

잠재의미분석을 활용한 성격검사문항의 의미표상과 요인구조의 비교

박 성 준 박 희 영 김 청 택[†]
한국과학기술원 전산학부 서울대학교 심리학과

본 연구는 수검자가 검사 문항을 어떻게 이해했는지를 조사하기 위해 검사문항의 의미표상을 탐구하였다. 잠재의미분석을 활용하여 성격검사문항과 성격요인의 의미표상 간 유사도를 나타내는 의미유사도 행렬을 제안하였고, 이를 기존의 탐색적 요인분석 결과와 비교하였다. 이를 위해 예비 연구에서 대학생 154명을 대상으로 제한된 맥락에서 성격의 5요인을 각각 묘사하는 지문을 수집하였고, 이를 바탕으로 5차원의 축소하여 의미공간을 구성하였다. 연구 1에서는 간편형 한국어 BFI의 요인부하량 행렬과, 예비 연구에서 구성된 의미공간에서 생성한 의미유사도 행렬을 비교하여, 두 행렬이 높은 정적 상관성이 있음을 보여주었다. 연구 2에서는 의미유사도를 기반으로 성격검사문항을 생성하고, 수검자의 반응을 수집하여 탐색적 요인분석을 통해 요인구조를 도출하여 두 행렬이 유사함을 보였다. 결론적으로 본 연구는 성격검사에 대한 수검자의 반응 없이 검사문항의 의미표상을 분석하여 구성타당도를 추론할 수 있는 방법을 제안하였고, 성격검사의 요인구조를 검사문항과 성격요인의 의미표상 간 유사도로 해석할 수 있음을 보여주었다. 이러한 결과는 성격 검사 개발에 실용적인 도움을 줄 수 있을 것이다.

주제어 : 잠재의미분석, 탐색적 요인분석, 성격의 5요인, 의미 표상, 구성타당도

[†] 교신저자: 김청택, 서울대학교 심리학과, 서울특별시 관악로1 서울대학교
연구분야: 심리측정, 인지모형
Tel: 02-880-6076, E-mail: ctkim@snu.ac.kr

서론

성격(personality)은 업무능력(Barrick & Mount, 1991), 학업능력(Noffle & Robins, 2007), 임상적 증상(Nigg, John, Blaskey, Huang-Pollock, Willicut, Hinshaw, & Pennington, 2002) 등과 관련이 깊다. 성격검사는 기업의 신입사원 공개채용 및 고등교육기관의 신입생 선발 시 인성검사 도구로써 널리 활용되고 있다.

성격검사 개발 시 구성타당도 검증을 위해 가장 널리 사용되는 전통적인 기법은 탐색적 요인 분석이다(Fabrigar, Wegener, MacCallum, & Strahan, 1999). 이는 각 문항과 대응되는 리커트(likert) 척도에 대한 수검자의 응답을 기반으로 문항 간의 공변관계에 내재한 공통요인과 특이요인을 분석한다(Tucker & MacCallum, 1997). 즉, 문항 간 수검자 응답의 상관관계를 분석하여 소수의 잠재요인과 문항 간의 관계를 요인구조를 통해서 나타낸다. 이처럼 탐색적 요인분석은 검사문항에 대한 수검자의 반응 자료를 토대로 진행된다.

Watson(1913)은 심리학을 자극과 반응 간의 관계 탐구라고 정의하였다. Tourangeau, Rips와 Rasinski(2000)는 검사문항에 응답하는 수검자의 인지과정은 이해(comprehension), 인출(retrieval), 판단(judgement), 응답(response)의 4단계로 구성됨을 주장하였다. 이에 따라 성격검사 문항과 그에 대응하는 수검자의 응답과 연결할 수 있다. 즉, 성격검사 문항에 응답할 때 수검자가 검사문항을 이해하여 문항에 대한 의미표상을 구성하고, 이를 바탕으로 관련 경험을 인출한다. 이후 앞서 구성한 의미표상과 인출한 경험을 비교하여 해당 문항이 자신의 내적 속성과 일치하는 정도를 판단하여 수검자는 최종적으로 반응을 산출한다. 탐색적 요인분석을 사용하는 전통적인 접근법에서는 성격검사를 하는 수검자가 문항에 대한 자신의 반응을 인출하는 과정에 대한 설명을 간과하고 있다. 다시 말해, 성격검사의 타당도 검사 시 전통적인 접근법은 수검자의 반응을 토대로 요인구조를 탐색하지만, 반응 산출 이전의 인지적 단계에 초점을 맞추어 검사문항을 분석하는 연구는 매우 드물다.

본 연구에서는 반응 산출 이전의 단계에 초점을 맞추어, 수검자가 검사 문항을 어떻게 이해했는지를 조사하기 위해 검사문항의 의미표상을 탐구하였다. 검사문항의 의미는 수검자가 이해한 표상으로써, 인지과정 중 첫 번째 단계에 해당한다. 수검자가 문장으로 구성된 문항의 표상을 구성하기 위해서는, 1) 문항의 논리적 구조를 명확히 하고, 2) 문항이 무엇을 묻는지 초점을 맞추는 다음, 3) 명사 혹은 고유명사의 표상을 구성한 다음, 4) 그것을 문항의 술부(predicate)와 연결 짓는다(Graesser, Singer, & Trabaso, 1994; Tourangeau et al., 2000). 이처럼 수검자가 문장으로 이루어진 문항에 대한 표상을 구성하는 과정은 다소 복잡하다. 반면 문장이 아닌 단일 어휘로 이루어진 검사문항에 대한 의미표상을 구성하는 방식은 비교적 단순하다. 일반적인 검사에서 문장은 나열된 단어(sequence)의 의미적, 구문적 조합인데 반해 단일 어휘의 의미표상이 좀 더 분석이 편리하기 때문이다.

성격검사는 사람을 형용하는 어휘를 정리함으로써 시작된 바 있다. Thurston(1934)은 1300명의 사람들이 성격특성을 기술하는데 자주 사용한 형용사 0개를 고른 후 탐색적 요인분석을 통해 5개의 독립적인 성격요인을 추출해냈다. Galton(1884)은 ‘사전에는 인간의 성격을 나타내는 1,000개의 표현이 들어있고, 그것들은 서로 어느 정도 공통된 의미를 공유하고 있다’고 언급한 바 있다. Cattell(1943)은 성격을 기술하는 형용사를 찾는 작업에 본격적으로 착수하였다. 그는 사전에서 추출한 4,500개의 성격형용사를 분석하여 서로 사교하는 많은 수의 성격요인들을 찾아냈었다. 그러나 직교회전기법을 사용했을 때 그의 성격형용사들은 다섯 개의 공통요인으로 정리되었고, 이는 후속 연구자들에 의해 반복적으로 검증되었다(McCrae & Costa, 1985). Norman(1967)은 Cattell 보다 광범위하게 사전을 검토하여 2,800개의 성격형용사를 추출하여, 탐색적 요인분석을 이용해 1,431개로 정리한 뒤 75개의 군집으로 분류하였다. Goldberg(1990)는 선행연구를 모두 포괄하여 1,710개의 성격형용사들을 탐색적 요인분석 결과 다섯 개의 성격요인으로 분류됨을 보였고, 이에 Big Five라는 이름을 붙였다. 성격의 5요인(big five)은 횡문화적 일관성을 보인다(민경환, 2002). 5요인은 다양한 문화권에서 발견되며(Hofstee, Kiers, de Raad, Goldberg, & Ostendorf, 1997), 높은 수준의 평정자간 일관성(McCrae & Costa, 1987), 그리고 높은 수준의 예측타당도(Ozer & Benet-Martinez, 2006)을 보인다.

Goldberg(1990)의 기본어휘가설(fundamental lexical hypothesis)에 따르면, 성격의 개인차는 하나의 단어(single term)로 부호화(encode)하는 것이 가능하며, 개인차를 기술하는 형용사들을 성격 5요인으로 정리될 수 있다. 따라서 어떤 사람을 기술하는데 쓰이는 어휘들의 의미표상을 Big Five 성격 차원이 포괄하며, Big Five 성격요인은 기존의 요인분석 방법을 통해 안정적으로 검증된 성격 차원이므로 의미표상과 요인구조 간 비교가 가능할 것이다.

