

컴퓨터 프로그래머 적성검사의 개발 및 타당화 연구*

김명소

호서대학교 경상학부 산업심리학과

일반적으로 심리검사가 본격적으로 활용되기 위해서는 잘 계획된 예비 연구들을 통하여 검사 문항들을 개발한 후, 신뢰도·타당도 검증 및 규준작성을 위해 다수의 해당 집단을 대상으로 조사하고 분석하는 다단계 연구가 필요하다. 본 연구에서는 정보화 시대에 요구되는 컴퓨터 관련 인력의 효율적인 인재선발을 위해서 컴퓨터 프로그래머 적성검사를 개발하고 예비검사와 타당화 작업을 통한 검사의 수정·보완 과정이 단계적으로 이루어졌다. 예비검사이 문헌 조사 및 기존의 검사를 검토한 후, 체계적 직무분석을 통해 도출된 선발 차원들을 측정하는 250개의 예비문항을 제작했다. 예비검사 결과 문항분석을 통해 각 문항의 난이도, 변별도, 내적 일치도를 고려한 후 총 100문항이 선정되었고, 다양한 집단의 컴퓨터 프로그래머들을 대상으로 위 검사에 대한 4차례의 타당화 조사가 실시되었다. 그 결과 대체적으로 양호한 내적일치도 신뢰도 지수를 얻었을 뿐 아니라 기대할 수 있는 방향으로 학력별, 계열별, 직군별로 전체검사 및 하위검사에서 집단간 유의한 차이를 보이고 있어, 본 검사는 이 분야의 학문적 적성검사로써 의미 있는 구성개념 타당도를 나타내고 있었다. 또한 본 검사는 준거변인의 제한점에도 불구하고 준거변인(상사지도관찰표)과 .36의 의미 있는 상관이 있어 높은 예측력을 보이고 있다. 이밖에도 요인구조를 분석한 결과, 자료처리/전환능력 및 분석/평가능력 요인의 두 요인이 도출되어 앞으로 이 분야의 검사개발 방향을 제시해 주고 있다. 위와 같은 여러 차례의 타당화 연구가 진행되면서, 본 컴퓨터 프로그래머 적성검사는 그 문제점과 제한점이 다각도로 검토되어 하위검사 및 개별 문항이 수정되고, 또 하위검사 시간도 조정되어 처음 제작시보다 더욱 신뢰롭고 타당한 검사로 발전되었다. 마지막으로 본 연구의 의의 및 제한점, 그리고 앞으로의 연구방향을 논의해 보았다.

일반적으로 적성검사를 사용하여 인재선발을 하는 상황에서 가장 기본적으로 중요한 문제는 어떤 종류의 검사를 예측변인으로 사용할 것인가 하는 것이다. 즉, 다양한 직종에 걸쳐 널리 적용할 수 있는 기존의 일반적성검사를 사용할 것인지 아니면 관심

의 대상이 되는 특정 직무수행에 필요한 특수한 지식, 기술, 능력들을 측정하는 특수적성검사를 사용할지를 결정해야한다. 물론 이 결정은 어떤 논리적, 이론적 혹은 경험적 근거를 바탕으로 이루어져야 하는데(Society for Industrial and Organizational Psychology, 1987), 일반적성검사 결과로부터 특수한

*본 연구는 학술진흥재단에서 부분적으로 지원 받은 연구임

전문지(예: 컴퓨터 프로그래머)의 성공여부를 제대로

예측할 수 있다고 보기는 어려울 때가 많다. 특정 직무의 성공적 수행에 요구되는 적성은 모든 사람이다 소유하기보다는 일정수의 사람들만이 소유하고 있다고 가정되어 왔기 때문이며, 최소한 일반인에게는 검사 사용에 대한 설득력이 약해서 안면타당도(face validity)가 낮다고 할 수 있다. 또 응시자 편에서도 지원하는 직종과 관련 있어 보이는(즉, 안면 타당도가 높은) 문항들로 구성된 특수적성검사가 회사 편에서 인적자원에 많은 노력과 관심을 보인다고 생각하여 좀 더 우호적으로 반응하고 최선을 다해 검사를 치르겠다는 동기가 커질 수 있다(Schmidt, Greenthal, Hunter, Berner & Seaton, 1977; Schmitt & Klimoski, 1991; Smither & Peariman, 1991). 따라서 특정 직무수행에 적합한 특성을 소유한 사람들을 선발하기 위해 지능검사와 유사한 일반적성검사보다는 특수한 적성검사가 필요한 것이다. 그런데 지금까지의 적성검사는 지능검사처럼 일반적 형태로 개발되는 경우가 많았고 특정 직무를 위한 특수한 적성검사는 몇몇 분야(예: 예술직, 비서직) 외에는 체계적으로 시도되지 못한 편이다.

최근 국내외 정보산업의 확대와 함께 각 기업들의 컴퓨터 관련 전문인에 대한 수요가 급증하고 있다. 우리 나라가 소프트웨어 선진국으로 성장하기 위해서는 컴퓨터 분야의 전문인력이 뒷받침되어야 하며, 필요한 인력을 제대로 선발해야 하는 것이 자명한 사실이다. 최근 사회적으로 열린 교육시대를 맞이함에 따라 대기업을 중심으로 학력(일류대), 경험, 연령의 의미가 축소된 '열린 인사'가 실시되면서, 적성검사가 신입사원 채용의 최대 변수로 등장했다. 적성검사가 전공시험이나 대학 학점을 대체하는 신입사원 채용의 변별 수단으로 자리잡아가고 있기 때문이다(매일경제 신문, 1995년 9월 25일). 컴퓨터 분야에도 이 분야의 필요한 인력을 제대로 확보하기 위해

서는 이 분야에서 특수한 끼, 즉 적성이나 잠재력이 있는 인재를 파악하여 선별해 줄 수 있는 선발검사의 개발이 무엇보다도 중요하다고 본다. 특히, 컴퓨터 프로그래머의 업무는 시간이 흐름에 따라 빠르게 계속해서 변화되고 있는데, 구조화된 프로그래밍, 새로운 세대의 프로그램 언어, 시스템 통합, 멀티미디어 등의 출현으로 인하여 업무의 특성이 점차 세분화, 다양화되어 가기 때문이다. 따라서, 컴퓨터 분야의 필요한 인력을 효과적으로 선발하기 위해서는 시간과 경비가 요구되는 어려운 작업이라 할지라도 컴퓨터 프로그래머라는 업무에 요구되는 특수한 적성에 초점을 두어야 하며, 이와 무관한 특성들은 선발 결정에 유리하게도 불리하게도 영향을 끼쳐서는 안 될 것이다(Anastasi, 1990).

그러나 우리 기업에서는 주로 인력을 그룹차원에서 전체적으로 선발한 후 연수 시나 원서상의 지망순서에 따라 부서로 배치하는 경향이 있으므로, 특수 전문인을 선발한다고 하기보다는 제너럴리스트를 선발·양성해 왔다고 볼 수 있다. 따라서 컴퓨터 직종의 독특한 직무특성을 반영하는 효과적인 인재선발이 이루어지지 못하고 있다. 몇몇 대기업들이 컴퓨터 관련 인재선발시 전산 적성시험 또는 프로그래머 적성시험이라는 것을 실시하고 있지만, 이들은 대부분 컴퓨터 프로그래머의 업무 특성에 관한 체계적 연구나 타당도 혹은 활용성에 관한 분석 없이 외국의 적성검사(주로 IBM이나 Univac 검사)를 단순히 번역하거나 원문 그대로 사용하고 있는 실정이다. 이러한 검사는 심리측정의 특성을 지닌 검사로 평가하기 어렵고, 또 이대로 사용한다면 검사를 오용하는 경우가 된다(김명소, 1994). 더욱이 기존의 컴퓨터 프로그래머 적성검사들은 국내외적으로 매우 적은 수에 불과한데, 컴퓨터와 관련된 적성검사는 현대 사회에서 컴퓨터의 활용범위가 넓어져 감에 따라 비교적

최근에 들어 관심을 갖게 된 분야로서 양적, 질적 연구가 상당히 미흡한 상황이다. 본 연구자가 조사한 바에 의하면 국내외적으로 지금까지 컴퓨터 프로그래머 직무와 관련되어 발표된 체계적 직무분석 연구는 단 1편에 불과하고(Page & Caskey, 1988), 국내에서 자체 개발된 컴퓨터 프로그래머 관련 적성검사는 전무한 실정이며, 한국 Saville & Holdsworth Ltd.(SHL)이 영국의 SHL에서 개발한 전산직 적성검사(Computer Aptitude Battery)를 번역하여 출판하고 있다. 외국에서 개발된 컴퓨터 프로그래머 적성검사로도 Psychological Corporation에서 제작한 Computer Aptitude, Literacy, and Interest Profile이 더 이상 판매, 사용되지 않게 되어 Aptitude Test For Programmer Personnel(IBM), Univac Programmer Aptitude Battery(Univac), Computer Programmer Aptitude Battery(Science Research Associates(SRA)), Computer Aptitude Battery(SHL) 등 소수에 불과하다.

