Developing a Measurement for Task Complexity in Flight

Yiyuan Zheng; Yanyu Lu; Zhen Wang; Dan Huang; Shan Fu

INTRODUCTION: Task complexity is regarded as an essential metric that is related to a pilot's performance and workload. Normally, pilots follow Standard Operating Procedures (SOPs) during a flight. In this study, we developed a measurement named Task Complexity in Flight (TCIF) to represent the task complexity in the SOPs.

METHODS: The TCIF measurement combined four complexity components into one index: actions logic complexity (ALC), actions size complexity (ASC), information control exchange complexity (ICEC), and control mode complexity (CMC).To verify the measurement, we calculated 11 tasks during the takeoff and landing phases from the SOPs, and invited 10 pilots to perform the same tasks in a flight simulator. After flight, the TCIF results were compared with two workload measurements: the Bedford scale and heart rate.

RESULTS: The results of TCIF and the 4 components of the 11 tasks were calculated. Further, the TCIF results showed a significant correlation with the Bedford scores ($R = 0.851$) and were also consistent with the difference in heart rate ($R = 0.816$). Therefore, with the increased TCIF results, both the Bedford scale and the difference in heart rate increased.

DISCUSSION: TCIF was proposed based on the flight operating conditions. Although additional studies of TCIF are necessary, the results of this study suggest this measurement could effectively indicate task complexity in flight, and could also be used to guide pilot training and task allocation on the flight deck.

KEYWORDS: task complexity, standard operating procedures, task complexity in flight.

Zheng Y, Lu Y, Wang Z, Huang D, Fu S. *Developing a measurement for task complexity in flight.* Aerosp Med Hum Perform. 2015; 86(8):698–704.

Task complexity during flight is an essential characteristic that could influence flight safety, since it determines pilot performance and behavior on the flight deck. Task complexity should not interfere with the pilot's workload and should be within his/her own capability to ensure the pilot remains in control during the flight. In order to simplify and standardize the pilot's operations in flight, normally Standard Operating Procedures (SOPs) are provided which typically offer a list of action items or criteria arranged in a systematic manner, allowing the operator to record the presence/absence of the individual items listed to ensure that all are considered or complete.[9] By using the SOPs, pilots only need to follow the detailed procedures and operations to revert a complex situation to a safe configuration, especially when encountering an emergency or system malfunction. Consequently, the task complexity in the SOPs should guarantee optimum performance of the aircraft equipment and standard operations for the pilots. In addition, task complexity analysis could guide pilot training. More complex tasks require more training processes. However, no research has been focused on task complexity in flight in accordance with the SOPs.

In other fields, many studies have already shed light on task complexity analysis. Tan, Ng, and Mak[22] studied the effects of task complexity on auditors' performance. Kelly[14] presented ways in which the science of complexity helped to achieve peak performance. Campbell [3] summarized that, in the goal-setting process, task complexity and performance were inversely related. Moreover, systematic and logical relationships were also found among task complexity, types of information, information channels, and sources.[1] Wilson and Caldwell[27] indicated that under fatigue effects, the decrements in pilots' performance developed sooner in simple, nonengaging tasks than in the complex and more engaged tasks. Some studies also focused on

From the School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai, China.

This manuscript was received for review in November 2014. It was accepted for publication in April 2015.

Address correspondence to: Yiyuan Zheng, Dongchuan Rd, #800, Minhang Area, Shanghai 200240, China; leodeisler@sjtu.edu.cn.

DOI: 10.3357/AMHP.4189.2015

the task complexity decomposition.[11] Liu and Li[16] summarized 27 salient factors and 10 dimensions of task complexity. Nevertheless, all of the above were qualitative analyses which were subjective and required sufficient experience of the researchers. Several other studies tried to use quantification indices to present task complexity; for instance, Step Complexity (SC) measurement efficiently quantified the step complexity of nuclear power plants,[18] and Zhang developed a measurement of operation complexity in spaceflight.[29]

However, unlike task characteristics in the nuclear industry, tasks in flight are more varied and complex, which requires more information exchange and control operations of the pilots. SC measurement only considered step complexity and could not describe task complexity in flight comprehensively, especially during the flight phases which required a significant amount of a pilot's physical and mental workload, such as takeoff and landing. Since human-machine interaction is critical during these flight phases and the interaction is strongly related to the systems complexity,[28] the information acquisition and process, the corresponding devices, and the operations of the pilots should be regarded as paramount when analyzing task complexity in flight.

