The QTG - Some Perspectives
(Past, Present and Future)

Presented to the FSEMC Exploratory Meeting
Boeing Flight Training Facility, Crawley, UK
26th - 27th March 2015

Malcolm I Blackwood, B.Tech., C.Eng., MRAeS
Independent Flight Simulation Consultant
Canterbury, UK

Brief Curriculum Vitae

Malcolm Blackwood gained an Honours Degree in Aeronautical Engineering and Design from Loughborough University of Technology in 1978. Since then he has worked continuously in the flight simulation industry for Redifon Simulation (later Rediffusion Simulation) in Crawley and Singer-Link in Binghamton, USA, culminating in a return to Rediffusion, when in 1988 he was made Flight Section Manager, responsible for all flight modelling and QTG production. Since 1994 he has been an independent industry consultant and engineer.

Over the last 36 years he has personally produced more than 30 QTG's (and been responsible for a much larger number) and, since becoming independent, has taken primary project management responsibility for nearly all update and upgrade work in which he has been involved.

He has taken part directly in several industry Working Groups, including:

- IATA Flight Simulator Design and Performance Data Requirements
- ICAO 9625 First Edition
- JAR-STD 1A

He is currently a member of the UK Royal Aeronautical Society Flight Simulation Group Committee, and is the primary author of the RAeS Aeroplane Flight Simulation Training Device Evaluation Handbook, Volume 1 (Objective Tests).

Introduction

Various parties, anxious to further the cause of FSTD testing usefulness, have embarked on an exploratory program with a view to massively changing or even removing the QTG as the primary document used for regulatory authority qualification purposes. This may be a valid step, but then again it may be misguided or mis-timed. This paper seeks to explore the issues involved, along with some historical perspective and possible options and ways forward to improve FSTD testing, and thereby assist the industry in reaching properly thought-out conclusions.

The invitation notes for the meeting were as follows:

> *"The intent of the meeting is to follow up from the success of last year's exploratory meeting held in Dallas, Texas and open up discussion to alternative means of validation of FSTDs- and share and promote new ideas and ways of optimizing regular testing and checking methods. We invite your organization to participate in this "idea generation" activity to ensure we consider ideas from across the industry. Topics will include alternate means to qualify devices in lieu or in addition to the QTG, validation data for new aircraft models that heavily rely on binary simulations, data requirements for maintenance devices.*
>
> *To kick off the discussion, the meeting will entertain the opinion about an over reliance on QTGs for annual validation for performance and handling of the FSTD and offer there may be better means to perform the same level of confidence of validating the simulations are true to the aircraft. The premise for this point is that the QTG is a valuable testing tool to validate the simulation software initially, but often does not lend itself to identify problems with the simulation software over a recurring period. However, as an industry we spend thousands of hours repeating tests - where the actual software has not changed one bit. The challenge is to develop a more optimized approach to validating the simulation software and focus more on the hardware in the loop, where there are regular areas that should be checked and where problems often occur. This approach could be used effectively over the lifecycle of the FSTD. The meeting will introduce the concept of using flight test data along with engineering data to objectively evaluate an FSTD from a holistic point and treat the FSTD as a system - a system that may require different testing methods.*"

Working from these notes, the issues, as stated, can be broken down into the following basic list:

1. New ideas and ways are being sought on of optimizing regular testing and checking methods.

2. There is perceived to be an over-reliance on QTGs for annual validation for performance and handling of the FSTD.

3. There may be better means to perform the same level of confidence of validating the simulations are true to the aircraft.

4. The QTG is a valuable testing tool to validate the simulation software initially, but often does not lend itself to identify problems with the simulation software over a recurring period.

5. A FSTD needs to be validated from a holistic point of view.

Modern QTG Breakdown

The following table describes the sections of a modern QTG:

Front Document	
1	Performance
2	Handling Qualities - including flight controls
3	Motion System
4	Visual System (including System Responses)
5	Sound System
6	Functions and Subjective Tests

Points of note:

1. A distinction needs to be made among these sections. To some it is as if the perceived issues with sections 1 and 2 only are being used to cast shadows on the usefulness of all the other sections. Certainly that is the perception of the FSEMC initiative that has been encountered on several occasions.