이 중 IBM이나 Univac 컴퓨터 프로그래머검사는 1960년도에 자회사 프로그래머들에 의해 제작되어 사용되고 있어 오늘날의 심리검사 기준에 비추어 볼 때, 신뢰도나 타당도 측면에서 문제점이 있으며 측정하고 있는 영역이 일반 지능검사와 너무나 비슷하다. 또 SRA나 SHL 컴퓨터 프로그래머 적성검사는 심리검사 연구 기관에서 개발된 것이므로, 전자의 검사들에 비해 신뢰도나 타당도에서 좀 더 안정된 도구로 생각 될 수 있으나, 각각 60년과 70년대에 제작된 검사로서 최근 다양화, 전문화되어 가는 추세인 컴퓨터 프로그래머직에 적합한 인재를 선발하는 도구로서 적합하기 위해서는 지속적인 타당화, 표준화 연구가 필요한 것으로 사려된다. 이러한 현황을 미루어 볼 때, 정보화 시대를 맞이하고 있는 이 시점에서 컴퓨터 관련 분야의 효과적인 인재선발을 위해서 우

리 나라에서도 자체적으로 신뢰도와 타당도가 검증된 측정도구의 개발과 이에 대한 표준화 작업이 필요하다.

본 연구자의 1993~94년도 학술진흥재단 자유공모연구에서는 현업관련도가 높은 검사의 개발을 목적으로 먼저 컴퓨터 프로그래머란 직무의 본질을 이해하기 위한 체계적인 직무분석을 실시하였고, 그 결과, 타업무와는 구별되는 컴퓨터 프로그래머 직종의 요구특성(지식, 기술, 능력, 성격 특성)을 파악하였다(김명소, 강혜련, 1994). 이는 선발 및 배치용 검사가 내용 타당도를 갖추기 위하여 필수적인 절차이며 검사에 대한 과학적 이해를 증진시키고 또한 새로운 동형검사의 개발을 용이하게 해 준다. 따라서 본 연구에서는 위 직무분석 결과로 도출된 특성들 중 10개의 기술 및 능력 선발차원을 측정하기 위한 컴퓨터 프로그래머 적성검사를 개발하려고 했다. 앞으로 이 검사를 컴퓨터 분야의 신입사원 선발, 부서 배치, 승진 등과 같은 인사관리에 폭넓게 활용하기 위해서는 장기적인 추후연구들이 필요한데, 특히 구체적이고 경험적인 타당성 검증이 우선되어야 하겠다. 검사의 타당도란 그 검사를 평가하는데 있어 가장 중요하게 고려되어야 할 사항이며, 이는 검사가 측정하고자 하는 특성을 제대로 측정하고 있는가와 검사점수에서 도출된 의미나 추론이 적절한지 그리고 유용한지를 나타내 주는 정도를 말한다. 따라서 특정한 검사의 평가는 이러한 추론에 근거가 될 수 있는 증거를 모으는 과정이라고 할 수 있으며(APA, 1985), 완벽하게 타당하거나 혹은 전혀 타당하지 못한 검사란 있을 수 없을 것이고, 단지 연속선상에서 높은 타당도를 향하여 증거를 모아 가는 과정이 타당화 연구라고 볼 수 있다. 이러한 타당화 연구 없이는 검사의 적절한 활용이 불가능해지므로, 본 연구는 ① 컴퓨터 프로그래머 적성검사①의 개발절차를 간단

히 기술하고, ② 본 컴퓨터 프로그래머 적성검사의 본격적인 표준화 및 사용에 앞서 검사의 신뢰도 및 타당도를 예비검증하고 제반 문제점을 사전 검토하는 것이 주된 목적이다. 특히 타당도의 세 가지 주요 유형인 내용 타당도, 준거관련 타당도, 구성개념 타당도를 여러 가지 방법으로 살펴보기 위해 4단계에 걸친 타당화 조사가 실시되었다.

제 1차 연구(예비 조사)

1차 연구는 직무분석을 바탕으로 검사문항들을 개발하고, 그 특성들을 분석한 뒤 난이도와 변별도가 적당한 문항들을 선별하여 1차적으로 검사를 구성하는데 그 목적이 있다.

연구 방법

연구대상

본 검사의 예비조사는 국내 3개 대학의 컴퓨터 관련 학과 대학생 150명을 대상으로 실시되었다. 이들의 인구 통계적 특성은 다음과 같다(표 1).

표 1. 조사 대상자 특성

소 속(명)	성 별(명)
서울소재 상위권대(48)	남 (139)
서울소재 중위권대(51)	여 (11)
지방소재 사립대(51)	
전 체 (150)	

1) 본래는 직무분석 연구에서 기술, 능력 뿐만 아니라 지식, 성격 특성들도 파악되었고, 또 이를 바탕으로 검사가 능력검사와 성격검사로 구성되어 있었는데, 본 연구에서는 성격검사에 관한 논의가 제외되었고 일차적으로 선발용 검사인 능력검사 부분에 대한 타당도 검증만 실시되었다.

도구

직무분석으로부터 얻은 요인들에 대한 예비문항들을 개발하기 위한 본 검사 개발의 기본 방향은 다음과 같았다.

- 컴퓨터 프로그래머로서의 기본 능력, 적성, 또는 잠재력에 대한 파악이 가능해야 함. 즉, 직무분석에서 도출된 능력 요인의 측정이 가능해야 함.
- 다수 응시자를 대상으로한 집단검사 형태여야 함. 즉, 문항형태는 다지선다(multiple choice)임.
- 응시자들이 지루하거나 피곤하지 않고 최선의 능력을 발휘할 수 있도록 검사 시간은 가급적 1시간 내지 1시간 반을 넘지 않아야 함.

본 검사는 직무분석 결과로부터 나온 요인들을 기초로 문헌조사 및 국내외의 검사도구를 검토하여 제작된 것으로 1차적으로 5개 척도(각 척도별로 50문항씩으로 구성된 250개의 예비문항들)로 구성되었다. 직무분석 조사결과(김명소, 강해련, 1994)와 검사구성의 관계는 아래의 표 2와 같다.

본 적성검사의 내용 타당도를 확인하기 위해서 컴퓨터 관련 전문가들(컴퓨터 관련 학과 교수, 기업체 컴퓨터 관련 중간 관리자 10명)에게 본 검사의 내용이 직무분석에서 도출된 컴퓨터 프로그래머의 특수한 영역들의 중요하고 대표적인 분야들을 포함하고 있는지를 검토해 달라고 요청하였다. 본 검사가 내용 면에서 타당성을 갖추기 위해 필수적 절차인 직무분석이 이미 실시된 바 있고, 이 때 20여명의 컴퓨터 관련 대기업, 중소기업, 대학부설 전산원의 중간관리자들 및 컴퓨터 프로그래머들과 인터뷰를 하였다. 여기서 도출된 컴퓨터 프로그래머 직무수행에 요구되는 능력에 관한 인적 특성들을 바탕으로 검사가 제작되었으므로, 이미 전문가의 전문적 판단이 어느 정도는 검사제작과정에 기여했다고 볼 수 있다.

