

## 품질 개선 식품의 시장성 평가를 위한 차이-선호 융합 평가 방법론의 수행력 검증

김민아<sup>1</sup> · 이혜성<sup>2\*</sup>

<sup>1</sup>전남대학교 식품영양과학부 조교수, <sup>2</sup>이화여자대학교 식품생명공학과 교수

### Performance of the Difference-Preference Test for Marketability Evaluation of Reformulated Food Product

Min-A Kim<sup>1</sup> and Hye-Seong Lee<sup>2\*</sup>

<sup>1</sup>Assistant Professor, Division of Food and Nutrition, Chonnam National University, Gwangju 61186, Republic of Korea

<sup>2</sup>Professor, Department of Food and Biotechnology, Ewha Womans University, Seoul 03760, Republic of Korea

#### ABSTRACT

Successful product reformulation requires knowledge on consumer perception of the sensory differences between the original and reformulated products as well as an understanding of the impact such changes exert on consumer preference. In the present study, two consumer sensory tests simultaneously measuring consumers' group sensitivity and preferences (duo-trio difference-preference test and ABX preference-difference test) were compared to the paired-preference test with a no-preference option. Two pairs of commercial fermented milk products were evaluated by 200 consumers. Each consumer performed all three test methods in a balanced order. The results of *d'* analyses showed no difference in the preference responses among the three test methods. Regarding sensory discriminability, the duo-trio difference-preference test had higher test sensitivity than the ABX preference-difference test. The present findings suggest that the duo-trio difference-preference test is an efficient consumer sensory test method that enables consumer-based business decision making.

**Key words:** reformulation, difference-preference test, tried first bias, *d'*, signal detection theory

#### 서론

현대의 식품 업계는 인류의 건강을 위해 설탕, 소금, 지방 등 현대 식습관에서 과다하게 섭취하고 있는 성분들의 함량을 줄이고자 많은 노력을 기울이고 있다. 그러나 이들 성분은 식품의 맛 품질에 밀접하게 영향을 주고 있는 성분들이기에, 소량만 줄여도 소비자의 기호도에 큰 영향을 미칠 수 있다. 따라서, 성공적인 제품 개선을 위해서는 기존 제품과 개선 제품 간의 감각 차이를 정확하게 계측할 수 있는 예민한 감각평가 방법이 요구된다. 그러나 감각 차이가 지각된다 하더라도, 지각된 감각 차이가 반드시 선호 행동의 변화를 초래하는 것은 아닐 수 있다. 즉, 소비자가 제품 간의 감각 차이를 느끼더라도 그 차이의 정도가 허용할 만한 수준이면 해당 감각 차이는 선호 행동의 변화를 일으키지는 않을 것이다. 따라서, 감각·소비자 과학 분야에서는 소비자의 선호 행동 변화를 일으키지 않는 최대 감각 차이의 크기를 비즈니스

의사결정을 위한 '행동 기준(action standard)', '허용된 감각 차이(allowed difference)', '유사성 한계(similarity limit)' 등의 개념으로 일컬어 왔다(Bi J 2005; Bi J 2011; Jesionka V 등 2014; van Hout D 2014). 즉, 식품의 맛 품질에는 큰 영향을 주지 않는 범위 내에서 해당 성분들을 줄이거나 다른 성분으로 대체하는 것이 매우 중요하며, 이를 위해서는 제품 간의 감각 차이가 유의한지를 측정하는 것뿐 아니라 이러한 감각 차이가 선호에 영향을 미치는지를 측정하는 것이 비즈니스 의사결정을 위해 매우 중요한 정보를 제공할 수 있다.

최근 감각 차이와 선호 차이를 동시에 측정할 수 있는 감각평가법이 제안되어 왔다(Kim IA 등 2015; Kim MA & Lee HS 2015; Kim MA 등 2022). 일·이점 차이-선호 검사(duo-trio difference-preference test)는 그 한 종류로써, 일·이점 차이 검사에 선호도 질문을 융합한 형태이다(Kim MA & Lee HS 2015). 따라서, 시료의 제시 형태가 일·이점 차이 검사와 동일하기 때문에 한 번의 평가 시에 기준 시료와 두 개의 평가 시료를 포함하여 총 세 개의 시료를 제공한다. 평가원은 세 개의 시료를 순차적으로 맛 본 후, 두 개의 평가

\* Corresponding author : Hye-Seong Lee, Tel: +82-2-3277-6687, Fax: +82-2-3277-6687, E-mail: hlee@ewha.ac.kr

시료 중 기준 시료와 동일한 시료를 선택하고, 이어서 두 평가 시료 중 선호하는 시료를 선택한다. 전통적으로 감각 평가 분야에서는 평가 시료에 대한 감각 특성과 기호도를 동시에 평가하는 것을 지양해 왔다(Lawless HT & Heymann H 2010; Stone H 등 2020). 이는 감각 특성에 대한 평가는 보다 객관적이고 분석적인 사고를 요구하고, 기호 평가는 주관적이고 감정적인 사고를 요구하기에 두 질문을 동시에 하는 것은 소비자로 하여금 혼동을 일으킬 수 있다고 여겨왔기 때문이다. 소비자는 일상생활에서 식품을 즐기기 위해 섭취한다(Prescott J 1999). 따라서, 객관적인 식품의 감각 차이를 측정하는 목적이 아닌, 해당 식품에 대한 소비자의 차이식별력을 측정하기 위해서는 소비자 본연의 인식전략을 유도하여 평가를 진행하는 것이 보다 타당성 있는 결과를 도출할 수 있다. 이와 같은 맥락에서 소비자의 차이식별력을 평가할 때 기호를 개입시킨 평가방법들이 개발되어 왔고, 분석적인 사고를 유도한 평가방법들과 비교하였을 때 유의적으로 높은 차이식별력이 보고되었다(Frandsen LW 등 2007; Chae JE 등 2010; Kim MA 등 2014; Kim MA 등 2015, Kim MA 등 2016). 일·이점 차이-선호 검사 또한 선호 질문이 융합되지 않은 일·이점 차이 검사와 평가 수행력을 비교하였고, 일·이점 차이-선호 검사에서 선호 질문을 융합하였음에도 평가 수행력이 낮아지지 않았다는 결과가 보고되었다(Kim MA & Lee HS 2015). 이는 차이 검사와 선호도를 융합시켜서 평가하는 것이 가능할 수 있음을 나타내는 결과이다. 그러나 일·이점 차이-선호 검사의 평가 수행력은 차이식별 검사 질문이 융합되지 않은 기존의 선호도 검사와는 비교 평가되지 않았기 때문에 선호도 측면에서의 평가 수행력을 검증되는 것이 요구된다.

