A Tool for Easing the Cognitive Analysis of Design Prototypes of Aircraft Cockpit Instruments

The Human Efficiency Evaluator

Sebastian Feuerstack, Andreas Lüdtke, Jan-Patrick Osterloh
OFFIS – Institute for Information Technology
Escherweg 2, 26121 Oldenburg, Germany
{Feuerstack, Luedtke, Osterloh}@offis.de

ABSTRACT

Development and evaluation of dynamic and complex systems require new techniques and tools to evaluate the risks of Human and Systems Error, especially for safety critical systems. Established techniques like the cognitive workload analysis that can be used to assess the individual perceived operator workload for sets of tasks these are not widely used in industrial development. That is, because cognitive analysis of dynamic systems depends on complex architectures and simulations to evaluate workload over time, and is still driven by proprietary notations for cognitive models that require in-depth cognitive modeling skills and is currently only accessible to experts. In this paper we present an extension to CogTool, the Human Efficiency Evaluator (HEE) to ease the analysis of the impact of new instruments and new display designs with respect to human operator workload and task execution times. The tool is designed to make these cognitive analysis techniques available to non-experts, such as system analysts and engineers. We explain the cognitive modeling and analysis process supported by the HEE referring to an aeronautics scenario presented earlier by Hutchins. The cognitive analysis compares the task performance and workload of three generations of cockpit instrument designs to support pilots' with the slats/flaps settings during an aircraft approach with the current support in modern aircrafts and was performed by using the HEE.

Author Keywords

Cognitive Workload Analysis; Task Execution Time Prediction; Human-Machine-Interaction; Operator Workload Evaluation.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (HCI):

INTRODUCTION

Today the design of intuitive human-machine interaction (HMI) for safety critical applications is still a huge challenge for system designers. One reason is the lack of engineering methods that can readily be applied by designers. Typical questions that have to be answered by HMI designers during the development process of new assistance systems are:

1. What is the ideal position of an instrument?
2. How does a new instrument design affect the operators' task performance?
3. Does the design have an impact on the operators' workload?

These questions are usually answered by tests with human operators performing their task with realistic system prototypes and cognitive workload analysis, such as the NASA-TLX technique are then used to assess the individually perceived workload. Human operator testing can result in extensive information helping to discover common errors and usability problems and in getting feedback before the final system is being implemented. But human operator tests are also expensive in terms of time and money. Human operators that represent the targeted audience need to be recruited and paid, which is problematic in safety-critical system domains where extensively trained operators are needed (i.e. pilots). In those cases not only the participation, but also the absence of these experts in their regular position causes costs.

Further on, human operator testing can only be scaled to a very limited extend: Often, because of costs and time, only a few variants of a design can be tested, especially if these tests require a functional prototype to be implemented and also the impact of a new design can only be evaluated with respect to a small set of situations. But in commercial aviation hundreds of (standardized) procedures might be affected by an instrument change. Additionally those that are only relevant for very specific and rare situations (e.g. instrument damage or a specific weather situation) cannot be all considered in user tests and for all possible design variants. Finally, problems that are discovered while testing a functional working prototype frequently result in high re-engineering costs and time.

New methods and techniques are therefore needed to ease analyzing the impact of new instruments and new display designs. Engineers that design and implement HMI systems should benefit from a cognitive system analysis enabling them to predict task execution and operator workload in an early system design phase in that only preliminary design sketches are available. However, a Cognitive Analysis requires in depth cognitive modelling knowledge and is currently only accessible to experts.

In this paper we present an extension to CogTool [14], which is freely available as open source[1], to make Cognitive Analysis techniques available to HMI designers/engineers without requiring them to have a cognitive modelling background. We focus on the aeronautics domain as an example for a safety critical environment. Our extension is called Human Efficiency Evaluator (HEE). The HEE tool supports evaluation in early design phases to predict task performance and workload of different HMI designs by simulating the human behavior with a cognitive architecture based on low-fidelity prototypes such as photos, screenshots or sketches as input. With the HEE we contribute:

1. Extending the cognitive analysis capabilities by a tool to consider complex and dynamic HMIs of safety-critical systems.
2. Broaden the tool-supported cognitive analysis by also offering a workload-over-time assessment prediction.

This contribution does not focus on evaluating the model validity. For both contributions we integrate pre-existing architectures and prediction models: CASCaS (Cognitive Architecture for Safety Critical Task Simulation) which has been already evaluated earlier [7, 8] and the Visual, Cognitive, Auditory, and Psychomotor (VCAP) model prediction for workload evaluated in the aeronautics domain in [1, 2] and compared to other approaches in [18].