2. Since JAR-STD 1A Amendment 3 (1st July 2002), the Motion (Section 3) and Sound System (Section 5) tests have been much more comprehensive - aimed specifically at ascertaining the continuing functionality of the hardware. The Visual System (Section 4) has always had that functionality.

3. The intention, stated as of AC120-40B and further clarified in JAR-FSTD A and subsequent documents, was always that the tests be run in an integrated manner - this means the hardware, or as much of it is physically feasible, should be utilised when running tests. It should not be adequate to back-drive the controls merely for cosmetic reasons. THIS IS VITAL TO UNDERSTAND IF WE WANT TO MAKE THE TESTING MORE MEANINGFUL.

4. The issue of binary aircraft systems software may relieve the simulator manufacturer of much of the software design effort and also reduce the ability of operator maintenance staff to make alterations which may not be appropriate, but it does not alter the fact that it is the holistic interaction between the hardware and software that should really be under scrutiny in a QTG. This author's experience over the past 36 years is that, on the whole, operators' engineering staff do not generally make changes to aircraft system models unless there is a corresponding change to the aeroplane, or a new malfunction is requested; in either case there is usually zero effect on the QTG. Also, for non-Computer-Controlled Aeroplanes, the final tuning of the primary flight controls dynamics tests is almost always done by the simulator manufacturer flight controls engineer using the physical system in use on the FSTD; it is not automatically tuned merely by virtue of the fact that the aircraft manufacturer's models have been implemented in software.

5. Ultimately, the only truly valid way of testing that a FSTD is suitable for pilot training is for an experienced, competent pilot to be sure of its fidelity. This may mean the best way is to manually run the QTG tests - especially for Computer Controlled Aeroplanes. Initially this will be time-consuming and, depending on who does them, not necessarily accurate, so automatic running is an expediency. However it was never the intention that the industry expend all its effort on getting automatic tests to run while neglecting everything else.

6. Is it the case that running a test that always passes renders that test invalid? Are we not getting what we want if this happens? Do we stop doing a pre-flight walk-around of an aircraft because we haven't found any problem during the last walk-arounds, or do we stop training and testing a pilot because he/she has always previously passed their proficiency test? If we change the QTG tests, what would then happen? It is quite possible that we could end up with a different set of tests that almost always passes, thus achieving very little.

7. Much modern FSTD hardware contains significant portions which having nothing to do with that fitted in the aircraft. This hardware will not have 'aircraft' connections, often relying on fibre-optics or ethernet. This is another good reason for testing in an integrated manner.

8. From the above table, it is obvious that Sections 3, 4, 5 and the Functions and Subjective Tests require the hardware to be available and working as part of the simulator as a whole. Why would Sections 1 and 2 be exceptions to this rule?

9. If there is an over-reliance on QTG's, who is responsible for this and why is it happening?

10. We should not underestimate the power of confusion to wreak havoc amongst those who have less understanding or knowledge. The industry has been very successful in spreading its wings far and wide, but with that has come a much greater need for knowledge about the process. Most of the 'newcomers' rely on the more established members of the industry club to guide them down the correct route. The problem is that some of the established club members have lost track themselves of that route.

11. One other thing to remember is that the creators of ICAO 9625 Edition 3 have been very evangelistic in their portrayal of their new methodology, yet in 5½ years very little headway has been made. ICAO 9625 is the definitive document that has been adopted by the world's NAA's. As an industry, we may need to be careful about the message we would be sending if we embark on another round of 'QTG negotiations' before the previous round has been implemented.

A Summary of QTG Use, Misuse and Misunderstandings

A list of fundamental QTG issues which are sometimes not appreciated might look as follows:

- Importance and Value of Manual Tests/Pilot-in-the-Loop.
- Importance of Hardware in the Loop.
- Recognition that Sections 1 and 2 are not the complete QTG!
- There is arguably an over-reliance, not so much on the QTG as a concept, but on the degree of automation it now contains.

- Recognition that even a complete QTG is not thoroughly testing a FSTD for initial qualification.
- Standard operator maintenance and Daily Readiness should complement the QTG.
- Significance, or otherwise, of binary aircraft systems software.
- Significance, or otherwise, of non-aircraft hardware - face-plates with relatively simple electronics behind.