본 연구에서는 소비자들의 감각 예민도와 선호도를 동시에 측정할 수 있는 감각평가 방법론인 일·이점 차이-선호 평가 방법의 평가 수행력을 이점 선호도 검사(paired preference test)와 비교 평가함으로써 해당 평가방법의 효율성을 선호도 측면에서 입증하고자 하였다. 또한 감각 차이 질문과 선호 차이 질문의 순서에 따른 평가 수행력의 영향을 조사하기 위해 선호도를 먼저 평가하도록 하는 ABX 선호-차이 검사(ABX preference-difference test)를 새로운 변형 방법으로 소개하고 해당 방법의 평가 수행력을 함께 비교해보고자 하였다.

## 재료 및 방법

### 1. 재료

국내에 시판되고 있는 농후발효유 중 시장 점유율이 높은 A 제품과, 이와 감각특성이 유사한 두 종의 제품(B와 C)을

예비실험을 통해 선정하였다. 본 연구에서는 A 시료와 다른 두 시료 B와 C 사이의 감각 및 선호 차이, 즉 A-B 시료 쌍과 A-C 시료 쌍 간의 차이가 평가되었으며, 평가 방법의 효율성을 비교하고자 하는 본 연구의 목적을 위해 본고에서 해당 시료들은 A, B, C로 표기되었다. 냉장 보관 중이던 시료는 평가 직전에 투명한 일회용 플라스틱 컵 용기(65 mL)에 30 mL씩 담아 제공되었고(5±1 °C), 기준 시료를 제외한 모든 평가 시료는 무작위로 추출된 세 자리 난수로 표기되었다.

### 2. 소비자 패널

발효유 제품을 일주일에 한 번 이상 규칙적으로 마시는 소비자 200명(남성 54명, 여성 146명)이 본 연구에 참여하였다. 해당 연구는 이화여자대학교에서 진행되었으며, 실험 일주일 전부터 온라인과 오프라인 게시판에 홍보물을 공고함으로써 검사 참여를 안내하였다. 검사에 참여한 후 소정의 보상이 지급되었다.

### 3. 평가 방법

소비자 평가원은 세 가지 검사 방법, 즉, 이점 선호도 검사, 일·이점 차이-선호 검사, 그리고 ABX 선호-차이 검사를 모두 수행하였다. 선호도만을 평가하도록 하는 이점 선호도 검사와는 달리 두 종류의 차이-선호 융합 검사 방법은 선호도와 함께 차이식별력을 평가하도록 하는 검사이므로, 평가 방법에 대한 소비자 평가원의 혼동을 막기 위해서 소비자 평가원은 무작위로 두 그룹으로 분류되었고, 이 중 한 그룹은 이점 선호도 검사를 가장 먼저 수행하고 다른 한 그룹은 이점 선호도 검사를 가장 나중에 수행하도록 하였다. 각 그룹 내에서 두 종류의 차이-선호 융합 검사 방법은 균형된 순서로 수행되었다(Table 1). 모든 평가원은 각 평가 방법에 대하여 총 2회의 평가를 수행하였으며, 이는 A-B 시료 쌍에 대한 평가 1회와 A-C 시료 쌍에 대한 평가 1회로, 두 시료 쌍에 대한 평가 또한 균형된 순서로 수행되었다. 즉, 모든 소비자 평가원은 총 6회의 평가를 수행하였고, 약 5~10분 정도 소요되었다.

각 검사 방법의 구체적인 절차는 다음과 같다.

이점 선호도 검사 방법은 매 평가 시에 두 개의 평가 시료가 제공되고, 제공된 두 시료 중 선호하는 시료를 선택하는 방법으로 '선호도 없음' 항목이 함께 제공되었다(Fig. 1A). 각 시료 쌍에 대하여 모든 가능한 시료 제시 순서가 제공되었으며, 따라서 A-B 시료 쌍의 경우, <A-B>와 <B-A> 시료 제시 순서 모두가 균형되게 평가되었고, A-C 시료 쌍의 경우, <A-C>와 <C-A> 시료 제시 순서 모두가 균형되게 평가되었다.

일·이점 차이-선호 검사 방법은 일·이점 검사와 동일하

Table 1. The overview of the experimental design

Group	Session 1	Session 2	Session 3
1-1 (N=50)	Paired-preference test	Duo-trio difference-preference test	ABX preference-difference test
1-2 (N=50)	Paired-preference test	ABX preference-difference test	Duo-trio difference-preference test
2-1 (N=50)	Duo-trio difference-preference test	ABX preference-difference test	Paired-preference test
2-2 (N=50)	ABX preference-difference test	Duo-trio difference-preference test	Paired-preference test

계 기준 시료와 두 평가 시료가 제공되며, 두 평가 시료 중 기준 시료와 동일한 시료를 먼저 선택하고, 이어서 두 평가 시료 중 선호하는 시료를 선택하도록 하였다(Fig. 1B). 선호도 응답에는 ‘선호도 없음’ 항목이 포함되었다. 일·이점 차이-선호 검사 방법에서도 각 시료 쌍에 대하여 모든 가능한 시료 제시 순서가 균형되게 제시되었다. 즉, A-B 시료 쌍의 경우, A가 기준시료로 제시되는 <A:AB>, <A:BA> 제시 순서와 B가 기준시료로 제시되는 <B:AB>, <B:BA> 제시 순서 모두가 균형되게 평가되었다.