Therefore, for something more adapted to flight circumstances, we proposed a measurement called Task Complexity in Flight (TCIF). TCIF modified the three components of SC, step size complexity, step logic complexity, and step information complexity, to four different components: the Actions Size Complexity (ASC), Actions Logic Complexity (ALC), Information Control Exchange Complexity (ICEC), and Control Mode Complexity (CMC). The four components' complexities were deduced from three kinds of graphs: the Action Control Graph (ACG), Information Control Exchange Graph (ICEG), and Control Mode Graph (CMG).

As task complexity was related to the operator's workload, we implemented TCIF across 11 tasks in the takeoff and landing phases that were selected from the SOPs in a CRJ-200 flight simulator, and used 2 different workload measurements, which were the Bedford scale and heart rate (HR),[2] to verify it. The Bedford scale is one of the widely used subjective workload measurements and heart rate is considered the most common and reliable measure of workload.[12]

METHODS

Subjects

The 10 Chinese male commercial airline pilots who participated in the experiment were qualified captains or co-captains of the CRJ-200; mean total flight hours were 2368 ± 896 h (range from 1000 to 5000 h), and mean flight hours in the previous week before the experiment were 6.56 ± 5.66 h (range from 0 to 26 h). All the subjects signed the subject consent form for the present study before the experiment, which was approved by the Institutional Review Board of the School of Aeronautics and Astronautics, Shanghai Jiao Tong University.

Equipment

The experiment was carried out in a CRJ-200 full-flight simulator in Qingdao, China. It is a qualified commercial flight simulator (level C) conforming to the guidance presented in Federal Aviation Administration Advisory Circular AC 120-40B. The heart rate of each subject was recorded with physiological parameter monitoring equipment (Bio Harness, Zephry Technology, Annapolis, MD).

Procedure

TCIF consisted of four components which were deduced from three corresponding graphs of the task as shown in **Table I**. In order to show how to calculate TCIF, we used the initial climb task as an example. This task, which included four procedures, as shown in **Table II**, started from landing gear retraction and ended with flaps retraction in the takeoff phase. In the task, in addition to following the Flight Director (FD) displaying on the Primary Flight Display (PFD), the pilot needed to adjust the flight modes on the Flight Control Panel (FCP), such as Speed/Heading, which were displayed on the Flight Mode Annunciator (FMA) on the PFD, and supervise the gears and flaps configurations on the Engine Indication and Crew Alerting System.

Among the four components of TCIF in Table I, ALC and ASC, which described the amount and logical sequence of required actions in the task, were deduced from ACG. ACG of the initial climb task, which is shown in **Fig. 1,** was depicted based on Actions/Action number in Table II. To calculate the complexity of the given graph, we introduced two different kinds of order entropy. The first-order entropy calculated the regularity of the program control logic and the second-order entropy evaluated the number of hierarchical levels of the graph.[4] The first-order entropy considered the nodes that had the same number of incoming and outgoing paths as equivalent node classes and the second-order entropy regarded nodes as equivalent if they had the same number and type of neighbor nodes. Thus, ALC was geared to the first-order entropy and ASC belonged to the second-order entropy of ACG. The formula for the entropy value is below:

Table I. Components, Complexity Contribution, and Graphs of TCIF.

COMPONENTS	COMPLEXITY CONTRIBUTION	GRAPH
Actions Logic Complexity of task (ALC)	Logical sequence of the required actions	Actions Control Graph (ACG)
Actions Size Complexity of task (ASC)	The amount of required actions	
Information Control Exchange Complexity (ICEC)	HMI-Information Control Exchange	Information Control Exchange Graph (ICEG)
Control Mode Complexity (CMC)	HMI-Control Mode Cognition and Operations	Control Mode Graph (CMG)

Table II. Task Decomposition of Initial Climb.

