

식품 품질관리를 위한 신호탐지이론(SDT) 감각차이식별분석 이론과 생수 품질관리에서의 활용

Food quality management using sensory discrimination method based
on signal detection theory and its application to drinking water

김민아^{1,*} · 심혜민² · 이해성¹

Min-A Kim^{1,*}, Hye-Min Sim², and Hye-Seong Lee¹

¹이화여자대학교 식품공학과, ²농심 식품연구소

¹Department of Food Science and Engineering, ELTEC College of Engineering, EwhaWomans University,
Seoul 03760, South Korea

²NongShim R&D Nutrition Research Team, Seoul 07057, South Korea

Abstract

Sensory perception of food/beverage products is one of the most important quality factors to determine consumer acceptability and thus sensory discrimination methodology has been a vital tool for quality management. Signal detection theory(SDT) and Thurstonian modeling provide the most advanced psychometric approach to modeling various discrimination methods. In these theories, perceptual and cognitive decisional factors are considered so that, a fundamental measure of sensory difference (d') can be computed, independent of test methods used. In this paper, sensory discrimination analysis based on SDT and Thurstonian modeling is introduced for

more accurate and systematic applications of sensory and hedonic quality management in industry. Ways to realize the statistical power and relative sensitivity of sensory discrimination methods theorized in SDT and Thurstonian modeling in practice, are also discussed by using a case study of the Nongshim quality management program for drinking water in which SDT A-Not A test methodology was further optimized.

Key words: Quality management, signal detection theory, test power, action standard, consumer relevance

* Corresponding author: Min-A Kim, Department of Food Science and Engineering, ELTEC College of Engineering, EwhaWomans University, Seoul 03760, South Korea
Tel: 82-2-3277-4428
E-mail: mina_kim@ewha.ac.kr
Received February 11, 2019; revised March 10, 2019; accepted March 10, 2019



서론

1. 식음료 제품의 감각 품질 관리의 중요성

소비자가 식품을 섭취할 때 지각하는 식품의 감각적 특성과 이로 인한 즐거움과 쾌감(pleasantness)은 소비자의 식품에 대한 기호도(hedonic, liking, acceptability)를 결정 짓는 가장 중요한 요소이며, 이러한 소비자가 지각하는 식품의 감각적 및 기호적 품질은 소비자의 식품 선택을 결정 짓는 가장 중요한 품질 요소 중 하나이다. 따라서, 최근 식품 업계는 식품의 감각적 품질(‘맛 품질’)은 유지하면서 더 건강하고 지속 가능하며 비용 측면에서도 보다 효율적으로 제품을 생산하기 위해, 원료 배합물을 달리하거나 가공 공정을 변경하는 데에 많은 노력을 기울이고 있다. 일례를 들면, 소비자는 건강상의 이익을 위해 나트륨, 설탕, 또는 지방 함량이 저감화된 제품을 원하지만, 이로 인하여 맛 품질이 저하되는 것은 원하지 않을 수 있다. 다시 말해, ‘더 건강한 제품’이 소비자를 만족시키기 위해 ‘좋은 맛’을 유지해야 한다는 점은 여전히 중요하다.

따라서 산업 환경 및 소비자의 요구 변화로 인하여 제품 성분(또는 공정)을 변화시켰을 때에는, 소비자의 입장에서 소비자가 지각하는 감각적 및 기호적 품질을 이해하고, 소비자가 허용할 수 있는 기준에서 맛 품질이 변화되었는지 연구·검토할 수 있는 단계가 필수적으로 요구된다. 제품 성분(또는 공정)의 작은 변화로 인해 소비자가 지각하는 제품의 종합적인 감각 또는 기호 품질이 크게 달라질 수 있다. 또한 얼마만큼의 제품 차이가 소비자 입장에서 다른 (또는 동일한) 품질의 제품으로 간주되어 소비자의 선택에 영향을 미치는지는 제품 특성에 따라 다를 수 있다. 예를 들어, 품질이 균일하지 못한 신선식품과 품질이 균일하게 조절되는 가공식품에 대해서는 감각적 특성에 대한 품질관리에 다른 기준을 적용해야 한다. 따라서 제품 특성의 차이와 소비

자가 인지하는 제품의 감각적 품질 기준을 고려하여 식품 및 음료 제품의 감각 품질의 차이를 측정하는 것이 중요하며, 다양한 제품 군에 대해서 차별적인 품질관리 방안을 마련하고 활용하는 것이 식품 산업에서 성공적인 품질관리 프로그램을 운영하는 데 중요한 과제가 될 것이다.

2. 식음료 제품의 감각 품질 측정에 활용되는 감각 차이식별 분석 이론의 소개 배경

제품의 감각적 품질 변화를 측정하는 가장 주된 연구방법은 감각 차이식별 분석(sensory discrimination analysis) 방법이다. 차이식별 연구방법론은 일반적으로 감각 차이 검사(sensory difference testing)라고도 불리며, 이항 분포(binomial distribution)에 근거한 차이 검정을 바탕으로 분석하는 삼점 검사(triangle test) 방법이 산업계에서 가장 널리 사용되어 왔다. 그러나 이 방법은 1950년대 식품 산업에서 개발되었고(Peryam, 1958; Peryam과 Swartz, 1950), 그 당시에는 감각 평가 이론에 적용할 수 있는 인간 행동 및 인지 심리적 측면에 대한 과학적 정보가 존재하지 않았기 때문에 전통적인 방식의 삼점 검사를 통한 자료 분석은 매우 제한적이라고 할 수 있다. 이항 분포에 의한 차이 검정이란 감각 평가에서 사용되는 인지적인 결정과정에 대한 정보는 무시하고, 단순히 선택 항목의 수에 따라 결정되는 기회 수준에 근거해서 총 응답 결과가 두 시료 간에 유의적인 차이를 보여주는지를 판단하는 분석방법이다. 이 분석방법으로는 제품 간의 차이의 정도를 나타내는 결과를 도출할 수 없고, 차이가 나는지 여부만을 확인할 수 있다. 또한 이항 분포에 의한 유의성 판단은 감각 평가방법에 의존적이기 때문에 결과의 활용 측면에서도 제한적이다. 평가방법에 의존적이라는 것은 삼점 검사를 수행해서 얻은 결과가 다른 평가방법, 예를 들면 일·이점 검사(duo-trio test) 방법이나 삼자 택일(3-AFC; 3-Alternative Forced Choice) 방법을 수행

해서 얻은 결과와 동일한 결과를 도출한다고 볼 수 없다는 것이다. 실제로 삼점 검사보다 3-AFC 방법을 사용하면 더 높은 정답률을 얻을 수 있다.

