Part 4

Testing of the System

4.1 Introduction

In order to consider the secrecy and accuracy of the chosen electronic voting system, it was necessary for the Commission first to carry out a general examination of its design and operation, including documentation and procedures. The Commission also reviewed previous tests carried out on the system and carried out its own further tests. This work has already been described in *Part 2*.

This part develops the Commission's main conclusions arising from its work as regards the examination testing of the system and takes account of issues arising from the public submissions reviewed in *Part 3*. Conclusions in relation to accuracy and secrecy are developed in *Part 5*.

4.2 Review of System Generally

The Commission examined the general operation of the system as it would be used at elections. This involved a review of the hardware and the software from the viewpoint of voters and election officials. The review was carried out on a semi-formal basis but also as part of the formal tests undertaken by the Commission.

Usability

Although criticisms of some aspects of its usability have emerged, the Commission found the system to be easily understood, both in general concept and in practical use. For election personnel, its operation corresponds logically to the administrative electoral procedures currently in place for manual voting. From the voter's point of view, the "booth" design of the voting machine and the replica ballot paper interface maintain a useful and helpful linkage to the paper voting procedure. This is not the case with all electronic voting systems.

The design and implementation of the proposed system are thus straightforward, contributing to ease of use by voters and election personnel.

There are, however, specific usability issues in the case of people with certain disabilities for whom use of the machine renders voting more difficult (and less private) than in the paper voting system.

Additionally, the system presents issues for people who are unfamiliar with, or fear, technology but these issues are capable of being addressed by means other than the adaptation of the system itself.

Previous Use

A previous version of the chosen system was deployed in pilot tests in three constituencies at the Dáil general election and at the Nice II referendum in 2002. The results from these pilot tests suggest that the system in general operated successfully and without significant difficulty.

While it has not been possible to verify that the system correctly captured all the votes cast at the 2002 Dáil election and referendum because there was no parallel run of paper and electronic voting, the Commission has confirmed in its own testing that the votes recorded as having been cast at the Dáil election were correctly counted by the system.

However, as there have been alterations to both the voting machine and, in particular, to the system software since 2002, previous experience of the use of the proposed system in Ireland can only be indicative of its general suitability for use. In view of the alterations made, the system as proposed for use at the 2004 elections and referendum is effectively a different system and will consequently require to be fully tested before its accuracy and secrecy can be determined.

A voting machine designed by the same manufacturers has been in use for several years in the Netherlands and has also been used more recently in Germany. Although this usage may reflect well on the general characteristics of the voting machine, the fact that it has been modified and updated for use in an Irish context limits the value of this experience for the purposes of establishing its suitability for use at elections in Ireland.

In addition, since the rules for the conduct of elections in the Netherlands and Germany are substantially different from those in Ireland, the software used in conjunction with the voting machine in each case is so different as to make it impossible to draw any definitive conclusions regarding the suitability of the software of the system chosen for use in Ireland.

Physical Security

In its examination of the system as currently proposed for use, the Commission identified a number of actual and potential security weaknesses. These are detailed in the work of the Commission as described in *Part 2* and are summarised hereunder.

It was found that, as regards the voting machine, the programming/reading unit and ballot module, the physical security measures and the methods used for the storage and transfer of the votes may be susceptible to interference on account of the ease with which the internal components of these devices may be accessed and also on account of the technology used.

In the case of the hardened PC on which elections are configured before the poll and on which the votes are counted afterwards, it was found that the "hardening" measures were easily bypassed so as to allow the PC's suppressed functions to be re-enabled, possibly for purposes which might interfere with the functioning of the software or the conduct of the election.

Multiple versions of the application software are in circulation and, although guidelines and training are given, there is no express control over which version is used in any particular case. It is possible to load and run older versions of the software onto the hardened PC in parallel with newer versions while any version can be overwritten, quite possibly inadvertently, by another.

In addition, it appears that there is nothing in the overall deployment procedures that would force key aspects of running the election to be carried out on the hardened PC rather than on any other PC.

Security Policy and Software Assurance

The work carried out by the Commission has also drawn attention to the importance of adopting such recognised security standards and software assurance methodologies as are appropriate to the design, development and testing of critical computer applications such as electronic voting systems.

In the absence of external audit features, the role of such standards and methodologies assumes a heightened significance in the context of establishing the trustworthiness and reliability of an electronic voting system such as the chosen system.

On the basis of the documentation and information regarding the system that is available to the Commission, it has not been possible to confirm or evaluate the degree to which such objectives have been met in the case of the chosen system.

Procedures, Controls and Documentation

The Commission has determined that the general procedures and controls surrounding the use of the system at elections are of comparable importance to those internal to the system itself. In addition to the system operating manuals that have been issued, the guidelines for returning officers and their staff have had to be significantly revised to take account of the introduction of the system and the changes it brings to the conduct of elections generally.