표 2. 하위검사와 직무분석 요인과의 관계

하위검사	직무분석 요인
도식능력(Diagramming Ability)	자료 표현 능력 논리적 사고 및 통찰력
코딩능력(Coded Instructions Ability)	테크니컬한 기술
수리력(Numerical Ability)	수리적(계산, 판단 및 정보처리)능력
추리력(Analytical Ability)	분석, 판단, 평가 능력 논리적 사고 및 통찰력
기획력(Planning Ability)	기획력 및 창의력 분석, 판단, 평가 능력

전문가들의 판단결과 각 하위검사가 대체적으로 양호하다는 판단이 나왔으며, 도식능력에서 두 문제가 전문적인 지식(예: 알고리즘)을 요구하고 있어서 다른 문항들로 대체하였다. 즉, 본 검사는 성취검사가 아니므로 컴퓨터 전공자만 풀 수 있는 문항은 제외시키도록 하였다. 본 검사의 하위검사들의 내용은 다음과 같다(각 하위검사의 내용을 살펴보면 특정 하위 검사들과 특정 직무분석 요인들과의 연결성이 좀 더 명확해진다).

도식(다이아그램)능력 검사. 이 검사는 지시문이 주어지는 상황에서 논리적인 규칙을 이해하고 적용할 수 있는지를 평가하고 있다. 응시자는 플로우 차트의 형태로 배열된 지시 또는 비교명령에 따라 논리적이고 조직적인 사고력과 규칙을 가지고 지시된 문제를 부호로 변형시켜 나가는 능력을 평가받게 된다. 이는 프로그램 설계시 필수적인 능력이다.

코딩능력 검사. 이 검사는 부호화된 명령을 이해하고 수행할 수 있는지를 평가하기 위한 검사이다. 응시자는 지시문을 이해하고 주어진 문장을 정확하게 부호로 바꾸도록 요구되어 테크니컬한 프로그램 작성능력을 평가받고 있다. 이 검사는 언어적인 이해 능력과 자료전환 능력을 동시에 측정하고 있다.

수리력 검사. 이 검사는 수처리 능력, 기초 통계자료 해석 및 수리적 부호화 능력을 측정하고 있다. 응시자는 수리적 판단을 신속하게 하여야 하며 또한 주어진 상황을 이해하고 그것을 수학의 부호(기호)로 변형시킬 수 있는지를 평가받게 된다. 이 검사의 내용은 현실적이어서 프로그래머들이 실제로 다루는 문제들을 포함하고 있다. 즉, 실제 계산 없이 문제 해결에 가장 좋은 방법을 나타낸 식을 찾는 능력을 평가받게 된다. 검사 내용은 아래와 같은 두 가지 하위 척도로 구성되어 있다.

- 자료해석력 - 표나 그래프 혹은 척도로 제시된 정보의 수리적 계산 능력은 물론 수리적 판단능력, 시스템 분석작업 등 실제 업무현장에서 경험하게 되는 통계적 자료를 다루는 능력.

- 수리적 부호화 능력 - 수학적인 문맥을 이해하고 명확히 조직적으로 사고하며 기호화할 수 있는 능력. 이 검사는 실제 계산 없이 문제해결에 가장 좋은 방법을 나타낸 식을 찾는 능력.

추리력 검사. 이 검사는 지시된 자료에서 구성원리를 추리하고 해석하는 능력, 즉, 논리적 추론능력을 측정하고 있다. 응시자의 신속한 시각적 판단력과 분석력이 요구되는 검사이다. 검사내용은 제시된 일련의 수나 문자들의 규칙을 분석하여 다음에 올

것을 예측하는 것과 제시된 도형간의 관계를 찾거나 변화하는 그림들의 변화법칙을 찾아내는 것으로 구성되어 있다.

기획력 검사. 이 검사는 주어진 계획이나 전략을 정확히 판단, 분석, 평가하고 문제점을 수정할 수 있는지를 평가한다. 컴퓨터 프로그래머의 영역이 시간이 흐름에 따라 확장되고 업무의 특성이 세분화되어 가기 때문에 단순한 코딩 작업뿐 아니라 기획 및 관리 업무까지 요구되므로, 본 하위검사가 포함되었다.

조사절차

1994년 6월에 3개 대학교 컴퓨터 관련 학과(주로 컴퓨터 공학과) 대학생 집단을 대상으로 250개의 예비문항들을 실시하였다. 각 대학 학과장의 허락을 받고 각 학교별로 찾아가서 연구자가 조교의 도움을 받아 직접 검사를 실시하였다.

분석방법

문항분석을 위해서 PC용 ITEMAN(version 3.5)을 사용하였다. 검사의 신뢰도를 구해보기 위해 내적 일치도 계수(Cronbach's α)를 산출했고, 하위검사의 관계성을 살펴보기 위해 검사간 Pearson 상관계수를 구하였다.

연구 결과

문항 통계(난이도 및 변별도)

원칙적으로 문항 난이도(p)가 .20보다 작거나 .90보다 큰 문항은 삭제하거나 보기를 수정 보완하였고 문항별 변별도도 너무 낮은 문항($\alpha = .05$ 에서 유

의하지 않은 것)은 삭제하였다. 각 척도별로 난이도가 적당하고 변별도가 우수한 문항을 20개씩 선택하여 수정된 검사를 구성하였다. 전체적으로 볼 때 유지된 문항들의 문항 난이도의 범위는 .200~.933이었고 문항 변별도 계수는 .238부터 .894까지로 나타났다. 각 검사의 평균 문항 난이도와 변별도는 표 3에 나타나 있다.

표 3. 평균 문항 난이도 및 변별도

하위검사	평균난이도(p)	평균변별도 (양분점 상관계수)
도식능력	.53	.47
코딩능력	.51	.55
수리력	.46	.45
추리력	.44	.43
기획력	.45	.45
전체검사	.50	.47

전체적으로 .50의 난이도를 보이고 있어 양호한 난이도 수준으로 해석된다. 하위검사별로는 코딩능력 검사가 비교적 쉬웠던 반면에 수리력과 추리력, 기획력 검사는 약간 어렵다는 것을 알 수 있다.

본 검사의 변별도를 살펴보면, 모든 하위검사가 좋은 변별도의 일반적 기준인 .25를 넘는 변별도를 보이고 있어 본 검사의 변별력이 높다고 할 수 있다.

신뢰도

전체검사 및 각 하위검사별 내적 일치도 신뢰도 계수(α)는 다음의 표 4와 같다.

표 4. 내적 일치도 신뢰도 계수(α)

전체검사	도식능력	코딩능력	수리력	추리력	기획력
.851	.785	.830	.716	.688	.605

하위검사별 내적 일치도 신뢰도 계수는 일반적

으로 양호했으나, 코딩능력이 .830으로 가장 높았고, 기획력이 .605로 가장 낮았다. 전체적으로는 .851로 높은 신뢰도를 보이고 있다.

하위검사간 상관관계

하위검사간 상관계수는 .142~.464 사이로 낮거나 중간 크기의 상관을 나타내고 있고, 따라서 각 척도가 서로 많이 중복되지도 않으면서 동시에 같은 능력을 측정하고 있다(표 5). 수리력과 추리력, 도식 능력과 추리력, 코딩능력과 기획력은 비교적 높은 상관을 보이고 도식능력과 기획력, 추리력과 기획력은 비교적 낮은 상관 관계를 나타내고 있다.

표 5. 하위검사간 상관관계

	도식 능력	코딩 능력	수리력	추리력	기획력
도식능력	1.000				
코딩능력	.275	1.000			
수리력	.350	.292	1.000		
추리력	.427	.187	.464	1.000	
기획력	.142	.441	.203	.173	1.000

제 2차 연구

2차 연구의 목적은 예비연구 결과로부터 얻어진 100문항의 검사를 실제로 컴퓨터 프로그래머들을 대상으로 실시하여 문항들을 검토하고 신뢰도·타당도를 살펴 보는데 있다. 특히 집단간 평균점수차이를 분석하여 본 검사의 학문적 적성검사로서의 구성개념 타당도를 검토하였다.

연구 방법

연구대상

본 연구의 대상은 국내 2개 대학의 컴퓨터 관련학과 대학원생 58명과 2개 전자회사의 사원 및 대학부설 전산원 직원 58명으로 구성된 116명의 컴퓨터 프로그래머들이었다. 이들의 인구 통계학적 특성은 표 6에 제시되어 있다.

표 6. 조사 대상자 특성

소 속(명)	성 별(명)
석사과정 (43)	남(86)
박사과정 (15)	여(30)
회사·전산원(58)	
전 체 (116)	

도구

1차(예비) 연구에서 추려진 100문항(5개 하위 검사)으로 구성된 검사를 사용하였다.

