

ICS 67.240

XX XX



# 中华人民共和国国家标准

GB/T XXXX—202X/ISO 13299 :2016

---

## 感官分析 方法学 建立感官剖面的导则

Sensory analysis -- Methodology -- General guidance for establishing a sensory  
profile

(ISO 13299:2016, IDT)

(征求意见稿)

202X- - 发布

202X- 实施

---

中华人民共和国国家市场监督管理总局

中国国家标准化管理委员会

发布

## 目 次

前言.....	I
1 范围 .....	1
2 规范性引用文件.....	1
3 术语和定义.....	2
4 一般测试条件.....	4
4.1 设备和评价室.....	4
4.2 评价员.....	4
4.3 产品.....	5
4.4 样品.....	5
4.5 初步讨论.....	5
5 描述性方法：原则和主要特征.....	6
5.1 一致性剖面.....	6
5.2 参比剖面偏差（相对参比标度） .....	6
5.3 自选剖面.....	6
5.4 闪现剖面.....	6
5.5 定量的描述剖面.....	6
5.6 定性感官剖面.....	7
5.7 动态主导型感官属性测试（TDS） .....	7
6 感官剖面建立步骤.....	7
6.1 总则.....	7
6.2 测试准备.....	7
6.3 测试开展.....	9
6.4 统计解释.....	10
6.5 研究报告.....	10
附件 A（资料性附录）一致性剖面.....	12
附录 B（资料性附录）偏离参比法（或相对参比评价） .....	14
附件 C（资料性附录）自选剖面.....	16
附件 D（资料性附录）闪现剖面.....	18
附录 E（资料性附录）定性的感官剖面 .....	20
附件 F（资料性附录）定量的描述剖面.....	22
附件 G（资料性附录）动态主导型感官属性测试（TDS） .....	29
附件 H（资料性附录）评价小组量化一个属性时的单变量分析 .....	33
参考文献.....	43

## 前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准使用翻译法等同采用 ISO 13299:2016《感官分析—方法学—建立感官剖面一般导则》。

与本标准中规范性引用的国际文件有一致性对应关系的我国文件如下：

——GB/T 10221—2012 感官分析 术语（ISO 5492:2008, MOD）

——GB/T 15549—1995 感官分析 方法学 检测和识别气味方面评价员的入门和培训（ISO 5496:1992, IDT）

——GB/T 10220—2012 感官分析 方法学 总论（ISO 6658:2005, IDT）

——GB/T 16291.1—2012 感官分析 选拔、培训与管理评价员的一般导则 第 1 部分：优选评价员（ISO 8586-1:1993, MOD）

——GB/T 16291.2—2010 感官分析 选拔、培训和管理评价员的一般导则 第 2 部分：专家评价员（ISO 8586-2:2008, IDT）

——GB/T 12315—2008 感官分析 方法学 排序法（ISO 8587:2006, IDT）

——GB/T 13868—2009 感官分析 建立感官分析实验室的一般导则（ISO 8589:2007, IDT）

——GB/T 16861—1997 感官分析 通过多元分析方法鉴定和选择用于建立感官剖面的描述词（ISO 11035:1994 IDT）

本标准做了下列编辑性修改：

术语和定义 3.9 中，将“ISO 5492:2008/修改单 1（未出版）”改为于“ISO 5492:2008/修改单 1：2016”；

本标准由全国感官分析标准化技术委员会（SAC/TC566）提出归口。

本标准起草单位：。

本标准主要起草人：。

# 感官分析 方法学 建立感官剖面的导则

## 1 范围

本标准给出了建立感官剖面的导则。感官剖面适用于通过视觉、嗅觉、味觉、触觉或听觉来评价的所有产品或样品（例如，食品、饮料、烟草制品、化妆品、纺织品、纸张、包装、空气或水样品等）的感官分析。本标准同样适用于人类认知和行为的研究。

感官剖面的一些应用如下：

- 开发或改进产品；
- 根据感官特性来定义产品、生产标准或评分标准；
- 定义用于保质期测试的“新鲜度”参比样；
- 研究和改进产品的保质期；
- 将产品与参比样、市面上类似产品或研发中产品进行对比；
- 描绘产品的感知特性，并将它们与其仪器、理化性质或消费者接受度等因素相关联；
- 采用类型和强度来描述样品中臭味或异味（例如，污染研究）。

## 2 规范性引用文件

下列文件对于本文件的应用必不可少，其全部或部分在本文件中被正式引用。凡是注明日期的引用文件，仅注明日期的版本适用于本文件。凡是未注明日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

ISO 5492 感官分析 术语（Sensory analysis—Vocabulary）

ISO 5496 感官分析 方法学 检测和识别气味方面评价员的入门和培训（Sensory analysis — Methodology — Initiation and training of assessors in the detection and recognition of odours）

ISO 6658 感官分析 方法学 总论（Sensory analysis — Methodology — General guidance）

ISO 8586 感官分析 选拔、培训与管理评价员的一般导则（Sensory analysis — General guidance for the selection, training and monitoring of assessors）

ISO 8587 感官分析 方法学 排序法（Sensory analysis — Methodology — Ranking）

ISO 8589 感官分析 建立感官分析实验室的一般导则（Sensory analysis — General guidance for the design of test rooms）