Instead, after we have discussed the related work in the upcoming section, we focus on demonstrating how an incremental cognitive analysis and prediction of operator's workloads can contribute to design decisions. Therefore, after we have presented the HEE tool in the thirst section, we introduce the fourth section a scenario taken from the Hutchin's article "How a Cockpit Remembers Its Speeds" [13]. We demonstrate for this scenario how a cognitive analysis can support design decisions and report the results of a task performance and workload analysis of the Hutchins scenario and compare these results to a slats/flaps setting cockpit support in a modern aircraft, the B737-800.

RELATED WORK

In the recent years, several tools have been proposed. The CogTool [14] enables non-experts in cognitive system analysis to create predictive human performance models to estimate the task completion time for skilled human operators of classical Windows, Icons, Menu, Pointer (WIMP) interfaces. CogTool is used in an early design phase. Photos or screenshots are arranged into wired frames and then annotated with interactive widgets that offer frame navigation (i.e. links or buttons), or more complex interactions, such as menu-navigation.

The HEE is an extension of CogTool and used to identify performance and workload "hotspots" for human-machine interaction for safety-critical systems, such as aircrafts, control rooms and clinical healthcare systems. HMIs in these domains often do not rely on typical WIMP widgets (ie. Buttons, or menus) but are assembled based on specifically designed instruments. The HEE is open to the addition of domain-specific widgets (such as flight instruments) that are modeled based on state-charts.

The Hierarchical Task Mapper (HTAmap) framework [11] is another approach to simplify the development and analysis of cognitive models aiming to reduce the development effort. HTAmap implements a pattern-based approach: It transforms sub-goal templates gained by a preceding SGT task analysis [16] into a cognitive model by associated cognitive activity patterns (CAP). Several re-usable CAPs have been implemented to generate declarative and procedural ACT-R [3] structures e.g. to describe scanning, observation, monitoring or action execution of an operator. With CAP HTAmap implements a concept for re-using concepts within a cognitive model that is task-centric while the HEE implements re-usability on an instrument level by linking instrument designs to cognitive models.

The Automation Design Advisor Tool (ADAT) [19] supports comparing Flight Management System (FMS) designs in terms of their expected effects on human performance and also evaluates FMS designs based on guidelines. Like the HEE, ADAT is designed to be used by subject matter experts (e.g. to human factor experts in the aeronautics domain) but the conceptual foundation of both tools is different: ADAT extensively applies heuristics to evaluate the display layout, whereas the HEE relies on simulating an operator with a cognitive architecture. Bothe tools evaluate the design using evaluations scales. ADAT focuses on graphical design evaluation while the HEE invests in task and workload predictions.

COGENT [5] is a graphical modeling editor targeted to psychologists that allows programming cognitive models at a higher level of abstraction. It is based on box/arrow diagrams that link to a set of standard types of cognitive modules that implement theoretical constructs from psychological theory. Both COGENT, CogTool, and HEE share the idea of making cognitive modeling easier by allowing programming on a higher level of abstraction. Whereas COGENT focuses on psychologists, the HEE is targeted to be used by subject matter experts.

The Cognitive system analysis supported by HEE is based on computational models of human cognitive processes. Cognitive models usually consist of two parts: a cognitive architecture, which integrates task independent cognitive processes (like perception, memory, decision making, learning, motor actions) and a flight procedure model which describes procedures as a temporally ordered hierarchical tree of goals (e.g. landing the aircraft), sub goals (e.g. extend flaps/slats, extend landing gear, apply air brakes) and actions (e.g. press button, move lever). Computational models are executable in a simulation environment to simulate interaction between human and machine. In order to perform such a simulation the procedure model has to be

[1] https://github.com/cogtool/cogtool, last checked 05/19/15

'uploaded' to the architecture. Thus, a cognitive architecture can be understood as a generic interpreter that executes such formalized flight procedures in a psychological plausible way. An overview of cognitive computational models like ACT-R, SOAR, MIDAS and others is provided in [6] and [20]. OFFIS has developed the cognitive model CASCaS (Cognitive Architecture for Safety Critical Task Simulation) [7, 8]. In a recent report [20] prepared for the FAA under coordination of NASA leading cognitive pilot models have been analyzed and compared with each other. CASCaS was considered "to be the most comprehensive type of pilot performance model, addressing attention, interaction, and errors." (page 64).

