

SPA 심사원 능력과 효율성에 관한 연구

이종무* · 박철수** · 김만술***

<목 차>

I. 서론	IV. 연구방법
II. 이론적 배경과 선행연구	4.1 자료수집과 분석방법
2.1 표준의 주요내용	4.2 분석결과
2.2 SPA와 심사내용	V. 결론
2.3 관련 선행연구와 1차실험	참고문헌
III. 연구모형과 변수	<Abstract>
3.1 모형과 가설	
3.2 변수와 측정항목	

I. 서론

성공적인 시스템 개발을 위해 필요한 고품질 소프트웨어(SW)를 개발하려는 산업계의 노력은 최근 표준개발과 사례적용을 통하여 많은 결실을 맺고 있다. 그 가운데서도 프로세스 품질개선을 위한 노력은 실제 산업계의 요구에 부응하고, 또한 적극적인 적용의지의 표출로 이어져, ISO/IEC 15504 국제표준¹⁾의 탄생을 가져왔다. 국내에서는 90년대 후반이후 지속적으로 프로세스 품질개선을 위한 표준 개발회의와 산업계적용에 적극 참여하여 왔고, 최근 다수 현장에서

SW 프로세스 심사(SW Process Assessment: SPA)를 직접수행해 왔으며, 현재도 이에 필요한 표준교육과 공인심사원(Certified Assessor)의 국내 양성이 지속적으로 이뤄지고 있다.

SPA란 조직의 SW 프로세스를 일정 기준이나 기본 틀을 근간으로 만들어진 심사모형에 따라 평가하는 것이다(정호원, 1999; ISO, 2004). 이와 관련된 대표적 모형이 ISO/IEC 15504이며, 이는 1992년 이후 국제표준 제정을 목표로 진행된 SPICE 프로젝트의 명칭과 혼용되기도 한다. 이는 프로세스의 객관적 능력결정과 개선을 위해 ISO SC7/WG10이 승인한 프로젝트로서, 소

* 한라대학교 경영학과, 교수, jmlee@halla.ac.kr

** 한라대학교 경영학과, 교수, 교신저자, cspark@halla.ac.kr

*** 한라대학교 경영학과, 교수, mkim@halla.ac.kr

1) 이하 표준으로 통칭한다.

소프트웨어의 개발과정에서 보다 효과적인 프로세스 품질관리 능력을 획득하기 위한 산업현장의 절실한 요구에서 비롯되었고, 이미 국제적으로 3차에 거친 대규모 현장사례 적용을 위한 다수의 공식심사(Trial)와 표준 제정원칙에 따른 표준화가 이뤄진 바 있다(정호원, 2007).

이러한 표준 심사에서 자주 제기되는 문제 가운데 우선 해결해야 할 과제로는 표준적용의 신뢰성과 효과성, 그리고 심사과정의 효율성 등을 들 수 있다. 여기서 신뢰성이란 표준자체에 관한 신뢰성과 함께 결과도출을 위한 심사원간의 상호합의에 관한 신뢰성을 의미하며, 표준효과성은 개발된 심사모형의 기준이나 기본 틀에 대한 적합성(Conformity)과 개발 심사모형의 적용성(Applicability)에 관한 것이다(정호원, 1999; Emam, 1998; SPICE, 1998).

특히 이 가운데 경험적으로 실제 SPA에서 가장 많이 제기되는 문제는 표준 신뢰성이나 효과성보다는, 심사결과에 대한 객관적인 신뢰성의 문제이다. 이는 우선 심사를 수행하는 팀 단위의 복수심사원들의 심사결과가 동일한 배경 하에서 다른 시점에서도 지속적으로 같은 결과로 제시되는가에 관한 반복성(Repeatability)과 다른 심사원들의 경우에도 역시 동일한 결과도출이 가능해야 하는 재생성(Reproducibility) 등의 보장을 통해서 확인될 수 있다고 본다(정창신, 1999; 정호원, 1999; Jung, 2003). 또한 이와 함께 심사의 효율성 문제도 대두되는데 이는 SPA에서 심사일정과 비용 등의 제약조건(ISO, 2004)을 고려하여 투입된 심사자들이 과연 합리적으로 활용되고, 복수 심사원들의 합의를 통한 신뢰성 있는 심사결과의 도출이 가능한가에 관한 문제이다. 현재 국내 표준 심사에서는 이러한 이유에

서, 복수의 심사원들로 심사 팀을 구성해 심사수행을 하도록 권장하고 있다(정학중, 1998). 그러나 팀을 구성하는 심사원들의 개별특성, 예를 들어 사전 심사경험의 정도, 소프트웨어 공학지식의 수준, 기타 심사 대상조직에 관한 지식정도에 따라서는 심사 효율성과 신뢰성에 많은 차이를 나타낼 수 있지만, 이를 실증적으로 확인한 바는 없다.

본 연구에서는 표준(ISO, 2004)에서 제시한 심사원능력(Competence)과 특성에 따라 과연 효율적 심사진행과 신뢰성있는 합의도출이 가능한지를 탐색적 차원에서 확인해 보고자 한다. 이를 위해 SPA 심사원들이 인지하는 심사효율성에 영향을 주는 요소들을 표준의 정의에 따라 심사원 및 조직특성, 표준지식, 그리고 심사기술 등의 요인 가운데 주로 심사효율성에 많은 영향을 미치는 요인이 무엇인지, 또한 이들 간의 인과관계가 어떻게 존재하는지를 확인해 보고자 한다.

이를 위해 2장에서 이론적 배경과 선행연구들을 살펴보고, 3장에서는 연구모형과 관련된 변수 및 측정항목, 그리고 관련가설을 제시한다. 4장에서는 연구방법으로서 자료수집과 분석, 연구한계와 기타사항을 논의하고, 5장은 결론으로 구성한다.

II. 이론적배경과 선행연구

2.1 표준의 주요내용

일반적으로 IT 프로젝트의 실패는 낮은 품질의 SW, 비용의 증대, 일정의 지연 등과 같은 공

통적인 문제들로 요약해 볼 수 있는데, 이를 위한 해결책 가운데 가장 효과적인 대안의 하나로 제시되는 것이 바로 SW 프로세스 개선(SPI)이다 (Pfleeger, 1994; Rout, 2007). 최근 이를 위한 가장 큰 산업계의 노력은 크게 두 가지의 결실로 나타나 산업계에서 현재 활발히 적용되고 있다. 그 하나가 ISO에서 개발한 표준²⁾을 들 수 있으며, 또 다른 하나는 미 국방성의 지원에 따라 SEI가 개발한 CMMI-DEV(Capability Maturity Model Integration for DEvelopment)등이 있다 (CMMI-DEV, 2006). 현재 이들 결과물들의 주요 내용은 상당히 유사한 측면이 있으며, 모두 SPI를 위한 것으로 이해할 수 있다(정호원, 2007).

SPI란 객관적인 표준에 근거한 심사모형을 바탕으로 SPA를 실행하여 지속적인 프로세스 개선을 이뤄내고, 궁극적으로는 고품질의 SW개발을 지향해 나간다고 목적을 이루기 위한 SW공학적인 노력이라고 설명할 수 있다. 표준에 따르면 높은 SW프로세스 품질추구는 일련의 프로세스 목표의 인식, 방법의 정의, 프랙티스들의 확립, 그리고 지속적인 개선(Haase, 1996; Mackie, 1997) 등의 제 과정을 통해 이뤄지며, 구체적으로는 SW개발 생명주기에 근거한 프로세스 계획, 관리, 실행, 통제 및 개선에 관한 내용을 포함하고 있다.

참고적으로 표준탄생의 가장 큰 배경은 프로젝트의 규모나 조직특성에 관계없는 품질개선 활동의 마련과 보다 객관적인 개발 성숙도를 평가하기 위한 지침의 필요성(Emam, 1998;

Haase, 1996; SPICE, 1998; Mackie, 1997)을 들 수 있다. 이를 위해 참조 가능한 일반적인 심사 모형으로는 과거 Bootstrap, Trillium, ISO 9000, 그리고 CMM 등의 다수를 참고할 수 있지만, 국내의 경우 심사규모나 범위의 일반성, 목표의 포괄성 등의 장점 때문에 1997년 이후, 표준에서 제시하는 심사모형에 근거한 심사원교육과 함께 조직 내·외부 심사원들에 의한 산업현장의 SPA가 활발히 진행되고 있으며, 2006년 이후 KS표준으로 활용되고 있다(SPICE, 1998; 한국 SPICE, 1998; 한국SPICE, 2006; 정호원, 2007).

