

지필형 상황판단검사에 대한 비평적 고찰*

이 상 철† 이 순 목 조 영 일

성균관대학교 심리학과

지필형 상황판단검사는 저충실도 시뮬레이션에 속하는 것으로서, 종합평가센터(assessment center)와 같은 충실도가 높은 시뮬레이션에 비해 준비가 용이하고, 비용 및 경비가 절약되고, 준거타당도가 좋으며, 비교적 높은 안면타당도와 소수집단에 대한 불리효과가 작은 장점을 가지고 있다. 그러나 지금까지 지필형 상황판단검사의 연구는 준거타당도 중심으로 진행되어왔기 때문에 구성개념타당도 중심의 연구가 요망된다. 이에 따라 지필형 상황판단검사의 구성개념이 기초 실용지식에 기반한 실용지능이라는 주장과 다차원적인 구성개념이라는 이견이 제시되고 있다. 또한 지필형 상황판단검사는 측정도구로서의 방법효과만이 존재한다는 주장도 존재한다. 따라서 지필형 상황판단검사 연구방향으로는 지금까지 구성개념 파악이 어려웠던 한계점을 극복하기 위하여 연구영역으로서의 영역중심적 연구를 제안하였으며, 연구방법론적으로는 방법효과를 제거하기 위한 3단계 접근을 제안하였다. 또한 타당화의 방법으로서 과정모형을 통한 적극적 타당화를 하여 구성개념을 확립하는 연구방향을 제안하였다.

주요어 : 상황판단검사, 구성개념타당도, 적극적 타당화, 방법효과, 3단계 접근, 영역중심적 연구

* 이 논문의 작성을 위하여 2003년도 두뇌한국21사업으로부터 부분적인 지원이 있었음.

† 교신저자 : 이 상 철, 성균관대학교 심리학과, iopsy2000@skku.edu

본 연구는 국내 일부기업에서 선발용으로 사용되고 있는 지필형 상황판단검사에 대한 것으로서 그 검사형태의 출발지인 미국에서는 80여 년이나 된 역사에도 불구하고 최근에서야 그에 대한 연구가 본격적으로 시작되고 있다. 2003년도 미국 산업 및 조직심리학회에서는 지필형 상황판단검사에 대한 심포지엄이 여러개 운영될 정도로 그에 대한 연구가 가속화되고 있다. 국내에서도 이 검사형태가 선발장면에서 보다 보편화될 추세에 대비하여 이 검사의 장단점 및 연구방향에 대한 고찰이 필요하다. 일반적으로 지필형 상황판단검사는 저충실도 시뮬레이션(low-fidelity simulation)에 속하는 것으로서(Motowidlo, Dunnette, & Carter, 1990), 종합평가센터(assessment center)와 같은 고충실도 시뮬레이션(high-fidelity simulation)에 비해 제작이 용이하고, 비용 및 경비가 절감되고, 준거타당도가 있으며, 그리고 비교적 높은 안면타당도와 소수집단에 대한 불리효과가 작은 장점을 가지고 있다(Motowidlo et al., 1990; McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001). 한편, 지금까지 지필형 상황판단검사에 대한 연구들은 구성개념타당도 보다는 준거타당도에만 관심을 기울인 결과 과연 어떤 내용을 측정하는가에 대해서는 수렴되는 증거가 없다(Weekley & Jones, 1999; McDaniel et al., 2001). 따라서 지필형 상황판단검사의 유용성에 비추어 볼 때, 구성개념의 탐색을 통한 이론개발이 필요한 시점이다. 본 연구에서는 지필형 상황판단검사 연구에서의 문제점인 구성개념 탐색의 미진함을 개선하기 위한 영역중심접근의 개념화와 경험적 요인분석을 위한 3단계 접근을 제안하고, 보다 견실한 타당화 방향으로서 제 변수들간의 관계 및 업무성과에 대한 과정모형을 통한 적극적인 타당화를 제안하고자 한다.

지필형 상황판단검사의 준거타당도는 여러 연

구들을 종합적으로 살펴보았을 때 대략 0.20~0.62의 범위를 보인다. 이에 대하여 메타분석을 시행한 McDaniel 등(2001)은 지필형 상황판단검사의 준거타당도는 일반화가 가능하며 그 추정치는 0.34라고 보고하였다. 이러한 결과에도 불구하고, 개별 연구에서 준거타당도간 큰 차이를 보이는 원인에 대하여 Weekley 등(1999)은 지필형 상황판단검사를 통해서 측정하고자 하는 구성개념이 각 연구마다 서로 다르기 때문이라는 견해를 보였다. 즉, 수렴하는 구성개념들이 아니라 마치 오렌지와 사과를 섞는 것과 같은 방식으로 일반화 연구를 하고 있는 현실에 대한 지적이다.

이에 따라 Weekley 등(1999)은 지필형 상황판단검사의 구성개념에 대하여 주요한 세 가지 논쟁을 소개하고 있다. 우선, Sternberg와 동료들은(Sternberg, Conway, Ketrton, & Bernstein, 1981; Wagner & Sternberg, 1993) 지필형 상황판단검사가 영역내 기초실용지식(basic practical knowledge)을 측정하고 있으며, 그러한 측정치들이 실용지능을 구성개념으로 한다는 관점을 가지고 있다. 상황판단검사에서 측정되는 내용이 일반 생활이나 학교 장면에서 성공을 위한 지식이라는 견해(Oswald, Schmitt, Ramsay, Kim, & Gillespie, 2002)는 Sternberg와 동료들이 주장하는 기초실용지식과 유사한 점이 있다. 두 번째 관점은 지필형 상황판단검사가 업무 수행과 관련된 일반정신능력, 경험, 직무지식, 성격과 같은 여러 가지 개념들과 관련 있으므로 다차원적 구성개념이라는 것이다(예: Chan & Schmitt, 1997). 여기서 직무지식은 기초실용지식보다 전문적 내용이 된다. 기초실용지식이 영역중심의 기본지식이라면 직무지식은 영역내 특정 직무에서의 과제처리용 업무 지식 또는 전문지식이 된다. 끝으로 지필형 상황판단검사가 측정하고 있는 구성개념이 무엇인가에 대한 논쟁과는 달리, 지필형 상황판단검사가

구성개념을 반영하고 있는 것이 아닌 단지 측정의 기법(technique)일 뿐이라는 주장도 있다(예: McDaniel et al., 2001; Weekley & Jones, 1999). 이 여러 가지 주장은 그동안 상황판단검사의 구성개념에 대한 연구에 큰 진전이 없었음을 의미한다.

본 연구에서는 상황판단검사의 구성개념에 대한 이제껏 연구에서의 문제점을 지적하고 앞으로의 연구방향으로서 적극적인 타당화를 할 것을 제안하고자 한다. 이 글의 순서는 다음과 같다. 우선 상황판단검사에 대한 소개에서 개요, 지필형 상황판단검사의 형태 및 발달, 장점, 채점방식에 대한 요약을 제공한다. 다음은 문제점 및 구성개념에 대한 논란과 그에 대한 토론이 제시되고, 끝으로 연구방향에 대한 모형이 제시된다. 연구방향으로서는 영역중심적 접근, 구성개념도출을 위한 3단계 접근, 그리고 적극적 타당화를 제시한다.

지필형 상황판단검사에 대한 소개

상황판단검사의 개요

상황판단검사는 일련의 상황들이 지필, 언어, 또는 시각적인 형태로 제시되고 응답자의 반응이 요구되는 검사를 의미한다(Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001). 상황판단검사의 역사는 1920년대의 사회적 지능검사로 거슬러 올라갈 정도로 오래전부터 사용되어왔다(McDaniel et al., 2001). 일반적인 형태를 보면 일련의 직무관련된 상황들이 지필, 언어, 또는 시각적인 형태로 제시되고 응답자의 반응이 요구되는 방식의 검사들이다(Clevenger et al., 2001). 한편, 상황판단검사는 업무 수행의 예측에 사용되

고 있는 다른 검사들 즉, 일반정신능력검사, 성격검사, 또는 흥미검사 등과 같이 구성개념 중심으로 제작을 해온 것이 아니라 준거로서의 업무수행을 예측하느냐에 중점을 두고 발전한 것이다.

형태상으로 볼 때 상황 판단 검사들의 지필형이 오래전부터 등장(예: Moss, 1926)했음에도 불구하고, 보편적인 형태는 종합평가센터(assessment center)와 같이 실제 직무에 근접한 상황에서 피험자의 행동을 측정하는 형태였다. 즉, 지원자가 실제에 근접한 상황에 노출될수록 실제의 자기를 노출한다는 관점이었다. 종합평가센터에서 주로 시행되는, 고충실도의 시뮬레이션은 제작과 사용에 있어 비용이 많이 들고 편리성 저하라는 단점을 가지고 있었다(Weekley & Jones, 1999; Thornton, 1992). 따라서 고충실도 시뮬레이션의 타당도에 근접하면서 제작상의 경제가 있고 편리성을 높인 검사로서 지필형의 상황판단검사가 등장하게 되었다. 최근에 상황판단검사의 보편화된 사용은 Motowidlo, Dunnette, 및 Carter(1990)의 지필형으로 된 즉, 저충실도의 시뮬레이션에 대한 소개 이후로 본다(Clevenger et al., 2001). 선발 절차에 적절한 형태로의 발전이 이전부터 이루어졌지만(예: Millard, 1952), 이러한 초기 원형의 사용은 특정 관리자의 서류함에 배달된 작업요청에 대한 시뮬레이션을 시도한 서류함 검사(Fredericksen, Saunders, & Wand, 1957; Thornton, 1992에서 재인용)를 거쳐, Motowidlo 등(1990)의 지필형 검사의 개발로 한층 쉽게 사용되는 길이 열렸다. 즉, 역할연기(role-play) 또는 작업표본 등과 같은 고충실도 시뮬레이션과 달리 응답자들은 문서화되어 있거나 음성화되어 있는 기술문을 받고, 이 상황에 어떻게 반응할 것인가에 지필로 응답하도록 하는 형태를 갖추었다(Clevenger et al., 2001).

한편, 상황판단검사는 각 영역의 사회문화적 가치를 기반으로 제작되었으므로, 정답역시 정보 처리론적 검사에서의 특징인 객관적 해답이 아니라, 그 영역내 주제문제전문가(SME)들의 견해가 요약되어 간주관적인(inter-subjective) 경험적 해답을 설정하여 채점하게 된다. 따라서 각 영역에서 개발된 검사는 그 영역 특유의 가치와 규범을 반영하므로 상황판단검사가 무엇을 재느냐는 것보다는 “어떤 영역의” 상황판단검사가 그 영역에서의 성공에 필요” 무엇을 재느냐고 해야 할 것이다.

요약하자면 Motowidlo 등(1990) 이후 지필형 상황판단검사는 고충실도의 시뮬레이션이 가지는 장점인 타당도를 유지하면서 제작상의 경제성을 높이고 사용상의 편리성을 높인 지필형 검사로서 보편화되었다. 다음에는 이와 같은 지필형 상황판단검사가 가지는 형태 및 발달에 대하여 기술하고자 한다.