Several models exist to predict workload. The NASA Task Load Index (NASA-TLX) [10] is a popular subjective multi-dimensional technique to estimate mental workload. Its uses six dimensions to assess workload: mental demand, temporal demand, performance, effort, physical demand, and frustration. By a questionnaire an overall workload score is calculated based on a weighted average of ratings on each dimension. The NASA TLX is usually applied to users after they have experienced the system or by subject matter experts. Since the HEE is targeted to an early design phase evaluation (in that no systems to experience are available) and targeted to support non-experts, we have chosen a workload computation model proposed by Aldrich and McCracken [1]. VCAP uses subject matter experts to rate flight mission tasks for helicopters between 0 (no demand) to 7 (highest demand) to the following workload channels: visual; auditory; cognitive and psychomotor (movement). An recent extension of the approach [2] details the initial ordinal ratings into a cardinal scale to support summing up channel workloads to a score that enables to identify 'excessive workload'. Other workload prediction models are the Timeline-Line Analysis and Prediction workload model (TLAP) [17] and the Workload Index (W/INDEX) [15] that is based on Wickens' multiple resource theory (MRT) that describes the human operator as an information processor with fixed capacities.

A comparison of all three models based on a low-fidelity flight simulation with 16 participants [18] revealed that all three models could account for between 56% and 84% of the variance. Even though for another data pool with data from helicopter simulations collected by the same authors the predictions were less accurate. By a subsequent model fitting including the removal of the red-line assumption (which specifies the assumption of overload) the variance accounted for increased significantly for the VCAP model. The additional consideration of the assumption of cross-resource conflicts further improved the predictions of VCAP and TLAP by 12% and 9% respectively.

HUMAN EFFICIENCY EVALUATOR[2]

The Human Efficiency Evaluator (HEE) is a tool for cognitive system analysis of early design prototypes of aircraft cockpit instruments. It is based on CogTool [14]. The HEE re-uses the user interface of CogTool to compare designs based on tasks and corresponding scripts that the tool user generates by demonstrating a task on a design sketch.

The following subsection explains the overall cognitive analysis process supported by the HEE, which is quite similar to CogTool. The second subsection details our contribution: The support for adding new, domain-specific instruments based on state-chart models that rely on a set of cognitive operators with associated workload values and the automated creation of a virtual environment and a pilot model that then is simulated within CASCaS to predict the operator's workload-over-time.

Cognitive Analysis with the HEE[3]

Figure 1 illustrates the three basic activities: definition of designs, demonstration of procedures, and operator simulations of the procedure demonstrations.

Design Definition

The HEE requires photos or sketches as input. All further information required for the cognitive system analysis is then annotated directly to the photos. Figure 2 shows a screenshot of the design specification of an aircraft cockpit with the HEE: Different to CogTool that has a fixed set of predefined widgets to be used during interacting with a WIMP system (i.e. pull-down menus, buttons), we extended CogTool with a model-based backend that support the specification of arbitrary domain specific instruments (c.f. next subsection for details).

All instruments are available within the palette of instruments (the vertical bar at the left of figure 2) and can be dragged onto the photo to set their exact position in the environment. In addition to the annotation of the instruments a screenshot is annotated with the pilot's position and the photo resolution is set. The latter is done by marking one instrument on the photo with the correct physical size and enables to calculate hand and head movements of the operator more precisely. For the former, the location of the pilot's head and the pilot's distance to the control panel needs to be set to fix the pilot's initial line of sight (figure 2: looking out of the font window).

Procedure Demonstration

After the design has been defined, the pilot's procedures relevant to a design can be demonstrated. A procedure demonstration is a sequence of pilot actions that use the instruments to perform a certain task (e.g. flaps setting) towards an overall goal (e.g. landing the airplane). Figure 3 depicts a screenshot of the task demonstration editor of CogTool that has been extended to consider the dynamic instrument models and an explicit memory recall dialogue.

Like with CogTool, tasks are described by interacting with the annotated photo via a context-sensitive popup menu that is activated by clicking on an instrument with the right

[2] HEE website: http://multi-access.de/hee

[3] A video documentation of the modeling and evaluation of the Hutchins' scenario is available at: http://multi-access.de/128

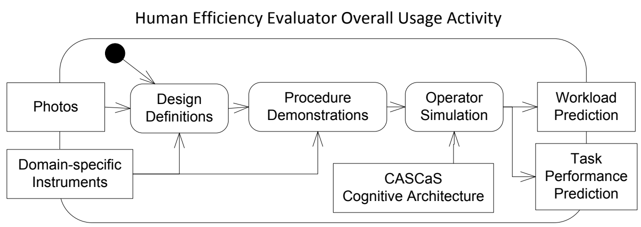

Figure 1. Activity chart illustrating the principal activities supported by the HEE.

mouse button. Different to CogTool that has pre-defined entries, the entries of the context-menu (showing the context-sensitive interaction possibilities with the instrument) of the HEE are generated by interpretation of the instruments' models.