The Commission's review of the guidelines and user manuals issued by the Department of the Environment, Heritage and Local Government and of the procedures at elections generally indicates that close attention must be given to procedural issues and controls regarding the storage, handling, deployment and use of the equipment by election personnel, including the following:

- the recruitment and training of personnel with the necessary competence in the use of the system;
- the heightened need for a "segregation of duties" between key officials running the election, in a context where some of their crucial functions will be removed from public scrutiny during the count;
- the need for control and verification procedures for the issue and use of all equipment and software;
- the need for the scrupulous documentation of "ballot paper" reconciliation accounts;
- procedures for the manual entry into the electronic system of postal and special votes cast on paper;
- the need for very clear instructions for dealing with the errors and failures in the system that experience in other jurisdictions has shown will occur in one form or another on election day;
- the need for absolutely unambiguous version control of the software, to ensure that only certified and up to date versions are actually used in the election;
- the provision of clear guidance and assistance to voters, including voters who may be fearful of, or inexpert at using, new technology.

4.3 Review of Previous Tests

The reports of a number of previous tests carried out on the system and its components were reviewed by the Commission and these are listed at *Appendix 1B*. These tests were mainly commissioned by the Department of the Environment, Heritage and Local Government and were carried out by independent agencies and companies, mostly outside Ireland.

The Commission also had access to voting information from the pilot tests of the system at the 2002 Dáil general election and to information relating to tests carried out by the Department of the Environment, Heritage and Local Government based on data from local elections held at Athy and Buncrana in 1999.

Scope and Context of Tests

Most of the main component parts of the proposed system have been reviewed or tested by reputable independent testing agencies. However, these components have been reviewed and tested to different degrees and for different purposes.

Some of the hardware tests relate to certification of compliance with safety and other requirements as regards the physical and electrical properties of the components of the system but do not specifically test their functionality as regards their use at elections. Typically the hardware devices are tested individually and not when connected to each other or when used as part of an overall test of the system as a whole.

The functionality of the system has also been independently evaluated against the prescribed requirements and functional specification of what it is supposed to do.

Testing of the software has related mainly to the counting module of the election management software. Although the embedded source code of the voting machine and the source code of the election management and counting modules of the election management software have been reviewed, not all have been tested. The reviews typically measure the code against the prescribed requirements and functional specification of what it is supposed to do. Although the quality and overall structure of the code are commented on in a general way, these (and other) design aspects of the code do not appear to have been specifically submitted for evaluation.

It would be desirable that the whole system should be evaluated, tested and certified by a single independent agency as being suitable for use in an Irish electoral context and that this should be carried out by an agency duly accredited to a recognised standard.

Once its suitability has been independently certified, it would also be desirable that there should be a segregation of functions as between the approval of the system for use in Ireland and its procurement and deployment for such use. This appears to have been the case in the Netherlands where the machine is type-approved by central government (on the basis of independent certification) as being suitable for use at elections by the authorities in local municipalities.

Relevance of Tests

Although the hardware and software elements were found to have been tested to the extent outlined

above, a significant issue arose from the fact that the software versions that were tested and examined had already been, or would be, superseded by more recent versions by the time of the June elections.

This issue is of particular relevance to the election management software which has gone through numerous versions since it was first reviewed and tested. While the changes made in subsequent versions have been submitted for further independent review, the reports of earlier reviews appear to have been relied upon as having continuing validity in respect of the unchanged portions of the software.

It has been indicated to the Commission that each new software version requires to be reviewed and tested in full because the changes made, however small, can have unforeseen effects on portions of the software that were not changed, giving rise to potentially serious consequences.

This issue also has a bearing on the usefulness of the 2002 pilot tests of the system since the software version as it then existed has been changed many times in the intervening period and must be regarded as being a different product to that which was used in 2002.

It was not possible in the context of this report for the Commission to carry out its own full review of the software. Nor would there be time for such a review to have been carried out sufficiently thoroughly before the elections in June 2004, particularly as the final version of the software had not yet been issued at the time of the Commission's first reporting date of 1 May 2004.

On the basis of the information available to it, the Commission is not aware of whether similar issues to those outlined above in relation to the election management software may also apply to the other software and hardware aspects of the remaining parts of the system outside of the election management software.

It is also a matter of some concern that new versions of the software continue to be issued in the run-up to the June elections as it suggests that the software is not fully prepared for its intended purpose.

Before the system can confidently be adopted for use at elections in Ireland, it will be necessary to settle on a final definitive version of the software and all related hardware and software components. There then needs to be a full independent review of the source code and testing of the final system to be used. Any subsequent software upgrade would require a further full system re-test.

Completeness of Test Programme

On the basis of the work done by the Commission in reviewing previous tests, it appears that some components of the system have not been independently tested and certified, other than in respect of their physical and electrical properties as mentioned above.

As regards hardware devices, the functionality of the programming/reading unit has not been independently tested while the hardened PC and the printer do not appear to have been tested at all.

As regards software elements, processes at the interface between tested components (such as the processes for the aggregation of the votes from individual ballot modules, including the

disaggregation of votes cast at separate elections, and their transfer to a CD for transportation to a remote counting venue) do not appear to have been independently tested.

Range of Tests

While not wishing to comment on the quality of the tests carried out to date, it appears to the Commission that the range of tests that has been carried out is limited.