본 연구에서는 검사문항의 의미표상에 접근하기 위해 잠재의미분석(latent semantic analysis)을 활용한다. 잠재의미분석은 문서 간 어휘들의 공통출현(co-occurrence)정보를 이용하여, 어휘 혹은 문서의 의미표상을 벡터로 표현하는 기법이다(Landauer & Dumais, 1997). 잠재의미분석은 문서 내에 등장하는 어휘 빈도를 행렬화하여 어휘-문서행렬(term-document matrix)를 구성한 다음, 출현빈도가 1이하인 어휘는 행렬에서 제외된 뒤, 행렬에 적절한 가중치를 주어 가중행렬을 구한다. 다음 단계로는 가중행렬을 특이값 분해(Singular Value Decomposition)하고 소수의 설명력이 높은 적절한 개수의 축을 이용하여 차원을 축소한다. 이 과정에서 원래의 행렬에서는 볼 수 없었던 어휘나 문서의 의미표상 간 고차상관이 드러나게 된다. 이렇게 축소된 공간은 의미공간(semantic space)이 되며, 의미공간 상에서 어휘 혹은 문서의 의미 표상은 벡터로 표현된다.

Landauer & Dumais(1997)는 잠재의미분석의 성능을 검증하기 위해 30,457개의 백과사전 표제어를 이용하여 모형을 학습시켰다. 그리고 모형으로 하여금 ETS에서 시행하는 TOEFL 동의어 찾기 문제 80문항을 풀게 하였다. 과제로 주어진 TOEFL 문항은 제시어와 선택지로 주어진 보기 단어 중, 제시어와 의미가 가장 유사한 선택지 하나를 고르는 것이었다. 모형은 제시어와 선택

지 간의 유사도를 모두 계산한 다음 가장 높은 유사도 값을 갖는 선택지를 정답으로 골랐다. 모형이 어휘 간 의미표상의 유사도 판단을 인간의 그것에 동일한 수준으로 할 수 있음을 보였다. 뿐만 아니라, 잠재의미분석은 문서 유사도 평가 과제에서도 모형은 뛰어난 결과를 보였다. TOEFL 시험 응시생들의 답안을 채점하는 여러 명의 채점자들 간 점수의 상관은 .77이었는데, 학습된 모형이 매긴 점수와 두 명의 채점자가 매긴 점수의 상관 또한 각각 .68과 .77로 나타났다. 더불어, 학습된 모형이 매긴 점수와 응시생들에게 부여된 평균 시험점수의 상관은 .77로 나타났는데, 이는 모형이 ETS 채점자와 동일한 수준의 성능을 보였음을 의미한다(Landauer, Darell, Bob, & Schreiner, 1997).

또한 잠재의미분석은 인간의 추론 과정을 매우 유사하게 모방하는 것으로 알려져 있으며, 문서의 분류(김청택, 이태현, 2002), 맥락적 의미를 고려한 동의어 판단(이태현, 2003; 이태현, 김청택, 2004), 텍스트 축약(Gong & Liu, 2001), 그리고 표절검사(지혜성, 조준희, 임희석, 2010)를 연구하기 적합하다. Kwantes, Derbentseva, Vartanian와 Marmurek(2016)은 실험자가 Big-Five 성격요인별 시나리오에 대해 작성한 에세이와 실험자가 응답한 BFI 척도 점수를 수집하여, 에세이의 의미표상이 실제 성격요인 점수와 유의미한 정적 상관이 있다는 것을 보였다. 즉, 잠재의미분석을 활용하여 작성자의 텍스트를 통해 성격요인을 예측하였다. 그러나 구체적으로 성격검사의 요인구조를 살펴보는 것은 없었다.

본 연구에서는 첫째, 잠재의미분석을 이용하여 어휘기반 성격검사문항의 의미표상을 벡터로 표현하고, 문항들의 의미표상을 다섯 개의 성격요인의 의미표상과 유사도를 측정함으로써 각 문항의 의미표상이 요인의 의미표상과 얼마나 유사한지 살펴볼 것이다. 둘째, 문항 간 의미표상의 관계를 체계적으로 분석하고 이를 요인부하량 행렬과 비교함으로써, 수검자가 문항들을 어떻게 이해하게 될지 살펴볼 것이다. 셋째, 문항들의 의미표상을 기반으로 검사를 구성하였을 때, 수검자들의 반응이 이러한 의미표상을 충실히 반영하고 있는지를 살펴볼 것이다. 이러한 연구는 수검자의 인지과정에 대한 함의를 얻을 수 있을 뿐만 아니라, 두 문항 간 의미표상과 이후의 인지과정이 유사하다면, 요인분석을 이용한 전통적인 구성타당도 검증 과정에 실용적인 도움을 줄 수 있을 것이다.

문항의 의미표상 분석들: 의미유사도 행렬

검사문항의 의미표상을 체계적으로 분석하기 위한 분석들과 절차가 아래에 제시되어 있다. 의미표상의 원자료가 되는 의미유사도행렬의 도출과정은 다음과 같다.

적절한 맥락에서 산출된 지문 수집

어휘-문서행렬의 구성하기 위해서 적절한 맥락에서 산출된 지문을 수집해야한다. 성격검사에

사용되는 어휘는 형용사로서 사람을 꾸미는 역할을 하는데, 이러한 맥락 이외에서 산출된 문서까지 잠재의미분석의 자료로 활용할 경우 다의어 표상 문제가 발생할 수 있다. 이러한 문제가 발생하지 않도록 지문의 산출 맥락을 제한해야 한다. 일반적인 성격 측정을 위한 검사문항은

나는 _____ 한 사람이다.

와 같은 형태를 가지며, 빈 칸의 형용사가 자신과 얼마나 일치하는지를 응답하게 된다. 즉, 수검자가 검사문항을 이해하는 맥락은 어떤 내적 속성을 지닌 사람을 평가하는 맥락이며, 수검자는 이러한 맥락에서 검사문항을 접하게 된다. 이처럼, 성격 측정 맥락에 걸맞은 지문을 산출하기 위해 수검자에게 적절한 지시를 제공하여 제한된 맥락에서 산출된 지문으로 어휘-문서 행렬을 구성한다. 구체적으로, 전체 지문의 수가 n 이고, 모든 지문에서 등장한 어휘의 종류의 수가 m 이라면, 이 행렬은 (m, n) 의 크기를 갖는 행렬이 되며, 행렬의 각 원소는 해당 문서에서 특정 어휘가 등장한 빈도를 나타낸다. 다음으로, 이 행렬에 적절한 가중치를 부여한다.

차원 수의 결정

다음은 어휘-문서 행렬을 특이값 분해(singular value decomposition)하고, 적절한 수의 차원 수를 결정하여 차원 축소를 수행한 뒤 의미공간을 구성하는 단계이다. 즉, 이 단계에서 의미공간의 적절한 차원 수를 결정한다. 이를 결정하는 방법은, 축소된 공간, 의미공간에 투사(projection)된 문서의 표상 벡터를 대상으로 k -평균 군집화와 같은 알고리즘을 적용하여, 문서의 특성이 특정 차원 수에서 잘 군집화되는지 평가하는 것이다. 의미공간의 적절한 차원 수는 높은 분류정확도를 보이는 차원 중에서, 오컴의 면도날(Occam's razor)에 따라 가장 적은 수가 의미공간의 적절한 차원 수로 결정한다. 직관적으로, 각 문서의 의미표상이 쉽고 명확하게 구별될 수 있다면, 저차원 벡터만으로도 충분히 문서의 표상을 성공적으로 표현할 수 있을 것이기 때문이다.

의미유사도 행렬 계산

이 행렬은 검사문항과 요인 간의 관계를 의미표상의 유사도를 이용해 나타낸 행렬이며, 앞 단계에서 구성한 적절한 수의 차원으로 이루어진 의미공간 상에서 어휘 간 유사도를 통해 계산되는 행렬이다.