Cognitive Simulation

Each procedure demonstration can be simulated to predict the pilots' execution performance and workload using a cognitive architecture. CogTool generates ACT-R [3] procedures that it executes in an embedded lisp interpreter. The results are then visualized inside CogTool using PERT-charts. We connected CASCaS to the HEE. CASCaS is specifically targeted to simulate cognitive processes that are relevant for the design of safety critical systems and has a long history in pilot performance evaluation in the aircraft domain [7, 8].

We extended the CogTool to generate the input data that CASCaS requires for simulating the pilot's behavior: First, a topology data file that captures the aircraft cockpit relevant instruments with their millimeter-exact position and their relevant variables (retrieved from the HMI design model) and second, the procedure files (derived from task demonstrations).

Figure 2. Instrument annotation with the HEE using a dynamic palette of predefined instruments (vertical bar at the left).

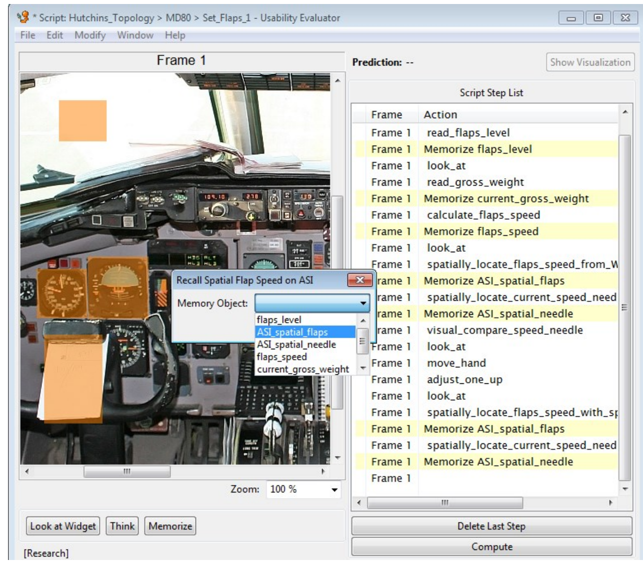

Figure 3. Working memory chunk recall during task demonstration.

Specification of Instruments' Models

New instrument for the cognitive system analysis are specified by two models: First, by a static data structure that describes data that can be read (perceived) from the instrument and/or physically manipulated by e.g. turning a wheel or pressing a button. Second, by a dynamic model that defines by a state chart the cognitive, perceptual and motor actions on the level of the Model Human Processor [4] relevant when using the instrument's functions.

Static Model

Figure 4 depicts an excerpt from the static class model definition of the Airspeed Indicator (ASI) instrument (c.f. figure 5), which supports the pilot in identifying airspeeds at which the slats/flaps setting needs to be adjusted. So called speed bugs (cf. figure 5 the white bugs, which can manually adjusted around the instrument's bezel) are used to remind a pilot to adjust the slats/flaps. The static model identifies data that can be perceived from the instrument, such as each speed_bugs' position or the current location of the speed needle for instance.

Airspeed Indicator
+current_speed : Integer +speed_needle : Spatial +speed_bug : Spatial +speed_bug_2 : Spatial +speed_bug_3 : Spatial +speed_bug_4 : Spatial
+set_speed_bug(in loc : Spatial)

Figure 4. Static airspeed indicator instrument model.

Dynamic Model

Complementary to the static model, the dynamic model defines how the instrument can be integrated in the pilot's procedures (i.e., how the instrument designer expects that the instrument will be used by the pilot). We use state chart models to specify this behavior. A dynamic instrument model connects an instrument's control action or perception of the pilot with a cognitive process specification. Control actions (such as toggling a button or adjusting the flaps lever) and perception (e.g. the pilot looks at the speed needle of the ASI) are modelled as events, whereas the relevant cognitive processes are specified using entry action triggers.

Figure 6 depicts an example from the dynamic state model of the airspeed indicator instrument. It defines two possible ways to check if the current aircraft speed requires a new slats/flaps setting: Either the pilot locates the speed bug and visually compares the spatial position of the speed bug with the position of the speed needle of the ASI (third scenario from above). Or the pilot recalls the slats/flaps speed from the working memory and performs a numerical mental comparison with the speed value read from the speed needle position.

