Test Pilot and Airline Pilot Differences in Facing Unexpected Events

Yiyuan Zheng; Yanyu Lu; Yuwen Jie; Zhiqiang Zhao; Shan Fu

BACKGROUND: Unexpected events in flight might decrease the transparency of the flying process and weaken the pilot's perception of the current state, or even erode manipulating skills. However, during the flight test of a new or modified aircraft, to verify the boundaries of aircraft aerodynamic performance and handling stability, unexpected events may be encountered that need to be handled by the test pilot. Therefore, studying the differences between test pilots and airline pilots could help improve flight safety.

METHODS: Two kinds of physiological parameters, eye blink rate and average fixation duration and task-related performance of test pilots and airline pilots, were analyzed in three abnormal scenarios. A total of 16 pilots participated. The study was carried out in an A320 flight simulator.

RESULTS: The differences were significant for both test pilots and airline pilots in eye blink rate and average fixation duration. Furthermore, the reaction time of test pilots (Mean = 23.38 s) was significantly shorter than airline pilots (Mean = 42.63 s) in Unreliable Airspeed condition, and the pitch angle deviations between them were significant in both Wind Shear and Unreliable Airspeed condition.

DISCUSSION: The uncertainty of environmental change could create more severe pressure and mental workload influence than actual system failure. For airline pilots, compared with test pilots, the importance of practicing manual flight should still be emphasized. Improving reactions to unexpected ambient conditions and unannounced fault status could also contribute to flight safety.

KEYWORDS: test pilots, airline pilots, flight performance.

Zheng Y, Lu Y, Jie Y, Zhao Z, Fu S. *Test pilot and airline pilot differences in facing unexpected events.* Aerosp Med Hum Perform. 2023; 94(1):18–24.

The increased use of advanced techniques and automated systems has not only reduced pilot's workload, but also improved aviation safety remarkably. For instance, the implementation of head-up displays significantly enhances the pilot's situational awareness during takeoff and landing under night flight conditions or flying in bad weather with poor visability.[24] Moir and Seabridge reported the integrated module avionics system provides a more concise display mode and more reasonable alarm logic, making it easier for pilots to operate and monitor the aircraft.[14] However, even for well-trained airline pilots, as Landman et al. stated, excessive automation may decrease the transparency of the flying process and weaken the pilot's perception of the current state, which may lead to automation surprises.[12] Moreover, the extensive use of automation may erode the pilots' manipulating skills. Plenty of aircraft accident reviewers have reported situations in which pilots encountered abnormal automation events. The latest disasters of Lion Air Flight 610 and Ethiopian Airlines flight 302 in 2018 and 2019, respectively, both revealed the flight crews were unable to effectively recognize and respond to undesired multiple airplane automated nose-down stabilizer trim movement and the effects of potential Angle of Attack (AOA) sensor failure.[20]

From the Shanghai Aircraft Airworthiness Certification Center of CAAC and Shanghai Jiao Tong University, Shanghai, P.R. China.

This manuscript was received for review in August 2022. It was accepted for publication in October 2022.

Address correspondence to: Yiyuan Zheng, Ph.D., NO.128 Konggang yi Rd., Shanghai, Shanghai, P.R. China 200335; zhengyiyuan@saacc.org.cn.

DOI: https://doi.org/10.3357/AMHP.6157.2023

Therefore, the novel technology and sophisticated automation design features must be subjected to strict airworthiness certification to ensure safety, as Wise and Hopkin suggested.[28] A large number of compliance activities need to be carried out in the validation and verification stage. The most commonly used methods of compliance (MOC) defined in airworthiness regulations are flight test (MOC 6) and simulator test (MOC 8).[26] Flight test is a kind of compliance activity which obtains and analyzes the required data through the test aircraft flying under real atmospheric scenarios and evaluates the design specifications and safety level of the aircraft.[3] Normally many boundary conditions of aircraft aerodynamic performance and handling stability characteristics should be verified by flight test. For instance, it needs to be demonstrated by flight test that Vlof (Liftoff Speed) shall not be less than 110% of Vmu (Minimum Unstick Speed) in full engines, and 105% of Vmu in single engine shutdown. In general, Perkins indicated that flight test is the preferred method to show compliance rather than simulator test, unless the demonstration is too risky, or the required environment or airplane conditions are too difficult to attain.[17]

Typically, for a newly designed aircraft, the test flight usually takes over 3000 flight hours, and even for a modification model, the duration lasts often more than 1000 h. Thus, it requires test pilots' participation to perform test flights in a relatively short period of time, and not all airline pilots are qualified or capable of completing a flight test. Test pilots, as Culick suggested, refer to the personnel who conduct specific maneuvering flight in a novel or modified aircraft, play an important role in the flight test certification.[5] They evaluate the flight performance and verify the compliance of specific airworthiness standards through acquiring measurement parameters. The minimum entry criterion for a test pilot is to reach the flight instructor level with no less than 7000 flight hours of route operating experience. In addition, he/she must complete a professional training course lasting 50 wk at a qualified test pilot school.[9] Meanwhile, to maintain the qualification, test pilots must satisfy the experience requirements of instrument flight and night flight.