ISO 11035 感官分析 通过多元分析方法鉴定和选择用于建立感官剖面的描述 (Sensory analysis — Identification and selection of descriptors for establishing a sensory profile by a multidimensional approach)

ISO 11136 感官分析方法—对受控区域的消费者开展喜好测试的一般导则 (Sensory analysis — Methodology — General guidance for conducting hedonic tests with consumers in a controlled area)

### 3 术语和定义

ISO 5492 中界定的术语和定义适用于本文件。

#### 3.1

##### **特性 attribute**

可感知的特征。

[来源：5492:2008，有修改]

#### 3.2

##### **感官剖面 sensory profile**

对样品感官特性 (3.1) 的描述，通常采用特性强度值表示。

#### 3.3

##### **部分感官剖面 partial sensory profile**

某些选定感官特性 (3.1) 的剖面，通常采用特性强度值表示。

实例：气味剖面，风味剖面，质地剖面。

#### 3.4

##### **定量的描述剖面 quantitative descriptive profile**

对样品感官特性 (3.1) 及其强度值的描述。

[来源：5492:2008，有修改]

#### 3.5

##### **定性的感官剖面 qualitative sensory profile**

对样品感官特性的描述，不包含强度值。

[来源：5492:2008，有修改]

#### 3.6

**一致性感官剖面 consensus sensory profile**

感官评价小组对产品各种感官特性（3.1）的描述，经集体讨论后形成的具有一致性的结果。

**3.7****偏离参比法（相对参比评价） deviation from reference method (relative-to-reference rating)**

将所有样品与参比样进行对比评价的定量描述感官剖面程序（3.2）。

**3.8****自选感官剖面 free-choice sensory profile**

每一评价员独立为一组产品选择的特性组成的感官剖面。

**3.9****闪现剖面 flash profile**

评价员选择自己对产品特性的描述词，并按照强度高低对产品排序的一种感官剖面法。

注 1：是自选感官剖面法的一种变体，区别在于使用了排序而不是评价。

[来源：ISO 5492:2008/修改单 1：2016]

**3.10****动态主导型感官属性测试 temporal dominance of sensations TDS**

要求每个评价员对产品评价过程中占主导的感觉进行持续测量并记录的过程。

**3.11****评价小组 sensory panel**

参加感官分析的评价员组成的小组。

[来源：ISO 5492:2008, 有修改]

**3.12****小组组长 panel leader**

主要负责管理小组活动并招募、培训和管理评价员的人员。

注 1：小组组长也需负责设计和开展感官测试、分析和解释数据。

[来源：ISO 13300（所有部分），有修改]

**3.13****优选评价员 selected assessor**

挑选出的具有较强感官分析能力的评价员。

[来源：ISO 5492:2008, 1.6]

### 3.14

#### **专家评价员 expert sensory assessor**

具有高度的感官敏感性、经过广泛的训练并具有丰富的感官分析方法经验，能够对所涉及领域内的各种产品做出一致的、可重复的感官评价的优选评价员（3.13）。

[来源：ISO 5492:2008, 1.8]

## 4 检验的一般条件

### 4.1 设备和评价室

实验室应配备 ISO 6658 中规定的用于样品制备的所需设备。

开展感官剖面时需要满足 ISO 8589 规定的要求。需要展开讨论（例如，关于结果、产品、参比样等）时，房间的布置需保证评价员间交流，小组组长需确保产品评价是在适当的条件（例如，合适的灯光）下开展。

应指定小组组长负责感官剖面工作开展。小组组长应：

- 培训评价员，
- 维护评价小组，
- 开展测试。

小组组长需达到 ISO 13300-1 和 ISO 13300-2 所规定的任职资格（例如，熟悉招募和培训步骤）。

### 4.2 评价员

本标准适用于采用优选评价员或专家评价员进行的剖面方法。评价员的选拔、培训和管理必须符合 ISO 8586 中的规定。

评价员的数量及培训内容根据采用的剖面方法而定。评价员的重复性和再现性随选择水平和培训时间而提升。结果的解释和产品间突出性的差异也取决于评价员的数量和培训。

可通过会谈、通告或个人联系等方式开展候选人招募。按照所需评价人数的 2~3 倍进行面试和筛选。须重视以下特征：

- 产品评价所需的健康状况；
- 兴趣和动机；
- 约定时间内时间安排和小组讨论时适用性；
- 及时性；
- 集中力；
- 记忆的能力；
- 如实沟通和感知汇报的能力；
- 所研究特征差别的辨别能力；
- 团队合作的能力。

感官敏感度可以通过建立 10 人或以上评价员的评价小组来平衡。

### 4.3 产品

明确研究的产品及其制备的条件。

例如：用水或牛奶制备的有糖或无糖的速溶咖啡。

### 4.4 样品

样品的准备和提供需符合 ISO 6658 中的规定。确保评价员不能通过样品提供方式得出样品性质的相关结论。例如，必要时使用有色测试眼镜或彩色光源掩饰样品的外观差异。