	ACTIONS / ACTION NUMBER	DISPLAYS	CONTROL DEVICES
Start	Landing gear retraction / 1	EICAS→Gear Indicator	
Procedure 1	Following FD / 2	PFD→FD	Control Wheel (2 dimensions)
	Setting Speed Mode on FCP / 3	PFD→FMA→SPD	Speed Mode Switch→On/Off
	Checking Speed Mode on PFD / 4		
Procedure 2	Following FD / 2	PFD→FD	Control Wheel
	Setting Heading Mode on FCP / 5	PFD→FMA→HDG	Heading Mode Switch→On/Off
	Checking Heading Mode on PFD / 6		
Procedure 3	Following FD / 2	PFD→FD	Control Wheel
	1000 feet, setting speed to 200 knots / 7	PFD→Altitude→1000; FCP	Speed knobs→rotation→200 knots
Procedure 4/End	Checking if the speed is increasing / 8	PFD→Vertical Speed→↑	Throttle (1 dimension)
	Retracting Flaps / 9	EICAS→Flaps Indicator	Flap Gate→array→up

EICAS: Engine Indication and Crew Alerting System; FD: Flight Director; PFD: Primary Flight Display; FMA: Flight Mode Annunciator; SPD: Speed; HDG: Heading; FCP: Flight Control Panel.

$$Entropy = H = -\sum_{i=1}^{h} p(A_i)\log_2 p(A_i) \qquad \text{Eq. 1}$$

where H was the number of node classes, and $p(A_i)$ was the frequency of the node class A_i.[17]

As in Fig. 1, node 3 and node 5 had same number of incoming and outgoing paths ({2; 1}), and so did node 4 and node 6 ({2; 3}). Otherwise, the remaining five nodes all had different numbers of incoming and outgoing paths. Therefore, nine nodes composed seven node classes of the first-order entropy in Fig. 1, and ALC equaled:

$$H_{ALC} = -\sum_{i=1}^{7} p(A_i) = \left[5 * \left(\frac{1}{9}\log_2\frac{1}{9}\right) + 2 * \left(\frac{2}{9}\log_2\frac{2}{9}\right)\right] = 2.726$$

Considering the second-order entropy of ACG, since none of the nodes had the same neighbors with other nodes in Fig. 1, therefore each node constituted one node class and there were nine node classes. ASC was equivalent to:

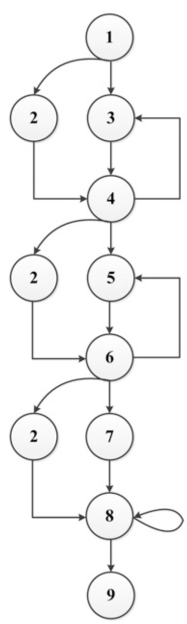

Fig. 1. Action Control Graph (ACG) of the initial climb task, where the numbers represent the action numbers in Table II.

$$H_{ASC} = -\sum_{i=1}^{9} p(A_i) = \left[9 * \left(\frac{1}{9}\log_2\frac{1}{9}\right)\right] = 3.170$$

Each flight task involved information and control flow exchange that constituted the most crucial part of the task load. System information required real-time control in aircraft piloting.[7] Furthermore, information resource channel conflicts would lead to human error.[24] In other words, the greater the requirements of the channels in the procedure, the higher the complexity they would spawn. Consequently, two factors should be taken into account when defining ICEC. First, the complexity of the hierarchical levels of the information and control exchange process should be described according to ICEG, and the second-order entropy was used to calculate H_{P_n}, where P_n was the procedure n in the task. The second factor was the occupation circumstance of information resource channels in each procedure. Five channels were considered, which were visual and auditory channels with the left hand, right hand, and feet.[21] We defined the influence factor of the channel's occupation for the procedure n was $log_2(C_{P_n})$. Therefore, as shown below, $ICEC_{P_n}$ equaled:

$$ICEC_{P_n} = log_2(C_{P_n}) * H_{P_n} \qquad \text{Eq. 2}$$

where P_n was the procedure n in the task, and the $ICEC_{all}$ of the task was calculated as:

$$ICEC_{all} = \sqrt{\frac{\sum (ICEC_{P_n})^2}{n}} \qquad \text{Eq. 3}$$

As in Procedure 1 of the initial climb task, pilots needed to follow the FD and select Speed Mode on the FCP. After selecting, they had to confirm the mode on the FMA. Therefore, the ICEG of Procedure 1 was depicted as in **Fig. 2**. Since there were nine nodes in Fig. 2 and none of the nodes had the same neighbors with other nodes, therefore nine nodes constituted nine node classes of the second-order entropy, and H_{P_1} equaled:

$$H_{P_1} = -\sum_{i=1}^{9} p(A_i) = \left[9 * \left(\frac{1}{9}\log_2\frac{1}{9}\right)\right] = 3.170$$

