TEST ACCREDITATION – MINIMISING RISK AND ADDING VALUE

Ray Finch and Sherobhi Rajamantri

ABSTRACT

The testing of software is critical to the attainment of software quality. Reliable, meaningful test results can only be obtained by a test facility operating within a properly structured management system, using validated tests run by staff who understand the limitations of the tests. Test laboratory competency can be confirmed through independent, third party accreditation of the laboratory against the requirements of ISO/IEC 17025 *General requirements for the competence of testing and calibration laboratories*. Such accreditation supports software quality life cycle initiatives.

CONTENTS

1.   INTRODUCTION

The development of software which meets the users' expectations without exposing the developer, supplier or user to unacceptable commercial or legal risk is generally a substantial task. It requires significant planning, organising, monitoring and control activities to ensure that quality software is delivered on time. Testing of the software is a critical element of the process used by management to monitor and control the quality of developed software.

How then can management be sure that the quality of the testing is adequate? Also how can the user of the software be confident that the testing of the software has been effective? This paper proposes third party testing accreditation as the means for developing heightened confidence in testing for developers, suppliers and users.

2.   SOFTWARE RISK

The risks associated with software are varied and range from the trivial to the critical. The consequences of flawed software range from inconvenience or irritation to the user, eg computer games, to major property damage and injury or death to persons who are not even direct users of the software, eg software controlling trains or aircraft. Even the impact of the trivial example may not be acceptable to the developer or supplier due to the potential for loss of image, good name, etc. The impact of the serious cases is usually acceptable to nobody.

Fortunately software has behaved quite well and the serious situations are relatively few. Presumably this can be attributed to the efforts of developers to manage the development process in a "quality" environment and to ensure an appropriate level of testing. However software is becoming even more pervasive in our lives. It is being applied in situations which were apparently satisfactorily handled in other ways in the recent past, eg. engine management systems in automobiles. Software is being used to create market needs which barely existed even five years ago, eg. DVD, WAP, and its development is being driven by market forces which are continually driving down the development and life cycle times. Software lifecycles pressures are increasing. New product releases are associated with market pressure for reduced development times and resulting pressure on the testing. Testing is often seen as the reason for software release delays - " there were no problems until the testers got hold of it!". An admission of poor management!

In this environment the need for "quality" testing becomes even more critical. We all hope that problems in the software development will be picked up by the testing, but if the testing is not adequate, or is not effective, what are the implications for software quality? Software quality cannot be guaranteed without effective testing. Poor testing will not prevent poor quality software being released.

Testing is usually the last chance that software developers have to discover software glitches before the users find them. Without effective testing, software developers would have a hard time demonstrating that they had adequately addressed due diligence issues. Effective testing also provides increased assurance for the suppliers and the users. They get to sleep at nights, as well as the developers. Finally the image and status of the testing community is enhanced by quality testing and the perceived value of testing is increased.

3.   WHAT CONSTITUTES A QUALITY TEST RESULT ?

Some developers on a tight deadline might consider that if testing shows no defects in the software then this is a good result. Not only is this an unlikely scenario but it is most likely indicative of inadequate testing. It is suggested that quality testing has, at least, the following major characteristics:

- **CAPABILITY**

    It must be capable, ie able to properly exercise the software being tested and cover the range of client's requirements, ie. it must be appropriate, relevant, applicable. For

example a Philips screwdriver works well on a Philips screw, ie. it is capable, but cannot be used on a screw with a straight slot, ie it is not capable. This seems obvious, but test tools are being bought which are not capable of meeting the purchaser's expectations. All test methods and test tools need to be capable.

- **VALIDITY**

  The test must be valid, ie the results achieved must reflect reality, eg no false positives or negatives. The test must produce results which are meaningful. They must be correct. It also should not produce indeterminate results. If it does any of these the test is not valid and needs to be reviewed and modified or discarded. Indeterminate or false results create unnecessary effort. An invalid test is a waste of time and effort.

- **COMPETENCY**

  Testing must be competent, ie performed by competent personnel. Testers must know what they are doing. They must understand the purpose of the software being tested. They must understand its operation and implementation. If additional skills, eg financial, taxation, security, superannuation, electrical safety, are required, they need to be brought on board.