이와 같은 이항 분포 분석의 한계를 극복하기 위해 감각 차이식별 연구 분야는 개별적인 차이 검사 방법들로의 발전이 아닌, 각 평가방법들의 연관성과 상대적인 수행력 차이도 분석하여 비교 선택할 수 있는, 신호 탐지 이론(signal detection theory) 및 Thurstonian 모델링을 활용한 통합적인 방법론이 개발되어 발전하였다. 이러한 이론에서는 인간의 두뇌에서 작용하는 인간의 감각 체계와 감각 평가 시에 활용되는 인지적 결정과정에 대한 연구가 진행되었고, 최근에는 선택된 감각 평가방법을 수행할 때 인간의 뇌에서 일어나고 있는 인지적 과정과, 이러한 요인들이 평가 결과에 어떻게 영향을 미치는지에 대해 많은 연구가 진행됨에 따라 상당한 지식이 축적되었다. 따라서 산업계에서는 제품 특성에 맞는 더 정확한 감각 품질 평가 및 관리 프로그램을 개발하고, 보다 정확한 비즈니스 의사 결정을 내리기 위해 이러한 최신 이론을 적절하게 활용하는 것이 필요하다.

본 논문에서는 식품 및 음료 제품들의 감각 품질에 대한 차이식별 측정을 가능하게 하는 최신 이론인 신호 탐지 이론과 Thurstonian 모델링을 소개하고, 식품의 감각적 및 기호적 품질관리를 위해 산업체에서 효과적으로 감각 차이식별 분석을 활용할 수 있는 방안을 논의하고자 한다. 또한 생수 제품의 효과적인 품질관리를 위해 최적화된 감각 품질관리 프로그램을 개발하여 운영했던 연구 사례를 소개하고자 한다.

본론

1. 감각·소비자과학의 발전과정과 감각 차이식별 분석 모델 및 이론의 필요성

과거 관능검사라 불리던 감각 평가가 과학으로

서 발전되기 위해서는 이론과 모델을 개발하는 것이 중요하다. 이론 없이는 감각 평가 결과의 분석을 통해서 신뢰할 수 있는 제품 품질과 소비자 반응을 예측할 수 없기 때문이다(Lee와 O'Mahony, 2007). 또한 기업마다 다양한 특성을 가진 제품들을 서로 다른 환경에서 생산하고 있고, 감각 평가 활용 시에 차별적인 제약 조건이 따르기 때문에, 기존의 감각 평가방법을 그대로 답습하는 것은 적절하지 않다. 따라서 감각·소비자과학의 배경 이론과 모델을 이해하고 자사 환경에 적합하게 응용하여 활용할 수 있는 전문성의 배양을 통해, 자사 제품에 최적 조건인 감각 평가 실험을 설계하고, 이에 따라 감각 품질관리를 실행하는 것이 필요하다.

이전에는 감각 및 소비자 평가 영역은 이론적인 배경 없이 평가방법들을 단순히 적용하는 경우가 일반적이었다. 감각 측정 자체의 근본적인 문제를 고려하지 않은 채, 수준 높은 통계 분석을 사용하기만 한다면 결과의 신뢰도가 높을 것이라고 낙관적으로 기대해왔다. 따라서, 더 나은 측정 방법을 적용하여 자료의 품질을 향상시키고, 감각 과정 및 감각 측정 과정을 보다 잘 이해하려고 노력하기보다는 손쉽게 얻을 수 있는 자료에 더 나은 통계 기법을 적용하는 데 더 많은 노력을 기울여 왔다. 그러나 최근, 감각·소비자과학 분야 연구원들은 감각 품질관리의 주요 문제가 감각 측정 자체와 관련되어 있음을 인식하여 통계 기법보다는 측정 모델에 큰 관심을 기울이기 시작했다(Ennis 등, 2014).

신호 탐지 이론과 Thurstonian 모델링은 현재 가장 진보한 감각 평가 이론이다(Green과 Swets, 1966; Macmillan과 Creelman, 2005; Thurstone, 1927a, 1927b). 신호 탐지 이론은 인간의 심리적 예민도를 이해하기 위한 이론적 방법론으로 정신 물리학자에 의해 개발되었다. 인간 감각의 예민도는 감각·소비자과학에서 사용되는 차이 검사 방법과 동일한 감각 차이식별 작업을 사용하여 측정되기 때문에 감각 차이 검사와 밀접한 관련이 있다. 정신 물리학



에서 이 접근법의 가장 중요한 장점은 반응 평가자의 주관적인 평가기준(일종의 응답 편향(response bias))에 의해 유동적으로 변하지 않는 절대적인 감각 예민도(sensitivity)의 측정을 가능하게 한다는 것이다. 한편, 감각 과학 분야에서 소개된 Thurstonian 모델링의 핵심 개념은 신호 탐지 이론의 주요 개념과 실질적으로 동일하다. 따라서, 신호 탐지 이론과 Thurstonian 모델링은 서로 다른 학문 분야를 토대로 하여 개별적으로 발전하여 왔으나, 핵심 개념이 공통되기에 신호 탐지 이론 및 Thurstonian 모델링이라 불리고 있다. 식품 분야에서 이 분야의 선구자는 University of California, Davis의 O'Mahony 교수와 버지니아의 지각 연구소(IFP; Institute for Perception)의 Ennis 박사와 Rousseau 박사이다. 현재는 이 이론이 널리 보급되어 감각·소비자과학 분야의 여러 응용통계 프로그램으로 활용이 가능하게 되었다. 예를 들어, IFP에서는 Thurstonian 모델링을 기반으로 차이식별 분석이 가능한 IFPrograms을 개발하여 제공하고 있고, Denmark Technical University의 Brockhoff 교수 연구팀은 무료 소프트웨어인 R 프로그램 내에서 구동 가능한 R 패키지 sensR을 개발하여 실용화하였다. 또한 청각에 대한 신호 탐지 이론 모델링 분야의 세계적인 권위자인 University of Auckland의 Hautus 교수는 SDT assistant 소프트웨어를 개발하여 제공 중에 있다. 이 중 R 패키지 sensR과 SDT assistant 프로그램은 무료로 배포되고 있는 통계 소프트웨어로써, 홈페이지에서 손쉽게 다운로드하여 사용할 수 있다.