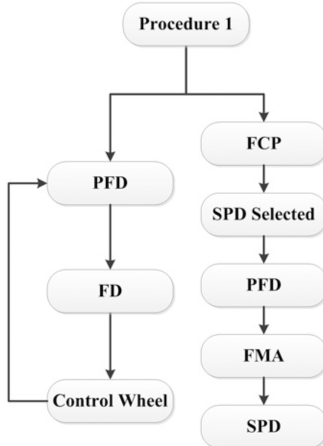

Fig. 2. Information Control Exchange Complexity (ICEC) of Procedure 1 in the initial climb task. FCP: Flight Control Panel; PFD: Primary Flight Display; SPD: Speed; FD: Flight Director; FMA: Flight Mode Annunciator.

Meanwhile, the requirements of information resource channels were the visual channel with the left hand and right hand in Procedure 1. Therefore, the influence factor of information resource channels was $log_2(3)$ and $ICEC_{P_1}$ was 5.0242. The $ICEC_{P_n}$ of the three other procedures and the overall $ICEC_{all}$ in the task were calculated according to Eq. 1–3. Similarly, the results were 5.024, 5.265, 5.265, and 5.146, respectively.

For different tasks, the required control devices varied. Even for the same device, the operating modes were multiple. For instance, throttle operation provided forward thrust in most conditions and could also provide reverse thrust when the aircraft was landing on the runway. Another example was the operations on the FCP. Pilots could select appropriate flight modes by activating buttons like Speed Mode, Heading Mode, etc. They could also set required airspeed, heading, and altitude values on the FCP. However, these two kinds of operations would generate two different complexities. The former one only needed a button push and the latter manipulation required more precise adjustment. Therefore, to obtain a comprehensive task complexity, CMC should also be taken into account. In Table II the required control devices are listed and CMG of the initial climb task is shown in **Fig. 3**.

Similarly, none of the nodes had the same neighbors with other nodes in accordance with the second-order entropy in Fig. 3, therefore all the 17 nodes formed 17 node classes. CMC of the initial climb task was:

$$H_{CMC} = -\sum_{i=1}^{17} p(A_i) = \left[17 * \left(\frac{1}{17}\log_2\frac{1}{17}\right)\right] = 4.088$$

TCIF combined ALC, ASC, ICEC, and CMC together to form an index to indicate the task complexity in flight. The calculation of TCIF is shown below, where α,β,γ, and δ were the weighting factors of ALC, ASC, ICEC, and CMC, correspondingly.

$$TCIF = \sqrt{(\alpha * ALC)^2 + (\beta * ASC)^2 + (\gamma * ICEC)^2 + (\delta * CMC)^2} \quad \text{Eq. 4}$$

The weighting factors were assigned according to the fuzzy analytical hierarchy process, which was used to determine the relative importance in multicriteria decision making.[23] There were 10 experienced male pilots (mean flight hours = 7173.2 h, SD = 5270.9 h) who participated in the determination. All of them were asked to grade relative weights among the four components. Based on the pilots' comparisons, the weighting factors for α, β, γ, and δ were 0.22, 0.23, 0.26, and 0.29. Therefore, TCIF of the initial climb task was 2.022 according to Eq. 4.

There were 11 tasks, 5 in the takeoff phase and 6 in the landing phase, that were selected from the CRJ-200 SOPs to calculate their TCIF. Each task consisted of three to six procedures, as shown in **Table III**. Furthermore, these 11 tasks were carried out in the CRJ-200 flight simulator by the 10 pilots. Each of them performed the 11 tasks in 1 experiment twice, once as pilot flying and once as pilot monitoring. Two indices, the Bedford scale and heart rate, were recorded to evaluate the TCIF results when the pilots performed as pilots flying. The Bedford scale, which is a 1-10 scale that assesses four levels of workload, was presented by a flight instructor when the pilot fulfilled each task during the experiment. Scales of 1 through 3 were scores of satisfactorily perceived workload of the task. Scales of 4 through 6 indicated a perceived workload that was tolerable to complete the task. Scales of 7 to 9 represented a perceived workload that was not tolerable for the given task. A scale of 10 was reserved for when the pilots' workload was too high to complete the task at all with all available effort given to the task.[6] In addition, in order to eliminate the individual variation of HR,[19] HR-D, which was equal to the difference between real-time HR and baseline HR of the pilot, was used:

$$\text{HR-D} = \text{Real time HR} - \text{Baseline HR} \quad \text{Eq. 5}$$

where the real time HR was the value recorded during the experiment and the baseline HR was the mean value recorded when each pilot performed a cruise task about 5 min before the experiment in the same flight simulator.