Many studies have investigated the impact of expertise on pilots' physiological characteristics and their flight performance. Undoubtedly, expertise casts light on establishing and maintaining situation awareness in the face of automation surprises or unexpected events. Kasarskis et al. found during VFR flight, experts had obviously shorter dwell times and more total fixations than novices.[11] Similar results were also found in glideslope control and dynamic target tracking tasks where experienced and novice pilots differed in scanning strategies and areas of interest.[10] Furthermore, Tsang identified expert pilots were also able to direct their attentions in a manner conducive to selecting flight-relevant diagnostic information.[27] Endsley expressed more expert pilots made better decisions, such as the future flight state projections, based on current aircraft attitude and speed.[7] However, for routine operations or frequent faults in actual flight, the difference is insignificant. Casner et al. found when abnormal events were presented in a familiar context, reactions were consistent with accepted standards and varied little from pilot to pilot.[2] Nevertheless, to our knowledge, rare study has experimentally examined the behavior of test pilots when facing the unexpected events in flight.

To establish more optimized coping strategies for unexpected events or automation surprises, and more effective training/retraining planning, we studied the differences between test pilots and commercial airlines pilots in their physiological characteristics and flight performance in three abnormal scenarios, including: Encountering Wind Shear after lifting off, Unreliable Airspeed during taking off, and Stabilizer Trim Failure during approach. In each scenario, two kinds of physiological parameters, eye blinks rate and average fixation duration, and task-related performance were analyzed. The experiment was carried out in a D-level A320 flight simulator, using 16 subjects–8 test pilots and 8 airline pilots.

METHODS

Subjects

For this study, 16 Chinese male pilots (8 test pilots and 8 airline pilots), ranging in age from 36 to 52 (Mean = 45.3 ± 4.96), participated. The mean total flight hours of those pilots were 8967 ± 3465 (range from 3000 to 15,000). Among them, eight test pilots, three from Civil Aviation Administration of China (CAAC) and five from Commercial Aircraft Corporation of China, ranged in age from 44 to 50 (M = 46, SD = 2.90), with average total experience of 10,682 h (SD = 2937). The other eight pilots, with 42.5 yr (SD = 5.24) average age and 7252 h mean flight experience (SD = 3275), were all from China Eastern Airlines. Furthermore, each pilot had been captain of Airbus 320, and simultaneously some of them had been recruited as captains for some other types of aircrafts (3 for A330, 2 for A350, and 2 for A380). Before the experiment, all subjects signed the consent form, which was approved by the Institutional Review Board of Shanghai Jiao Tong University.

Equipment

The experiment was carried out on one A320 D-level full flight simulator, which belonged to CAAC in Shanghai, China. The flight simulator conformed to the guidance published in Federal Aviation Administration Advisory Circular AC 120-40B (Airplane Simulator Qualification).[8] The flight simulator had also been used as pilot training and other airworthiness technology research. The checklist, quick reference handbook, and simulator configuration were provided to the pilots. In addition, one head-mounted eye tracker (Tobbi Glass III, Sweden) was used in this study to capture the required data of each subject's dominant eye. The eye tracker was calibrated by instructing participants to gaze at one fixed point before the experiment. Horizontal and vertical eye movement trajectories were interpolated to determine fixation point with a resolution of approximately less than 0.2 cm. The sample of the eye tracker was 100 Hz.

Procedures

For the sake of investigating the tests pilots and commercial airlines pilots' differences, three abnormal scenarios were designed, including Encountering Wind Shear after lifting off (WS), Unreliable Airspeed during taking off (UA), and Stabilizer Trim Failure during approach (STF). The relevant tasks configurations and the procedures of the crew operating are listed below.

Encountering Wind Shear After lifting off. This flight task was conducted in Shanghai Pudong International Airport. The task was initiated when the TOGA (Takeoff/Go-around) button was pressed. Then, the pilot increased the thrust and kept accelerating until the aircraft reached the speed of V1 (takeoff decision speed). Simultaneously, one moderate predicted wind shear at 400 feet was settled. When the corresponding alter appeared, the pilot was required to push the throttle to the maximum position immediately and rotate at the speed of VR. Subsequently, he should increase the pitch angle and maintain it at 18° until getting rid of wind shear (2000 ft). In this scenario, the reaction time to wind shear and pitch angle deviation during the climb were selected to reflect the pilot performance.

Unreliable Airspeed During taking off. This flight task was carried out in Shanghai Pudong International Airport. The pilot performed takeoff and initial climbing according to the standard operation procedures. At an attitude of 5000 ft, total pitot blockage occurred, resulting in unreliable airspeed. The pilot needed to recognize the current airspeed was inconsistent with the state of the aircraft, adjust the thrust, and maintain the height until the airplane reached the target pitch angle corresponding to Flight Crew Operating Manual (FCOM). Subsequently, he was required to keep climbing manually to 20,000 ft according to the weight and center of gravity of the aircraft and the appropriate pitch angels at different flight levels in FCOM. In this scenario, the response time to the unreliable airspeed and pitch angle deviation during the climbing was selected to reflect the pilot performance.