对同一温度下的样品制备和分装实行标准化操作。样品采用三位随机编码，提供顺序可采用适当的实验设计。

为了提高结果的可靠性和有效性，如果可能，任何样品或者样品组都应安排不同的测试日，且提供 2~3 次或者更多。重复的次数应根据要求的精确度，所观测结果的分散度以及评价员逐步熟悉样品的过程中分辨力提升的具体趋势为指导。重复评价可以实现对实验误差的估算。对同一批次产品的重复评价，也可以看出同一评价员评分值的分散情况，而不同批次产品的重复评价也可反映产品间的差异。评价方案中应给出重复哪些样品，以及在什么条件下准备和评价样品的要求。

评价员完成所有评价之前，不得公开样品的相关信息。

### 4.5 初步讨论

确保评价员完全熟悉所有的所需研究特性和 ISO 6658 中规定的评价机制。如有必要，



应安排关于测试问题和样品性质的初步讨论。一些典型性的产品需要被提供和讨论。小组组长应确保讨论结果不影响后续评价。

## 5 描述性方法：原则和主要特征

### 5.1 一致性剖面

一致性剖面分析中，评价员通过分享个人观点，就不同属性、属性出现顺序和属性强度达成一致结论。

通常，该标度仅限制几个标值，1个属性仅有1个分值（达成一致的得分）。如果某个评价员不同意小组分数时，应该如实记录在报告中。

### 5.2 偏离参比剖面（相对参比标度）

样品通常以成对形式提供。评价员依据通用术语表中的每个属性，直接对2个产品进行比较。或者由果及因，根据2个产品的分值进行比较。如需要比较2种以上产品时，则应在相同条件下将每种产品与参比样品进行比较。

针对样品和参比样间的差异进行数据分析。

### 5.3 自选剖面

自选剖面中，每位评价员使用他/她自己的术语表而不是通用术语表。

采用适当的多元分析方法对结果进行解释说明，如广义普鲁克分析。结果应以绘图形式呈现。

### 5.4 闪现剖面

闪现剖面是自由选择剖面法的一个变体，是同时提供所有样品，并通过排序对样品进行比较的方法。

采用适当的多元分析方法对结果进行解释说明，如广义普鲁克分析。结果常以绘图形式呈现。

### 5.5 定量描述性剖面

定量描述性剖面中，评价员采用通用属性表评价样品并对强度进行评分。

有几种方法可以建立定量描述性剖面，其中一些技术已经注册商标<sup>1</sup>。分析结果应包括各属性的强度得分，便于进行单变量或多变量分析。

## 5.6 定性感官剖面

定性感官剖面中，评价员只需采用通用术语表评价属性是否存在，不用给出感知的属性强度。

与定量描述性感官剖面相比，定性感官剖面的术语表范围更大，产品依赖性更小。小组培训应集中在对各类参比样的识别和记忆上。参比样应稳定且不随时间变化，以便于记忆。

评价员数量和/或重复次数应当高于定量描述性感官剖面。

结果通常以每个属性的出现频次表示。

## 5.7 动态主导型感官属性测试

TDS 是一种时间剖面法，要求每个评价员在评价产品过程中连续记录下占主导的感觉。

使用通用术语表选择占主导的感觉。此外，可以对所选主导属性的强度进行评分。

数据应以任一时刻占主导属性的比例呈现。数据通常转换成曲线，其中 x 轴为时间轴。测试产品的不同属性曲线应合并到一张图中。

# 6 建立感官剖面的步骤

## 6.1 总则

本条款给出了建立感官剖面的通用步骤。每种方法的详细说明，请参阅相应的附录。

## 6.2 测试准备

### 6.2.1 选择培训产品

培训产品的选择应符合 ISO 8586 中的规定。

### 6.2.2 筛选评价员

评价员的筛选应符合 ISO 8586 和 ISO 5496 中的规定。

---

<sup>1</sup> QDA®和 Spectrum™ 是目前适合定量描述剖面分析的商业程序实例。这些信息是为了方便本文件的使用者而提供的，并不表示 ISO 对这些程序的认可。

### 6.2.3 选择最佳属性

目的是确定一组不重叠、单一、客观、明确和可参照的属性，可能的话，尽量对研究样品开展完整的描述性分析。该重要步骤可单独或共同完成，通过感官剖面的方法。如果需要  
一个通用术语表，小组组长可采用表 1 中所列的任意一种或多种方法获得。

表 1 最佳属性的选择程序

编号	原则	方法	优点	缺点
1	使用现有术语和参照标准。	通过参考文献和咨询专家进行适当选择。采用规定的标准，教会评价员每个描述词的特性。需要时为描述词提供一个强度标度。	吸取专家多年累积的经验。其他评价组也可解释该剖面，便于与其他研究比较。	对于一组特定样品而言，使用现有术语或参照标准可能做出不准确或不适当的选择。建立新术语时可能遗漏本该发现的属性。
2	特殊评价环节中，使用评价小组建立后期将使用到的术语。	采用优选评价员小组；在有经验的小组组长指导下，通过圆桌讨论中获得术语。 参照标准由小组组长或测试要求者提供，也可由评价员在评价环节中提供。可与方法 1 相结合。	建立术语的过程比方法 3 更省时。	所得剖面仅适用于给定的评价小组和样品。 如果没有参照标准，其他评价小组不能对其进行解释。
3	特殊评价环节中，使用评价小组来形成后期将使用的术语。	参照 ISO 11035，采用一组预先准备好的训练样品来识别和选择判别术语；再通过统计分析逐步减少术语的数量。	使用一个完全客观的选择和删除过程，最大限度地减少由于传统误解或先入为主概念的术语，所选术语能最优表达评价员感知的样品特性。	所得剖面仅适用于给定的评价小组和样品。 如果没有参照标准，其他评价小组不能对其进行解释。该过程相对耗时，且需要一定经验，特别是数据分析方面。