ABX 선호-차이 검사 방법은 일·이점 검사와 동일하게 기준 시료와 두 평가 시료가 함께 제공되나, 기준 시료가 제일 마지막에 위치한다. 이를 정신물리학 분야에서는 ABX-매칭 차이 검사라 부르며(Macmillan NA & Creelman CD 2005), 따라서 본 연구에서는 ABX 선호-차이 검사 방법이라 명명되었다. 즉, 두 평가 시료가 먼저, 그리고 기준 시료가 가장 나중에 제공되며, 두 평가 시료 중 선호하는 시료를 먼저 선

택하도록 한 후, 마지막에 제공된 기준 시료가 앞서 맛본 두 평가 시료 중 어떤 시료와 동일한지를 선택하도록 하였다(Fig. 1C). 앞선 방법들과 동일하게 선호도 평가 시에 ‘선호도 없음’ 항목이 제공되었다. ABX 선호-차이 검사 방법도 일·이점 차이-선호 검사 방법과 동일하게 A 시료가 기준시료로 제시되는 두 경우(<AB:A>, <BA:A>)와 B 시료가 기준시료로 제시되는 두 경우(<AB:B>, <BA:B>) 모두 균형되게 평가되었다.

본 연구에서 비교하고자 하는 세 종류의 평가 방법은 모두 한 번의 평가를 위해 두 개 또는 세 개의 시료를 비교해야 하는 검사 방법이다. 따라서 한 번의 평가 내에서 제공되는 시료의 제시 순서에 따른 응답의 차이를 조사하기 위하여, 소비자 패널이 각 평가 내에서 제공된 시료들을 왼쪽부터 순차적으로 맛보도록 하였으며 한 번 맛본 시료를 다시 맛볼 수 없도록 하였다(Kim MA & Lee HS 2012; Kim IA 등 2015; Kim MA & Lee HS 2015). 또한 입을 헹글 수 있도록

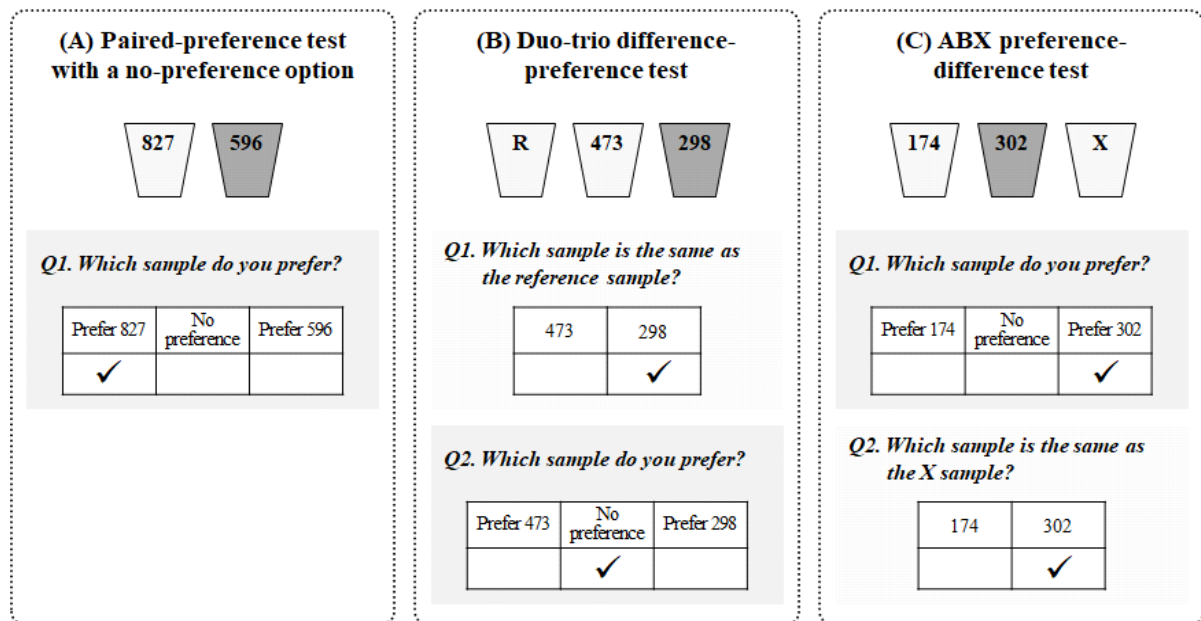


Fig. 1. Illustration of the sample presentation and the response format for (A) the paired-preference test with a no-preference option, (B) the duo-trio difference-preference test, and (C) the ABX preference-difference test.

제공된 입가심 물(20±2°C)은 각 평가를 시작하기 전에만 허용되었고 한 평가 내에서 시료를 맛보는 사이에는 허용되지 않았다. 감각평가 시에는 이전에 맛본 시료의 감각이 다음 시료에 영향을 줄 수 있기 때문에 입행굽을 수행하는 것이 일반적이나, 차이식별 검사에서는 한 번의 평가 내에 비교되는 시료 사이에 입행굽을 수행하는 것은 오히려 시료 간의 비교를 방해해서 평가 수행력을 감소시킬 수도 있다(Lee YM 등 2009).

