

Q–Q Test for Pilot Evaluation of Ergonomics in Aircraft Cockpit

Xiaochao Guo, Yanyan Wang, Qingfeng Liu and Duanqin Xiong

Abstract The Q–Q test (Qualitative and Quantitative test) was set up in the context of pilot evaluation of ergonomics in an iterative process of aircraft cockpit design to facilitate and optimize human-centred design with quantitative data of pilot use in simulated flight. First, a subjective rating scale was built on degrees of acceptability and satisfaction referring to Cooper–Harper Scale. Second, the quantitative measurements of user performances were simultaneously introduced, and the difference between the optimal design and the tested design on the basis of user performances and parameters of designs would support iterative design or redesign. Third, the consensus of pilot sample or subgroups was integrated by the Delphi technique to determine which problems should be solved on time after communication with stakeholders. One example was also discussed.

Keywords Q–Q test (Qualitative and quantitative test) • Pilot evaluation • Aircraft cockpit • Ergonomics • Human factors • Cooper–harper scale • Delphi technique

1 Introduction

The Cooper–Harper scale is a popular tool for pilot evaluation in handling qualities of piloted aircraft [1–3], in which pilot ratings determined the handling characteristics of the aircraft in terms of qualitative degrees of suitability and Levels (1-Satisfactory, 2-Acceptable, or 3-Controllable) with comments indicating what deficiencies had been discovered. However, quantitative data may be desired by designers in pilot evaluation of ergonomics because results of user performance test (UPT) will be very helpful to improve in human-centred design [4, 5]. So the Q–Q

X. Guo (✉) · Y. Wang · Q. Liu · D. Xiong (✉)
Institute of Aviation Medicine PLAAF, Beijing 100142, China
e-mail: guoxiaochao1965@163.com

D. Xiong
e-mail: xdqin123@163.com

test (Qualitative and Quantitative test) method was established in the context of pilot evaluation of ergonomics in an iterative process of aircraft cockpit design.

2 Method

2.1 Pilot Ratings as Qualitative Probe in Q-Q Test

A subjective rating scale was built on degrees of acceptability and satisfaction with 10 points in 4 levels similar to Cooper–Harper scale. In Q–Q test, pilot must judge according to acceptability or satisfaction and give his comments about What and How aiming at the points within levels with deficiencies (Fig. 1).

It is noted that scored 6 or 10 was not allowed if pilot rated a design in Level 2 which means “moderate acceptable” and vice versa (Fig. 1a), for instance.

2.2 Quantitative Measurement of User Performance in Q-Q Test

User performances were measured in two conditions, i.e., the optimal design versus the tested design in the context of flight use. The difference of $X_{\text{optimal}} - X_{\text{tested}}$ will indicate the deviation in the tested design and the range of improvement in the iterative design based on the user requirements or pilot needs of flight use.

2.3 Participants of Pilots in Q-Q Test

Pilot sampling is closely related to the purposes of pilot evaluation in Q–Q test and the stages of production design on the basis of anthropometric factors, flight

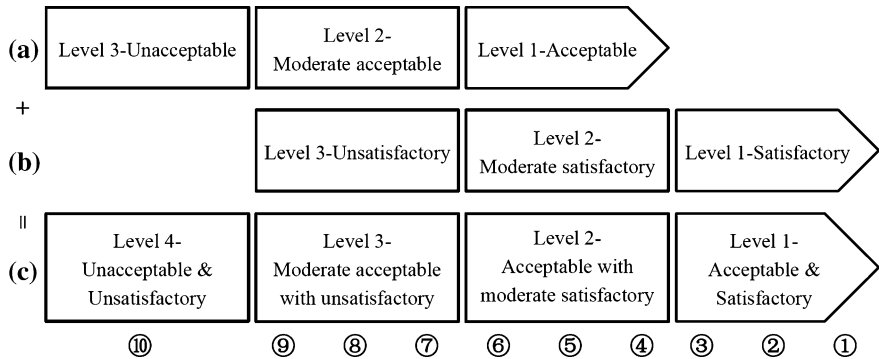


Fig. 1 Pilot rating scales in Q-Q test

experiences, and the like. Pilots could be selected referring to GJB4856 and GB/T12985 or ISO7250-3 [6–8] if design range was specified as smallest to largest 5th–95th percentile or 1st–99th value in stature (body height) and body mass (weight) according to anthropometric requirements. The size of pilot sample could be confirmed by GB/T23699 [9] at least 3 persons in screening test or 7 persons in detailed test for each subgroup.

2.4 Integration of Data in Q-Q Test

For qualitative data, integration of pilot evaluations was accomplished by the Delphi technique [10]. The quantitative data were analysed in statistics.

