

ISO/IEC 25023 소프트웨어 품질 특성 평가를 이용한 시험자간 비교시험에 관한 연구

이창희*, 김명섭, 김윤숙, 노아름, 양정심, 송성현
LG전자, SW센터, SW공학연구소

A Study on Proficiency Comparison Testing Between Testers Using ISO/IEC 25023 Software Quality Characteristics Evaluation

Chang-Hee Lee*, Myung-Sup Kim, Yoon-Sook Kim, A-Reum Noh,
Jung-Sim Yang, Sung-Hyun Song
Software Engineering Laboratory, Software Center, LG Electronics, Seoul, South Korea

요약 본 연구는 국제표준 ISO/IEC 25023 소프트웨어 품질 특성 평가를 이용하여 ISO/IEC 17043 숙련도 시험 요구 사항을 기준으로 시험자간 숙련도 시험의 결과를 비교 분석한 연구이다. 시험 분야 신청서는 KOLAS에서 정의하는 전기 시험 분야의 소프트웨어 시험으로 한국계량측정협회에 제출을 하였다. 시험자간 숙련도 시험 평가 항목은 국제표준인 ISO/IEC 25023 SW 품질 특성 중 기능 완전성과 시간 효율성으로 선정을 하였다. 시험자간 숙련도 비교 시험은 서울 소재의 LG전자 SW공인시험소에서 2019년 5월 28일에 KOLAS 평가사가 현장에 입회하여 실시되었다. 시험 평가 방법은 기능 완전성으로 시험자 1인당 1회 측정을 하였고 시간 효율성으로 시험자 1인당 10회를 반복 측정을 하였다. 시험의 수행 결과 기능 완전성의 경우는 12개의 시험 단계를 시험자들이 모두 통과하여 만족으로 평가되었다. 시간 효율성은 10회 반복 측정한 평균 결과 시험자 1과 3은 9.45 ± 0.01 초, 시험자 2는 9.52 ± 0.0 초로 나타났으나 통계적으로 유의하게 차이가 없었다. 결론적으로 시험자들의 숙련도는 동등하였으며 해당 시험기관 시험 결과의 신뢰성을 확인할 수 있었다.

Abstract This study compared and analyzed the results of proficiency testing between testers based on the international standard ISO/IEC 17043 proficiency testing requirements using ISO/IEC 25023 software quality characteristics evaluation. Application of the test was submitted to the KASTO as a software test in the electric test field defined by KOLAS. Among ISO/IEC 25023, the evaluation items for proficiency testing between testers were selected as functional completeness and time behavior. The comparison test between testers was carried out on May 28, 2019 at the accredited LG Electronics SW Laboratory in Seoul and attended by KOLAS assessors. The test evaluation method was measured once per tester with functional completeness, and measurement was repeated 10 times per tester according to time behavior. As a result, in the case of functional completeness, all 12 test steps were passed, and the satisfaction was evaluated. Time behavior was repeated 10 times, and the mean was 9.45 ± 0.01 seconds for testers 1 and 3 and 9.52 ± 0.0 seconds for tester 2, but there was no statistically significant difference. In conclusion, the proficiency of the testers was equal, and the reliability of the test results of the test institute was confirmed.

Keywords : ISO/IEC 17043 proficient testing, ISO/IEC 25023 software quality evaluation, Functional completeness, Time behavior, Software testing

*Corresponding Author : Chang-Hee Lee(LG Electronics)

email: chp.lee@lge.com

Received March 17, 2020

Accepted June 5, 2020

Revised April 21, 2020

Published June 30, 2020

1. 서론

끊임없이 급변하는 현대 시대는 생존과 번성을 위하여 정보 기술 (Information Technology)에 점점 더 의존하고 있다. 정보 기술의 발전은 개인의 삶의 형태뿐만 아니라 모든 산업 구조를 변화시키며 기업들의 경쟁력과 이익에도 영향을 미치고 있다[1]. 이러한 정보 기술 분야의 연구들을 보면 소프트웨어의 품질 문제는 사회 전반에 걸쳐 인적, 물질적, 재산적 피해와 직결될 뿐만 아니라 소프트웨어 산업의 경쟁력 확보를 위해서 해결되어야 하는 주요 이슈이다[2].

소프트웨어의 품질은 어떠한 종류의 소프트웨어가 그 특성에 맞게 가지고 있어야 하는 속성을 말하는 것이다. 소프트웨어 품질 특성을 평가하는 국제 표준 규격이 ISO/IEC 25023 이다. ISO/IEC 25023에서는 소프트웨어 품질 특성을 기능성 (Functionality), 신뢰성 (Reliability), 사용성 (Usability), 효율성 (Efficiency), 유지보수성 (Maintainability), 이식성 (Portability), 보안성 (Security), 상호운용성 (Interoperability)으로 8 가지 품질 평가 모델을 매트릭스 형태로 제시하고 있다[3-7].

