

## 특집: 감각과학의 이해와 산업적 활용

## 관능평가에 대한 제언

곽 한 섭

한국식품연구원 감각인지연구단

## Precautions for Sensory Evaluation

Han Sub Kwak

Research Group of Cognition and Sensory Perception, Korea Food Research Institute, Gyeonggi 13539, Korea

최근 식품 연구는 Research & Business Development에 기반을 두어 단순히 연구로 끝나는 것이 아니라 연구 결과를 사업적으로 연결하는 연구 목표 설정이 많이 이루어지고 있다. 아울러 농림축산식품부의 ‘농식품6차 산업’에서도 농산물의 ‘부가가치의 증대’를 강조하고 있다. 농산물의 부가가치 증대는 식품 산업과 필연적으로 이어진다. 농산물을 가공하여 고부가가치 식품을 생산하는 기존의 관점에서, 이제는 제조된 식품에 대해서 소비자가 어떻게 가치를 부여하고 어떻게 받아들이는가를 측정하여, 소비자가 원하는 식품을 시장에 출시하여 실질적인 매출로 이어지는 진정한 부가가치의 증대로 인식의 변화가 이루어지고 있다. 최근 식품 연구 및 정책 동향과 발맞추어 식품에 대한 정량적인 관능 분석 및 소비자들의 평가 결과의 필요성이 높아지고 있다. 이러한 분야를 담당하는 관능평가의 중요성이 식품 연구 및 산업에서 급격히 증가하고 있으나, 국내에서 발표되는 학술논문의 관능평가 수준은 국제적인 기준에 미치지 못하는 현실이다.

유럽과 미국에서는 관능평가를 기존의 식품가공학, 식품미생물학 및 식품화학 등과 더불어 하나의 세부 학문분야로 인정하고 있다. 그러나 한국에서는 단순히 개발 제품에 대하여 소비자 기호도를 평가한다는 수준의 인식을 가지고 있으며, 이러한 인식이 관능평가 연구에서 많은 오류를 범하고 발전을 저해하고 있다고 생각한다. 미국의 경우 식품과학 전공이 설치된 대부분 대학에서 관능평가를 전공한 교수가 1~3명 있으며, 관능평가 방법 개발, 제품과 연계한 관능평가 연구 등 다양한 연구를 진행하고 있다. 또한, North Carolina State University, University of Arkansas, Kansas State University의 경우 학과 내 관능평가 센터를 설립하여 기업체 및 외부의 과제를 대행하며 다양한 식품 및 소비재 용품의 관능평가를 진행하고 있다. 아울러 미국의 식품 및 소비재 회사는 오래전부터 독립된 관능평가 부서를 운영하고 있다. 필자가 2009년에 인턴으로 일했던 세계 최대의 향신료 회사인 McCormick & Company의 경우 약 40년 전부터 관능평

가를 진행하였으며, 현재는 R&D 관능평가 부서, 소비자 관능평가 부서 및 품질관리 관능평가 부서로 나누어 10명이 넘는 관능평가 전공자들이 일하고 있다. 이들은 자사의 향신료, 시즈닝 등에 대해서 다양한 관능평가를 진행하고 제품개발에 참여하고 있다. 또한, 연구소에서 진행된 다양한 연구결과를 세계적인 학술지에 발표하여 자사의 연구 활동을 홍보하고 있다(1,2). 미국의 Target과 같은 대형 유통회사도 관능평가 전공자를 채용하여 자사 PB상품의 발굴 및 품질관리에 활용하고 있으며, Tragon company와 Sensory Spectrum과 같은 관능평가 대행업체의 규모도 나날이 증가하고 있다. 현재 국내 산업체의 경우 대형 식품과 화장품 회사부터 서서히 관능평가를 전공한 인력을 채용하여 생산하는 제품의 지속적인 평가 및 모니터링을 하고 있으나, 아직 관능평가에 대한 경영진의 인식은 부족한 편이다. 또한, 학계의 경우 관능평가를 전공한 교수의 숫자가 타 세부전공과 비교해서 절대적으로 부족한 현실이다.