요인부하량 행렬은 공통요인분석의 가정 상, 문항들이 측정하고자 하는 소수의 공통요인은 문항과 위계적 관계를 갖는 것처럼 해석되지만, 수학적으로 요인부하량 행렬의 원소 λ_{ij} 는 j 요인에 대한 문항 i 의 상관계수이다. 즉, 의미유사도 행렬에서도 이와 같이 각 원소 λ_{ij} 는 j 요인에 대한 문항 i 의 의미표상 벡터 간의 유사도이다.

유사도를 평가하는 척도(similarity measure)는 코사인 유사도를 사용했다. 이는 다차원 공간상의

벡터 간 각도의 코사인 값을 이용하여 측정된 두 벡터간의 유사한 정도를 의미한다. 계산된 값이 1에 가까운 값을 가질수록 더 높은 유사도를 의미하며, 아래와 같은 식으로 계산된다.

$$\cos(a, b) = \frac{a \circ b}{\|a\| \|b\|}$$

이 방식은 어휘의 의미표상을 비교할 때 자주 사용되는 척도이다(Finkelstein et al., 2002; Hill, Reichart, & Korhonen, 2015). 의미유사도 행렬에서 요인은 의미공간을 구성하는 축에서의 좌표로 정의되었다.

정리하면, 의미유사도 행렬의 각 행은 단어어휘로 이루어진 문항을, 각 열은 요인을 뜻하며, 각 원소는 의미공간 상에서 계산된 검사문항과 요인명의 의미표상 간 유사도이다. 요인분석에서 요인부하량 행렬은 공통요인분석의 가정 상, 문항들이 측정하고자 하는 소수의 공통요인은 문항과 위계적 관계를 갖는 것처럼 해석되지만, 수학적으로 요인부하량 행렬의 원소 λ_{ij} 는 j요인에 대한 문항 i의 상관계수이다. 본 연구가 제안하는 의미유사도 행렬에서도 이와 유사하게, 각 원소 λ_{ij} 는 j요인에 대한 문항 i의 의미표상 벡터간의 유사도이다.

따라서 의미유사도 행렬과 요인부하량 행렬은 동일한 형태를 가지므로 요인부하량 행렬과 비교가능하다. 즉 이를 이용해 검사문항의 의미표상 구조와 요인구조를 비교할 수 있다. 구성타당도 검증 시 탐색적 요인분석을 통해 계산한 요인부하량 행렬을 해석하는 것과 같이 의미유사도 행렬 또한 각 잠재요인에 대하여 높은 의미유사도를 보이는 문항들끼리 묶어서 그것의 의미표상이 요인명의 의미표상과 유사하다고 해석할 수 있게 된다.

이러한 분석들에 기반을 두고, 본 연구에서는 전통적인 탐색적 요인분석을 통해 도출한 요인부하량 행렬과 의미유사도 행렬을 비교함으로써 성격검사의 요인구조와 의미표상의 관계를 탐구하였다. 예비 연구에서는 잠재의미분석에 필요한 의미공간을 구성하고, 연구 1과 연구 2에서는 의미공간을 이용하여 아래 제시된 두 가지 질문에 답하였다.

1. 성격검사문항들의 의미표상이 요인구조와 합치될 것인가? 다시 말하면 의미유사도 행렬과 요인부하량 행렬은 유사한 형태를 보일 것인가?

2. 의미표상이 서로 유사한 문항으로 구성된 성격검사의 요인구조는 의미표상에서의 의미유사도 행렬과 유사하게 나타날 것인가?

성격검사문항 중 어떤 문항들이 동일한 요인으로 묶인다는 것은 곧 문항에 대한 수검자의 반응이 유사하다는 것을 의미하기 때문에, 이런 경우 의미표상 또한 서로 유사한지를 연구 1에서

탐구하였다. 다음으로, 이와 반대로 연구 2에서는 의미표상이 유사한 문항들에 대한 수검자의 응답 패턴이 유사한지를 탐구하였다. 두 연구를 통해, 수검자가 이해한 문항의 의미와 리커트 척도에 대한 응답 간의 관계를 살펴보고, 나아가 검사 응답 과정에서 일어나는 인지과정에 대한 함의를 얻고자 하였다.

예비 연구: 의미공간의 구성

방 법

연구대상

서울대학교 심리학과 수업 수강생 154명의 자료가 수집되었다. 실험에 참여한 수검자는 심리학과 R-point 시스템을 통해 모집되었다. 과제가 한국어 글쓰기를 요구하므로 한국어 글쓰기가 자유로운 수검자만 실험에 참여할 수 있었다.

도구

수검자들에게 성격 5요인과 그에 대한 간략한 설명을 아래와 같이 제시하였다.

성실하다 맡은 일을 책임감 있게 완수하고, 계획을 철저히 세운다.

개방적이다 다양한 경험을 하는 것을 좋아하고, 변화와 도전을 즐긴다.

친화적이다 다른 사람의 입장을 존중하고 공감하며 배려한다.

정서적으로 안정적이다 걱정이 없고 태평하며 낙천적이다.

활동적이다 매사에 열정적이고 최선을 다하며, 적극적이다.

성격 요인명에 대한 간략한 설명은 BFI-K, NEO-PI-R, NEO-FFI 의 성격요인 정의를 참고하였다(McCrae & Costa, 1989; Saucier, 2002). 다섯 개의 성격 요인명 중 외향성과 신경성은 일상적으로 더 자주 사용되는 어휘로 교체되었으며 각각 활동적이다, 정서적으로 안정적이다로 제시하였다. 이는 어휘를 교체하지 않고 원래의 명칭대로 제시할 경우, 실험 참가자들이 가지고 있는 심리학 관련 지식을 동원하여 교과서인 지문을 일률적으로 산출하는 것을 막기 위함이었다. 다음으로 수검자들은 다섯 개의 성격요인에 대하여 각각 이러한 성격을 지닌 사람은 일반적으로 어떤 특성을 지니거나 지니지 않을지, 어떤 행동을 할지를 구체적으로 100자 이상 기술하였다.

분석

154명이 산출한 성격요인에 대한 지문 중, 불성실한 응답자로 분류된 9명의 자료(동일문장반복, 분량 미달, 분량 중 공백의 과도한 사용 등)와, 지시문을 따르지 않은 1명의 자료를 합하여 총 10명의 자료가 분석에서 제외되었다. 144명이 산출한 720개의 지문 중, 데이터 기록 오류로 미기록된 지문 하나를 제외한 719개의 지문이 최종 분석 대상이 되었다.

지문은 형태소 분석을 포함한 전처리 과정을 거쳤다. 의미상 동일하나 어미의 변화 혹은 조사, 보조용언과의 결합 등으로 인해 어휘 간 구문적 차이가 있는 경우, (예) ‘활기찬’, ‘활기차’, ‘활기차다’, ‘활기’, ‘활기를’ 등) 이를 별개의 어휘로 다루지 않기 위해 형태소 분석을 수행하였다. 이를 위해 강승식(2002)의 형태소 분석기가 사용되었다. 분석기 사용 전, 사용자정의사전에 국립국어원에서 제시한 한국어 학습을 위한 표준어휘 5,780개를 추가하였고, 지문의 어휘의 기본형은 명사/복합명사/형용사/부사를 구분하여 추출하였다. 형태소 분석 이외의 전처리 과정에서는 파이썬(python)3.3 정규식(re) 모듈 등이 사용되었다. 또한 변형된 한글 모음 및 자음, 비속어 및 오타자는 어휘를 교정하였다.