Stabilizer Trim Failure. This task was carried out during the approach phase (Position: PDL, N31 07.8, E121 40.3), and the terminal point was runway 35R in Shanghai Pudong International Airport. The initial status of aircraft in this task was 210 kts speed, 8900 ft altitude, and 168-degree heading. Then, a failure of horizontal stabilizer jamming was set, and the primary flight control system was degraded to the direct mode. Meanwhile, one 'STAB FAULT' warning appeared on EICAS display instantly.

The pilot performed a manual trim by pressing STAB TRIM to try to restore the failed state. After an invalid attempt, he pressed the CUT OUT button to switch off the stabilizer trim tunnel, and adopted the current speed as maximum flight speed and VRef Full +15 kt as reference landing speed to land the airplane with a 3-detent flaps configuration. In this scenario, the reaction time to the alert, and the deviation between actual landing speed and reference speed was selected to reflect the pilot performance.

The research subjects were in the pilot flying role from the left seat. The experiments were carried out from 8:00 a.m. to 4:00 p.m., local time, and all the participants reported being well rested. Each pilot was involved for a maximum of 2 h. Before the experiment, each subject was trained with normal flight profile for half an hour to become familiar with the simulator configurations and the procedures, and was instructed to deal with the unexpected events based on alarm system, display information, and FCOM in the formal test. An experienced A320 type rated flight instructor acted as the nonflying support pilot.

Statistical Analysis

SPSS 17.0 for Windows was used to process the experiment data. ANOVA analysis was implemented in this study. When $P < 0.05$, the results were considered statistically significant.

RESULTS

The results of the experiment would be described in two dimensions. Due to individual differences, physiological parameters would be analyzed considering same subjects. On the other hand, the flight performance of two types of pilots would also be compared based on different trials.

Eye Blink Rate

For test pilots, the difference was significant (F (2, 21) = 5.799, $P = 0.010$) in three scenarios. In UA, the average eye blinks rate was maximum (Mean = 12, SD = 2.62), followed in STF (Mean = 11.25, SD = 2.49), and the minimum was in WS (Mean = 8.13, SD = 2.10). Further, post hoc tests showed a significant difference between WS and UA ($P < 0.01$), and between WS and STF ($P = 0.017$). For airline pilots, most results of eye blinks rate were similar. The most frequent average blink rate was found in UA (Mean = 10.63, SD = 2.13), then in STF (Mean = 8, SD = 1.77), and the least was in WS (Mean = 7.88, SD = 1.55), as shown in **Fig. 1**. The difference was also significant [F (2, 21) = 5.726, $P = 0.010$]. However, post hoc tests showed a significant difference between WS and UA ($P < 0.01$), and between UA and STF ($P < 0.01$).

Average Fixation Duration

Considering average fixation duration in three abnormal events, the results of test pilots and airline pilots were similar. The minimum average fixation duration both occurred in STF, which was 1.77 s (SD = 0.35) and 2.19 s (SD = 0.26) respectively. The medium duration was in UA, which was 2.12 s (SD = 0.24) and 2.33 s (SD = 0.18) separately, and the maximum duration appeared in WS, which was 2.46 s (SD = 0.26) and 2.63 s (SD = 0.22), as shown in **Fig. 2**. In addition, the difference is significant for both test pilots (F (2, 21) = 11.519, $P < 0.01$) and airline pilots (F (2, 21) = 8.614, $P < 0.01$). For test pilots, post hoc tests showed a significant difference between WS and UA ($P = 0.028$), WS and STF ($P < 0.01$), and between UA and STF ($P = 0.024$). However, for airline pilots, only between WS and UA ($P = 0.010$), and between WS and STF ($P < 0.01$), the differences were significant.