지필형 상황판단검사의 형태 및 발달

지필형 상황판단검사는 광범위한 작업 상황에서의 판단을 측정하기 위한 지필검사를 말한다(McDaniel et al., 2001). 지필형 상황판단검사의 형태는 크게 나누어 두 종류로 볼 수 있다. 첫째, 시나리오형태로서 작업 상황에서 판단능력을 측정하기 위해 작업상황시나리오와 이에 대한 행동대안으로 구성된 정답형 지필검사로서 논리적 유추보다는 상식, 경험 그리고 일반적인 기초지식을 도출하기 위한 측정을 중심으로 하였다. 1942년에 제작된 미육군용 판단검사(Army judgment test)가 대표적인 예이다. 둘째, 단일문장형으로서 상황에 따른 시나리오를 주지 않는 대신에 다양한 작업 관련 행동을 나타내는 문장을 제시하고, 그 행동의 적절성에 대하여 동의하는 정도를 측

정하는 검사이다(McDaniel et al., 2001). 예를 들자면 “효과적인 상사는 신입사원의 작업에 대해 세밀하게 감시를 해야한다”라는 기술문에 대해 동의의 여부를 측정하는 것이다. 다음은 지필형 상황판단검사의 발달을 시간적 흐름에 따라 개발되어 온 형태의 변천에 초점을 맞추어 표 1에 정리하였다.

1920년대에 제작된 “조지 워싱턴 사회적 지능 검사(George Washington Social Intelligence Test)”의 하위검사로서 “사회 상황들에 대한 판단(Judgment in Social Situations)”이라 부르는 단일문장형의 상황판단검사가 포함되었고, 이것이 역사적으로 최초의 상황판단검사가 된다.

이후에, 제 2차 세계대전 당시 군 심리학자들은 “군인들의 판단을 측정하기 위한 검사(Judgment Test of Soldiers)”를 개발하였다(McDaniel et al., 2001). 이 검사는 시나리오형으로서 다양한 행동대안에 대한 판단을 요구하였다. 따라서 문항들은 논리적 유추보다는 상식, 경험 그리고 일반적인 지식을 도출하기 위한 능력에 기초하였다. Northrop(1989)는 이 검사에는 일반정신능력이 상당정도 반영되어 있다고 지적하였다

1940년대부터 관리자의 잠재력을 측정하기 위해 다양한 형태의 지필형 상황적 검사들이 등장하였으며, 이들은 저충실도 시뮬레이션과 유사한 형태로 개발되었다. 대표적인 것 중 하나는 “실용상황판단검사(Practical Judgment Test)”이다(Cardall, 1942). 이 검사는 일상적인 비즈니스와 사회적 상황과 같은 실생활장면을 기술한 단일문장형의 중다선택 문항들로 구성되었다. 한편, Cardall (1942)은 “실용 판단(Practical Judgment)”은 지능이나 학업 배경과 같은 요인들과 통계적으로 독립적인 요인이라고 주장하였다. 그러나 Carrington (1949)과 Taylor(1949)등의 연구자들은 “실용 판단 검사(Practical Judgment Test)”가 일반 지능 검사와

표 1. 지필형 상황판단검사의 발달

지필형 상황판단 검사의 명칭	연구자 (연대)	목적 및 형태	영역	(관련된) 구성개념
George Washington Social Intelligence Test	Moss(1926)	- 작업상황에 대한 판단능력 측정 - 단일문장형	작업영역	일반정신능력
Judgment Test of Soldiers	US Army (2차세계대전, Northrop, 1989 재인용)	- 군인의 판단능력 측정 - 시나리오형	군내영역	일반정신능력
Practical Judgement Test	Cardall(1942)	- 일상적 비즈니스와 사회적 상황에 대한 판단능력 측정 - 단일문장형	작업영역	일반정신능력
How Supervise?	File(1945)	- 작업관련부분에서 판단능력 측정 - 단일문장형	기업 관리자 영역	일반정신능력
LEADS	Mowry(1957)	- 대인관계 기술문에 대한 판단능력 측정 - 단일문장형	생산관리자 영역	대인관계능력 의사소통능력
Supervisory Practice Test	Bruce & Learner(1958)	- 관리자가 접할 수 있는 작업관련상황 판단 또는 대처능력에 대하여 측정 - 단일문장형	관리자 영역	일반정신능력
Early Identification of Management Potential	Campbell et al(1970)	- 관리자의 성공잠재력에 대한 확인 - 단일문장형	기업 관리자 영역	일반정신능력
Test 905	Corts(1980)	- 작업자들과 상호작용하는 상황에 대한 중다선택문항중 최선의 해결책 모색 - 단일문장형	인사 관리자 영역	대인관계능력 의사소통능력
Low-Fidelity Simulation	Motowidlo et al(1990)	- 관리자 선발용 - 시나리오형	일선 관리자 영역	대인관계능력 문제해결능력
TKIM	Wagner & Sternberg (1993)	- 관리자의 기초실용지식 측정 - 시나리오형	기업 관리자 영역	실용지능

유의한 상관을 보인다는 증거를 제시하였다 (McDaniel et al., 2001에서 재인용).

그리고 그 이후에 “감독기술검사(How Supervise?)” (File, 1945; File & Remmers, 1948)가 개발되었다 (McDaniel et al., 2001에서 재인용). 이 검사는 사람들 사이에서 의사결정이 요구되는 실용적 상황에서 효과적으로 기능하기 위해 관리자들에게

요구되는 능력을 측정하기 위해 설계되었다. 즉, 산업 및 조직과 같은 실용장면에서 인간관계에 대한 관리자의 지식과 통찰을 측정하기 위한 것이 목적이었다. “감독기술검사”의 문항들은 관리자들이 직면하는 전형적인 작업관련 상황에서 어떤 판단을 내리는가에 대한 반응을 요구하였다. “감독기술검사”의 문항들은 감독활동 부분과

회사정책 부분으로 나누어 볼 수 있다. 감독활동 부분은 감독자들이 부하들을 다룰 때 경험하는 문제 상황들과 관련되며, 각 문항들은 간단한 문장의 형태로서 3점 척도로 구성된 단일문장형이었다. 결과적으로 “감독기술검사”에 대한 점수는 관리자들과 비관리자들로 구성된 집단간 차이에 대하여 성공적으로 변별을 보였다(Bruce & Learner, 1958). 예를 들어 Millard(1952)는 관리자로 구성된 두 집단에서 이 검사와 일반정신능력검사의 상관이 각각 .71과 .62로 나타났으나, 그렇지 않은 집단에서는 유의하지 않은 상관을 보고했다. 따라서 Millard는 “감독기술검사”가 본질적으로 지능검사라고 결론을 내렸다.

또한 “리더십 평가 및 개발 척도(Leadership Evaluation and Development Scale: LEADS)”는 Mowry(1957)로부터 개발된 방식으로서 여덟 개의 대인 관계문제들로 구성된 기술문과 그에 따른 복수선택형의 응답대안들로 구성된 단일문장형이었다.

“관리행동검사(Supervisory Practices Test)”(Bruce & Learner, 1958)는 관리자로서 판단이 요구되는 상황에서 효과적으로 관리자의 역할을 담당할 수 있는 능력을 파악하고자 단일문장형의 50문항으로 설계되었다(McDaniel et al., 2001; Motowidlo et al., 1990). 즉, 문항들은 관리자가 접할 수 있는 전형적인 작업 관련의 상황에서 어떻게 판단하고 대처할 것인가에 대한 행동대안을 선택할 수 있도록 구성되었다.

1950년대 말에서 1960년대 초에 지필형 상황 판단검사들은 대규모 조직에서 관리자의 성공을 예측하기 위한 선발검사의 일부분으로 사용되기 시작하였다. 예를 들어, “관리 잠재력에 대한 초기 식별(Early Identification of Management Potential)”은 관리자로서 성공 잠재력을 확인하기 위해 제작되었다(Campbell, Dunnette, Lawler, & Weick, 1970). 이 검사의 형태는 문제상황이 단일문장으로

로 주어지고 이를 해결하기 위한 다양하고 가능한 행동 또는 의사결정 중에서 가장 적절하다고 생각하는 것을 선택하도록 하였다.

한편, 미국 인사 관리국(U.S. Office of Personnel Management)에서 사용한 “Test 905”는 작업자들과 상호 작용하는 상황을 단일문장으로 제시하고 다수의 행동대안들 중 최상의 해결책을 선택하도록 하였다(McDaniel et al., 2001에서 재인용).

Motowidlo 등(1990)이 개발한 관리자 선발 수준 측정을 위한 지필형 상황판단검사 절차는 시나리오형으로서, 작업 상황의 시나리오가 제시되고 각 시나리오 아래에 다양한 행동대안들이 지원자들에게 제시된다. 이에 대해 응답자들은 각 상황에서 가장 적절한 조치와 가장 적절치 않은 조치 모두를 선택하도록 요구되었다. 어떤 행동이 가장 적절한지 부적절한지는 미리 조직내 전문가 모임에서 설정된다. 이 방식이 현재 산업 및 조직심리학에서 가장 많이 쓰이고 있는 방식이다.

Wagner와 Sternberg(1993)는 “관리자를 위한 묵시지 질문지(Tacit Knowledge Inventory for Managers (TKIM))”를 제작하였다. 그러나 여기서 실제로 묵시지가 측정되지는 않으므로 “기초실용지식”이라고 함이 더 정확한 명칭이다(Wagner, 2002). 이 검사결과는 그 저자들이 1980년대 초부터 연구해온 실용지능에 대한 측정치로서 사용되었다. 이 검사에서 높은 점수를 보이는 즉, 실용지능이 높은 사람들이 관리자 또는 경영자의 경력에서 성공적인 수행을 보일 수 있는지의 여부를 연구하기 위해서였다. 따라서 영역중심적인 상황 시나리오를 사용하여 영역내 사람들이 각 시나리오에서 제시하는 행동대안들을 수집한 후 실용지능이 높은 관리자 또는 경영자들의 응답프로파일을 경험적인 해답으로 선정해 놓고 일반 응답자의 응답을 그 경험적 해답에 비교하여 채점

하였다. 이 검사는 형태상으로는 지필형 상황판단검사이며 이전에 소개되었던 검사들에 비해서 길고 자세한 시나리오를 사용했다는 점에서 차별된다.

지필형 상황판단검사의 장점

저충실도 시뮬레이션인 지필형 상황판단검사는 다음과 같은 장점이 있다. 첫째, 상대적으로 개발하고 관리하며 점수화하는데 있어서 용이하고, 경비가 적게 들어간다. 그러면서도 여러 연구들에서 어느 정도의 준거 타당도를 보인다고 보고하고 있다(Motowidlo et al., 1990; Motowidlo & Tippins, 1993; Pulakos & Schmitt, 1996). 또한 최근의 메타분석연구는 범위축소와 준거의 비신뢰도에 의한 영향을 수정한 결과 지필형 상황판단검사의 준거타당도는 0.34라고 보고하였다(McDaniel, et. al., 2001).