전처리된 지문은 R 3.0.2를 이용해 분석되었고, lsa package v.0.73(Wlid, 2014)가 핵심적으로 사용되었다. 이를 이용하여 문서-단어 행렬(term-document matrix)을 구성하고, 적절한 가중치를 부여하였다. 문서단어 행렬에서 각 문서의 단위는 전처리된 성격 요인 기술문이다. 국소 가중치(local weight)로는 행렬의 각 원소에 1을 더한 후 자연로그함수를 취했다.

$$w_l = \ln(x_{ij} + 1)$$

여기서 x_{ij} 는 j문서에 나타난 i어휘의 출현빈도를 의미한다. 이는 각 어휘에 대한 학습곡선을 반영한 것이다(Laudaur & Dumais, 1997). 잦은 노출로 인해 학습된 어휘의 재인율은 노출 빈도와 선형적으로 비례하지 않는다는 점을 반영한 것이다. 전역 가중치(global weight)는 각 어휘마다 서로 다른 가중치를 부여하는 것으로서 엔트로피의 역수가 사용되었다. 특정 문맥에서만 나타나는 어휘에 높은 가중치를 주고, 모든 문맥에서 고르게 나타나는 어휘는 상대적으로 낮은 가중치를 주는 것이다.

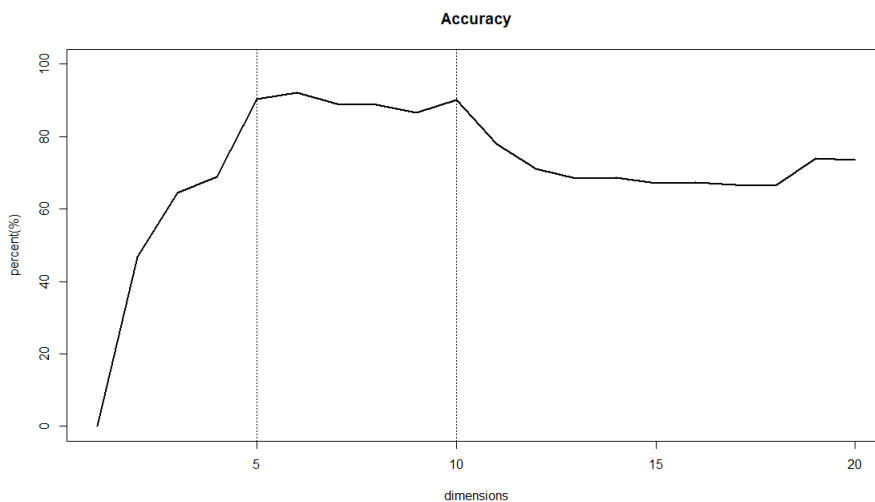
$$w_g = -\log c / \sum_j p_{ij} \log p_{ij}$$
$$p_{ij} = x_{ij} / \sum_j x_{ij}$$

여기서 c 는 전체 문서의 개수이며 $\log c$ 는 모든 어휘에 동일하게 주어지는 상수로써, 정규화(normalizing)를 위한 상수이다. p_{ij} 는 어떤 어휘가 해당 문서에서 출현할 확률이며 각 문서에서 해당 어휘가 나타난 빈도를 모든 문서에서 그것이 나타난 출현빈도의 총합으로 나눈 것이다. 전역 가중치를 이와 같은 방식으로 부여하면 특정 문맥에서만 나타나는 어휘는 해당 문맥에 대한 의미를 더 잘 반영하므로 높은 가중치를 얻고, 모든 문서에서 고르게 나타나는 대명사, 조사 등의 어휘에 대한 가중치는 줄어든다. 이외에도 여러 가중법들이 있지만, 이 방법이 다른 방법들보다 우수한 결과를 보였다(Nakov, Popova, & Matcev, 2001). 이러한 가중치가 부여된 문서단어 행렬을 특이값분해(singular value decomposition, SVD)로 분해하고, 어휘 혹은 문서의 의미표상 벡터를 구현하기 위해 의미공간을 구성하였다. 이 과정에서 유사도를 기반으로 문서가 잘 분류되는지 차원 수에 따른 분류정확도를 계산하였다. 문서 분류 기법은 k-평균 군집화 기법을 사용하였다. K-평균 군집화 기법은 주어진 자료를 k개 군집으로 묶으며 각 군집과의 거리 차이의 분산을 최소화하는 방식으로 계산된다.

연구 결과

문서단어 행렬은 2678x719차원의 행렬로 구성되었다. 이를 1~718차원으로 축소하여 의미공간을 구성한 뒤, 문서벡터간 유사도로 문서를 분류하였다.

문서의 분류정확도는 매우 적은 수의 차원으로 축소하였을 때 가장 높게 나타났으며, 이는 일반적으로 차원을 축소시키면 정보의 손실로 인해 정확도가 감소하는 경향성이 있음에도 불구하고, 자료가 소수의 차원만으로도 잘 표현되었음을 뜻한다. 정확도는 그림 1과 같이 5~10차원에



(그림 1) 성격기술문의 차원별 분류 결과 정확도 (1-20차원)

서 가장 높게 나타났으며, 이 경우 평균 분류정확도는 90%로 나타났다. 이는 충분히 높은 수준의 문서 분류정확도라고 할 수 있다. 따라서 본 연구에서는 수집자들이 산출한 성격요인 기술문은 5차원의 의미공간에 표현될 수 있다. 즉, 5차원의 의미공간 상에서 각 성격요인별로 지문들의 의미표상이 잘 분리됨을 뜻한다.

연구 1

연구 1에서는 예비 연구에서 구성한 5차원의 의미공간을 바탕으로 성격검사문항들의 의미표상이 요인구조처럼 분류될 것인지 알아보았다. 이를 위해, 기 타당화된 성격검사의 요인부하량 행렬과 의미유사도 행렬을 비교하였다.

도구

간편형 한국어 5요인 성격검사(김지현, 김복환, 하문선, 2011)의 검사문항과, 보고된 요인부하량 행렬을 이용하였다. 이 척도는 총 15문항으로 다섯 가지 성격요인을 각각 3문항씩 측정한다. 척도 타당화를 위해 성인 및 청소년 720명을 대상으로 수집한 자료를 이용하여 탐색적 요인분석을 한 결과, 고유값이 1이 넘는 요인이 5개였고, 모형적합도는 GFI=.946, CFI=.916, RMSEA=.064로 보고되었다.

분석

간편형 한국어 5요인 성격검사 검사문항 간의 의미표상을 비교하기 위해 각 문항에서 주요 어휘를 추출하였다(표 1). 이 검사가 참고한 원척도 BFI(John & Srivastava, 1999)는 각 문항이 단일 어휘로 구성되어 있으며, 본 연구 또한 검사문항으로써 어휘에 주목하였으므로 예비 연구에서 수집된 지문 속에서 최소한 두 번 이상 등장했던 어휘 중에서 추출한 어휘를 표 1에 제시하였다. 만약 두 번 이상 등장했던 어휘가 두 개 이상이었다면, 추출한 어휘는 단일 어휘로써 원문항을 가장 잘 대표하는 어휘로 선정하였다. 5번 문항의 경우, ‘능률적’이라는 어휘가 예비 연구에서 수집된 전체 지문 상에서 존재하지 않아서, 가장 의미가 유사하다고 판단된 ‘효율적’이라는 어휘로 교체하였다.

표 1의 우측 어휘를 이용하여, 연구 1에서 구성한 의미공간 상에서 검사문항에서 추출한 어휘와 요인명 간의 유사도를 계산하였다. 각 검사문항과 요인명의 의미표상 간 유사도 계산을 위해 사용된 요인명은 ‘성실성’, ‘개방성’, ‘외향성’, ‘친화성’, ‘정서적 안정성’이었다. 본 연구에서 정

〈표 1〉 간편형 한국어 BFI 검사문항 및 추출 어휘

검사문항	추출 어휘
1. 창의적인	창의
2. 독창적이며 새로운 생각을 잘 떠올리는	독창적
3. 활발한 상상력을 가진	상상력
4. 일을 완벽하게 하는	완벽히
5. 일을 능률적으로 하는	효율적
6. 믿을만하게 일을 하는	믿음직
7. 쉽게 침울해 지는	침울
8. 우울한	우울하다
9. 걱정을 많이 하는	걱정
10. 조용한	조용하다
11. 수다스러운	수다
12. 외향적이며 사교적인	외향적
13. 사려 깊고 거의 모든 사람에게 친절한	친절하다
14. 다른 사람을 잘 도와주며 이타적인	도와주다
15. 용서를 잘 하는	용서하다

서적 안정성은 신경성(neuroticism)을 의미하며, 정서적 안정성 요인의 값이 클수록 신경성이 높음을 의미한다.