#### 4. 통계 분석

세 평가 방법 간의 평가 시료에 대한 선호 응답 빈도는 카이 제곱 검정을 통해 유의적 차이 여부를 분석하였다. 또한 각 평가 방법으로부터 도출된 감각 및 선호 차이 정도를 지표화하기 위하여 신호 탐지 이론과 Thurstonian Modeling을 적용하여  $d'$ 으로 정량화하였다.  $d'$ 은 차이 정량 지표로써, 두 시료 간의 감각 차이 또는 선호 차이를 나타낸다. 먼저, 감각 차이  $d'$ 은 일·이점 차이-선호 검사와 ABX 선호-차이 검사로부터 수집된 차이식별 검사 결과로부터 계산되었다. 모든 평가원이 각 시료 쌍에 대해 반복 없이 1회씩 수행하였기 때문에 모든 평가원의 결과를 합하여 그룹  $d'$ 을 산출하였으며, 거리 비교 전략을 사용하는 일·이점 차이 검사를 가정하여 계산되었다(Macmillan NA 등 1977; Macmillan NA & Creelman CD 2005). 선호 차이  $d'$ 은 ‘선호도 없음’ 응답을 포함하여 각 시료 쌍에 대한 선호도 응답 결과를 분석하는 것으로 Thurstonian 2-AC model을 근거로 계산되었다(Glenn WA & David HA 1960; Braun V 등 2004; Christensen RHB 등 2012; Ennis DM & Ennis JM 2012). 도출된 감각 차이 또는 선호 차이  $d'$ 에 대한 유의성은 카이 제곱 검정을 활용하여 분석되었고(Bi J 등 1997), 모든 분석은 R 프로그램 내에 sensR 패키지를 활용하였다(Christensen RHB & Brockhoff

PB 2015; R Core Team 2021).

## 결과 및 고찰

### 1. 선호도 응답에 대한 차이식별 검사의 영향

선호도 검사 시에 차이식별 검사를 함께 수행하는 것이 선호도 응답에 영향을 미치는지를 확인하기 위해 이점 선호도 검사 방법과 감각 예민도 차이 및 선호도 검사를 동시에 평가하는 두 종류의 검사 방법(일·이점 차이-선호 검사, ABX 선호-차이 검사)의 선호도 응답 결과를 Table 2에 제시하였다. A-B 시료 쌍에 대한 선호도 검사 결과, 세 검사 방법 모두에서 선호 차이  $d'$ 이 유의적으로 0보다 높게 예측되었다. 선호 차이  $d'$ 이 양의 값을 나타내는 것은 A 시료에 대한 선호도 응답율이 더 높다는 것을 의미하며, 해당 값이 0과 유의적으로 차이 난다는 것은 이와 같은 응답율의 차이가 유의미한 차이라는 것을 의미한다. 다시 말해서, A-B 시료 쌍에 대한 선호 차이  $d'$  결과로부터 A 시료에 대한 선호도가 유의적으로 더 크다는 것을 알 수 있으며, 세 방법 모두에서 동일한 결과가 도출되었다. A-C 시료 쌍에서도 세 검사 방법 모두에서 0보다 유의적으로 큰  $d'$ 이 도출되었으며, 이 또한 A 시료에 대한 선호도가 C 시료보다 유의적으로 크다는 것을 나타낸다. 선호도 응답 비율을 봤을 때, 세 검사 방법 모두에서 ‘선호도 없음’ 응답 비율은 8% 내외로 낮게 나타났으며 (Table 2). 종합적으로, 선호도 응답 비율과 선호 차이  $d'$  모두를 고려하였을 때, 이점 선호도 검사 방법과 감각 차이식별 검사를 융합한 차이-선호 (또는 선호-차이) 방법 간의 선호도 응답에 있어서 유의적인 차이를 보이지 않았다( $\chi^2$  statistics,  $p>0.05$ ). 이와 같은 결과는 선호도 검사에 차이식별 검사를 융합하는 것이 선호도 응답에 영향을 미치지 않을 수 있음을 나타내는 바이다. 즉, 일·이점 차이식별 검사에 선

Table 2. Preference difference  $d'$  estimates (and its 95% CI) for the three comparison methods

Test method	Sample pair	Preference responses (and %)			Total	$d'$ (95% CI)
		Prefer 'A'	No preference	Prefer 'B' (or 'C')		
Paired-preference test	A vs. B	126 (63.00)	11 ( 5.50)	63 (31.50)	200 (100.00)	<b>0.58 (0.33, 0.82)</b>
	A vs. C	131 (65.50)	20 (10.00)	49 (24.50)	200 (100.00)	<b>0.77 (0.53, 1.02)</b>
Duo-trio difference-preference test	A vs. B	112 (56.00)	11 ( 5.50)	77 (38.50)	200 (100.00)	<b>0.31 (0.07, 0.56)</b>
	A vs. C	129 (64.50)	21 (10.50)	50 (25.00)	200 (100.00)	<b>0.74 (0.50, 0.99)</b>
ABX preference-difference test	A vs. B	133 (66.50)	7 ( 3.50)	60 (30.00)	200 (100.00)	<b>0.67 (0.42, 0.92)</b>
	A vs. C	130 (65.00)	17 ( 8.50)	53 (26.50)	200 (100.00)	<b>0.72 (0.47, 0.96)</b>

<sup>1)</sup> Values in bold indicate that 95% CIs did not contain zero.

호도 질문을 융합했을 때 차이식별 검사의 감각 예민도가 저하되지 않았고(Kim MA & Lee HS 2015), 본 연구에서 선호도 응답 또한 차이식별 검사의 영향을 받지 않았기 때문에 차이식별 검사와 선호도 검사를 융합한 검사 방법은 효율적인 검사방법으로써 제안될 수 있다.