이는 기존의 프로세스 개선이나 능력결정을 위한 심사방법들과는 달리 심사대상 프로젝트의 크기나 심사목적에 제한성을 두지 않으며, 심사 결과 또한 ISO 9000과 같이 인증에 그 목적을 두지 않는다는 장점이 있다. 예를 들어, 과거 SEI의 CMM이나 BELL Canada의 Trillium, 혹은 ESPRIT의 Bootstrap 등의 대상이 주로 대형 혹은 중형 프로젝트에 한정되는 데 비해, SPICE는 프로젝트의 크기에 제한되지 않으며, 또한 심사목적 역시 Bootstrap이나 STD가 프로세스 개선에 목적을 두는 데 비해서 프로세스 개선과 능력결정 모두를 목적으로 삼고 있다. 또한 소프트웨어 획득, 공급, 개발, 운영, 발전, 지원 능력개발 및 지속적 개선을 원하는 모든 소프트웨어 조직에 적용할 수 있고, 특정 기술이나 방법론 등을 사전에 가정하지 않는 등 포괄적인 심사지표를 포함(한국SPICE, 2006)하고 있어, 매우 유연하고 경제성이 확보되는 SPA 방법이라고 할 수 있다.

2) 표준개발 초기부터 SPICE(Software Process Improvement and Capability determination)란 약칭을 함께 사용해왔으며, 현재는 system 프로세스의 개념으로 변경되었다. 본 연구에서는 국가 규격(KS)인 SPICE 표준을 적용한 사례를 중심으로 설명하고, 주요 용어는 대부분 약어로 표기됨으로 본 연구에서도 이를 사용한다.

일반적으로 표준 제정에서 앞서 요구되는 기업 현장에서의 유용성 확인과정을 위해서, 해당 표준 역시 90년대 이후 전 세계적으로 3차에 걸친 시범적용과 성과물의 DB화, 그리고 이의 연구분석과 이를 반영한 후속 개정작업 등이 이루어져 왔다(Jung, 2001).

그 결과 프로세스 개선을 위한 전 세계 산업현장의 성공사례에서 파악된 기본프랙티스(BP)들을 중심으로 3개 생명주기 범주(기본, 조직, 지원)의 9개 그룹, 총 48개 프로세스들의 도출되었고, 이를 SPI에 적용하기 위한 참고모형과 예비심사모형 등이 표준으로 제시되었다. 참고로 범주별 세부 프로세스를 살펴보면, 기본 생명주기 프로세스 범주로는 획득, 공급, 엔지니어링, 운영 그룹과 하위 22개 프로세스가 있으며, 조직 범주에는 관리, 개선, 자원 및 기반구조, 재사용 그룹과 하위 16개 프로세스들이, 그리고 지원 범주에는 지원 그룹의 10개 하위 프로세스들이 있다(ISO, 2004).

2.2 SPA와 심사모형

SPA 수행을 통한 프로세스 개선을 이루려면, 우선 현 상태의 파악과 프로세스 수행능력의 적합성 등을 판단하기 위한 심사모형의 활용이 필요하다. 표준에 기반한 심사수행에서는 참조모형과 구체적인 심사지표들로 구성되는 심사모형을 기반으로 이뤄지는데, 여기서 참조모형은 ISO/IEC 12207(ISO, 1995)에 기초한 프로세스와 각 프로세스들의 능력을 특징짓는 프로세스 속성(PA)을 통해 능력수준을 설명하는 2개 차원으로 구성된다.

전자는 소프트웨어 생명 주기상의 프로세스

0범주별 구분과 프로세스 목적을 정의하고, 이를 또한 측정 가능한 용어들로 설명하며, 후자는 심사 대상조직의 특정 프로세스를 달성하거나 혹은 달성 목표로서 사용 가능한 프로세스 능력을 설명한다(ISO, 2004).

2.2.1 프로세스 및 프로세스 능력 차원

참조모형에서는 소프트웨어의 개발, 유지, 획득, 공급 및 운영 관련 조직의 프로세스들을 전체적으로 3개 주요 프로세스 그룹과 5개 범주 그리고 하위 프로세스들로 구분해 정의하고 있다. 3개 생명주기 프로세스 그룹이란 주요 생명주기, 지원 생명주기, 그리고 조직 생명주기 프로세스들이며, 활동 유형별 구분으로는 고객-공급자, 공학적, 지원, 관리, 그리고 조직 등의 5개 범주별 프로세스 등이 있다.

주요 생명주기 프로세스 그룹에는 고객-공급자 범주와 공학적 범주의 프로세스들이 관련되는데, 전자는 고객에 직접 영향을 주는 프로세스들로서 소프트웨어 개발과 대 고객 전달관련 지원 프로세스, 그리고 시정활동의 제공과 소프트웨어 제품 그리고/혹은 서비스 사용 관련 프로세스들로서 구성된다. 후자의 프로세스는 소프트웨어 제품, 시스템과의 관계, 고객 문서화 등을 직접 명세하거나, 구현 혹은 유지 활동과 관련된 프로세스들로서 시스템 전체가 하나의 소프트웨어로 구성되는 환경 하에서는 단지 소프트웨어의 구축 및 유지 활동에 관련되는 프로세스들이 해당된다. 지원 생명주기 프로세스 그룹은 소프트웨어 생명주기상의 다른 프로세스들에 의해 사용될 수 있는 지원 프로세스 범주의 프로세스들로 구성되고, 조직 생명주기 프로세스 그룹은 생명주기상의 프로젝트 혹은 프로세스 관리에

사용될 수 있는 일반적 성격의 활동들을 가지는 프로세스들로 구성되는 관리 프로세스와 전체 조직의 경영목표의 수립과 조직의 경영목표 달성에 관련되는 조직차원의 프로세스의 2개 범주의 프로세스들로서 구성된다.

한편 프로세스 능력수준과 그 발전은 능력수준별로 해당 속성들로 표현된다. 이는 프로세스의 목적과 경영목표 효과성을 관리 및 개선하는 전반적인 능력 수준별 각 측면을 설명한다. 따라서 프로세스 수행능력 개선에 따른 수준의 향상은 해당 속성(들)의 결합으로 나타나게 되며, 세부 능력수준은 프로세스 목적 달성이 전반적으로 실패한 수준인 Level 0(불완전수준)에서 현재와 미래 경영목표에 맞는 최적화된 프로세스를 수행하는 수준인 Level 5(최적 수준)까지의 6개 수준으로 구분된다.

2.2.2 심사지표와 등급결정

참조모형의 각 프로세스는 정의된 목적 기술문을 통해 그 내용을 알 수 있는데, 이는 특수여건 하에서 프로세스의 독특한 기능적 목표를 포함하며, 또한 프로세스 산출물 식별에 필요한 추가적인 자료를 모두 포함한다. 기본적으로 프로세스 목적의 만족은 프로세스 능력 구축의 첫 단계이며, 따라서 이를 위한 적절한 지표들의 확인은 해당 프로세스 수행 확인의 첫걸음이 된다. 이에 관한 심사지표로는 해당 프로세스별 BP와 입출력 작업산출물(WP)들이 활용된다.

프로세스 목적 만족 이후는 프로세스 능력수준별로 해당 관리지표들의 확인을 통해서 수준의 향상 정도를 파악할 수 있다. 즉, 공통프랙티스(GP)와 공통자원(GR), 그리고 공통작업산출물(GWP)의 증거들을 통해 수준별 PA의 달성정

도를 수준별로 파악할 수 있다. 이들 심사모형, 참조모형의 관계 및 구성, 그리고 심사지표들의 내용을 도식화하면 <그림 1>과 같다.