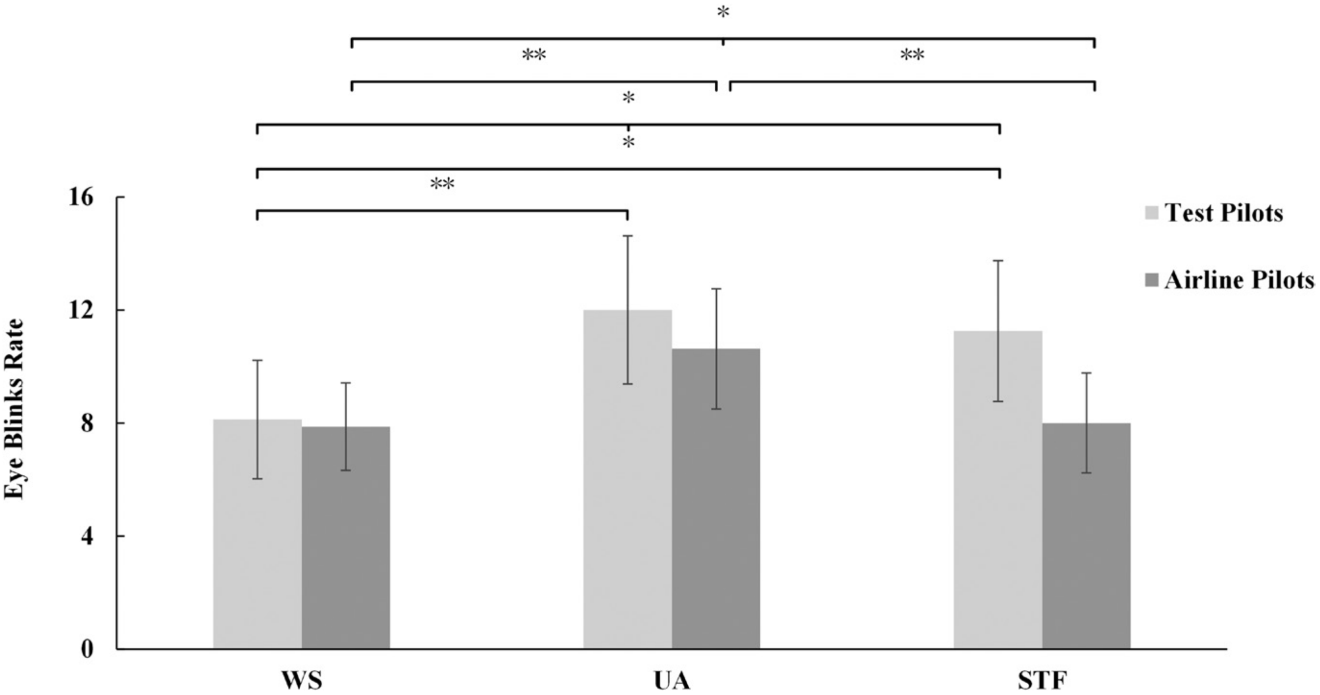

Fig. 1. The results of Eye blinks rate of test pilots and airline pilots in three flight tasks, which were Encountering Wind Shear after lifting off (WS), Unreliable Airspeed during taking off (UA) and Stabilizer Trim Failure (STF) (* $P < 0.05$, ** $P < 0.01$) . The error bars stand for the SD of eye blinks rate of the subjects either for test pilots or for test pilots.

Encountering Wind Shear After Lifting Off

In this scene, we were interested in reaction time to wind shear and pitch angle deviation during the climb. The reaction time was the interval from 'wind shear' flashing and voice warning appeared to the pilots pushing the throttle to the maximum position. The pitch angle deviation was equal to the difference between the pilot's average pitch angle and 18° during disengagement from wind shear. The mean reaction time of test pilots to wind shear was 3.96 s (SD = 0.72), and for airline pilots, the average reaction time was 4.05s (SD = 0.72). Comparing their reaction time revealed no significant difference [F (1, 14) = 0.06, $P = 0.811$]. Further, there was significant