최근 여러 국가나 기업들의 동향을 보면 ISO/IEC 25023 소프트웨어 품질 특성 평가 모델을 활용하여 적합한 평가체계를 구축하고 소프트웨어 제품 특성에 맞는 시험방법 개발에 경쟁들을 하고 있다. 이는 소프트웨어의 공신력 있는 시험 결과를 통하여 경쟁력 있는 품질 확보를 위한 목적이다. 이러한 소프트웨어 평가 체계나 시험 방법 개발의 공신력 있는 구체적인 실행 방안이 숙련도 시험 평가 프로그램이다. 숙련도 시험 평가 프로그램은 실험실에서 시험 절차를 준수하여 생산되는 시험 데이터들의 신뢰성을 평가하는 외부 평가 프로그램이다. 특히 숙련도 시험 프로그램은 시험기관과 교정기관이 공인 시험 기관으로 적합성을 인정 획득하기 위해서는 의무적으로 실적을 제출하도록 되어 있다[8]. 다시 말해, 숙련도 시험은 해당 기관의 실험실에 대한 시험 역량을 평가하기 위한 필수 도구인 것이다. 숙련도 시험은 동일 시험 재료로 여러 시험기관들 간에 숙련도에 차이가 있는지 평가하는 시험기관간 숙련도 시험 평가와 동일 시험기관에 속해 있는 시험자들 간에 차이가 있는지를 평가하는 시험자간 숙련도 시험 평가를 대표적인 형태로 들 수 있다[9].

이러한 숙련도 시험 체계에도 불구하고 소프트웨어 시험기관들은 시험 결과의 신뢰성과 품질 확보에 어려움을 겪고 있다. 이는 산업 현장에서 소프트웨어 국제 표준

을 어떻게 활용해서 시험 평가에 적용해야 하는지가 명확하지 않은 경우나 시험 재료가 없는 경우, 시험 방법을 수립하지 못하는 경우, 그리고 특수 시험수행 전문가가 부족한 경우들이 신뢰성과 품질확보를 어렵게 하는 원인이다[10]. 더욱이 하드웨어의 경우는 공인 시험기관 운영이 활발하게 이뤄지고 있는 반면에, 소프트웨어의 경우에는 시험기관의 공인 인정 실적이 상대적으로 부족하고 소프트웨어 그 자체의 특성으로 인하여 숙련도 평가 방법이 다양하게 개발되지 않고 있다. 이러한 이유로 소프트웨어 분야에서 품질 보증을 위하여 국제 표준에 적합한 시험자의 숙련도 시험 프로그램 개발과 평가가 필요하다.

이전 소프트웨어 시험의 연구 사례를 보면 시험기관 간에 시험 데이터를 통계적 기법을 활용하여 측정된 결과의 정확도 및 성능을 비교 분석한 사례[11]가 있었고 제품별 융합 소프트웨어에 대한 보안 및 성능 관점에서 소요되는 시간들을 분석한 연구[12], 그리고 정보통신 (Information & Communication Technology) 분야에서는 웹 서비스에서 시간 반응성의 평가 방법 개발에 관한 연구가 있었다[13]. 그리고 국제 표준 규격 ISO/IEC 25023 소프트웨어 품질 평가를 기준으로 시험기관의 데이터를 분석하고 인구학적 성별에 따라 소프트웨어 시험 결과에 차이가 있는지를 통계적으로 분석한 연구도 있었다[14].

본 논문에서는 이전 논문과는 다르게 ISO/IEC 25023의 소프트웨어 품질 특성과 ISO/IEC 17043 숙련도 요구사항을 이용하여 시험 기관의 수행 능력을 평가하는 시험자간 숙련도 프로그램의 평가 절차 및 방법, 그리고 시험 결과를 비교 분석한 시험 적용 사례를 제시하였다. 이러한 사례를 통하여 본 논문에서 제시된 내용들은 소프트웨어 분야에서 공인 시험기관으로 최초 인정을 신청하는 기관들과 소프트웨어 품질 개선 활동을 수행하는 기관들에게 선행 사례로 활용 될 수 있을 것이다.

본 논문의 구성은 2장에서 국제 표준 규격 ISO/IEC 25023의 소프트웨어 품질속성 평가를 기준으로 해당 연구의 목적을 설명하였다. 3장에서는 소프트웨어 품질 평가와 비교숙련도 시험의 표준 규격을 설명하고 본 연구의 시험 방법 및 자료의 수집에 대하여 설명하였다. 4장에서는 시험자간 숙련도 시험 결과로 기능 완전성과 시간 효율성의 시험 결과를 분석하고 정리하였다. 5장에서는 시험자간 숙련도 시험 결과에 대한 논의 사항과 제약 사항에 대하여 설명을 하였고 6장에서는 본 연구의 결론 및 앞으로의 연구방향과 과제에 대하여 제시를 하였다.

2. 연구 목적

본 연구의 목적은 소프트웨어 시험자간 숙련도 비교시험을 통하여 해당 시험기관의 시험수행능력 및 기술능력을 파악하고 시험 결과의 신뢰성 및 정확성을 확보하는데 있다.

해당 시험의 구체적 평가 수단으로는 국제 표준 규격 ISO/IEC 25023의 소프트웨어 품질속성 평가를 기준으로 해당 시험에 대한 기능 완전성과 시간 효율성을 측정하여 시험자간의 능력을 비교 평가하고자 한다(Table 1).