State transitions and entry actions of a state chart can contain cognitive operators. Table 1 lists some of them, which are relevant for the cognitive analysis of the Hutchins scenario that will be discussed in the next section. With the cognitive operators, the specific pilot's behavior while using the instrument is detailed for the later simulation of the operator in the cognitive architecture. Further on, they embed workload components based on the component scales originally derived by McCracken and Aldrich [1] (cf. table 1).

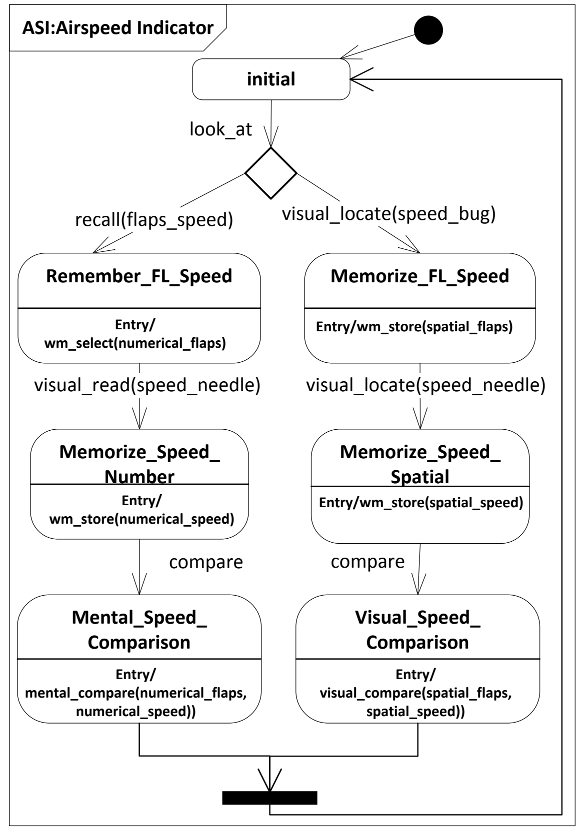

Figure 6. Dynamic model excerpt of the ASI instrument.

Case Study: Results of using HEE to analyze workload during airport approach with different assistance technologies

We choose the cognitive analysis of a pilot task in a cockpit of a commercial airliner as described in Hutchin's article "How a Cockpit Remembers Its Speeds" [13] as the scenario to explain the application of the HEE and also to test our approach for two reasons: First, it is a real-world scenario and the pilots' procedures as well as the cognitive processes are analyzed in detail. Second, the scenario, which is about the pilots' procedure of configuring the lift characteristics of

Figure 5. Airspeed Indicator with speed bugs 4

Component	*Scale Value*	*Descriptor*	*Annotation*
Visual	3,7	Visually Discriminate	visual_compare
	4	Visually Inspect	visual_inspect()
	5	Visually Locate	visual_locate()
	5,9	Visually Read	visual_read()
Cognitive	5,3	Recall, Memorize	recall(), wm_store()
	7	Estimation, Calculation	mental_compare(), calculate(),
Psycho-motor	2,2	Discrete Actuation	adjust()

Table 1. Workload components with scale values from [2] used for the Hutchins scenario.

[4] Photo by Christian Schröder licenced under CC BY-SA 3.0 http://creativecommons.org/licenses/by-sa/3.0/

the aircraft wings (via extending the flaps/slats) during approach to an airport, illustrates the continuous advancement in cockpit design.

Different to the original scenario from Hutchins that was used to explain distributed cognition and also considered the interaction between pilot and first officer we currently focus on evaluating the interaction between the cockpit instruments and the pilot flying.

During an approach the pilot has to extend the flaps and slats incrementally to different positions until the final position for landing has been achieved. Slats at the leading edge of the wing and flaps at the trailing edge of the wing are moveable parts which are extended to different positions during approach to guarantee enough lift despite decreasing speed. It is the pilots' task to achieve the different position at prescribed speed intervals by moving the Flaps Lever to the corresponding lever positions. If a flaps/slats position is achieved at a speed which is too low or too high a stall situation or a physical damage of the slats/flaps can result in a substantial risk for the aircraft to crash. The relevant speed intervals have to be computed based on the current weight of the aircraft.

For the wing configuration several instruments and sources of information are relevant: (1) The airplane weight that can be read from the Fuel Quantity Indicator (FQI). (2) The current speed of the aircraft, which is displayed by the Air Speed Indicator (ASI) and (3) the Flaps Lever (FL) instrument that is used to extend the slats and flaps subsequently to match the current aircraft speed.