Statistical Analysis

SPSS 17.0 for Windows was used to process the experiment data. Linear regression analysis was implemented to establish the relationship between TCIF with the Bedford scale and HR-D. The significance of the regression coefficient was analyzed by *t*-test and, when $P < 0.05$, the result was considered statistically significant.

RESULTS

The results of the TCIF and the corresponding 4 different components of the 11 tasks are listed in **Table IV**. The calculation process was similar for the initial climb task. Besides, the mean value and SD of the Bedford scale and HR-D for the 10 pilots who

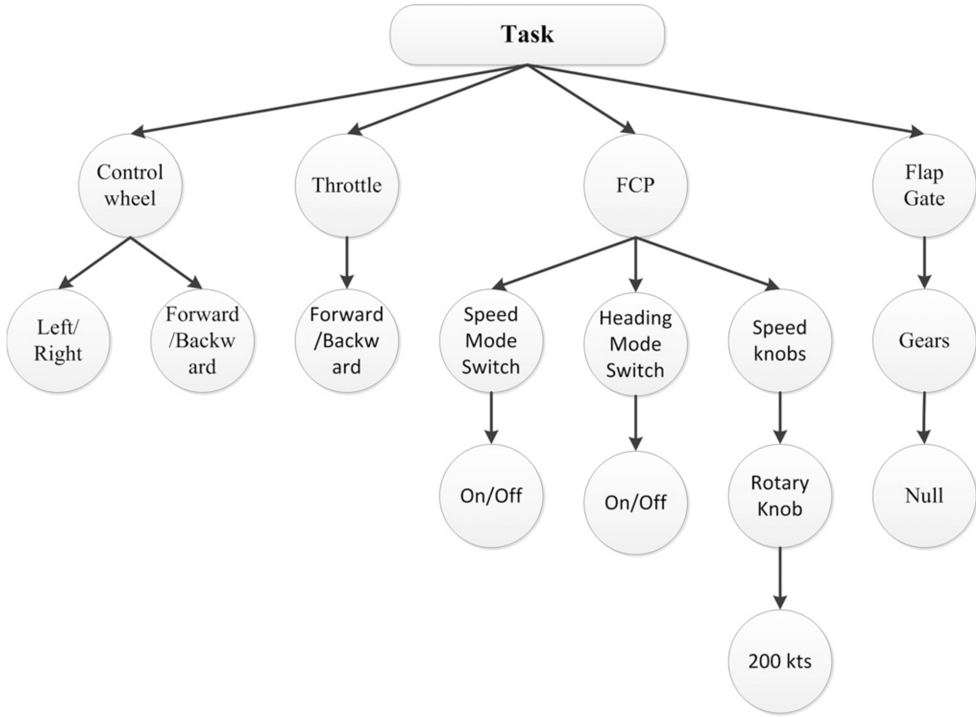

Fig. 3. Control Mode Complexity (CMC) of the initial climb task.

performed the 11 tasks in the experiment are also shown in Table IV**.**

The linear regression curves of the TCIF results with the Bedford scale and HR-D are shown in **Fig. 4**. The results indicated that, with a more complex task, the higher the Bedford scale would be (R = 0.851, $P = 0.001$) in the experiment. The coefficients passed the t-test ($t = 4.918$, $P = 0.001$). Therefore, the results from TCIF measurement reasonably match with the Bedford scale. Similarly, the TCIF results also presented a great correlation with HR-D (R = 0.816, $P = 0.002$) and the coefficients passed the t-test ($t = 4.238$, $P = 0.002$), indicating that the model was statistically significant. Thus, the TCIF results were consistent with HR-D. Otherwise, the two workload measurements also showed a good correlation (R = 0.859, $P = 0.001$) in the experiment. With the increased task complexity in flight, both the Bedford scale and HR-D increased correspondingly.

Table III. Eleven Flight Tasks and the Corresponding Number of Procedures.

