55$	$W = 45$	$W = 55$	$W = 18$
		$M = (1, 2.5)$	$M = (1, 4)$	$M = (1, 5)$	$M = (2.5, 4)$	$M = (2.5, 5)$	$M = (4, 5)$
	2	$p = 0.04$	$p = 0.004$	$p = 0.016$	$p = 0.062$	$p = 0.25$	$p = 0.50$
		$W = 33.5$	$W = 45$	$W = 28$	$W = 26$	$W = 22$	$W = 4$
		$M = (2, 3)$	$M = (2, 4)$	$M = (2, 4)$	$M = (3, 4)$	$M = (3, 4)$	$M = (4, 4)$
	3	$p = 0.002$	$p = 0.002$	$p = 0.0001$	$p = 1.0$	$p = 0.56$	$p = 0.44$
		$W = 55$	$W = 55$	$W = 78$	$W = 25$	$W = 24$	$W = 11$
		$M = (2, 4)$	$M = (2, 4)$	$M = (2, 4)$	$M = (4, 4)$	$M = (4, 4)$	$M = (4, 4)$
Usability	1	$p = 0.004$	$p = 0.004$	$p = 0.002$	$p = 1.0$	$p = 0.008$	$p = 0.016$
		$W = 45$	$W = 45$	$W = 55$	$W = 2$	$W = 36$	$W = 28$
		$M = (2, 3.5)$	$M = (2, 4)$	$M = (2, 5)$	$M = (3.5, 4)$	$M = (3.5, 5)$	$M = (4, 5)$
	2	$p = 0.016$	$p = 0.004$	$p = 0.002$	$p = 0.25$	$p = 0.008$	$p = 0.094$
		$W = 28$	$W = 45$	$W = 55$	$W = 6$	$W = 36$	$W = 25$
		$M = (2.5, 3.5)$	$M = (2.5, 4)$	$M = (2.5, 4.5)$	$M = (3.5, 4)$	$M = (3.5, 4.5)$	$M = (4, 4.5)$
	3	$p = 0.008$	$p = 0.98$	$p = 0.002$	$p = 0.002$	$p = 0.008$	$p = 0.0001$
		$W = 36$	$W = 12.5$	$W = 55$	$W = 0$	$W = 36$	$W = 78$
		$M = (3, 3)$	$M = (3, 4.5)$	$M = (3, 5)$	$M = (3, 4.5)$	$M = (3, 5)$	$M = (4.5, 5)$

Note.  $p$ -values in Table 2 were not adjusted for multiple pair-wise comparisons. A critical probability value of  $p < 0.00046$  was calculated using a Bonferroni correction test. Due to the explorative design of the evaluation this moderate  $p$ -value was omitted. A conservative 5% level of significance ( $p < 0.05$ ) was consequently selected for the interpretation of the results. However, as can be seen in the table above, several  $W$ -values on the 1% level of significance ( $p < 0.01$ ) are present. HUD 1 = Descriptive alpha-numerical; HUD 2 = Descriptive, predictive, alpha-numerical; HUD 3 = Descriptive, predictive, alpha-numerical, graphical, semi-circular; HUD 4 = Descriptive, predictive, alpha-numerical, graphical, vertical tape.





**Figure 11.** Self-assessment medians for Pilot Group 1–3 on HUD 1–4. Note: For variables Workload and Stress, a lower median value is favourable.

### 3.5. Support

Differences were found for all pilots between HUD 1 and HUD 2, between HUD 1 and HUD 3, and between HUD 1 and HUD 4. A difference was found only for novice pilots between HUD 2 and HUD 3 and between HUD 2 and HUD 4. Again, the two graphical displays were ranked higher by all pilots but neither of the two displays was preferred when compared with each other. Expert pilots ranked HUD 2, HUD 3 and HUD 4 equally high.

### 3.6. Usability

Differences were found for all pilots between HUD 1 and HUD 2, between HUD 1 and HUD 4, and between HUD 2 and HUD 4. Differences were also found between HUD1 and HUD 3 but only for novice and semi-experienced pilots, between HUD 2 and HUD 3 only for expert pilots, and between HUD3 and HUD 4 for novice and expert pilots. The graphical vertical tape display (HUD 4) was ranked higher by all pilots. The descriptive/predictive only display (HUD 2) was ranked higher than HUD 3 by novice and expert pilots but not by semi-experienced pilots who ranked HUD 3 higher than HUD 2.

## 4. Discussion

In the present evaluation the methods of evaluating HUD 1–4 were based on subjective self-assessments from six-point Likert scale questionnaires and interviews, all adjusted to match the simulated taxi-runs. The purpose of the study was to evaluate *which* information provided the best support by collecting self-assessments on six variables. Based on the prerequisites of the evaluation (non-flight simulator, self-assessments), inferential interpretations must be used cautiously, especially for workload, stress and situational awareness. Statistical significance for these variables shall therefore be regarded more as an indication.

When assessing workload, no specific factor is a reliable measurement. Therefore, a SWAT (subjective workload assessment technique) is useful since it contains a multidimensional scale (Wilson and Corlett, 1995). The three self-assessment steps (minimal, average, maximal) in the SWAT scale constitute a risk of selecting the middle level, thus avoiding an active choice (Karlsson et al., 2009). A six-point (even) Likert scale however rules out the average option thus forcing an active choice (Karlsson et al., 2009). In addition to the SWAT, the NASA-TLX (NASA Task Load Index) method assessing cognitive and

physiological demands can be used (Hart and Staveland, 1988). NASA-TLX is similar to SWAT but also measures physiological dimensions, a variable not applicable for the present evaluation. Both methods require extensive pre-test preparations as well as concluding post-test efforts, for researchers as well as for participants. The resources for the study at the time did not match these required efforts.

Stress can be measured by heart rate (Regula et al., 2014), perspiration (Torrente-Rodríguez et al., 2020), finger temperature (Oka et al., 2001) or cortisol levels (Selye, 1950), however, these require laboratory equipment, surveillance and means of evaluating the observations.

Situational awareness can be assessed by freeze probe techniques (Endsley, 1995) using e.g., SAGAT (Situation Awareness Global Assessment Technique) (Endsley, 1996) or self-rating tools such as SART (Situation Awareness Rating Tool) (Taylor, 1990). SAGAT uses objective measurements, where system status is compared with operators' assessments of the situation, while SART uses subjective measurements. Both tools are commonly used in simulator trials. However, since SAGAT requires freezes (where the displays go blank) in the simulation and the present evaluation consisted of freezes *only* it was deemed not suitable. SART collects operator assessments, usually post-trial. Collecting assessments post-trial can have adverse effects on these assessments due to poor recall or confounding with other variables such as workload (Salmon et al., 2009).

Information and support can be measured by completion rates, e.g., where effectiveness, efficiency and satisfaction is assessed by fail/pass compared with accomplishing set goals (ISO, 2016). The present evaluation contained no set goals in terms of accomplishing a task but only to assess if the HUDs were perceived useful as such.

Usability can be assessed by, e.g., task success rate (Nielsen, 2001), which relates the number of successful attempts with the number of attempts, or by task completion time using an analysis of variance between task completion times (Tullis and Albert, 2008). Usability was therefore evaluated by self-assessments and not objective measurements in the present evaluation.