作为剖面评价的一部分，可以要求评价员提供一个或多个的整体评价。例如：

——整体果味或辣味，和

——整体风味强度。

**注意：**受训的感官评价小组的全面评价不属于喜好，因为可能会有偏差：受训过的评价员对产品的描述是客观的，且会有意或无意识采用与代表性目标消费者不同的评价策略。如果是用于指导新产品开发的喜好测试，需按照 ISO 11136 的规定进行。

## 6.2.4 确定评价顺序

对于有些产品（例如饮料），某些属性的感知顺序是产品剖面的重要特征。对于其他类产品，属性的评价顺序可以改变，例如与质地或物理性质相关的属性。

例子：一块融化的巧克力；一张湿润的面巾纸。

通常，评价顺序应取决于评价属性的顺序。后味或后感觉通常最后评价；这些属性在样品被食用或使用后仍然可以被感知。与肤感和织物感评价类似，口腔质感的评价顺序可以通过产品预处理方式提前确定。小组组长通过设定操作（一个咀嚼，一个手动挤压）来确定属性的评价顺序。

## 6.2.5 选择适当的响应标度

当剖面是由样品每个属性的强度组成时，应仔细选择响应标度。感官剖面采用的响应标度可以是数字的或语义的，连续的或不连续的，单极的或双极的，如 ISO 4121 所述。

注：响应标度获取的数据可能会受到所谓的“末端效应”的影响：评价员不愿意使用接近任意末端的数字，以免后来的样品更加极端。通常认为连续响应标度，例如线性标度，不太可能产生末端效应。

## 6.2.6 培训评价员

应按照 ISO 8586 的规定开展评价员培训。建议提供给评价员反馈意见。

## 6.3 测试开展

### 6.3.1 记录表

使用预打印好的记录表，表中应包含所用标度说明；或使用计算机或平板电脑进行数字化记录。记录表上应留下空白，并要求评价员对附加属性提供意见或建议。

注：单张纸或同一屏幕上出现多个评价时可能会相互影响或曲解（例如光环效应：对一个特性的正面或负面评价可能影响到下一个特性的评价）。

### 6.3.2 样品评价

安排评价员在独立的评价间开展评价。

根据采用方法的不同，样品应单个（连续或一个接一个）或同时提供。同时供样有利于样品间的比较，但会使剖面分析不独立。对于这 2 种情况，评价员间的样品递送顺序应平衡或随机分配，以减少顺序效应影响。

每个评价员和每轮次所评价的样品数量应根据每轮次所需时长、产品性质、属性数量和预期差异进行调整：如果样品差异很小，且样品具有强烈或持久的味道时，可以限制样品的提供数量。

## 6.4 统计解释

进一步分析前，始终建议对原始数据进行检查和图形化处理，旨在：

- 检查并纠正明显的错误，
- 如果可以，删除异常值，
- 获得对整个数据结构的一个初步印象，
- 发现后续有意思的可能趋势。

多种分析手段均适用：全部数据的平均值及标准偏差计算，或关注单个评价员或一些特殊属性，如箱形图、直方图和线形图。

统计处理手段取决于采用的评价方法。通常，数据解释包括 3 个步骤。第一步是重点关注评价员的表现，检查数据是否存在实验性错误。第二步是确保评价使用的描述词是最能区分所研究产品的描述词。这一步通常也被称为单变量分析（见附件 H）。最后一步是评估第一阶段所有描述词的有效性。通常也被称为多变量分析。通常在描述词分类后进行，如视觉描述词、风味描述词、味道描述词和质地描述词。

注：本标准无法对所有多变量分析技术一一赘述，但附录讲述不同剖面方法时有所提及。

## 6.5 研究报告

研究报告应包含以下信息：

- a) 开展研究的公司或实验室名称，小组组长和研究主管姓名；
- b) 研究目的；
- c) 样品的完整信息；
- d) 日期、测试时间和评价所需时长；
- e) 研究的操作条件（包括任何与本标准中方法不一致的条件）：
  - 1) 注明采用本标准进行评价，如 ISO 13299 和质量控制步骤；
  - 2) 详细的操作步骤：样品提供顺序和方式，重复性的定义（相同/不同的日期，相同/不同的批次）；
  - 3) 评价员的数量和类型（优选评价员或专家评价员）；

- 4) 评价属性的定义列表、评价方案和参比样品;
- 5) 使用的响应标度;
- 6) 研究期间提供给评价员的任何其他信息 (例如, 产品的类型或品牌信息);
- f) 研究结果, 统计分析后采用适当形式 (图形、图表和绘图等) 呈现。如果与合作伙伴或赞助者有相关协议, 报告内容无需特别详细。

附件 A  
(资料性附录)  
一致性剖面

### A.1 原则

一致性剖面分析中，评价员对产品展开评价，讨论属性和/或强度，最终得出一个达成一致的描述结果。该方法要求结果要达成一致，而不是将数据平均化。小组必须接受大量训练去确定属性特征，并基于所尝味道判断属性和强度。