  Testers must be able to develop test plans and test cases and understand the limitations of these. They must be able to correctly apply test tools and understand their limitations. This applies regardless of whether in-house or commercial tools are being used. Testers must be able to identify and pursue suspect test results. They must be able to assess the impact of fixes on previous test results. Testers need an enquiring mind which does not readily accept conclusions without supporting evidence.

- **CONTROLLABILITY**

  Testing must be performed under controlled conditions, ie hardware and software configurations and operating states must be known to the testers and cannot be changed without the knowledge of the testers. Anything which has potential to impact the result must be controlled by the tester. This not only implies the software under test but also the hardware, operating system software, application software, test tools, etc, involved in the testing. Unauthorised changes to hardware or software must be excluded. Without control the test result cannot be treated as reliable. If the conditions under which testing is done are unknown the test results must be treated, at least, with care. When performing "live" testing, eg on the internet or over a network, it may not be possible to control the load on the system and this is something which would need to be considered when reviewing test results.

- **CHAIN OF EVIDENCE**

  Testing must be documented, ie the actual methods and test cases applied and the test results together with a full record of the hardware, software configurations and conditions under which the test was performed. There must be a record of the requirements of the software under test, the agreed test plan, what was tested, what tests were performed, how the tests were validated, the hardware and software configurations used, the test results and the criteria used to decide pass/fail conclusions, and the test personnel. Without this it cannot be demonstrated that effective testing was performed and there is no way that anyone can repeat the tests. Also if there is any legal challenge such evidence strengthens the tester's case.

- **REPEATABILITY AND REPRODUCIBILITY**

  A test is repeatable if the results obtained when a repeat test is performed under identical conditions are consistent with the original results. A test is reproducible if the results obtained when a repeat test is performed by another tester or laboratory are consistent

with the original results. Clearly if these conditions cannot be met the test is, at best, doubtful. Repeatability and reproducibility is sometimes thought to be an outcome from all the other characteristics of a quality test, but should not be treated as the indicator of quality software testing as it is theoretically possible to achieve repeatability and reproducibility with an invalid test.

- **IMPARTIALITY AND OBJECTIVITY**

    Test results must be impartial, ie not biased towards any particular result. Also the results must not be coloured by the feelings or opinions of the testers. It is possible for a test to be inherently partial by having a bias toward a particular outcome; this should have been detected during its validation and is one reason for doing validation. Of course tests can be biased by the testers due to pressure applied by the client, or by management.

    In the case of tests performed within a development organisation there is a need for management to allow the testing to be performed in a way which minimises any tester bias. In this situation management needs to ensure that it is receiving effective testing and therefore a true picture of the quality of the product. It needs to ensure that it is not inadvertently applying pressure to the test staff which results in less than ideal test outcomes. The use of development staff to test a product which they have developed is to be avoided as it is very difficult to ensure effective testing in these circumstances.

The achievement of testing which displays all of the above characteristics requires the application of considerable management and technical effort. This will not occur by chance. Quality testing cannot occur just by good luck; it does require good management. Reliable test results require a combination of an appropriate management system and technically competent personnel using appropriately validated test methods, controlled equipment, software and test tools, and controlled recording techniques, etc.

4.   HOW DO YOU ENSURE RELIABLE TEST RESULTS AND QUALITY TESTING?

The answer to how quality testing is achieved is simply that you perform the tests in a testing laboratory which meets the requirements of ISO/IEC 17025, *General requirements for the competence of testing and calibration laboratories.* This international standard specifies general requirements for the competence of testing laboratories regardless of the technical discipline in which they operate. It has been adopted as an Australian Standard, AS ISO/IEC 17025, without change.

This standard reflects the lessons learned from fifty five years of international experience in the operation of many thousands of testing laboratories in many technical disciplines. Operation of a testing facility in accordance with this standard ensures that the testing characteristics described above are being addressed.

 A common initial reaction from IT industry personnel is "Software testing is different!" or "There is no commonality between software testing and the more traditional test disciplines such as chemical testing, mechanical testing or electrical testing!". It may be a surprise to many to learn that the similarities far outweigh the differences. Any differences lie mainly in the technical test techniques applied. On the surface there are significant differences between software testing and chemical testing, for example, or between electrical testing and microbiological testing, yet ISO/IEC 17025 still applies, and works.