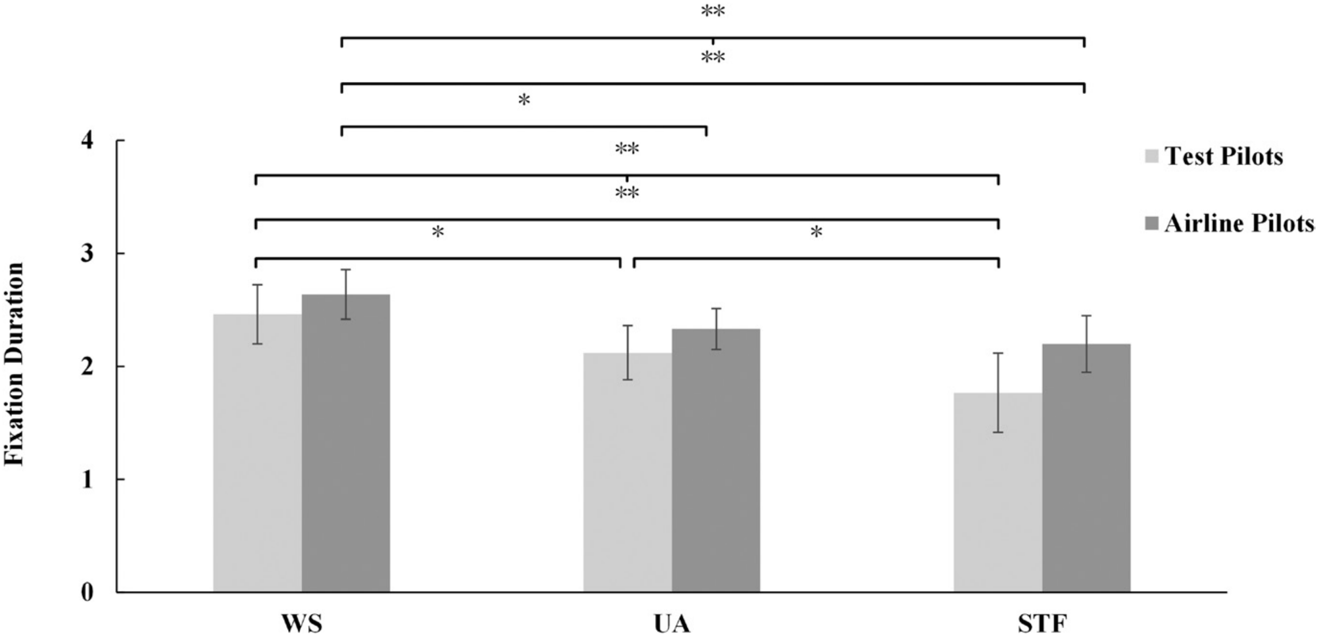

Fig. 2. The results of fixation duration of test pilots and airline pilots in three flight tasks, which were Encountering Wind Shear after lifting off (WS), Unreliable Airspeed during taking off (UA) and Stabilizer Trim Failure (STF) (* $P < 0.05$, ** $P < 0.01$) . The error bars stand for the SD of fixation duration of the subjects either for test pilots or for test pilots.

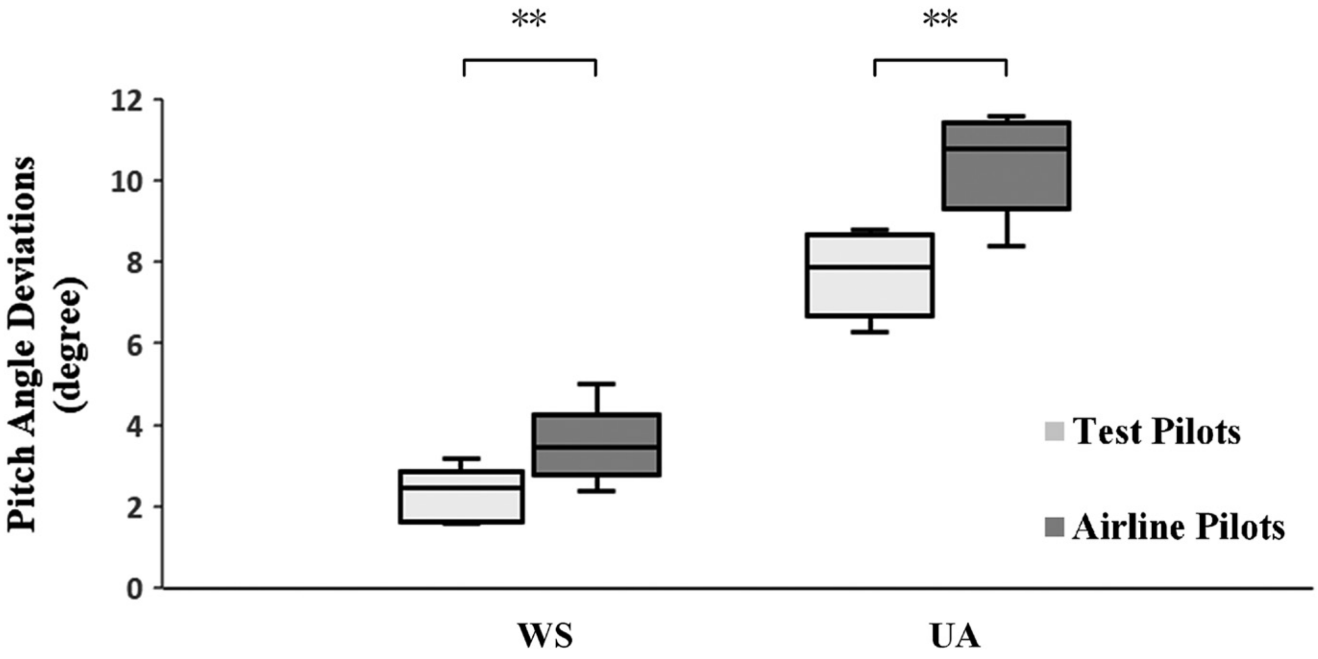

Fig. 3. Pitch angle deviation of test pilots and airline pilots in two flight tasks, which were Encountering Wind Shear after lifting off (WS), Unreliable Airspeed during taking off (UA) (** $P < 0.01$).

difference [F (1, 14) = 10.08, $P < 0.01$] of the average pitch angle deviation between the test pilots and airline pilots (test pilots: 2.3° ± 0.64; airline pilots: 3.6° ± 0.89).

Unreliable Airspeed During Taking Off

In this scenario, we also paid attention to the reaction time to the unexpected event, which was from the pitot blockage occurred to the pilots leveling the airplane, and the average pitch angle deviation during the climb. The mean reaction time of test pilots to unreliable airspeed was 23.38 s (SD = 6.46), and for airline pilots, the average reaction time was 42.63 s (SD = 6.89). The difference between them was significant [F (1, 14) = 33.269, $P < 0.01$]. Otherwise, the difference of average pitch angel deviation between two types of pilots was also significant [F (1, 14) = 23.353, $P < 0.01$], test pilots (M = 7.74, SD = 1.03) were more precise in manipulating than airline pilots (M = 10.40, SD = 1.17). The pitch angel deviations of two types of pilots in WS and UA are shown in **Fig. 3**.

