
UK IT SECURITY EVALUATION & CERTIFICATION SCHEME

UK CC Interpretation - UK/3.1/006

12 March 2007

Status: Endorsed for use in UK Scheme

Subject: Selecting a sample for repeating developer tests

REFERENCES CEM, Version 3.1, Revision 1, September 2006 - ATE_IND.2-4

Issue

Sample Size

1. The CEM Annex A.2 provides principles on how the evaluators should go about selecting a sample to examine a subset of evaluation evidence assumed to be representative of the entire set to gain confidence in the correctness of the evidence.
2. Whilst the principles for selecting a sample of developer tests to repeat are comprehensive, they do not expand upon the considerations underlying the principles, which may help the evaluators in their selection of developer tests to sample.

Interpretation

Sample Size

3. The following approach should be used (as outlined in the CEM paragraph 1822):
 - a) The sample used must be sufficient to represent all distinct aspects of developer testing, such as different test regimes.
 - b) The sample used must be sufficient to detect any systematic problem in the developer's functional testing process.
 - c) The evaluator contribution resulting from the combination of repeating developer tests and performing independent tests must be sufficient to address the major points of concern for the TOE. Performance of sufficient independent evaluator testing, as required by ATE_IND.2.2E, is thus a prerequisite to point (b) above.
 - d) The sample used should be justified in the ETR.
4. The above approach is derived from the following considerations:
 - a) Whilst individual problems in the implementation of the TOE are of concern, it is the implementation itself, rather than the testing of it, where this concern is primary. Developer functional testing gives the developer a first opportunity to check the security

of the TOE. In order to obtain assurance in the implementation through class ATE testing, the evaluator first needs to obtain confidence that the developer's testing process does not exhibit any systematic problems (e.g. associated with inaccurate performance or recording of the tests), but having done so can add most value by performing independent tests under ATE_IND.2.2E. Note that:

- Not only does class ATE testing check the implementation of the TOE, but a correct TOE implementation gives confidence in the validity of a set of tests which is claimed to demonstrate correctness of the implementation. By demonstrating correctness of the implementation, the independent tests performed by the evaluator can thus add to the confidence in the validity of the developer's tests.
- Once confidence has been obtained in the developer's testing process, problems in the implementation are more likely to be identified through independent tests formulated by the evaluator. This builds on the principle of CEM para 1374(c) (referenced by CEM para 1369) that there should be a balance of evaluation activities.

b) Flexibility is given for the sample percentage to be determined as appropriate. Note that:

- Typically, the greater the amount of developer testing, and the greater the uniformity of the developer's test process, the lower the percentage sample that will suffice to represent distinct aspects of developer testing and detect a systematic problem. In the same way, if the amount of developer testing is small, or if the developer's test process is highly variable, then a greater percentage sample may be required (e.g. it will usually be insufficient for the evaluators to repeat only a single test).
- Some tests may be compound or more complex than others, so any sample should be viewed as a percentage of 'the testing' as opposed to a percentage of 'the tests'.

c) This approach is consistent with other valid CEM principles:

- It employs the principle of cost effectiveness
- It employs the principle that there should be a balance of evaluation activities
- It employs the principle that the evaluator should make a sufficient contribution to assurance in the area of class ATE testing.

d) The CEM para 1371 note, to the effect that a larger sample than originally planned may be needed if the original sample reveals problems, also remains valid.