<그림 1> 심사모형과 심사지표(ISO, 2004)

일반적으로 Level 2와 Level 5 사이의 능력수준과 관련된 PA를 성취하려면, 일반적으로 지원, 관리, 그리고 조직차원의 하나 이상의 프로세스 범주에 속하는 프로세스의 수행이 관련된다. 또한 Level 2 속성들은 관리 및 지원차원 프로세스 범주의 프로세스들과 관련이 있으며, 그 이상 수준의 속성들은 조직차원 범주의 프로세스들과 관련이 있다. 결국 프로세스 능력 수준의 향상에 따라 각기 다른 범주의 프로세스들의 중복 연관성이 나타날 수 있으며, 이를 합리적인 우선순위에 따라 심사 수행과정에서 확인한다면, 보다 효율적인 심사의 진행과 원활한 합의도출이 가능할 수 있다. 즉, 낮은 능력 수준의 심사의 경우에는 상대적으로 높은 수준을 가진 조직의 심사보다는 쉽게 BP와 GP의 중요도 우선순위에 따라 결과를 도출할 수가 있다.

2.2.3 심사원 능력과 팀 구성

앞서 언급한 대로 SPA는 일반적으로 외부심사원 중심의 팀에 의한 심사를 수행한다. 그 이

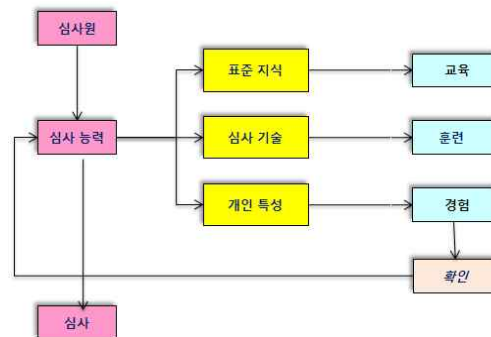
유는 개방적인 면담과 토의, 그리고 합의에 의한 심사가 신뢰성 확보에 유리하다는 기대에서 비롯된 것이지만, 사실 이를 국내 개발현장의 심사 사례에서 직접 실증적으로 확인한 바는 없다. 그러나 팀을 구성하는 심사원들의 개별특성, 예를 들어 심사조직특성, 이전 심사경험, 소프트웨어공학 지식수준, 기타 심사 대상조직에 관한 이해정도에 따라서는 심사 신뢰성과 효율성에 많은 차이를 나타낼 수 있을 것이며, 이는 표준에서의 심사원능력과 심사 효율성의 관계 정의에서도 나타나 있다.

SPICE심사에서 책임심사원의 주요한 책임과 역할은 참여한 심사원들의 적절한 심사수행의 확인과 감독, 그리고 합의도출 등이다. 특히 책임심사원 뿐만 아니라 참여심사원들의 심사경험과 지식, 심사 참여의 적극성 정도 등이 각 심사 사례별로 서로 다르므로, 제한된 심사일정을 고려할 때 심사원들의 합리적인 팀 구성은 심사성패에 많은 영향을 줄 수 있다. 이는 실제로 심사에 앞서 우선 해결해야할 문제이지만, 정규심사원 교육에서는 경험자와 유경험자를 적절히 고려한 팀 구성만을 통상 권고할 따름이다(한국 SPICE, 2006).

표준에서 제시하는 심사원 능력에 관한 지침(ISO, 2004)에 따르면, 심사원은 심사를 수행하면서 심사원 능력을 나타내 보일 수 있어야 한다. 여기서 능력이란 “첫째, 표준 프로세스에 관한 지식, 둘째 프로세스 참조모형, 심사모형, 심사방법과 도구, 그리고 등급결정 프로세스 등을 포함한 표준의 능숙한 심사기술, 셋째 SPA의 효과적인 수행에 기여하는 개인특성 등을 통하여 기인하는 것”으로 정의한다. 또한 이러한 능력은 표준교육과 SPA훈련, 그리고 SE 및 IT경력을 포

합한 개인적 경험특성에 의해 갖춰질 수 있으며, 따라서 객관적인 심사원들의 능력은 심사원 표준교육, 심사관련 기술훈련과 기타 전문경험 등을 통해 확인이 가능하다고 본다.

따라서 SPICE 심사에서는 일반적으로 심사원들의 자격심사 기준으로 이를 직접채택 적용하고 있으며, 또한 심사 팀의 구성시 심사원의 평균적 능력을 객관화하기 위해서 위의 정의에 입각한 예비심사원 시험 및 훈련, 그리고 보수교육을 지속적으로 적용해 나가고 있는데, 이에 관한 세부사항 및 이들의 관계는 표준에서 정의한 도표를 참고해 볼 수 있다.



<그림 2> 심사원 능력과 특성(ISO, 2004)

2.3 관련 선행연구와 1차 실험연구

SW공학 연구에서는 표준개발을 추진하는 기관의 초기 연구목적으로서 소위 일반화를 공통적으로 거론하고 있다. 이는 표준의 객관적 유용성 및 앞서 언급한 표준의 신뢰성과도 밀접한 관련이 있다. 예를 들어 SPICE 표준개발 및 연구는 국제 연구컨소시엄(IRC)이 주축이 되어 객관적 일반화를 추진해 왔으며, 이에 관한 내용은 Jorgensen의 연구(2004)를 참고해 알 수 있다.

또한 SPICE의 실증적 유용성에 관한 연구들은 과거 시범적용 결과보고서, Jung(2001)의 연구, 그리고 후속 연구들(Rout, 2007; Jung, 2003)을 참고해 볼 수 있다.

이밖에 SW 제품품질에 관한 선행연구들은 비교적 더 다양하게 이뤄지고 있다. 예를 들면, 최근의 제품품질 표준의 유용성과 시범적 적용을 위한 선행연구(Al-Kilidar, 2005; Valenti, 2002) 들을 통해 표준적용의 유용성에 관한 내용을 참고할 수 있으며, 기타 품질특성별 모형개발, 특성과 해당 속성간의 연결을 통한 개발평가, 도구구현, 모형과 타당성 등 다양한 연구(Boehm, 1978; Chua, 2004; Jung, 2004; Jung, 2007)도 표준의 적용방법을 확인할 수 있다.

그러나 최근 표준개발 이후 프로세스품질의 실증분석(Emam, 2000; Jung, 2001; Jung, 2003)과 SPA 산업현장 적용에 관한 연구결과(Jorgensen, 2004; Rout, 2007)는 지속적으로 축적중이지만, 본 연구와 직접 관련되는 표준을 적용한 SPA나 심사원 특성과 효과성에 관한 실증 연구는 현재까지 미진한 상태이다.

본 연구에 앞서 우선 1차 실험연구를 진행하였는데, 실제 K기업의 현장심사에 참여한 선임 심사원 포함 외부심사원 6명과 내부심사원 4명에게 개별 설문조사를 통해 심사원특성과 함께 해당 심사원능력 요소들의 중요도를 측정하였다. 해당 조직은 비교적 전반적인 담당자들의 소프트웨어 공학적 지식과 프로세스에 대한 지식이나 경험 수준이 아직은 높지 않으며, 기존 시스템 개발 및 유지보수 경험이 비교적 낮은 수준이지만, 프로세스 개선 및 능력 개발에 관한 목표의식은 매우 높은 특징을 아울러 가지고 있었다.

우선 심사 대상 프로세스의 각 프랙티스별 특성을 ID와 함께 제시하고, 개별 심사참여자들에게 순위법에 따라 심사지표의 중요도를 평가하도록 하였다. 이는 순위화 평가가 등급화나 기타 다른 평가방법에 비해 사용의 편리함과 아울러 평가의 변별력도 뛰어난 장점을 갖고 있으며, 순위자료는 다수 평가자의 다수 대상에 대한 평가자료로서 통계적 다변량 분석이 가능하기 때문이다(Johnson, 1992).

본 연구에서는 순위자료들의 상관관계를 알아보기 위해서, 몇 개의 가상 변수와의 관계를 가지고 상관관계를 재 정의한 인자분석을 통한 최적변환 처리(Johnson, 1992)로 1차 연구결과를 분석해 보았다. BP를 중심으로 분석한 결과, 프로세스 심사경험과 SE경력, 그리고 심사원간의 사전 문서검토를 행한 경우와 그렇지 않은 경우, 그리고 외부심사원과 내부심사원간의 견해 차이가 일부 존재하는 것으로 나타났다.

따라서 이를 고려해 예를 들어, 심사원간의 사전검토나 표준에 관한 이해가 보다 손쉬운 합의와 의견일치를 이끌어 낼 수 있을 것이며, 이는 결국 심사효율성을 높여 줄 수 있을 것이라고 판단하였다. 본 연구에서는 이러한 1차 사전연구와 분석결과를 참고하여, 실증적으로 일반화된 연구결과의 도출이 가능한지 확인해 본다.