ISO/IEC 17025 works in the various areas of testing because it is a high level document which does not specify the detail of discipline specific test procedures; rather it defines a generic infrastructure which is essential for all testing regardless of discipline. It provides for specific interpretation in the various technical disciplines. As a result of differing needs between disciplines some aspects may have stronger impact in some disciplines than others, or may even be non-existent, eg. the requirement for sampling is important when testing food, or environmental waters, but may be irrelevant when performing a type test on a custom built electrical switchboard for a major building. These variations between technical disciplines are

addressed via discipline specific interpretive documents which are prepared by the various bodies which provide ISO/IEC 17025 accreditation services around the world.

5.   WHAT DOES ISO/IEC 17025 REQUIRE ?

ISO/IEC 17025 specifies the requirements for testing laboratories under two major headings, viz, Management Requirements and Technical Requirements. The clause headings are listed below in table 1.

Clause 4: Management Requirements		Clause 5: Technical Requirements	
4.1	Organisation	5.1	General
4.2	Quality system	5.2	Personnel
4.3	Document control	5.3	Accommodation and environmental conditions
4.4	Review of requests, tenders and contracts	5.4	Test and calibration methods and method validation
4.5	Subcontracting of tests and calibrations	5.5	Equipment
4.6	Purchasing of services and supplies	5.6	Measurement traceability
4.7	Service to the client	5.7	Sampling
4.8	Complaints	5.8	Handling of test and calibration items
4.9	Control of nonconforming testing and/or calibration work	5.9.	Assuring the quality of test results
4.10	Corrective action	5.10	Reporting the results
4.11	Preventive action		
4.12	Control of records		
4.13	Internal audits		
4.14	Management review		

Table 1: ISO/IEC 17025 Table of contents.

These requirements are regarded as the necessary range of issues which need to be addressed by any testing facility to ensure reliable, "quality", results, regardless of the technical discipline in which it is operating. The range of requirements listed ensures that the testing characteristics described in section 3 above are achieved.

These management and technical requirements are a package, ie. an effective test result will not be achieved by one without the other. In fact a technically competent tester can produce ineffective results unless his or her testing is supported by an adequate management infrastructure. Obviously the reverse is also true - no quality management system can produce reliable test results unless the technical aspects are also adequately addressed.

Those of you who are familiar with the requirements of ISO 9000 for quality systems will by now recognise a strong similarity between these requirements and those of clause 4 of ISO/IEC 17025. This is not coincidence; in fact it was intentional. Clause 1.6 states that laboratories which comply with ISO/IEC 17025 do in fact comply with ISO 9001 or ISO 9002 as relevant. The reverse is not correct as ISO/IEC 17025 clearly includes technical requirements which are not addressed by either ISO 9001 or ISO 9002.

It is not the intent of the authors to detail the requirements of each clause of ISO/IEC 17025. Instead the focus will be on only the following four clauses which are considered to be of

particular significance in the achievement of reliable test results and which have been observed to be often handled by software testers in a less than satisfactory fashion.

5.1   Review of requests, tenders and contracts (ISO/IEC 17025, clause 4.4)

This is the first stage of the test process and is absolutely critical to the success of the testing. At this stage the test facility needs to clarify the requirements with its client and ensure that the extent of testing is fully understood by both parties. This applies regardless of whether tests are done in-house by the developer's own test staff, or by an independent, third party test facility. The software requirements specification must be available to the testers who need to develop a reasonable understanding of the purpose of the software and how it has been implemented.

The test laboratory must review its capability, ie technical resources, equipment, personnel, skills, etc. to ensure that it can perform the tests required. The client must understand the limitations of the proposed testing and the ensuing remanent risk.

The desired outcome from this process is the development of the scope of the testing agreed between the testers and the client. This will often take the form of a high level test plan which will also form the basis for any commercial arrangements between the laboratory and its client.

The commercial relationship between an independent third party test laboratory and its client may tend to ensure that this process is handled in a far more structured and complete way than might be the case when testing facilities are in-house. In the extreme case, when the testing staff are also the development staff, then it is likely that this stage of the process will be very difficult to handle effectively because of the lack of separation between developer and tester.