The detection and the discrimination of displayed information can be improved by using different shapes, colours and sizes (Howarth and Bullimore, 2005). When assessing the two HUDs with combined graphical and alpha-numerical information, the pilots' self-assessments for stress, workload, situational awareness and support were positive. The HUDs' presentation of redundant information can explain this result (Howarth and Bullimore, 2005). However, for some pilots this was not the case, since stress and workload increased. This can be caused by differences in experience, training and perception, and is an indication that individual pilot performance may differ from crew performance. In fact, the concept behind a multi-crew is based on the combined performance of the crew and not on the individual pilot. Furthermore, it was found that displaying both descriptive as well as predictive information was a preferred condition of display properties. This finding is consistent with Meister (1985), who suggested that such properties can reduce cognitive workload. Lewis et al. (1990) also suggest that redundancy as well as a decreased need for vigilance is a result of such a presentation. Redundancy may explain the good ratings for HUD 3 and 4 since they contained descriptive, predictive and alpha-numerical information combined with graphical information.

The four HUDs were designed and evaluated from an expert-novice perspective. The reason for this is the ever-changing composition of cockpit crew, e.g., pilots with different expertise, variations in cognitive and sensory motor abilities, age and daily form must (at the same time) be able to interact with the proposed HUDs. The outcomes from this study also indicate that differences between novice and expert pilots can be resolved by incorporating combinations of graphics, alpha-numeric presentations and intuitive design based on Gestalt laws (Wertheimer, 1923/1938). Such combinations can cater for more experienced pilots who may use heuristics to a higher extent than novice pilots, who may use calculation as a strategy when interacting with a system. This is in line with research performed by Rasmussen (1983, 1985) and the Skills, Rules and Knowledge (SRK) framework.

Experience may however be expressed in several ways. Pilot flight time is a more exact way to estimate flying experience compared with number of years of employment or age. Since pilots must log all flying activities they undertake, using flight time is therefore an easy way to enable comparisons on pilot-experience. Flight time experience has also been consistently found to be a relevant factor in

epidemiologic studies of aircraft accident rates (Taylor et al., 2007) as well as for recruiting purposes. However, flight time as a measurement needs to be further dissected since different flight operations (e.g., high performance jet operations, crop dusting, commercial, military, multi-pilot) have implications for what a flight hour actually contains.

The study indicated that a semi-circular display, compared with a speed-tape design, reduced workload to a greater extent for novice pilots. This can be explained by the previous experience novice pilots have from the beginning of their training in using such gauges, displays and dials.

The lower (better) rating on stress for HUD 3 and 4 is valuable since one of the main advantages of HUDs is increased situational awareness (Foyle et al., 1996; Proctor, 1999, 2000). The three variables of stress, information and situational awareness are related since information increases situational awareness and lack of (or abundant) information may decrease situational awareness due to cognitive limitations. Such a situation is disadvantageous and may increase the amount of stress in the cockpit (Wiener, 1988, 1989; Sarter and Woods, 1992).

Furthermore, the HUDs developed in this study are per se limited to being an interface for cockpit use. If stakeholders such as pilots and ATC share information by using joint systems, e.g., SMS, such information could be used to present predictions to all pilots and ATC at the same time. Such integrated systems can provide means for improved planning and situational awareness as well as reduced fuel consumption. Additionally, such a system can offer ATC means to manage the traffic flow simply by assigning target speeds direct to taxiing aircraft.

One way of assessing the proposed HUD is to regard it as a demonstrator of a more efficient way of conducting taxi under time limitations. Future studies could include evaluations of task success rate between target speed HUD presentations and non-target speed presentations. Furthermore, the use of HUDs for navigational support (Foyle et al., 1996; Cheng et al., 2009) has been assessed and incorporating a target speed presentation as well may present means of resolving both navigation and speed during taxi. Future applications of the HUD also need to meet further requirements since the HUD may similarly be used as flight guidance support. That is, during taxi, the pilot at some point needs to switch the HUD from ground mode to flight mode. An additional application may provide possibilities for the individual pilot to customise the HUD for their personal requirements and techniques.

## 5. Conclusion

The ground segment (taxi to departure runway) of a 4D flight trajectory is generally controlled by set time constraints. This study aimed to design and evaluate information on a display, to improve time-keeping. Conclusions from focus group interviews with pilots as well as from prevailing taxi procedures show that pilots generally perform rule-of-thumb estimations built on distance-time-speed calculations when managing taxi speeds. Based on the pilots' self-assessments on six variables, a graphical vertical tape was proposed with descriptive/predictive HUD presentation including alpha-numerical information on target taxi speeds. Graphical information attracted the participating pilots in the usability test more than alpha-numerical information only. Results from the study also indicated that differences between novice, semi-experienced and expert pilots can be balanced by displaying a combination of alpha-numerical and graphical information, thus providing redundancy. The results from this study could be further elaborated in future research in a flight simulator. In order to derive improved inferential conclusions, self-assessments could be complemented by measurements such as NASA-TLX and SAGAT. Furthermore, evaluating navigational taxi support in combination with taxi target speed support could also be beneficial.

## References

- Academio. (2021). *Wavelength to colour relationship calculator*. <https://academo.org/demos/wavelength-to-colour-relationship/>. Accessed 1 May 2021.
- Annett, J. and Duncan, K. D. (1967). Task analysis and training design. *Occupational Psychology*, **41**, 211–221.