둘째, 일반정신능력 검사를 포함한 다양한 능력검사에서 보고된 인종별, 성별에 따른 평균점수의 차이에 비하여 지필형 상황판단검사에서는 인종별, 성별에 따른 평균점수 차이가 전반적으로 작다는 점이 보고되었다(예: Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Motowidlo et al., 1990; McDaniel et al., 2001). 이는 곧 소수집단에 대한 불리효과(adverse impact)가 작다는 것을 의미한다.

셋째, 지필형 상황판단검사는 지원자들이 비교적 높은 안면타당도가 있다고 지각하여, 산업장면에서 입사 지원자에게 직무관련 선발절차로서 보다 적절하다는 평을 받는다(예: Motowidlo et al., 1990; McDaniel et al., 2001). 이는 지필형 상황판단검사의 문항제작 과정상 문항내용들이 실제 활동상황 또는 직무상황으로 구성되기 때문이다. 이와 같은 안면타당도에 대한 지각은 피검

사자들로부터 긍정적인 반응을 유발한다(Bauer, Maertz, Dolen, & Campion, 1998; Kluger & Rothstein, 1993; Ployhart & Ryan, 1998; Smither, Reilly, Millsap, Pearlman, & Stoffey, 1993). 이러한 긍정적인 반응은 검사를 받을 때의 동기화에 기여한다.

이상의 논의를 통하여 지필형 상황판단검사는 종합평가센터에 비해 제작이 용이하고, 비용 및 경비가 절약되며, 일반화가 가능한 준거타당도를 가지며, 그리고 비교적 높은 안면타당도와 소수집단에 불리효과가 작은 장점을 가지고 있음을 알 수 있다.

지필형 상황판단검사의 채점방식

지필형 상황판단검사에 대한 지금까지의 채점 방식은 원형성(prototype) 방식과 시나리오점수 방식으로 대별된다. 원형성 방식은 업무수행이 좋은 사람들의 응답 프로파일(profile)을 정답으로 하여 일반사람들의 응답 프로파일과의 거리계산을 하는 방식으로서 Sternberg와 동료들에 의해서 1980년대부터 사용되어 왔다(Sternberg, Conway, Ketron, & Bernstein, 1981; Sternberg, Wagner, & Okagaki, 1993; Wagner & Sternberg, 1993). 이 방식은 각 시나리오별 대안적 행동 또는 반응들을 수집하여 전문가들이 볼 때 그 영역에서 성공적인 사람을 그 대안이 묘사하는 정도, 이른바 특징도를 구한후 각 문항에 대한 이들의 평균치를 경험적인 해답으로 설정한다. 이후 입사지원자들의 점수계산시 이 경험적 해답으로부터의 거리를 계산하여 점수화한다.

시나리오점수 방식은 지금까지 상황판단검사에 대한 대다수 연구들(예: Motowidlo et al, 1990; Weekley et al, 1999)에서 주로 사용되어온 방식으로서 가장 좋은 대안이나 가장 나쁜 대안 또한 둘 모두를 선택하도록 한 후, 그에 대한 응답형

태에 따라 그 시나리오에 대한 총괄적 채점이 되는 방식이다. 따라서 각 시나리오에 대한 점수를 전체 시나리오에 걸쳐서 합산하면 개인의 점수가 된다. 이와 같은 시나리오점수 방식이 상황판단검사 연구에 있어 산업심리학에서 널리 사용되어 왔다. 시나리오점수 방식에서의 정답은 각 시나리오 아래의 행동대안에 대하여 최상(best)의 대안과 최악(worst)의 대안을 전문가들간의 협의를 통하여 설정한다. 단, 최상과 최악의 행동간 차이가 없는 시나리오는 삭제하도록 한다. 이후 정리된 각 시나리오별 적절한 수효의 행동대안이 포함된다. 이후 점수화 전략에 따라 각 시나리오에서의 점수를 구한다.

시나리오 점수 방식에서 점수화 전략은 Weekley와 Jones(1999)의 전략을 참고 한다. 응답자가 최적/최악의 대안을 골라야 하는 경우를 보자. 최적/최악의 대안을 모두 올바르게 제시하면 +2점, 최적/최악의 대안 중 하나는 제대로 맞추었고, 하나는 최적도 최악도 아닌 것을 답하면 +1점, 어느 것도 정답을 맞추지 못했으나 거꾸로(예: 최적 대안으로서 최악의 대안을 제시, 최악의 대안으로서 최적 대안을 제시) 답한 것은 없을 때 0점을, 최적/최악의 대안중 하나는 거꾸로 응답하고, 하나는 거꾸로가 아닌 그러나 오답을 제시할 경우 -1점, 최적/최악의 대안을 모두 거꾸로 제시하면 -2점을 준다. 이와 같은 점수계산 전략에 따라 지필형 상황판단검사의 총 점수는 시나리오별 점수의 합으로 계산한다.

지필형 상황판단검사의 문제점

증분타당도의 결여

산업심리학에서 새로운 검사가 지지받기위해

거치는 절차인 증분타당도는 지필형 상황판단검사에 대해서 아직 보고되지 않고 있다. 즉, 일반정신능력검사의 준거타당도에 비해, 지필형 상황판단검사의 추가로 얼마나 타당도가 증가하는지의 연구가 없는 것이다. 타당도의 증분이 0에 가까울 경우, 현장에서의 인기에 무관하게 선발용도구로서의 근거를 상실하게 된다. Schmidt(1996)의 메타분석에서 제시된 바와 같이 미국에서의 평가센터(Assessment Center)는 선발장면에서 일반정신능력에 추가하는 증분타당도가 매우 낮아서, 최근에는 선발용보다는 오히려 개발용으로 더 많이 쓰이게 되었다. 지필형 상황판단검사는 그러한 연구 이전에, 높은 안면타당도로 인하여 실제장면에서 널리 쓰이게 된 것이다.

영역중심의 개념화 결여

또한 이 검사가 무엇을 재는가에 대해서도 충분한 연구가 없다. 그 이유 중 하나는 지필형 상황판단검사 점수의 활용에 있어서, 많은 직무와 상황에 걸쳐 타당한 예측변수로서 알려진 일반정신능력검사와 같은 방식으로 개념화하였다는 데 있다고 본다. 즉, 지필형 상황판단검사는 제목이 나타내듯이 상황이나 장면에 따라 적용여부가 달라지는 것임을 인식하지 못하였다는 점을 지적할 수 있겠다. 상황판단검사에서 측정되는 판단력은 일반정신능력에서 언급되는 탈맥락적 인지가 아니라, 맥락의존적인 상황인지(situated cognition)이다. 따라서 상황판단검사는 전통적인 정보처리론이 아니라, 심리학에서의 사회문화적 혁명에 부응하여 부각되는 사회문화적 가치 기반의 검사라고 할 수 있다. 따라서 타당화는 주어진 영역에 제한되고 탈맥락적인 일반정신능력 과 같은 폭넓은 일반화를 적용할 수가 없다. 일반정신능력의 경우에서와 같이 많은 직무, 상황

에 걸친 일반화를 하고자 할 것이 아니라 영역별 타당도를 보아야 할 것이다. 그렇다면 McDaniel 등(2001)의 전통적인 '타당도일반화'를 목적으로 한 메타분석은 상황판단검사에는 해당되지 않을 것이다. 오히려 영역내에서 여러 직무에 적용된 연구들에 대한 메타분석이 의미있을 것이다.

그러나 영역의 정의는 일정한 틀이 없는 상태라고 할 수 있다. 예로서 Motowidlo 등(1990)에서는 미국의 원격통신산업내 7개의 회사로 된 거국적 컨소시엄의 지원을 받아 5개 기능분야(관리/엔지니어링/마케팅/영업/프로그래밍/감독)를 경영영역으로 간주하고 상황판단검사를 개발하였다. 한편 Borman, Hanson, Oppler, Pulakos, 및 White (1993)는 군대장면에서, Wagner와 Sternberg(1985)는 몇 가지 직업장면에서, 그리고 Williams 등(1986)은 학교장면에서 지필형 상황판단검사를 개발하였다. 이러한 연구들에서 얻은 측정치들에 대해서, 일반정신능력 검사와 같은 타당도일반화를 요구해서는 안될 것이다.

구성개념타당도의 결여

상황판단검사들의 타당화 과정에서 이들이 무엇을 측정하는가에 대한 경험적 구조의 제시 없이, 총점을 사용해서 준거와의 관련을 제시한 것이 지금까지의 연구경향이다. 이러한 타당도를 좋게 보면 준거타당도이지만, 구성개념이 연구되지 않은 상태이므로 단지 측정기법의 타당도일 뿐이다. 이로 인하여 지필형 상황판단검사는 단지 기법일 뿐 구성개념이 존재하지 않는다는 견해(Weekley & Jones, 1999)를 가져왔다. 이것은 영역중심의 개념화 결여와도 관련이 있다. 즉, 지필형 상황판단검사들이 개발된 상이한 영역 차이에도 불구하고 많은 영역에 일반화 될 수 있

는 구성개념을 보고자 한 것이 영역내 구성개념 타당도의 연구에 지연을 가져왔을 수가 있다. 종합평가센터(assessment center)에서 실시되는 상황판단검사들에 대한 고찰에서 Landy와 Shankster (1994)가 내용요인보다는 방법효과로서 활동(exercise)요인만이 나타나고 있다고 결론내린 것 역시 영역중심의 구성개념 연구 대신에 보다 일반적인 구성개념을 찾고자 하는 가운데 그러한 구성개념은 도출되지 않고, 상황판단검사에 대하여 측정기법으로서의 타당도만이 있다는 보고를 하게 된 것이다.

그러나 구성개념 타당화 이전에 널리 사용되고 있는 검사 또는 측정기법은 상황판단검사만 이 아니다. Rothstein, Schmidt, Erwin, Owens, 그리고 Sparks(1990)는 면접이나 전기자료 역시 같은 처지라고 주장하고 있다. 그러므로 지필형 상황판단검사의 유용성에 비추어 볼 때 이의 구성개념을 탐색해 보는 것은 앞으로의 이론개발에 중요한 의의가 있다. 즉, 구성개념이 분명해지고 명세화될 때 비로소 업무수행 및 회사적응이라는 준거들을 설명하는 과정적 모형을 통한 설명이 가능해 질 수 있다. 다음은 지필형 상황판단검사의 구성개념에 대한 주장들을 개관하고자 한다.

지필형 상황판단검사의 구성개념 논쟁

실용지능이라는 관점

Sternberg 등(1995)은 지능에 대하여 전통적인 학업 지능이상의 개념으로 확장되어야 한다고 주장한다. 따라서 일반정신능력검사에서 일반적인 기본지식을 테스트하듯이 실용지능의 측정은 기초 실용지식검사를 통해서 가능하다고 보았다.

그들이 제작한 검사들은 일종의 상황판단검사였으며, 그들은 그 검사가 실용지능을 측정하는 척도라고 주장하므로, 이것은 그들의 상황판단검사가 재는 구성개념을 단일한 의미의 실용지능을 재는 것으로 간주하게 한다(Weekley & Jones, 1999). 그러나 과연 Sternberg등이 실용지능을 단일개념으로 보았는가에 대해서는 의문의 여지가 있다. 그들은 실용지능의 여러 측면으로서 자기 관리, 타인 관리, 및 과제관리를 제시하였으며 다차원 척도법을 통해 몇가지 차원(예: Horvath et al., 1996; Sternberg, 1985)을 제시한 바 있다.