This stage of the testing process is critical to the remainder of the process. It requires the involvement of testing personnel who are able to understand the purpose of the software, how it has been implemented and any relevant regulatory requirements, eg Australian Taxation Office, industry requirements. Staff performing this review must understand the limitations of their testing.

The testing facility will often prepare an assessment of the remanent software risk associated with the proposed testing. An iterative process may result as the client attempts to find an acceptable balance between risk and testing cost. This process of negotiation and agreement between client and laboratory is critical to the success of any testing. For many testing projects this can involve a significant proportion of the total resources used on the project. The need for the more critical areas of the software to be identified and adequately tested is crucial for the client.

The agreed test plan, howsoever named and presented, is the basis for the testing and for development of test cases and is a major link in achievement of satisfactory outcomes. Without proper management and technical input at this stage reliable test results cannot be achieved.

5.2   Test and calibration methods and method validation (ISO/IEC 17025, clause 5.4)

This is a very significant area which is often handled poorly. Normal practice is to devolve the high level test plan down into a series of tests and test cases. Extreme care needs to be taken to ensure that the tests and test cases chosen do, in fact, meet the principles of capability, validity and repeatability. ISO/IEC 17025 requires validation of test methods. This activity aims to check that the tests and test cases are capable of testing the software, do produce valid results and are repeatable. The effort required can be in proportion to the risk involved. Some more trivial tests can be validated "by inspection", whereas complex tests will need to be validated using more detailed techniques. Validation can be done in various ways

including:

- code inspection,
- using the proposed method to test some known, or reference, software with well known performance and inter-comparing the results,
- running the proposed test on software containing a suitable range of known "seeded" errors.
- comparing its performance against another proven test technique.

Ideally, validation should be performed by someone other than the person who designed the test.

The use of automated test tools is a critical issue. The ability of a commercial test tool to be used within its intended range of application is not required to be validated by the test laboratory. The real issue is whether the software under test actually falls within the intended scope of application of the test tool, ie is the test tool capable of testing this particular software. Similarly any test tool developed in-house by the laboratory needs to be validated, ie its ability to test the software under test to produce valid results repeatedly, must be confirmed.

Test methods and test cases must be documented. They form the instructions to the tester. If the tester follows the method they are also part of the record of the test. Records of validation are required to be maintained. These can be an aid to analysis of suspect results.

The performance of invalid tests is a waste of time. Time lost analysing flawed results can be better spent validating the test in the first place. It is important to note that a test which is valid for the testing of a particular software package may not be valid for another package which is similar. Some further validation may be necessary to confirm its suitability.

Caution needs to be exercised to ensure that a sufficient level of validation is performed commensurate with the level of remanent software risk which is acceptable to the client. Similarly, care is needed to ensure that the range and selection of data values used in test cases is sufficient to adequately test the software.

5.3   Control of records. (ISO/IEC 17025, clause 4.12)

Record control is all about the collection, approval, change control, storing and archiving of records. Proper control of the following records is required to ensure that test results are reliable:

- the requirements specification for the software being tested,
- the agreed test plan,
- the test methods, test procedures, test cases developed to implement the test plan,
- records of validation of test cases,
- records of validation of test tools,
- records of hardware and software configurations (both software being tested and software on which it is running) including records of any changes,
- records of test personnel - who did the test - are they competent, qualified, authorised?
- test results - how did the system behave? Compliance outcomes and criteria.
- records of checking of results.
- test reports - interim, final. Records of checking of reports. Issuing authority.

The need to control records is paramount. Inadequate record control prevents back-tracking and analysis. Inadequate records can lead to incorrect conclusions or confusion regarding what needs to be done or what was done. In the worst case scenario, proper records are needed as a defence in case of litigation.

5.4   Assuring the quality of test results. (ISO/IEC 17025, clause 5.9)

Any process, once established, needs monitoring to ensure that it continues to deliver the expected outcomes and testing is no different. Even when a capable, validated, repeatable test method is applied, it is still possible for the results to be incorrect due to unpredictable changes in hardware or software or due to changes in environmental or human factors. Therefore it is considered necessary to implement a quality assurance regime whereby some form of check is made on an occasional basis to ensure that the testing process is continuing to produce expected results.

The level of checking can be dependent on the risk involved. Checks can include:

- Repeat testing by another person
- Testing of a reference sample
- Testing by alternative means

An experienced tester will not assume that a test will continue to deliver the correct result and will take some action to monitor the ongoing validity of the test results.