이점 선호도 검사에서 자주 언급되는 편향 요인 중 하나인 첫 시료 효과(tried first bias)는 가장 먼저 맛본 시료를 선호하는 경향을 의미한다(Greenberg A 1958; Greenberg A & Collins S 1966; Moskowitz HR 등 2012). 식품 평가의 경우, 제시되는 시료를 순차적으로 맛봐야 하기 때문에 이와 같은 제시 순서와 관련한 편향 요인은 감각 평가에서 자주 보고되고 있다. 따라서, 각 평가 방법 내에서 시료 제시 순서에 따른 선호도 응답의 안정성을 비교하기 위해서 처음 제시된 시료의 종류에 따라 선호도 응답을 구분하여 Table 3에 나타냈다. 일·이점 차이-선호 검사에서 B 시료가 첫 시료로 제공된 경우를 제외하고는 모든 경우에서 유의적으로 0보다 큰  $d'$ 이 나타났으며, 이는 Table 2에서 본 바와 같이 A 시료에 대한 선호도가 유의적으로 크다는 것을 의미한다. 이와 같은 A 시료에 대한 선호도는 A 시료가 첫 시료로 제공되었을 때 B 시료가 첫 시료로 제공되었을 때보다 유의적으로 높게 관측되었다( $\chi^2$  statistics=17.77,  $p<0.01$ ). 특히, 일·이점 차이-선호 검사에서는 B 시료가 첫 시료로 제공되었을 때 B 시료

에 대한 선호도 응답이 증가하여 A 시료와 B 시료 간의 유의적인 선호도 차이가 상쇄되는 결과를 나타냈다. Schutz HG & Bradley JE(1954)는 삼점 검사에 선호도 검사를 융합시킨 검사에서 가장 다르다고 응답한 시료를 덜 선호하는 경향인 외 시료 효과(odd sample bias)가 발생할 수 있음을 보고하였다. 외 시료 효과는 일·이점 차이-선호 검사에서 기준 시료를 선호하는 효과로 나타날 수 있으며(Kim MA & Lee HS 2015), 기준 시료가 처음 제공되는 일·이점 차이-선호 검사에서는 첫 시료 효과와 동일한 결과를 나타내어 궁극적으로 첫 시료인 기준 시료에 대한 선호 결과를 높일 수 있다. 따라서, Kim MA & Lee HS(2015)의 연구에서는 일·이점 차이-선호 검사에서 선호도 응답 시에 기준 시료에 의한 영향을 상쇄시키기 위하여, 한 명의 평가원이 평가하고자 하는 두 시료를 모두 기준 시료로 하여 평가를 수행하도록 하는 방법을 제안하였다. 그러나 본 연구에서는 개인차에 의한 영향을 최소화하기 위하여 모든 평가원이 세 평가 방법을 모두 수행하도록 실험을 설계하였고, 따라서 각 시료 쌍에 대해서는 한 번의 평가만을 수행하도록 하였다. 즉, 본 연구 결과에서는 처음 맛보는 시료가 선호 응답에 매우 큰 영향을 줄 수 있음이 다시 한 번 확인되었으며, 이는 감각 평가 시에 모든 평가 제시 순서가 균형되게 제시되어야 할 필요성을 강조하는 결과이다. 이와 같은 편향 요인은 평가원마다도 다르게

**Table 3. Preference difference  $d'$  estimates (and its 95% CI) for tests using each type of the first tasted sample for the three comparison methods**

Test method	Sample pair	First tasted sample	Sample presentation	Preference responses (and %)			Total	$d'$ (95% CI)
				Prefer 'A'	No preference	Prefer 'B' (or 'C')		
Paired-preference test	A vs. B	A	<AB>	66 (65.35)	8 ( 7.92)	27 (26.73)	101 (100.00)	<b>0.72 (0.37, 1.07)</b>
		B	<BA>	60 (60.61)	3 ( 3.03)	36 (36.36)	99 (100.00)	<b>0.44 (0.09, 0.79)</b>
	A vs. C	A	<AC>	67 (67.68)	8 ( 8.08)	24 (24.24)	99 (100.00)	<b>0.82 (0.47, 1.18)</b>
		C	<CA>	64 (63.37)	12 (11.88)	25 (24.75)	101 (100.00)	<b>0.72 (0.38, 1.07)</b>
Duo-trio difference-preference test	A vs. B	A	<A:AB> and <A:BA>	66 (66.00)	5 ( 5.00)	29 (29.00)	100 (100.00)	<b>0.68 (0.33, 1.04)</b>
		B	<B:AB> and <B:BA>	46 (46.00)	6 ( 6.00)	48 (48.00)	100 (100.00)	-0.04 (-0.37, 0.30)
	A vs. C	A	<A:AC> and <A:CA>	72 (71.29)	10 ( 9.90)	19 (18.81)	101 (100.00)	<b>1.02 (0.67, 1.39)</b>
		C	<C:AC> and <C:CA>	57 (57.58)	11 (11.11)	31 (31.31)	99 (100.00)	<b>0.48 (0.14, 0.82)</b>
ABX preference-difference test	A vs. B	A	<AB:A> and <AB:B>	69 (68.32)	0 ( 0.00)	32 (31.68)	101 (100.00)	<b>0.67 (0.30, 1.06)</b>
		B	<BA:A> and <BA:B>	64 (64.65)	7 ( 7.07)	28 (28.28)	99 (100.00)	<b>0.67 (0.32, 1.03)</b>
	A vs. C	A	<AC:A> and <AC:C>	70 (69.31)	11 (10.89)	20 (19.80)	101 (100.00)	<b>0.96 (0.61, 1.31)</b>
		C	<CA:A> and <CA:C>	60 (60.61)	6 ( 6.06)	33 (33.33)	99 (100.00)	<b>0.49 (0.15, 0.84)</b>

<sup>1)</sup> Values in bold indicate that 95% CIs did not contain zero.