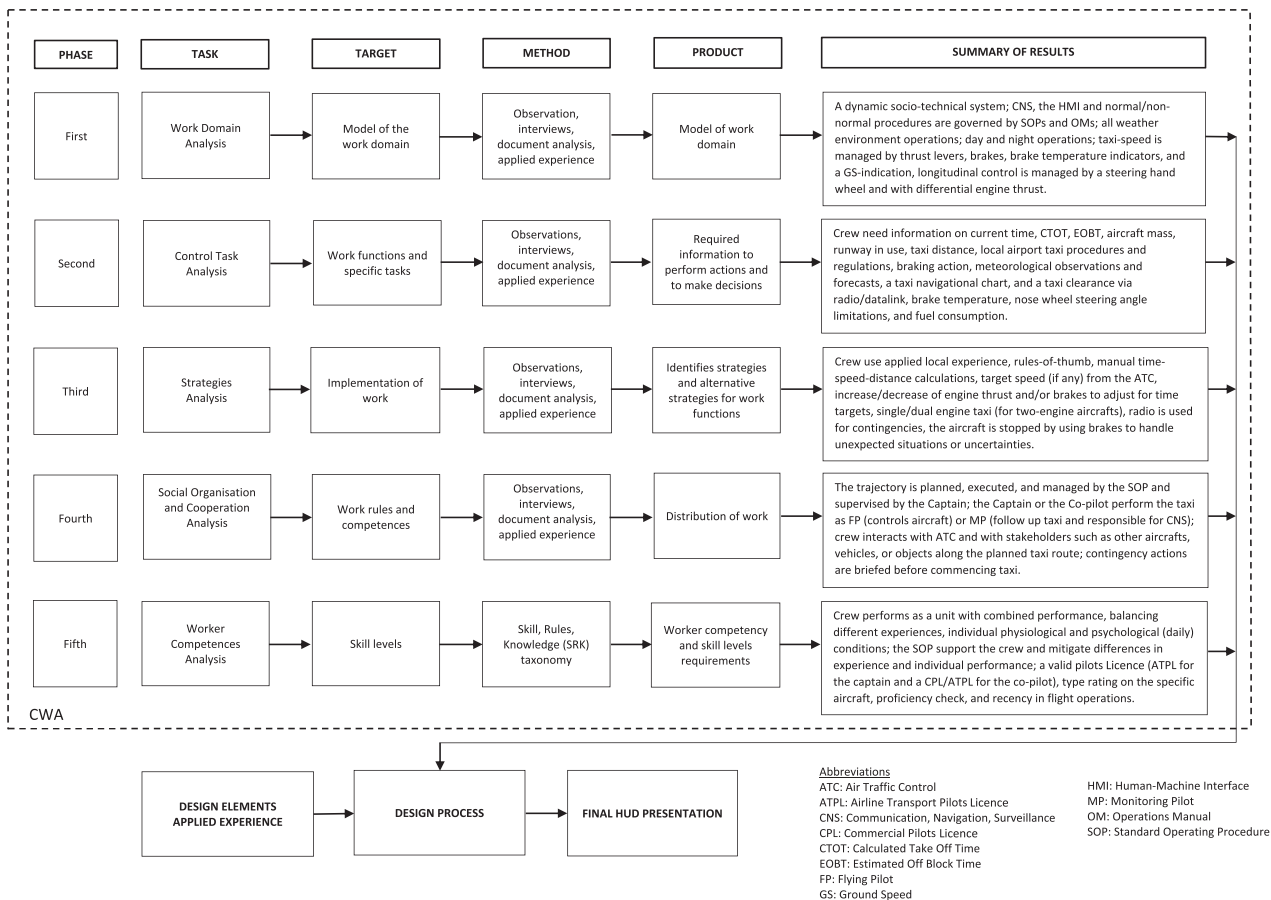
- Ashoori, M., Burns, C. M., d'Entremont, B. and Momtahan, K. (2014). Using team cognitive work analysis to reveal healthcare team interactions in a birthing unit. *Ergonomics*, **57**(7), 973–986. doi:10.1080/00140139.2014.909949
- Atkin, J. A. D., Burke, E. K. and Ravizza, S. (2011). A More Realistic Approach for Airport Ground Movement Optimisation with Stand Holding. *Proceedings of the 5th Multidisciplinary International Scheduling Conference (MISTA 2011)*, Phoenix, AZ.
- Atkins, S. and Brinton, C. (2002). Concept description and development plan for the surface management system. *Journal of Air Traffic Control*, **44**(1), 1–8.
- Balakrishnan, H. and Jung, Y. (2007). A Framework for Coordinated Surface Operations Planning at Dallas-Fort Worth International Airport. *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Hilton Head, SC.
- Benyon, D., Turner, P. and Turner, S. (2005). *Designing Interactive Systems: People, Activities, Contexts, Technologies*. New York: Pearson Education.
- Bisantz, A. M., Stone, R. T., Pfautz, J., Fouse, A., Farry, M., Roth, E. M., Nagy, A. and Thomas, G. (2009). Visual representations of meta-information. *Journal of Cognitive Engineering and Decision Making*, **3**(1), 67–91. doi:10.1518/155534309X433726
- Brady, A., Naikar, N. and Treadwell, A. (2013). Organisational Storytelling with Work Domain Analysis: Case Study of Air Power Doctrine and Strategy Narrative. In: Ahamed, S., Piantadosi, J., Agrawal, M. and Boland, J. (eds.). *MODSIM 2013, 20th International Congress on Modelling and Simulation*, Adelaide, Australia, 1138–1144.
- Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, **3**(2), 77–101.
- Brickner, M. S. and Foyle, D. C. (1990). Field of View Effects on a Simulated Flight Task with Head-Down and Head-up Sensor Imagery Displays. *Proceedings of the Human Factors Society 34th Annual Meeting*, Santa Monica, CA, 1567–1571. doi:10.1177/154193129003401931.
- Cheng, V. H. L., Andre, A. D. and Foyle, D. C. (2009). Information Requirements for Pilots to Execute 4D Trajectories on the Airport Surface. *Proceedings of the 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)*. Reston, VA, 1–2.
- Collins Aerospace. (2021). *HGS 6000 with enhanced flight vision system for the Boeing 737*. [https://www.collinsaerospace.com/-/media/project/collinsaerospace/collinsaerospace-website/product-assets/marketing/h/head-up-display/collins\\_evs-3600-hgs-6000\\_4pg\\_bro\\_final-web.pdf?rev={mathsurround=\vskip\\$=\\$}9088c989e09646be869197d152eb705c](https://www.collinsaerospace.com/-/media/project/collinsaerospace/collinsaerospace-website/product-assets/marketing/h/head-up-display/collins_evs-3600-hgs-6000_4pg_bro_final-web.pdf?rev={mathsurround=\vskip$=$}9088c989e09646be869197d152eb705c). Assessed 1 May 2021.
- Couluris, G. J., Fong, R. K., Downs, M. B., Mittler, N., Signor, D., Stassart, A. and Hsiao, T. (2008). A New Modeling Capability for Airport Surface Traffic Analysis. *2008 IEEE/AIAA 27th Digital Avionics Systems Conference*. St. Paul, MN. doi:10.1109/DASC.2008.4702818.
- Craig, C. (2012). Improving flight condition situational awareness through human centered design. *Work*, **41**(3), 4523–4531. doi:10.3233/WOR-2012-0031-4523
- Dashevsky, S. G. (1964). Check-reading accuracy as a function of pointer alignment, patterning, and viewing angle. *Journal of Applied Psychology*, **48**, 344–347.
- de Winter, J. C. F. and Dodou, D. (2014). Why the Fitts list has persisted throughout the history of function allocation. *Cognition, Technology & Work*, **16**, 1–11.
- Eclipse. (2021). *B612 font family*. <https://projects.eclipse.org/working-group/polarsys>. Assessed 1 May 2021.
- Effken, J. A., Brewer, B. B., Logue, M. D., Gephart, S. M. and Verran, J. A. (2011). Using cognitive work analysis to fit decision support tools to nurse managers' workflow. *International Journal of Medical Informatics*, **80**(10), 698–707.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, **37**(1), 32–64.
- Endsley, M. R. (1996). Automation and situation awareness. In: Parasuraman, R. and Mouloua, M. (eds.). *Automation and Human Performance: Theory and Applications*, Mahwah, NJ: Erlbaum, 163–181.
- Endsley, M. R. (2015). Situation awareness: Operationally necessary and scientifically grounded. *Cognition, Technology and Work*, **17**, 163–167. doi:10.1007/s10111-015-0323-5
- ESSE. (2021). Color contrast checker & ratio calculator. <https://contrastchecker.online/>. Assessed 23 March 2021.
- EUROCONTROL. (2017). *Airport collaborative decision making (A-CDM)*. <https://www.eurocontrol.int/concept/airport-collaborative-decision-making>. Assessed 10 September 2020.
- EUROCONTROL. (2020). *Daily traffic variation – states 2020*. <https://www.eurocontrol.int/Economics/DailyTrafficVariation-States.html>. Assessed 10 September 2020.
- Fadden, S., Ververs, P. M. and Wickens, C. D. (1998). Costs and Benefits of Head-up Display Use: A Metanalytic Approach. *Proceedings of the Human Factors and Ergonomics Society*. Santa Monica, CA, 16–20.
- Foyle, D. C., Sanford, B. D. and McCann, R. S. (1991). Attentional Issues in Superimposed Flight Symbolology. In: Jensen, R. S. (ed.). *Proceedings of the Sixth International Symposium on Aviation Psychology*, Columbus: Ohio State Press, 577–582.
- Foyle, D. C., McCann, R. S. and Shelden, S. G. (1995). Attentional Issues with Superimposed Symbolology: Formats for Scenelinked Displays. *Proceedings of the 8th International Symposium on Aviation Psychology*, Columbus, OH, 98–103.
- Foyle, D. C., Andre, A. D., McCann, R. S., Wenzel, E., Begault, D. and Battiste, V. (1996). Taxiway navigation and situation awareness (T-NASA) system. Problem, design, philosophy, and description of an integrated display suite for low visibility airport surface operations. *SAE Transactions: Journal of Aerospace*, **105**, 1511–1418.
- Gish, K. W. and Staplin, L. (1995). *Human factors aspects of using head-up displays in automobiles: A review of the literature (Report DOT HS 808 320)*. Washington, DC: U.S. Department of Transportation, Federal Highway Administration.
- Gotteland, J. B., Durand, N. and Alliot, J. M. (2003). Handling CFMU Slots in Busy Airports. *Proceedings of the 5th USA/Europe Air Traffic Management R & D Seminar*. Budapest, Hungary.