With different cockpit generations and increasing cockpit automation several advances have been implemented in different aircraft types:

First scenario: "Rule of Thumb"

First cockpits did not offer specific support for the pilot to calculate the speed. Thus, by a "rule of thumb" a pilot calculated each speed limit that requires slats/flaps reconfiguration by using mental arithmetic. This calculation is based on one value, the reference landing approach speed (VRef). Even for a modern aircraft, like a Boing 737, these speeds (SPD) can be still calculated (in case of an instrument failure) based on VRef and mental arithmetic:

UP_SPD = Vref*(30/40)+70

SPD_F1 = UP_SPD – 20

SPD_F5 = SPD_F1 – 20

SPD_F15 = Vref+5

As soon as the UP_SPD is reached, the flaps lever needs to be adjusted to the "1" position and then subsequently to flaps position "5" (*SPD_1*), "15" (*SPD_F5*), and finally to "30" for *SPD_F15*. The pilot's procedure for setting the flaps can be demonstrated based on the design annotation of the cockpit (cf. fig 3): *look_at* flaps lever, *visual_read* "current_flaps_setting" (which then triggers a *memorize* current_flaps_setting" into working memory), *look_at* FQI, *visual_read* "current_weight". Then, by the custom cognitive operator *calculate_flaps_speed* the mental arithmetic of the pilot is simulated and memorized in "flaps_speed". With the flaps speed in mind, the pilot checks the ASI to *visual_read* the current airspeed and *mental_compares* it with the "flaps_speed", which follows the specification of the left branch of the ASI state chart depicted in fig. 6. Finally the pilot *looks_at*, *graps* and *adjusts* the flaps lever one level up.

The simulation of such a procedure with CASCaS results in execution time predictions for each single action of the pilot's procedure fed into the cognitive architecture. Figure 7 depicts the workload distribution all four scenarios of the Hutchins use case separated into visual, cognitive, and psycho-motoric workload components. For the first scenario peeks in the visual workload reflect the pilot collecting the information, such as the airplane weight and current flaps setting to do the mental arithmetic. The high amount of cognitive workload is the result of the pilot's mental arithmetic and the need to store and recall information during these calculations.

Second scenario: "Speed Cards"

To get rid of the high amount of cognitive workload in the first scenario, speed cards were invented. Speed cards list for a specific airplane type and the most common gross weight levels the slats/flaps configuration speeds. The speed card for the airplane weight is usually positioned in the view of the pilot (c.f. fig.3 – were it is positioned on top of the york) and implements Don Norman's "put information to the world" principle to liberate the pilot from recalling the speeds from memory. The effect on the cognitive workload of the pilot can be observed in figure 7, whereas the visual workload does not change significantly compared to the first scenario.

Third scenario: "Speed Bugs"

To reduce the visual workload, speed bugs were invented. Figure 5 depicts an ASI instrument of the MD-83 aircraft. By four moveable pointers (speed bugs) at the bezel of the ASI the pilot indicates the aircraft speeds that require a slats/flaps adjustment. The result of the simulation of the new procedure with speed bugs (c.f. fig. 7) reveals that speed bugs reduce visual workload during the approach since they indicate relevant information close to where they are required (i.e. no head turns) and also shift the effort to calculate or review the correct speeds to a different, less stressful flight phase (i.e. during cruise).

Fourth scenario: "Primary Flight Display"

Modern aircraft cockpits integrated the air speed indicator information in the Primary Flight Display (PFD). It is an instrument that aggregates large amount of current flight data and predictive data about the anticipated aircraft state with a single glance. Figure 8 depicts the PFD of a modern B737-800 aircraft. The vertical moving speed stripe on the left illustrates the current aircraft speed. Like in the MD-83 ASI, speed bugs along the speed stripe indicate slats/flaps adjustments. They are computed automatically. At any point in time the next relevant speed bug is displayed (confirm figure 8: the "-1" symbol (light green) indicating to "set flaps to position 5" at an airspeed of 192). The effect of the new integrated PFD instrument to the visual and cognitive workload is visible in fig.7: The PFD condenses relevant information for adjusting the flaps/slat setting to one single area of interest. Fig 7 also depicts the psychomotor workload, which remains the same in all scenarios (adjusting the flaps lever) but happens much earlier for the last scenario. We could also observe that the overall time to perform the slats/flaps setting procedure is reduced from approx. 2751ms for scenario 1 2649ms (2), 2419ms (3), to 906ms when using the PFD in scenario 4.