식품 감각 과학과 심리학 간의 이러한 융합은 최근 식품 업계에서 제품 평가에 유리한 많은 평가방법의 개발을 이끌었다. 전 세계적으로 감각 과학자들 사이에서 이러한 과학분야의 융합을 통해서 감각·소비자과학 분야가 신뢰성을 지니면서도 활용성이 높은 식품과학으로 발전할 수 있다는 인식이 증가하고 있다. 이러한 이론은 감각 지각과 예민도 측정에 대해서 전통적인 관능검사 접근법과는 완전히

다른 견해를 제공하기 때문에 대부분의 식품 회사에서 신호 탐지 이론과 Thurstonian 모델링은 신속하게 활용되지 않았으나 점차적으로 이를 활용해서 얻을 수 있는 이점들이 소개되면서 국내에서도 동서식품, 롯데중앙연구소, 농심 등 대기업을 중심으로 신호 탐지 이론 및 Thurstonian 모델링에 기초한 감각 차이식별 분석 방법론이 활용되고 있다.

본 논문에서는 정신물리학에 익숙하지 않은 식품 분야 전공자들이 감각 측정 이론에 대해서 쉽게 이해할 수 있도록 수학적 이론의 설명이 아닌 간단한 도식화를 통한 개념적 설명을 제시하고자 한다.

2. 신호 탐지 이론(signal detection theory) 및 Thurstonian 모델링

2.1. 신호 탐지 이론 및 Thurstonian 모델링의 기본 개념

신호 탐지 이론은 자극에 대한 탐지율이 평가자의 예민도(sensitivity)와 평가기준(decision criteria)에 의해 결정된다는 이론으로, 예민도와 평가기준을 분리하여 개념화하기 때문에 평가기준에 의한 영향을 배제하고 예민도만을 측정할 수 있다. 이 이론에서는 평가자가 지각할 수 있는 자극, 즉 '신호'(signal)가 없을 때에도 인간의 신경계는 계속적으로 활동하고 있기 때문에 '배경 잡음'(background noise)이 존재함을 전제한다. '배경 잡음' 수준은 시간이 지남에 따라 무작위로 변화하지만 일반적으로 평균값을 중심으로 정상 분포를 나타낸다(Fig. 1 (a)). 즉, Fig. 1 (a)의 도식은 자극이 없을 때 '잡음' 지각의 확률 분포(probabilistic distribution)로 생각할 수 있으며, 여기서 수직축은 신경 자극의 발생 확률을 나타내고, 수평축은 신경 자극의 강도를 나타낸다. '배경 잡음'보다 강도가 더 높아서 지각되는 자극이 존재하면, 평가자가 지각하는 자극 수준에는 '신호' 수준이 이 '잡음' 수준에 추가된다. 즉, 신경계의 활동이 증가하여 Fig. 1 (b)에서 볼 수 있듯이 '신호 + 잡음'(signal + noise) 수준이 발생한

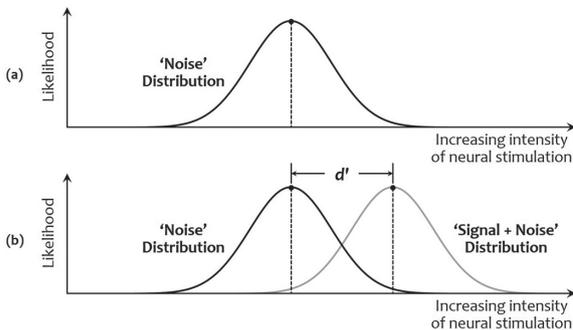


Fig. 1. (a) 'Noise' and (b) 'Signal+Noise' probability distributions in Signal Detection Theory

		Stimulus	
		Signal presented	Signal not presented
Response	"Yes"	Hit	False alarm
	"No"	Miss	Correct rejection

Fig. 2. Classification of the responses by presentation of the signal

다. '신호 + 잡음'에 나타난 확률 분포는 강도가 큰 방향으로 분포가 오른쪽으로 이동한다는 점을 제외하고는 '배경 잡음'의 형태와 동일하다. 신호 강도가 클수록 신경계의 활동이 증가한다는 의미이고, 이는 곧 평가자가 강한 자극으로 지각한다는 뜻이다. 신경 활동의 평균 증가(두 분포의 최고점 사이의 거리)를 분포의 변동수준(표준편차)으로 나눈 값이 d' (d -prime)으로 평가자의 예민도를 나타내는 정량 지표이다. 구체적으로 d' 은 자극이 제시되었을 때와 제시되지 않았을 때의 평가자의 응답 비율을 이용하여 계산될 수 있다. Fig. 2에 나타난 바와 같이, 평가자의 응답은 자극의 제시 여부에 따라 적중(hit), '누락'(miss), '오경보'(false alarm), 그리고 '정기각'(correct rejection)으로 분류될 수 있다. 기본 모델에서는 정규 분포로 각 확률 분포를 나타내기 때문에 다음의 수식에서 나타난 바와 같이 적중률(hit rate)과 오경보율(false alarm rate)을 이용하여 평가

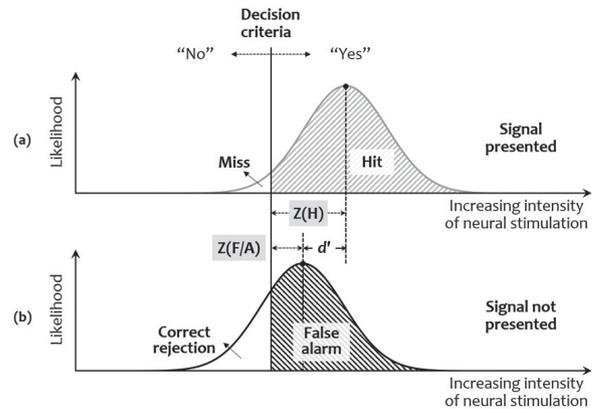


Fig. 3. How to calculate d' based on Signal Detection Theory

자의 예민도를 d' 으로 정량화 할 수 있다(Fig. 3).

$$d' = z(\text{hit rate}) - z(\text{false alarm rate})$$

평가자의 평가기준은 Fig. 3에 제시된 수평축을 따라 이동 가능하다. 평가자는 자신의 평가기준보다 더 큰 자극이 지각된다면 신호가 있다고 응답할 것이며, 평가기준에 미치지 못하는 경우에는 신호가 없다고 응답할 것이다. 만약, 평가자가 보다 엄격한 평가기준을 가지고 있다면, 예를 들어 Fig. 3에 제시된 평가기준이 수평축을 따라 우측으로 이동한다면, 적중률이 감소하게 되나, 동시에 오경보율 또한 감소하게 되므로 d' 은 일정한 값으로 계산된다. 즉, 서두에 언급했듯이, 신호 탐지 이론은 평가기준에 의한 영향을 배제한 채 예민도만을 측정할 수 있다.