Stabilizer Trim Failure

In this task, we focused on two indicators, the reaction time to warning 'STAB FAULT', which was the period from when the alert took place to when STAB TRIM was pressed, and the landing speed deviation, which was equal to the difference between actual landing speed and VRef Full +15kts. The mean reaction time of airline pilots (M = 4.86 s, SD = 0.58) was slightly shorter than test pilots (M = 5.54 s, SD = 0.77), however, the difference was insignificant [F (1, 14) = 3.927, $P = 0.068$]. Moreover, statistically different landing speed deviations was found (test pilots: 3.50 knots ± 1.60; airline pilots: 3.63 knots ± 1.92), but of no practical significance [F (1, 14) = 0.020, $P = 0.890$], as shown in **Fig. 4**.

DISCUSSION

In this study, three abnormal scenes were carried out in an A320 D-level flight simulator, one is where the aircraft was in an unexpected ambient environment (WS), and the other two were system failures (UA, STF). Considering the eye movement data, it seemed that the uncertainty of environmental change would give rise to more severe pressure and mental workload influence than system failure with minimum eye blink rate and maximum fixation duration in wind shear condition both for test pilots and airline pilots. This result is reasonable, according to findings of National Research Council, when encountering severe weather conditions, such as low altitude wind shear, the aircraft may deviate from the normal trajectory or even lose stability rapidly, which poses a great safety risk for flight, especially in take-off and landing phase.[4] Comparatively, the failure of a single system would not lead to disastrous consequences, as the important systems on the aircraft have redundant design.[1] Even for the failure or jamming of one control surface, pilots can still manipulate the aircraft through other controls.

This also explains why strong crosswind and natural icing test flights are the most challenging high-risk test subjects in the certification progress.

In the scenario of WS, pilots could recognize the unexpected event immediately based on warning information, and only needed to increase the pitch angle and maintain it at a constant degree until eliminating wind shear (2000 ft). The duration of the scene was relatively short, and the angle manipulating requirement was fixed. However, in the scenario of UA, the pilot was required to identify the failure on his own initiative and adjust the pitch angles at different flight levels until

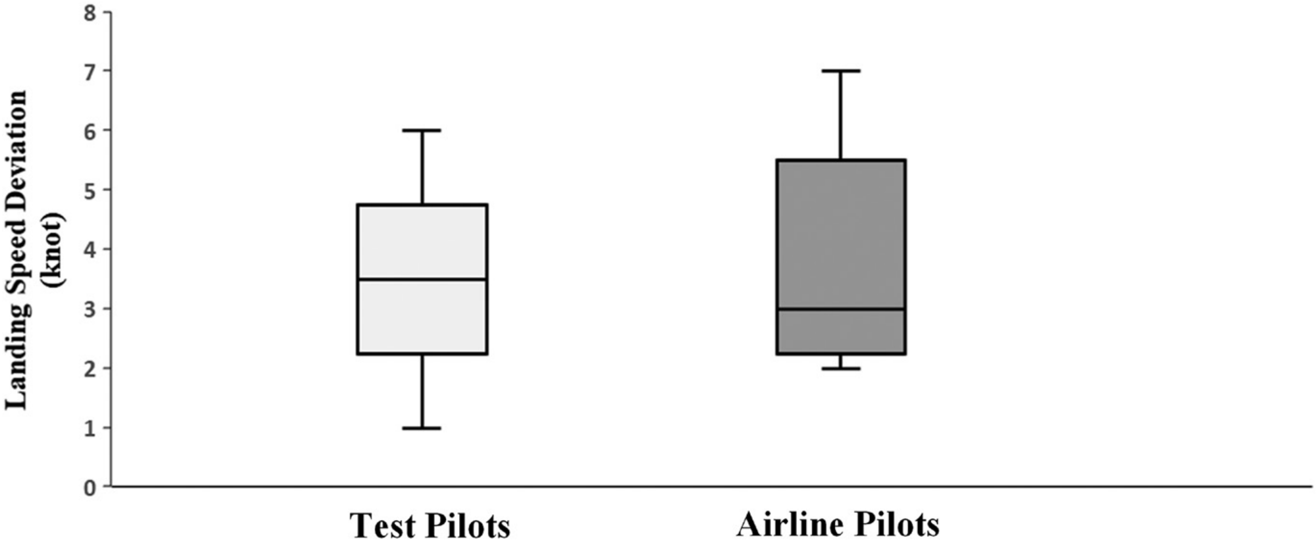

Fig. 4. Landing Speed deviation of test pilots and airline pilots in Stabilizer Trim Failure.