Arthur D Little有限公司于40年代末开发的风味剖面法（FPM），通用食品公司产品评价和质地技术小组于60年代开发了质地剖面法（TPM）。结果达成一致均是2种方法的关键重要因素。这2种方法后发展为90年代的Spectrum<sup>TM</sup>法<sup>2</sup>。但是，这2种技术间存在一个根本区别：FPM需要达成一致，而TPM并不强制最终达成一致（见第6条）。

### A.2 感官属性

通过评价小组组长帮助提供和维护参比样，评价员开发和定义整个小组使用的通用术语。每个描述词的参考标准和定义在训练阶段创建。在FPM中，评价员描述风味并估计其强度，同时评价由产品带来的“大致印象”（称为振幅）。TPM使用标准化术语来描述质地特征；这些都是从物理和感官2个方面来描述。

### A.3 标度

起初，FPM为5点标度。

0：无特征体现；

)（：特征处于阈值级别；

1：轻微特征；

2：中等特征；

3：强烈特征。

但这个标度随后被扩展了，包括使用箭头，1/2s，或者加减号。最初的TPM使用了扩展的13点版本的味道剖面方法。如今，TPM最常见的可能是15点标度；也可能是一个10或

<sup>2</sup> QDA<sup>®</sup>和 Spectrum<sup>TM</sup> 是目前适合定量描述剖面分析的商业程序实例。这些信息是为了方便本文件的使用者而提供的，并不表示 ISO 对这些程序的认可。

15厘米的标尺。

#### A.4 评价员

参与建立一致剖面的评价员具有以下2个典型特征：

——他们只占少数（FPM中的4-6）。

——他们训练有素。在FPM中，会提供给评价员一套参比样，它们代表该类产品的范围、成分和处理变量。在TPM中，评价员需要学习在标度上使用不同的分数，并能再现其评分。

参见参考文献[16]。

#### A.5 程序

——评价员对样品进行逐一单独评价，并记录属性的强度和出现顺序。

——产品评价结束后，小组z负责人收集结果，并领导评价小组进行总体讨论，就属性、强度等方面（如属性的顺序）形成一致剖面。讨论可能需要几分钟或几十分钟。可提供样品参比样以丰富讨论。一旦小组达成一致，最终的感官描述性“剖面”就创建了。

——若未达成一致，小组组长可能会在新评价过程中提供产品或记录评价员间的差异。

——重复该过程直到所有产品评价完毕。

——小组组长必须小心避免高级成员或主导性人物强加他们的观点。

以上是基本步骤，实测时可以有多种变化。

#### A.6 统计分析和结果解释

FPM中不包含统计分析，数据主要以表格形式呈现，也可采用图形表示。TPM中，小组最终决议是通过小组协商一致或数据统计分析得出。最终报告中，数据可采用表格或图形形式呈现。

可以使用多元统计分析方法（例如“树”图或PCA）分析产品之间的相似性和差异性。



## 附录 B

## (资料性附录)

## 偏离参比法（或相对参比评价）

**B.1 原则**

偏离参比法，也称为相对参比评价，由Larson和Pangborn（1978）首次提出。该方法是通用描述性分析的变体，采用参比样来评价所有样品（样品成对呈现）。该方法适用于有明确参比样或区分困难的情况。通常有2种方式：评价员将直接样品与参比样比较并评分（参比样设在标度中心点）；或者评价员通过计算样品和参比样之间分值差进行样品比较。通常，评价员并不知道哪个样品是参比样。

**B.2 感官属性**

评价员采用的通用感官属性表的创建类似于定量描述分析。感官属性表必须在培训期间制定完成，可使用已知术语组成的“封闭式列表”，或者由评价小组完全自主创建。

**B.3 标度**

响应标度可以采用结构式或非结构式。

当样品直接与参比样对比评价时，通常采用双极标度，参比样作为中心点。这种情况下，对于给定的属性，评价员对样品的评价是比参比样更多或更少（图B.1）。



图B.1 参比样作为中点时的非结构式标度示例

当样品不直接与参比样对比评价时（评价完成后再计算差异），采用的标度与定量描述分析标度相同，对于给定的属性，2种标度非常相似（图B.2）。



图B.2 参比样和样品间差异计算时的非结构式标度示例

## B.4 评价员

通常情况下，根据待测产品，参照定量描述剖面法进行评价员的筛选和培训。每次评价时采用同一个参比样，能有效提高评价员的一致性以及多次评价时评分稳定性。

## B.5 步骤

训练过程与定量描述剖面类似。包括多个主题的训练环节，如使评价员熟悉产品类别的环节，开发和掌握感官属性表的环节，以及使用标度对每个属性强度评分的环节。

评价时，评价员会收到一组2个样品。建议评价员和样品轮次间采用平衡呈现设计。

——直接比较样品，1个样品定为参比样，另1个样品与其对比进行评价（图B.1）。参比样评价通常需要在待测样品前进行。待测样评价时，评价员可根据需要反复比较参比样。

——评价员对2个样品(图 B.2)进行独立评分。通常允许评价员多次品尝样品，评价员也可同时评价2个样品(例如，同时在面颊两侧涂抹化妆产品)。评价完成后再进行2个样品的差异计算。