식품을 섭취하는 상황을 생각해 보면, 인간이 지각하는 식품의 감각 자극은 항상 일정하지 않다. 식품 자체의 변동성, 이전에 섭취한 식품이 구강 내에 잔존하는 정도와 침 성분과 같은 구강 내의 상이한 환경, 그리고 감각 신호가 두뇌로 전달되는 과정에서 신경 활동 정도의 차이 등과 같은 다양한 요인들로 인하여 인간이 지각하는 식품의 감각적 특성의 강도는 다를 수 있다. 따라서, 인간 두뇌의 지각 변동성을 고려하는 신호 탐지 이론 및 Thurstonian

모델링이 식품의 감각 차이식별 분석에 효과적으로 적용될 수 있다. 감각 차이식별을 위한 다양한 평가 방법에서는 두 개 이상의 시료를 개별적으로 또는 순차적으로 다른 순서로 조합하여 평가하는 경우가 일반적이다. 예를 들어, 삼점 검사와 3-AFC 방법의 경우에는 매 평가 마다 두 개의 다른 시료가, 한 시료는 한 번 제시되고 다른 시료는 두 번 제시되어 순차적으로 평가된다. 이 때, 신호 탐지 이론 및 Thurstonian 모델링에서는 한 시료 자극은 ‘배경 잡음’ 수준으로 간주되고, 다른 시료 자극은 ‘신호’ 수준으로 간주된다. 즉, 두 개의 시료 자극에 의한 감각 지각 분포가 활용되고, 이 두 분포가 얼마나 서로 가깝게 존재하여 겹쳐지는지의 정도가 감각의 차이, 즉 d' 으로 나타나게 된다.

신호 탐지 이론 및 Thurstonian 모델링은 평가 요원이 평가를 수행할 때 사용하는 인지적 결정 전략(cognitive decision strategy)을 중요하게 고려하여 분석한다. 따라서 각 차이식별 평가방법마다 고유의 심리측정 함수를 포함한 분석 모델(psychometric function)을 연구하여 사용한다. 다양한 감각 차이식별 평가방법들 간에 연구된 심리측정 함수를 비교해보면, 이점 비교검사(2-AFC; 2-Alternative Forced Choice)와 3-AFC가 일·이점 검사와 삼점 검사보다 정답률이 높다. 즉 같은 시료에 대해서 동일한 평가자들을 활용해서 평가를 수행하였을 때, 상대적으로 일·이점 검사와 삼점 검사보다 2-AFC와

3-AFC를 사용했을 때 더 높은 정답률을 얻을 수 있다. 이러한 현상은 앞서 언급된 인지적 결정 전략의 차이에서 비롯된다. 예를 들어, 삼점 검사의 경우에는 평가자가 ‘거리 비교’(comparison of distances) 전략이라 불리는 인지 전략을 사용하여, 강도 축 상에서 어떤 두 시료 자극이 가장 비슷하게 느껴졌는지를 판단한다(Fig. 4 (a)). 이러한 모델은 신호 탐지 이론에서는 차이 평가 모델(differencing model)이라고 불리며, 이때 사용되는 인지 기준은 tau-criterion(τ -criterion)이라고 일컫는다. 3-AFC의 경우에는 평가자가 가장 강한 자극으로 지각한 시료를 선택한다. 맥주의 상위 표면에서 거품을 걷어내는 것과 같은 방식으로 가장 강한 자극을 전달한 시료가 선택되기 때문에 이것을 ‘스키밍’(skim off) 전략이라고 한다(Fig. 4 (b)). ‘스키밍’ 전략은 식품 시료 평가에 대한 Thurstonian 모델 적용 시에 개발된 이름으로, 신호 탐지 이론에서는 절대 평가 모델(independent judgment model)이라고 불리며, 이때 사용되는 인지 기준은 beta-criterion(β -criterion)이라고 일컫는다. Fig. 4에서 볼 수 있듯이 ‘스키밍’ 전략은 ‘거리 비교’ 전략보다 효율적이다. 즉, ‘스키밍’ 전략을 사용하면 ‘거리 비교’ 전략을 사용하여 비교하는 것보다 정답을 제시할 확률이 높아진다.

신호 탐지 이론 및 Thurstonian 모델링은 평가자의 의사 결정 과정을 고려하기 때문에 다양한 유형의 감각 평가방법으로부터 얻은 자료를 분석하여

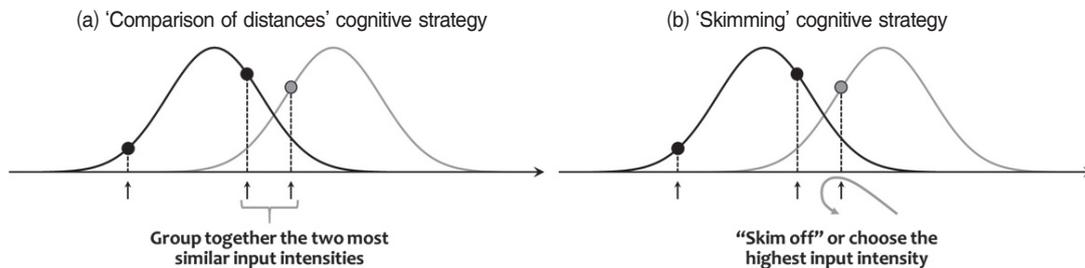


Fig. 4. Illustration of (a) the ‘comparison of distances’ strategy used in the duo-trio and triangle methods and (b) the ‘skimming’ strategy used in the paired comparison and 3-AFC methods

평가 제품 간의 감각 차이의 정도 또는 식별력의 정도를 d' 이라는 표준화된 단위를 활용하여 정량화할 수 있다. d' 은 0부터 무한대까지 가능한 수치이지만, 식품 시료에서는 0부터 4까지의 범위가 일반적이다. d' 은 앞서 설명한 바와 같이 감각 차이를 측정하는 데 사용된 평가방법의 영향을 받지 않는 감각 차이의 절대 척도이다. 따라서, 동일한 사람들이 5개의 다른 감각 차이식별 평가방법을 사용하여 동일한 시료들을 식별하는 과정을 수행한다면, 5개 다른 방법들의 차이점을 고려한 이론에 근거해서 모두 동일한 d' 값이 분석되어야 한다. 즉, 5가지 방법으로 얻은 결과를 d' 단위로 비교할 수 있다. 정답 비율(P_c ; proportion of correct responses)과 같은 간단한 정량지표, 또는 차이식별 비율(P_d ; proportion of discriminators)과 같은 기존 통계 분석방법을 사용하여서는 서로 다른 방법 간의 결과 비교는 가능하지 않다. 따라서 신호 탐지 이론 및 Thurstonian 모델링 연구법은 많은 잠재적 이점을 제공하므로 감각·소비자과학 분야에서 감각 및 기호 측정방법론의 향후 개발 방향으로 간주되고 있다.