3. 연구 방법

본 연구의 시험 방법은 국제 표준 ISO/IEC 17043 숙련도 시험 요구사항을 기준으로 절차 및 방법에 대한 모델을 제시하고 국제 표준 ISO/IEC 25023 소프트웨어

품질 특성 평가를 기준으로 소프트웨어 시험의 평가 지표를 이용하여 시험자들이 수행한 시험 결과를 비교하는 연구이다.

3.1 연구 도구

3.1.1 소프트웨어 품질 평가

국제 표준 ISO/IEC 25023 소프트웨어 품질 평가는 ISO/IEC 25010에서 정의하는 소프트웨어의 품질 모델을 적용하여 사용되었다. 소프트웨어 품질 평가는 다양한 다른 국제 표준 규격 또는 지침에서 제시하는 방법을 참조하여 사용할 수 있다. 예를 들어, 기능의 크기 측정은 ISO/IEC 14143에서 정의하고 있으며 시간 효율 측정의 경우는 ISO/IEC 14756을 참조할 수도 있다. 이러한 ISO/IEC 25023 소프트웨어 품질 평가의 특성은 소프트웨어 시험뿐만 아니라 개발 및 제품 생산 단계에서도 다양하게 사용될 수가 있다.

Table 1. Summary table for the usage of quality measures

Subcharacteristic	Quality measure name (ID)	Description
	Measurement function	
Functional completeness	Functional coverage (FCp-1-G)	What proportion of the specified functions has been implemented?
	$X = 1 - A/B$ A = Number of functions missing, B = Number of functions specified	
Time behavior	Mean response time (PTb-1-G)	How long is the mean time taken by the system to respond to a user task or system task?
	$X = \sum_{i=1}^n (Ai) / n$ Ai = Time taken by the system to respond to a specific user task or system task at i-th measurement, n = Number of responses measured	
	Response time adequacy (PTb-2-G)	How well does the system response time meet the specified target?
	$X = A/B$ A = Mean response time measured by PTb-1-G, B = Target response time specified	
	Mean turnaround time (PTb-3-G)	What is the mean time taken for completion of a job or an asynchronous process?
	$X = \sum_{i=1}^n (Bi - Ai) / n$ Ai = Time of starting a job i, Bi = Time of completing the job i, n = Number of measurements	
	Turnaround time adequacy (PTb-4-G)	How well does the turnaround time meet the specified targets?
$X = A/B$ A = Mean turnaround time measured by PTb-3-G, B = Target turnaround time specified		
Mean throughput (PTb-5-G)	What is the mean number of jobs completed per unit time?	
$X = \sum_{i=1}^n (Ai / Bi) / n$ Ai = Number of jobs completed during the i-th observation time, Bi = i-th observation time period, n = Number of observations		

ISO/IEC 25023 소프트웨어 품질 평가의 주특성 항목 중 기능 적합성(Functional suitability)의 평가 항목은 제품 또는 시스템이 제공하는 정도를 평가하는데 사용된다. 기능 적합성의 정의를 보면 주어진 환경 조건에서 시스템과 소프트웨어의 명시적인 요구와 암묵적인 요구를 충족하는지를 평가하는 것으로 되어 있다. 기능 적합성의 부특성으로 기능 완전성 (Functional completeness)의 평가 항목은 소프트웨어의 기능들의 집합이 모든 요구사항을 포함하는 정도를 평가하는 것이고 기능 정확성(Functional correctness)은 구현된 모든 기능이 정상적으로 동작하는지 여부를 평가하고 기능 적절성(Functional appropriateness)은 사용 목적에 따른 기능 제공이 적절한가를 평가하는데 사용된다.

성능 효율성 평가 항목은 명시된 조건에서 사용되는 자원의 양에 대한 성능을 평가하는데 사용된다. 자원의 양은 소프트웨어 제품 또는 시스템의 소프트웨어, 하드웨어 구성, 그리고 저장 매체와 같은 자료들이 포함될 수 있다. 성능 효율성의 부특성으로 시간 반응성 항목은 기능을 수행 할 때 제품 또는 시스템의 응답, 처리, 작업의 시간과 속도가 요구 사항을 충족하는지에 대한 정도를 평가하는데 사용된다.

3.1.2 비교속련도 시험

국제 표준 ISO/IEC 17043 비교속련도 시험에서는 비교속련도에 대한 정의를 둘 이상의 시험기관이 사전에 정해진 조건에서 동일 또는 유사한 시험 대상 제품에 대하여 시험을 수행하고 설정된 평가 기준으로 시험기관들의 비교를 통하여 그 기관들의 능력을 평가하는 것으로 정의하고 있다[15].

ILAC (International Laboratory Accreditation Cooperation, 국제 시험기관 인정 협력체)과 APLAC (Asia Pacific Laboratory Accreditation Cooperation, 아시아 태평양 시험기관 인정 협력체)은 속련도시험 프로그램 운영에 대하여 가이드를 제시하고 있으며, 한국의 경우는 KOLAS (Korea Laboratory Accreditation Scheme, 한국 인정 기구)에서 속련도 시험 프로그램에 대한 운영 규정과 시험기관들에 대한 속련도 시험 참가 주기를 규정하고 있다.