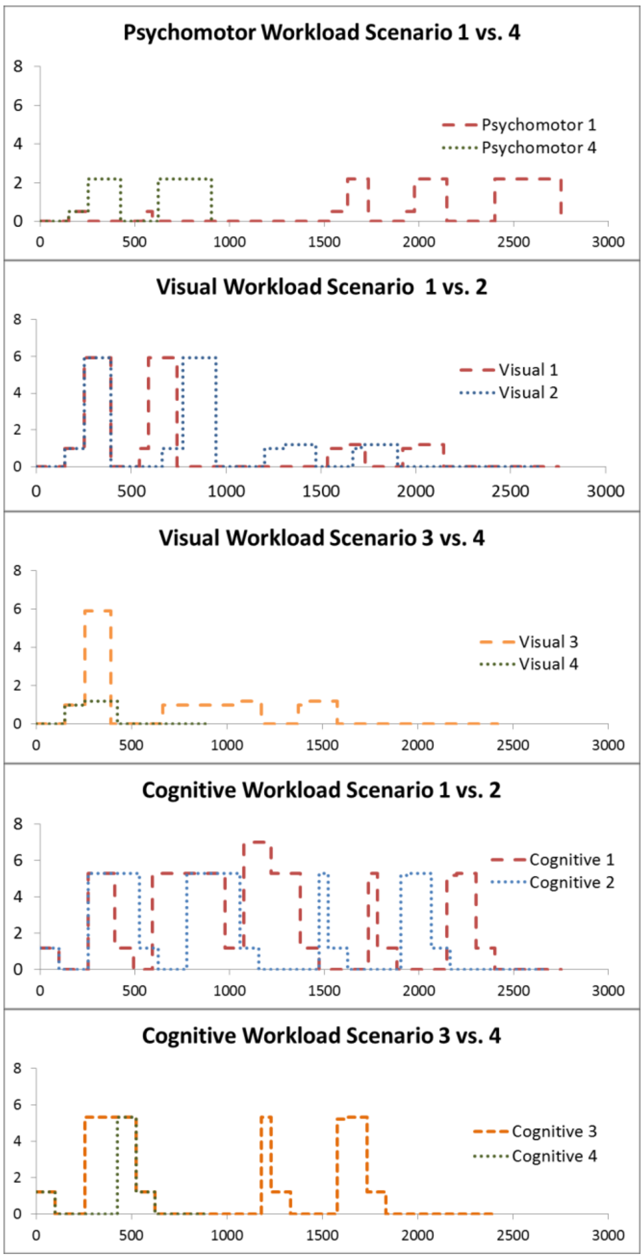

Figure. 7. Comparison of Psychomotor, Visual, and Cognitive workload distribution over time (ms).

DISCUSSION AND CONCLUSION

By analyzing pilot's task performance and workload with HEE for different cockpit designs, we were able to demonstrate the benefit of the different pilot assistance systems. The analysis shows that task performance and workload could be significantly reduced with each new instrument setup for managing the slats/flaps setting task during an approach.

However, our current approach still suffers several limitations and the results have to be treated therefore with caution: With the current state of the HEE we can only generate predictions of single tasks, which do not reflect a realistic situation in a cockpit, in that typically several procedures are performed in parallel (i.e. speed and altitude monitoring). Attention shifts between parallel tasks can have a significant impact on the performance of an individual procedure.

The workload metric by McCracken, Aldrich and Bierbaum [1, 2] is a subjective workload measurement that is based on data gained by Subject Matter Experts. Initial validation has focused on studies with military helicopter pilots [1]. Recently the workload metric has been extended to consider commercial airliner pilots [9]. Hollnagel and Woods compared several measurements employed in empirical research and balanced them regarding two dimensions: The ease of measurement and the meaningfulness of a measurement [12]: Expert ratings have been identified to have a stronger theoretical basis as direct workload measurements while also being easier to measure.

The intention of our approach is to offer an indication of workload and task performance "hot spots" for human machine interfaces in an early design phase. In these stages where no functional prototype is available not the absolute measurements, but the comparison of design variants is the most interesting aspect.

Figure 8. B737 Primary Flight Display (PFD) with vertical speed stripe at the left.

ACKNOWLEDGMENTS

The HEE is developed in the EU ARTEMIS JU project HoliDes (http://www.holides.eu/) SP-8, GA No.: 332933. Any contents herein reflect only the authors' views. The ARTEMIS JU is not liable for any use that may be made of the information contained herein.