III. 연구 모형과 변수

3.1 모형과 가설

이전 심사경험들을 통해 심사원 다수가 공감하는 SPA의 성패는 몇 가지로 요약해 볼 수 있

다. 과연 일반화되고 신뢰성 있는 표준의 제공이 가능한가, 또한 이를 적절히 교육 및 훈련받아 산업현장의 SPA에 적용할 수 있는 심사원의 능력이 갖춰져 있는가, 그리고 심사의 효과성이 객관적인 실증자료로 분석·제공될 수 있는가 등에 달려 있다고 본다.

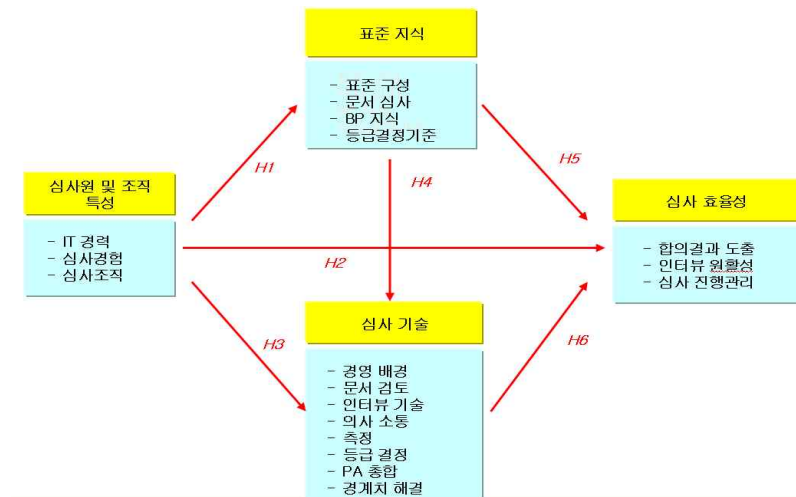
2장에서 살펴본 바와 같이, SPICE는 성공적인 산업현장의 사례적용과 분석을 통해 실증적으로 프로세스품질에 관한 객관적이고 신뢰성 있는 일반화된 표준(Jorgensen, 2004)으로서 개발되어 활용중이며, 심사의 효과성에 관한 실증적 결과는 예산, 일정, 고객만족, 생산성, 업무만족/직원사기, ROI 등의 성과와 관련된 분석연구(Jung, 2001; Emam, 2000)를 통하여 다수 제공되고 있다. 그러나 아쉬운 것은 심사자체의 신뢰성과 효율적 심사진행에 영향을 줄 수 있는 심사원들의 특성과 능력에 관한 실증적 연구이다.

따라서 본 연구에서는 과연 표준에서 제시하는 심사원 능력에 심사원 개인 및 조직특성이 영향을 주는지와 심사원 능력들이 심사효율성에

어떠한 영향을 주는지 분석해 보고자 한다. 심사원 능력이란 표준에 관한 이해와 지식, 그리고 심사수행을 위한 기술 등이 갖춰지는 수준이라 정의하고, 이들 능력에 따라 만족스런 합의결과와 도출과 원활한 인터뷰 진행, 그리고 전반적인 효율적 심사진행과 관리 등으로 측정되는 효율적인 SPA가 과연 수행되는지를 확인해 보고자 한다.

현재 대부분의 SPICE 심사에서는 팀단위 심사가 진행되는 관계로 심사일정의 구성을 사전 문서검토와 팀별 필요 인터뷰, 그리고 그 이후 결과도출을 위한 등급결정 과정을 갖도록 되어 있다. 따라서 능력과 관련되는 심사기술도 문서검토, 인터뷰, 그리고 최종 등급결정에서 필요로 하는 다양한 기술들을 대상으로 심사효율성에 미치는 영향을 살펴봐야만 한다.

표준의 내용에 관한 이해와 지식에 관한 항목 들로는 SPICE의 구성과 이해, 문서심사 지식, 프로세스 수행과 능력에 관한 지표인 BP와 GP의 이해, 프로세스 등급수준의 세분화측정을 위



<그림 3> 연구 모형

<표 2> 연구 가설

가설	내 용
H1	심사원 및 조직특성은 표준지식에 긍정적 영향을 준다.
H2	심사원 및 조직특성은 심사효율성에 긍정적 영향을 준다.
H3	심사원 및 조직특성은 심사기술에 긍정적 영향을 준다.
H4	표준지식은 심사기술에 긍정적 영향을 준다.
H5	표준지식은 심사효율성에 긍정적 영향을 준다.
H6	심사기술은 심사효율성에 긍정적 영향을 준다.

한 PA 지식, 그리고 등급체계(Rating Scheme)에 관한 지식 등을 측정해 본다.

심사기술과 관련한 세부 측정문항은 심사원들의 문서검토와 인터뷰실시와 관련된 경영배경(Business Context) 이해, 문서검토 방법과 기술, 일정시간 준수능력과 인터뷰 제약조건 등으로 구분하여 측정하고, 등급결정과 관련된 기술은 측정과 합의, 등급결정, PA의 총합, 그리고 시범적용에서 제시된 경계치문제(Boundary Problem) 해결 방법 등을 통하여 측정하도록 하며, 이들이 과연 상호관련성이 있는지와 최종 독립변수로서 SPA 비효율성에 미치는 영향을 분석해 보도록 한다.

질문형식은 실제 심사 중 신속한 답변을 얻기 위해서, 표준 용어정의(ISO, 2004)에 기초하여 1차 사전연구에서 확인된 항목들을 중심으로, 설문에서는 심사원 및 조직특성에 따라 심사원 능력과 심사효율성에 미치는 영향정도를 5점 척도로 측정하였다³⁾. 앞서 구성된 연구모형에 따라 실증적으로 확인할 가설을 요약해 보면 다음 <표 2>과 같다.

3.2 변수와 측정항목

모형에서 정의된 변수와 관련된 측정항목은 표준과 심사원 심사실습 및 실제 심사에서 자주 제시되는 주요 문제점과 심사결과 신뢰성에 영향을 주리라 기대되는 항목 가운데, 이전 1차 실험 연구결과 확인된 요소들을 선택하였다(Emam, 1998; ISO, 2004; 한국SPICE, 2006).

이에는 전반적인 심사결과와 확신, 표준의 이해도와 심사대상 프로젝트의 내용이해, 문서 심사, 문서 심사시간, 인터뷰 기술과 적용능력, 인터뷰 시간, 심사 일정 및 진행, 팀원 간의 의사소통, 등급결정(주관적 판단과 객관적 정량화)문제들과 BP, GP, PA의 이해와 신뢰성, 평점 결정, PA총합 결정, 등급 정량화를 위한 경계문제해결 등의 최종등급 결정방법 이해와 적용 등에 관한 것이 모두 포함된다. 해당변수와 지표명, 측정항목과 관련지표의 구분은 변수명, 지표명, 측정항목과 지표유형 열로 나뉘 <표 2>와 같이 요약 정리할 수 있다.

3) 설문에서는 신속한 응답을 위해 선임심사원의 권고를 참고하여, 심사원능력 가운데 어떤 항목의 능력이 부족하여 심사 비효율성을 초래하는가를 질문한 후, 코딩과정에서는 역수값으로 분석하였다.