조사절차

1995년 1월에 대학원생들과 대학부속 전산실 직원들을 각 학교별로 찾아가 연구자가 직접 검사를 실시하였고, 모 그룹의 전자 계열회사에 속한 프로그래머들은 담당 부서 과장에게 협조를 의뢰하여 부서별로 담당과장이 검사를 실시하였다.

분석방법

문항분석을 위해서 PC용 ITEMAN(version 3.5)이 사용되었고, 나머지 자료는 SAS 패키지를 이용하여 통계 처리하였다. 검사의 신뢰도로서 내적 일치도(Cronbach's α)를 산출하였으며, 하위 검사의 영역별 독립성을 살펴보기 위해 하위검사간 Pearson 상관계수를 구하였다. 아울러 검사의 구성개념 타당도 증거를 수집하기 위해 F검증을 통한 집단간 평균 차이를 검토하였다. 집단에 대한 구분은 자신들이 체크한 직명(job title)에 따라 코딩을 주업무로 하는 단순

직 프로그래머 집단과 시스템 디자인을 주로 하거나 코딩과 디자인을 함께 하는 시스템 디자이너(어널리스트) 집단으로 구분하였다.

연구 결과

문항 통계 및 신뢰도

위 컴퓨터 프로그래머 집단에서 문항 난이도와 변별도, 신뢰도를 검토해 본 결과 표 7과 같이 나타났다.

표 7. 평균 문항 난이도, 변별도 및 신뢰도계수

하위검사	평균난이도(p)	평균변별도 (양분점 상관계수)	α 계수
도식능력	.68	.46	.78
코딩능력	.84	.52	.81
수리력	.70	.42	.67
추리력	.63	.42	.70
기획력	.54	.37	.63
전체검사	.67	.44	.85

전체적으로 볼 때 문항 난이도(p)가 상대적으로 높아서 쉬운 듯 보이지만, 이것은 본 연구의 조사대상자들 특성 때문일 것이다. 즉 조사대상자들은 이미 대기업 및 중소기업의 프로그래머로 선발된 사람들이거나 우수 대학원의 석·박사과정 학생들이나 대학원생인 것을 생각해 볼 때, 일반 프로그래머 응시자들을 대상으로 검사를 실시한다면 정상분포에 가까운 점수분포가 형성될 것으로 기대된다. 또한 변별도는 모든 하위검사들에서 양호한 것으로 보여진다. 특히, p값이 높은 상황에서 .44의 평균변별도는 의미가 크다고 할 수 있다.

각 하위검사별 내적일치도는 코딩능력 검사에서 제일 높았고, 상대적으로 기획력 검사에서 낮았다.

그 이유로는 코딩능력 검사는 가장 단일한 특성을 측정하고 있다고 볼 수 있는 반면에, 기획력 검사는 현재까지 타당성이나 활용성에 관한 연구가 부족하다는 점을 들 수 있겠다. 또한 수리력 검사는 두 가지 하위척도로 구성되어 이들간 일치성이 다소 낮은 듯 싶다.

하위검사간과 전체검사와의 상관관계

각 하위검사와 다른 하위검사들간의 상관관계와 또 전체검사와의 상관관계를 분석해 본 결과 다음과 같다(표 8).

표 8. 하위검사 및 전체검사와의 상관관계

	도식 능력	코딩 능력	수리력	추리력	기획력
도식능력	1.00				
코딩능력	.33	1.00			
수리력	.35	.35	1.00		
추리력	.25	.38	.43	1.00	
기획력	.25	.32	.36	.46	1.00
전체검사	.64	.62	.70	.72	.50

위 표에서 하위검사간 상관계수는 .25~.46의 상관을 보여서 영역별 구별성이 어느 정도 나타나고 있다. 추리력은 다른 하위검사들과 비교적 높은 상관을 보이고 또 검사 총점과도 가장 높은 상관관계를 나타내고 있다. 물론, 하위검사간에 중간정도의 상관관계가 있어서 총점과 각 검사의 고유한 관계는 계산되지 않았음을 감안해야 한다.

집단간 점수차이 분석

성별, 소속별, 직명(프로그래머/시스템 디자이너) 별, 대학원별 평균 점수를 비교해 본 결과 다음과 같다(표 9).

집단 비교에 있어서 남/여 별로는 큰 차이가 없었고 단지 도식능력에 있어서 남자가 우수하게 나타

표 9. 집단간 평균 점수 차이 분석

(괄호 안은 표준편차, 각 하위검사는 20문항으로 되어 있고 모두 1점의 배점을 받음)

	전체검사	도식능력	코딩능력	수리력	추리력	기획력
전체(N=116)	66.88(10.67)	13.59(3.39)	16.27(3.38)	13.56(3.13)	12.58(3.42)	10.69(2.86)
남(N=86)	66.98(10.69)	13.91(3.76)*	16.13(3.40)	13.66(3.00)	12.51(3.45)	10.78(2.88)
여(N=30)	64.81(10.72)	11.56(4.15)*	17.13(3.16)	12.94(3.89)	13.06(3.34)	10.13(2.73)
석사과정(N=43)	71.33(7.90)*	15.40(2.55)*	17.05(2.29)	14.05(2.81)*	13.65(3.16)*	11.19(2.81)*
박사과정(N=15)	76.33(9.82)*	14.67(3.54)*	17.67(1.32)	16.33(1.32)*	14.33(3.81)*	13.33(1.87)*
회사(N=56)	63.47(9.72)*	12.59(3.80)*	16.80(3.52)	13.11(3.31)*	11.37(3.38)*	9.66(2.79)*
프로그래머(N=88)	64.87(10.5)*	13.19(4.06)*	16.22(3.42)	13.02(3.10)*	12.17(3.27)*	10.30(2.85)*
디자이너(N=28)	72.52(9.63)*	14.78(3.06)*	16.41(3.37)	15.30(2.69)*	13.96(3.65)*	12.07(2.50)*
서울소재 중위권 대학(N=33)	71.43(8.10)	16.00(2.31)	17.54(1.35)	13.57(2.52)*	13.46(2.87)	10.86(3.08)*
서울소재 상위권 대학(N=26)	73.64(9.03)	14.60(3.03)	16.76(2.91)	15.60(2.74)*	14.32(3.73)	12.36(2.14)*

* 최소 .05수준 이하에서 유의한 차이나 상관을 나타냄; 이하 동일

났다. 유의미한 차이는 아니지만 코딩능력에서 여성
의 점수가 높은 점은 언어적으로 표현된 지식을 이
해하여 꼼꼼하고 정확하게 코딩하는 능력은 여자가
남자보다 잘 할 수 있기 때문인 것 같다.

두 번째로 석사/박사과정 대학원생들, 그리고 회
사나 전산원 직원들의 비교에서 전체검사 및 모든
하위검사에서 박사>석사>회사, 전산원 순으로 점수
가 높게 나타났다. 회사원이나 전산원 직원의 최종학
력이 모두 대학인 점을 고려해 볼 때 이 결과는 학
력의 차이로 설명될 수 있다.

세 번째로 프로그래머와 시스템 디자이너의 평
균 점수치를 살펴 본 결과 5개 검사 모두에서 시스
템 디자이너가 높은 점수를 보였고, 특히 코딩능력을
제외한 4개의 검사에서 유의한 차이($p<.05$)를 보였
다. 코딩능력은 프로그래머와 디자이너 모두에게 필
수적인 능력이고, 특히 단순 프로그래머의 업무 중
대부분을 차지하는 것이므로 두 집단에서 큰 차이가
없는 것이 당연하다고 볼 수 있다.

마지막으로 서울소재 상위권과 중위권 대학원생
들을 비교해 보았는데, 예상대로 상위권 대학원생들
이 중위권 대학원생들보다 높은 점수를 보이고 있다.
특히 수리력과 기획력에서는 유의한 차이를 나타내

고 있다. 이상에서 살펴본 바와 같이 집단간 점수 차
이가 기대할 수 있는 방향으로 나타나고 있어 본 검
사의 타당성을 시사하고 있다.