KOLAS 규정에 의하면 속련도 시험의 주요 요구사항은 공인 시험기관으로 인정을 획득하기 위해서는 반드시 사전에 속련도 시험에 참여하고 평가 기준에 만족하는 결과를 제출하게 되어 있다. 또한 시험기관이 공인 시험기관으로 지위를 인정 획득한 경우라도 해당 인정 시험

분야의 측정수행능력을 지속적으로 입증하기 위하여 시험 중분류별로 3년에 1회 이상의 속련도 시험에 참가한 실적의 결과를 요구하고 있다. 만약 국내에 비교 가능한 시험기관이 없는 경우나 시료의 기준값 설정이 불가능한 경우에는 KOLAS에서 지정한 평가사의 입회하에 시험자간 비교시험도 가능하도록 규정을 하고 있다.

3.2 시험 방법 및 자료 수집

본 연구의 시험 분야는 KOLAS에서 정의하는 전기시험 분야 (대분류, 03)의 소프트웨어 시험 (소분류, 012)으로 KASTO (Korea Association of Standards & Testing Organizations, 한국계량측정협회)에 신청서 (PC-19-23)를 제출 하였다. 시험자간 속련도 시험 평가 항목은 국제표준인 ISO/IEC 25023 SW 품질 특성 중 기능 적합성의 기능 완전성과 수행 효율성의 시간 효율성을 시험 평가 항목으로 선정을 하였다.

시험자간 속련도 비교 현장 시험은 KOLAS 평가사가 입회하여 2019년 5월 28일에 서울 소재의 LG전자 SW 공인시험소 (KT874)에서 실시되었으며 시험 모델은 기능 완전성으로 시험자 1인당 1회 측정 하였고 시간 효율성으로 시험자 1인당 10회 반복 측정을 하였다(Fig. 1).

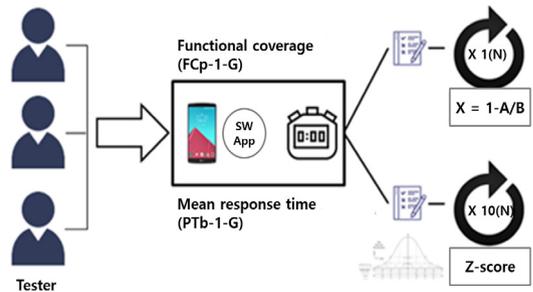


Fig. 1 Model of Testing

시험자간 속련도 비교 평가 시험은 안정성이 확보된 모바일 폰 1개 시료로 진행하였다. 시험에 사용된 모바일 폰 기기는 중국 샤오미(Xiomi) 업체에서 제작한 Redmi Note 7 (Android 9) 모델을 사용하였으며 소프트웨어 어플리케이션은 GIF 메이커-편집기(제조사: GIF Maker & GIF Editor & Video Maker, 버전: 1.5.26)를 사용 하였다.

시험자간 속련도 시험의 참가자는 동등한 시험 능력을 보유한 시험자들로 선정하기 위하여 대학 전공 및 경력 연수를 고려하여 최종 선정하였다. 시험에 참여하는 시험

자는 모두 여성이었으며 경력은 최소 20년 이상이었고 대학 전공은 컴퓨터공학과를 졸업하였으며 학력은 학사 2명과 석사 1명으로 선정하였다(Table 2).

Table 2. Characteristics of participants in the test

Participants	Gender	Education	Major	Duration of job
Tester 1	Female	Bachelor Degree	Computer Engineering	21
Tester 2	Female	Master Degree	Computer Engineering	20
Tester 3	Female	Bachelor Degree	Computer Engineering	21

시험자간 숙련도 비교시험은 KOLAS 숙련도 시험 기술위원회에서 선임한 평가사가 사전에 시험 계획서의 적합성을 평가하였고 현장 시험 당일에 입회하여 진행하였다. 시험자들은 일괄된 시험 수행을 위하여 사전에 준비된 시험 절차를 제공 받았으며 시료의 사용방법과 시험 범위를 확인하고 시험을 수행하였다. 시험 절차는 시료를 시험자들에게 순차적으로 전달하고 시험자는 시료의 번호 및 기기명, 제조사, 모델, 소프트웨어 버전이 문제가 없는지 확인하고 시험을 진행하는 것이다. 시험 진행 중 시료에 문제가 있으면 시험 관리자에게 즉시 보고하여 조치할 수 있도록 하였으며 시험수행의 결과 기록은 표준화된 기록지를 이용하여 작성하도록 하였고 완료된 결과 기록지는 입회 참석한 KOLAS 평가사에게 제출하도록 하였다.

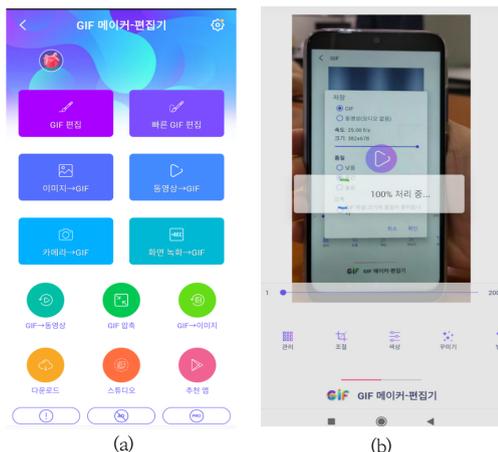


Fig. 2 Test procedure: (a) Application screen of functional conformity assessment, (b) Video screen to measure the time of mean response time.