TASK NUMBER	TASK NAME	NUMBER OF PROCEDURES
Task 1	Preparation before taxiing	6
Task 2	Rotation	3
Task 3	Gear up	3
Task 4	Initial climb	4
Task 5	Climbing to 10,000 ft	3
Task 6	Flaps down	3
Task 7	Localizer intercepting	4
Task 8	Gear down	3
Task 9	Glide slope intercepting	4
Task 10	Setting the altitude to go around	4
Task 11	Landing	5

DISCUSSION

The primary purpose of task analysis was to compare the demands of the system on the operator with the capabilities of the operator and, if necessary, to alter those demands, thereby reducing error and achieving successful performance.[15] As an essential characteristic of task, task complexity should be considered carefully when performing task analysis. In this study, TCIF measurement was developed based on four components, ALC, ASC, ICEC, and CMC, to indicate the task complexity in flight. ALC and ASC represented the complexities of the operations size and logic in the task. ICEC and CMC were newly introduced, comparing with SC measurement. Aircraft flight operations require a large amount of manipulation, not only surveillance. Therefore, ICEC was crucial when considering the task complexity, as it expressed the information acquisition and process, and the device operations of pilots. ICEC integrates information and the control flow exchange of each procedure in the task according to the SOPs. Appropriate information and control interaction could guarantee flight safety. According to accident investigators, aviation accidents are usually the result of a chain of unexpected events, culminating with the unsafe controls of operators.[20] Furthermore, ICEC included the influence of information resource channels occupation on complexity. According to multiple resource theory, if different resource channels were occupied simultaneously, the mental workload of operators increased dramatically, even jeopardizing the task.[25] However, ALC, ASC, and ICEC still could not reflect the whole task complexity in flight. To be more comprehensive, task complexity should contain three aspects: functional, behavior, and structural.[10] Therefore, CMC was also necessary, both in functional and behavioral aspects. It indicated the corresponding operating mode complexity of the control device in the task.

Task complexity in flight correlated with the operator's workload and the four components of TCIF reflected different workload aspects. ASC was related to the physical workload and ALC was relevant to the mental workload. Since ICEC indicated the human-machine interaction process, it reflected the mental workload and time pressure. CMC was also related to the physical workload.

The Bedford scale was implemented as a subjective workload measurement, rather than the generally used NASA-TLX, because NASA-TLX is a post hoc measurement, while the Bedford scale was used during the task. Consequently, the Bedford

Table IV. Results of TCIF and the Four Different Complexity Components.

TASK NUMBER	ASC	ALC	ICEC	CMC	TCIF	HR-D		BEDFORD	
						MEAN	SD	MEAN	SD
Task 1	2.000	2.000	2.585	2.585	1.191	2.3	4.14	1	0.67
Task 2	2.000	2.000	4.097	2.585	1.506	6	3.77	2.6	1.07
Task 3	1.585	1.585	5.024	3.000	1.728	7	4.08	3.3	0.82
Task 4	3.170	2.726	5.146	4.088	2.022	8.7	4.64	2.3	1.06
Task 5	3.585	1.793	6.329	3.807	2.277	15.2	6.32	3.2	0.79
Task 6	2.320	1.520	3.680	2.000	1.344	8.1	4.25	2.4	0.97
Task 7	3.807	1.149	5.114	3.322	1.943	12.3	5.19	3.2	1.03
Task 8	2.807	1.149	2.000	3.459	1.275	5.4	2.22	1.6	0.84
Task 9	3.170	2.059	6.035	4.170	2.231	16.5	6.80	3.6	1.07
Task 10	3.700	3.085	6.479	3.322	2.337	13.5	4.74	4.2	1.14
Task 11	2.807	2.522	6.644	3.459	2.290	26.5	6.69	4.6	0.97

scale could provide a better real-time performance indicator. Although it might result in some distractions, the Bedford scale attempted to minimize the extra workload by only asking questions for yes or no answers from pilots; for instance, was the workload satisfactory and was there any spare time for additional tasks, etc. Meanwhile, the Bedford scale was more connected to task complexity, since it could assign the workload to four different levels. According to the TCIF results and the linear regression curve with the Bedford scale, it could predict what kind of task complexity was acceptable or insufferable for pilots.

Furthermore, TCIF results could guide pilot training. The more complex tasks need more training processes for pilots to become familiar with them. Only with enough practice and experience could the pilots' resources demanded to accomplish a task decrease until the response became automatic.[26] Thus, their reactions could be consistent, fast, and error free, and automatically triggered. However, excessive complexity might still result in the omission of procedural steps in a task, which is a form of human error with serious consequences in many complex work settings.[13] In other words, the tasks, which are allocated to pilots during their flight, should not override their limitations. Therefore, TCIF measurement might also direct task allocation on the flight deck.