제 3차 연구

3차 연구는 2차 연구에서 수정된 검사를 또 다
른 컴퓨터 프로그래머집단에 실시하여 하위검사들의
특성을 살펴보고 성별, 직군별, 학력별 집단간 차이
검증을 실시하여 구성개념 타당도를 검토해 보는데
그 목적을 두고 있다.

연구 방법

연구 대상

3차 연구의 연구대상은 국내 모그룹의 컴퓨터
관련 계열회사 직원 223명이었다. 이들의 인구통계학
적 특성은 표 10과 같다.

표 10. 연구대상의 인구 통계적 특성

성별(명)	직군(명)	출신대학(명)
남(145)	소프트웨어직군(146)	고등학교, 전문대 (9)
여(78)	연구개발직군 (77)	지방대 (38)
		서울소재 대학 (176)
전 체 (223)		

도구

2차 연구와 동일(2차 연구결과 문항분석을 통해 난이도가 지나치게 높거나 낮아서 변별력에 문제가 있는 5개 문항이 수정되었음).

절차

1995년 6월 국내 모그룹의 전산계열회사에 속한 프로그래머 223명을 대상으로 제 2차 연구에서 5개 문항이 수정된 검사를 실시하였다. 회사 담당자의 협조를 얻어 연구자와 보조연구자의 감독 하에 오후 업무시간(1시간 30분간)을 사용하여 검사를 실시하였

다.

분석방법

수집된 자료는 PC용 ITEMAN(version 3.5)과 SAS패키지를 이용하여 통계 처리되었다. 즉, 본 검사의 타당화를 위해 ITEMAN을 이용하여 문항분석을 재실시하였고, SAS를 이용하여 집단간 평균을 비교하였다.

연구 결과

하위검사와 문항 통계 및 신뢰도

위 조사대상자 집단에서 하위검사 평균, 표준편차, 평균 난이도, 변별도 및 신뢰도를 측정해 본 결과 다음과 같다(표 11).

전체적으로 볼 때, 기획력을 제외하고는 평균과

표 11. 하위검사와 문항 통계 및 신뢰도 계수

하위검사	평균(SD)	α 계수	평균난이도	평균 변별도 (양분점 상관계수)
도식능력	15.03(2.77)	.68	.75	.35
코딩능력	13.75(3.92)	.84	.68	.46
수리력	11.68(2.36)	.51	.58	.31
추리력	13.17(3.10)	.68	.65	.35
기획력	9.37(2.53)	.43	.46	.28

표 12. 출신 학교별(학력) 점수 차이 분석

		도식능력	코딩능력	수리력	추리력	기획력	전체검사
고졸/전문대	평균 (SD)	11.1(5.9)*	13.0(4.8)	10.0(4.0)	13.1(3.3)	9.2(2.0)	56.4(12.1)
	N	9	9	9	9	9	9
지방사립대	평균 (SD)	14.5(4.1)*	14.4(3.6)	11.8(1.7)	13.4(3.1)	9.9(1.9)	64.0(9.5)
	N	22	22	22	22	22	22
지방국립대	평균 (SD)	15.0(2.3)*	15.0(3.6)	12.4(1.9)	12.8(3.5)	9.7(2.9)	64.9(8.6)
	N	16	16	16	16	16	16
서울기타대	평균 (SD)	14.2(2.9)*	13.2(4.5)	11.0(2.1)	13.0(3.2)	9.3(2.5)	60.4(9.8)
	N	90	90	90	90	90	90
서울일류대	평균 (SD)	15.6(2.7)*	14.0(8.9)	12.0(2.1)	13.3(2.7)	9.7(2.7)	64.6(9.3)
	N	89	89	89	89	89	89
전 체	평균 (SD)	14.8(3.4)*	13.8(3.9)	11.7(2.4)	13.2(3.1)	9.4(2.5)	62.8(9.6)
	N	226	226	226	226	226	226

연구 방법

문항난이도(p)값이 높아서 문항들이 쉬운 듯 보인다. 그러나 2차 연구에서와 마찬가지로 본 연구 대상자들은 이미 대기업에 선발된 프로그래머들이었음을 기억해야 할 것이다. 각 하위 검사별 내적 일치도는 1, 2차 조사에서도 나타난 바와 같이 코딩능력 검사에서 제일 높았고, 상대적으로 기획력에서 가장 낮았다.

집단간 점수차이 분석

컴퓨터 프로그래머의 출신학교별 차이를 검증해 본 결과 도식능력에서만 유일한 차이를 나타내고 있다(표 12). 다른 하위검사들도 기대할 수 있는 방향으로 학력에 따른 차이를 보이고 있는데, 코딩능력 검사에서는 지방대학 출신의 컴퓨터 프로그래머들이 약간 높은 점수를 받았다. 전반적으로 지방대학 출신과 서울소재대학, 특히 서울 기타대 출신 컴퓨터 프로그래머들 간 차이가 크지 않은데, 이는 본 연구대상자들이 한 회사에 모두 속해 있고, 우수한 지방출신 프로그래머들이 속한 부서를 대상으로 검사를 실시했기 때문으로 생각된다. 반면에, 성별이나 직군별로는 유의미한 차이를 보이지 않아 표를 제시하지 않았다. 또한 소프트웨어직과 연구개발직도 비슷한 업무를 수행하는 직군이기에 때문인지 직군 별로 의미 있는 차이를 보이지 않았다.

제 4차 연구

4차 연구는 하위검사의 특성에 관한 제조사 및 본 검사의 준거관련 타당도를 조사하기 위해 실시되었다. 준거변인 측정을 위해 상사지도관찰표를 사용하였다.

연구대상

4차 연구는 모그룹 계열사 소속 소프트웨어 관련 부서의 사원 146명을 대상으로 수행되었다. 이 집단의 평균 근무시간은 2년 4개월이었고 평균 연령은 27세 9개월이었다.

도구

3차 연구에서 언급된 검사와 동일하나 단지 수리력 검사만 자료 해석력과 수리적 부호화능력 검사로 분리시켰다. 이는 두 검사가 모두 수리적 능력을 다루고 있지만 앞의 연구들에서 수리력 검사의 내적 일치성이 다른 검사들에 비해 상대적으로 낮게 나왔고, 정해진 시간에 함께 실시했을 때 순서효과에 의해 후반에 위치한 수리적 부호화능력 검사는 몇 문제밖에 풀지 못하는 결과를 야기시키므로, 이 두 검사를 분리하여 실시하였다. 수리력 검사의 문항수는 각각 15개씩으로 하였고, 나머지 검사들은 20개씩 그대로 유지하여 총 110문항으로 구성된 검사가 되었다. 또 앞의 연구들에서 문항들의 p값이 다소 높은 것으로 드러나 검사 시간을 평균 2분씩 줄였다.

조사절차

본 검사의 준거관련 타당도를 조사하기 위해서 146명의 컴퓨터 프로그래머들을 대상으로 위 검사를 실시하였다. 모그룹 인력개발위원회 직원들의 협조를 얻어 본 연구자와 담당 직원 2명의 감독 하에 1시간 30분간 검사가 실시되었다. 준거변인으로 사용된 상사지도관찰표는 모그룹의 21세기 인재상 조사결과로 나타난 바람직한 인재의 특성들을 반영하는 총 49개의 문항으로 구성되었다. 즉, 컴퓨터 프로그래머의 독특한 업무 능력만을 평가하기보다는 전반적인 능

력, 성격특성, 대인관계 기술 등을 다루고 있다. 관찰표는 검사에 참가한 프로그래머들의 상사들에게 평가설문지를 작성하도록 하여 인력개발원 담당자를 통해 회수하였다.

분석방법

통계처리를 위해 PC용 SAS 패키지를 이용하였다. 하위척도간 관련성을 살펴보기 위하여 Pearson 상관계수를 구하였고, 준거관련(공인)타당도를 확인하기 위해 검사 총점과 준거변인(상사지도관찰표) 점수와의 상관계수를 구한 후 각 하위검사와 준거변인 점수간의 상관계수도 산출해 보았다.

연구 결과

하위검사별 통계

위 연구대상자들에 대한 평균 및 표준편차는 다음과 같았다(표 13).