2015년 한국식품영양과학회지에 출판된 논문을 검토한 결과 관능평가를 진행한 논문이 총 40편으로 나타났다. 필자의 판단 기준으로 이들 40편에 대한 관능평가 오류에 대해서 Table 1과 같이 나누었다. 한국식품영양과학회지에 출판된 논문에서 SCI(E) 논문 수준에 있는 관능평가 논문은 없는 것으로 나타났다. 명문화된 규정은 없으나 일반적으로 SCI(E) 논문에 출판되기 위해서는 소비자 조사의 경우 100명 이상의 소비자를 필요로 하고, 실험과정에 대해서 자세히 서술되어 있어야 한다. 묘사분석은 훈련된 패널을 6~12명 사용하여야 하며, 실험 방법에 실험 과정 및 표준시료가 자세히 서술되어 있어야 한다. 또한, 다변량 통계분석을 이용하여 분석 결과에 대하여 구체적으로 서술하여야 한다. 관능평가의 선진국으로 알려진 미국과 유럽의 경우 이미 약 20년 전에도 관능평가 논문에 이러한 사항들을 자세히 서술하고 있다(3,4). 그러나 현재 국내에서 발표된 연구의 경우 1990년대 유럽/미국 수준의 관능평가에도 못 미치는 실험 방법 및 결과 해석을 자주 접할 수 있다.

**Table 1. 2015년도 한국식품영양과학회지에서 관능평가를 진행한 논문(40편)의 오류 분석 결과**

구분	비율
SCI 논문 수준	0.0%
문제없음	20.0%
실험 방법 표기 오류	52.5%
소비자 조사 시 인원수(30명) 미만	52.5%
척도 사용 오류	52.5%
실험 방법 서술 부족	45.0%
훈련-비훈련 패널 사용 오류	27.5%
묘사분석 및 소비자 조사에 같은 패널 사용	12.5%

40편의 논문 중 오류가 없는 논문은 8편으로 전체 논문 중 20%를 차지하고 있다(Table 1). 오류가 없다는 것은 단지 필자가 분류한 대표적인 오류를 범하지 않은 것으로 명확한 실험 방법 표기, 소비자 조사의 경우 30명 이상의 패널 사용 및 9점 척도 사용, 묘사분석의 경우 훈련된 패널의 사용 및 올바른 척도 사용을 한 실험을 언급한다. SCI(E)급 학술지의 수준에는 못 미치지만 관능평가의 실험 설계 및 방법 서술이 올바르게 진행된 것으로 볼 수 있다. 그중에서 Im 등(5)의 논문은 관능평가의 실험 방법에 대하여 묘사분석 및 소비자 기호도 조사로 명확하게 나누는 것을 보여준다. 묘사분석에서 패널 선정, 패널 훈련 및 표준 시료 개발, 묘사분석 과정에 대하여 구체적인 서술을 하고 있다. 또한, 소비자 조사에서는 패널 정보, 시료 제시 방법, 사용 척도를 명확하게 언급하고 있다. 닭가슴살의 품질 특성에 대한 연구를 진행한 Yang 등(6)의 논문도 묘사분석 및 소비자 기호도 조사에 대하여 어떻게 실험이 진행되었는지 충분한 설명을 하고 있다. 관능평가 실험 방법 서술에 대해서 Im 등(5) 및 Yang 등(6)의 논문을 참고한다면, 관능평가 실험 방법에 대하여 충분한 설명을 논문에 작성할 수 있을 것이다.

관능평가 실험 오류에 대해서 분석한 결과 실험 방법 표기 오류, 소비자 조사 시 인원수 부족(30명 기준) 및 척도 사용 오류가 각각 관능평가를 진행한 논문의 52.5%에서 나타났다(중복집계 포함). ‘실험 방법 표기 오류’는 ‘실험 방법’의 소제목으로 ‘관능평가’라고 작성한 것이 대부분이다. ‘관능평가’는 하나의 실험 방법이 아닌 세부 학문 단위라는 것을 인식하여야 한다. Stone과 Sidel(7)은 ‘관능평가’의 정의를 아래와 같이 내리고 있다.

*A scientific discipline used to evoke, measure, analyze, and interpret those responses to products that are perceived by the senses of sight, smell, touch, taste, and hearing.*

이러한 정의에 비추어 볼 때, 실험 방법 서술에 ‘관능평가’라고 소제목을 붙이는 것은 너무 광범위한 내용을 담고 있다고 생각한다. 일반적으로 관능평가라는 것은 크게 소비자 기호도 조사, 묘사분석, 차이식별의 세 가지 정량

적인 실험 방법 및 초점집단 인터뷰(Focus Group Interview) 등의 정성적인 실험 방법을 포함하는 광범위한 범위를 나타낸다. 관능평가(Sensory Evaluation)의 제목을 가진 교과서를 보아도 위에서 언급된 모든 실험 방법을 ‘관능평가’ 안에 설명하고 있다. 따라서 논문 작성 시 어떠한 관능평가 실험을 하였는지를 명확하게 서술하여야 한다. 세부적으로 소비자 패널을 이용하여 기호도 평가를 하였다면 ‘소비자 기호도 조사’, 훈련된 패널을 이용하여 제품의 강도 및 특성을 평가하였다면 ‘묘사분석’, 제품의 차이에 대하여 평가하였다면 ‘차이식별’ 등의 어떠한 관능평가 방법을 사용하였는지 구체적으로 서술하는 것이 올바르다고 생각한다.