2.2. 통계적 파워와 실효적 파워

산업체에서 감각 차이식별 검사를 사용하고자 할 때, 가장 기본적으로 고려해야 하는 요인이 평가방법의 통계적 파워(statistical power)이다. 통계적 파워는 차이식별 평가방법의 효율성을 나타내는 중요한 지표 중의 하나로, 제품 간의 차이가 존재할 때 그 차이를 감지하는 차이식별 평가방법의 평가력을 의미한다(Ennis, 1993; Ennis와 Jesionka, 2011; Lawless와 Heymann, 2010). 통계적 파워 측면에서, 비교하는 제품 간의 차이를 명시하는 ‘특정 감각 차이 비교 평가 방법’(specified method)이 차이를 명시하지 않는 ‘절대 감각 차이 비교 평가 방법’(unspecified method)보다 우수하다고 알려져 있다. 그러나 산업체에서의 품질관리를 목적으로 할

때, 제품 간의 차이를 명시하는 것은 결코 쉬운 일이 아니다. 예를 들어, 생산 과정에서 이미나 이취를 야기할 수 있는 성분이 혼입되었다고 가정했을 때, 감각 특성이 어떻게 변화할 것인지를 연구자가 선제적으로 특정하여(specify) 측정하는 것은 매우 어려운 일이며, 특정 감각 특성의 변화를 타겟으로 삼아 품질의 변화 여부를 측정한다 할지라도 항상 같은 감각 특성의 변화가 일어날 것이라고 확신할 수 없다. 따라서, 산업체에서는 ‘절대 감각 차이 비교 평가 방법’이 일반적으로 활용되고 있다.

차이식별 평가방법의 통계적 파워는 평가 요원이 감각 자극에 대하여 온전히 기억하며, 지각된 감각 자극이 왜곡됨 없이 두뇌로 명확히 전달됨을 가정하기 때문에, 실제 감각 평가를 수행하는 상황에서는 이론적으로 계산된 통계적 파워가 실현되지 못할 가능성이 존재한다. 따라서 최근에는 평가방법의 효율성을 논의하기 위해, 통계적 파워와 함께 실효적 파워(operational power)에 대한 논의가 활발하게 이루어지고 있다(Jeong 등, 2016, 2017; Kim과 Lee, 2012; Kim 등, 2015a; van Hout 등, 2011). 실효적 파워란 생리적 또는 인지적인 지각 과정과 관련하여 발생할 수 있는 요소들의 영향을 고려한 차이식별 평가방법의 평가력으로, 시료 제시 순서와 같은 요인들이 중요하게 작용한다. 사점 비교검사(unspecified tetrad test)의 경우, 삼점 검사, 일·이점 검사 등 다른 ‘절대 감각 차이 비교 평가 방법’에 비해 통계적 파워가 높기 때문에 우수한 방법으로써 제안되고 있으나(Ennis, 2012, 2013; Ennis와 Christensen, 2014; Ennis와 Jesionka, 2011; Ennis 등, 2014; Ishii 등, 2014; Rousseau와 Ennis, 2013), 평가 시에 맛봐야 하는 시료의 수가 많기 때문에 상대적으로 실효적 파워가 떨어질 수 있다(Jeong 등, 2016). 따라서, 신호 탐지 이론과 Thurstonian 모델링을 적용하여 제품 간의 감각 차이를 과학적으로 정량화하기 위해서는 이 이론이 적용될 수 있는 방



향으로 평가방법을 최적화하는 것이 매우 중요하다고 할 수 있다.

3. 차이식별 평가방법의 최적화: 생수 연구 사례

최근에는 생수 제품을 포함하여 다양한 음료제품들이 산업화되고 있다. 다른 식품 군과 마찬가지로 생수 및 음료 제품의 경우에도 소비자가 지각하는 맛은 기호적인 측면에서도 중요하지만 안전성 평가의 중요한 요소이기 때문에, 이러한 제품의 품질관리를 위해서는 감각적 품질 특성이 더욱 중요하게 평가되어야 한다(Zoeteman, 1980). 생수는 일반 식품과는 달리 특별한 맛이 없는, 즉 무미, 무취의 특성을 가지고 있는 특수한 성격의 식품군에 해당하기 때문에 소비자가 마시기에 성분이 안전하다고 하더라도, 감각적으로 불쾌한 맛이나 냄새가 느껴진다면 소비자는 해당 제품을 음용수로 받아들이지 않을 것이다(Zoeteman, 1980; WHO, 1971; McGuire, 1995). 즉, 소비자들은 생수에서 심미적으로 좋지 않은 맛이 느껴진다면 안전성에 문제가 없는 제품이라 하더라도 마시는 물로써 적합하지 않은 제품이라고 간주하기 때문에 보다 예민한 평가방법이 요구된다고 할 수 있다. 생수와 같이 예민하게 감지하기 쉬운 특수한 식품군에 적합한 차이식별 평가방법을 개발하기 위해 실제 산업체에서 신호 탐지 이론을 적용하여 개발한 사례에 대해서 소개하고자 한다.

식품 회사 농심에서는 직접 취수하여 판매하는 생수 제품의 효과적인 생산을 위해 새로운 공장을 설립하면서, 설비, 생산, 유통 및 보관 등 전반에 걸친 품질관리 방법을 수립하던 중 세부적인 관리 측면에서 품질관리를 위해 생수에 최적화된 감각평가를 어떻게 선정할 것인지를 논의해왔다. 기존에는 삼점 검사, 일·이점 검사, 2-AFC, 또는 3-AFC 등의 평가방법으로부터 얻은 결과를 이항분포에 기

초하여 통계적인 유의차를 확인하였으나, 이항 통계 분석방법은 기회확률만을 고려할 뿐 평가방법의 예민도 차이를 반영하지 않기 때문에 비교하고자 하는 제품 간의 감각 차이의 수준을 정량화 할 수 없다는 한계가 있었다. 이에 따라 신호 탐지 이론 및 Thurstonian 모델링에 기초한 분석방법의 필요성이 대두되었다. 또한 식품의 감각적인 특성은 매우 복잡적이기 때문에 안정적인 차이식별 결과를 얻기 위해서는 무엇을 평가하는지에 대한 평가 기준, 즉 제품 감각 품질 비교를 위한 평가 축의 안정화가 요구된다. 이를 위해 최근에 많은 연구가 이루어지면서도 산업적으로도 많이 활용되고 있는 리마인더(reminder) 시료를 제공하는 감각 차이식별 방법론(Kim 등, 2015a Shin 등, 2016; Jeong 등, 2016)을 생수 품질관리 평가에 적용할 수 있도록 연구를 진행하였다.