climbing to 20,000 ft. The operation was more complicated and took longer, resulting in a larger deviation of control accuracy. Specifically, when encountering an unexpected environment condition (WS), the test pilots performed significantly better than airline pilots in terms of aircraft manipulation. It is undeniable that all pilots who can serve as captains have undergone strict training and assessment and are capable of ensuring flight safety. However, since the daily work of test pilots is to be exposed to uncertain surroundings and medium or high-risk conditions, identifying potential inconspicuous risks is a rule-based behavior for them, while for ordinary pilots, it may still be in the knowledge-based level according to Rasmussen's human performance model.[19] Therefore, the disposal measures for some unexpected events, based on configurations of control flight elements (attitude and power), might have become the content of test pilot's long-term working memory.[22] This allows them to achieve more sophisticated and accurate operations according to the standard procedures, which not only can bring better safety margin, but also make the aircraft operate in an economical state. The indicators in the flight manual are the relatively optimal values calculated based on parameters such as aircraft weight and center of gravity. Such accurate control is particularly evident in the long-term operation requirements, i.e., manual angle control. Nevertheless, for both test pilots and airline pilots, the precise control requirements of single point could be well met, for instance, the landing speed. Orlady assumed as takeoff and landing phases are the most common scenes in flight training, qualified pilots have an intimate knowledge of the parameters that might affect safety at critical flight moment and could implement them at the optimal time.[16]

Otherwise, reaction time was also selected as pilot's performance indicator in this study, as one hallmark of expertise is the speed at which experts work, as Masunaga and Horn suggested.[13] In two scenarios with warnings (WS, STF), the differences in reaction time of test pilots and airline pilots were insignificant. In detail, in wind shear condition, the average reaction time was shortest both for two types of pilots, and test pilots was significantly faster than that of airline pilots. Surprisingly, when stabilizer trim fault occurred, the test pilots' reaction was slightly slower. The phenomenon might be a result of any of the following reasons. Firstly, in case of warning, a corresponding alert tone and a red highlight information would appear in flight deck to ensure the immediate pilot's awareness and immediate action.[18] Therefore, all the well-trained pilots were able to respond in time in scenarios with alarm prompts. Secondly, when encountering wind shear, a flashing light in the primary flight display and speech warning would emerge, which could grab the pilot's attention more quickly than the warning only display in crew alerting system with the unified auditory indication. Single tones provide no information as such, so it is not surprising that a speech warning system would out-perform such meager nonverbal signals.[6] Smith et al. also found speech warnings provided an advantage in reaction time and response accuracy over auditory icon warnings.[21] Thus, speech warning should be used in the most common emergency situation, such as stall, because it requires very little cognitive processing and has the ability to alert and to inform the nature of the hazard.[15] Thirdly, although test pilots were slower in stabilizer trim fault condition, their reaction time was in an acceptable range and still could control the aircraft appropriately. Moreover, the trim failure was a kind of appearance, which would be triggered by a variety of reasons. The test pilot might spend more time exploring the root fault to enhance the situation awareness, rather than simply follow the flight manual.

Conversely, without alarm, pilots need to identify the differences of numerical information on displays and determine whether the flight parameters were reliable by themselves, resulting in a sharp increase of the reaction time in scenario UA. Stanton and Edworthy found an auditory warning would lead to a quicker response than visual stimulus.[23] However, the performance of test pilots was significantly better than that of ordinary pilots, and they could detect the occurrence of unreliable airspeed in a relatively short time. In real flying, the airline pilots might perform even less efficiently, since test pilots encounter sudden changes, unusual attitudes, and aircraft performance extremes more often in research and development test flight or certification test flight in new or modified aircrafts. The states and configurations of such aircrafts are usually not as stable as that of aircrafts normally operated by airlines. Furthermore, test pilots fly close to the safety boundaries more frequently to test and validate the performance and characteristics of the aircraft in test flight. Some flight test scenarios with rare system faults or with rigorous surroundings might never be met by airline pilots.[25] This kind of boundary detection requires test pilots to be more circumspect and sensitive to changes in aircraft status and enables them to respond more quickly to unannounced faults, thus effectively improving the safety level of test flight.

By comparing the physiological reactions and performance of test pilots and airline pilots when facing the unexpected events in this research, there are three aspects would be enhanced in airline pilots training. First, although a great quantity of automated equipment could be used, the importance of manual flight should still be emphasized, which allows pilots to precisely control the aircraft without the help of automation system for a long period. Secondly, increasing the training of flight in unexpected ambient conditions, especially during takeoff and landing phase, because changes in the environment are more likely to cause pilots to startle, and such critical flight phases require them to respond more timely. Last, but not least, improving the reaction to the unannounced fault status, would allow pilots to cope with deviations more calmly and reduce safety risks.

ACKNOWLEDGMENTS

Yiyuan Zheng, Ph.D., Professor Yuwen Jie, Ph.D., and Zhiqiang Zhao, Test Pilot, Shanghai Aircraft Airworthiness Certification Center of Civil Aviation Administration of China, Shanghai, P.R. China; and Profesor Shan Fu and Yanyu Lu, Shanghai Jiao Tong University, Shanghai, P.R. China.