각 검사문항을 대표하는 어휘와 성격요인간의 유사도를 계산하여 의미유사도 행렬을 도출한 다음, 간편형 한국어 BFI척도의 요인부하량 행렬과 이를 비교하였다.

연구 결과

표 2는 김지현, 김복환 & 하문선, 2011의 간편형 한국어 BFI척도의 요인부하량 행렬이다. 이는 최대우도법(maximum likelihood)을 이용하여 공통요인분석을 수행하였으며, 사각회전(oblique rotation)법을 사용하였다. 표 3은 검사문항의 의미유사도 행렬이다. 이 행렬의 각 원소는 각 문항과 요인명에 대응되는 의미표상 벡터 간의 코사인 유사도이다.

표 3을 통해 의미표상의 유사성 또한 요인부하량 행렬의 요인구조와 매우 유사한 것을 확인할 수 있다. 15개 검사문항에서 추출한 어휘 모두 측정하고자 하는 성격요인명과 가장 높은 의미유사도를 나타냈다. ‘상상력’은 개방성뿐만 아니라 외향성 요인과도 높은 의미유사도를 보였다.

〈표 2〉 간편형 한국어 BFI척도의 요인부하량 행렬

	개방성	성실성	외향성	친화성	정서적 안정성*
1. 창의적인	0.87	0.29	0.25	0.31	-0.16
2. 독창적이며 새로운 생각을 잘 떠올리는	0.79	0.23	0.27	0.26	-0.10
3. 활발한 상상력을 가진	0.68	0.15	0.35	0.27	-0.02
4. 일을 완벽하게 하는	0.14	0.71	-0.10	0.18	-0.10
5. 일을 능률적으로 하는	0.31	0.70	-0.05	0.35	-0.19
6. 믿을만하게 일을 하는	0.24	0.66	-0.07	0.25	-0.13
7. 쉽게 침울해 지는	-0.08	-0.12	-0.16	-0.10	0.79
8. 우울한	-0.15	-0.15	-0.21	-0.16	0.74
9. 걱정을 많이 하는	-0.04	-0.08	-0.02	0.04	0.53
10. 조용한	0.20	-0.19	0.73	0.04	-0.20
11. 수다스러운	0.26	-0.17	0.71	0.09	0.03
12. 외향적이며 사교적인	0.34	0.21	0.55	0.29	-0.24
13. 사려깊고 거의 모든 사람에게 친절한	0.16	0.34	0.03	0.68	-0.04
14. 다른 사람을 잘 도와주며 이타적인	0.27	0.34	0.08	0.55	-0.05
15. 용서를 잘 하는	0.21	0.02	0.16	0.48	-0.06

며, ‘조용하다’는 외향성 요인뿐만 아니라 친화성 요인과의 다소 높은 의미유사도를 보였다.

검사문항이 특정 성격요인에 대해 가장 높은 부하량을 보일 경우, 그것의 의미표상 또한 해당 성격요인의 의미표상과 가장 높은 의미유사도를 보였다. 예를 들면, 문항 ‘창의적인’의 경우, 개방성 요인에 대해 가장 높은 요인부하량을 보였고 나머지 성격요인에 대해서는 낮은 값을 가졌으나 해당 문항의 의미표상은 ‘개방성’과 가장 높은 의미유사도를 보였고 나머지 성격요인과는 다소 낮은 의미유사도를 보였다.

표 4는 각 성격요인별로 표 2의 요인부하량과 표 3의 의미유사도의 상관을 계산한 결과이다. 표 4에는 각 성격요인별 요인부하량과 의미유사도 간의 상관계수를 제시하였다. 전반적으로 각 성격요인에 높은 요인부하량을 보인 문항들은 그에 대응되는 요인명과 높은 의미유사도를 보였다. 둘 간의 피어슨 상관계수를 계산한 결과, 개방성($r=.900, p<.01$), 성실성($r=.784, p<.01$), 외향성($r=.808, p<.01$), 친화성($r=.536, p<.01$), 정서적 안정성($r=.886, p<.01$)으로 나타났고, 검사문항의 의미표상과 요인구조의 유사도 상관계수의 5요인 평균은 0.784로 높은 정적 상관을 보였다. 이러한 결과는 성격의 5요인의 요인부하량 행렬과 의미유사도 행렬이 매우 유사한 패턴을 보이며,

〈표 3〉 간편형 한국어 BFI척도 검사문항의 의미유사도 행렬

	개방성	성실성	외향성	친화성	정서적 안정성
1. 창의	0.995	-0.014	0.357	0.096	0.014
2. 독창적	0.974	0.014	0.404	0.035	0.120
3. 상상력	0.796	-0.033	0.690	-0.002	0.101
4. 완벽히	0.081	0.965	-0.119	-0.189	0.099
5. 효율적	-0.093	0.897	0.297	0.114	0.264
6. 믿음직	-0.040	0.981	-0.155	-0.117	0.042
7. 침울	-0.041	0.067	0.376	0.449	0.876
8. 우울하다	-0.067	0.027	0.372	0.094	0.924
9. 걱정	0.042	0.081	0.148	0.035	0.987
10. 조용하다	0.146	0.299	0.915	0.541	0.269
11. 수다	0.067	-0.026	0.875	0.032	0.068
12. 외향적	0.272	0.016	1.000	0.475	0.045
13. 친절하다	0.102	0.127	0.352	0.976	0.108
14. 도와주다	0.099	0.466	0.468	0.865	0.168
15. 용서하다	-0.030	0.154	0.126	0.776	0.540

〈표 4〉 간편형 한국어 BFI척도의 의미유사도와 요인구조 간의 피어슨 상관계수

상관계수	개방성	성실성	외향성	친화성	정서적안정성
<i>r</i>	0.90**	0.78**	0.81**	0.54*	0.89**

높은 정적 관련성을 갖는다는 것을 보여 주며 성격 5요인 검사에서 수검자의 반응이 하나의 요인으로 묶이면 의미표상이 문항 간 유사함을 보인다.

연구 2

연구 2에서는 각 성격요인 의미표상과 가장 유사한 의미표상을 가지는 문항들을 그 요인의 문항들로 선택하여 성격검사를 구성하였으며 해당 성격검사에 대하여 응답한 자료를 분석한 요인분석에서의 요인구조가 의미표상의 그것과 유사하게 나타나는지를 탐구하였다. 즉, 성격 5요

인의 각 요인명과 의미표상이 유사한 어휘를 이용해 새로운 성격검사문항을 만들고 이에 대한 수검자의 반응을 수집하였다. 수집한 자료를 바탕으로 탐색적 요인분석을 수행하여 요인부하량 행렬을 계산하였고, 의미표상의 유사성 패턴이 요인구조를 통해 드러날 수 있는지 살펴보았다.

방 법

연구대상

서울대학교 심리학과 수업 수강생 154명의 자료가 수집되었고, 의미공간 구성을 위한 예비연구에 참여했던 수검자와 동일했다. 실험에 참여한 수검자는 심리학과 R-point 시스템을 통해 모집되었고, 수검자들은 실험참여 대가로 1 연구점수(credit)을 받았다.

측정도구

예비연구에서 구성한 의미공간 상에 존재하는 어휘 중에서, 연구 1에서 사용된 성격 5요인의 요인명과 의미유사도가 충분히 높은(0.8 이상) 어휘를 각 요인별로 3개씩 선택하였다. 그와 동시에 이 어휘들은 대응되는 요인명이 아닌 다른 요인명 4개에 대해 의미유사도가 평균 0.3으로 충분히 낮았다. 이러한 기준으로 선정된 어휘 목록은 표 5와 같았으며, 의미유사도 행렬은 표 6에 제시되었다. 요인명은 연구 1와 동일하게 구성되었다.