영향을 미칠 수 있기 때문에 한 명의 평가원이 가능한 평가 제시 순서를 모두 평가하도록 하는 것이 바람직할 수 있다.

반복 평가에 따른 시료 노출에 의한 선호도 응답의 변화 양상을 조사하기 위하여 이점 선호도 검사를 가장 먼저 수행한 소비자 그룹과 가장 나중에 수행한 소비자 그룹의 이점 선호도 응답 결과를 비교 평가하였다(Table 4). 시료 제시 순서에 따라 응답 결과를 분리하여 보았을 때, 이점 선호도 검사를 가장 먼저 수행한 소비자 그룹에서만 유의적인 첫 시료 효과가 나타났다. 즉, A-B 시료 쌍에서는 A 시료가 먼저 제시되었을 때 A 시료에 대하여 유의적으로 높은 선호도가 관찰되었으나 B 시료가 먼저 제시되었을 때는 두 시료 간의 유의적인 선호 차이가 관찰되지 않았다. A-C 시료 쌍에서도 동일한 양상이 확인되었다. 이와 같은 결과는 이점 선호도 검사의 경우, 처음 평가 시에 첫 시료 효과에 의해 선호도 응답의 안정성이 낮을 수 있음을 나타내며, 궁극적으로는 평가 시료에 대한 선호 차이를 잘못 예측하는 결과를 야기할 수 있다. A-B 시료 쌍에 대한 그룹별 선호 차이  $d'$  결과를 살펴보면, 이점 선호도 검사를 가장 먼저 수행한 소비자 그룹은 0.32였고, 이점 선호도 검사를 가장 나중에 수행한 소비자 그룹은 0.85였다. 특히, 이점 선호도 검사를 먼저 수행한 그룹의 선호 차이  $d'$ 이 0과 유의적으로 다르지 않은 값을 나타냈다. A-C 시료 쌍에 대한 결과 또한 이점 선호도 검사를 먼저

수행한 소비자 그룹과 나중에 수행한 소비자 그룹 각각의 선호 차이  $d'$ 은 0.49와 1.09로 두 시료 쌍 모두에서 이점 선호도 검사를 나중에 수행했을 때 선호 차이  $d'$ 이 높은 경향을 보였다. 즉, 첫 시료 효과에 의한 선호도 응답에의 영향은 궁극적으로 평가 시료에 대한 선호 차이를 과소평가할 수 있고, 따라서 보다 정확한 선호 차이를 측정하기 위해서는 선호도 검사에서도 반복 평가가 필요할 수 있음을 나타내는 바이다. 이와 같은 반복 평가에 따른 선호 응답의 안정성은 차이-선호 융합 검사방법에 대해서도 추후 확인될 필요가 있다.

## 2. 두 종류의 감각 차이-선호 융합 검사방법의 차이식 별력 비교

감각 예민도 차이와 선호도 차이 질문의 순서를 달리 구성하여 개발된 일·이점 차이-선호 검사와 ABX 선호-차이 검사간의 감각 차이식별 예민도를 비교 평가하기 위하여 A-B 시료 쌍과 A-C 시료 쌍에 대한 감각 차이 정도를  $d'$ 으로 정량화하였다(Table 5).  $d'=1$ 은 이점 비교 검사(2-alternative forced choice; 2-AFC)에서 76%의 정답율을 보이는 값으로 한계값 정도의 감각 차이를 의미한다. 본 연구에서는 A-B 시료 쌍과 A-C 시료 쌍 모두  $d'$ 으로 2 이상의 감각 차이가 나타났다, 이는 A 시료와 다른 두 시료 간의 감각 차이가 상당히 크다는 것을 의미한다. 일·이점 차이-선호 검사와 ABX

**Table 4. Preference difference  $d'$  estimates (and its 95% CI) of the paired-preference test which were performed at the first and last**

Group	Sample pair	Sample presentation	Preference responses (and %)			Total	$d'$ (95% CI)
			Prefer 'A'	No preference	Prefer 'B' (or 'C')		
1 (who performed the paired-preference test first)	A vs. B	<AB>	31 (60.78)	2 ( 3.92)	18 (35.29)	51 (100.00)	<b>0.46 (0.02, 0.95)</b>
		<BA>	27 (55.10)	0 ( 0.00)	22 (44.90)	49 (100.00)	0.18 (-0.32, 0.68)
		Pooled	58 (58.00)	2 ( 2.00)	40 (40.00)	100 (100.00)	0.32 (-0.02, 0.67)
	A vs. C	<AC>	31 (62.00)	4 ( 8.00)	15 (30.00)	50 (100.00)	<b>0.59 (0.10, 1.08)</b>
		<CA>	28 (56.00)	5 (10.00)	17 (34.00)	50 (100.00)	0.40 (-0.08, 0.88)
		Pooled	59 (59.00)	9 ( 9.00)	32 (32.00)	100 (100.00)	<b>0.49 (0.15, 0.83)</b>
2 (who performed the paired-preference test last)	A vs. B	<AB>	35 (70.00)	6 (12.00)	9 (18.00)	50 (100.00)	<b>1.02 (0.52, 1.53)</b>
		<BA>	33 (66.00)	3 ( 6.00)	14 (28.00)	50 (100.00)	<b>0.70 (0.21, 1.21)</b>
		Pooled	68 (68.00)	9 ( 9.00)	23 (23.00)	100 (100.00)	<b>0.85 (0.50, 1.21)</b>
	A vs. C	<AC>	36 (73.47)	4 ( 8.16)	9 (18.37)	49 (100.00)	<b>1.08 (0.56, 1.62)</b>
		<CA>	36 (70.59)	7 (13.73)	8 (15.69)	51 (100.00)	<b>1.10 (0.60, 1.61)</b>
		Pooled	72 (72.00)	11 (11.00)	17 (17.00)	100 (100.00)	<b>1.09 (0.73, 1.45)</b>

<sup>1)</sup> Values in bold indicate that 95% CIs did not contain zero.