<표 3> 변수와 측정항목

변수명	지표명	지표유형	측정항목
PERSON ¹⁾	IT_Career	조형	IT 혹은 SE경력 수준 (상중하)
	SPA_Exp	조형	프로세스심사의 과거경험 수준 (상중하)
	Group	조형	심사 대상조직의 특성 차이 (프로젝트별)
IS ²⁾	Standard	반영	SPICE 표준 구성과 전반적 이해
	Document	반영	문서심사에 관한 이해
	BP	반영	수행 프랙티스와 작업산출물에 관한 지식
	Rating_Criteria	반영	등급결정 기준체계의 이해
SKILL ³⁾	Bus_Context	반영	심사조직의 경영배경에 관한 인터뷰 평가
	Time_Doc	반영	문서검토를 통한 심사관리기술 (시간과 투입요소활용)
	Time_Interview	반영	인터뷰에 관한 심사관리기술 (시간과 투입요소활용)
	Team_Comm	반영	팀별 심사원 간의 의사소통과 문제해결
	Measuring	반영	팀별 프로세스 측정
	GP_Rating	반영	공통 프랙티스의 등급 결정
	PA_Aggregation	반영	프로세스 속성의 총합과 적용
	Rating_Boundary	반영	등급결정 경계치 관련 해결방법
SPA ⁴⁾	SPA_Conf	조형	효율적인 팀 합의에 관한 확신
	Interview	조형	원활한 인터뷰 시간관리와 진행
	SPA_MG	조형	전반적인 프로세스심사의 진행과 관리

주) ¹⁾심사원 및 조직특성; ²⁾표준지식; ³⁾심사기술; ⁴⁾심사효율성

IV. 연구 방법

4.1 자료수집과 분석방법

본 연구에서는 L사 심사에 참여한 15개 팀 전체 30명 심사원들을 대상으로 현장설문을 통하여 2개 프로젝트에 관련된 자료수집이 이뤄졌다. SPA 비효율성 측정을 위한 개인적 특성, 표준이해, 심사기술과 지식, 그리고 기타 심사결과에 관한 신뢰수준, 심사조직특성 등 총 20개 측정항목을 각각 5점 척도로 수집되었다.

심사 대상조직인 L사는 각각 정보통신 및 SI 관련 개발을 주사업으로 삼고 있으며, 오랜 기간

자체적인 품질보증 지원체계와 관리경험을 갖추고 있어, 1차 실험연구 심사조직과는 달리 다양하고 높은 능력수준과 등급결정이 예상되었다. 실제 L사의 심사사례는 상대적으로 높은 관리능력과 성숙수준을 보유한 것으로 평가되었고, 심사과정에서도 서로 상이한 2개 프로젝트를 동일한 심사일정에 따라 15팀의 심사원들이 독립적인 심사를 수행하였어도, 원만한 합의와 일정준수가 지켜진 성공적인 심사사례였다.

수집된 자료의 특징을 살펴보면, 참여한 SPICE 심사원들의 평균 심사횟수는 약 3회, 평균 SE 경력 9.6년 수준이었고, 자료처리와 관련하여 수집된 일부자료의 결측치 대체는 산술평

균에 의한 단순결측치 대체방법을 사용하였으며, 연구모형상의 경로분석을 위하여 인과관계 규명에 매우 유용한 연구방법의 하나인 PLS(Partial Least Squares)를 사용하였다. PLS는 다른 구조방정식 모형도구들에 비하여 다양한 장점을 갖는 것으로 평가되는데(Chin, 1999; Gefen, 2000; Hulland, 1999), 우선 변수측정을 위한 지표설정과 오류문제 해결이 상대적으로 용이한 장점을 들 수 있다. 또한 탐색적 차원의 이론과 모형개발에 매우 적합하며, 적은 표본수로도 분석이 가능하고, 다양한 분석도구의 사용도 가능하여 최근 관심이 매우 높다.

본 연구에서는 앞서 언급한 대로 반영지표 뿐 아니라 조형지표에 의한 분석연구가 요구되며, 연구표본의 제한성⁴⁾을 고려한 분석방법이 필요하고, 표준의 심사원능력에 관한 변수 간 경로관계에 관한 실증분석과 심사원 개인적 특성의 간접적 영향도 확인하기 위해, PLS를 통한 모형검증과 인과분석을 수행하였다.

기존 인과관계 측정에 관한 구조방정식 연구(Chin, 1998a; Chin, 1998b; Diamantopoulos, 2001)에 따르면 두 가지 측정지표, 즉 조형지표와 반영지표의 구분에는 매우 세심한 주의가 필요하며, 경험적으로 잠재변수와 지표설정이 연구의 성패에 미치는 영향은 매우 크다는 것을 알 수 있다. 전자는 변수를 구성하는 측정항목들이 각각 변수를 통해 측정하려는 구성개념을 설명하고, 따라서 원인제공의 역할을 하는 경우에 해당되며, 후자는 측정항목이 구성개념의 측정결과가 반영되어 나타나는 특징을 갖는 경우에 해당된다.

예를 들면, 개인별 특성을 구성하는 요인으로는 심사경험이나 IT경력 등이, 그리고 심사효율성에 대해서는 팀합의 확신이나 전반적인 심사 진행 등이, 그리고 심사기술과 관련하여서도 다양한 문서, 인터뷰, 측정기술들이 영향을 주는 것으로 볼 수 있는데, 이는 그 특성상 조형지표로 가정해 볼 수 있다. 이에 비해, 심사원 능력을 구성하는 변수인 표준지식이 부족하여 낮은 수준이라면 그 결과 구성항목 값도 다 함께 낮아질 수 있는 즉, 반영지표로 이해할 수 있을 것이다.

4.2 분석 결과

본 절에서는 과연 연구모형의 내용을 지지할 수 있는가를 세부적 경로분석을 통해 분석해 본다. 이를 위해서는 우선 수집된 자료의 경로분석 결과는 측정모형 분석을 통한 신뢰성 및 타당성 확인을 거쳐서, 설정된 인과관계의 경로에 대한 효과 및 전반적인 적합도, 그리고 추가적인 조절 효과 등을 평가하기 위한 구조모형 분석 및 최종 결과 논의 순으로 해석해 본다.

4.2.1 측정모형의 분석

우선 변수에 대한 타당성을 검증하기 위하여 탐색적 요인분석을 실시하였다. 주성분 분석법을 이용한 요인추출을 수행하였는데, 직교회전법을 이용한 결과 적재량의 값이 최소 0.4이상인 항목들로서, 아이겐값이 1 이상인 요인들을 묶어 연구모형을 적용하도록 하였다(Gefen and Straub, 2005). 그 결과 GP와 PA에 관한 2개 항목의 값이 미달되어, 나머지 18개 항목만을 대

4) 표본수집의 어려움은 제3자 SPA에서 요구되는 심사의 기밀유지원칙(ISO, 2004)에서 대부분 기인한다. 따라서 해당 심사에 참여한 심사원들의 현장심사 자료를 수집하는 것은 일반적으로 매우 어렵다.

상으로 연구모형을 적용분석해 보도록 하였다.

본 연구는 이전연구 분석(Chin, 1998a; 1998b; 이종무, 2009)에 근거하여, 측정모형의 분석을 위한 변수간 근사가중치 결정방법으로 인과모형 연구에서 주로 적용되는 경로가중치 결정방법과 부스트랩방식을 적용하였다. 측정모형의 타당성을 확인하기 위해서 각 구성개념별 적재항목의 요인값과 t-값을 바탕으로 변수와 측정항목간의 집중타당성과 판별타당성, 그리고 내적일관성 등의 검증을 우선 수행하였다.

항목의 요인집중의 적절성을 확인하기 위한 집중타당성은 확인적요인분석(CFA)의 결과 값인 항목별 요인적재량이 0.7이상이면서, 다른 측정항목들과의 교차요인적재량보다 클 경우를 기준에 타당한 것으로 확인할 수 있다. 또한 판별

타당성은 각 요인의 추출된 평균분산(AVE)의 제곱근이 다른 구성개념의 상관계수보다 큰 값을 갖는 경우 이를 확인할 수 있다고 본다 (Fornell, 1981; Wixom, 2005; Hulland, 1999). 또한 신뢰성분석은 복합신뢰도값이 0.7이상이며, AVE값이 최소 0.5이상인 경우 신뢰성이 있는 즉, 내적일관성이 있는 것으로 판단한다 (Fornell, 1981; Hulland, 1999; Chin, 1998b). 본 연구의 결과값은 <표 3>과 같이 정리해 볼 수 있다.