What does 'Integrated Testing' mean ?

> *"Historically, the tests provided in the QTG to support FSTD qualification have become increasingly fragmented. During the development of the ICAO Doc 9625 Manual of Criteria for the Qualification of Flight Simulators, 1993 by an RAeS Working Group, the following text was inserted:*
>
> *"It is not the intent, nor is it acceptable, to test each Flight Simulator subsystem independently. Overall Integrated Testing of the Flight Simulator should be accomplished to assure that the total Flight Simulator system meets the prescribed standards."*
>
> *This text was developed to ensure that the overall testing philosophy within a QTG fulfilled the original intent of validating the FSTD as a whole whether the testing was carried out automatically or manually.*
>
> *To ensure compliance with this intent, QTGs should contain explanatory material that clearly indicates how each test (or group of tests) is constructed and how the automatic test system is controlling the test e.g. which parameters are driven, free, locked and the use of closed and open loop drivers.*
>
> *A test procedure with explicit and detailed steps for completion of each test must also be provided. Such information should greatly assist with the review of a QTG that involves an understanding of how each test was constructed in addition to the checking of the actual results.*
>
> *A manual test procedure with explicit and detailed steps for completion of each test should also be provided."*
>
> [CS-FSTD(A), Initial Issue, AMC1 FSTD(A).300 Qualification basis, b.1.iii (pages 41-42)]

There are some suggestions for the way ahead further on, but it would seem that, for some organisations, the premise appears to be that the purpose of the QTG is to check the simulation software. This may often be how it is done in practice, but to say that the QTG concept was designed that way is absolutely not true, and was never the intent of the regulations. ***Note particularly the third paragraph above, which was also present in JAR-FSTD A, issued in 2008.*** The wording used by both the JAA and EASA states the requirement very clearly. If, in 2015, the industry is generating the Performance and Handling Qualities tests by merely exercising the software, then we have regressed 20 years, and QTG's are not being run correctly in accordance with the above text. Is the requirement being ignored (if so, by whom?), or merely not understood?

In many cases, the use of so-called Fokker tests to prove actual control forces and positions only for initial evaluations has been taken as licence to ignore the flight controls hardware and only use software model forces and positions in subsequent testing.

If the only way to do a test in an integrated manner is to run it manually with a pilot in the loop (e.g. Computer Controlled Aeroplanes), then perhaps this is how it should always be done.

One area that has surprised me to see it on so many FSTD's, is the way the 3 autopilot tests are being run. Most manufacturers seem to have gone to great lengths to make the required selections using software only so that the tests can be run automatically by their test system. This author does not believe this was the intention - these 3 tests are essentially manual tests, and a human operator should be physically selecting the switches on the MCP. With this kind of over-automation, it is understandable why some think that what is being done is pointless. They may be right, but that is because tests are not being done properly, and the Regulators do not always notice.

The Regulators have always left it up to the industry to come up with acceptable ways of fulfilling the requirements. From the very start, the rules allowed for an alternative means to qualify a FSTD, yet in 35 years no-one has come up with a viable alternative; so is the existing process really flawed? However, many Regulators have their own issues, as we will come to later.

It is also worth pointing out that the QTG was never designed to fully "validate" the FSTD, it is merely a benchmark - albeit an important one - that is used for future reference; a surrogate for running the much more comprehensive ATM's. However, it may be that there is an "over-reliance" by many manufacturers of the QTG, or at least specifically the automatic validation tests it contains, to initially "validate" the FSTD. But a QTG, even if it is 100% correct, cannot be considered to be a comprehensive evaluation of an completed FSTD, and it was never meant to be so. Further to this aspect, it is noteworthy that EASA AMC & GM to Part ORA (page 47) clearly states that the FSTD must be subject to annual Functions and Subjective tests fly-outs.

Historical Background to the QTG

What we now recognise as the QTG came about as a direct result of the FAA Advanced Simulation Plan, issued as Appendix H of FAR 121, June 1980. It was this document that heralded the birth of modern advanced simulators and pointed the way forward for eventual Zero Flight Time training.