다음으로, 선정된 15개 어휘에 대하여 “나는 ___한 사람이다”라는 검사맥락에 어울리도록 각 어휘별로 간단한 검사문항을 구성하였다. 일부 어휘가 형용사가 아닌 명사 등의 형태를 띠고 있었기 때문에, 이를 성격을 기술하는 형용사 혹은 형용사구의 형태로 변형하였으며 연구에 사용된 검사문항들은 표 7에 제시되어있다. 수검자가 15개의 검사문항에 응답하는 시간은 평균 5분 가량 소요되었다.

〈표 5〉 의미유사도기반 성격의 5요인 측정도구 제작을 위해 선정된 어휘 목록

요인명	어휘
정서적 안정성	예민, 긴장하다, 스트레스
성실성	부지런하다, 책임감, 약속
외향성	활력, 나서다, 적극적
친화성	배려하다, 공감하다, 도와주다
개방성	관대, 진취적, 배우다

〈표 6〉 검사문항에 포함된 어휘 간 의미유사도 행렬

	정서적 안정성	성실성	외향성	친화성	개방성
1. 예민	0.905	0.243	0.094	0.315	0.044
2. 긴장하다	0.934	0.311	0.015	-0.058	0.018
3. 스트레스	0.942	0.254	0.178	-0.010	0.068
4. 부지런하다	-0.054	0.999	0.004	-0.052	0.003
5. 책임감	0.020	0.993	0.095	-0.027	0.055
6. 약속	-0.034	0.994	-0.086	-0.024	-0.010
7. 활력	0.118	0.034	0.847	0.019	-0.030
8. 나서다	-0.013	0.379	0.913	0.336	0.121
9. 적극적	-0.023	0.186	0.932	0.202	0.120
10. 배려하다	0.072	0.089	0.469	0.993	0.046
11. 공감하다	0.075	0.059	0.454	0.987	0.205
12. 도와주다	0.103	0.409	0.566	0.892	0.089
13. 관대	0.147	0.054	0.337	0.122	0.984
14. 진취적	0.059	0.015	0.564	0.083	0.931
15. 배우다	-0.017	0.070	0.602	-0.041	0.860

〈표 7〉 선정된 어휘를 이용해 구성한 검사문항

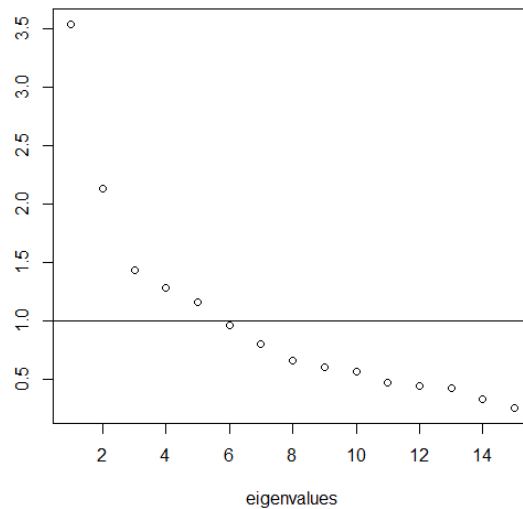
어휘	제시 문항
1. 예민	1. 예민한
2. 긴장하다	2. 긴장을 잘 하는
3. 스트레스	3. 스트레스를 잘 이겨내는
4. 부지런하다	4. 부지런한
5. 책임감	5. 책임감 있는
6. 약속	6. 약속을 꼭 지키는
7. 활력	7. 활력있는
8. 나서다	8. 나서는 것을 좋아하는
9. 적극적	9. 적극적인
10. 배려하다	10. 배려하는
11. 공감하다	11. 공감능력이 뛰어난
12. 도와주다	12. 남을 돕는 것을 즐기는
13. 관대	13. 관대한
14. 진취적	14. 진취적인
15. 배우다	15. 배우는 것을 좋아하는

분석

연구 1에서 제외되었던 10명을 제외한 144명의 검사문항에 대한 응답을 분석하였다. MPLUS 7.0을 이용해 탐색적 요인분석을 수행하였다. 요인부하량 행렬 계산 시, 직교회전(varimax)이 사용되었다. 다음으로 R 3.02를 사용하여 표 5에 제시된 어휘와 요인명의 유사도를 계산하여 구성한 의미유사도 행렬과 두 요인부하량 행렬을 각각 비교하였다.

결과 및 논의

15개의 검사문항에 대한 수검자들의 반응을 수집한 자료를 바탕으로 탐색적 요인분석을 수행한 결과이다. Kaiser(1960)의 기준에 따라, 고유값이 1인 이상의 요인이 다섯 개로 나타나는 것을 확인하였고 최대우도추정법(MLE)이 적용되었다. 5개 요인은 전체 분산의 63.5%를 설명하였고, 각 축별 고유값은 그림 2의 고유치 검사(scree test) 결과로 제시되었다. 5요인 모형의 모형적합도는 표 8과 같다. 적합도는 $\chi^2=54.13(df=40, n.s.)$ 로 나타났다. RMSEA는 .050(90% 신뢰구간에서 .000~.081), SRMR은 .031, CFI와 TLI는 각각 .969과 .918로 나타나 5요인 모형은 자료에 부합하는 것으로 판단하였다. 다음으로 각 요인에 대한 해석을 위해 직교회전을 수행하였고, 회전된 요인 부하량 행렬은 표 9에 제시되어있다.



(그림 2) 공통요인분석 스크리 도표(scree plot)

〈표 8〉 5요인 모형의 모형적합도 지수

모형 적합도	χ^2	df	CFI	TLI	RMSEA (90% C.I.)	SRMR
	54.13	40	.969	.918	0.050 (0.000-0.081)	.031

〈표 9〉 5요인 모형의 요인부하량 행렬

	정서적 안정성	성실성	외향성	친화성	개방성
1. 예민한	0.638	0.132	0.046	-0.084	-0.128
2. 긴장을 잘 하는	0.409	-0.126	-0.124	0.244	0.010
3. 스트레스를 잘 이겨내는	-0.194	0.296	0.132	-0.085	0.111
4. 부지런한	0.181	0.511	0.104	0.076	0.270
5. 책임감 있는	-0.001	0.603	0.084	0.282	0.111
6. 약속을 꼭 지키는	-0.102	0.563	0.005	0.138	0.028
7. 활력있는	-0.213	0.177	0.626	0.242	0.027
8. 나서는 것을 좋아하는	0.095	-0.129	0.736	-0.220	0.290
9. 적극적인	-0.072	0.063	0.858	0.086	-0.192
10. 배려하는	0.035	0.416	0.057	0.596	0.183
11. 공감능력이 뛰어난	0.067	0.120	0.046	0.797	0.196
12. 남을 돕는 것을 즐기는	-0.190	0.225	0.035	0.105	0.639
13. 관대한	-0.048	0.139	0.031	0.218	0.453
14. 진취적인	0.086	0.365	0.560	-0.016	0.153
15. 배우는 것을 좋아하는	0.184	0.456	0.089	0.032	0.339

표 10은 5요인 모형이 의미유사도와 요인구조 간의 피어슨 상관계수를 계산한 결과이다. 외향성($r=.796, p<.01$)과 친화성($r=.556, p<.05$)은 통계적으로 유의미했으며, 정서적 안정성($r=.447, p<.10$)과 성실성($r=.508, p<.01$)은 근사적으로 유의미(marginally significant)했다. 개방성($r=.387, p>.10$)은 통계적으로 유의미하지 않았다.