REFERENCES

1. Boglietti A, Cavagnino A, Tenconi A, Vaschetto S, editors. The safety critical electric machines and drives in the more electric aircraft: A survey. 35th Annual Conference of IEEE Industrial Electronics. Piscataway (NJ): IEEE; 2009: IEEE.
2. Casner SM, Geven RW, Williams KT. The effectiveness of airline pilot training for abnormal events. Hum Factors. 2013; 55(3):477–485.
3. Corda S. Introduction to aerospace engineering with a flight test perspective. New York (NY): John Wiley & Sons; 2017.
4. National Research Council. Low-altitude wind shear and its hazard to aviation. Washington (DC): National Academies Press; 1983.
5. Culick FEC. The Wright brothers: first aeronautical engineers and test pilots. AIAA J. 2003; 41(6):985–1006.
6. Edworthy J, Hellier E. Auditory warnings in noisy environments. Noise Health. 2000; 2(6):27–40.
7. Endsley MR. Situation awareness. In: Sarvendy G, Karwowski W, editors. Handbook of human factors and ergonomics, 5th ed. New York (NY): John Wiley & Sons; 2021:434–455.
8. FAA. Airplane Simulator Qualification. Washington (DC): FAA; 1991: 120–140B.
9. Gray W. Handling qualities evaluation at the USAF Test Pilot School. AIAA Atmospheric Flight Mechanics Conference; August 10–13, 2009; Chicago, IL; Reston (VA) AIAA; 2009: 6317.
10. Jacobs DM, Morice AH, Camachon C. Montagne G. Eye position affects flight altitude in visual approach to landing independent of level of expertise of pilot. PLoS One. 2018; 13(5):e0197585.
11. Kasarskis P, Stehwien J, Hickox J, Aretz A, Wickens C, editors. Comparison of expert and novice scan behaviors during VFR flight. Proceedings of the 11th International Symposium on Aviation Psychology; March 5–8, 2001; Columbus, OH; Columbus (OH): Ohio State University; 2001:6.
12. Landman A, Groen EL, Van Paassen M, Bronkhorst AW, Mulder M. Dealing with unexpected events on the flight deck: a conceptual model of startle and surprise. Hum Factors. 2017; 59(8):1161–1172.
13. Masunaga H, Horn J. Characterizing mature human intelligence: Expertise development. Learn Individ Differ. 2000; 12(1):5–33.
14. Moir I, Seabridge A. Aircraft Systems: Mechanical, electrical, and avionics subsystems integration. New York (NY): John Wiley & Sons; 2011.
15. Noyes J, Hellier E, Edworthy J. Speech warnings: a review. Theor Issues Ergon Sci. 2006; 7(6):551–571.
16. Orlady LM. Airline pilot training today and tomorrow. Crew resource management. Cambridge (MA): Academic Press; 2010:469–491.
17. Perkins CD. Stability and Control: Flight Testing. Amsterdam: Elsevier; 2014.
18. Pritchett AR. Reviewing the role of cockpit alerting systems. Human Factors and Aerospace Safety: An International Journal: No. 1. London: Taylor and Francis; 2017:5–38.
19. Rasmussen J. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Trans Syst Man Cybern. 1983; SMC-13(3):257–266.
20. Sgobba T. B-737 MAX and the crash of the regulatory system. Erratum published: https://doi.org/10.1016/j.jsse.2021.01.003. J Space Saf Eng. 2019; 6(4):299–303.
21. Smith, Sean E., Karen L. Stephan, and Simon P. Parker. Auditory warnings in the military cockpit: a preliminary evaluation of potential sound types. Edinburgh (Australia): Defence Science And Technology Organisation, Air Operations Division; 2004.
22. Sohn YW, Doane SM. Memory processes of flight situation awareness: Interactive roles of working memory capacity, long-term working memory, and expertise. Hum Factors. 2004; 46(3):461–475.
23. Stanton NA, Edworthy J. Auditory warnings and displays: An overview. In: Stanton NA, Edworthy J, editors. Human Factors in Auditory Warnings. Oxford: Routledge; 2019:3–30.
24. Stanton NA, Plant KL, Roberts AP, Allison CK. Use of Highways in the Sky and a virtual pad for landing Head Up Display symbology to enable improved helicopter pilots situation awareness and workload in degraded visual conditions. Ergonomics. 2019; 62(2):255–267.
25. Stoliker FN. Introduction to flight test engineering. Neuilly-sur-Seine (France): NATO-AGARD; 1995.
26. Tischler MB. System identification methods for aircraft flight control development and validation. Oxford: Routledge; 2018.
27. Tsang PS. Assessing cognitive aging in piloting. Human Error in Aviation. Oxford: Routledge; 2017:425–464.
28. Wise JA, Hopkin VD, editors. Human factors in certification. Boca Raton (FL): CRC Press; 2000.