집중타당성 확인을 위한 CFA 값을 분석해 보면, 반영지표로서 구성개념 표준지식과 심사기술 관련된 항목 포함 총 10개 측정항목의 요인적재량은 기준값 0.7을 상회하였다. 그러나 나머지 8개 측정항목은 1차연구와는 달리 기준값 이하

<표 4> 신뢰성과 타당성 분석

측정항목<->구성개념*	요인 적재량	t-값	확인적요인분석				판별타당성분석			
			PERSON	IS	SKILL	SPA	PERSON	IS	SKILL	SPA
Group → PERSON	0.1945	1.1693	0.1945	0.2747	0.1840	-0.0757	1			
IT_Career → PERSON	-0.0420	0.2599	-0.0420	-0.2409	-0.0275	0.1547				
SPA_Exp → PERSON	0.9840	5.1379	0.9840	0.5371	0.7214	0.5859				
BP ← IS	0.9181	39.4552	0.4057	0.9181	0.7048	0.3605	0.4679	1		
Document ← IS	0.8904	33.3120	0.4930	0.8904	0.7643	0.4395				
Rating_Criteria ← IS	0.7592	10.8632	0.1785	0.7592	0.4450	0.1149				
Standard ← IS	0.7671	15.9625	0.5346	0.7671	0.6915	0.4249	0.4879	0.8885	1	
Bus_Context ← SKILL	0.7400	11.4069	0.4821	0.7692	0.7400	0.5279				
GP_Rating ← SKILL	0.8001	18.4135	0.3826	0.8001	0.7068	0.4989				
Measuring ← SKILL	0.6199	7.8797	0.3085	0.5200	0.6199	0.5794				
PA_Aggregation ← SKILL	0.2922	2.8590	0.0789	0.5615	0.2922	0.0856				
Rating_Boundary ← SKILL	0.3183	2.7812	-0.0063	0.5152	0.3183	0.2091				
Team_Comm ← SKILL	0.7469	10.3457	0.4228	0.7964	0.7469	0.5520				
Time_Doc ← SKILL	0.6353	7.1747	0.3506	0.7601	0.6353	0.4162				
Time_Interview ← SKILL	0.8137	19.6015	0.6533	0.5050	0.8137	0.7556	0.6478	0.4723	0.6934	1
Interview → SPA	0.2589	1.3722	0.2409	0.2920	0.3065	0.2589				
SPA_Conf → SPA	0.4432	2.5202	0.2759	-0.0529	0.2515	0.4432				
SPA_MG → SPA	0.7252	5.0993	0.3812	0.2696	0.5756	0.7252				

* ← 는 반영지표를, → 는 조형지표를 각각 의미함.

로 나타나, 측정항목으로서 타당하지 못한 것으로 나타났다. 그러나 이들 항목들의 지표상 특성을 고려해 최종 연구결과의 확인은 가능할 것으로 판단된다⁵⁾.

판별타당성의 경우에는 표준지식과 심사기술의 AVE의 값이 0.7086과 0.5663으로 나타났으며, 결과 제공근은 기타 계수들의 값 이상으로 나타나 적절하였다. 그리고 복합신뢰도값도 기준값을 상회하는 결과(0.9058과 0.9054)로 나타났고, AVE값도 최소기준값 이상으로 나타나 신뢰성 확인도 가능했다.

4.2.2 구조모형의 분석

본 절에서는 연구목적인 표준에서 제시한 심사원 능력과 심사 비효율성에 영향을 주는 요인들의 관계를 살펴보기 위해, 심사원 및 조직특성에 관한 독립변수와 심사원 능력과 심사효율성에 관한 종속변수간의 관계에 근거한 모형을 분석해 보도록 한다.

연구모형별 PLS 경로분석의 최종 결과값은 <표 5>와 같이 요약해 볼 수 있다. 일반적으로

구조모형의 적합성은 각종 적합도 지표와 복수의 종속변수들의 평균적 적합도, 그리고 경로상의 적합도 등을 통해서 확인한다. PLS 구조모형의 적합성 판단기준으로는 관련 Redundancy값이 모두 양수임과 Community값이 최소 0.5 이상임을 확인해 알 수 있다(Chin, 1998b; Tenenhaus, 2005). 그러나 아쉽게도 본 연구모형의 경우 종속변수로서 심사효율성의 구성항목에 대한 Redundancy값이 음수로 나타났으며, Community값 역시 표준지식을 제외하고는 기준에 적합하지 못한 것으로 나타났다.

이와 함께 각 종속변수들의 개별 적합도는 R2 값을 통하여 확인할 수 있는데, Cohen(1988)의 기준⁶⁾을 적용하면 심사기술과 심사효율성은 높은 수준을, 그리고 표준지식은 중간 수준의 효과성을 갖는 것으로 나타났다. 이를 상대적으로 해석해 보면, 심사원 및 조직특성에 따라 심사기술이나 심사효율성에 미치는 영향이 표준지식의 습득에 미치는 영향보다는 더 크다는 것을 알 수 있다.

그리고 경로모형 전체적인 적합도⁷⁾는 적절한

<표 5> 기본모형의 분석 결과값

구성개념	R-Square	Cronbach's Alpha	Community	Redundancy
PERSON	0.0000	-	0.3281	0.0000
IS	0.2189	0.8588	0.7086	0.1476
SKILL	0.7962	0.8832	0.5663	0.4474
SPA	0.7256	-	0.2124	-0.2564

- 5) 지표특성으로 인해 선행연구에서 제시한 타당성과 신뢰성 판단기준을 필수적으로 적용해야 하는 것은 아니라 해도(Chin, 1998a), 보다 완벽한 이론모형의 검증을 위해서 추가 수정된 연구가 필요할 것으로 판단된다. 본 연구에서는 이들 항목을 제외시킨 수정모형의 재분석은 지면관계상 제외하였다.
- 6) 효과성을 각각 높은(0.26이상), 중간(0.13~0.26), 그리고 낮은(0.02~0.13) 수준으로 구분해 판단하고 있다.
- 7) 경로모형 적합도값은 R2와 Community의 각 평균값의 곱의 제곱근으로 계산해 얻을 수 있다.

<표 6> 경로 분석의 결과

경로	경로계수	t-값	결과
PERSON → IS	0.5182	9.2907***	가설 H1 채택
PERSON → SPA	0.1133	1.5660	가설 H2 기각
PERSON → SKILL	0.4340	3.9992***	가설 H3 채택
IS → SKILL	0.5837	5.1879***	가설 H4 채택
IS → SPA	0.7061	3.9422***	가설 H5 채택
SKILL → SPA	1.4855	10.6461***	가설 H6 채택

주) * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

값인 0.5937로 나타났는데, 이는 선행연구 (Tenenhaus, 2005)에서 제시한 높은 검정력 기준을 초과하는 값이다.

종합적인 구조모형상의 경로분석 결과는 <표 5>와 같이 요약해 볼 수 있다. 비록 구성항목의 적절성과 구조모형의 적합성에 문제가 있지만, H2를 제외한 대부분의 가설들이 채택되었다. 즉, 심사원 개인과 조직특성은 표준지식과 심사기술에 영향을 주는 것으로 볼 수 있으며, 표준지식과 심사기술은 각각 심사기술과 심사효율성에도 영향을 주는 것으로 나타났다.

결론적으로, 반영지표의 특성을 고려해 측정항목의 타당성 확인을 배제한다는 가정 하에서, 심사원 특성에 따라 표준지식과 심사기술은 영향을 받을 수 있으며, 표준지식도 심사기술에 영향을 줄 수 있다고 볼 수 있다. 이는 표준에 근거한 국제심사원 교육과정에서 실제 반영되어, 표준이론 강의 후 필수 심사기술 실습의 요구사항의 타당성을 뒷받침하는 것으로 이해할 수 있으며, 보다 효과적인 심사수행을 위해서는 결국 적절히 심사원능력을 고려한 팀 구성이 필수적이라는 것을 알 수 있다. 또한 이 결과는 현행 국제심사원 교육과정에서도 실제 수행하듯이, 비록 심사원 개인의 SE경험이 높더라도 철저한 사전

표준지식 습득과 심사기술의 훈련수준이 높아야만 성공적인 심사가 가능함을 추측케 하며, 특히 표준지식의 이해가 높은 심사기술을 통해 심사효율성에 주로 영향을 끼친다고 볼 수 있다. 그러나 다음 절에서 논의하듯이 비록 조정지표라 하더라도 심사효율성에 미치는 영향은 측정항목상의 문제가 있는 것으로 나타났으므로, 보다 개선된 후속연구를 통해서만 확실한 설명이 가능할 것이다.

4.2.3 연구한계와 기타논의

앞서 기본모형의 결과 분석에서 언급했듯이 몇 가지 연구모형상의 문제점들이 나타났다.