However, some questions still remain about TCIF measurement. First, since all the SOPs that are provided to pilots must be acceptable to warrant task completion, the range of tolerable complexity of the SOPs should be determined. This might be solved based on correlation with the Bedford scale. According to existing four grading, more tasks in the SOPs would be required to determine the precise correlation between TCIF and the Bedford scale to determine the acceptable level of task complexity. If the task complexity exceeds human capability, it is better to allocate the whole task or parts of the functions to the automatic system. Second, task procedures are typically developed without considering that operations might be perturbed and disrupted.[8] In the experiment, there were no disruptive conditions during the tasks. However, as interruptions from the ATC or flight attendants can happen at any moment in real flight, pilots can be distracted from regular operating procedures.[5] Therefore, TCIF is an ideal value regarding a given task which is desirable to be accomplished without interference and can only represent the expected operating process. When implementing TCIF in evaluating the task complexity of the SOPs or in some other implementations, some redundancies or other measures might be taken into account.

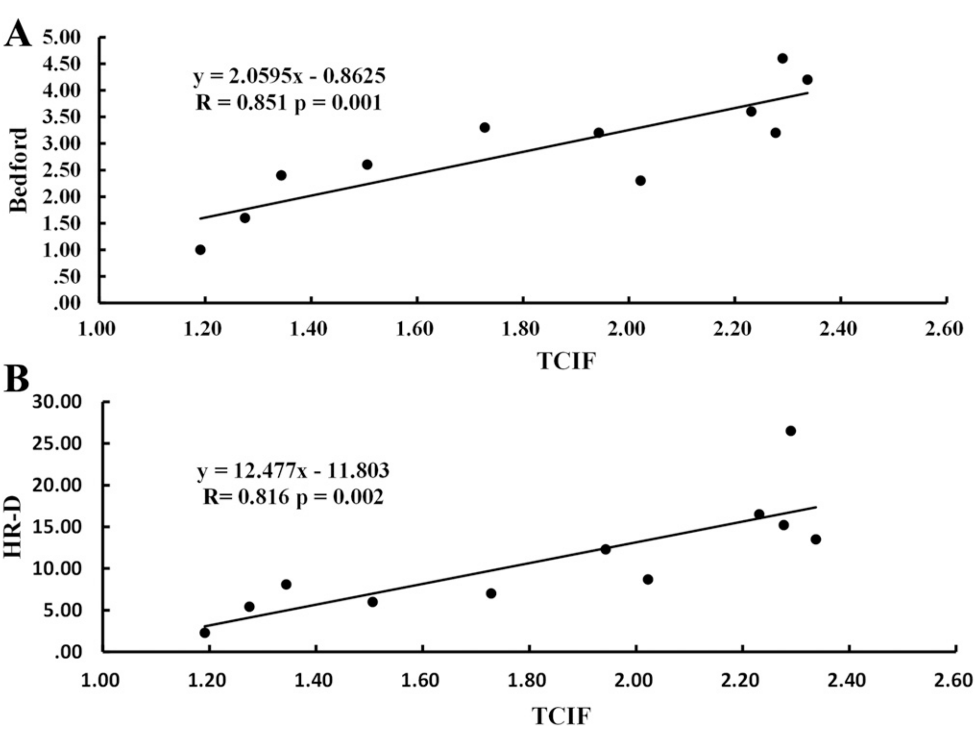

Fig. 4. The linear regression results of the Task Complexity in Flight (TCIF) values with two different categories of workload measurements. A) Regression result of the Bedford scale with TCIF. B) Regression result of HR-D with TCIF.

In conclusion, in this study, we developed a measurement named TCIF. Although additional implementing studies are needed, the results of this study suggest TCIF could effectively indicate task complexity in flight by comparing it with two workload measurements. TCIF measurements could also be used to guide pilot training and task allocation on the flight deck.

ACKNOWLEDGMENTS

This research work was supported by the National Basic Research Program of China (973 Program No. 2010CB734103) and the National Natural Science Foundation of China (61305141).

Authors and affiliations: Yiyuan Zheng, Ph.D. Candidate, Yanyu Lu, Doctor, Zhen Wang, Ph.D. Candidate, and Dan Huang, Associate Professor, School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai; and Shan Fu, Professor, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China.