Horvath등(1996)은 군인들의 리더십연구에서 단일문장형의 상황판단검사를 제작하고, 문항간 유사성/상이성 자료를 가지고 다차원척도법을 통해서 복수의 차원(예: 소대장의 경우 대인스킬에 자신감 얻기, 리더십스타일의 정의, 입장을 분명히 하기, 책무성 강화)을 제시한 것은 그들의 상황판단검사에서의 개념인 실용지능을 단일개념으로만 보지는 않았음을 의미한다. 그러나 그들이 제시한 대부분의 경험연구에서 실용지능을 단일개념으로 제시한 것이 Weekly등(1999)에 의해서 '단일개념으로 보고 있다'는 오해를 불러일으킨 것으로 이해된다.

직무 관련된 다차원적인 구성개념이라는 관점

상황판단검사의 구성개념에 대한 또 다른 관점은 다차원적인 구성개념(multidimensional construct)으로 구성된다는 관점이다(예: Weekley et al., 1999). 이와 같은 관점에 입각하여 상황판단검사를 직무관련 변수들인 일반정신능력, 성격, 경험, 그리고 직무지식과 비교하는 연구들이 존재한다. 우선 인지능력에 초점을 맞춘 연구로는 Chan과 Schmitt(1997), 그리고 McDaniel, Finnegan, Morgeson, Campion, 그리고 Braverman(1997)의 연구가 대표적

이다. Chan과 Schmitt는 상황판단검사가 개인들이 판단해야 하는 문제에 접근해가는 적절한 해법 또는 다양한 능력과 스킬(skill) 차원들을 포함한다는 점에서 성질상 다차원적이다라는 점을 지적하고 있다. 따라서 이런 검사들의 점수는 인지능력의 영향을 나타낸다고 주장하였다. 또한 McDaniel 등(1997)은 메타분석(meta-analysis)을 통해 지필형 상황판단검사가 일반정신능력과의 추정된 상관이 0.53이라고 보고함으로써 Chan과 Schmitt의 주장을 지지하였다. 물론 상관계수는 어떠한 종류의 구체적 관계인지 명시하는 것이 아니지만 적어도 관련성이 있음에 대한 증거로서는 가능하다.

둘째로 작업경험(work experience)이 상황판단검사 점수와 관련되어 있다는 연구결과들이 있다(Sternberg et al., 1995; Weekley & Jones, 1997). 이는 직무 관련된 경험이 많은 사람들일수록 상황판단검사에서 기술된 특정 문제상황에서 어떻게 처리해야할 지에 대하여 적응과 학습을 했을 가능성이 높으므로 작업경험과 상황판단검사 점수는 정적인 관계가 있음을 알 수 있다(Weekley & Jones, 1999).

셋째로 직무지식과 상황판단검사 점수와 관련되어 있다는 연구들이 존재한다. Hunter(1983)가 인지능력과 경험이 직무지식을 매개로 업무성공에 영향을 미친다는 이론 모형을 제안한 이후로 직무지식의 매개효과는 널리 수용되어왔다(Borman et al., 1995; Quinones, Ford, & Teachout., 1995). 즉, 인지능력과 경험은 직무지식의 획득 및 학습에 정적인 영향을 미쳐 업무성공에 영향을 미친다는 이론 모형이다. Schmidt(1994)는 상황판단검사가 업무성공과 관련있는 다른 구성개념들인 인지능력(예: Chan et al., 1997; McDaniel et al., 1997), 작업 경험(예: Sternberg et al., 1995; Weekley & Jones, 1997)등과 관계가 있고 상관계수의 크기로

볼 때 직무지식과 관련이 많을 것이라고 주장하였다. 여기서의 직무지식은 상황판단검사에서 측정되는 기본지식에 비해서는 보다 전문성이 있는 지식이다. 상황판단검사에서는 특정직무보다 넓은 범주인 영역을 대상으로 하므로 특정직무에 대한 지식을 묻게 되지 않기 때문이다. 또한 상황판단검사가 주로 사용되는 신입사원 선발상황에서 전문지식을 측정하는 것은 신입사원 직무의 관점에서 볼 때 내용타당도의 시비를 일으킬 수 있고, 반드시 전문지식을 측정해야한다면 검사가 아닌 다른 방법, 즉 자격증, 졸업장, 또는 실기시험 등이 있을 수 있다. 그렇다면 상황판단검사는 직무지식과 관계있으며, 영역내 적성에 가까운 구성개념을 가지게 될 것이다. 즉, 인지적/비인지적 성격의 기초역량들을 구성개념으로 가질 가능성이 있다.

넷째로 상황판단검사와 성격과의 관련성에 대하여 초점을 맞춘 연구로서 다음과 같은 연구들이 있다. Mullins와 Schmitt(1998)는 상황판단검사가 5요인검사(Costa & McRae, 1992)의 NEO Five Factor Inventory 측정치중에 성실성($r=0.26$), 호감성($r=0.22$)과 가장 강한 상관을 보인다고 보고하였다(Clevenger et al., 2001에서 재인용). 그러나 Smith와 McDaniel(1998)은 상황판단검사가 성실성($r=0.32$), 정서안정성($r=0.22$)과 상관이 있음을 보고하였으나, Mullins와 Schmitt(1998)의 연구에서처럼 상황판단검사와 호감성 간의 관련성은 발견하지 못하였다. 따라서 이 두 개의 연구가 완전하게 일치하는 것은 아니지만, 상황판단검사가 성격의 측면과 어느 정도 관련된다는 충분한 증거를 제시하고 있다.

상황판단검사의 두 얼굴

상황판단검사에서 대비되는 두 측면들을 여러

방향에서 조명해 본다. 많은 산업심리학자들은 상황판단검사에서 측정되는 구성개념이 무엇이라고 단정은 못 내리지만 인지능력, 성격, 동기, 경험, 직무지식 등과 상호관련성 있는 다차원요인이라고 주장하였다. 그러나 상황“판단검사”라는 제목에서 보듯이 상황에 대한 판단을 요구하는 검사이므로 인지능력을 기본으로 한 검사임에는 틀림없다. 그런데 상황관련의 인지능력이므로 주어진 상황에서의 인지로 불수가 있고, 이것은 Cronbach(1990)가 말하는 최대수행과 전형적수행의 두 측면을 함께 가지고 있는 검사라고 할 수 있다. 즉, 단순히 최대수행만 보는 탈맥락적 판단 즉, 추론검사도 아니고 전형적수행만 보는 성격/동기 검사도 아닌 복합적 성질의 검사이다. McDaniel, Hartman, 및 Grubb(2003)는 자신들의 메타분석에서 검토된 상황판단검사들에서 물음방식을 보면 시나리오물에 주어진 행동대안들 중 ‘~가장 좋은/나쁜’ 행동대안을 묻는 경우와 ‘당신이라면 (통상적으로) 어떻게 하시겠습니까?’의 두 가지임을 보고하였다. McDaniel 등(2003)은 전자를 지식에 대한 질문이고, 후자를 행동성향에 대한 질문이라고 하였는데, 지식형 질문의 경우 인지검사들과 관계가 높았고, 행동성향 질문의 경우 비인지검사들과 관계가 높았다. 이 경우 역시 요인을 추출한 것이 아니고 지식형 질문에서의 총점, 행동성향 질문에서 총점을 사용해서 조사된 것이다. 이것은 상황판단검사에서 물음방식에 따라 최대수행에 관련된 개념이 측정되거나 전형적수행에 관련된 개념이 측정됨을 의미한다. 그러나 산업심리학에서 사용되어온 질문방식은 행동대안 중 하나 또는 두개를 선택하고 그 결과에 의해 시나리오에 대한 점수를 매기는 것이므로 행동대안이 아닌 시나리오가 채점의 단위로 되어 시나리오효과와 내용효과를 어떤 방식으로든 분리할 수 없는 채점이 되어버

린다. 즉 요인구조를 파악하기가 어려워진다.

그러면 Sternberg 등의 실용지능과 McDaniel 등 (2003)의 지식형 척도 및 행동성향척도가 측정하는 구성개념간에는 어떤 관계가 있을 것인가? Sternberg 등은 지식형이나 행동성향의 방식으로 묻지 않았다. 암묵적 이론(implicit theory)의 관점에서 물음이 제시되었다. 즉, 실용지능에 대해서 사람들이 가지고 있는 개념을 Rosch(1978)방식으로 접근하여 포괄적으로 정의하고자 하였으므로 지식형도 행동성향의 형식도 아니었다. 예로서 '아래의 시나리오 상황에서 주어진 행동대안들이 실용지능이 높은 사람이 취할 특징적 행동이라고 생각되는 정도를 주어진 평정척도상에서 표시하시오'라고 하는 형식이었으므로 그것은 지식일수도 있고 행동성향일수도 있다. 이러한 문항별 응답자료에 대해서 전문가의 경험적 해답과 비교한 거리를 응답자의 점수로 하였으나 Sternberg 등은 그 자료에 요인분석을 하지 않은채로, 총점 자체를 실용지능의 측정치라고 하고 있다. 이들은 문항을 채점단위로 하고 있으므로 만일에 그 응답자료에서 시나리오효과를 제거하고 요인분석을 하였다면 요인중 일부는 McDaniel 등이 제시한 지식형 질문점수에, 나머지는 행동성향질문점수에 관련을 보일 것이다. 그러나 이러한 요인들이 상황에 대한 판단을 요구하므로서 측정되는 것이지 실제 행동은 아니므로 전반적으로는 "상황판단"으로 또는 "상황인지"라고 할 수 있다.

연구방향제안

상황판단검사의 영역중심적 특성을 감안한다면 어떤 특정의 상황이나 영역을 설정하고서 그 범위내에서 지필형 상황판단검사가 재고 있는 구성개념에 대한 탐색을 시도해야할 것이다. 상

황판단검사에 잠재한 요인들을 포괄적으로 "상황판단"으로 부르기로 한다. 지식형이 될 수도 있고 행동성향이 될 수도 있으므로 이 글에서는 잠정적으로 '상황판단'으로 해놓고, 앞으로의 경험연구에서 좀더 구체화되기를 기대한다. 또한 상황판단검사의 구성개념은 기존의 많은 구성개념들이 얽혀서 산업장면에서의 수행에 대한 설명모형을 이루는 논리적 연결망(nomological network)속에서 위치가 확인될수 있어야 한다. 이를 위해 Cronbach(1988, 1990)가 정리한 적극적 타당화(strong program of construct validation)의 방법을 제안 한다. 적극적 타당화는 소극적 타당화(weak program of construct validation)에 대비된다.