요인부하량이 각 요인에 대해 .32이상의 값을 가진다면 이는 대응되는 요인에 대해 10% 이상의 설명력을 가진다는 의미이다(Costello & Osborne, 2005). 이러한 기준에 따라 표 9의 요인부하량 행렬을 살펴보았을 때 15문항 중 10문항은 문항별로 의미유사도가 가장 높았던 요인에서 가

〈표 10〉 5요인 모형의 의미유사도와 요인구조 간의 피어슨 상관계수

상관계수	정서적 안정성	성실성	외향성	친화성	개방성
<i>r</i>	0.447 ⁺	0.508 ⁺	0.796**	0.556*	0.387

장 높은 요인부하량을 보인 반면, 5개 문항은 그렇지 않았다. 예컨대 ‘스트레스를 잘 이겨내는’ 문항은 어떤 요인에도 높은 요인부하량을 보이지 않았으며, ‘배려하는’과 ‘진취적인’, ‘배우는 것을 좋아하는’의 문항은 2개 이상의 요인에 교차부하(cross-loaded)되었고, ‘남을 돕는 것을 즐기는’ 문항은 친화성 요인이 아닌 개방성 요인에 대해 가장 높은 요인부하량을 보였다.

연구 1에서 요인분석을 통해 기 타당화된 검사문항의 의미표상은 대응되는 요인명의 의미표상과 유사하고, 그렇지 않은 요인과는 의미표상의 유사성이 낮다는 것을 보였음에도 불구하고, 일부 문항과 개방성 요인에서 의미유사도와 요인구조가 다소 일관적이지 않은 결과를 보였다. 이러한 결과에 대한 단서를 얻기 위해 요인구조와 일관적인 경향을 보였던 10개 문항과 그렇지 않았던 5개 문항을 나누어 보았다. 일관적인 결과를 보인 10개 문항과 나머지 4요인의 의미유사도의 평균값은 0.094였던 반면, 나머지 5개 문항의 대응되지 않는 4요인과의 의미유사도 평균값은 0.183으로, 일관적인 결과를 보인 문항의 값보다 약 두 배 가량 높았다. 종합하자면, 어떤 문항이 그에 대응되는 요인과 의미유사도가 높다고 할지라도 대응되지 않는 나머지 요인과의 의미유사도가 충분히 낮지 않을 경우 요인부하량이 두개 요인에 교차부하 되거나, 측정할 것으로 기대한 성격요인에 대해 낮은 요인부하량을 보일 수 있음을 의미한다.

논 의

본 연구는 성격검사문항에 대한 반응 이전의 수검자의 인지적 단계에 초점을 맞추어, 검사 문항을 어떻게 이해했는가에 대한 단서를 얻기 위해 검사문항의 의미표상에 접근하고자 하였다. 이를 위해 예비 연구에서는 의미표상을 벡터로 표현하기 위한 의미공간을 구성하는 절차를 제시하고, 의미유사도 행렬을 제안하였다. 이를 기반으로 연구 1에서는 타당화된 성격검사의 요인부하량 행렬과 의미유사도 행렬을 비교하였다. 연구 2에서는 검사문항과 성격요인명 간의 의미유사도를 기반으로 선정한 검사문항들의 요인구조를 탐색하였다.

예비 연구에서는 성격검사 맥락을 기반으로 산출한 지문은 잠재의미분석 의미표상에서 소수의 차원만으로도 잘 표현될 수 있다는 것을 보여주었다. 이를 기반으로 성격 5요인을 기술하는 지문과, 지문에 등장하는 어휘들을 5차원 벡터로 표현할 수 있는 기반을 마련하였다. 이를 이용하여 연구 1에서는 간편형 한국어 BFI(김지현 외, 2011)의 요인부하량과 의미유사도 행렬을 비교한 결과, 모든 성격 요인에서 의미유사도와 요인부하량이 높은 일치도를 보였다. 이는 성격검사

문항에 대한 응답이 유사하다면 문항의 의미표상이 유사하다는 것을 시사한다.

연구 2에서는 이와 반대로, 수검자가 이해한 성격 5요인에 대한 의미표상을 이용해 검사문항을 구성하였을 때 검사의 반응을 분석한 탐색적 요인분석 결과는 성격요인의 요인구조와 합치됨을 보였다. 검사문항을 구성할 때, 각 문항을 측정하고자하는 성격요인과 의미유사도가 높고 나머지 요인과는 유사도가 충분히 낮도록 구성하면, 전반적인 요인구조 또한 유사한 패턴을 나타냄을 보였다. 연구 2의 결과는 성격의 5요인 중 4개의 요인에서 요인부하량과 의미유사도가 유의한 정적 상관을 보였다. 이러한 결과는 전반적으로 수검자가 검사문항을 어떻게 이해하는지를 알 수 있다면, 수검자의 검사문항에 대한 반응 패턴을 예측할 수 있음을 의미한다. 다만 두 개의 이상의 성격요인과 유사한 의미표상을 가진 애매한 문항들은 요인구조의 예측이 다소 어려웠다. 예를 들어, 문항 중 ‘진취적인’의 경우, 개방성과는 .931, 외향성과는 .564의 의미유사도를 보였으나 요인구조는 개방성보다는 외향성에 더 높은 요인부하량을 보였다. 또한 마지막 문항인 ‘배우다’의 경우, 개방성과는 .860, 외향성과는 .602의 의미유사도를 보였으나 요인부하량은 개방성이나 외향성의 두 요인이 아닌 성실성에서 가장 높았다. 이러한 결과로 미루어볼 때, 어떤 문항이 그에 대응되는 요인과 의미유사도가 높더라도 그에 대응되지 않는 나머지 요인과의 의미유사도가 충분히 낮지 않으면 수검자의 반응을 이용한 요인구조를 탐색할 때 의미유사도 행렬이 재현되기 어려움을 의미한다.

연구 2에서 보고된 요인구조는 검사 구성 시 엄격한 기준을 적용하여 애매한 문항을 제거하는 과정을 통하여 더욱 의미표상과 일관적인 요인부하량 행렬을 도출할 수 있을 것으로 기대된다.

본 연구의 의의는 다음과 같다. 첫째, 성격검사문항 중에서 하나의 요인으로 묶인 문항들의 의미표상들이 서로 유사했으며, 반대로 의미표상이 유사한 문항들에 대한 반응 또한 하나의 요인으로 묶이는 경향성을 관찰하였다. 즉, 성격검사문항들의 의미표상이 어떤 성격요인의 의미와 유사할수록, 그리고 다른 성격요인과 상이할수록, 대응되는 성격요인을 측정하는 문항에 대해 일관적으로 반응한다는 것을 의미한다. 수검자가 문항을 어떻게 이해했는지에 따라 문항에 대한 반응이 달라질 것이며, 문항의 내용이 많은 수검자들에게 정확하게 이해될 수 있다면 전반적인 수검자들의 반응패턴을 예측할 수 있음을 시사한다. 이는 검사개발 과정에서 실용적인 도움으로 연결된다. 어떠한 성격검사의 구성타당도를 검증하기 위해서 수검자의 반응 자료에만 의존할 경우, 검사문항을 추가하거나 교체할 때마다 수검자의 반응 자료를 새롭게 수집해야 하는 제약이 있다. 그런데 의미공간을 구성하는 절차에 따라서 적절한 검사 맥락에서 수집된 지문을 활용하여 그것을 구성했다면, 이는 성격검사 문항을 개선 및 보완하는 과정에서 요인부하량 행렬을 예측하는 분석 도구로써 기능할 수 있을 것이다. 나아가 본 연구가 제시한 의미유사도 행렬은 다양한 척도에 적용될 수 있는 방법론이므로 성격검사 이외에 다양한 척도를 개발하는데 활용될 수 있다.

둘째, 잠재의미분석을 이용한 의미유사도 행렬은 전통적인 탐색적 요인분석의 결과를 해석하는 데 기여할 수 있다. 의미유사도 행렬에서 요인명 또한 어떤 어휘의 의미표상 벡터임에 주목한다면, 이는 성격요인명을 결정하는 데 객관적 기준을 부여함으로써 기존의 요인분석을 보완한다. 탐색적 요인분석을 통해 추출된 요인들의 이름을 정하고 이를 설명하는 과정은 척도의 내용 타당도를 검증하는 과정을 통해 결정되었으나, 여기에 요인명과 검사문항의 의미표상 간 유사도를 계산하여 양적인 통계치를 제공함으로써 요인명을 결정하는데 도움을 줄 수 있다.