- Green, R. G., Muir, H., James, M., Gradwell, D. and Green, R. L. (2001). *Human factors for pilots*. (2nd rev. ed.), Aldershot: Ashgate Publishing Limited, 113–128.
- Green, J. E. (2011). The potential for reducing the impact of aviation on climate. *Technology Analysis and Strategic Management*, 21(1), 39–59.
- Grote, M., Williams, I. and Preston, J. (2014). Direct carbon dioxide emissions from civil aircraft. *Atmospheric Environment*, 95, 214–224.
- Harris, D. (2011). *Human Performance on the Flight Deck*. Boca Raton, FL: CRC Press.
- Hart, S. G. and Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In: Hancock, P. A. and Meshkati, N. (eds.). *Human Mental Workload*, Amsterdam: North Holland Press, 139–183.
- Hawkins, F. H. (1987). *Human Factors in Flight*. 2nd rev. edition, Burlington, VT: Ashgate Publishing Company.
- Hecker, P., Doehler, H. U., Korn, B. and Ludwig, T. (2001). Extending Enhanced Vision Capabilities by Integration of Advanced Surface Movement Guidance and Control Systems (A-SMGCS). *Proceedings of SPIE, the International Society for Optical Engineering*, Orlando, FL, 58–66. doi:10.1117/12.438033.
- Helander, M. G. (2006). *A Guide to Human Factors and Ergonomics*. 2nd rev. edition, Boca Raton, FL: CRC Press.
- Helton, W. S., Hollander, T. D., Warm, J. S., Tripp, L. D., Parsons, K. S., Matthews, G., Dember, W. N., Parasuraman, R. and Hancock, P. A. (2007). The abbreviated vigilance task and cerebral hemodynamics. *Journal of Clinical and Experimental Neuropsychology*, 29, 545–552.
- Hirsch, R. (2004). *Exploring Colour Photography: A Complete Guide*. London: Laurence King Publishing. ISBN 1-85669-420-8.
- Holder, E. and Pecota, S. R. (2011). Maritime head-up display: A preliminary evaluation. *Journal of Navigation*, 64(4), 573–594.
- Howarth, P. A. and Bullimore, M. A. (2005). Vision and visual work. In: Wilson, J. R. and Corlett, N. (eds.). *Evaluation of Human Work*. 3rd rev. edition, Boca Raton, FL: CRC Press, 337–338.
- ICAO. (2007). International Civil Aviation Organization. *PANS-ATM Procedures for Air Navigation Services—Air Traffic Management* [Doc 4444], 15th ed. (Amdt 4, 15/11/2012).
- ISO. (2016). International Organization for Standardization. *ISO/IEC 25022:2016*. <https://www.iso.org/standard/35746.html>. Assessed 9 April 2021.
- Jenkins, D. P., Stanton, N. A., Salmon, P. M., Walker, G. H. and Young, M. S. (2008). Using cognitive work analysis to explore activity allocation within military domains. *Ergonomics*, 51(6), 798–815. doi:10.1080/00140130801915246
- Karlsson, S., Osvalder, A. L., Rose, L., Eklund, J. and Odenrick, P. (2009). Design processes. Chapter 10. In: Bohgard, M., Karlsson, S., Lovén, E., Mikaelsson, L.-Å., Mårtensson, L., Osvalder, A.-L., Rose, L. and Ulvengren, P. (eds.). *Work and Technology on Human Terms*, Stockholm: Prent, 579.
- Kragt, H. (1992). *Enhancing Industrial Performance*. London: Taylor & Francis.
- Krueger, R. A. and Casey, M. A. (2000). *Focus Groups: A Practical Guide for Applied Research*. Thousand Oaks, CA: SAGE Publications Inc., 21–33.
- Lawson, B. R. (2006). *How Designers Think: The Design Process Demystified*. Oxford: Architectural Press, 33–49.
- Lazarus, R. S. and Folkman, S. (1984). *Stress, Appraisal, and Coping*. New York, NY: Springer.
- Lee, H., Simaiakis, I. and Balakrishnan, H. (2010). A Comparison of Aircraft Trajectory-Based and Aggregate Queue-Based Control of Airport Taxi Processes. *Proceedings of the 29th Digital Avionics Systems Conference*, Salt Lake City, UT, 1.B.3-1-1.B.3-15.
- Lewis, R. F., Kelland, D. Z. and Kupke, T. (1990). A normative study of the repeatable cognitive-perceptual-motor-battery. *Archives of Clinical Neuropsychology*, 5, 201–214.
- Linell, P. (1994). *Transkription av tal och samtal: Teori och praktik (Transcription of Talk and Conversation: Theory and Practice)*, Arbetsrapport från Tema K, Rapport nr 1994:9 (Working Reports from Theme-K, No. 1994:9), Linköping: Linköping University.
- Marín, A. G. (2006). Airport management: Taxi planning. *Annals of Operations Research*, 143(1), 191–202.
- Markus, T. A. (1969). The role of building performance and appraisal in design method. In: Broadbent, G and Ward, A. (eds.), *Design Methods in Architecture*, London: Lund Humphries, 109.
- Martin-Emerson, R. and Wickens, C. D. (1997). Superimposition, symbolism, visual attention, and the head-up display. *Human Factors*, 39, 581–601.
- Maver, T. W. (1970). Appraisal in the building design process. In: Moore, G. T. (ed.). *Emerging Methods in Environmental Design and Planning*, Cambridge, MA: MIT Press, 196.
- Megaw, T. (2005). The definition and measurement of mental workload. In: Wilson, J. R. and Corlett, N. (eds.). *Evaluation of Human Work*. 3rd edition, Boca Raton, FL: CRC Press, 337–338.
- Meister, D. (1985). *Behavioral Analysis and Measurement Methods*. Chichester: J. Wiley and Sons.
- Merritt, A. C. and Helmreich, R. L. (1996). Human factors on the flight deck: The influence of national culture. *Journal of Cross-Cultural Psychology*, 27(5), 5–24.
- Moray, N. (1996). A Taxonomy and Theory of Mental Models. *Proceedings of the Human Factors Ergonomics Society 40th Annual Meeting*, Santa Monica, CA, 164–168.
- Morrison, J. G., Kelly, R. T., Moore, R. A. and Hutchins, S. G. (1998). Implications of decision making research for decision support and displays. In: Cannon-Bowers, J. A. and Salas, E. (eds.). *Making Decision Under Stress: Implications for Individual and Team Training*, Washington, DC: American Psychological Association Press, 375–406.