기능 완전성의 시험절차는 다음과 같다.

- step 1) 시료 폰의 바탕화면에 있는 시료 어플리케이션 (GIF 메이커-편집기)을 실행한다 (Fig. 2a).
- step 2) GIF 메뉴를 실행한다.
- step 3) 이미지 리스트에서 시험에 사용할 파일 4 개를 선택한다.
- step 4) 시험에 사용될 이미지 파일 4 개가 선택되었는지 체크박스 표시로 확인한다.
- step 5) 시료 어플리케이션 상단의 체크 버튼(√)을 선택한다.
- step 6) 전환된 화면에서 오른쪽 상단 체크버튼을 누른다.
- step 7) 저장 옵션으로 [GIF, 속도 (20.00 f/s), 크기 (자동), 품질 중간, 압축은 선택하지 않음]을 확인한 후 확인 버튼을 누른다.
- step 8) 이미지가 100% 처리됨을 확인하고 재생버튼을 눌러 GIF 파일이 생성됨을 확인한다.
- step 9) 생성된 GIF 이미지 정보 (파일명, 파일크기, 해상도, 프레임 번호, 지속시간, 파일경로)를 확인한다.
- step 10) 해당 파일 경로로 진입하여 GIF 파일이 생성되어 있음을 확인한다.
- step 11) 시험 결과 기록지에 Pass와 Fail 중 하나를 선택하여 작성하고, Fail의 경우는 특이 사항 란에 자세히 기록한다.
- step 12) 시료 어플리케이션(GIF 메이커-편집기)의 종료 버튼을 눌러 시험을 종료한다.

시간 효율성의 시험절차는 다음과 같다.

- step 1) 시료 폰의 바탕화면에 있는 시료 어플리케이션(GIF 메이커-편집기)을 실행한다.
- step 2) 동영상의 GIF 메뉴를 실행한다.
- step 3) 이미지 리스트에서 시험에 사용할 파일 4 개를 선택한다.
- step 4) 오른쪽 상단의 완료 버튼을 누른다.
- step 5) 처리 진행 표시가 100% 처리됨을 확인한다.
- step 6) 카메라 어플리케이션으로 동영상 촬영을 시작한다.
- step 7) 시험자는 촬영시작 확인 후 오른쪽 상단의 체크 버튼(√)을 누른다.
- step 8) 저장 옵션으로 [GIF, 속도 (20.00 f/s), 크기 (자동), 품질 중간, 압축은 선택하지 않음]을

- 확인한 후 확인 버튼을 누른다.
- step 9) 진행 상태가 100% 처리됨을 확인 후 동영상 촬영을 완료(Stop)한다(Fig. 2b).
- step 10) 시험자는 촬영한 동영상 재생 어플리케이션을 실행한다.
- step 11) 재생버튼을 클릭하고 동영상에서 체크버튼을 클릭한 시점과 종료 시점의 프레임(Frame)시간을 측정한다.
- step 12) 측정된 최종 프레임 시간을 확인하여 시험 결과 기록지에 기록한다.
- step 13) 시간 효율성의 시험 절차를 step 1) 부터 step 12)까지 10회 반복하여 기록한다.

3.3 결과 분석 및 평가

수집된 시험 결과 자료는 엑셀(Microsoft Excel 2010)을 이용하여 기능 완전성을 분석하였고 시간 효율성은 통계 프로그램 Minitab 17 (Minitab Incorporated, Pennsylvania, USA)을 이용하여 일원배치 분산분석(One-way Anova)을 하였다.

기능 완전성 평가 방법은 시험 소프트웨어 어플리케이션 기능을 12 단계의 시험 케이스로 작성하였으며 각 단계별로 기능 동작에 문제가 없는지 확인하고 ISO/IEC 25023에서 제시하는 소프트웨어 품질 평가 기준을 적용하였다[6].

$$X = 1-A/B \tag{1}$$

Where, A number of functions missing, B number of functions specified

기능 완전성 시험 평가의 X 값은 시험자 3명이 모두 같은 결과인 경우는 만족으로 평가를 하였고 시험자 3명 중 1명이라도 결과가 다른 경우는 의심으로 평가를 하였으며 시험자 3명 모두의 결과가 다른 경우는 불만족으로 평가를 하였다(Table 3).

Table 3. Criteria of functional completeness evaluation

Result	Evaluation
$(X1 = X2 = X3)$	Satisfaction
$(X1 \neq X2 = X3)$ or $(X1 = X2 \neq X3)$ or $(X1 \neq X3 = X2)$	Suspicion
$(X1 \neq X2 \neq X3)$	Dissatisfaction

Tester 1: X1, Tester 2: X2, Tester 3: X3

시간 효율성 평가는 소프트웨어 어플리케이션의 동영상 생성 처리 기능을 11단계로 나누어 각각 10회 반복 수행하였으며 3명의 시험자들이 수행한 동영상 생성 처리의 평균 시간을 통계적으로 차이가 있는지를 기각치 F 값을 유의수준 5% 기준으로 일원배치 분산분석을 수행하여 판단하였다.

$$X = \sum_{i=1}^n (Ti)/n \tag{2}$$

Where, Ti time taken by the system to respond to a specific user task or system task at i-th measurement, n number of responses measured

시간 효율성 시험 평가 X 값은 시험자 3명의 모분산이 모두 같은 결과인 경우는 만족으로 평가를 하였고 시험자 3명 중 1명이라도 결과가 다른 경우는 불만족으로 평가를 하였다(Table 4).