**Table 5. Sensory difference  $d'$  estimates (and its 95% CI) for the duo-trio difference-preference test and the ABX preference-difference test**

Test method	Sample pair	No. of correct responses (out of 200)	$d'$ (95% CI)
Duo-trio difference-preference test	A vs. B	174	<b>2.94 (2.47, 3.44)</b>
	A vs. C	178	<b>3.15 (2.65, 3.68)</b>
ABX preference- difference test	A vs. B	156	<b>2.22 (1.81, 2.63)</b>
	A vs. C	164	<b>2.50 (2.08, 2.94)</b>

<sup>1)</sup> Values in bold indicate that 95% CIs did not contain zero.

선호-차이 검사 간의 감각 차이식별 예민도를 비교하였을 때, A-B 시료 쌍의 감각 차이는  $d'=2.94$ 와  $2.22$ , A-C 시료 쌍의 감각 차이는  $d'=3.15$ 와  $2.50$ 으로, 카이 제곱 검정 결과 일·이점 차이-선호 검사에서의 평가 예민도가 유의적으로 높게 나타났다( $\chi^2$  statistics=9.40,  $p<0.001$ ). 선호도 질문을 동일·비동일 검사에 융합시켜 평가 예민도를 비교했던 Kim IA 등(2015)의 연구에서는 선호도를 먼저 수행했던 선호-차이검사가 동일·비동일 차이검사를 먼저 수행했던 차이-선호 검사보다 더 예민한 평가 결과를 보였던 것과 상반되는 결과이다. 두 연구 결과 간의 차이는 선호 검사와 융합한 차이 검사의 특성 때문일 수 있다. 본 실험에서 사용한 일·이점 검사와 ABX 매칭 검사는 평가 시료 이외에 기준 시료를 제공하며, 따라서 평가 예민도에 기준 시료가 상당히 중요하게 영향을 미칠 수 있다. 기준 시료의 위치를 달리하여 시료 제시 순서를 변형시킨 다양한 일·이점 차이검사의 평가 예민도를 비교 평가한 이전 문헌들에서, 기준 시료가 제일 먼저 제공되는 전통적인 일·이점 차이검사가 기준 시료가 중간에 제공됨으로써 기준 시료와 두 평가 시료가 직접 비교 가능하도록 고안된 기준시료 중간 일·이점 차이검사보다 높은 평가 예민도를 보였다(Lee HS & Kim KO 2008; Kim MA 등 2010, Kim MA & Lee HS 2012). 이는 일·이점 검사에서 기준 시료를 먼저 제공하는 것이 비교 평가를 위한 기준을 제시할 수 있어 평가의 안정성을 높일 수 있음을 나타낸다. 따라서 본 연구에서는 기준 시료가 가장 먼저 제시된 일·이점 차이-선호 검사가 기준 시료가 가장 나중에 제시된 ABX 선호-차이 검사보다 높은 평가 예민도를 보였을 수 있다.

## 요 약

본 연구에서는 가장 일반적인 선호도 검사 유형인 이점 선호도 검사방법과의 비교 평가를 통해 차이식별 검사에 선호도 검사를 융합한 감각평가 방법의 효율성을 입증하고자 하였으며, 차이식별 검사와 선호 검사의 평가 순서를 달리한

두 종류의 평가 방법(일·이점 차이-선호 검사, ABX 선호-차이 검사)을 비교 평가하였다. 일·이점 차이-선호 검사와 ABX 선호-차이 검사 모두 선호도 응답에 있어서 이점 선호도 검사와 유의적인 차이를 보이지 않았으며, 이는 선호도 검사에 차이식별 검사를 융합시킨 검사 방법의 활용이 가능할 수 있음을 보여주는 결과이다. 두 종류의 차이-선호 검사 방법의 차이식별력을 비교하였을 때, 일·이점 차이-선호 검사가 보다 예민한 결과를 나타냈고, 이는 기준 시료가 먼저 제공됨으로써 평가 시료에 대한 비교 기준이 마련될 수 있었기 때문인 것으로 사료된다. 결론적으로, 차이식별 검사와 선호도 검사를 융합한 평가방법은 두 가지 정보를 동시에 제공할 수 있는 효율적인 검사방법으로 활용될 수 있으며, 이때 차이식별 검사를 먼저 수행하는 것이 일·이점 검사의 평가수행력을 높일 수 있는 것으로 확인되었다. 본 연구에서는 감각 차이가 다소 큰 시료를 이용하여 평가가 진행되었다. 이와 같은 결과는 시료 간의 감각 차이와 같은 다양한 시료 특성에 의해서도 영향을 받을 수 있기 때문에 다양한 시료를 이용해서 재 검증될 필요가 있다.

## 감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구입니다(No. 2022R1C1C1012074).

## REFERENCES

- Braun V, Rogeaux M, Schneid N, O'Mahony M, Rousseau B (2004) Corroborating the 2-AFC and 2-AC Thurstonian models using both a model system and sparkling water. *Food Qual Prefer* 15(6): 501-507.
- Bi J (2005) Similarity testing in sensory and consumer research. *Food Qual Prefer* 16(2): 139-149.