- Nagaoka, S., Muto, P. and Yoshioka, E. (1978). Estimating the average taxi speed of departing aircraft. *Journal of Navigation*, **31**(3), 494–499. doi:10.1017/S0373463300042016
- Newman, R. L. (1987). Improvement of Head-up Display Standards (Vol 1): Head-Up Display Design Guide (Appendix). Report prepared for the Air Force Wright Aeronautical Laboratories Flight Dynamics Laboratory. (AFWAL-TR-87-3055).
- Newman, R. L. (1995). *Head-up Displays: Designing the Way Ahead*. Aldershot: Ashgate.
- Nielsen, J. (2001). *Success rate: the simplest usability metric*. <https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/>. Assessed 1 March 2021.
- Nikoleris, T., Gupta, G. and Kistler, M. (2011). Detailed estimation of fuel consumption and emissions during aircraft taxi operations at Dallas/Fort Worth International Airport. *Transportation Research Part D: Transport and Environment*, **16**(4), 302–308. doi:10.1016/j.trd.2011.01.007
- Oka, T., Oka, K. and Hori, T. (2001). Mechanisms and mediators of psychological stress-induced rise in core temperature. *Psychosomatic Medicine*, **63**(3), 476–486. doi:10.1097/00006842-200105000-00018
- Osva, A. L., Rose, L. and Karlsson, S. (2009). Methods. Chapter 9. In: Bohgard, M., Karlsson, S., Lovén, E., Mikaelsson, L.-Å., Mårtensson, L., Osva, A.-L., Rose, L. and Ulvengren, P. (eds.). *Work and Technology on Human Terms*, Stockholm: Prevent, 477.
- Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology*, **24**, 436–477.
- Palmer, M. T. (1995). Interpersonal communication and virtual reality: Mediating interpersonal relationships. In: Biocca, F. and Levy, M. (eds.). *Communication in the Age of Virtual Reality*, Hillsdale, NJ: Lawrence Erlbaum, 277–299.
- Proctor, P. W. (1999). Enhanced head-up symbology builds situational awareness. *Aviation Week & Space Technology*, April **19**, 64–65.
- Proctor, P. W. (2000). New head-up tool aims to cut runway incidents. *Aviation Week & Space Technology*, August **14**, 48–50.
- Proctor, R. W. and Vu, K. P. L. (2016). Principles for designing interfaces compatible with human information processing. *International Journal of Human-Computer Interaction*, **32**, 2–22. doi:10.1080/10447318.2016.1105009
- Rasmussen, J. (1983). Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, **13**(3), 257–266. doi:10.1109/TSMC.1983.6313160
- Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on Systems, Man and Cybernetics*, **SMC-15**(2), 234–243. doi:10.1109/TSMC.1985.6313353
- Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. New York, NY: Elsevier Science Ltd.
- Rasmussen, J., Pejtersen, A. M. and Goodstein, L. P. (1994). *Cognitive Systems Engineering*. Hoboken, NJ: John Wiley & Son, Inc.
- Ravizza, S., Chen, J., Atkin, J. A., Burke, E. K. and Stewart, P. (2013). The trade-off between taxi time and fuel consumption in airport ground movement. *Public Transport*, **5**(1–2), 25–40.
- Regula, M., Socha, V., Kutlík, P., Socha, L., Hána, K., Hanáková, L. and Szabo, S. (2014). Study of Heart Rate as the Main Stress Indicator in Aircraft Pilots. *Proceedings of the 16th International Conference on Mechatronics, Mechatronika 2014*, 639–643.
- Roth, E. M. and Bisantz, A. M. (2013). Cognitive work analysis. In: Kirlik, A. and Lee, J. (eds.). *Handbook of Cognitive Engineering*. New York: Oxford University Press, 240–260.
- Rubenstein, R., Hersh, H. M. and Ledgard, H. F. (1984). *The Human Factor: Designing Computer Systems for People*. Burlington, MA: Digital Press.
- Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D. P., Ladva, D., Rafferty, L. and Young, M. S. (2009). Measuring situation awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, **39**, 490–500.
- Salmon, P. M., Jenkins, D., Stanton, N. and Walker, G. (2010). Hierarchical task analysis vs. cognitive work analysis: Comparison of theory, methodology and contribution to system design. *Theoretical Issues in Ergonomics Science*, **11**, 504–531.
- Sandelowski, M. (1995). Focus on qualitative methods. Sample size in qualitative research. *Research in Nursing & Health*, **18**(2), 179–183. doi:10.1002/nur.4770180211
- Sarter, N. B. and Woods, D. D. (1992). Pilot interaction with cockpit automation: Operational experiences with the flight management system. *International Journal of Aviation Psychology*, **2**(4), 303–321.
- Selye, H. (1950). *Stress: The Physiology and Pathology of Exposure to Stress*. Montreal, QC: Acta Medica Publication.
- SESAR. (2015). SESAR WPE TREE Demonstration Session. *Eleventh USA/Europe Air Traffic Management Research and Development Seminar (ATM2015)*. Palma de Mallorca, Spain.
- SESAR. (2021). HP Repository. [https://ext.eurocontrol.int/ehp/?q{\mathsurround=\lopskip\\$=}node/1615](https://ext.eurocontrol.int/ehp/?q{\mathsurround=\lopskip$=}node/1615). Assessed 23 March 2021.
- Sexton, J. B. and Helmreich, R. L. (2003). Using language in the cockpit: Relationships with workload and performance. In: Dietrich, R. (ed.). *Communication in High Risk Environments*, Hamburg: Helmut Buske Verlag GmbH, 57–74.
- Sexton, J. B., Thomas, E. J. and Helmreich, R. L. (2000). Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. *BMJ*, **320**, 745–749. doi:10.1186/cc2872
- Shackel, B. (1981). The Concept of Usability. *Proceedings of IBM Software and Information Usability Symposium*, Poughkeepsie, NY, 1–30.

- Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A. and Wiegmann, D.A. (2007). Human error and commercial aviation accidents: an analysis using the human factors analysis and classification system. *Human Factors*, **49**(2), 227–242.
- Simaiakis, I., Balakrishnan, H., Khadilkar, H., Reynolds, T. G., Hansman, R. J., Reilly, B. and Urllass, S. (2014). Demonstration of reduced airport congestion through pushback rate control. *Transportation Research Part A: Policy and Practice*, **66**, 251–267.
- Smith, K. and Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors*, **37**(1), 137–148. doi:10.1518/001872095779049444
- Soo, K., Mavin, T. and Roth, M. (2016). Mixed-fleet flying in commercial aviation: A joint cognitive systems perspective. *Cognition, Technology & Work*, **18**(3), 449–463.
- Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics*, **37**, 55–79.
- Stettler, M. E. J., Eastham, S. and Barrett, S. R. H. (2011). Air quality and public health impacts of UK airports. Part I: Emissions. *Atmospheric Environment*, **45**, 5415–5424.
- Stokes, A. and Kite, K. (1997). *Flight Stress: Stress, Fatigue, and Performance in Aviation*. 2nd rev. edition, Aldershot: Ashgate Publishing.
- Stuart, S. W., McAnally, K. I. and Meehan, J. W. (2003). Head-up displays and visual attention: Integrating data and theory. *Human Factors and Aerospace Safety*, **1**(2), 103–124.
- Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design. *Situational Awareness in Aerospace Operations (AGARD-CP-478)*, 3/1–3/17. Neuilly sur Seine, France: NATO-AGARD.
- Taylor, J. L., Kennedy, Q., Noda, A. and Yesavage, J. A. (2007). Pilot age and expertise predict flight simulator performance: A 3-year longitudinal study. *Neurology*, **68**(9), 648–654. doi:10.1212/01.wnl.0000255943.10045.c0
- Teichner, W. H. (1974). The detection of a simple visual signal as a function of time on watch. *Human Factors*, **16**(4), 339–353. doi:10.1177/001872087401600402
- Torrente-Rodríguez, R. M., Tu, J., Yang, Y., Min, J., Wang, M., Song, Y., Yu, Y., Xu, C., Ye, C., IsHak, W. W. and Gao, W. (2020). Investigation of cortisol dynamics in human sweat using a graphene based wireless mhealth system. *Matter*, **2**, 921–937. doi:10.1016/j.matt.2020.01.021
- Tullis, T. and Albert, B. (2008). *Measuring the User Experience. Collecting, Analyzing, and Presenting Usability Metrics*. New York: Morgan Kaufmann.
- U.S. Department of Defense. (2012). Design criteria standard. *Human Engineering MIL-STD-1472G*. <http://www.everyspec.com>. Accessed 18 September 2020.
- U.S. Department of Transportation. (2014). Federal Aviation Agency: *Advisory circular on Electronic Flight Displays*. <https://www.faa.gov/documentLibrary/media/Advisory-Circular/AC-25-11B.pdf>. Accessed 15 March 2018.
- van Cott, H. P. and Kinkade, R. G. (1972). *Human Engineering Guide to Equipment Design*. Washington, DC: DCUS, Government Printing Office.
- van Veen, H. and van Erp, J. B. F. (2001). Tactile information presentation in the cockpit. In: Brewster, S. A. and Murray-Smith, R. (eds.). *Haptic Human Computer Interaction (LNCS2058)*, Vol. 2058, Berlin, Germany: Springer, 174–118.
- Vicente, K. J. (1999). *Cognitive Work Analysis: Towards Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, **34**(4), 457–468. doi:10.1177/001872089203400407
- Weintraub, D. J. and Ensing, M. (1992). Human factors in head-up display design. In: *The Book of HUD*, Dayton, OH: CSERIAC, 1–2.
- Wertheimer, M. (1923/1938). Untersuchungen zur lehre von der Gestalt II. *Psychologische Forschung*, **4**, 301–350. (Excerpts translated into English as ‘Laws of organization in perceptual forms’ in W. D. Ellis (Ed.), *A Source Book of Gestalt Psychology*. New York, NY: Hartcourt, Brace and Co., and as ‘Principle of perceptual organization’ in D. C. Beardslee and M. Wertheimer (Eds.), *Readings in Perception*. Princeton, NJ: D. van Nostrand Co., Inc.)
- Wickens, C. D. and Alexander, A. L. (2009). Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology*, **19**(2), 182–199.
- Wiener, E. L. (1988). Cockpit automation. In: Wiener, E. L. and Nagel, D. C. (eds.). *Human Factors in Aviation*, San Diego, CA: Academic Press, 433–461.
- Wiener, E. L. (1989). Human Factors of Advanced Technology (“Glass Cockpit”) Transport Aircraft. NASA Contractor Report 177528. Moffet Field, CA: NASA Ames Research Center.
- Williams, K. W. and Ball, J. D. (2003). *Usability and Effectiveness of Advanced General Aviation Cockpit Displays for Instrument Flight Procedures*. Washington, DC: U.S. Department of Transportation, Federal Aviation Administration. Office of Aerospace Medicine, Publication No. DOT/FAA/AM-03/17.1.
- Wilson, J. R. and Corlett, E. N. (1995). *Evaluation of Human Work. A Practical Ergonomics Methodology*, 2nd edition, London: Taylor & Francis.
- Wood, R. B. and Howells, P. J. (2001). Head-up displays. In: Spitzer, C. R. (ed.). *The Avionics Handbook*, Boca Raton: CRC Press, Chapter 4.
- Zhang, J., Johnson, K. A., Malin, J. T. and Smith, J. W. (2002). *Human-Centered Information Visualization*. Tubingen, Germany: International Workshop on Dynamic Visualization and Learning. doi:10.1.1.121.6064&rep=rep1&type=pdf.

## APPENDIX A



**Figure A1.** Cognitive work analysis (CWA) adapted from Rasmussen et al. (1994) and its relation to the design process.



**Table A1.** Overview of the design elements included in the four HUD presentation designs.

Design element	HUD1	HUD2	HUD3	HUD4	Explanation
– Consistency – Relative location	RTA, GS and DTG located upper left on the display.	RTA, ETA, RGS, GS and DTG located upper left.	RTA, ETA, RGS, GS and DTG located upper left. Semi-circular moving point dial located upper right.	RTA, ETA, RGS, GS and DTG located upper left. Moving ball vertical tape located centre right.	The content, arrangement, symbology and format of the information on the HUD should be sufficiently compatible with the HDDs to preclude pilot confusion, misinterpretation, increased cognitive workload or flight crew error (U.S. Department of Transportation, 2014).
– Control movement stereotype	Increased GS results in increased alpha-numerical GS-indication. Decreased GS results in decreased alpha-numerical GS-indication. ETA will vary accordingly.	Increased GS results in increased alpha-numerical GS-indication. Decreased GS results in decreased alpha-numerical GS-indication. ETA will vary accordingly.	Increased GS results in increased alpha-numerical GS-indication. Decreased GS results in a decreased alpha-numerical GS-indication. Increased GS makes GS-indication move clockwise. Decreased GS makes GS-indication move counter-clockwise. ETA will vary accordingly.	Increased GS results in increased alpha-numerical GS-indication. Decreased GS results in a decreased alpha-numerical GS-indication. Increased GS makes GS-indication move up. Decreased GS makes GS-indication move down. ETA will vary accordingly.	The response of a display to control movements shall be consistent, predictable and compatible with the operator's expectations (Van Cott and Kinkade, 1972; U.S. Department of Defense, 2012).

*Continued.*

*Table A1. Continued.*

Design element	HUD1	HUD2	HUD3	HUD4	Explanation
– Control response compatibility	Thrust variations result in changes of GS indicated by alpha-numerical numbers.	Thrust variations result in changes of GS indicated by alpha-numerical numbers. ETA will vary accordingly.	Thrust variations result in changes of GS indicated by alpha-numerical numbers and movement of graphical GS-indication. ETA will vary accordingly.	Thrust variations result in changes of GS indicated by alpha-numerical numbers and movement of graphical GS-indication. ETA will vary accordingly.	Controls and displays must be organised so that a control input is aligned with the response (Helander, 2006; Proctor and Vu, 2016).
– HUD screen layout	Assigned runway centre top, actual time bottom right, taxi clearance bottom left, RTA, GS, and DTG upper left.	Assigned runway centre top, actual time bottom right, taxi clearance bottom left, RTA, ETA, RGS, GS, and DTG upper left.	Assigned runway centre top, graphical GS-indication upper right, actual time bottom right, taxi clearance bottom left, RTA, ETA, RGS, GS, and DTG upper left.	Assigned runway centre top, graphical GS-indication centre right, actual time bottom right, taxi clearance bottom left, RTA, ETA, RGS, GS, and DTG upper left.	Law of proximity: objects or shapes that are close to one another appear to form groups. See <a href="#">Wertheimer (1923/1938)</a> and <a href="#">Palmer (1992)</a> on Gestalt laws.
– Emergent features	N/A	N/A	Alignment of the GS moving point dial with RGS index vertically allows for quick detection of divergences.	Alignment of the GS moving ball vertical tape with the RGS index horizontally allows for quick detection of divergences.	The interaction of small elements forming together to make a larger one. Facilitates rapid detection of divergences ( <a href="#">Dashevsky, 1964</a> ).