3 Results

An example was presented in Table 1, in which there were 5 pilots in each subgroup of smallest or largest percentile, i.e., 5th (P_5) or 95th (P_{95}) percentile of pilot stature according to GJB4856 during screening test of cockpit design mentioned by GB/T23699/ISO15537 above.

The meanings of levels in Table 1 were shown as in Fig. 1a for acceptability.

4 Discussions

4.1 Pilot Sampling as User Representatives

Participants of pilots took a key role in Q-Q test because the pilot sample was representative of user population such as that in GJB4856. So pilot sampling should be either from the intended users or as similar as possible to the population. In stratified sample, participants may be representatives of 5th, 50th, 95th subgroups in GJB4856, the valid cases of sample were generally not less than 5 persons in user performance test (UPT) for statistical power (Table 1), and the differences of gross sample, i.e., whole of 5th, 50th, 95th subgroups to user population should be analysed statistically.

Considering that the overweight trends of pilots show an obvious rise with time and age [11] and some human dimensions seemed to augment generally in view of body growth [12], the analysis of participant characteristics of pilot sample (Table 2) could help the designers of aircraft cockpit to make correct decision in the iterative process of design.

Table 1 Results of Q-Q test as an exemplar in screening test of cockpit design

Pilot subgroup	Index	Vote in the 1st round of Delphi			Vote in the 3rd round of Delphi			Pilot comments in consensus	Value to change from UPT
		Level 1	Level 2	Level 3	Level 1	Level 2	Level 3		
P_5	1	5			5				
	2	3	2			5		More retrusive	66 mm
	3	4	1		5				
	4	4	1		5				
	5	1	3	1		5		Much higher	
	6			5			5	More retrusive	80 mm
	7	1	3	1		5		Much higher	
	8		1	4			5	More retrusive	188 mm
	9	5			5				
	10	1	1	3		5		More spacious	
$P_{9.5}$	1	5			5				
	2	2	2	1		5		More retrusive	52 mm
	3	4	1		5				
	4	4	1		5				
	5	4	1		5				
	6		2	3			5	More retrusive	35 mm
	7	5			5				
	8		2	3			5	More retrusive	159 mm
	9	5			5				
	10	1	1	3			5	More spacious	

Table 2 Participant characteristics of pilot sample in a Q–Q test (in mm)

Anthropometric codes and items in GJB4856		Users in GJB4856		Pilots in Table 1		T-test
		\bar{X}	s	\bar{X}	s	
2.1	Stature (body height)	1705.0	36.6	1703.4	64.1	
2.88	Length of upper extremity	727.5	22.7	753.3	30.6	$P < 0.05$
3.1	Sitting height	924.3	21.9	909.5	27.0	
3.3	Eye height II, sitting	810.0	21.4	750.9	176.1	
3.25	Lower extremity length, sitting	979.5	28.5	997.3	47.7	
3.15	Forearm-plus-hand length	453.1	14.4	459.7	32.2	
5.1	Body mass (weight) Kg	68.0	7.6	76.7	10.0	$P < 0.05$

4.2 The Consensus by the Delphi Technique

4.2.1 Personal Judgement by Individual Pilot

Participants of pilots must have been trained prior to the Q–Q test to be familiar with the tested design and its requirement of use in simulated flight. Each of pilots must first manipulate prototype of the design independently, then give pilot rating on one’s own.

Pilot ratings were given by two steps. The gross judgement was made at first such as Acceptable, Moderate acceptable, or Unacceptable (Fig. 1), a rating score was marked in succession within the Level.

It suggested that the tested design had some deficiencies if a pilot judged the design as “not acceptable” or “not satisfactory,” the pilot must point out what deficiencies to be discovered and how to improve. Discount without cause should not be allowed by testers in principle.

4.2.2 Group Judgement by Pilot Sample

In stratified sample, each of subgroup such as P_5 , P_{50} , P_{95} pilots in Table 1 should make their group decision after round by round with the Delphi technique. The consensus perhaps emerged or few exceptions probably stuck to which is no matter for Q–Q test. The group comments of pilots must also present as in Table 1 if pilot sample judge the tested design as “not acceptable” or “not satisfactory” mentioned above.

4.3 Communication Within Stakeholders

The conclusions in Q–Q test were classified into “improvement suggested,” “improvement warranted,” “improvement required,” and “improvement mandatory” on

the basis of group judgement by pilots after communication of ergonomics specialists (the tester group as an example), participants of pilots (user representatives) to the designers of aircraft cockpit while product managers attended the conference both from users and manufacturers.