以上2种方式中，参比样通常被当做样品（“盲参比样”，未明确指明哪个样品是参比样），并作为评价员可靠性的内部测定。

某些条件下，2种以上样品的比较在技术上是可行的。

——所有样品与参比样的比较必需在相同条件下。

——应仔细筛选参比样，保证其与评价样品是同类产品。

正如在定量描述剖面中一样，评价通常需要重复。

## B.6 统计分析和结果解释

以上2种方式均是针对样品和参比样间的差异进行数据分析

与中心点相比，得分低于参比样的样品通常为负值，高于参比样的样品通常为正值，采用经典的单变量统计分析。当仅2个样品时（其中1个是参比样），采用双尾配对t检验。所有样品需要相互比较以及与参比样比较，只要满足了前面所提条件，均可采用单变量和多变量分析。这种情况下，为了参照，可以使用盲参比样的分值替代每对样品中参比样的分值。

附件 C  
(资料性附录)  
自选剖面

### C.1 原则

自选剖面 (FCP) 是每个评价员创建其自己的描述性术语表并通过事先预备好或自选的标度对感觉强度量化的过程。

FCP 可以进行 3 个或以上的多属性产品的比较及差异分析。通常用于产品空间的感知描绘的创建 (例如, 用于市场研究), 也可用于属性创建。

### C.2 感官属性

评价员创建自己的感官属性描述词, 且不需要对术语含义进行解释。评价员相互讨论属性可能有助于产品的描述, 但讨论结果并不需要达成一致, 因为每位评价员都有自己的术语表。这些术语可以不同, 只要评价员自己能够理解即可。每位评价员采用自己的术语表进行所有产品的评价。可以选择所有或者部分属性, 如形状、颜色、味道和气味等。部分评价员仅使用很少几个属性。特殊属性使用没有规定。评价员可以使用其希望的任何词汇和/或词语组合进行描述 (例如“闻起来像露营一样”)。

### C.3 标度

评价员通过事先预备好的或自选标度对感觉强度进行量化。允许评价员采用不同标度对产品进行不同评价。但是, 每个评价员需始终使用自己的标度评价所有产品。无需对强度或标度进行培训。

### C.4 评价员

评价员需具有良好的视觉、嗅觉、味觉、触觉和听觉, 事先没有进行过某一产品的特定培训。要求筛选具有较好描述能力的评价员, 能使用一些全面综合的词汇对具有非常多属性的产品展开描述。无需开展进一步培训。

建议至少需要 10 名评价员。

### C.5 步骤

第一步, 评价员各自收集属性, 选择他们想要或事先预备的标度方法。

第二步，评价员在标准条件下对需比较的产品属性及强度展开评价，注意确保整个评价过程的一致性。样品提供顺序随机。评价员评价产品需重复2次或以上。

方格法可以代替FCP法，他比FCP法生成的属性更加结构化，是先选择简单属性再通过一系列比较不断改进来完成。第一轮，将2到4个产品（通常3个）进行比较，生成属性。再将上一轮的1个产品与另外2个新产品进行比较。评价员需回答2个新产品与第1个产品有何相似或不同。依此类推，3个一组的测试一直持续进行。

### C.6 统计分析和结果解释

产品的完整描述不能通过ANOVA、PCA或不适当的线性回归方法进行统计评价。FCP数据通常采用广义普鲁克分析法(GPA),通过对每位评价员的单次评价和重复性展开统计，通过数据转换（旋转、平移和缩放）获取具有一致性的产品轮廓。

附件 D  
(资料性附录)  
闪现剖面

### D.1 原则

闪现剖面法由Sieffermann在2000年首次提出。它是FCP和排序法的结合，评价员通过描述一组产品间的差异单独创建属性。它通过感官描绘，在感官空间中显示类似的产品分组。直接比较属性强度排序，强调相对差异。

### D.2 感官属性

每位评价员独立开展组内产品间差异描述，并独立获得属性。这是成功开展闪现剖面的重要要求，不允许使用喜好类术语。

评价员间可以对属性展开讨论，有利于产品的描述。但，每位评价员可以使用自己的术语表，讨论结果无需达成一致。

### D.3 标度

由于要求评价员对每个属性的不同样品进行排序，所以没有具体的标度可用。

### D.4 评价员

评价员应具有良好的视觉、嗅觉、味觉、触觉和听觉，之前没接受过针对某一产品的专门培训。要求筛选具有较好描述能力的评价员。能使用综合全面词汇来描述具有大量属性的产品，后期无需进一步培训。

建议至少需要10名评价员。

### D.5 步骤

首先通过1到2次培训，指导评价员方法的使用以及如何开展产品描述和差别分析。在此过程中，每位评价员通常会获得一组产品并独自进行描述。可以通过样品分类找到产品间的差异。然后，评价员在小组组长协助下删除无关的或喜好属性，并添加相关的属性。