*Continued.*

Table A1. Continued.

Design element	HUD1	HUD2	HUD3	HUD4	Explanation
– Contrast – Character, number, and symbol attributes	Colour contrast ratio of 2.5:1. PC set to 100% brightness.	Colour contrast ratio of 2.5:1. PC set to 100% brightness.	Colour contrast ratio of 2.5:1. PC set to 100% brightness.	Colour contrast ratio of 2.5:1. PC set to 100% brightness.	Contrast for displaying alpha-numeric characters on a HUD at least 1.5:1 (U.S. Department of Defense, 2012), and at least 1.2:1 to 1.3:1 (Wood and Howells, 2001). Colour contrast ratio calculations for HUD 1–4 (ESSE, 2021).
Fonts, lower/upper case	Calibri 18 pt. fat upper case.	Calibri 18 pt. fat upper case.	Calibri 18 pt. fat upper case.	Calibri 18 pt. fat upper case.	Airbus B612 font e.g., equals to Calibri 18 pt. (Eclipse, 2021). Fat upper/lower case characters/numbers used (U.S. Department of Defense, 2012.)
Colour	Red: 0. Blue: 173. Green: 255.	Red: 0. Blue: 173. Green: 255.	Red: 0. Blue: 173. Green: 255.	Red: 0. Blue: 173. Green: 255.	Reflective HUDs are usually monochromatic, using a narrow-band green emitting phosphor, with a wavelength of 500–550 nm since the eye is most sensitive to that wavelength (Wood and Howells, 2001; U.S. Department of Defense, 2012). A wavelength of 500–550 nm is equal to 173/255/0 and 163/255/0 on the RGB Colour Scale (Hirsch, 2004; Academio, 2021).

Continued.

**Table A1.** *Continued.*

Design element	HUD1	HUD2	HUD3	HUD4	Explanation
Size	6.4 mm at a distance of 500 mm.	6.4 mm at a distance of 500 mm.	6.4 mm at a distance of 500 mm.	6.4 mm at a distance of 500 mm.	The height for HUD alpha-numeric characters should be not less than 8.1 mrad (28 min) 4.05 mm of visual angle. The height for HUD non-alpha-numeric characters should be not less than 9.9 mrad (34 min) 4.95 mm of visual angle (U.S. Department of Defense, 2012.) Font size = $\frac{254\text{mm}}{72} \times 18\text{ pt.}$ text = 6.4 mm.
– HUD screen dimensions	Width: 200 mm. Height: 150 mm.	Width: 200 mm. Height: 150 mm.	Width: 200 mm. Height: 150 mm.	Width: 200 mm. Height: 150 mm.	Typical dimensions of a commercial jet aircraft HUD is approximately a width of 200 mm to 250 mm and a height of 100 mm to 200 mm. See e.g., Collins Aerospace (2021).
– Display–eye distance	Up to 500 mm.	Up to 500 mm.	Up to 500 mm.	Up to 500 mm.	Maximum display–eye distance set to 500 mm (U.S. Department of Defence, 2012).

*Continued.*

*Table A1. Continued.*

Design element	HUD1	HUD2	HUD3	HUD4	Explanation
– Design Eye Position (DEP)	Up to 500 mm.	Up to 500 mm.	Up to 500 mm.	Up to 500 mm.	From this point, a pilot can view all relevant HDDs and the outside world scene through the aircraft windshield, while being able to access all required cockpit controls. A distance of 241–343 mm, dependent on cockpit geometry, is suggested by Wood and Howells (2001).
– Head motion box	Width: 115 mm. Height: 65 mm. Length: 150 mm.	Width: 115 mm. Height: 65 mm. Length: 150 mm.	Width: 115 mm. Height: 65 mm. Length: 150 mm.	Width: 115 mm. Height: 65 mm. Length: 150 mm.	Three-dimensional space surrounding the DEP with a width not less than 115 mm, a height not less than 65 mm and a length not less than 150 mm (U.S. Department of Defence, 2012).
– Field of view (FOV)	Vertical: 25°. Horizontal: 30°.	Vertical: 25°. Horizontal: 30°.	Vertical: 25°. Horizontal: 30°.	Vertical: 25°. Horizontal: 30°.	Maximum angular extent over which symbology from the image source can be viewed by the pilot allowing vertical and horizontal head movement within the head motion box. A 24–28° vertical and 30–34° horizontal angle suggested by Wood and Howells (2001).

Note: HUD: head-up display; RGS: required ground speed; DEP: design eye position; HDD: head-down display; GS: ground speed; FOV: field of view; RTA: required time of arrival; S, slow; ETA: estimated time of arrival; F: fast; DTG: distance to go; RBG: red, blue, green colour scale.

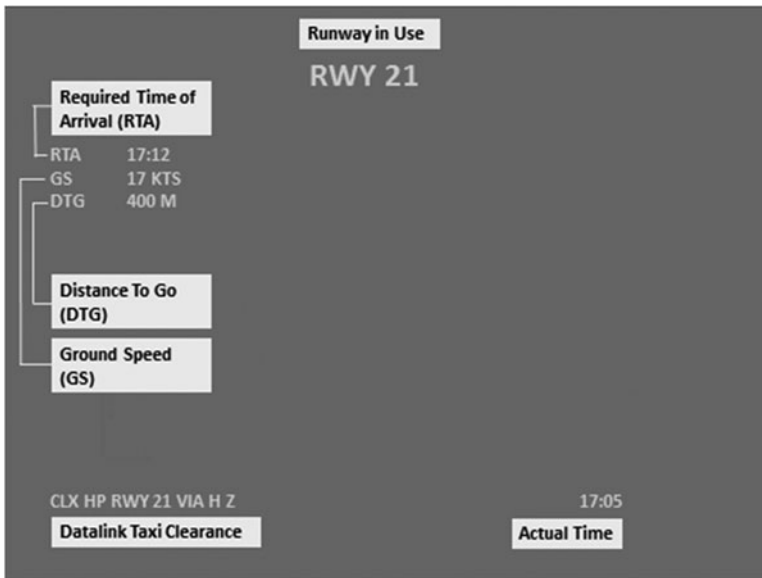


Figure A2. User description of HUD 1.

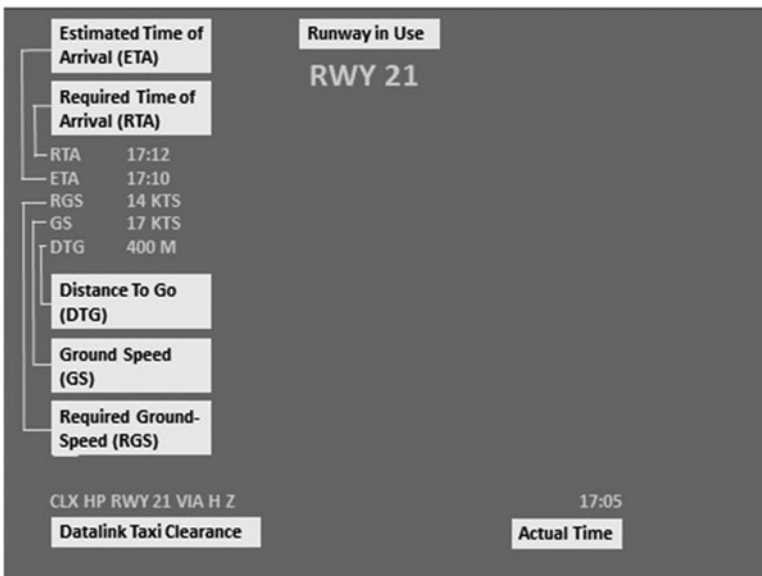


Figure A3. User description of HUD 2.