Table 4. Criteria of time behavior evaluation

Result	Evaluation
$(\mu_1 = \mu_2 = \mu_3)$	Satisfaction
not $(\mu_1 = \mu_2 = \mu_3)$	Dissatisfaction

Variance of tester 1 : μ_1
 Variance of tester 2 : μ_2
 Variance of tester 3 : μ_3

4. 연구 결과

4.1 기능 완전성의 평가 결과

이 연구의 ISO/IEC 25023 소프트웨어 품질 평가에 대하여 시험자간 숙련도에 참여한 시험자 3명의 기능 완전성 시험 결과는 Table 5와 같다. 소프트웨어 어플리케이션의 기능 완전성 시험은 12단계의 기능들이 제대로 작동하는지를 각 단계별로 시험 케이스로 만들어 평가를 하였으며 시험자 3명 모두 12단계를 통과하였다. 평가에 사용된 전체 기능의 수 B값 12개와 문제가 발견된 기능의 개수 A값은 하나도 없는 것으로 확인되었고 최종 평가 X값은 1-A/B로 시험자 3명의 X값이 모두 1로 계산되어 만족으로 평가가 되었다.

Table 5. Result of functional completeness

Stage of measure	Tester 1		Tester 2		Tester 3	
	Pass	Fail	Pass	Fail	Pass	Fail
Step 1	1	0	1	0	1	0
Step 2	1	0	1	0	1	0
Step 3	1	0	1	0	1	0
Step 4	1	0	1	0	1	0
Step 5	1	0	1	0	1	0
Step 6	1	0	1	0	1	0
Step 7	1	0	1	0	1	0
Step 8	1	0	1	0	1	0
Step 9	1	0	1	0	1	0
Step 10	1	0	1	0	1	0
Step 11	1	0	1	0	1	0
Step 12	1	0	1	0	1	0
X=1-A/B	1		1		1	

Table 6. Results of difference in measurement evaluation between testers

구분	N	Mean	SD	F	p_value
Tester 1	10	9.45	0.01	3.02	0.065
Tester 2	10	9.52	0.01		
Tester 3	10	9.45	0.01		

SD : standard deviation

4.2 시간 효율성의 평가 결과

이 연구의 시간 효율성 결과는 Table 6과 같다. 소프트웨어 어플리케이션의 시간 효율성 시험은 11단계의 기능을 수행하고 4개의 이미지 파일을 선택하여 동영상 파일이 생성되는 시간을 시험 케이스로 만들어 10회 반복 측정하여 평가를 하였다. 각 시험자들이 수행한 동영상 생성되는 평균시간의 최종 시험 결과는 시험자 2가 9.52 ± 0.01로 평균 시간이 가장 높게 나타났으며 시험자 1과 시험자 2는 9.45 ± 0.01로 평균 시간이 같은 것으로 나타났다. 시험자들의 시험 결과 평균 시간에 차이가 있는지 일원배치 분산분석을 실시한 결과 통계적으로 유의하지 않은 것으로 나타났다(F=3.02, p = 0.065).

5. 논의

이 연구는 국제 표준 ISO/IEC 17043 숙련도 시험 요구사항을 이용하여 시험자간 숙련도 시험 절차 및 방법을 제시하였고 ISO/IEC 25023 소프트웨어 품질 특성을 이용하여 시험에 대한 평가 방법을 제시하였다. 시험 방

법은 동일한 한 개의 모바일 폰과 소프트웨어 어플리케이션 시료를 이용하였고 시험자는 시험 경력 및 전공을 고려하여 3명을 선정하였으며 시험 항목은 기능 완전성과 시간 효율성으로 시험을 완료한 후에 그 결과를 분석하여 평가를 하였다.

숙련도 시험의 평가로 기능 완전성의 경우는 12 단계의 어플리케이션 기능 테스트 케이스를 수행하였고 시험자 3명 모두 12 단계를 통과하여 시험자들의 숙련도에는 차이가 없는 것으로 평가가 되었다. 시간 효율성의 경우는 시험자 3명이 동영상 생성 기능 11 단계를 10회 반복 측정하였고 동영상 생성 처리 평균 시간들은 통계적으로 유의하게 차이가 없는 것으로 나타났으며 시간 효율성에 대한 시험자들의 숙련도에는 차이가 없는 것으로 평가가 되었다. 이러한 결과들로 숙련도 비교 시험에 참여한 시험자들의 숙련도는 동일하다는 결과를 도출 하였고 시험자들이 속한 그 시험기관 또한 동일한 결과를 기대할 수 있다는 것으로 해석할 수 있다.