소비자 기호도 평가 시 패널 수(30명 이상)가 부족한 경우가 52.5%의 논문에서 발견되었다. 소비자 조사 시 통계적으로 유의미한 결과를 얻기 위해서는 일정 수준 이상의 소비자 패널이 필요하다. 일반적으로 SCI(E) 논문에서는 100명 이상의 패널을 이용해서 실험을 진행하는 경우가 많다. 상황에 따라서 100명 미만의 패널도 인정을 해주나, 이러한 것은 평가의 특이성 또는 제한성이 있거나 시료 제시 방법상의 문제로 100명에 근접한 패널을 대상으로 하는 경우가 많다. 필자의 경우도 89명의 패널로 소비자 조사를 진행한 결과를 포함한 논문을 SCI(E)급 저널에 투고하였으나, 에디터가 마지막에 ‘소비자 패널 숫자 부족’을 이유로 논문 심사자들의 ‘논문 게재 가능’ 결과를 뒤집은 결과를 통보받은 경험이다. 그만큼 소비자 조사 시 패널의 숫자는 논문의 통과 여부를 결정할 만큼 중요하며, 실험 결과의 신뢰도를 나타낸다고 할 수 있다. SCI(E)급이 아닌 국내 학술지에 100명이라는 기준을 동등하게 적용하기에는 무리가 있다고 생각한다. 그러나 2015년 한국식품영양과학회지 일부 논문의 경우 10~20명의 인원을 가지고 소비자 조사를 진행하였다. 이러한 경우는 소비자 조사에 대한 이해가 부족한 것으로 판단된다. 통계학에서 t-분포와 정규분포를 나눌 때, 일반적으로 30개 표본을 기준으로 한다. 관능평가 결과 분석은 데이터의 정규분포를 가정하고 결과를 분석하기에(8) 국문지의 경우 단순 평균 비교의 경우 최소 30명 이상의 패널이 있어야 한다고 생각한다. 덧붙여 principal component analysis, factor analysis, agglomerative hierarchical clustering analysis 및 partial least squares regression analysis와 같은 다변량 분석을 위해서는 50명 이상의 소비자 패널을 이용한 실험을 진행하여야 통계적으로 의미 있는 결과를 나타낼 수 있다고 생각한다. 향후 관능평가가 국내 식품 연구에 자리를 잡아 나간다면 소비자 조사에 참여하는 최소 패널의 수도 현재 제시한 30명에서 점차 증가하여야 한다고 본다.

척도 사용의 오류는 소비자 조사 시 가장 대표적으로 사용되는 9점 척도를 사용하지 않은 경우와 묘사분석 시 15~16점 척도 또는 10 또는 15 cm 선척도를 사용하지 않은 경우를 잘못된 실험으로 판단하였다. 소비자 조사

시 5점 또는 7점 척도를 사용한 논문도 몇 편 나타났다. 특히, 여러 제품을 평가할 때 5점 척도를 사용할 경우 제품 간에 기호도를 분별하는 분별력이 떨어질 우려가 있기에 부득이한 경우가 아니면 사용하지 않기를 권한다. 현재 대부분의 소비자 기호도 평가에서 9-point hedonic scale이라 불리는 9점 척도를 사용한다. 이 척도는 1940~50년대 미국에서 군용 식량 및 급식 평가를 위해 연구가 진행되었으며(9), 세계적으로도 가장 많이 사용되고 있다. 대부분의 SCI(E) 논문에서 9점 척도를 사용하고 있으며, 소비자 조사에서 가장 안정된 척도라 할 수 있다. 9점 척도에 대한 다양한 연구가 수십 년간 진행되었으며, 대표적으로 어떻게 9점 척도를 사용해야 하는가에 대한 연구(10) 및 국가별 9점 척도 비교 연구(11) 등이 있다. 따라서 소비자 기호도 평가를 진행할 시 관능평가 분야에서 가장 널리 쓰이고 있는 9-point hedonic scale을 사용하는 것을 권장한다. 묘사분석의 경우도 조사된 논문에서 7점 척도, 9점 척도 등 다양한 척도가 사용되고 있으나, 세계적으로 많이 쓰이고 있는 15~16점 척도 또는 10 또는 15 cm 선척도를 사용하는 것이 시료 사이의 묘사 특성의 강도를 가장 잘 구분할 수 있다고 본다.