6.   How do we know that the testing meets ISO/IEC 17025?

ISO/IEC 17025 is used worldwide as the yardstick for independent, third party accreditation of testing laboratories operating in a wide range of technical disciplines. Such accreditation is accepted as prima facie evidence that the accredited laboratory has been assessed against the requirements of the standard by an independent third party laboratory accreditation body, such as NATA.

Accreditation is the outcome of a process of assessment of the operation of a test facility against the requirements of ISO/IEC 17025 by a laboratory accreditation body. The laboratory's operation is assessed against both the management system aspects and the technical aspects of the standard. Upon successful completion the laboratory is issued with an accreditation certificate and is able to use this to:

- confirm to its management that it is operating in accordance with ISO/IEC 17025, ie it has the necessary basic management and technical controls in place to enable it to produce reliable test results
- demonstrate to software developers, suppliers and users that it has the necessary systems in place to produce reliable test results
- demonstrate that it not only has systems which enable it to perform effective testing, but that it is sufficiently confident in its procedures that it has been prepared to have them scrutinised by an independent third party accreditation body
- support marketing initiatives by software developers and suppliers to enhance the market perception of software products
- increase confidence of purchasers and users in the reliability of the software they purchase
- complement and support general software quality initiatives and life cycle models.

7.   THE ASSESSMENT PROCESS

Accreditation of software testing facilities provides management with increased confidence in testing, reduces risk resulting from inadequate testing, supports due diligence initiatives, and increases the quality of the delivered product.   Some of this comes from the knowledge that the laboratory has undergone the process of assessment against ISO/IEC 17025.

Assessment generally comprises an initial "advisory" visit to the laboratory by the accreditation body's staff officer. The purpose of this is to gain an understanding of the test facility's operation, to inform the laboratory about the assessment process and to identify any obvious, major, deficiencies in its compliance with the standard which need to be addressed . This visit can occur either before or after a desktop review of the laboratory's management

system documentation. The laboratory's documentation is reviewed against the requirements of ISO/IEC 17025. Much of this review will be done prior to the formal on-site assessment. Implementation of documented procedures will be confirmed during the formal on-site assessment.

The initial on-site assessment will take place after the laboratory has addressed any issues identified during the advisory visit or the review of documentation. It would typically take several days depending on the scope of the test facility's operation. It is done by a team comprising a staff officer from the laboratory accreditation body and usually at least two technical assessors. The staff officer's role is to coordinate the assessment and to audit the management requirements. The role of the technical assessors is to review staff and procedures against the technical requirements. This is somewhat of a simplification as, in reality, the roles do have some overlaps, eg both will need to examine test records and reports.

The technical assessors are drawn from test laboratories, industry, regulatory bodies, academia, etc. They are peers of the staff of the laboratory being reviewed, in the sense that they are involved in similar activities. For example, a recent assessment by NATA was performed by a team of three comprising the NATA staff officer, a generalist software "tester" from another laboratory and someone from the regulatory body involved, who also had current "hands-on" testing experience.

Some readers will no doubt react negatively to the thought of being assessed by someone from another test facility. The usual response is "They're a competitor!" or "They'll find out all our secrets". The NATA staff officer's role is to ensure that assessors do not access commercially sensitive information. Usually such information is irrelevant to the assessment process. Also assessors are subject to confidentiality agreements. In extreme cases procedures can be implemented to ensure that assessors do not take away any hard or soft copy information from the laboratory premises. NATA has been using this approach successfully for over fifty years. Care is taken to minimise potential problems. However laboratories find that the benefits of discussing testing issues, techniques, etc, with true external peers, far outweighs any perceived commercial negatives.

It should be noted that assessors are independent, ie, they have no involvement in the outcome of the testing. It differs from internal peer review performed in some test laboratories. These normally consist of one staff member reviewing the work of another. In these situations the reviewer may be under significant pressure to "approve" the results, eg, the software release date has already passed. It should also be noted that the technical assessors are not NATA employees; they are true peers in the sense that they are involved in some way with software testing on a daily basis and are knowledgeable about the work done by the facility being assessed.