시험기관에서는 숙련된 시험자의 객관적인 시험 수치를 통하여 결과를 제공해야 하고 때로는 숙련된 시험자의 눈을 통하여 의심스러운 결과들을 찾을 수도 있을 것

이다[10]. 이러한 숙련된 시험자의 능력을 객관적으로 입증할 수 있는 유용한 방법이 숙련도 시험 프로그램이며 객관적으로 입증된 숙련도 프로그램은 인정기구 및 시험기관, 이해관계자들에게 결과에 대한 신뢰성을 보장하게 되는 것이다. 따라서 시험기관은 신뢰성 있는 시험의 결과 제공을 위하여 지속적인 시험 능력을 반영할 수 있도록 숙련도 시험 프로그램에 자발적으로 참여하고 프로그램 개발에 노력해야 할 것이다[15].

우리의 연구는 몇 가지 제한점이 있다. 첫째, 본 연구는 국제 표준인 ISO/IEC 25023 소프트웨어 품질 특성의 평가를 이용하여 시험자간의 숙련도를 측정하는 방법을 제시했다는 점에서 연구의 중요성과 의의를 들 수 있다. 그러나 ISO/IEC 25023의 8가지 주특성 중에서 기능 완전성과 시간 효율성만을 대상으로 평가를 진행하였기 때문에 다른 주특성 평가 항목들도 추가로 연구가 진행되어야 할 것이다. 둘째, 일개 시험기관의 시험자들로 시험자간 숙련도 비교 시험을 진행하였기 때문에 여러 시험기관들 간의 비교 평가와 더 나아가 국가들 간의 비교 평가 연구도 진행되어야 할 것이다. 마지막으로 본 연구의 결과를 기초로 하여 다양한 인구학적 특성을 고려한 숙련도 비교 평가 연구와 비교 숙련도의 향상을 위한 전략적 프로그램을 개발하는 연구들도 추가로 제안하는 바이다.

이러한 제한점에도 불구하고 본 연구는 소프트웨어 시험 인정분야에서 시험자간 숙련도 비교 평가를 국제 표준 기반으로 적합한 시험 절차와 평가 방법을 제시하였으므로 향후 진행되는 연구들의 문제점을 보완하거나 기초 자료가 될 것으로 사료된다.

6. 결론

본 연구는 시험자간 숙련도 비교 평가를 위한 시험 절차 및 방법을 제시하였고 소프트웨어 애플리케이션의 기능 완전성과 시간 효율성 시험 결과로 시험자들은 동등한 결과를 획득할 수 있다는 능력을 보여 줌으로써 해당 시험기관의 소프트웨어 시험 결과는 신뢰성이 있다는 것을 보여 주었다.

오늘날 소프트웨어 분야는 융/복합 기술로 빅데이터와 인공지능, 자율주행 등과 결합하여 그 한계를 가늠하기가 어려울 정도로 복잡해지고 빠르게 진화를 하고 있다[16]. 이러한 현실에서 소프트웨어에 대한 품질 요구사항은 지속적으로 늘어나고 있으며 신뢰성 있는 시험 결

과를 획득하기에 더욱더 어려워지고 있는 실정이다.

우리나라는 글로벌 수준의 정보 기술과 인프라를 보유하고 있으며 정부 및 민간에서는 소프트웨어를 전략산업으로 선정하여 막대한 지원을 하고 있다. 하지만 이러한 노력에도 불구하고 소프트웨어 시험에 대한 지원은 상대적으로 부족한 것이 현실이다. 이러한 점을 고려하여 국내 산업 및 글로벌 시장에서 소프트웨어에 대한 경쟁력 있는 품질 및 신뢰성 확보를 위하여 시험분야에 대한 지속적인 관심과 연구가 진행되어야 할 것이다.

마지막으로 향후 연구 계획으로는 숙련도 시험 전용 애플리케이션을 개발하고 다수의 시험자들을 숙련도 시험에 참여하도록 하여 시험자들의 숙련도에 미치는 요인을 분석할 예정이다.

References

- [1] Nilaranjan Barik, Puspanjali Jena, Narayana Sethy. Assessing information communication technology (ICT) skills of degree science students of An autonomous college of ODISHA : A survey. VSRD International Journal of Library & Information Science, Vol. 1, Issue 1, 2015. Available From: <http://hdl.handle.net/10760/31112>
- [2] Bo Kyung Park, Woo Sung Jang, Ki Du Kim, Young Chul Kim. A Simplified Test Maturity Model (sTMM) for Small and Midsize Test Organization. Journal of KIISE, Vol. 45, No. 6, pp. 522-532, 2018. DOI: <http://dx.doi.org/10.5626/JOK.2018.45.6.522>
- [3] ISO/IEC 25000 "Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Guide to SQuaRE", 2014. Available From: <https://www.iso.org/obp/ui/#iso:std:iso-iec:25000:ed-1:v1:en>
- [4] ISO/IEC 25010, "Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models", 2011. Available From: <https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-1:v1:en>
- [5] ISO/IEC 25020 "Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality measurement framework", 2019. Available From: <https://www.iso.org/obp/ui/#iso:std:iso-iec:25020:ed-2:v1:en>
- [6] ISO/IEC 25023 "Software engineering: Software product Quality Requirements and Evaluation (SQuaRE) - Measurement of system and software product quality", 2016. Available From:

- <https://www.iso.org/obp/ui/fr/#iso:std:iso-iec:25023:e-d-1:v1:en>
- [7] ISO/IEC 25030 "Systems and software engineering — Systems and software quality requirements and evaluation (SQuaRE) — Quality requirements framework", 2019. Available From: <https://www.iso.org/obp/ui/fr/#iso:std:iso-iec:25030:e-d-2:v1:en>
- [8] Almira S, Nermina Zaimović-U, Hazim B. Proficiency testing and interlaboratory comparisons in laboratory for dimensional measurement. Journal of Trends in the Development of Machinery and Associated Technology, Vol. 16, No. 1, 2012. Available From: <https://www.tmt.unze.ba/zbornik/TMT2012Journal/24.pdf>
- [9] A. Shakhshiro, A. Fajgelj, U. Sansone. Comparison of Different Approaches to Evaluate Proficiency Test Data. Conference Paper, The Royal Society of Chemistry, Special Publication No. 307, Cambridge, pp. 220-228, 2007. DOI: <https://doi.org/10.1007/s00769-008-0413-7>
- [10] Ewa Szewczak, Adam Bondarzewski. Is the assessment of interlaboratory comparison results for a small number of tests and limited number of participants reliable and rational?, Accreditation and Quality Assurance volume 21, pp. 91-100, 2016. DOI: <https://doi.org/10.1007/s00769-016-1195-y>
- [11] A Diniță. Conformity assessment of the measurement accuracy in testing laboratories using a software application. 13th International Conference on Tribology, Materials Science and Engineering, 174, 2017. Available From: <https://iopscience.iop.org/article/10.1088/1757-899X/174/1/012069>
- [12] H.Y.Lee, H.S. Yang, Convergence Performance Evaluation Model for Intrusion Protection System based on CC and ISO Standard. The Journal of Digital Police & Management, Vol. 13, No. 5, 2015. DOI: <https://doi.org/10.14400/JDC.2015.13.5.251>
- [13] Jung-Hun Lee, Mi-Jin Lee. Time behaviour evaluation of software products considering service types. ICT Convergence (ICTC), Conference Paper, October 2012. Available From: <https://ieeexplore.ieee.org/document/6387129>
- [14] Jung, Hye-Jung. The Software Quality Testing on the basis of the International Standard ISO/IEC 25023. Journal of the Korea Convergence Society, Vol. 7, No 6, pp.35-41, 2016. DOI: <http://dx.doi.org/10.15207/JKCS.2016.7.6.035>
- [15] Kyunghwan Park, Cheolgyu Lee. A study on proficiency test sample development for enviromental and reliablility testing field. J Korean Soc Qual Manag, Vol. 45, No.3, pp. 447-462, 2017. DOI: <https://doi.org/10.7469/JKSQM.2017.45.3.447>
- [16] Chang-Hee Lee, Byeong-Yun Chang. The effect of the

self-measurement frequency levels on SmartCare obesity management. Journal of the Korea Academia-Industrial cooperation Society, Vol. 16, No. 3 pp. 1972-1980, 2015. DOI: <http://dx.doi.org/10.5762/KAIS.2015.16.3.1972>

이 창 희(Chang-Hee Lee)

[정회원]



- 2007년 8월 : 연세대학교 보건대학원 보건정보 전공 (보건학 석사)
- 2016년 2월 : 아주대학교 경영대학 MS/OM 전공 (경영학박사)
- 2015년 2월 ~ 현재 : 보건정보통계학회 이사
- 2010년 5월 ~ 현재 : LG전자 SW공학연구소 책임연구원

<관심분야>

SW 공학, SW Quality, Robot SW, Health Informatics, u-Health, CDSS (Clinical decision support system)

김 명 섭(Myung-Sup Kim)

[정회원]



- 2005년 2월 : 서강대학교 컴퓨터공학과 (공학석사)
- 2006년 1월 ~ 현재 : LG전자 SW공학연구소 책임연구원

<관심분야>

SW 공학, SW Quality, Automotive SW

김 윤 숙(Yoon-Sook Kim)

[정회원]



- 2004년 2월 : 홍익대학교 컴퓨터 공학 전공 (학사)
- 2011년 4월 ~ 현재 : LG전자 SW 공학연구소 책임연구원

<관심분야>

Test techniques and method, Functional Safety SW, SW metric, R&D Quality

송 성 현(Sung-Hyun Song)

[정회원]



- 2004년 2월 : 연암공과대학교 디지털 정보 전자과(전자)정보 통신 전공 (공업전문학사)
- 2003년 12월 ~ 현재 : LG전자 SW공학연구소 선임연구원

<관심분야>

SW 공학, SW Quality, Robot SW

노 아 름(A-Reum Noh)

[정회원]



- 2002년 2월 : 성균관대학교 전기 전자 및 컴퓨터공학부 (학사)
- 2008년 7월 ~ 현재 : LG전자 SW 공학연구소 책임연구원

<관심분야>

SW 공학, SW Quality, Robot SW, 클라우드

양 정 심(Jung-Sim Yang)

[정회원]



- 1999년 2월 : 인하대학교 공과대학 전자계산공학과 (학사)
- 2006년 5월 ~ 2016년: LG CNS
- 2016년 ~ 현재 : LG전자 SW공학 연구소 책임연구원

<관심분야>

SW 공학, SW Quality, Robot SW