소극적 타당화는 탐색적 경험주의라고 표현할 수 있다. 즉, 검사 개발자가 검사점수와 다른 변수들간의 상관과 같은, 보다 많은 정보를 제공할 수록 타당화에는 도움이 된다는 다다익선(多多益善)의 입장을 취하고 있다. 예를 들어, 스탠포드-비네 검사(Stanford-Binet scale)의 제4판에서는 검사를 이해하는데 도움을 주는 정보들은 검사의 구성개념 타당도를 이해하는데 기여할 수 있다고 하고 있는데, 이는 이른바 소극적 타당화의 개념을 반영한다. 이와 같은 소극적 타당화는 검사에 대한 해석과 이해를 증진시키고자, 탐색적 요인분석으로부터 극단치를 나타낼 수 있는 사례 연구에 이르기까지 다양한 방법들이 포함된다. 한편, 적극적 타당화는 정교한 이론적 해석에 초점을 맞춘다. 즉, 변수간 관계에 대한 예측을 제시하는 논리적 연결망에 기초한 경쟁가설들을 검증하므로서 타당화하는 방식이다. 본 연구에서는 적극적 타당화 방식을 사용하기 위해 일반정신능력, 상황판단, 작업경험, 직무지식, 성격, 개인사, 동기, 조직몰입, 조직지원, 과업환경 그리고 업무성과와 같은 변수들과의 관계 및 업무성과에 대한 과정모형의 개발 및 검증하는 과정을

통해 지필형 상황판단검사의 구성개념을 보다 적극적으로 검토할 것을 제안한다. 적극적 타당화를 위한 전제조건은 영역중심적 검사를 제작해서 구성개념이 있는지 알아보는 것이다.

영역중심적 지필형 상황판단검사의 제작

영역중심적 지필형 상황판단검사의 제작은 특정 영역을 중심으로하여 중요상황 또는 직무상황에 대한 시나리오제작과 그 상황에서의 행동대안 제작으로 구성된다. 여기서의 직무상황은 제작될 상황판단검사가 사용될 영역에서 세분화되고 구체화된 상황이다. 이 검사가 신입사원대상의 선발도구가 될것이라면 입사후 훈련가능하거나, 다른방식(예: 졸업장, 자격증)으로 검토 가능한 전문지식은 제외될 것이다. 따라서 선발도구에 반영될 수 있는 상황은 영역내 상식이나 초임 직무의 수행 또는 학습에 대한 준비도의 성격을 가질 것이다. 그렇다고 일반정신능력, 성격, 동기 등의 기존에 정착된 구성개념을 대체하는 것도 아닐 것이다. 한편 각 시나리오에서의 물음방식은 Sternberg등에서의 평정척도 방식이 “최적/최악” 대안만을 선택하게하는 방법보다 유리하다. 즉, 시나리오가 아닌 문항을 채점단위로 하므로 각 대안에 대한 응답에서 시나리오 효과를 제거하고 요인분석을 할 수 있기 때문이다 (Lee & Kim, 2000). 또한 수집한 시나리오 및 대안적 행동 또는 반응을 평정척도로 제작하는 과정에서 한 시나리오 아래의 문항들이 가급적 독립적 내용일 경우 나중에 문항점수에서 시나리오 효과의 제거에 용이하다(Lee & Kim, 2000). 현재 산업 및 조직심리학에서 널리 사용되고 있는 “최적/최악” 대안의 선택은 채점단위가 문항이 아니라 시나리오가 되므로 구성개념 연구에는 어려움이 있으나 전체 시나리오에 대한 총점을

가지고, 지금까지의 관행과 같은 준거타당화에는 사용가능하다.

상황판단검사의 구성개념 도출

지금까지 상황판단검사에 대한 연구들을 살펴 보았을 때 구성요인을 밝히는 연구는 존재하지 않았다. 이는 주로 준거타당도에 초점이 맞추어져 연구가 이루어져왔기 때문이기도 하지만, 각 시나리오내 문항점수에 반영되는 시나리오 효과(방법효과)를 제거하지 못한 이유에도 기인한다(이순목, 김아영, 조영미, 2000). 따라서 상황판단검사의 구성개념 도출을 위해서 이순목 등(2000)과, Lee와 Kim(2002)의 방식을 따라 1단계로서 각 시나리오내에서 시나리오 효과를 제거한 잔차화된 문항점수들을 구한후, 2단계에서 잔차화된 점수들에 대하여 탐색적 요인분석을 하고, 3단계에서는 원자료로 돌아가 2단계에서의 내용요인구조와 검사제작시의 시나리오 설정에 따른 방법효과를 가설로 한 확인적 요인분석을 통해 지필형 상황판단검사의 구성요인을 재검토한다. 이때 확인적 요인분석시에 동일한 시나리오에 속한 문항들간에 방법효과를 공유하므로 방법요인의 설정이 가능하다. 그러나 방법효과를 명시적으로 설정할 경우(예: Widaman, 1985의 CTCM모형) 자료분석시에 수렴이 안되거나 부적절한 해가 나오는 것이 오랜 실험을 통해 밝혀졌다(Marsh, 1989; Kenny & Kashy, 1992). 따라서 방법효과를 공유하는 문항들의 고유요인간에 상관을 시키는 모형(Marsh, 1989; Kenny & Kashy, 1992)을 사용할 것이 권고된다. 이렇게 해서 일단 구성개념의 도출이 이루어지면 새로운 자료에 그 결과를 다시 적용하는 교차타당화가 필요할 것이다.

다음은 적극적 타당화의 연구방향으로서 지필형 상황판단검사의 구성개념에 대한 연구와, 기

존의 예측변수와 준거들간 관계속에서 상황판단 요인의 이론적 위치탐색을 토론하고자 한다.

지필형 상황판단검사의 구성개념에 대한 연구

앞서 이론적 배경에서 상황판단검사의 구성개념에 대한 상이한 견해들을 소개하였다. 이와 같은 견해들 가운데 실용지능으로 보는 견해는 물론, 일반정신능력, 경험, 직무지식, 성격과 같은 여러 가지 개념들과 관련된 다차원적 구성개념이라는 견해 역시 경험적 연구를 기다리고 있다. 우리는 상황판단검사의 요인구조와 그를 통한 수행설명에 이론개발에 관심을 가지므로 그 요인구조를 실용지능이라고 해야할지, 다른 무엇으로 할지는 관심의 대상이 아니다. 따라서 상황판단검사의 구성개념에 잠정적으로 인지 및 비인지적 측면을 모두 수용하기위하여 “상황역량”으로 부르기로 한다. 여기서의 역량은 McClelland (1973)가 제시한 개념으로서 인지/비인지를 포함한다.

Sternberg와 동료들은 1981년의 초기연구에서 실용지능의 구성개념을 제시한 후 계속적으로 하나의 총점을 사용해왔으므로 다차원적 성질을 경험적으로 보이지는 못했다. Sternberg와 동료들이 자신들의 실용지능 척도에서 다차원적 구조가 아닌 단일차원만이 도출된 것은 요인분석의 방법상 문제였다. 즉, 그들의 척도에서 제시된 상황 시나리오의 효과를 배제한 후의 요인분석을 하지 않고, 상이한 시나리오에 속한 여러 문항들을 모두 합해서 요인분석을 하다보니 “시나리오”라는 측정형태 즉, 방법의 효과가 내용을 압도한 것으로 본다. 따라서 지필형 상황판단검사의 결과에서 시나리오의 효과를 배제하고 요인분석을 하는 3단계 접근을 한다면 다차원 구조를 얻을 것이다(Lee & Kim, 2002). 그리고

McDaniel등(2003)에서 지식형 질문이 인지검사에, 행동성향 질문이 비인지검사에 유의한 상관을 보이는 것을 감안할 때, 상황역량은 크게 지식요인과 행동성향요인으로 나누어 볼 수 있을 것이다.

그러면, 지필형 상황판단검사에 대한 요인분석을 통해 확인된 구성요인들이 여타변수들과 어떤 관련을 가지는지 연구문제를 설정하기로 한다.

연구문제 1: 지필형 상황판단검사의 지식요인은 성격이나 동기에 비해 다른 인지개념들(일반정신능력, 직무지식)과 더 큰 관계를 가질 것이다.

연구문제 2: 지필형 상황판단검사의 행동성향요인은 일반정신능력이나 직무지식에 비해 성격과 유의한 관계를 가질 것이다.

기존의 예측변수들과 준거들간 관계속에서 상황역량의 이론적 위치 탐색

예측변수로서 일반정신능력, 상황역량, 및 직무지식과 준거변수로서의 업무성과 및 회사적응간의 관계에 대한 연구문제를 제시하고자 한다. 산업심리학에서는 일반정신능력이 업무성과에 대한 최상의 예측변수라고 보고 있다(예: Ree & Earlse, 1992). 예를 들어, Hunter(1983)가 제시한 상사 또는 동료 평정에 대한 과정보형에 따르면 일반정신능력이 직무 지식의 획득에 영향을 미쳐서 결국 직무 수행에 대하여 상당한 정도로 예측 타당도가 있음을 밝히고 있다. 또한 Ree, Carretta, 및 Teachout(1995)의 연구에서는 훈련과정중의 일반정신능력과 사전 직무 지식, 그리고 훈련중에 획득한 직무지식과 수행과의 관계를

살펴본 결과 일반정신능력이 직무지식의 획득에 영향을 미쳐서 결국 직무 수행에 유의한 영향을 미치고 있음을 밝히고 있다. 그런데 직무상 지식은 Borman, White, 및 Dorsey(1995)가 말하듯이 과업지식(technical job knowledge)과 맥락지식(contextual job knowledge)으로 나뉜다. 일반정신능력은 인지능력으로서 상대적으로 과업지식에 관계가 크고 이는 결과적으로 업무성과에 영향을 미친다. 이와 같은 논의를 통해 다음과 같은 연구문제를 설정할 수 있다.

연구문제 3: 일반정신능력은 과업지식을 향상시키고, 과업지식은 과업수행에 영향을 미친다. 즉, 과업지식의 획득 없이는 일반정신능력 자체로서 과업수행에 대한 예측이 불가능하다.

적용범위로서 일반정신능력은 다양한 학습과제에 일반적 적용이 가능하고 상황판단은 영역 범주에 속하는 것으로 구분할 수 있다. 일반정신능력은 광범위한 영역에 걸치므로 특정 영역에서의 예측력은 그 영역내에서 충분한 분산을 가지는 다른 변수에 의해 보완되어야 한다. 상황판단은 특정의 문화와 가치로 정의되는 구체적 영역중심적 특징을 가지므로 일정영역내에서의 적용에서는 일반정신능력을 상당정도 보완하는 중요한 예측변수가 될 것이다.

업무수행에 대하여 과업수행과 맥락수행으로 구분한 Borman 등(1993)의 견해를 보면, 구체적인 업무에서의 숙련 및 실적을 나타내는 과업수행에 비해, 사회적 및 동기적인 맥락을 지원하는 대인간 행동의 집합으로서 정의되는 맥락수행은 과업이외에 대인간 촉진 및 직무현신의 요소로 구성되는 환경적응적인 수행이다. 따라서 조직적응을 준거변수로 보았을 때 일반정신능력과 상황역량에 대한 다음과 같은 연구문제를 설정할

수 있을 것이다.

연구문제 4: 조직적응과 같은 맥락수행을 준거로 할 때 상황역량은 일반정신능력에 비하여 더 큰 관계를 가질 것이다.

연구문제 5: 상황역량은 맥락지식의 형성에 영향을 미치고 맥락지식은 조직적응과 같은 맥락수행에 영향을 미칠 것이다.