The first FAA requirements document after 1980, AC121-14C, put the flesh on the bones of Zero Flight Time, and this document and its successors rapidly became a statement of work for the simulator manufacturers, who in turn began to bring pressure to bear, along with the airlines and operators, on the airframe manufacturers to produce flight test data from the aircraft as a basis for comparison in the simulator. The QTG was originally called an ATG (Approval Test Guide) and most, perhaps all, of the tests were run manually in the early days.

As time progressed, it became clear that these flight tests which were to be run in the simulator needed to be run automatically, primarily for the following reasons:

1. As a regulatory body requirement, the test had to be re-run at regular intervals for recurrent testing of the simulator. Repeatability was therefore of great importance.

2. When flying the simulator, pilots had difficulty matching the exact inputs of the flight test pilot for a given test. With early data packages containing flight test data that often consisted of some very complex flight control manipulations, this proved to be very time intensive and even then rarely gave accurate results.

3. Shorter simulator delivery timescales meant that less time was available to engineers to use

the complete simulator, hence simulator manufacturers needed to transfer much of their testing (including that performed for an ATG) to their in-house computing facilities, which were primarily set up for automatic rather than manual testing.

It may be that the third point here was of the greatest significance since, by definition, tests that have been developed on an in-house engineering computer have not been developed on an actual simulator and therefore some means must have been employed to 'simulate' certain aspects of the simulator itself, especially with regard to the pilot controls (both primary and secondary) and the cockpit indications. Originally, tests were driven using elevator, rudder, etc., instead of column, wheel, pedals because this could be done easily while the simulator was still in construction elsewhere in the factory.

This in itself did not mean that when the test was run in-house it could not be carried across to the simulator at all, or even with a low degree of confidence. Indeed for many ATG/QTG tests it makes little difference whether it is run on a fully integrated and functioning simulator or not, if one ignores any requirement to include the hardware or to make the test 'look good' from an aesthetic point of view in the cockpit. The issue here should not be whether the test works for the initial qualification - which is usually all that the FSTD manufacturer is concerned about - but whether the test is a valid and useful measure of the overall FSTD performance in the same manner throughout the service life of the device.

Regulatory Authority Issues

During the years after 1980, the FAA, with industry participation, built on their foundation. The table below tracks the progress of the FAA, because they have been the leaders of the industry, and at least in the early days theirs was the standard to meet. In simplified terms, the FAA regulatory requirements evolution progressed as follows:

Requirements	Year Enacted	Time Difference	Comment
AC121-14C	1980	-	
AC120-40	1983	3 years	Evolutionary Change
AC120-40A	1986	3 years	Evolutionary Change
AC120-40B	1991	5 years	Evolutionary Change
14 CFR Part 60	2008	17 years	Evolutionary Change
ICAO 9625 Edition 3	20??	?? years	Revolutionary Change

Note that in the early days all FFS manufacturers were involved in the requirements formulation, so they perhaps understood them better.

Other regulators followed the FAA lead, notably the UK, France, Canada, Australia and Japan, each making its own changes and modifications - but always increasing the testing versus the FAA. Note that the involvement of other countries increased dramatically after 1991, and also that there are many more types of device now.

Remember the variation of Moore's Law:
"Regulatory requirements never diminish!" See the table below for numbers of tests in each FAA document, for a FFS:

Test Section	AC120-40A	AC120-40B	CFR 14 Part 60
Taxi	2	2	2
Takeoff	7	7	9
Climb	3	3	4
Cruise/Descent	0	0	5
Stopping	2	4	6
Engines	0	2	2
Static Controls	7	8	10
Dynamic Controls	3	8	14
Longitudinal	17	22	18
Lateral/Directional	11	13	15
Landing	4	5	11
Ground Effects	0	1	1
Protection Functions	0	0	11
Totals	56	75	108

And remember, whichever way you turn, there is a lawyer standing there to guide your path.....
The QTG is signed off by the regulators and as such becomes a very significant legal document in any issues/disputes that may arise due to an inspector failing a simulator on a recurrent check or a post accident/incident investigation.

The other significant problem, which is major and is one that has been highlighted in the FSEMC concerns, is regulatory authority inconsistencies. This is a far greater issue than whether a QTG is doing its job or not, because whatever the industry decides to do about the QTG, this is a problem that will not be solved, and may well be exacerbated, by making significant changes to the qualification criteria.