식품 시료에 대한 전통적인 감각 평가에는 앞서 설명한 2-AFC, 3-AFC, 일·이점 검사, 삼점 검사와 같은 강제선택법(forced-choice)을 사용한 평가방법들이 주로 사용되어져 왔으나, 신호 탐지 이론에서 사용하는 기본 평가방법은 이원 응답을 활용하는 A-부A 방법 형태이다. 이 A-부A 방법에서는 평가자가 제시된 여러 시료 중 하나를 선택하는 것이 아니라, 제시된 한 평가 시료(A시료 또는 A가 아닌 시료(=부 A))에 대해서 두 가지 응답 옵션('A이다' 또는 'A가 아니다') 중 하나를 선택하게 된다(Fig. 5 (a)). 이와 같은 형태의 평가를 수행하기 위해서는 평가자가 기준 시료인 A시료에 대해 알고 있어야 하므로, 기준 시료에 대한 친숙화 과정이 수반되는 것이 일반적이다. A-부A 검사방법에 리마인더가 융합된 검사방법인 리마인더 A-부A (A-Not AR; A-Not A with Reminder) 검사법은 기준 시료를 평가 시료와 함께 제공함으로써, 추가적인 친숙화 과정 없이도 평가를 안정적으로 수행할 수 있다는 장점이 있다(Fig. 5 (b)) (Jeong

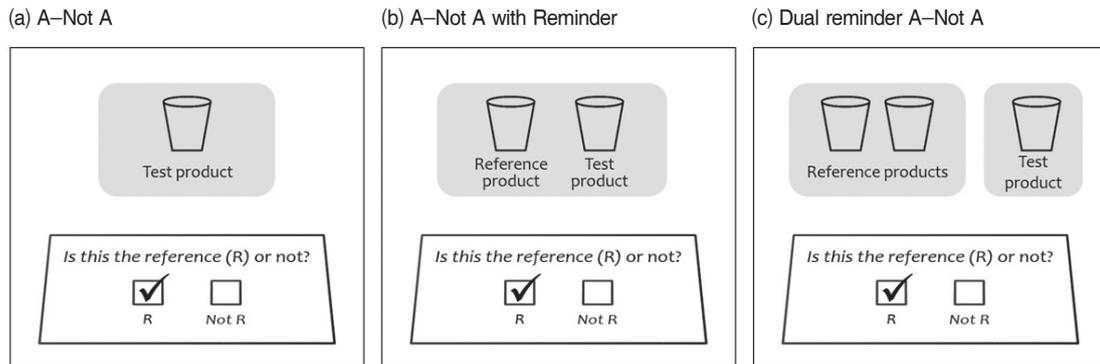


Fig. 5. Illustration of sample presentation and question for a test product for (a) A-Not A, (b) A-Not A with reminder, and (c) dual reminder A-Not A

등, 2016).

생수 품질 평가를 위해, 이와 같은 리마인더 디자인을 활용한 형태로써, 평가 제품을 검사하기 전에 기준 제품을 두 번 맛보게 하는 듀얼 리마인더 A-부A (DR A-Not A; dual reminder A-Not A) 평가 방법을 새로이 개발하였고(Fig. 5(c)), 강제선택법 (forced-choice) 중 통계적 파워가 높은 평가방법 중의 하나인 3-AFC와 평가 수행력을 비교함으로써 두 평가법 간의 효용성을 확인하였다(Moon 등, 2016). 비교 연구 결과, 평가 요원들은 DR A-Not A를 수행하는 동안 자연스럽게 기준 제품에 대하여 노출되기 때문에 추가적인 친숙화 과정 없이도 효율적인 beta-인지전략을 활용하는 것과, 최적화된 시료 순서로 구성되어 있기 때문에 평가의 실효적 파워를 높임으로써 3-AFC에 상응하는 평가 수행력을 나타냄을 확인하였다.

위와 같은 연구결과를 토대로 농심에서는 정기적인 품질관리 운영모델로 목적에 맞게 개발된 감각 차이식별 검사방법인 DR A-Not A를 생수 모니터링 평가방법으로 채택하였다. 지속적인 평가를 통해 제품 품질의 차이를 분석하여 수치화 할 수 있다는 점과 규격으로 정한 품질관리 수준(action standard) d' 을 통해 적합, 부적합을 판단할 수 있어 유용하게 활용하고 있다. 리마인더 디자인을 활용

한 차이식별 평가방법은 감각 차이가 작거나 중간 정도 차이가 나는 제품까지 ($d' = 0.5 - 2.0$) 모두 최적으로 평가가 가능하고, 상당히 적은 표본 수로도 제품 간의 차이를 정확하게 평가할 수 있기 때문에 소수의 훈련된 전문패널을 통해 평가를 수행하고 있으며 분석결과는 앞에서 소개된 R 패키지 sensR을 통해 d' 을 도출하고 있다. 또한 물의 품질이 항상 비슷한 수준으로 유지되는지 확인할 수 있도록 정기적인 모니터링을 운영하여 관리하고 있다. 이처럼 식품 군에 맞게 최적화된 차이식별 평가방법의 적용은 결과 신뢰도를 제고하므로 기업으로 하여금 잘못된 결과를 도출할 가능성을 줄여준다. 따라서, 기업들이 품질관리를 과학적이고 체계적으로 운영하기 위해서는 제품 특성에 맞게 최적화된 감각 차이식별 검사방법을 개발하고 적용할 수 있어야 할 것이다.

4. 체계적인 품질관리를 위한 기준 설정

4.1. 종합적 감각 차이 측정에 의한 품질관리

효과적인 품질관리를 위해서는 이를 위한 행동 기준(action standard)을 설정하는 것이 우선되어야 한다. 이 기준은 기업의 이익과 손실에 직결되기 때문에 매우 신중하고 정확하게 설정되어야 한다. 매

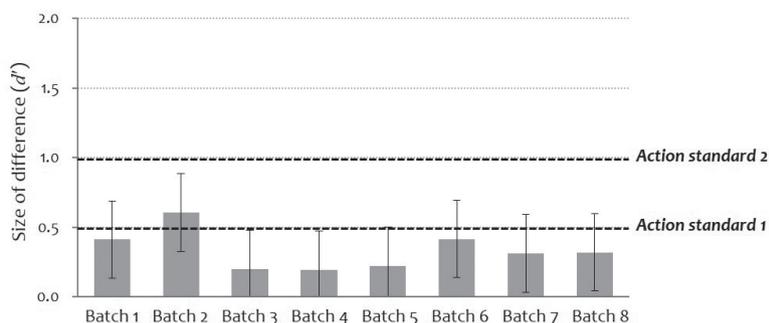


Fig. 6. Examples of action standard for food quality management

우 엄격하게 기준이 설정된 경우, 새로이 개발된 제품의 출시가 까다로울 것이며, 이 기준에 부합하지 못하여 새로 개발된 제품이 출시되지 못했을 때, 실제로 소비자들은 기존 제품과 새 제품 간의 차이를 지각하지 못할 가능성이 존재하고, 기업은 투자 비용에 대해 손실을 입게 된다. 반면, 기준을 느슨하게 설정한 경우에는 소비자들이 기존 제품과의 차이를 지각하여 새로이 개발된 제품의 구매를 거절하게 되어 이 또한 기업에 막대한 손실을 불러 일으킬 수 있다.