Hunter와 Hunter(1984)의 메타분석결과, 직무 수행과 작업 경험간의 상관인 0.18로 나타나 성공적인 업무성과를 위해 작업경험이 영향을 미치는 것으로 나타났다. 또한 McDaniel, Schmidt, 및 Hunter(1998)의 연구에서는 다양한 직업들을 통털어 작업 경험과 직무 수행간에 수정된 평균 상관이 0.32로 보고되었다. 그리고 Schmidt, Hunter, 및 Outerbridge(1986)는 작업 경험이 직무 수행의 예측에 미치는 영향을 경로분석(path analysis)을 통해 보다 이론적인 가설들을 검증해 본 결과 유의한 관련성을 보인 연구결과를 보고하였다. 그런데 Schmidt 등(1986)의 연구에서 살펴보았을 때 경험의 주요 효과는 업무 수행에 대한 직접 효과가 아닌 직무 지식의 획득 또는 학습이었다. 따라서 경험의 증가는 직무지식에 대한 획득 또는 학습의 증가를 가져와 업무 수행에 영향을 준다고 말할 수 있다. 경험의 증가는 상황에의 참여를 의미하고 상황역량의 증가를 가져온다. 그렇다면 이 상황역량이야말로 경험과 직무지식간의 중개변수가 될 수 있다. 즉 “경험→상황역량→직무지식”으로 생각할 수 있다. 이것은 일반정신능력이 직무지식을 매개로 직무 수행에 간접효과를 미친다고 말하는 Hunter(1983), Borman 등(1995), 그리고 Ree 등(1995)의 연구결과와 평행하다.

연구문제 6: 경험은 상황역량의 형성에 직접적 영향을 미칠 것이다.

연구문제 7: 상황역량은 직무지식 즉, 과업지식 및 맥락지식의 형성을 통해서 과업수행과 조직적응에 영향을 미칠 것이다.

성격은 맥락수행 요인이다(Borman & Motowidlo, 1993). 조직에서의 맥락수행은 과업활동과 과업 과정에 중요한 촉매로서 작용하여 조직적, 사회적, 그리고 심리적 맥락을 형성하고 조직 효과성에 이바지한다(Borman et al, 1995). Barrick, Mount, 및 Strauss(1993)의 연구결과에 의하면 성격 중 성실성의 경우 업무 수행과 큰 상관($r=0.22$)을 보여 주고 있다.

한편, Hurtz와 Donovan(2000)은 성격과 업무성과와의 관계에 대한 연구에서 여러 직업 범주에 걸쳐 성격차원에 대하여 메타분석한 결과의 평균 타당도는 크기로 볼 때 성실성(0.22), 정서안정성(0.14), 호감성(0.13), 외향성(0.10), 그리고 개방성(0.07)의 순서였다. 이를 통해 볼 때 성격차원중 성실성의 준거타당도가 가장 높다고 볼 수 있다.

또한 Barrick과 Mount(1993)의 성격모형에 의하면 앞서 언급한 바와 같이 성실성은 다양한 직업에 걸쳐 준거에 대한 타당한 예측변수임이 밝혀졌으며, 외향성은 관리자들과 판매원들과 같은 사회적 계약과 관련된 직업을 포함하는 다양한 준거에 대한 타당한 예측변수라고 말하고 있다. 또한 지적 개방성은 훈련 숙달에 타당한 예측변수이고, 호감성은 메타분석을 통해 업무수행에 대하여 유의한 예측변수임이 밝혀졌다. 그런데 이렇게 유의한 타당도의 이면에 감추어진 구조를 가설화하는 것이 이론개발에 의미있다. 즉, 성격이 업무수행에 기여하는 것은 직무지식을

통해서라고 할때 성격은 무엇을 통해서 직무지식의 향상을 가져올것인가? 성격은 환경이라는 폭넓은 상황에서 획득되는 측면이 크다는 것을 감안할 때, 일반정신능력보다는 상황역량의 형성에 기여할 것이다. 즉, 환경적응적인 요소로서의 성격은 역시 환경적응적인 상황역량의 향상에 이바지하고 그로 인하여 맥락지식의 획득을 촉진한다고 말할 수 있다. 따라서 다음과 같은 연구문제가 도출된다.

연구문제 8: 성격은 상황역량의 향상을 통해 맥락지식의 획득을 촉진한다.

업무수행에 대한 타당한 또 한 가지 예측변수로서 개인사(biographical data) 검사를 꼽을 수 있다(Rothstein, Schmidt, Erwin, Owens, & Sparks, 1990). 예를 들어, Hunter와 Hunter(1984)는 정량적인 문헌검토를 통해 업무수행에 대한 일반정신능력검사의 준거타당도가 0.47이라고 한 반면, 동일한 준거에 대하여 개인사 검사의 타당도는 0.37이라고 보고하였다. 또한 Reilly와 Chao(1982)는 개인사 검사의 평균적인 준거타당도가 0.35라고 하였으며, Asher(1972)는 개인사 검사에 대한 준거타당도가 밝혀진 문헌들의 90%가 0.30이상이라고 밝히고 있다. 이를 통해 볼 때 개인사 검사는 업무수행에 대하여 타당한 예측변수라고 말할 수 있다.

그러나 일반정신능력검사가 영역 일반적인 특성을 보이며 일반화가 가능한 반면, 개인사 검사는 상황적으로 특수한 형태를 보여 상이한 영역에 대하여 일반화하기가 어렵다(Rothstein et al, 1990). 예를 들어 Hunter 등(1984)은 개인사 검사가 상황 제한적이라는 견해를 밝히고 있으며, Thayer(1977)는 개인사 검사에 대한 타당도는 연령, 조직 체계, 준거의 적용, 일시적인 해당 직무

에 있어서의 변화 등에 따라 상이하다는 견해를 보이고 있다. 따라서 개인사 검사는 영역중심적 또는 영역특수적인 형태를 가지며 특히 판매직의 업무수행 예측에 많이 이용되고 있다(Vincur, Schippmann, Switzer, & Roth, 1998). 그 이유로서는 판매직의 경우 우선 준거가 명확하다는 점, 그리고 판매 업무 수행시 감독자가 없으므로 판매사원의 자율성을 발휘해야 하고 거절 경험을 많이 겪으므로 과거 자신의 경험에 의지해 상황을 판단해 나가야 하는 능력이 요구되며 이는 개인사 검사에 의해 예측이 될 수 있기 때문이다.

연구문제 9: 개인사변수들은 특정 영역에서 (예: 판매직) 상황역량의 획득을 예측해주고 그에 따른 직무지식의 축진을 통해서 과업 및 맥락수행에 기여한다.

이상과 같은 연구문제1에서 연구문제9까지를 그림 1로 표현하면 아래와 같다.

동기변수는 지금까지 심리학분야에서 직무수행에 가장 가까운 근접변수로서 논의되어왔다(예: Kanfer, 1990). 즉 모든 류의 개인적 자원은 동기화 없이는 활용될 수도 없고, 업무성과를 가져올 수도 없을 만큼 중요한 조절변수인 것이다. 예를

들어 Schmidt 와 Hunter(1992)의 연구는 정보처리론적 입장에서 성격과 동기를 직무수행에 가까운 근접변수로서 설정한 결과를 보이고 있다. 이와 같은 동기변수는 훈련 장면에서의 학습과 관련하여 많은 연구가 존재한다. 예를 들어 Quinones(1995)의 연구에서는 훈련동기부여가 학습에 정적인 영향이 있음을 보여주었다. 또한 Colquitt와 Simmering(1998)은 훈련 중 동기부여와 학습간에 직접적이고 정적인 관계가 있음을 보고하였다. 지식획득에 동기화가 조절변수로 작용함을 보이는 것이다. 이러한 연구에서 지식은 주로 과업지식을 중심으로 한다. 그리고 Kanfer (1990)는 복잡한 과업지식 및 스킬의 획득에는 지속적인 주의와 연습이 요구된다는 점을 명시하며 내적동기부여를 통한 동기유지의 개념을 강조하였다. 이를 통해 볼 때 동기변수는 과업지식의 획득 및 과업수행에 조절변수 역할을 하는 것으로 가설화된다. 따라서 다음과 같은 연구문제의 설정이 가능하다.

연구문제 10: 노력 투입의 강도와 방향성을 나타내는 동기화는 과업지식획득을 촉진하며, 한편 노력투입을 꾸준히 강도있게 유지하는 동기유지는 과업 수행을 촉진할 것이다.

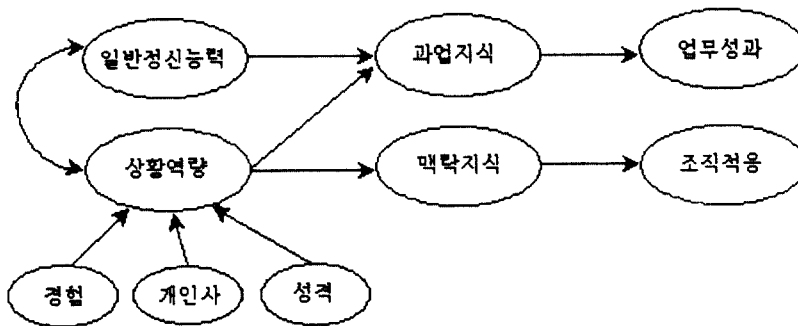


그림 1.

한편, 맥락수행에 관련된 변수로서 조직몰입을 생각해볼수 있다. 조직몰입과 직무수행간 관련성에 대한 연구들은 정적인 관련성을 보고하고 있다(Mowday, Steers, & Porter, 1979). 한편, 조직몰입은 대인간 촉진과 직무헌신이라는 두 개의 요소로 이루어진다. 전자는 협력적이고 동료를 이해하며 도움을 주는 요소이며 후자는 조직목표를 지원하기 위해 규율을 따르고 솔선수범하며 열심히 일하는 행위이다. 이와 같은 조직몰입은 과업중심적이기 보다는 환경적응적이고 조직맥락 중심적인 특성을 가진다. 따라서 조직몰입은 업무수행에 대하여 과업수행과 맥락수행으로 구분한 Borman 등(1993)의 견해에 따라 맥락지식 및 맥락수행에 영향을 미칠 것으로 예상된다.

또한 종업원들이 볼 때 조직의 지원이 있다고 지각할 경우 조직에 대한 심리적 애착은 증가하며, 이는 조직몰입을 가져오게 되거나(Gaertner & Nollen, 1989), 조직충성도, 조직규정준수 등의 맥락수행을 촉진하게 된다. 즉, 지각된 조직지원은 맥락지식의 활용으로 맥락수행이 이루어지는 과정에 영향을 미칠 것으로 예상된다.

한편 과업수행에 지원적인 과업환경은 과업지식이 수행에 활용되는 과정에서 조절효과를 가져오고, 마찬가지로 종업원이 지각하기에 조직의 지원이 있다고 할 때 맥락지식은 조직적응과 같은 맥락적 수행에 활용될 것이다.

연구문제 11: 조직몰입의 높고 낮음에 따라 맥락지식획득 및 맥락수행의 과정에 촉진 또는 부진을 초래할 것이다.