표 13. 하위검사별 통계치: 평균 및 표준편차

하위검사(문항수)	평균(SD)
도식능력(20)	11.00(2.93)
코딩능력(20)	14.57(3.93)
수리적부호화능력(15)	6.26(2.56)
자료해석력(15)	6.43(2.15)
추리력(20)	6.14(1.91)
기획력(20)	8.93(2.43)

검사시간을 줄인 이유 때문인지 모든 하위검사에 있어서 대체적으로 평균점수가 낮아졌는데, 특히 수리력 검사가 시간조정의 영향을 가장 많이 받았다. 3차 연구와 비교해 볼 때, 평균점수가 반 이상 떨어졌고, 특히 표준편차가 큰 쪽으로 줄어들었다(예, 추리력: 3.10 → 1.91). 점수의 범위가 이처럼 줄어들

면 다른 하위검사나 총점과의 상관계수 및 준거변인과의 상관계수에 영향을 미치게 될 것이다.

하위검사간과 전체검사와의 상관관계

하위검사간의 독립성을 검증하기 위해 검사간 상관계수 및 검사 총점과의 상관계수를 구하였다(표 14).

표 14. 하위검사간 및 전체검사와의 상관계수

	도식 능력	코딩 능력	수리적 부호화 능력	자료 해석력	추리력	기획력
도식능력	1.00					
코딩능력	.34	1.00				
수리적 부호화능력	.38	.22	1.00			
자료해석력	.21	.25	.26	1.00		
추리력	.25	.19	.27	.32	1.00	
기획력	.18	.33	.20	.19	.13	1.00
전체검사	.63	.72	.62	.58	.54	.56

표에 나타난 상관계수를 자세히 살펴보면 제2, 3차 연구 결과와는 차이가 있음을 알 수 있다. 특히 추리력이 이전 연구들에서보다 총점과의 상관이 낮아지고 하위검사간 관계성에 약간의 변화가 있었다. 이는 문항난이도에 비해 추리력 검사 시간이 너무 짧아 상대적으로 적은 분산이 산출되었기 때문으로 해석된다.

준거변인과의 상관관계

상사지도관찰표 점수는 4~9점까지의 부적으로 편포된 분포를 보였으며, 평균은 6.44였고, 표준편차는 1.43이었다. 따라서 준거변인에 있어서 범위의 제한요인이 작용하여 상관계수가 낮아질 것으로 예상되었다. 본 검사와 준거변인인 상사지도관찰표와의 상관계수는 표 15와 같다.

표 15. 하위검사와 상사지도관찰표와의 상관계수

상사지도관찰표와의 상관	
도식능력	.21*
코딩능력	.08
수리적부호화능력	.19*
자료해석력	.14
추리력	.22*
기획력	.19*
전체검사(중다회귀)	.36*

앞의 표를 살펴보면 상사지도관찰표 평점 점수와 관련이 높은 것은 추리력, 도식능력인데 이는 SRA 컴퓨터 프로그래머 적성검사 manual (SRA, 1993)에 보고된 결과와도 일치하고 있어, 이 능력들은 컴퓨터 프로그래머 업무수행에 있어서 필수적인 것으로 추정된다. 상사평가점수와 6개 하위검사간의 중다회귀결과 $R^2=.13$ 이 산출되어 예측변인과 준거변인과의 상관계수는 약 .36정도로 추정된다. 이는 준거변인이 구체적 업무능력보다는 다방면의 종합적인 특성을 반영하고 있는 점과 양 변인 모두 범위의 제한에 영향을 받고 있다는 점 등을 고려할 때, 검사가 상당히 예측력이 있음을 나타내 준다.

위 연구결과를 토대로, 즉, 각 하위검사의 평균 및 분산, 총점과의 상관계수, 그리고 상사 지도관찰표와의 상관계수(예측력 계수)를 분석한 후, 각각의 중요도를 재검토하여 하위검사의 문항수와 시간이 재조정되었다(표 16).

표 16. 재조정된 하위검사의 문항수 및 시간

	문항수(개)	시간(분)
도식능력	30	12
코딩능력	30	10
수리적부호화능력	20	10
자료해석력	20	10
추리력	25	10
기획력	20	10
전체검사	145	62

제 5차 연구

5차 연구에서는 실제 선발상황에서 응시자를 대상으로 본 검사의 심리측정적 특성들을 전반적으로 검토해 보고자 하였다. 특히 요인분석을 실시하여 6개 하위검사의 구성개념도 살펴 보았다.

연구 방법

연구대상

제 5차 교차타당화 연구의 대상은 국내 모기업 소프트웨어직 선발 채용시험 응시자 1087명이었다. 이들 중 중·고등학교와 전문대학을 졸업한 사람도 10%정도 포함되어 있었으며, 평균 연령은 26세였다.

도구

4차 연구결과로부터 시간과 문항수가 재조정된 검사가 사용되었다(물론 실제 채용에서는 본 검사 외에 다른 검사들도 사용되었다).

조사절차

1995년 10월에 국내 모기업 소프트웨어직 응시자 1087명을 대상으로 검사가 실시되었고, 표준화된 절차(예: 녹음기 사용)에 의해 그 기업 인사 담당자들의 감독 하에 검사가 진행되었다.

분석방법

본 검사는 실제 채용상황에서 선발도구로 사용될 목적을 가지고 제작되었으므로, 실제 응시자를 대상으로 전반적인 사항에 대해 타당화 연구를 실시해 보는 것은 그 의의가 크다고 할 수 있다. 따라서 하위검사별 기초 통계치(평균, 표준편차) 및 평균 난이

표 17. 하위검사와 문항 통계 및 신뢰도 계수

	문항수(배정)	평균(SD)	평균난이도(P)	평균변별도	α 계수
도식능력	30	12.90(3.58)	.43	.31	.70
코딩능력	30	14.13(4.54)	.47	.36	.78
수리적부호화능력	20	6.75(2.37)	.24	.31	.53
자료해석력	20	6.48(2.14)	.32	.26	.47
추리력	25	16.27(3.03)	.65	.30	.64
기획력	20	12.04(3.50)	.60	.39	.71
전체검사	145	68.57(12.8)	.45	.32	.80

도(p), 변별도(양분점 상관계수)는 물론 내적일치도 계수 α 도 다시 구해보았다. 또한 하위검사간 상관계수와 각 검사가 총점에 고유하게 기여하는 정도를 산출하였으며, 구성개념 타당도를 확인하기 위해 F검증을 통해 집단간의 평균을 비교하였다. 또한 6개 하위검사의 구성개념을 살펴보기 위해 요인분석도 실시하였다.

연구 결과

하위검사와 문항 통계 및 신뢰도

위 응시자들을 대상으로 평균, 표준편차, 난이도, 변별도 및 신뢰도(α)를 구해본 결과 표 17과 같다.

위 표를 살펴보면, 각 하위검사들이 대부분 응시자들에게 적절한 수준을 유지하고 있다. 특히 추리력 검사는 평균이나 표준편차, 난이도 등에서 이전 연구에서보다 바람직한 방향으로 변화되었다. 전체검사의 난이도를 살펴보면 전체적으로 .50에 가까운 난이도를 보이는 반면, 수리적 부호화 능력과 자료해석력은 이미 고용된 프로그래머 집단에 비교해 볼 때 응시자들에게는 너무 어려운 듯하다. 시간을 좀 더 배정하거나(10분에서 12분으로 변경) 내용 면에서 좀 더 쉽게 수정할 필요가 있다.

또한 검사의 변별도나 α 를 하위검사별로 살펴

보면, 대체로 대부분의 하위검사들이 양호한 변별도와 신뢰도를 보이고 있으며, 특히 기획력과 코딩능력 검사에서 높은 변별도 및 신뢰도를 나타내고 있다. 그러나, 추리력, 특히 자료해석력의 경우는 상대적으로 낮은 변별도와 신뢰도를 보이고 있다. 물론 문항이 너무 어려웠기 때문에 예상될 수 있는 결과이기는 하지만 이 검사에 대한 문제검토와 개선이 요구된다.

하위검사간과 전체검사와의 상관관계

하위검사간 상관관계와 각 하위검사와 전체 검사와의 관계를 살펴보기 위해 상관계수를 구해보았다(표 18).