In particular, it appears that there has been no independent end-to-end testing of the full system as it would be operated at elections and those tests that have been carried out by the Department of the Environment, Heritage and Local Government are based on small data sets only.

The need for end-to-end testing – encompassing all the software and hardware components, as well as the administrative systems for managing these – is accentuated by the fact that the use of the system at the June elections would involve up to 4 polls being taken simultaneously and by the added complexity of the multi-level administrative structure that is proposed to operate it. This appears to be a particularly high-risk environment in which to roll out the first live use of the proposed system in a national context.

It is also the case that there has been no parallel testing of the proposed system in a real or simulated election context, either as compared with the paper system or as compared with alternative methods of electronic voting. Although the system was deployed on a pilot basis in 2002, these elections were not run in parallel with a paper ballot and the software has been modified many times since then. Parallel testing of critical systems is regarded as an important test to ensure their reliability and accuracy.

4.4 Further Tests Carried Out

Having identified a number of gaps in the range of tests carried out to date, the Commission decided to carry out further tests of the system. While these tests are not an adequate substitute for formal independent testing of the system, they serve to illustrate the range and the nature of the additional testing that is required.

Three main types of test were carried out by the Commission:

(1) input-output testing of a representative sample of the voting machines deployed nation-wide (*Appendix 2C*);
(2) miniature end-to end testing of the whole system at simultaneous elections (*Appendix 2D*);
(3) parallel testing of the counting software using a large number of data sets (*Appendix 2E*).

In addition, the work carried out by the Commission included a preliminary code review of the source code for the counting software only.

Test Method

The above tests were carried out as “black box” input/output tests. A black box test is one that is carried out without knowing how the system processes the data between input and output points.

Data inputs are thus entered into the system and the outputs are evaluated against known or expected values.

More comprehensive testing would involve "glass box" testing whereby the internal processing of the system is known and data can therefore be designed to test it in a more structured fashion. In addition, because glass box testing involves access to the source code of the system software, it is possible to review the structure and content of the code to identify behaviours, characteristics and possible weaknesses of the system as well as assessing the quality of its design.

Sample Test of Voting Machine

The test carried out by the Commission on a representative sample of the voting machines deployed to returning officers confirms that the deployed system can accurately and consistently record voter preferences at a single-election situation. This test also confirms the reliability in use of the ballot modules as a corresponding number of these were used in the test.

End-to-End Test

The miniature end-to-end testing of the system involved the configuration of polls and the casting of sample votes at three simultaneous elections. The votes were then transferred to separate centres for counting and the results were verified to be correct. This test confirms that the system can accurately record and count the votes in the context of multiple simultaneous elections.

Parallel Test

The parallel test of the counting software using a similar counting program developed for the Commission to process voting information from the pilot tests at the 2002 Dáil elections confirms that the votes recorded at these elections were accurately counted.

This test was also applied to a large number of sample data sets including data sets involving unusual or difficult electoral situations and the counting software produced the same results as the test program in the majority of cases.

In a small number of cases in which a discrepancy was noted between the results produced by the counting software and those produced by the test program, the discrepancy was found to be attributable to an error in the counting software involving the calculation of fractions for the purposes of distributing votes in the transfer of surpluses.

While this error is significant in itself and should be remedied, it also reduces confidence in both the programming and the previous testing of the system and there is a possibility that further and more extensive testing will uncover further software errors.

Source Code Review

Although the Commission sought access to the full source code of the election management software and the voting machine for the purpose of carrying out a preliminary review of its structure

and content, it was not possible to achieve this within the timeframe of the Commission's interim report. However the Commission obtained access to the source code relating to the counting function (only) of the election management software and this was reviewed accordingly.

In the absence of the full source code, it was therefore not possible to carry out even the preliminary review that might have been possible within the timeframe of the Commission's interim report.

It has become clear to the Commission that a comprehensive review of the full source code of the system is necessary to establish its trustworthiness to a level compatible with the critical importance of voting at elections.

Such a comprehensive code review was also outside the scope of the Commission's reporting deadline of 1 May, indeed, there was not sufficient time remaining in any event before the June elections to allow a full code review of the final version of the software that would actually run in those elections.

4.5 Conclusion

From its general review of the system, from its review of previous tests and from the results of the further tests it has carried out, the Commission has identified a number of items that need to be addressed before the chosen system can be put into use.

Furthermore, the general issues in relation to testing discussed in this part draw attention to what the Commission considers to be the principal weakness of the proposed system. This is that the level and comprehensiveness of testing to date – in particular as regards the complete system as proposed to be deployed at the June elections – means that it is not possible to establish its trustworthiness and reliability at this time and that further testing, including a full source code review, is thus required.

This conclusion must be seen in a context where the trustworthiness and reliability of a system for counting votes in public elections must be established with a very high degree of confidence – a confidence far higher than in many other applications of new technology.

From this it further follows that it is not possible at this time to draw definitive conclusions as to the accuracy and secrecy of the proposed system and such limited conclusions as can be drawn in this regard are set out in *Part 5*.