A revolutionary change to the FSTD objective testing requirements is not compatible with regulatory authority consistency; we need education, stability of staffing and experience among the regulators to create a level playing field where many more countries trust each other's inspectors.

Regulators rarely have sufficient resources to keep the industry satisfied. It is unfortunate that money, specifically the lack of it, is the most consistent problem facing most regulatory authorities, and the industry suffers in varying degrees because of it.

Regulators often do not have sufficient knowledge or experience, and many are not given sufficient training to truly understand the intent of the rule, standard or evaluation process. We all know that some regulators have more experience than others, but the simple fact is that QTG's are not being done properly in many cases, and the regulators often do not notice. Many countries would like to have their own trained people rather than relying on regulators from other countries, but this takes time, effort and of course more money.

Some Possibilities

Notwithstanding the above comments, there may be some possibilities for ongoing workload reduction without reducing vigilance. However, be cautioned that 'new' or 'different' is not necessarily 'better', and that these are ideas, not necessarily recommendations. The rationale for including them here is that, taken in context, with certain caveats they may facilitate less onerous, yet more thorough integrated testing of the complete FSTD without sacrificing overall fidelity:

1. For certain tests and certain operators and certain FSTD's, enforce running bi-annually rather than annually. But - the criteria for this needs to be well defined and strictly enforced!

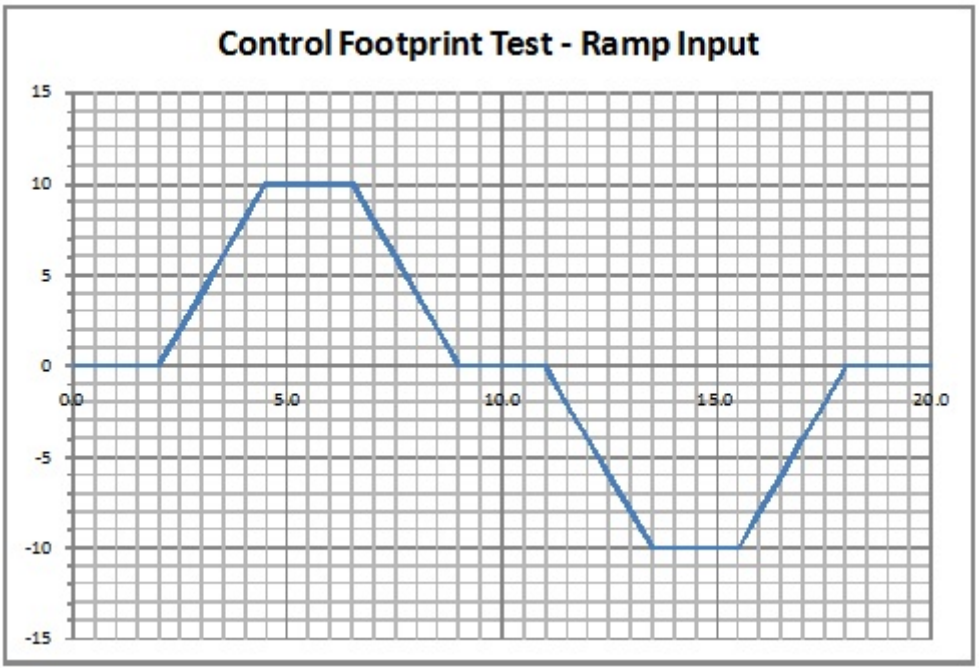
2. Combine certain tests, e.g. engine inoperative climbs with trims, roll response with roll overshoot, rudder response with Dutch roll, longitudinal trims with all other tests, ground acceleration and takeoff, ground acceleration and rejected takeoff, etc.