가장 높은 상관을 보이고 있는 하위검사들은 도식능력, 코딩능력, 수리적 부호화능력, 자료해석력이다. 반면에, 기획력과 추리력은 서로 상관이 매우 높으면서 앞의 4개 검사들과는 매우 낮은 상관관계를 나타내고 있다. 즉, 위 상관계수행렬은 하위검사들이 2개의 그룹 또는 cluster로 나누어지는 현상을 보이고 있다. 또 검사 총점에 미치는 고유한 설명력을 알아내기 위해 전체를 100%로 생각하여 각 하위검사의 분산을 %로 바꾸어 보았다. 총점과의 상관이 가장 높은 두 검사 즉, 코딩능력과 도식능력 검사의 기여도가 각각 18.36%, 15.38%로 높게 나타나서 그 중요성을 실감케하고 있다.

표 18. 하위검사간 상관계수 및 각 검사와 전체검사간 상관계수

	도식능력	코딩능력	수리적 부호화능력	자료해석력	추리력	기획력
도식능력	1.00					
코딩능력	.50	1.00				
수리적부호화능력	.40	.37	1.00			
자료해석력	.37	.44	.28	1.00		
추리력	.09	.14	.06	.04	1.00	
기획력	.12	.14	.05	.07	.44	1.00
전체검사	.70	.77	.55	.56	.50	.52
다중상관계수의 제곱을						
100%로 놓고 대체한 값	15.38(%)	18.38(%)	10.23(%)	11.59(%)	7.12(%)	9.29(%)

집단간 점수차이 분석

각 하위검사에 있어서 응시자간 학력별 평균차이가 있는지를 살펴보기 위해 평균값과 F값을 계산해 보았다(표 19).

이 표에서 나타난 바와 같이 6개 하위검사 모두 학력에 따른 유의한 차이를 보이고 있으며, 전체검사 총점 또한 유의한 차이를 보이고 있다. 특히 다른 하위검사들과 비교해서 코딩능력, 추리력, 기획력이 학력에 따른 차이가 큰 하위검사들이었다. 따라서 특수 직무능력검사인 컴퓨터 프로그래머 적성검사가 학력에 따른 높은 변별력을 가지는 것으로 해석될 수 있다. 물론, 고등학교나 전문대학 졸업자 중에서도 우수한 응시자가 많아 전체 합격자의 12%정도를 차지했다. 이는 종전에 4년제 대학 졸업자에게만 응시 가능했던 채용제도에서는 생각할 수 없던 일로 열린

인사제도의 고무적인 결과로 해석될 수 있다.

다음으로는 각 하위검사에 있어서 계열별 응시자간 평균차이가 있는지 살펴보았다(표 20).

전체 총점을 비교하면 이공계(70.5), 인문계(66.8), 예능계(58.8)로 소프트웨어직에서는 이공계 출신의 총점이 가장 높았음을 알 수 있다. 이러한 결과는 컴퓨터 프로그래머 적성검사가 이공계 성향이 짙은 소프트웨어직에 적절한 적성을 잘 반영해 주었던 것으로 해석할 수 있다. 이공계열이 적성에 맞다고 생각한 학생들이 고등학교에서 이과로 진학하고 전문대나 대학전공을 이공계열로 선택하는 경향이 있을 것이다.

하위검사별로는 특히, 도식능력, 수리적 부호화능력, 코딩능력, 추리력, 기획력에서 이공계 출신이 유의하게 높은 점수를 얻었던 것으로 나타났다.

표 19. 학력별 검사점수 분포(평균 및 SD)

하위검사	고졸	전문대졸	대졸이상
	평균(SD)	평균(SD)	평균(SD)
도식능력	11.1(3.4)*	12.0(3.5)*	13.5(3.3)*
코딩능력	11.6(4.2)*	12.9(4.1)*	15.0(3.9)*
수리적부호화능력	3.7(2.0)*	4.2(2.2)*	5.0(2.3)*
자료해석력	5.8(2.1)*	5.9(1.9)*	6.8(2.2)*
추리력	15.5(3.5)*	16.6(2.9)*	17.8(2.8)*
기획력	9.9(3.6)*	11.0(3.5)*	12.8(3.3)*
전체검사	57.6(12.2)*	62.6(11.9)*	70.9(12.9)*

표 20. 계열별 검사 점수분포(평균 및 SD)

하위검사	인문계	이공계	예능계
	평균(SD)	평균(SD)	평균(SD)
도식능력	12.7(2.7)*	13.5(3.3)*	9.3(3.8)*
코딩능력	14.4(4.3)*	14.8(4.4)*	12.7(4.0)*
수리적부호화능력	4.2(2.0)*	5.1(2.4)*	4.5(2.4)*
자료해석력	6.3(1.9)*	6.7(2.1)*	5.7(1.9)*
추리력	16.9(2.6)*	17.8(2.8)*	15.9(3.1)*
기획력	12.3(3.1)*	12.6(3.3)*	10.7(3.3)*
전체검사	66.8(8.9)*	70.5(11.2)*	58.8(9.9)*

요인분석

끝으로 본 프로그래머 적성검사의 구성개념 타당도를 살펴보기 위해 6개 하위검사간 상관계수를 이용하여 탐색적 요인분석을 실시하였다. 요인추출방식은 공통요인분석(common factor analysis)방법을 사용하였으며, eigenvalue와 스크리 검사(scree test)를 함께 고려하여 요인수를 선택하였다. 최종 요인구조를 얻기 위한 회전방식은 직각회전 방법인 varimax를 이용하였는데, 그 결과를 표 21에 제시하였다.

표 21. 회전된 요인구조(요인 부하량)

	요인 1	요인 2
코딩능력	.73	.13
도식능력	.69	.08
자료해석력	.56	.05
수리적부호화능력	.53	.02
추리력	.08	.66
기획력	.06	.66
eigenvalue	1.98	1.08
설명력변량의 비율(%)	64%	36%

표를 살펴보면, 본 검사의 구조는 2개의 요인으로 명확하게 구분되는데, 제 1 요인은 전체 공통분산의 64%를 설명하고 있으며 코딩능력, 도식능력, 자료해석력, 수리적 부호화능력 검사가 높은 부하량을 나타내 주어 “자료처리 혹은 자료전환 능력”과 관련된 검사들로 묶여있는 것을 알 수 있다. 즉, 실제 주어진 상황이나 자료를 신속히 이해·판단·해석하고

지시에 따라 자료를 부호화 하거나 또는 자료를 통해 의사 결정하는 능력들이 요인 1을 대표하고 있다. 제 2 요인의 경우에는 전체 공통분산의 36%를 구성하며 추리력과 기획력 검사로 이루어져 있다. 이 요인은 두 검사가 공통으로 측정하고 있는 특성인 “관계분석·평가 능력”을 나타낸다. 따라서, 본 검사에서 측정하고 있는 기본적인 능력은 자료처리/전환 능력과 분석/평가능력이라 볼 수 있다.

결론 및 논의

정보화 시대를 맞이하여 컴퓨터 관련 전문직에 관한 수요가 급격히 늘고 있는 가운데 이 분야의 효과적인 인재선발을 위해서 신뢰롭고 타당한 특수 적성검사의 필요성이 절실해지기 시작했다. 특히 현재 우리의 실정은 몇몇 기업들에서 사용하고 있는 컴퓨터 관련 적성검사가 외국의 검사를 그대로 번역한 것으로 검사의 활용도나 타당성에 관한 연구가 매우 부족한 실정이므로, 컴퓨터 프로그래머 적성검사의 개발 및 타당화 연구는 그 필요성 및 의의가 크다고 하겠다.

본 연구는 체계적인 직무분석을 통해 컴퓨터 프로그래머 적성검사를 개발하고, 개발된 컴퓨터 프로그래머 적성검사를 선발 도구로 활용하려는 목적을

위한 예비연구로서, 여러 차례에 걸쳐 다양한 집단의 컴퓨터 프로그래머들을 대상으로 검사의 타당도를 검증하였다. 그 결과 대체적으로 양호한 내적 일치도 신뢰도 지수를 얻었을 뿐 아니라, 컴퓨터 관련 전문가들에 의해 검토된 각 하위검사들이 높은 변별력을 보이고 있으므로 내용타당도가 지지되었다고 할 수 있다. 또 기대하는 방향으로 학력별, 계열별, 직급별로 전체검사 및 하위검사에서 집단 평균점수간 유의한 차이를 나타내어, 이 분야의 학문적 지능/적성검사의 성격을 갖는 특수 적성검사로서 높은 구성개념 타당도를 갖는다고 할 수 있다.