본 연구의 제한점과 후속 연구를 위한 제언은 다음과 같다. 첫째로 일반적인 성격검사는 검사 문항이 단일 어휘가 아닌 문장의 형태를 갖는다는 점이다. 의미유사도 행렬 구성 시 문장의 의미표상을 성격요인의 의미표상과 비교할 수 있어야 본 연구의 결과가 다른 성격검사까지 확장될 수 있을 것이다.

어휘의 의미표상을 활용해 문장의 표상으로 만드는 방법에 대해서는 단순 어휘벡터의 덧셈(Laundaur & Dumais, 1997)이 있었으나 이는 문장의 의미표상을 구현하는 올바른 방법이라 보기 어렵다. 문장의 의미표상은 단순한 어휘의 의미표상의 합으로 볼 수 없으며, 문장의 해석에는 어휘의 등장 순서 또한 큰 영향을 미친다. 그밖에도 문장의 문법구조 및 다양한 품사들이 수검자의 의미표상에 영향을 미치므로 적절한 문장의 의미표상벡터 계산법이 필요하다. 최근 들어 활발히 연구되고 있는 문장 및 문서의 분산표상을 학습(Le & Mikolov, 2014)하는 방법론을 적용하여 본 연구의 결과를 확장하는 후속 연구가 요구된다.

둘째, 성격검사뿐만 아니라 다양한 구성개념을 측정하는 척도를 타당화 할 때 본 연구의 결과가 검사개발 과정에서 도움을 주려면 성격요인이 아닌 다른 구성개념을 측정하는 검사문항들의 의미표상 또한 탐구해야한다. 대표적으로 정적정서 및 부정적정서를 측정하는 척도(PANAS)는 각 문항이 단일어휘로 이루어진 검사(Watson, Clark, & Tellegen, 1988)로써 한국어로 변안되었으며(이현희, 김은정 & 이민규, 2003), 최근 박홍석과 이정미(2016)에 의해 재타당화 되었다. 이 척도는 정적 정서 및 부정 정서를 측정하기 위해 20개의 문항을 이용하는데, 이러한 척도들은 본 연구가 제안한 방법론을 성격요인을 측정하는 검사가 아닌 다른 척도에 적용하기 용이한 척도로서 이를 활용한 후속 연구가 기대된다.

참고문헌

- 강승식 (2002). **한국어 형태소 분석과 정보 검색**. 서울: 홍릉과학출판사.
- 김지현, 김복환, 하문선 (2011). 간편형 한국어 BFI(Big Five Inventory) 타당화 연구. **인간이해**, 32, 47-65.
- 민경환 (2002). **성격심리학**. 서울: 법문사.

- 박홍석, 이정미 (2016). 정적 정서 부적 정서 척도(PANAS)의 타당화. **한국심리학회지: 일반**, 35(4), 617-641.
- 이태현 (2003). **ELSA를 이용한 주제별 문서분류 및 다의어 의미 해소**. 석사학위논문, 서울대학교, 서울.
- 김청택, 이태현 (2002). 뇌와 인지 모형: 잠재의미 분석을 사용한 문서분류. **한국심리학회: 인지 및 생물**, 14, 309-319.
- 이태현, 김청택 (2004). LSA 모형에서 다의어 의미의 표상. **인지과학**, 15, 1-53.
- 이현희, 김은정, 이민규 (2003). 한국판 정적 정서 및 부적 정서 척도(Positive Affect and Negative Affect Schedule; PANAS)의 타당화 연구. **한국심리학회지: 임상**, 22(4), 935-946.
- 지혜성, 조준희, 임희석 (2010). 한국어 문장 표절 유형을 고려한 유사 문장 판별. **한국컴퓨터교육학회논문지**, 13(6), 79-89.
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta analysis. *Personnel Psychology*, 44(1), 1-26.
- Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, 38, 476-506.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1-9.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272.
- Galton, E. (1884). Measurement of character. *Fortnightly Review*, 36, 179-185
- Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216-1229.
- Gong, Y., & Liu, X. (2001). Generic text summarization using relevance measure and latent semantic analysis. In *Proceedings of the 24th annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 19-25.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371-395.
- Hill, F., Reichart, R., & Korhonen, A. (2015). Simlex-999: evaluating semantic models with (genuine) similarity estimation. *Computational Linguistics*, 41(4), 665-695.
- Hofstee, W. K., Kiers, H. A., De Raad, B., Goldberg, L. R., Ostendorf, F. (1997). A comparison of big-five structures of personality traits in Dutch, English, and German. *European Journal of Personality*, 11(1), 15-31.

- John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, 2(1999), 102-138.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151.
- Kwantes, P. J., Derbentseva, N., Lam, Q., Vartanian, O., & Marmurek, H. H. (2016). Assessing the big five personality traits with latent semantic analysis. *Personality and Individual Differences*, 102, 229-233.
- Tucker, L. R., & MacCallum, R. C. (1997). Exploratory factor analysis. *Unpublished manuscript*, Ohio State University, Columbus.
- Landauer, T. K., Laham, D., Rehder, B., & Schreiner, M. E. (1997). How well can passage meaning be derived without using word order? A comparison of latent semantic analysis and humans. In *Proceedings of the 19th Annual Meeting of the Cognitive Science Society*, 412-417.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211.
- Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, 1188-1196.
- Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G., & Ruppin, E. (2002). Placing search in context: The concept revisited. *ACM Transactions on Information Systems*, 20(1), 116-131.
- McCrae, R. R., & Costa, E. T. (1985). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences*, 6, 587-597.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
- McCrae, R. R., & Costa, P. T. (1989). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology*, 56(4), 586-595.
- Nakov, P., Popova, A., & Mateev, P. (2001). Weight functions impact on LSA performance. *EuroConference RANLP*, 187-193.
- Nigg, J. T., John, O. P., Blaskey, L. G., Huang-Pollock, C. L., Willcutt, E. G., Hinshaw, S. P., & Pennington, B. (2002). Big five dimensions and ADHD symptoms: Links between personality traits and clinical symptoms. *Journal of Personality and Social Psychology*, 83(2), 451-469.
- Noffle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, 93(1), 116-130.
- Norman, W. T. (1967). *2800 personality trait descriptors: Normative operating characteristics for a university population*. Ann Arbor: Department of Psychology, University of Michigan.

- Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*, 401-421.
- Saucier, G. (2002). Orthogonal markers for orthogonal factors: The case of the big five. *Journal of Research in Personality, 36*, 1-31.
- Thurstone, L. L. (1934). The vectors of mind. *Psychological Review, 41*, 1-32.
- Tourangeau, R., Rips, L. C., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology, 47*, 1063-1070.
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*(2), 158-177.
- Wild, F. (2014). lsa: Latent Semantic Analysis: R package version 0.73.

1차 원고 접수: 2018. 11. 06
1차 심사 완료: 2019. 5. 31
2차 원고 접수: 2019. 8. 2
2차 심사 완료: 2019. 9. 9
3차 원고 접수: 2019. 9. 16
3차 심사 완료: 2019. 9. 25
4차 원고 접수: 2019. 10. 01
최종 게재 확정: 2019. 10. 02

(Abstract)

A Comparison between Factor Structure and Semantic Representation of Personality Test Items Using Latent Semantic Analysis

Sungjoon Park
KAIST
School of Computing

Heeyoung Park
Seoul National University
Department of Psychology

Cheongtag Kim

To investigate how personality test items are understood by participants, their semantic representations were explored by Latent Semantic Analysis. In this thesis, *Semantic Similarity Matrix* was proposed, which contains cosine similarity of semantic representations between test items and personality traits. The matrix was compared to traditional factor loading matrix. In preliminary study, semantic space was constructed from the passages describing the five traits, collected from 154 undergraduate participants. In study 1, positive correlation was observed between the factor loading matrix of Korean shorten BFI and its semantic similarity matrix. In study 2, short personality test was constructed from semantic similarity matrix, and observed that its factor loading matrix was positively correlated with the semantic similarity matrix as well. In conclusion, the results implies that the factor structure of personality test can be inferred from semantic similarity between the items and factors.

Key words : LSA, EFA, construct validity, Big Five, semantic representation