실험 방법 서술의 부족도 45%의 논문에서 발견되었다. 묘사분석의 경우 단순히 “훈련된 패널을 이용하여 어떠한 특성에 대해서 측정하였다”라고 간략히 서술한 논문도 있으며, 소비자 조사의 경우 “몇 명의 패널이 기호도 조사에 참여하였다” 정도로 재료 및 방법을 작성한 논문도 발견되었다. 특히, 묘사분석의 경우 표준시료로 사용된 제품 및 강도특성을 밝혀주는 것이 필요하나, 그렇지 못한 논문이 다수 보였다. 이화학 및 미생물 분석 논문의 경우 많은 연구자가 기준에 출판되었던 논문을 참고하여 실험을 진행한다. 그러나 한국식품영양과학회지의 관능평가 논문 중 상당수는 다른 연구자가 실험 방법을 읽고 재현 실험을 진행하기 어려울 정도로 간략하게 실험 방법이 서술되어 있었다. 묘사분석에 대한 자세한 실험 방법 서술 및 통계분석은 Kappes 등(12)과 N'Kouka 등(13)의 논문을 참고한다면 완성도 높은 묘사분석 서술 및 통계분석을 할 수 있을 것이다. 국내 학술지에서는 무균포장죽에 대하여 묘사분석 및 소비자 조사를 진행한 Kwak 등(14), 닭강정에 대한 묘사분석 및 소비자 조사를 진행한 Lee 등(15) 및 순두부찌개 소스의 묘사분석과 소비자 조사를 진행한 Im 등(5)의 논문을 참고하여 관능평가와 관련된 논문을 작성한다면 보다 구체적이고 정확한 실험 방법 나아가 결과까지 논문에 서술할 수 있을 것으로 생각한다.

훈련-비훈련 패널 사용 오류는 묘사분석을 진행하는데 훈련되지 않은 패널을 사용하거나, 소비자 조사 시 훈련된 패널을 사용한 것을 의미하는 것으로 27.5%의 논문에서 발견되었다. 특히, 소비자 조사 진행 시 ‘훈련된 패널을 사용하여 기호도 평가를 진행하였다’와 같은 서술이 오류의 주를 이루고 있다. 소비자 조사와 묘사분석에 사

용되는 패널은 서로 전혀 다른 성격을 지니고 있다. 소비자 조사는 특정한 식품에 대하여 훈련을 받지 않은 일반 사람을 대상으로 하는 것으로, 연구의 목적에 따라 나이, 성별, 섭취 빈도 등에 대해서는 제한을 둘 수 있다. 반면에, 묘사분석에서는 충분한 훈련을 받은 패널을 이용하여야 하며, 이들은 소비자 조사에 참여하여서는 안 된다. 대표적인 관능평가 교과서인 Sensory Evaluation Practices에서 ‘묘사분석’에 대해서 아래와 같이 서술하고 있다(7).

*Descriptive analysis is a sensory methodology that provides quantitative descriptions of products, based on the perceptions from a group of qualified subjects.*

Stone과 Sidel(7)이 서술한 것과 같이 묘사분석은 ‘qualified subjects’를 필요로 한다. 여기서 ‘qualified subjects’란 묘사분석을 위해서 충분히 훈련받은 패널이라 할 수 있다. 묘사분석을 진행할 때 어떻게 패널을 선발하여 훈련을 진행하였는지 구체적으로 서술을 해주어야 한다. 묘사분석의 패널 선발 및 훈련에 대해서는 Meilgaard 등(16)이 출판한 Sensory Evaluation Techniques를 참고하길 추천한다. 잘못된 패널의 사용은 잘못된 이화학 분석기기 선택과 동일함을 기억해야 한다.

묘사분석 및 소비자 조사에 같은 패널을 사용한 경우가 12.5%로 나타났다. 훈련-비훈련 패널 패널 사용과 유사하나, 이 경우는 하나의 패널로 두 가지 평가를 한 것에 해당한다. 묘사분석과 소비자 조사에 사용되는 패널은 위에서 언급하였듯이 전혀 다른 성격을 가지고 있으며, 절대로 같은 패널로 실험을 진행해서는 안 된다. 일부 논문의 경우 ‘10명의 패널을 이용하여 전반적인 기호도에 대하여 평가하였으며, 그 후 시료의 특성강도에 대해서도 평가를 하였다’와 같이 서술된 논문이 보였다. 이것은 ‘HPLC를 이용하여 유기산과 휘발성 향미 성분을 분석하였다’와 유사한 실험 방법 서술이라고 생각하면 이해하기 쉬울 것이다. 따라서 위에서 언급한 바와 같이 비훈련 패널은 소비자 조사에 이용하고, 훈련된 패널은 묘사분석에 이용하는 실험 설계를 하여야 하며, 어떠한 경우에도 같은 패널로 소비자 조사와 묘사분석을 동시에 진행해서는 안 된다. 최근에는 소비자 패널을 이용하여 묘사분석을 진행하는 방법에 대하여 많은 연구가 되고 있으나(17, 18), 이 부분은 좀 더 논의가 필요한 하다고 생각한다.