또한 본 검사는 범위의 제한과 같은 준거변인의 문제점 등이 있음에도 불구하고, 준거변인인 상사지도관찰표와 의미 있는 상관(.36)을 보이고 있다. 이는 검사가 예측력을 가지고 있음을 의미하는 것으로서, 본 검사 점수와 프로그래머들이 조직에서 성공하는 것과는 높은 관련이 있다고 말할 수 있으며, 이러한 결과는 본 검사의 준거관련 타당도를 지지하고 있다고 볼 수 있다. 특히 일반지능/적성검사가 일반적으로 업무수행능력과 .20정도의 낮은 상관을 보이고 있는 점을 생각해 볼 때(Ghiselli, 1973; Schmitt, Gooding, Noe & Kirsch, 1984; Wigdor & Garner, 1982), 본 컴퓨터 프로그래머 적성검사의 준거관련 타당도 정도는 고무적이라고 평가할 수 있겠다.

그 외에도 6개의 검사로 구성된 본 검사의 구조를 밝혀 보기 위해 탐색적 요인분석을 실시한 결과, 2개의 이론적 요인, 즉 자료 처리/전환능력과 분석/평가 능력이 도출되었다. 물론 각 연구에서 사용된 대상 집단마다 하위검사들간 상관계수의 변동이 있었으므로 타집단을 대상으로 추후 요인분석을 실시하여 2요인 모델을 재검토해야 하겠지만, 위와 같은 1087명이라는 비교적 큰 집단을 대상으로 2개의 이론적 요인이 추출된 것은 앞으로 이 분야의 검사개

발 방향을 제시해 주고 있어, 그 의미가 크다고 하겠다.

또한 본 연구는 네 차례의 타당화 연구를 진행하면서, 본 컴퓨터 프로그래머 적성검사의 문제점 및 제한점을 다각도로 검토하여 문항 내용 및 하위검사 수정은 물론 문항수와 하위검사시간을 재조정하였다. 그 결과 처음 제작시보다 더욱 신뢰롭고 타당한 검사로 발전되었다. 그러나 실제 응시자를 대상으로 실시된 제 5차 연구결과에서 수리력 검사, 즉 자료해석력과 수리적 부호화능력 검사가 너무 어려운 것으로 드러나 본 검사가 본격적으로 선발상황에 사용된다면 이 하위검사들 중 몇몇 문항들은 조금 쉬운 문항들로 대체하거나 검사시간을 늘이도록 조정해야 할 것이다.

결론적으로 본 연구에서 개발되고 검증한 컴퓨터 프로그래머 적성검사는 수정되어야 할 부분이 있지만, 타당도는 비교적 만족스러운 수준이라고 판단된다. 물론, 본 검사가 인사분야에 활발히 적용되기 위해서는 앞으로 더 많은 분석과 연구를 수행해야 할 것이다. 특히 본 연구에서는 구성개념 타당도 검증을 위해 집단별 차이검증과 요인분석을 실시했는데, 추후 연구에서는 구성개념 타당도를 좀 더 다각적으로 검토해보기 위해 일반지능검사 및 SHL이나 SRA의 컴퓨터 프로그래머 적성검사들과의 상관(범위타당도) 및 면접이나 실기, 성격검사와의 상관(변별 타당도)도 살펴보아야 할 것으로 생각된다. 또한 준거관련 타당도를 확증하기 위해서는 상사관찰표 외에도 실제 업무성적과의 상관정도를 관찰해 봄으로써 본 검사의 업무 예측력에 관한 연구가 활발히 진행되어야 할 것이다. 마지막으로 본 연구의 대상자들은 첫 번째 연구를 제외하고는 단일 그룹에 속해 있는 사원들이거나 단일 집단에 응시한 지원자들이었기 때문에 그들이 비록 여러 회사에 분포되어 있

기는 하지만, 연구결과를 일반화시키는데는 조심스럽다. 향후에는 보다 다양한 인구통계적 변인으로 검사를 검증해 볼 필요가 있으므로 교차타당화 연구가 계속되어져야 하겠다.

참고 문헌

- 김명소, 강혜련 (1994). 컴퓨터 프로그래머 직무분석. *정보과학회지*, 12(3), 63-71.
- 김명소 (1994). 컴퓨터 프로그래머 적성검사 개발에 관한 연구. *호서대학교 사회과학연구*, vol. 13(pp 333-334).
- 매일경제신문 (1995). 적성, 인성검사 새 "복병". 9월 25일자.
- American Psychological Association (1985). *Standards for educational and psychological testing*. Washington, DC: APA, Inc.
- Anastasi, A. (1990). *Psychological testing*. New York: Macmillan.
- Page, R. C & Caskey. D. T. (1988). Computer programmer in S. Gael(Ed.), *The job analysis handbook for business, industry, and government*, Vol. 2, New York: John Wiley & Sons.
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461-477.
- Schmidt, F. L., Greenthal, A. L., Hunter, J. E., Berner, J. G & Seation, F. W. (1977). Job sample vs. paper-and-pencil trades and technical tests: Adverse impact and examinee attitudes. *Personnel Psychology*, 30, 187-198.
- Schmidt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Meta-analysis of validity studies published between 1964-1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Schmitt, N. & Klimoski, R. J. (1991). *Research methods in human resources management*. Cincinnati, OH: Southwestern.
- Science Research Associates. (1993). *Examiner's manual of computer programmer aptitude battery* (4th ed.), Rosemont, IL: Author.
- Smither, J. W & Peariman, K. (1991, April). Perceptions of the job-relatedness of selection procedures among college recruits and recruiting/employment managers. In R. R. Reilly(Chair), *Perceived validity of selection procedures: Implications for organizations*. Symposium conducted at the Sixth Annual Conference of Society for Industrial and Organizational Psychology, St. Louis, MO.
- Society for Industrial and Organizational Psychology. (1987). *Principles for the validation and use of personnel selection procedures*(3th ed.), Washington, DC: Author.
- Wigdor, A. K & Garner, W. R. (Ed.), (1982). *Ability testing: Uses, consequences and controversies*. Washington, DC: National Academy Press.

논문 초고 접수 : 1997. 2. 17.
최종 수정본 접수 : 1997. 5. 6.

ABSTRACT

A Study on the Development and Validation of Computer Programmer Aptitude Inventory (CPAI)

Myoung-So Kim
Hoseo University

Computer technology is being applied everyday to increasingly diverse and complex problems. These vast applications create a demand for able computer programmers and for valid selection tools to identify persons with aptitudes for this position. However, there has been little effort to systematically define the job of computer programmers and develop job-related, specific aptitude instruments. The purpose of the present study was (1) to develop a computer programmer aptitude inventory reflecting the selection dimensions resulted from a worker-oriented job analysis, and (2) to investigate the validity of the inventory for various groups of computer programmers. The 250 experimental items were pilot tested with 150 college students majoring in computer science and analyzed for difficulty, discrimination, and internal consistency, and finally 100 items were selected for the experimental form of the CPAI. Four field studies were then conducted with approximately 1600 computer programmers to revise and validate the experimental CPAI. It was found that the reliability coefficients of most of the subtests were acceptably high. As expected, there were also significant differences between computer programmers on the bases of education level, job level and college major, reflecting adequate construct validity of the CPAI. For the concurrent validation of the inventory, the CPAI was administered to currently employed computer programmers of a large software company, and the scores correlated with supervisors ratings of each programmer on overall performance (i. e., ability, personality, attitude, and the like). The validity coefficient of the CPAI was .36 for this group. The CPAI total score, as well as four of its subtests, were found to significantly relate to the ratings, although the group appeared to be highly selected and restricted in range of both predictor and criterion, and the ratings were on the overall characteristics, not on specific job-performance. Factor analysis revealed two significant factors, the data transformation factor and analysis/evaluation factor, which could provide test developers with future directions for developing psychological tests in the areas of software. In shot, the problems and limitations of the experimental CPAI has been thoroughly examined and revised thorough the validation studies, and this resulted in an enhanced, i. e., more reliable and valid, aptitude instrument consisting of 6 subtests with 110 items. The findings of the present study, therefore, indicate that the CPAI can be used as a selection tool in industrial settings. The implications of these findings, limitations of this study, and the direction of future research were discussed.