- Bi J (2011) Similarity tests using forced choice methods in terms of Thurstonian discriminial distance,  $d'$ . *J Sens Stud* 26(2): 151-157.
- Bi J, Ennis DM, O'Mahony M (1997) How to estimate and use the variance of  $d'$  from difference tests. *J Sens Stud* 12(2): 87-104.
- Chae JE, Lee YM, Lee HS (2010) Affective same-different discrimination tests for assessing consumer discriminability between milks with subtle differences. *Food Qual Prefer* 21(4): 427-438.
- Christensen RHB, Brockhoff PB (2015) SensR -An R-package for Sensory Discrimination. R Package Version 1.5-2. <http://www.cran.r-project.org> (accessed on 2022. 10. 18).
- Christensen RHB, Lee HS, Brockhoff PB (2012) Estimation of the Thurstonian model for the 2-AC protocol. *Food Qual Prefer* 24(1): 119-128.
- Ennis DM, Ennis JM (2012) Accounting for no difference/preference responses or ties in choice experiments. *Food Qual Prefer* 23(1): 13-17.
- Frandsen LW, Dijksterhuis GB, Brockhoff PB, Nielsen JH, Martens M (2007) Feelings as a basis for discrimination: Comparison of a modified authenticity test with the same - different test for slightly different types of milk. *Food Qual Prefer* 18(1): 97-105.
- Glenn WA, David HA (1960) Ties in paired-comparison experiments using a modified Thurstone-Mosteller model. *Biometrics* 16(1): 86-109.
- Greenberg A (1958) Paired comparisons in consumer-product tests. *J Mark* 22(4): 411-414.
- Greenberg A, Collins S (1966) Paired comparison taste tests: Some food for thought. *J Mark Res* 3(1): 76-80.
- Jesionka V, Rousseau B, Ennis JM (2014) Transitioning from proportion of discriminators to a more meaningful measure of sensory difference. *Food Qual Prefer* 32(Part A): 77-82.
- Kim IA, Yoon JY, Lee HS (2015) Measurement of consumers' sensory discrimination and preference: Efficiency of preference-difference test utilizing the 3-point preference test precedes the same - different test. *Food Sci Biotechnol* 24(4): 1355-1362.
- Kim MA, Chae JE, van Hout D, Lee HS (2014) Higher performance of constant-reference duo-trio test incorporating affective reference framing in comparison with triangle test. *Food Qual Prefer* 32(Part A): 113-125.
- Kim MA, Lee HS (2012) Investigation of operationally more powerful duo-trio test protocols: Effects of different reference schemes. *Food Qual Prefer* 25(2): 183-191.
- Kim MA, Lee HS (2015) Duo - trio difference-preference test with two replications: Use of psychological biases for measuring meaningful preference. *J Sens Stud* 30(3): 211-224.
- Kim MA, Lee YJ, Kim MS, Lee HS (2022) Reminder - preference test, affective difference-preference test using reference framing with a brand: 1. Sensitivity comparisons with the same - different difference - preference test. *Food Res Int* 155: 111065.
- Kim MA, Lee YM, Lee HS (2010) Comparison of  $d'$  estimates produced by three versions of a duo-trio test for discriminating tomato juices with varying salt concentrations: The effects of the number and position of the reference stimulus. *Food Qual Prefer* 21(5): 504-511.
- Kim MA, Sim HM, Lee HS (2015) Affective discrimination methodology: Determination and use of a consumer-relevant sensory difference for food quality maintenance. *Food Res Int* 70: 47-54.
- Kim MA, Sim HM, Lee HS (2016) Idiographic duo - trio tests using a constant-reference based on preference of each consumer: Sample presentation sequence in difference test can be customized for individual consumers to reduce error. *Food Res Int* 89(Part 1): 454-462.
- Lawless HT, Heymann H (2010) *Sensory Evaluation of Food: Principles and Practices*. 2nd ed. Springer, NY, USA. p 306.
- Lee HS, Kim KO (2008) Difference test sensitivity: Comparison of three versions of the duo-trio method requiring different memory modes and taste sequences. *Food Qual Prefer* 19(1): 97-102.
- Lee YM, Chae JE, Lee HS (2009) Effects of order of tasting in sensory difference tests using apple juice stimuli: Development of a new model. *J Food Sci* 74(6): S268-S275.
- Macmillan NA, Creelman CD (2005) *Detection Theory: A User'S Guide*. 2nd ed. Lawrence Erlbaum Associates, Inc., NJ, USA. pp 380-399.
- Macmillan NA, Kaplan HL, Creelman CD (1977) The psychophysics of categorical perception. *Psychol Rev* 84: 452-471.
- Moskowitz HR, Beckley JH, Resurreccion AVA (2012) *Sensory and Consumer Research in Food Product Design*



- and Development. 2nd ed. Blackwell Publishing, Iowa, USA. p 233.
- Prescott J (1999) Flavour as a psychological construct: Implications for perceiving and measuring the sensory qualities of foods. *Food Qual Prefer* 10(4-5): 349-356.
- R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org> (accessed on 21. 12. 2022).
- Schutz HG, Bradley JE (1954) Effect of bias on preference in the difference and preference test. pp 85-92. In: Food acceptance testing methodology. Peryarn DR, Pilgrim FS, Peterson MS (eds). National Academy of Sciences, Washington, DC, USA.
- Stone H, Bleibaum RN, Thomas HA (2020) Sensory Evaluation Practices. 5th ed. Academic Press, CA, USA. p 331.
- van Hout D (2014) Measuring meaningful differences: Sensory testing based decision making in an industrial context; applications of signal detection theory and Thurstonian modelling. Ph D Dissertation Erasmus Research Institute of Management (ERIM), Rotterdam, The Netherlands. pp 31-32.

---

Date Received	Nov. 7, 2022
Date Revised	Dec. 12, 2022
Date Accepted	Dec. 13, 2022