첫째, 변수 측정을 위한 일부항목의 타당성과 신뢰성이 부족한 것으로 나타났다. 이는 일부 선행연구에서 지적하는 모형의 재구성과 후속연구의 필요성이 나타난 것으로 볼 수 있다.

둘째, 자료의 측정항목 작성과 관련해서는, 지식이나 기술의 문제 혹은 부족한 정도를 묻는 질문형식은 심사 진행 중의 직접적인 설문수집의 어려움과 신속한 반응을 얻기 위한 기법으로 적용한 것이었으나, 응답자들의 인식의 혼란을 초래할 수도 있을 것이다.

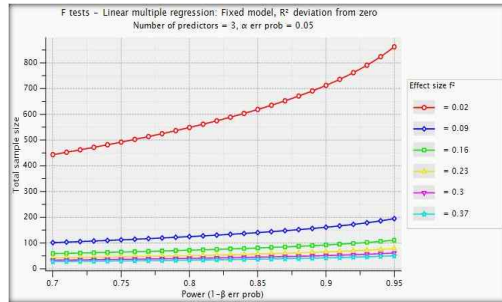
셋째, 본 실증연구에서 수집한 표본수의 적절

성과 검정력의 보장에 관한 문제점이 있을 수 있다. 최근 많은 경로모형 관련 실증연구에서 PLS 사용이 선호되는 주된 이유의 하나는 표본수와 정규분포에 관한 가정에 구애됨이 없이 적용하기 쉽다는 점이다.

본 연구에서도 이러한 장점 때문에 PLS를 적용하여 이론적 모형의 실증적 확인을 시도하였다. 그러나 최근 많은 연구에서는 최소 적정표본수와 관련하여 과거 단순한 경험적 직관이나 권고에 따라 표본수의 적절성을 확인하는 문제점과 결과도출에서의 오류가능성을 지적하고 있다. 본 연구의 경우 실제로 경험적 권고대로 최소 표본수를 참고하여 실증분석을 수행하였다. 참고로 관련적인 최소 표본수의 결정은 종속변수에 연결되는 독립변수의 최다 경로수와 개당 10개 표본이상을 사용하도록 권하고 있는데 (Chin, 1998a; Chin, 1998b), 본 연구에서도 이를 참고하여 선임심사원의 응답을 제외한 일반 심사원들의 표본 30개로 연구를 수행하였지만, 많은 검정력의 문제(Goodhue, 2006)를 일으킬 수 있음도 확인하였다.

따라서 손쉬운 검정력(power) 분석의 한 방법으로 기존연구를 참고하여, 각각 채택된 기본 유의수준, 검정력(1-β), 그리고 영향도와 채택된 표본수의 결정이 적절한가를 확인해 볼 필요가 있다. 보통 더 높은 유의수준을 채택하거나 혹은 더 많은 표본을 사용할수록 검정력을 높일 수 있으며, 그 결과의 신뢰성도 더욱 커짐을 알 수 있는데, 선행연구에 따르면 유의수준 0.05 이하와 검정력 0.8 이상을 일반적으로 권고하고 있다. 또한 이러한 유사 연구결과의 표본수를 비교하는 방법 외에 회귀분석의 검정력 분석도구를 사

용하는 방법을 적용해 볼 수 있는데, 본 절에서는 앞서 제시한 경험적 최소 표본수와 검정력 분석 도구를 이용한 시뮬레이션 값을 비교해 추가는 의해 보고자 한다⁸⁾.



<그림 4> 검정력 시뮬레이션 결과 (α=0.05; Max_No_of_Predictor=3)

<표 7> α=0.05 의 검정력 수준별 최소 표본수: 괄호안 α=0.001)

$1-\beta$ \ F	0.02	0.15	0.35
0.7	444 (940)	63 (133)	30 (62)
0.8	550 (1086)	77 (152)	36 (70)
0.9	713 (1305)	99 (182)	45 (83)
0.95	863 (1501)	119 (208)	54 (94)

도출된 결과는 유의수준 α=0.05를 기준으로 계산하여 이를 높은 유의수준(α=0.001)과 함께 비교하였는데, 그 결과 선행연구의 경험적 권고에 의한 최소 표본수와는 다르게, 검정력 분석결과에 의한 결과값은 조절효과 영향도 대, 중, 소를 기준(검정력 0.8 이상)으로 각각 최소 36, 77, 550 이상이 필요한 것으로 나타났다. 따라서 선행연구에서 권고하는 검정력수준을 얻기 위해서는, 본 연구에서 제시한 모형의 경로분석과 일부

8) 이를 위해 G*Power의 회귀분석을 기준으로 유의수준과 검정력에 따른 추정표본수를 비교한다.

조절효과의 입증을 위한 표본의 수는 약간 부족하다고 볼 수 있다.

종합적으로 평가하면, 검정력기준을 0.7 수준으로 낮춰 평가하거나 혹은 기존 선행연구의 경험적 권고를 참고한다면 본 연구결과의 신뢰성을 주장할 수는 있으나, 보다 높은 검정력과 분석결과의 신뢰성 확보를 위해서는 변수 측정항목의 추가보완과 충분한 수의 표본에 의한 추가연구가 필요하다고 판단된다.

V. 결 론

효율적인 프로세스 심사란 초기 심사 투입요소의 정의 단계에서부터 심사 제약조건 즉, 주요 심사자원의 가용성, 최대 심사 수행시간, 심사 프로세스의 특성, 심사 최대/최소 표본 수, 심사 참여자들의 특성 등이 두루 고려된 심사가 일정과 예산 내에서 원활히 수행됨을 의미한다.

본 연구에서는 표준에서 제시하는 심사원 특성과 능력과의 이론적 관계를 확인하기 위해 산업현장의 실제 심사사례를 통해 수집된 자료를 바탕으로, 과연 심사원과 조직특성이 표준지식과 심사기술의 습득에 영향을 줄 수 있으며, 또한 이들이 어떤 경로로 효율적인 심사진행에 영향을 줄 수 있는지를 경로분석방법을 통하여 확인해 보았다. 그 결과 심사원 SE경험에 따라서는 표준지식과 심사기술에 미치는 영향이 있으며, 표준지식도 심사기술에 선행하여 영향을 줄 수 있는 것으로 나타났다. 그리고 표준지식과 심사기술의 수준에 따라 전반적인 효율적인 심사에도 영향을 줄 수는 있을 것으로 기대했으나, 본 연구모형으로는 여러 문제점이 있는 것으로 나

타났다. 그리고 연구한계점에 근거하여, 추가적으로 통계적 검정력을 만족시키는 적정 표본수를 도출하여 비교해 보았다. 이는 향후연구 및 산업현장에서 효율적인 심사팀의 구성에서 직접 참고할 수 있을 것이며, 심사원 교육 및 훈련 프로그램의 이론적 타당성을 뒷받침하는 증거가 될 수 있을 것이다.

참고문헌

문태수, 서기철, “환경, 조직, 정보시스템 특성이 ERP시스템의 성공적 구축에 미치는 영향에 관한 실증연구,” 정보시스템연구, 제15권, 제1호, 2006, pp.73-96.

이상근, 강민철, 김보연, “혁신확산이론을 바탕으로 한 정보통신기술의 수용요인에 관한 분석적 실증연구,” 정보시스템연구, 제14권, 제2호, 2005, pp.257-276.

이종무, 박철수, 유영관, “IT도입요소를 고려한 외주 SW프로세스품질과 제품품질,” 정보시스템연구, 제18권, 제4호, 2009, pp.131-153.

정창신, 송정범, 이종무, “SPICE와 심사사례,” 한국정보과학회지, 제17권, 제1호, 1999, pp.13-22.

정학중, Audit Skill, KSPICE 2차 교육, ISO/IEC JTC1/ SC7/ WG10 한국위원회, 1998.

정호원, 황선명, “소프트웨어 프로세스 심사의 이해: SPICE를 중심으로,” 한국정보과학회지, 제17권, 제1호, 1999, pp.6-12.