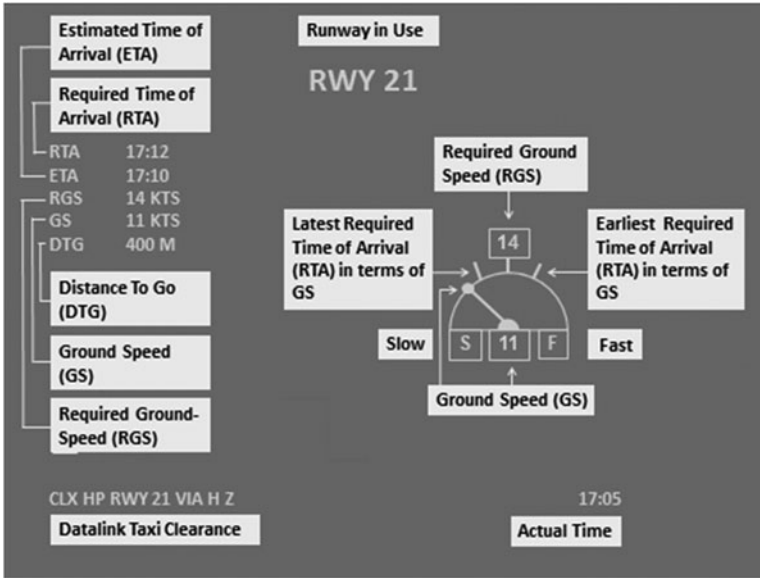


Figure A4. User description of HUD 3.

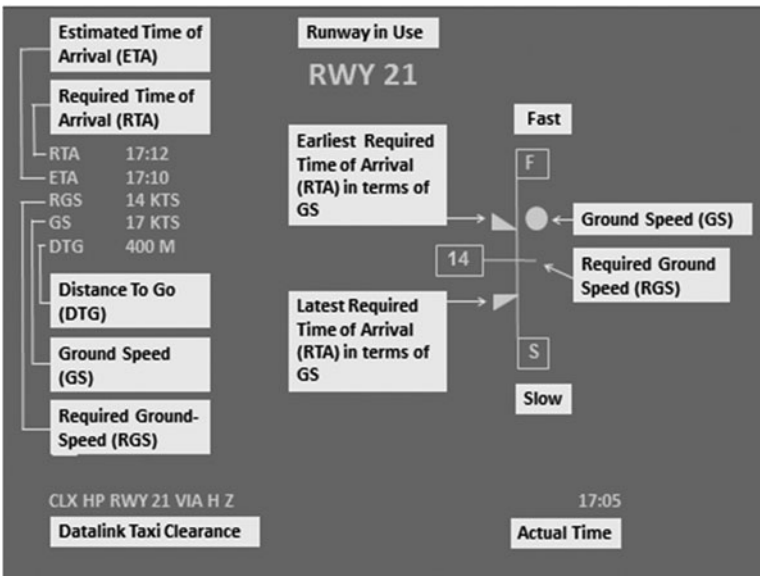
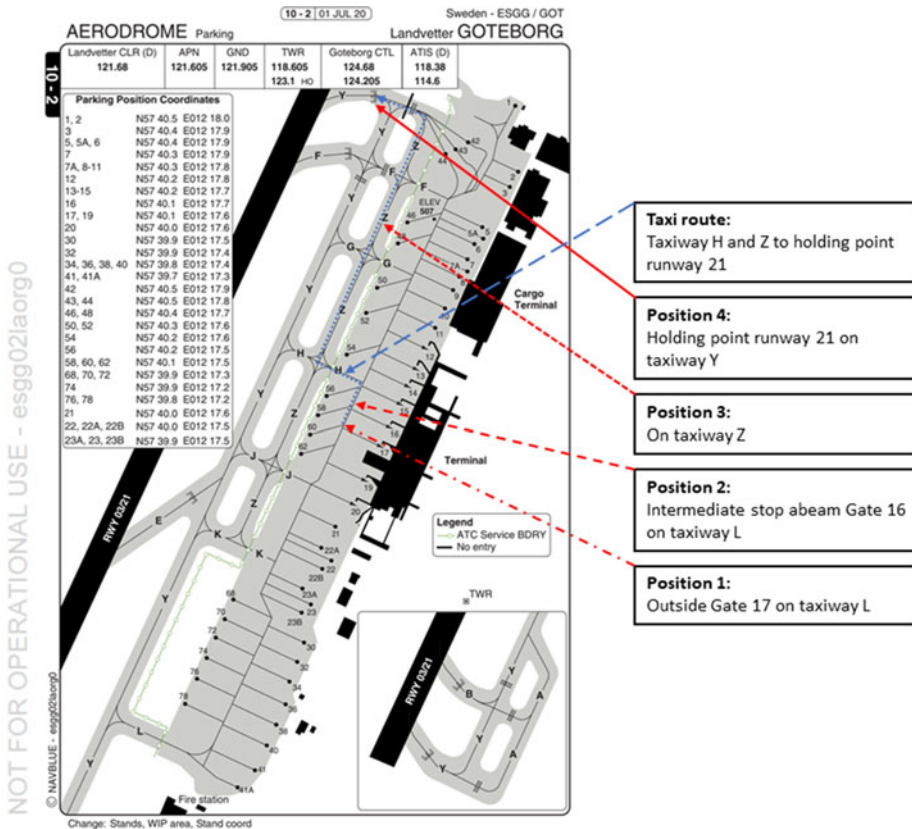


Figure A5. User description of HUD 4.

**Table A2.** Acronyms and explanations for HUD 1–4.

Acronym	Explanation
RWY 21	Runway 21 (approximately 210 degrees true direction)
CLX HP RWY 21 VIA H Z	Taxi clearance (CLX) using data link to Holding Point Runway 21 via Taxi Way Hotel (H) and Zulu (Z)
Data link	Communication between ATC and pilot without using radio telephone
17:05	Actual Coordinated Universal Time (UTC) in hours and minutes using 24 h format
RTA	Required time of arrival at holding position in hours and minutes using 24 h format
ETA	Estimated time of arrival at holding position in hours and minutes using 24 h format
GS	Ground speed in knots
RGS	Required ground speed in knots (equal to the black dot) to timely arrive at the holding point
DTG	Distance to go measured in metres (m) from gate to holding point
S	Slow, minimum GS to arrive at holding point according to RTA
F	Fast, maximum GS to arrive at holding point according to RTA



**Figure A6.** Airport navigation chart for Gothenburg Landvetter airport (ESGG/GOT) including taxi-run Position 1–4 and schematic taxi-route. Note: Chart is not for operational use. Chart date differs from the original chart date. Chart approved for publication by NAVBLUE.

Cite this article: Eklund R, Osvalder A-L (2021). Optimising aircraft taxi speed: Design and evaluation of new means to present information on a head-up display. *The Journal of Navigation* 74: 6, 1305–1335. <https://doi.org/10.1017/S0373463321000606>