REFERENCES

[1] Craddock W. Aldrich, T. B. and J. H. McCracken. A computer analysis to predict crew workload during lhx scout-attack missions:. Technical Report (MDA903-81-C-0504/ASI479-054-I-84(B), United States Army Research Institute Field Unit, Fort Rucker, AL, USA:, 1984.

[2] Theodore B Aldrich, Sandra M Szabo, and Carl R Bierbaum. The development and application of models to predict operator workload during system design. In *Applications of human performance models to system design*, pages 65–80. Springer, 1989.

[3] John R. Anderson and Christian J. Lebiere, editors. *The Atomic Components of Thought*. Lawrence Erlbaum Associates, Mahwah, NJ, 1998.

[4] Stuart K. Card, Allen Newell, and Thomas P. Moran. *The Psychology of Human-Computer Interaction*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1983.

[5] Richard Cooper and John Fox. Cogent: A visual design environment for cognitive modeling. *Behavior Research Methods, Instruments, & Computers*, 30(4):553–564, 1998.

[6] Chris Forsythe, Michael L. Bernard, and Timothy E. Goldsmith, editors. *Cognitive Systems: Human Cognitive Models in Systems Design*. Psychology Press, 2006.

[7] Andreas Lüdtke; Jan-Patrick Osterloh; Florian Frische. Multi–criteria evaluation of aircraft cockpit systems by model–based simulation of pilot performance. In *Proceedings of ERTS - Embedded Real Time Software and Systems 2012*, 02 2012.

[8] F. Frische, J.-P. Osterloh, and A. Lüdtke. Simulating visual attention allocation of pilots in an advanced cockpit environment. In *Modelling and Simulation (MODSIM) World Conference & Expo*, Hampton, Virginia, USA, 13– 15 October 2010.

[9] Brian F. Gore, Becky L. Hooey, Nancy Haan, Deborah L. Bakowski, and Eric Mahlstedt. A methodical approach for developing valid human performance models of flight deck operations. In Masaaki Kurosu, editor, *HCI (16)*, volume 6776 of *Lecture Notes in Computer Science*, pages 379–388. Springer, 2011.

[10] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[11] Dzaack-J. Wiesner A. & Urbas L. Heinath, M. Simplifying the development and the analysis of cognitive models. In *Proceedings of EuroCogSci07*, Delphi, Greece., 2007.

[12] Erik Hollnagel and David D. Woods. *Joint cognitive systems : foundations of cognitive systems engineering*. Taylor & Francis, Boca Raton, FL, 2005.

[13] Edwin Hutchins. How a cockpit remembers its speeds. *Cognitive Science*, 19(3):265–288, 1995.

[14] Bonnie E. John, Konstantine Prevas, Dario D. Salvucci, and Ken Koedinger. Predictive human performance modeling made easy. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 455–462, New York, NY, USA, 2004. ACM.

[15] RobertA. North and VictorA. Riley. W/index: A predictive model of operator workload. In GrantR. McMillan, David Beevis, Eduardo Salas, MichaelH. Strub, Robert Sutton, and Leo Van Breda, editors, *Applications of Human Performance Models to System Design*, pages 81–89. Springer US, 1989.

[16] Thomas C. Ormerod and A. Shepherd. *Using task analysis for information requirements specification: The SGT method.* Lawrence Erlbaum Associates, 2004.

[17] A.M. Parks and G.P. Boucek. Workload prediction, diagnosis and continuing challenges. In G.R. McMillan et al., editor, *Applications of Human Performance Models to System Design*. New York: Plenum Press, 1989.

[18] Kenneth J. Sarno. Role of multiple resources in predicting time-sharing efficiency: Evaluation of three workload models in a multiple-task setting. *The International Journal of Aviation Psychology*, 5(1):107–130, 1995. PMID: 11541493.

[19] Angelia Sebok, Christopher Wickens, Nadine Sarter, Stacey Quesada, Connie Socash, and Brian Anthon. The automation design advisor tool (adat): Supporting flight deck design in nextgen. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting*, 2010.

[20] Christopher Wickens, Angelia Sebok, John Keller, Steve Peters, Ronald Small, Shaun Hutchins, Liana Algarín, Brian F. Gore, Becky L. Hooey, and David C. Foyle. Modeling and evaluating pilot performance in nextgen: Review of and recommendations regarding pilot modeling efforts, architectures, and validation studies. Technical report, Human Centered Systems Lab, 2013.