3. Reduce the number of CCA tests which have to be run in a degraded mode.

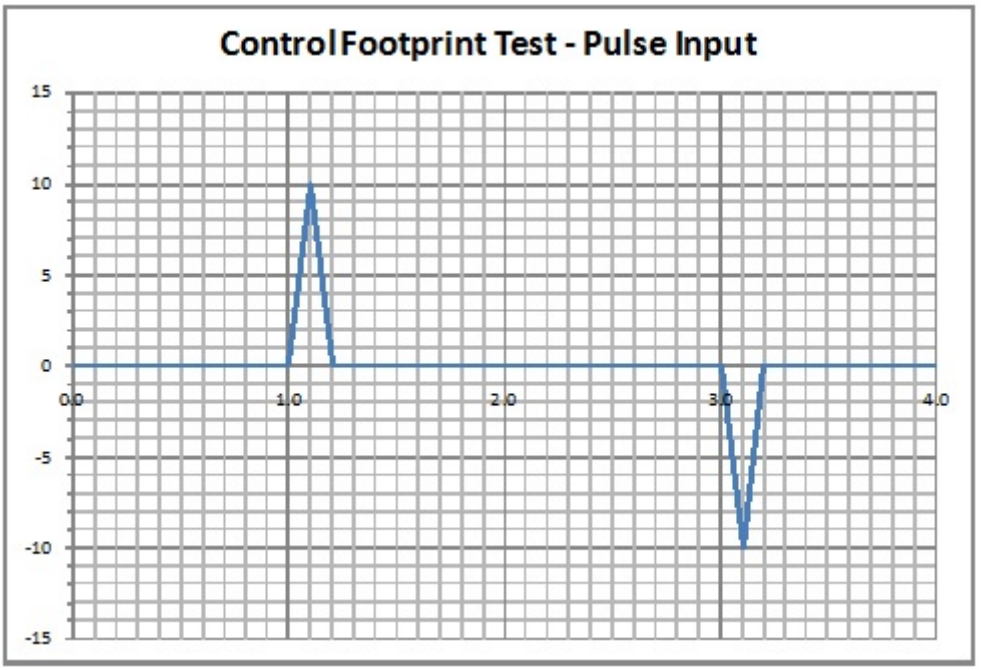
4. Allow a greater tolerance on some tests, e.g. recurrent sound frequency responses.

5. Force full QTG testing against flight test data only once every 3 years, with intervening testing, on a weekly or monthly basis, being done manually as part of a 'Daily Readiness' routine. The suggestion here would be to use simplified/stylised flying techniques which are not onerous or overly time-consuming to run. These could be run against those same simplified tests performed additionally to the full aircraft-based set, during the initial evaluation. There follow some possible examples:

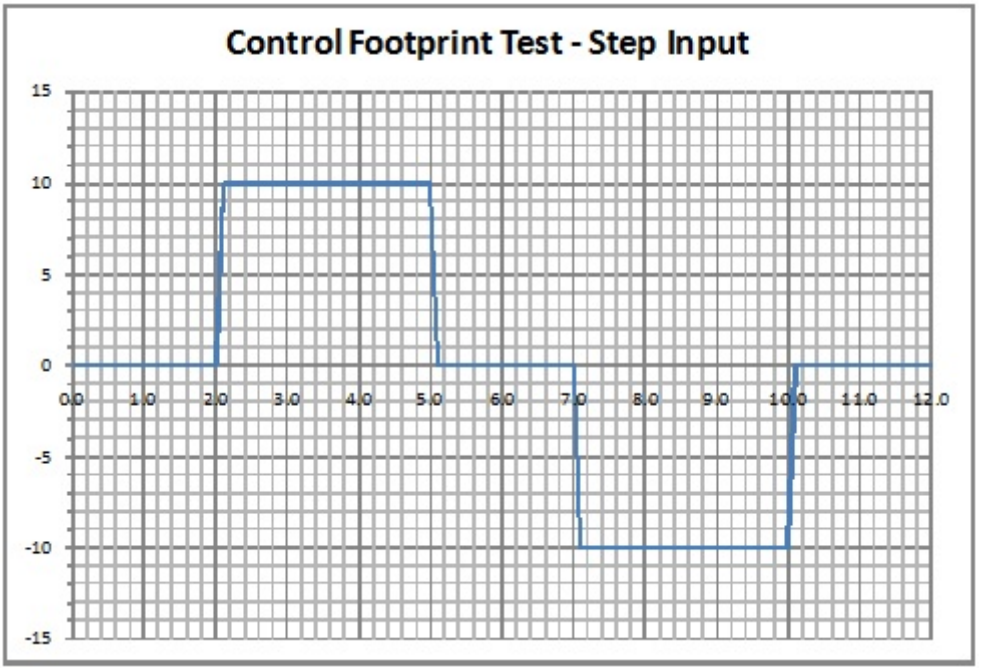
   A RAMP input, performed manually over, for example, a 20 second duration:



A PULSE input, performed manually over, for example, a 4 second duration:



A STEP input, performed manually over, for example, a 12 second duration:



The reason for using this type of profile is that ramp, pulse and step inputs are those used classically to test control systems. They are simple to apply and easy to repeat by either an experienced pilot or simulator technician and the response of the FSTD can be easily measured (by plotting the appropriate parameters). As part of a regular maintenance or daily (weekly? monthly?) readiness program, such tests could increase confidence that the FSTD is indeed performing as it should using inputs that pilots will be interested in. Note however, that I am not suggesting these

replace the Performance and Handling Qualities QTG tests, merely that they be used in an extended interim period between having to run the entire set of Performance and Handling tests. Of course, NAA support would need to be sought, as with all regulatory changes.

The QTG - Dinosaur, Tool or Aberration?

Some personal conclusions:

1. In the modern world, evolutionary change takes long enough for people and methods to stabilise, never mind revolutionary change! ICAO 9625 3rd Edition has shown that a revolutionary approach may not be well received by the industry.

2. Taking too long for implementation of any new methodology would mean there is a good chance the changes will be obsolete even before they are implemented.

3. Find out whether QTG tests are being run properly, especially with respect to hardware inclusion. Experience says that a very significant number are not, and that even extends to the flight controls static and dynamic tests.

4. The QTG has arguably been automated too much, to the point where some of the testing has ceased to be as useful as it could and should be.

5. Consider the role of the regulatory authorities - are they already confused enough?

6. Ensure that any reasons for ditching the QTG - or even part of it - are not seen as being merely financial/economic.

7. If QTG's, or even sections of them, are dispensed with, a great deal of education will have to be given to get the industry to be more consistent with whatever comes to replace it. Essentially, that will mean starting the educational process again from scratch.

8. Many (most?) simulator technicians/engineers working for operators do not believe that the QTG can be trusted to give them guidance as to what may be wrong at a given time, even though this may be untrue. Recently an engineer was contacted by an operator who said that they had just had a recurrent evaluation and that the inspector had stated that all the buffet amplitudes were far too high. The simulator had Level D, so the operator was asked whether they had run the QTG vibrations tests..... needless to say that was the first thing they then did!

9. The QTG, ATM, and manual testing are all mentioned in the rules and any attempt to change this will require a rule change that must include reasons for the change and suggested alternatives.

A Final Brief Summation

1. The industry has placed too much emphasis on getting automatic tests to work automatically, partly because there is considerable effort involved in even doing that.

2. In accomplishing the QTG tests, too little regard has been paid to the requirements of CS-FSTD(A) concerning integrated testing and perhaps also to the requirements of AMC & GM to Part ORA concerning annual functions and subjective testing.

3. Binary aircraft systems software may help in this area, in that it allows engineers to concentrate more on what really matters, and that is pilot confirmation that the FSTD as a whole conforms to the aeroplane. However, this should encompass greater emphasis on pilot interaction with the FSTD systems, include proper Manual Test Procedures.

4. Concerning the annual running of QTG tests, it is possible that some relief might be gained by reducing the frequency at which certain Performance and Handling Qualities tests need to be run, by substituting some rudimentary, yet effective, 'Daily/Weekly/Monthly Readiness' tests, especially where conventional flight controls testing does not apply (CCA).

5. However, the industry needs to beware of the inertia needed amongst the world's National Aviation Authorities to make these kind of changes, and a great deal of further education will be required whatever is decided.

The primary standard against which the FSTD should be measured must always be the aeroplane, and that means involvement by pilots.

Malcolm Blackwood, B.Tech., C.Eng., MRAeS
Canterbury, England, March 2015

References:

1. EASA CS-FSTD(A) Initial Issue, July 2012

2. EASA AMC & GM to PART ARA, April 2012

3. EASA AMC & GM to Part ORA, April 2012

4. JAA JAR FSTD-A, May 2008

5. FAA 14 CFR Part 60, May 2008