The communication should validate the group judgement of pilot sample or clarify further about some exceptions, and facilitate the improvement of the tested design after trade-off with an eye to technology, cost, and timing.

4.4 UPT Supporting the Design Improvement

It should be said that the degraded pilot ratings and user performances could be derived from the tested design in Q–Q test such as some parameters unsuitable for users or some of indexes mismatch to use requirement. To improve the tested design therefore aimed to correct the distinct departure.

A perfect design, i.e., an optimal design for users may be rated as grade 1, 2, or 3 in Level 1 as shown in Fig. 2, and participants of pilots could suggest few of improvement, although deficiencies in design were negligible or mildly unpleasant. A deviation from grade 1–3 was ignorable.

For example, if the tested design was graded as 5 in Level 2 by group judgement of pilot sample with comments of improvement warranted (Table 1), the parameters of the iterative design were warranted to adjust backwards about 52–66 mm. The value to change from UPT equated to the difference between the optimal design and the tested design on the basis of law of causation relative to parameters of designs and user performances.

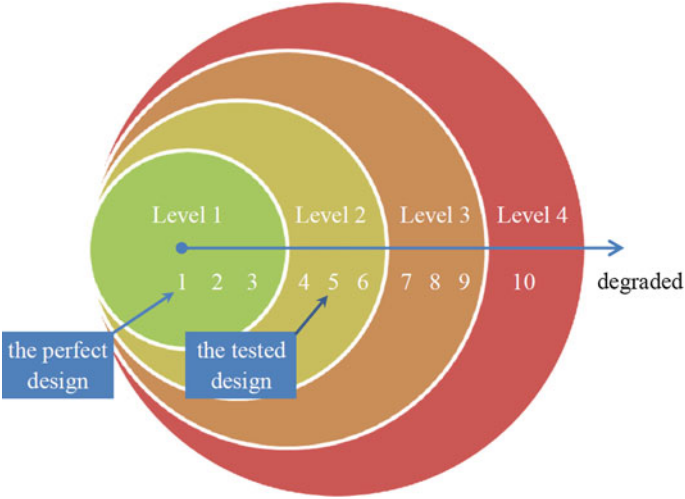


Fig. 2 To improve the tested design into Level 1

The values listed in Table 1 just were the differences of $X_{\text{optimal}} - X_{\text{tested}}$ as mentioned above, which could be reported in mean, range, or percentile as a result of data analysis in UPT.

5 Conclusions

The Q–Q test had been used in several types of aircraft for pilot evaluation of ergonomics, made better in facilitating human-centred design and supporting iterative design by scientific and quantitative data from users.

References

1. MIL-STD-1797A (2004) Flying qualities of piloted aircraft. U.S. Department of Defense
2. AC25-7C (2012) Flight Test guide for certification of transport category airplanes. Department of Transportation Federal Aviation Administration, USA
3. GJB2874-1997 Flying qualities standard for airplane with fly-by-wire control system. Military Standard Publishing House of the General Armament Department, Beijing
4. ISO9241-210:2010 Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems. International Organization for Standardization, Switzerland
5. GB/T21051-2007/ISO/TR16982:2002. Ergonomics of human-system interaction—Usability methods supporting human-centred design. China Standard Press, Beijing
6. GJB 4856-2003 Human dimensions of Chinese male pilot population. Military Standard Publishing House of the General Armament Department, Beijing
7. GB/T12985-1991 General rules of using percentiles of the body dimensions for products design. China Standard Press, Beijing
8. ISO7250-3:2015 Basic human body measurements for technological design—Part 3: Worldwide and regional design ranges for use in ISO product standards. Switzerland: International Organization for Standardization
9. GB/T23699-2009/ISO15537:2004 Principles for selecting and using test persons for testing anthropometric aspects of industrial products and designs. China Standard Press, Beijing
10. Linstone Harold A, Turoff Murray (2011) Delphi: a brief look backward and forward. *Technol Forecast Soc Chang* 78(9):1712–1719
11. Guo X et al (2015) The body mass index of taikonaut candidates. In: Long S, Dhillon BS (eds) *Proceedings of the 14th international conference on man–machine–environment system engineering*. Springer, Berlin, pp 129–134
12. Guo X et al (2015) Human dimensions of Chinese male pilots in sitting posture and some changes. In: Long S, Dhillon BS (eds) *Proceedings of the 14th international conference on man–machine–environment system engineering*. Springer, Berlin, pp 135–140

Man-Machine-Environment System Engineering
Proceedings of the 16th International Conference on
MMESE

Long, S.; Dhillon, B.S. (Eds.)

2016, XXII, 635 p. 170 illus., Hardcover

ISBN: 978-981-10-2322-4