국내의 짧은 관능평가의 역사, 이해 부족 및 부족한 연구 인력을 고려할 때 국내 학술지에 출판되는 관능평가의 질적 수준은 국제적인 기준에 한참 못 미치고 있다. 천릿길도 한걸음이라는 속담이 있듯이 다양하게 발견되고 있는 관능평가에 대한 오류를 점차 고쳐나가다 보면 국내 학술지에서도 수준 높은 다양한 관능평가 연구 결과를 볼 수 있을 것이다.

## 참고문헌

1. Hottenstein AW, Taylor R, Carr BT. 2008. Preference segments: A deeper understanding of consumer acceptance or a serving order effect?. *Food Qual Prefer* 19: 711-718.
2. King SC, Meiselman HL, Hottenstein AW, Work TM, Cronk V. 2007. The effects of contextual variables on food acceptability: A confirmatory study. *Food Qual Prefer* 18: 58-65.
3. Hough G, Sánchez R. 1998. Descriptive analysis and external preference mapping of powdered chocolate milk. *Food Qual Prefer* 9: 197-204.
4. Prindiville EA, Marshall RT, Heymann H. 1999. Effect of milk fat on the sensory properties of chocolate ice cream. *J Dairy Sci* 82: 1425-1432.
5. Im P, Han JH, Kim YC, Lee B, Kim MY, Chang Y, Yu S, Lee Y. 2015. Effect of guar gum on quality of soft tofu stew sauce. *J Korean Soc Food Sci Nutr* 44: 442-448.
6. Yang I, Im P, Kang J, Kwak H, Jeong Y, Kim M, Lee Y. 2015. Quality of commercial broiler breast meat retailed in Korea. *J Korean Soc Food Sci Nutr* 44: 1693-1699.
7. Stone H, Sidel JL. 2004. *Sensory Evaluation Practices*. 3rd ed. Academic Press, San Diego, CA, USA.
8. O'Mahony M. 1986. *Sensory Evaluation of Food: Statistical Methods and Procedures*. Marcel Dekker, New York, NY, USA.
9. Peryam DR, Pilgrim FJ. 1957. Hedonic scale method of measuring food preferences. *Food Technol* 11: 9-14.
10. Nicolas L, Marquilly C, O'Mahony M. 2010. The 9-point hedonic scale: Are words and numbers compatible?. *Food Qual Prefer* 21: 1008-1015.
11. Yeh LL, Kim KO, Chompreeda P, Rimkeeree H, Yau NJN, Lundahl DS. 1998. Comparison in use of the 9-point hedonic scale between Americans, Chinese, Koreans, and Thai. *Food Qual Prefer* 9: 413-419.
12. Kappes SM, Schmidt SJ, Lee SY. 2006. Descriptive analysis of cola and lemon/lime carbonated beverages. *J Food Sci* 71: S583-S589.
13. N'Kouka KD, Klein BP, Lee SY. 2004. Developing a lexicon for descriptive analysis of soymilks. *J Food Sci* 69: 259-263.
14. Kwak HS, Oh YJ, Kang HB, Kim TH. 2013. Descriptive profile and liking/disliking factors for aseptic-packaged rice porridge. *J Korean Soc Food Sci Nutr* 42: 1878-1885.
15. Lee SM, Bae SJ, Kim KO. 2015. Sensory characteristics and cross-cultural acceptability of sweet crispy chicken (Dak-gangjeong) prepared using sauces with different ethnic Korean style flavors among Korean and Chinese consumers. *Korean J Food Sci Technol* 47: 623-632.
16. Meilgaard MC, Civille GV, Carr BT. 2006. *Sensory Evaluation Techniques*. 4th ed. CRC Press, Boca Raton, FL, USA.
17. Ares G, Jaeger SR. 2013. Check-all-that-apply questions: Influence of attribute order on sensory product characterization. *Food Qual Prefer* 28: 141-153.
18. Dooley L, Lee YS, Meullenet JF. 2010. The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. *Food Qual Prefer* 21: 394-401.