연구문제 12: 과업환경은 “과업지식 및 스킬 → 과업수행”의 과정에서 조절변수 역할을 한다.

연구문제 13: 지각된 조직지원은 “맥락지식 및 스킬 → 맥락수행”의 과정에서 조절변수 역할을 한다.

이상의 연구문제들을 종합하면 그림 2와 같다. 그림 1과 차별을 두기 위하여 연구문제 10에서부터 연구문제 13까지에 해당하는 모형은 점선으로 표시하였다.

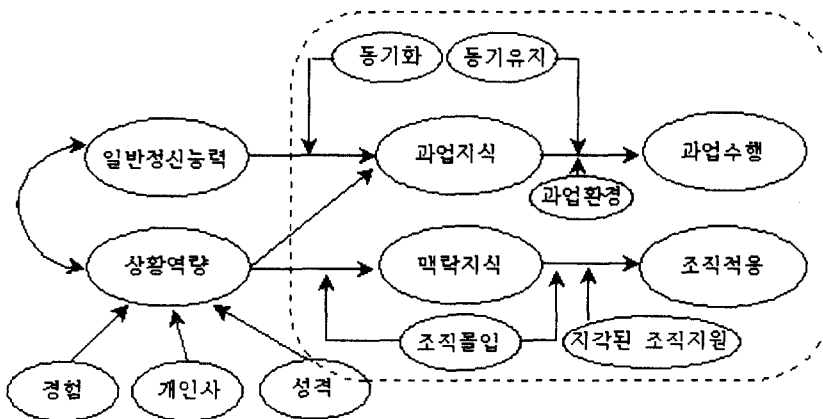


그림 2.

토론 및 결론

이 연구에서는 지필형 상황판단검사의 오랜 역사에 비추어 연구가 미흡한 점들에 대한 지적 및 앞으로의 연구방향을 제시하였다. 연구상 미흡한 점들은 일반정신능력에 대비한 증분타당도의 연구, 그리고 영역별로 개념화되어야 할 구성개념이라고 할 수 있다. 이러한 문제를 해결하기 위해서 가장 시급한 것은 인식의 전환이다. 즉, 전통적인 정보처리이론에 관점을 둔 일반정신능력검사와 달리, 상황판단검사는 사회문화적 혁명 또는 상황인지이론에 기반을 두므로 검사에 응답하는 사람들이 살고있는 삶의 현장에서의 가치가 반영되는 맥락의존적 검사로서 개념화되어야 하는 것이다. 이제 1970년대 80년대에 정착된 타당도일반화의 영향에서 벗어날 때임을 의미한다. 60년대 70년대의 인지혁명속에서 정보처리이론이 부각되면서 인지역량의 여러 영역에 걸친 일반화가 추세였고 탈맥락적인 일반정신능력이 모든 예측변수 가운데 제1의 자리를 굳혔으며 모든 다른 변수들은 일단 일반정신능력이 예측방정식에서 분석된 후에 추가되면서 타당도의 증분이 얼마인지를 보는 것이 관행이 되었다 (Schmidt & Hunter, 1998). 그러나 일반정신능력이 폭넓은 예측변수가 되는 대신에 어떤 영역에 대해서는, 그 영역에서 개인들간의 차별화를 잘 보여주는 다른 변수가 있을 수 있음을 간과한 것도 사실이다(Cascio, 1998). 이제 사회문화적 혁명속에서 개인이 처한 활동체계에서의 개인역할이 중요해지고 '머리속에 무엇이 들어있느냐'보다는 '너의 머리가 어떤 환경에 들어있느냐'에도 주의를 돌리는 상황인지의 관점은 개인차와 문화차이를 인정하지 않던 정보처리이론을 벗어나 영역간 차이를 인정하고 상황과의 상호작용속에서 인지역량을 정의하는 관점을 추가하게 되었다.

상황판단검사는 이러한 추세에 부응하는 검사중 한가지로서 객관적 정답보다는 맥락의존적 정답을 가지며, 구성개념 역시 영역간에 일반화되기 보다는 영역내 구성개념을 가지게 될 것이다.

그런데 이러한 상황판단검사의 실제 제작에 있어서 직무관련성에 대한 수준 조절의 필요성이 제기된다. 즉, 지필형 상황판단검사가 성취도검사가 아닌 신입사원 선발을 위한 적성검사일 경우 직무관련성의 수준이 낮아야만 검사의 목적이 달성될 것이기 때문이다. 그러나 직무관련성의 수준 결정에는 주관적 판단을 배제할 수 없기 때문에 주제 전문가와의 의견 조율이 요구된다. 또한 전문지식이 아닌 기본적 실용지식의 문체가 차지하는 정도에 따라서 상황판단검사의 요인들은 그 분야의 실용지능이라고 할 수 있을 것이다. 또한 종래에 사용되어 온 메타분석 역시 타당도를 일반화하는 범주에 어느 정도의 제약을 허용해야 할 것이다. 일반정신능력에 적용되는 넓은 범주의 일반화보다는 영역중심으로 제한되는 좁은 범주의 일반화 개념이 필요하게 될 것이다.

영역중심이라고 할때의 영역의 정의는 크게 두 가지로 접근할 수 있다. 첫째는 전통적인 명시적 개념화의 방식에서는 필요충분한 조건들을 제시하고 이를 충족할 때 정의에 맞는다고 한다. 예로서 의자를 정의할 때 평평한 바닥이 있고, 등받이고 있으며, 네 개 또는 세 개의 다리가 있어 사람이 앉을수 있게 되어 있는 사물을 의자라고 한다. 그러나 한편으로는 필요충분조건 없이 원형성 방식(Rosch, 1978)으로도 가능하다. 우리가 의자라고 할 때 관찰되는 것은 앞서 정의한 원형에 아주 가까운 것에서부터 주변적인(marginal) 것에 이르기 까지 다양하다. 버스정류장 같은 경우 일반적인 의자는 파손되기 쉬우므로 직경 10~20cm 정도의 쇠파이프를



으로 바닥에 고정시켜놓아 꼭 앉아야할 사람만 걸터 앉게 되어 있다. 소극장에 가보면 관객들이 앉을 의자는 계단에 방식같은 것을 깔아놓은 정도일 뿐이고, 별장에 가보면 안락한 분위기에서 쿠이 들어있는 자루가 의자로 쓰이기도 한다. 의자의 범주에는 이렇게 다양한 형태의 '앉을 것'이 포함된다. 그렇다면 의자는 정의하기에 필요 충분한 내적인 특질이 있다기보다 단지 원형에 유사한 정도에 따라 의자로 불린다고 볼 수 있다.

Neisser(1979)는 이런 정의 방식을 '로쉬방식'이라고 하면서 지능의 정의에 응용하였다. 즉, '지능적인 사람'이라고 할 때 우리는 여러 차원에서 원형적인 모습을 떠올리고, 어떤 개인들이 이 원형에 유사하면(similarity) 지능적인 사람이라고 한다는 것이다. 물론 유사성의 정도는 원형이 가지고 있는 각 차원별로 높은데서 낮은데까지 다양할 것이다.

현대의 인지심리학에서 개념에 대한 개념화의 방식으로서 로쉬방식이 나온 배경에는, 하나의 개념화에 여러 가지 특성(feature)이 포함되며 각 특성들은 분포를 가지고, 분포의 가장자리에서 다른 개념들과 경계가 애매하게 되는 현실이 반영된 것으로 본다. 산업장면에서 영역이라고 할 때 지금껏 상황판단검사가 적용된 예들은 군대, 심리학자, 관리자, 경영관리영역 등으로서 특별이 정의없이 사용되어왔다. 즉, 특별히 영역을 어떻게 범주화하자고 한적은 없다. 명시적으로 정의하는 전통적 접근으로 하자면 표준직업분류 사전에서의 대분류, 중분류, 또는 소분류로 내려가면서 각 분류내의 직업집단들을 포괄 정의하는 기술문들로 그 영역을 정의하게 될 것이다. 이러한 접근이 필요하기도 하지만, 아직까지는 상황판단검사가 제작되는 해당 영역의 의미를 로쉬방식으로 개념화하였다고 볼수가 있다. 또한

아직은 상황판단검사가 충분히 많은 영역에서 제작된 것이 아니므로 명시적인 영역분류의 필요성이 강하게 제기되고 있지 않다고 볼수도 있다.

또 하나의 추후연구문제로서는 지필형 상황판단검사 제작시 정답선별의 정확성이다. 상황판단검사는 시나리오점수 방식, 원형성 방식 모두 경험적 해답을 기준으로 점수산정이 이루어진다. 그러나 영역일반적인 정답이 존재하지 않으므로 영역중심적인 경험적 해답의 정확성에 있어 논란이 발생할 수 있다. 원칙적으로는 주제 전문가들간에 분산이 적은 문항에서의 평균을 경험적 해답으로 해야 하겠지만, 어느 정도까지를 수렴하는 의견으로 보려는 영역내 주관성의 범위를 벗어날 수가 없다.

본 논문에서는 구성개념 도출을 위해 문항을 채점단위로 하는 점수방식을 소개하였다. 그러나 종래에 산업심리학에서 준거타당화의 용도로 사용되어 온 시나리오점수 방식 역시 보다 연구될 필요가 있다. 따라서 준거타당도를 위한 시나리오점수방식 사용시, 요구되는 반응에 따른 타당도 연구의 필요성이 제기된다. 보통은 최적/최악의 대안을 선택하게 해서 시나리오에 대한 점수를 매기지만 적용 영역의 특성상 최적(best), 최악(worst), 최적/최악(best/worst)중 어느 것을 요구하느냐에 따라 준거타당도의 차이가 발생할 수 있다. 즉, 판매영역(sales domain)의 경우 최적 대안을 선택하기보다는 최악 대안을 회피하는 것이 더 중요할 수 있다. 예를 들어 까다로운 고객과의 부적절한 대화의 경우 고객에게는 깊이 각인되며 이에 따라 고객 인적 네트워크에는 심한 타격을 받을 수 있기 때문이다. 따라서 판매영역에서는 피해야 할 최악방식이 무엇인지 알고 있는지를 테스트하는 것이 준거타당도 연구에 적절할 수 있음을 시사하고 있다.