오랜 기간 동안 많은 식품 산업체에서 차이식별 분석은 이항 분포에 기초한 통계적 유의성 검증을 통해 제품의 차이 여부를 판명해 왔다. 이는 평가 방법의 예민도를 고려하지 않은 분석법이기 때문에 어떠한 평가방법을 사용했는지에 따라 유의적인 결과를 도출할 수도 그렇지 않을 수도 있으며, 제품 간의 감각 차이가 존재한다 하더라도, 실제로 소비자는 해당 차이를 지각하지 못할 수도 있다. 따라서, 단순히 통계적 유의성에 기초하여 제품 간의 감각 차이 여부를 판정함으로써 제품의 품질관리를 하는 것보다는 제품 간의 차이를 수치화함으로써 식품 비즈니스에서 보다 효과적으로 활용될 수 있다.

신호 탐지 이론 및 Thurstonian 모델링을 적용한 분석을 통하여 제품의 감각 품질 차이를 d' 이라는

지표로 정량화할 수 있으며, 이 표준화된 지표를 식품의 감각 품질 기준으로 활용하는 것이 가능하다. Kuesten(2001)은 산업체에서 활용될 수 있는 실제적인 품질관리 기준으로 $d'=1$ 의 기준을 적용할 수 있다고 제시하였고, Bi(2011)는 식품 회사의 전략에 따라 품질관리 기준을 다르게 설정할 수 있다고 하였으며, $d'=0.2, 0.4$, 그리고 1.0을 엄격한 기준, 보통의 기준, 허용 범위가 다소 넓은 기준으로 각각 정의하였다. 이와 같은 기준을 내부적으로 설정함으로써, 효과적으로 품질을 관리할 수 있으며, 이는 제품 군에 따라, 목적에 따라 다르게 적용할 수 있다 (Fig. 6).

4.2. 종합적인 감각 차이 측정값과 소비자 기호도/선호도 연관성에 근거한 품질관리

식품의 감각 품질관리를 위한 기준은 제품의 감각적 품질에 대한 소비자의 기호 변화로 정의하여 측정될 수 있다. 이러한 접근 방법은 실제 식품 소비자 행동은 감성적·기호적인 소비자 반응에 의해 예측되기 때문에 산업계에서도 실제 타겟 소비자를 대상으로 하여 연구하고 제품을 위한 품질관리에 반영해야 한다는 것이다. 예를 들어, 원가 절감을 목적으로 하여 기존의 원료 성분을 바꾼 경우를 생각해 보자. 소비자들은 기존 제품과 원가 절감 제품 간의 감각 차이를 감지할 수도 있고, 감지하지

못할 수도 있다. 후자의 경우, 즉 두 제품 간의 감각 차이가 매우 미미하여 소비자들이 그 차이를 감지하지 못한다면, 원가 절감 제품의 출시가 가능할 것이다. 그렇다면, 전자의 경우는 새로이 개발된 제품으로의 출시가 불가능한 것인가? 소비자들이 두 제품의 감각 차이를 감지하는 경우에는 “그 차이가 소비자가 허용할 수 있는 정도의 차이인가”와의 비교가 필요하다. 즉, 소비자들이 그 차이를 감지하지만, 그 차이 수준이 허용할 수 있을 정도라면 역시 원가 절감 제품으로의 출시가 가능하다고 할 수 있다. 이를 위해서는 소비자가 허용하는 감각 차이(allowable difference) 또는 소비자에게 있어 의미 있는 감각 차이(consumer-relevant sensory difference)를 설정하는 것이 매우 중요하다(Ennis 등, 2014; Ishii 등, 2014; van Hout, 2014).

만약 소비자가 두 제품 간에 감각 차이가 없다고 느낀다면, 두 제품 중 한 제품에 대한 선호 또한 발생하지 않을 것이다. 제품 간 감각 차이가 커질수록 한 제품을 선호할 가능성이 커지게 되며, 차이의 크기가 일정 수준(즉, 허용하는 감각 차이의 수준)을 넘어서게 되면, 한 제품에 대한 선호가 나타나게 될 것이다. 따라서, Ishii 등(2014)은 소비자가 두 제품 중 한 제품에 대하여 선호를 나타낸다면, 그 감각 차이의 크기를 소비자에게 있어 의미 있는 감각 차이로 정의할 수 있다고 보고하였다. 비슷한 맥락에서, Kim 등(2015b)은 소비자가 허용하는 감각 차이를 측정하기 위해 차이식별 분석과 선호도 검사를 융합한 평가방법을 제안하였다. 평가하고자 하는 두 제품 중 한 제품에 대하여 유의적인 선호도를 보이는 소비자를 규명하고, 이들의 제품 감각 차이 식별력을 소비자가 허용할 수 있는 감각 차이로 정의하였다. 다시 말해서, 유의적인 선호를 보이는 소비자 집단의 제품 감각 차이식별력을 d' 으로 계산했을 때, 계산된 d' 의 하위 신뢰 구간이 소비자가 허용할 수 있는 감각 차이 수준으로 정의될 수 있다

고 보고하였다.