The assessment process is a mixture of discussion, demonstration, and "show and tell". Records will be examined to confirm compliance with the standard. The technical assessor will try to confirm staff competence via discussion and by demonstration of testing. Discussion will often revolve around the records of some completed series of tests, eg "Why did you do it this way?" " How do you know that result is valid?, etc.

The process is intended to be friendly, and non-confronting. Discussions are intended to be constructive. If the environment is right, one actually finds out more about the laboratory and this is to the long term benefit of the laboratory. Issues identified should be regarded as opportunities for improvement, rather than failures.

8.   WHO IS NATA ?

NATA is the National Association of Testing Authorities, Australia. It has been in the business of providing laboratory accreditation for 55 years. In fact it was the first such body in the world.

It has a Memorandum of Understanding with the Commonwealth Government by which the Government recognises NATA as the sole national laboratory accreditation body.

NATA has mutual recognition agreements (MRAs) with other laboratory accreditation bodies all around the world. Under these agreements tests performed in a laboratory accredited by one MRA partner will be accepted by the other MRA partners as though the tests were performed in one of their own accredited laboratories.

NATA is an association whose members comprise laboratories, government, regulators, professional associations etc. Members are represented on NATA's Council which provides general advice to the Board which directs the operations of the association. Specialist technical committees are established in each testing discipline to provide NATA with specific technical advice.

9.   WHAT DOES ACCREDITATION ACHIEVE ?

Accreditation results in a range of benefits for software test facilities, software developers, suppliers, users and regulatory bodies. These include the following:

- Many laboratories have observed that the process of preparation for the initial assessment enables them to identify problems in their operation. You can react to this by saying " Well I can do this on my own without seeking accreditation". Correct! You can! The reality is that most people don't or can't get around to it. Priorities overtake the intention and with the best will in the world, any detailed review of your own system gets postponed. Laboratories find that the decision to achieve accreditation becomes a driver for improvement of their management system and enhancement of their technical procedures.

- The assessment helps to remove internal "blinkers". We are all relatively unable to see the obvious due to our closeness to our own systems. It often takes someone with virtually no direct knowledge of our system to identify holes in it.

- There are known problem areas in testing, which tend to be common amongst laboratories working in similar technical disciplines. The assessment process can bring these to the attention of the testers for correction. Another opportunity for continuing improvement!

- Accreditation also helps to provide a more level playing field. It ensures that accredited laboratories are performing tests to the same standard using similar procedures, test techniques and competent personnel, operating within a proper management system .

- Accreditation also helps test laboratory management to confirm that due diligence issues have been adequately dealt with. Management can rest easy knowing that they have done everything which can be done to ensure that their test results are meaningful.

- It achieves a similar outcome for software developers, suppliers and users, by giving them assurance that the software has been tested in an environment of properly implemented management procedures by competent test staff using properly validated test procedures and equipment.

- In obtaining accreditation to ISO/IEC 17025 the laboratory is establishing a formal framework in which competent testing can be performed and reliable results achieved. It is worthwhile noting that this framework is consistent with the various software life cycle models and, in fact strongly supports them. It does this by ensuring that the testing is done in a quality environment which can ensure reliable test results and so add to the quality and value of the software.

- Accreditation provides a simple mechanism for laboratories to demonstrate that they do have the necessary competence and systems to meet the requirements of an internationally recognised standard for the competence of testing laboratories. This

information is readily promulgated through the accreditation certificate issued by NATA on successful conclusion of the assessment.

10.   CONCLUSION

NATA has recently built on the experience of over fifty years accreditation of testing facilities by establishing a program for the accreditation of software testing laboratories. This program is equally applicable to in-house testing facilities as well as third party test houses. It requires software test facilities to operate in accordance with ISO/IEC 17025 and, therefore, provides increased confidence in the effectiveness of the testing for laboratory management, software development managers, software suppliers and their clients alike.

In addition the development of this program took account of the lessons learnt from NATA's program for accreditation of AISEFs (Australian Information Security Evaluation Facilities). This program has been in place for five years and covers laboratories performing tests on software security.

The NATA Accreditation Service for Software Testing Laboratories is supported by the Commonwealth through the Testing and Conformance Infrastructure Program of the Department of Communications, Information Technology and the Arts. There was also considerable input from representatives of the IT testing industry.