REFERENCES

1. Bystrom K, Jarvelin K. Task complexity affects information seeking and use. Inform Process Manag. 1995; 31:191–213.
2. Cain B. A review of the mental workload literature. Toronto, Ontario (Canada): DRDC Toronto; 2007. DTIC Document2007.
3. Campbell DJ. Task complexity: a review and analysis. Acad Manage Rev. 1988; 13(1):40–52.
4. Davis JS, LeBlanc RJ. A study of the applicability of complexity measures. IEEE Trans Software Eng. 1988; 14(9):1366–1372.
5. Dismukes K, Young G, Sumwalt R. Cockpit interruptions and distractions. Flight Safety. 1999; 55(6):16.
6. Ellis KKE. Eye tracking metrics for workload estimation in flight deck operations. [Thesis.] Iowa City (IA): University of Iowa; 2009.
7. Endsley MR. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics. 1999; 42(3):462–192.
8. Greeno JG. Situations, mental models, and generative knowledge. In: Klahr D, Kotovsky K, editors. Complex information processing. Hillsdale (NJ): Lawrence Erlbaum; 1989:285-318.
9. Hales BM, Pronovost PJ. The checklist-tool for error management and performance improvement. J Crit Care. 2006; 21(3):231–235.
10. Ham D-H, Park J, Jung W. A framework-based approach to identifying and organizing the complexity factors of human-system interaction. IEEE Syst J. 2011; 5(2):213–222.
11. Ham D-H, Park J, Jung W. Model-based identification and use of task complexity factors of human integrated systems. Reliab Eng Syst Saf. 2012; 100:33–47.
12. Hankins TC, Wilson GF. A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. Aviat Space Environ Med. 1998; 69(4):360–367.
13. Hobbs A, Williamson A. Associations between errors and contributing factors in aircraft maintenance. Human Factors. Hum Factors. 2003; 45(2):186–201.
14. Kelly S, Allison MA. The complexity advantage: how the science of complexity can help your business achieve peak performance. New York (NY): McGraw-Hill; 1999.
15. Kirwan B, Ainsworth LK. A guide to task analysis: the task analysis working group. Boca Raton (FL): CRC Press; 2004.
16. Liu P, Li Z. Task complexity: a review and conceptualization framework. Int J Ind Ergon. 2012; 42(6):553–568.
17. Mowshowitz A. Entropy and the complexity of graphs. Bull Math Biophys. 1968; 30(1):175–204.
18. Park J, Jung W, Ha J. Development of the step complexity measure for emergency operating procedures using entropy concepts. Reliab Eng Syst Saf. 2001; 71(2):115–130.
19. Roscoe AH. Heart rate as a psychophysiological measure for in-flight workload assessment. Ergonomics. 1993; 36:1055–1062.
20. Shappell S, Detwiler C, Holcomb K, Hackworth C, Boquet A, Wiegmann DA. Human error and commercial aviation accidents: an analysis using the human factors analysis and classification system.Hum Factors. 2007; 49(2):227–242.
21. Wickens CD, Holland JG. Engineering psychology and human performance, 3rd ed. London (UK): Pearson; 1999.
22. Tan H-T, Ng TB-P, Mak BW-Y. The effects of task complexity on auditors' performance: the impact of accountability and knowledge. Auditing: A Journal of Practice & Theory. 2002; 21(2):81-95.
23. Tang Y-C, Beynon MJ. Application and development of a fuzzy analytic hierarchy process within a capital investment study. J Econ Manag. 2005; 1(2):207–230.
24. Wickens CD. Multiple resources and mental workload.Hum Factors. 2008; 50(3):449–455.
25. Wickens CD. Multiple resources and performance prediction. Theor Issues Ergon Sci. 2002; 3(2):159–177.
26. Wickens CD, McCarley JS. Applied attention theory. Abingdon (UK): Taylor & Francis; 2007.
27. Wilson GF, Caldwell JA, Russell CA. Performance and psychophysiological measures of fatigue effects on aviation related tasks of varying difficulty. Int J Aviat Psychol. 2007; 17(2):219–247.
28. Wilson JR. Fundamentals of systems ergonomics/human factors. Appl Ergon. 2014; 45(1):5–13.
29. Zhang Y, Li Z, Wu B, Wu S. A spaceflight operation complexity measure and its experimental validation. Int J Ind Ergon. 2009; 39(5):756–765.