결론

효율적인 식품 품질관리를 위해서는 이론적 체계에 근거한 차이식별 분석의 적용이 요구된다. 신호탐지 이론과 Thurstonian 모델링은 감각 차이의 정도를 표준화된 단위인 d' 으로 정량화 할 수 있게 하는 이론으로써, 다양한 감각 차이식별 평가방법들 고유의 심리측정 함수를 이용하여 본 수치가 산출되므로 다른 평가방법으로부터 도출된 값 간의 비교가 가능하다는 장점이 있다. 이 이론을 바탕으로 산출된 값의 신뢰성을 확보하기 위해서는 이 이론이 적용될 수 있는 방향으로 평가방법을 최적화하는 것이 매우 중요하다. 또한 체계적인 품질관리를 위해서는 제품(또는 공정)을 변화시켰을 때에 표준 제품과 비교하여 감각 및 기호적 품질에서 동등성(equality, matching 또는 similarity)을 유지할 수 있는 제품의 감각 차이의 정도를 결정하는 것이 매우 중요하다. 즉, 어느 수준의 감각 차이가 소비자의 행동에 영향을 미칠 것인지를 과학적으로 결정하는 것이 요구된다. 이러한 최대 감각 차이의 개념은 소비자가 허용하는 감각 차이(allowable difference) 또는 소비자에게 있어 의미 있는 감각 차이(consumer-relevant sensory difference)라고 정의되며, 기업에서 제품 품질 유지를 위한 중요한 지침으로 사용될 수 있다. 신호 탐지 이론과 Thurstonian 모델링을 활용한 감각 차이식별 분석 방법론은 산업체에서 제품 품질관리의 행동 기준(action standard)으로 소비자가 허용하는 감각 차이의 범주를 설정하고, 지속적으로 제품의 감각적 특성을 평가하면서 감각적 또는 기호적 품질 변화를 체계적으로 모니터링할 수 있는 전략적 연구법을 제공하며, 정보화시대에 발맞춰 식품 산업에서 자료의 통합과 활용을 위해서 앞으로 더욱 연구되고 발전되어야 할 분야이다.

참고문헌

- Bi J. Similarity tests using forced-choice methods in terms of Thurstonian discriminial distance, d' . *J. Sens. Stud.* 26: 151–157 (2011)
- Ennis DM. The power of sensory discrimination methods. *J. Sens. Stud.* 8: 353–370 (1993)
- Ennis JM. Guiding the switch from triangle testing to tetrad testing. *J. Sens. Stud.* 27: 223–231 (2012)
- Ennis JM. The year of the tetrad test. *J. Sens. Stud.* 28: 257–258 (2013)
- Ennis JM, Christense RHB. A thurstonian comparison of the tetrad and degree of difference tests. *Food Qual. Prefer.* 40: 263–269 (2014)
- Ennis JM, Jesionka V. The power of sensory discrimination methods revisited. *J. Sens. Stud.* 26: 371–382 (2011)
- Ennis JM, Rousseau B, Ennis DM. Sensory difference tests as measurement instruments: A review of recent advances. *J. Sens. Stud.* 29: 89–102 (2014)
- Green DM, Swets JA. *Signal detection theory and psychophysics*. John Wiley, Oxford, England. (1966)
- Ishii R, O'Mahony M, Rousseau B. Triangle and tetrad protocols: Small sensory differences, resampling and consumer relevance. *Food Qual. Prefer.* 31: 49–55 (2014)
- Jeong YN, Kang BA, Jeong MJ, Song MJ, Hautus MJ, Lee HS. Sensory discrimination by consumers of multiple stimuli from a reference: Stimulus configuration in A-Not AR and constant-ref. duo-trio superior to triangle and unspecified tetrad? *Food Qual. Prefer.* 47: 10–22 (2016)
- Jeong YN, van Hout D, Groeneschild C, Lee HS. Comparative identification method: Using 2-AFC strategy in constant-reference duo-trio for discrimination of multiple stimuli from a reference. *Food Qual. Prefer.* 62: 284–295 (2017)
- Kim IA, Kim HL, Cho HY, Lee HS. Optimal difference test sequence and power for discriminating soups of varying sodium content: DTFM version of dual-reference duo-trio with unspecified tetrad tests. *Food Res. Int.* 76: 458–465 (2015a)
- Kim MA, Lee HS. Investigation of operationally more powerful duo-trio test protocols: Effects of different reference schemes. *Food Qual. Prefer.* 25: 183–191 (2012)
- Kim MA, Sim HM, Lee HS. Affective discrimination methodology: Determination and use of a consumer-relevant sensory difference for food quality maintenance. *Food Res. Int.* 70: 47–54 (2015b)
- Kuesten CL. Sequential use of the triangle, 2-AC, 2-AFC, and same-different methods applied to a cost-reduction effort: Consumer learning acquired throughout testing and influence on preference judgements. *Food Qual. Prefer.* 12: 447–455 (2001)
- Lawless HT, Heymann H. *Sensory evaluation of food: Principles and practices* (2nd ed). Springer, NY, USA. pp. 83–84; 101–123 (2010)
- Lee HS, O'Mahony M. The evolution of a model: A review of Thurstonian and conditional stimulus effects on difference testing. *Food Qual. Prefer.* 18: 369–383 (2007)
- Macmillan NA, Creelman CD. *Detection theory: A user's guide* (2nd ed). Lawrence Erlbaum Associates Publishers, NJ, USA. (2005)
- McGuire MJ. Off-flavor as the consumer's measure of drinking water safety. *Water Sci. Technol.* 31: 1–8 (1995)
- Moon JW. Investigation of test performance of dual reminder A-Not A (DR A-Not A) in comparison to 3-AFC for sensory discrimination of drinking water. (Master dissertation). Ewha Womans University, Korea (2016)
- Peryam DR. Sensory difference tests. *Food Technol.* 12: 231–236 (1958)
- Peryam DR, Swartz VW. Measurement of sensory differences. *Food Technol.* 4: 390–395 (1950)
- Rousseau B, Ennis JM. Importance of correct instructions in the tetrad test. *J. Sens. Stud.* 28: 264–269 (2013)
- Shin HK, Hautus MJ, Lee HS. Unspecified duo-trio tests as powerful as the specified 2-AFC: Effects of instructions and familiarization procedures on cognitive decision strategies. *Food Res. Int.* 79: 114–125 (2016)
- Thurstone LL. A law of comparative judgment. *Psychol. Rev.* 34: 273–286 (1927a)
- Thurstone LL. Psychophysical analysis. *Am. J. Psychol.* 38: 368–389 (1927b)
- van Hout D. Measuring meaningful differences: Sensory testing based decision making in an industrial context; applications of Signal detection theory and Thurstonian modelling. (Doctoral dissertation). Erasmus Research Institute of Management, Rotterdam, Nederland (2014)
- van Hout D, Hautus MJ, Lee HS. Investigation of test performance over repeated sessions using signal detection theory: Comparison of three non-attribute specified difference tests 2-AFCR, A-Not A and 2-AFC. *J. Sens. Stud.* 26: 311–321 (2011)
- WORLD HEALTH ORGANIZATION (WHO). *International standards for drinking-water* (3rd ed). (1971)
- Zoeteman BCJ. Sensory assessment of water quality. Pergamon Press, Oxford, England. pp. 19–33 (1980)