참고문헌

- 이순목, 김아영, 조영미. (2000). 목시지 요인추출을 위한 2단계 방식: 고등학생자료에 응용. *한국심리학회지: 일반*, 19(2), 123-149.
- Barrick, M. R., Mount, M. A., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: test of the mediating effects of goal setting. *Journal of Applied Psychology*, 78, 715-722.
- Bauer, T. N., Maertz, C. P., Jr., Dolen, M. R., & Campion, M. A. (1998). Longitudinal assessment of applicant reactions to employment testing and test outcome feedback. *Journal of Applied Psychology*, 83, 892-903.
- Borman, W. C., Hanson, M. A., Oppler, S. H., Pulakos, E. D., & White, L. A. (1993). Role of early supervisory experience in supervisor performance. *Journal of Applied Psychology*, 78, 443-449.
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. Borman(Eds.), *Personnel selection in organizations* (pp. 71-98). New York: Jossey-Bass.
- Borman, W. C., White, L. A., & Dorsey, D. W. (1995). Effects of ratee task performance and interpersonal factors on supervisor and peer performance ratings. *Journal of Applied Psychology*, 80(1), 168-177.
- Bruce, M. M., & Learner, D. B. (1958). A supervisory practices test. *Personnel Psychology*, 11, 207-216.
- Campbell, J. P., Dunnette, M. D., Lawler, E. E., & Weick, K. E. (1970). *Managerial behavior, performance and effectiveness*. New York: McGraw-Hill.
- Candall, A. J. (1942). *Preliminary manual for the Test of Practical Judgment*. Chicago: Science Research.
- Cascio, W. N. (1998). *Applied psychology in human resource management*. NJ: Prentice-Hall.
- Carrington, D. H. (1949). Note on the Cardall Practical Judgment Test. *Journal of Applied Psychology*, 33, 29-30.
- Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, 82, 143-159.
- Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: the relationship between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology*, 82, 300-310.
- Clevenger, J., Pereira, G. M., Wiechmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology*, 86(3), 410-417.
- Corts, D. B. (1980). *Development and validation of a test for the ranking of applicants for promotion to first-line federal trades and labor supervisory positions*(Report No. PRR-80-30). Washington, DC: U.S. Office of Personnel Management, Personnel Research and Development Center.
- Colquitt, J. A., & Simmering, M. J. (1998). Conscientiousness, goal orientation, and motivation to learning during the learning process. *Journal of Applied Psychology*, 83(4), 654-665.
- Costa, P. T., Jr., & McRae, R. R. (1992). *Revised*

- NEO Personality Inventory(NEO-PI-R) and NEO Five Factor Inventory(FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.
- Cronbach, L. J. (1988). *Test validity.* NJ: Lawrence Erlbaum Associates.
- Cronbach, L. J. (1990). *Essentials of psychological testing.* NY: Harper & Row.
- File, Q. W. (1945). The measurement of supervisory quality in industry. *Journal of Applied Psychology*, 29, 381-387.
- File, Q. W., & Remmers, H. H. (1948). *How Supervise? manual 1948 revision.* New York: Psychological Corporation.
- Gaertner, K., & Nollen, S. (1989). Career experiences, perceptions of employment practices, and psychological commitment to organization. *Human Relations*, 42, 975-991.
- Horvath, J. A., Sternberg, R. J., Forsythe, G. B., Sweeney, P. J., Bullis, R. C., Williams, W. M., & Dennis, M. (1996). Tacit Knowledge in Military Leadership: Supporting Instrument Development, U. S. Army Technical Report 1042. U. S. Army Research Institute for the Behavioral and Social Sciences.
- Hunter, J. E. (1983). A causal analysis of cognitive ability, job knowledge, job performance. *Journal of Vocational Behavior*, 29, 340-362.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: the big five revisited. *Journal of Applied Psychology*, 85(6), 869-879.
- Kanfer, R. (1990). *Motivation theory and industrial organizational psychology.* In *handbook of industrial and organizational psychology.* Vol. 1. Ed by Dunnett, M. D. & Hough, L. M.
- Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimatrix by confirmatory factor analysis. *Psychological Bulletin*, 112, 165-172.
- Kluger, A. N., & Rothstein, H. R. (1993). The influence of selection test type on applicant reactions to employment testing. *Journal of Business and Psychology*, 76, 698-707.
- Landy, F. J., & Shankster, L. J. (1994). Personnel selection and placement. *Annual Review of Psychology*, 45, 261-297.
- Lee, S. & Kim, A. (2002). A Three-Step Approach to Factor Analysis on Data of Multiple Testlets. In *Measurement and Multivariate Analysis.* S. Nishisato, Y. Baba, H. Bozdogan, & K. Kanefuji(eds.). Tokyo, Japan: Springer-Verlag.
- March, H. W. (1989). Confirmatory factor analysis of multitrait-multidata: many problem a few solutions. *Applied Psychological Measurement*, 9, 1-22.
- McClelland, D. C. (1973). Testing for competence rather than for "Intelligence". *American Psychologist*, Jan, 1-14.
- McDaniel, M. A., Finnegan, E. B., Morgeson, F. P., Campion, M. A., & Braverman, E. P. (1997). Predicting job performance from common sense. Paper presented at the 12th Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: a clarification of the literature.

- Journal of Applied Psychology*, 86(4), 730-740.
- McDaniel, M. A., Schmidt, F. L., & Hunter, J. E. (1988). Job experience correlates of job performance. *Journal of Applied Psychology*, 73, 327-330.
- McDaniel, M. A., Hartman, N. S., & Grubb III, W. L. (2003). Situational Judgement Tests, Knowledge, Behavioral Tendency, and Validity: A Meta-Analysis. Paper presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology. Orlando, FL. April, 2003.
- Millard, K. A. (1952). Is how supervise? an intelligence test? *Journal of Applied Psychology*, 36, 221-224.
- Moss, F. A. (1926). Do you know how to get along with people? Why some people get ahead in the world while others do not. *Scientific American*, 135, 26-27.
- Mowday, R., Steers, R., & Porter, L. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior*, 14, 224-247.
- Mowry, H. W. (1957). A measure of supervisory quality. *Journal of Applied Psychology*, 41, 405-408.
- Motowidlo, S. J., & Tippins, N. (1993). Further studies of the low-fidelity simulation in the form of a situational inventory. *Journal of Occupational and Organizational Psychology*, 66, 337-344.
- Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: the low-fidelity simulation. *Journal of Applied Psychology*, 75(6), 640-647.
- Mullins, M. E., & Schmitt, N. (1998). *Situational judgment testing: Will the real constructs please present themselves?* Paper presented at the 13th Annual Conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.
- Neisser, U. (1979). The concept of intelligence. *Intelligence*, 3, 217-227.
- Northrop, L. C. (1989). *The psychometric history of selected ability constructs*. Washington, DC: U.S. Office of Personnel Management.
- Oswald, F. L., Schmitt, N., Ramsay, L. J., Kim, R. H., & Gillespie, M. A. (2002). Noncognitive predictors of college student success. Unpublished manuscript.
- Ployhart, R. E., & Ryan, A. M. (1998). The relative importance of procedural and distributive justice in determining applicants' reactions. *Journal of Applied Psychology*, 83, 3-16.
- Pulakos, E. D., & Schmitt, N. (1996). An evaluation of two strategies for reducing adverse impact and their effects on criterion-related validity. *Human Performance*, 9, 241-258.
- Quinones, M. A. (1995). Pretraining context effects: training assignment as a feedback. *Journal of Applied Psychology*, 80(2), 226-238.
- Quinones, M. A., Ford, J. K., Teachout, M. S. (1995). The relationship between work experience and job performance: a conceptual and meta-analytic review. *Personnel Psychology*, 48, 887-910.
- Ree, M. J., & Earles, J. A. (1993). Intelligence is the best predictor of job performance. *Current Directions in Psychology*, 1, 86-89.
- Ree, M. J., Carretta, T. R., & Teachout, M. S. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology*, 80, 721-730.
- Reilly, R. R., & Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1-62.

- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd(Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.
- Rothstein, H. R., Schmidt, F. L., Erwin, F. W., Owens, W. A., & Sparks, C. P. (1990). Biographical data in employment selection: can validities be made generalizable? *Journal of Applied Psychology*, 75(2), 175-184.
- Schmidt, F. L. (1996). The effect of hiring methods on the output and productivity of employees. *Paper presented in Seoul, Korea, June*.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection method in personnel psychology: theoretical and practical implications. *Psychological Bulletin*, 124(2), 262-275.
- Schmidt, F. L., & Hunter, J. E. (1992). Development of a causal model of process determining job performance. *Current Directions in Psychological Science*, 1(3), 89-92.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71, 432-439.
- Smith, K. C., & McDaniel, M. A. (1998). Criterion and construct validity evidence for a situational judgment measure. Paper presented at the 13th Annual Convention of the Society for Industrial and Organizational Psychology, Dallas, TX.
- Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, 46, 49-76.
- Spearman, C. (1927). *The abilities of man*. New York: Macmillan.
- Sternberg, R. J. (1985). *Beyond IQ: a triarchic theory of human intelligence*. New York: Cambridge University Press.
- Sternberg, R. J. (1997). Managerial intelligence: why IQ isn't enough. *Journal of Management*, 23(3), 475-494.
- Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology*, 41, 37-55.
- Sternberg, R. J., Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: the nature and role of tacit knowledge in work and at school. In H. Reese & J. Puckett(Eds.), *Advances in Lifespan Development*(pp.205-227). Hillsdale, NJ: Earbaum.
- Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychology*, 50, 912-926.
- Taylor, H. R. (1949). *Test of Practical Judgment*. In O. K. Buros(Ed.), *The third mental measurements yearbook*(pp. 694-695). New Brunswick, NJ:Rutgers University Press.
- Thornton III, G. C. (1992). *Assessment centers in human resource management*. NY: Addison-Wesley.
- Vinchur, A. J., Schippmann, J. S., Switzer, F. S., & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology*, 83(4), 586-597.
- Wagner, R. K. (2002). Personal communication. May 25, 2002.
- Wagner, R. K. (1997). Intelligence, training, and employment. *American Psychologist*, 52(10), 1059-1069.

- Wagner, R. K., Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: the role of tacit knowledge. *Journal of Personality and Social Psychology*, 49(2), 436-458.
- Wagner, R. K., Sternberg, R. J. (1993). *Tacit knowledge inventory for managers: user manual*. San Antonio, TX: Psychological Corporation.
- Weekley, J. A., & Jones, C. (1997). Video-based situational testing. *Personnel Psychology*, 50, 25-49.
- Weekley, J. A., & Jones, C. (1999). Further studies of situational tests. *Personnel Psychology*, 52, 679-700.
- Widaman, K. F. (1985). Hierarchical method covariance structure models for multitrait-multidata. *Applied Psychological Measurement*, 9, 1-22.
- Williams, W. M., Blythe, T., White, N., Li, J., Sternberg, R. J., & Gardner, H. (1986). *Practical Intelligence for School*. New York: Harps Collins.
- 1 차원고접수 : 2003. 8. 28.
수정원고접수 : 2003. 11. 8.
최종게재결정 : 2003. 11. 23.

Critical Review on Paper and Pencil Situational Judgement Test

Sangchul Lee

Soonmook Lee

Youngil Jo

Department of Psychology SungKyunKwan University

Situational judgement tests(SJTs) of paper and pencil type are low-fidelity simulation of work situations. These tests have some advantages over high-fidelity simulations such as assessment centers. SJTs of paper and pencil type are easy to prepare, economic, of good criterion validity, high face validity, and of less adverse effect on minority groups. However, validation studies of SJTs have been focused on criterion validity rather than construct validity. We still do not know what are measured by SJTs. Some argue that practical intelligence is measured and others argue that multidimensional attributes are measured. To make it worse, there are critiques that SJTs are valid only as measurement techniques. In order to overcome these problems, we propose domain-centered studies, three-step approach for factor analysis, and strong program of construct validation.

key words : Situational judgement test, construct validation, strong program of construct validation, three step approach, domain-centeredness study