정호원, 김길조, 김도관, ISO/IEC 15504 기반의

- 중소기업을 위한 소프트웨어 표준프로세스, 기술표준원, 2007.
- 한국 SPICE 위원회, A Phase 2 Trial Report: KSPICE, 소프트웨어 프로세스 심사원 워크샵, 1998.
- 한국 SPICE 위원회, KSPICE 심사원 교육자료, 2006.
- Al-Kilidar, H., Cox, K., and Kitchenham, B., "The Use and Usefulness of the ISO/IEC 9126 Quality Standard," *Proc. of ISESE 2005*, Noosa Heads, Australia, 2005, pp.126-132.
- Boehm, B. W., Brown, J. R., and Kaspar, J. R., *Characteristics of Software Quality*, TRW Series of Software Technology, Amsterdam, North Holland, 1978.
- Chin, W., "Issues and Opinion on Structural Equation Modeling," *MIS Quarterly*, Vol.22, No.1, 1998, pp.7-16.
- Chin, W., *The Partial Least Square Approach to Structural Equation Modeling*, Modern Methods for Business Research, ed. by G. Marcoulides, LEA, 1998, pp.295-336
- Chin, W., and Newsted, P., *Structural Equation Modeling Analysis with Small Samples Using Partial Least Squares*, Statistical Strategies for Small Sample Research, Sage Pub. Co., 1999.
- Chua, B. and L. Dyson, "Applying the ISO 9126 Model to the Evaluation of an e-Learning System," *ASCILITE Conference Proceeding*, 2004, pp.184-190.
- CMMI for Development, Ver. 1.2, Technical Report CMU/SEI-2006-TR-008, 2006.
- Cohen, J., *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Lawrence Erlbaum, Hillsdale, NJ, 1988.
- Diamantopoulos, A., and Winklhofer, H. M., "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research*, Vol.38, No.2, 2001, pp.269-277.
- Emam, K., Drouin, J., and Melo, W., *SPICE: The Theory and Practice of Software Process Improvement and Capability Determination*, IEEE Computer Society Press, 1998.
- Emam, K., and Birk, A., "Validating the ISO/IEC 15504 Measure of Software Requirement Analysis Process Capability," *IEEE Transaction on Software Engineering*, Vol.26, No.6, 2000, pp.541-566.
- Fornell, C. and Larcker, D. F., "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, Vol.18, No.1, 1981, pp.39-50.
- Gefen, D., and Straub, D., "The Relative Importance of Perceived Ease of Use in IS Adoption: A Study of e-Commerce Adoption," *Journal of Information Systems*, Vol.1, No.8, 2000.
- Gefen, D., and Straub, D., "A Practical Guide to

- Factorial Validity Using PLS-GRAPH,” *Comm. of the Association for Information Systems*, Vol.16, 2005, pp.91-109.
- Goodhue, D. and Thompson, R., “PLS, Small Sample Size, and Statistical Power in MIS Research,” *The 39th Hawaii International Conference on System Sciences*, 2006.
- Haase, V., “Software Process Assessment Concepts,” *Journal of Systems Architecture*, Vol.42, 1996, pp.621-631.
- Hulland, J., “Use of Partial Least Squares in Strategic Management Research: A Review of Four Recent Studies,” *Strategic Management Journal*, Vol.20, 1999, pp.195-204.
- ISO/IEC 12207, *Information Technology - Software Life Cycle Processes*, ISO, 1995.
- ISO/IEC IS-15504, *Information Technology - Software Process Assessment*, (All parts) ISO, Geneva, Switzerland, 2004.
- Johnson, R. and Wichern, D., *Applied Multivariate Statistical Analysis*, 3rd Ed., Prentice Hall, New Jersey, 1992.
- Jorgensen, M. and D. Sjoberg, “Generalization and Theory Building in Software Engineering Research,” *Empirical Assessment in Software Engineering Proceeding*, 2004, pp.29-36.
- Jung, H., R. Hunter, D. Goldenson, and K. Emam, “Findings from Phase 2 of the SPICE Trials,” *Software Process Improvement and Practice*, Vol.6, No.5, 2001, pp.205-242.
- Jung, H. and R. Hunter, “Evaluating the SPICE Rating Scale with Regard to the Internal Consistency of Capability Measures,” *Software Process Improvement and Practice*, Vol.8, No.3, 2003, pp.169-178.
- Jung, H., S. Kim, and C. Chung, “Measuring Software Product Quality: A Survey of ISO/IEC 9126,” *IEEE Software*, Vol.21, No.5, 2004, pp.88-92.
- Jung, H., “Validating the External Quality Subcharacteristics of Software Products according to ISO/IEC 9126,” *Computer Standards & Interfaces*, Vol.29, 2007, pp.653-661.
- Mackie, C., “Process Excellence and Capability Determination,” *BT Technology Journal*, Vol.15, No.3, 1997, pp.130-139.
- Pfleerger, S. and N. Fenton, S. Page, “Evaluating Software Engineering Standards,” *Computer*, Vol.27, No.9, 1994, pp.71-79.
- Rout, T., K. Emam, M. Fusani, D. Goldenson, and H. Jung, “SPICE in Retrospect: Developing a Standard for Process Assessment,” *Journal of Systems and Software*, Vol.80, No.9, 2007, pp.1483-1493.
- SPICE, Phase 2 Trial Interim Report, Version 1.0, ISO/IEC JTC1/SC7/WG10, 1998.
- Tenenhaus, M., V. Vinzi, Y. Chatelin, and C.

Lauro, "PLS Path Modeling," *Computational Statistics and Data Analysis*, Vol.48, No.1, 2005, pp.159-205.

Valenti, S., A. Cucchiarelli, and M. Panti, "Computer Based Assessment Systems Evaluation via the ISO9126 Quality Model," *Journal of Information Technology Education*, Vol.1 No.3, 2002.

Wixom, B., and Todd, P., "A Theoretical Integration of User Satisfaction and Technology Acceptance," *Information Systems Research*, Vol.16, No.1, 2005, pp.85-102.

관심사는 인공지능 응용, 지능형 재무예측시스템, 데이터마이닝, Business Intelligence, 지식자산가치평가 등이며, 현재 한라대학교 경영학과 교수로 재직하고 있다.

김만술(Kim, Mahn Sool)



저자는 서울대학교 대학원에서 경영학 석사학위와 알라바마주립대학에서 재무학 박사 학위를 취득하였다. 주요 관심사는 IT가치평가, 금융공학, 그리고 재무정보시스템이며, 한라대학교 교무처장과 기획처장을 역임하였고, 한라대학교 경영학과 교수로 재직하고 있다.

이종무(Lee, Jong Moo)



저자는 고려대학교 대학원에서 경영학박사 학위를 취득했으며, 미국 CompGraph사와 ASI사에서 SA/프로그래머로 근무하였고, 현재 KASPA 운영위원 및 심사위원으로 활동 중이며, 한라대학교 경영학과 교수로 재직하고 있다.

수로 재직하고 있다.

박철수(Park, Cheol Soo)



저자는 KAIST에서 회계/재무정보시스템을 전공하여 경영공학 석사와 박사를 취득하였다. KAIST 테크노경영연구소 및 동양경제연구소에서 연구원으로 재직했으며, 산업자원부(현 지식경제부) 전자거래진흥

원 자문위원을 역임하였다. Expert System with Application 등 국제학술지에 논문을 포함하여 다수의 논문을 주요 학술지 및 학술대회에 발표하였다. 주요

<Abstract>

A Study on Assessor's Competence and Efficiency of Software Process Assessment

Jong Moo Lee · Cheol Soo Park · Mahn Sul Kim

In this study, we verify the relationship, which is presented by ISO/IEC 15504, between the assessor's competence and the efficiency of assessment. The ISO standard has been developed by many formal trial experiences for the software process assessment. The software process assessment determines the process capability, and makes the software process improvement possible in the development cycle. In the field assessment, third-party and team-based assessments are generally recommended. Therefore assessors' competences are essential to achieve the effectiveness and the efficiency of assessment. Survey data are collected from the third-party process assessment of a real IT industry, and the analysis is done by a causal modelling technique - PLS. Some findings are as follows. Depending on assessors' SE experiences, effects of the assessor's competence are existed, and their standard knowledge and assessment skill are effected each other. And the efficiency of assessment depends on the assessor's competence. In addition, the power analysis is presented and future research is suggested, because of many limitations of this study.

Keyword : SPA, Assessor Competence, Assessment Efficiency, ISO/IEC 15504, PLS

* 이 논문은 2010년 3월 29일에 접수되어 1차수정(2010년 4월 26일)과 2차수정(2011년 3월 24일)을 거쳐 2011년 6월 9일 게재 확정되었습니다.