接下来进行样品测定。依据属性强度对产品排序，对属性依次进行测定。建议同一轮次进行部分样品的重复评价或多个轮次中进行所有样品的重复评价。

与FCP不同，所有产品均采用平衡区组或BIB（平衡不完整区组）设计同时提供给评价

员。样品呈送设计取决于评价产品的数量,建议至少5到6个样品。评价样品数量超过12个时,应该采用BIB法。

#### **D.6 统计分析和结果解释**

类似于FCP法,闪现剖面的结果也是采用多元统计分析(例如,广义普鲁克分析(GPA))进行处理。计算重复次数相关性考察评价员的可靠性(斯皮尔曼秩相关性)。

## 附录 E

### (资料性附录)

#### 定性感官剖面

##### E.1 原则

定性感官剖面是基于术语被选用频次的一种方法，主要用于表征复杂的芳香类产品，如葡萄酒或奶酪。通常给评价员提供一个巨大的描述词汇表，评价员必须选择最相关的描述词进行待测产品描述。然后根据评价员选择次数（选用次数）来计算评分。足够数量的可用术语确保评价员能够选择最合适的词汇来准确描述他们的感知。通常由专家评价员（受过训练的评价员）进行测试。

##### E.2 感官属性

属性列表是一个由明确定义的术语或专门针对产品开发和调整的词汇组成的“封闭列表”。列表的建立是在培训期间进行。表中属性数量需大于定量剖面的数量：不同术语数量约为100个。属性主要集中在香气或气味上，能精确与芳香参比样进行对应。每个术语都必须对应1个标准的芳香参比样（商业用参比样或天然产品）。

##### E.3 标度

定性剖面不使用标度。评价员只需根据感知的产品属性，在术语列表中选择：测到或未测到。

##### E.4 评价员

###### E.4.1 评价员数量

评价员的数量需高于定量剖面所需数量：通常是定量剖面评价小组的2倍（16名到30名合格评价员）。

###### E.4.2 资格和培训

培训分为2个阶段：一般培训和特定产品培训。一般培训期间（根据需要记忆的属性数量，进行5~15次或更多），评价员熟悉术语表，能从多个标准芳香参比样中选出最适合每个术语定义的参比样。

针对产品进行术语表的建立并调整时，一般培训期间，评价员需对术语表进行讨论并修订，删除认为不相关、不明确或冗余的词汇，并增加认为相关的词汇。

特定培训期间，评价员还需嗅闻标准芳香参比样，同时熟悉研究产品。

### E.5 步骤

受训的评价员接收到一整组产品。平衡设计每个产品被品尝的顺序。每种产品都要求每位评价员在术语表中选择他们感受的香味或气味，但他们能选择术语的数量有限，最大限定数目取决于以下两点：（1）允许评价员以相对较多的术语来表达他们感知的可能性；（2）人们对混合物中气味特性的辨别能力有限。

评价通常需要重复。

### E.6 统计分析和结果解释

收集完所有数据后，根据术语表的选用频率对其进行排序，确定每个产品最相关的描述词。最少评价员（通常为小组的10%到15%）选用的描述词可考虑用于后续的统计分析。建立包含选用次数最多的术语表（产品为横列和描述词为纵列）。

结果分析可选用多种统计分析方法，最常见的有：

——对每个术语的引用频率（2次重复）进行卡方（ $\chi^2$ ）分析，找出产品间属性“强度”的显著差异；

——多元分析：对术语表进行相应统计分析，该表中至少含有一个产品/重复中被15%以上评价员选择术语的选用频率。



## 附件 F

## (资料性附录)

## 定量描述剖面

**F.1 原则**

定量描述剖面 (QDP) 法是一种使用受训的评价小组的方法, 描述产品所有感官特性并对强度进行量化。

感官性质即属性, 由评价小组专门针对待测产品的类别建立培训小组, 。

所得剖面是专属评价小组和专属产品类别。如果没有参比样, 其他组不能对其进行解释。

**F.2 感官属性**

定量描述剖面使用受训过的评价员选择和定义产品区分所必需的属性。

因此, 第一次培训通常侧重生成感官特性 (例如食品的外观、香味、味道、质地和回味等属性), 并确保评价员能正确使用相应的属性标度。

通常, 属性的建立首先是通过讨论达成一致或通过多元分析确定大量属性, 这些属性被删减为一个有数量限制的术语表 (参见ISO 11035)。

**F.3 标度**

确定相关属性后, 下一步选择一个适当的响应标度测定样品每个属性的强度。

用于感官分析的响应标度可以是连续 (10~15厘米) 或不连续的。通常, 极限是无 (或非常弱) 和密集 (或非常强)。

**F.4 评价员**

该方法采用受训过的评价员, 通过选拔测试后, 针对待测产品的类别对培训员进行深入培训 (参见ISO 8586)。

**F.4.1 评价员资格****F.4.1.1 招募和选拔**

合格的招募者首先必须是待测产品的使用者/接受者。选拔测试中, 他们需证明具有以下能力:

——能够区分待测产品间的差异，

——能够记住感官特性的本质。

评价员需承诺参加完成项目所需的所有环节。

#### **F.4.1.2 训练**

针对产品开展专门性培训。

小组组长带领优选评价员建立新感官用语，或修订现有要用语，对产品所有感官特性展开描述。

评价的产品必须包含有全部需要评价的属性范围。此环节需重点关注产品的类别及偏差的代表性范围（所有类型产品，包括非典型产品及各类产品...）。

小组制定每个属性的定义及标准化的评价步骤，并开展产品评价分练习。

组长或测试要求者提供标准参比样，或评价员在训练期间提供。通常，标准参比样与测试样品不能相同。

对于1个产品类别的培训时间需要10~20个小时或更长。培训时间长短取决于产品具体情况以及分配给评价小组的任务。

小组组长需组织培训、提供产品、需要时介绍参比样和协助活动开展，而不仅是作为回答问题者。

#### **F.4.1.3 表现评价**

培训结束时，监测表现数据有助于小组组长提高小组和评价员的表现，考察培训效果，筛除表现不佳、不能继续下去的评价员。

小组评价产品时，需开展重复实验。大多数测试中，一般重复2或3次。

采用统计分析以检查整个小组和每个评价员的一致性和重复性。

#### **F.4.2 评价员数量**

至少需要8位评价员，建议使用12~15位评价员。

#### **F.5 步骤**

——通常，样品呈送采用依次递送顺序。

——建议进行2到3次重复。

## F.6 统计分析和结果解释

### F.6.1 总则

数据核查后有多种统计方法进行结果分析。最常见是方差分析（ANOVA），双因素方差分析使用广泛，可以分析产品间以及评价员间的差异。其他差异来源如“重复”差异也需考虑。具体统计分析方法见附件H。

此外，采用多元统计方法可以探索属性间的交互情况。能同时开展几个产品对多个属性影响的分析方法很多。PCA（主成分分析）是最常用和知名的方法，特别适用于数据量大且前期背景信息较少的情况。PCA可在方差—协方差矩阵或相关矩阵上进行。第1种方法中所有属性权重并不相同，而是根据每种属性的差别力而变化。第2种方法中不考虑属性差别力，直接赋予所有属性相同的权重，前提是先删除差别力小的区分属性。由于描述性差导致属性被拒绝的临界概率一般是0.05，通常在0.10~0.50间变化。

### F.6.2 结果呈现

结果通常以数字和图表形式呈现（如蛛网图，剖面曲线及多元分析的相关图表）。

## F.7 变量

### F.7.1 定量描述分析<sup>3</sup>

定量描述分析（QDA<sup>3</sup>）法包括测量和定量分析。许多出版物和书籍中对QDA都进行了介绍，包括参考文献[9]、[27]和[28]。通过建立QDA，研究人员获得了人类感官测量方面的最佳实践和知识，且随着实践和信息基础的发展还在不断提升。QDA法从最初已发展成为今天的Tragon QDA<sup>3</sup>。

QDA法应用广泛，包括理解消费者描述和区分产品的语言、绘制感知到的产品相似性和差异、成分替换、新品开发、对比评价、索赔证明和广告等用途。当与消费者的情感测量（例如喜欢、偏好、态度、情感、包装、使用）相关联时，可以使用QDA数据确定影响消费者选择行为的最主要关键感官属性，并理解消费者独特偏好的原因。开发人员通过这些信息制定针对特定消费者利益和需求的产品，市场营销人员通过QDA得到的语言和感官特性来帮助传达这些利益。

<sup>3</sup> QDA<sup>®</sup>和 Spectrum<sup>™</sup> 是目前适合定量描述剖面分析的商业程序实例。这些信息是为了方便本文件的使用者而提供的，并不表示 ISO 对这些程序的认可。

### **F.7.1.1 感官属性**

建立感官属性的描述语通常需要8~12小时的小组讨论。这是一个迭代过程，每次讨论都是基于上一次讨论基础，直至创建出全面的描述语。部分描述语建立步骤中可能包括观察行为，包括在家和/或其他使用环境下。在描述语初步建立和试点测试后，可进行描述语补充。描述语使用前、使用中和使用后，创建30~40个或更多的感官属性充分地描述感觉都是正常情况。

### **F.7.1.2 标度**

QDA使用非结构化或半结构化的线性标度对感知的差异和强度进行测量和标度。

QDA使用非结构化或半结构化（6英寸/15厘米）标尺，从任一端锚定0.5英寸，用于感知的差异和强度的测量和标度。这些等间距标度在众多心理物理学出版物中均有介绍。

### **F.7.1.3 评价员**

#### **F.7.1.3.1 评价员数量**

建议采用12位评价员。

#### **F.7.1.3.2 资格和培训**

##### **F.7.1.3.2.1 招募与筛选**

候选者是所测产品类别的爱好者，招募数量大于产品所需平均人数，按感官敏锐度、语言流畅度以及可用性筛选约30名候选人/消费者。

##### **F.7.1.3.2.2 训练**

培训由主持人或小组组长组织开展。小组组长负责引导和推动小组展开讨论，制定小组活动时间表，协助小组一起建立待测产品的常用描述词汇。QDA主持人不用现场教学，但需要观察消费者在典型消费者应用情境下的消费行为和对产品的看法。

感官敏感度筛选是基于多达30项的测试，使用差别检验进行重复测试。测试包括了产品间、感兴趣类别、从简单、适中到难区分范围中可能存在的差异。

##### **F.7.1.3.2.3 表现评价**

QDA法提供了小组表现评价的统计方法，包括单个评价员在属性、重复性上的表现以及

相对于整个评价小组的整体差异。

下一步分析重点是感知产品的相似性和差异性。方差分析被列为指定分析方法，详细介绍见参考文献[5]。

#### F.7.1.4 步骤

建立感兴趣产品类别的典型评价步骤并标准化。

评价步骤需模拟消费者对此类产品的典型使用行为，可能需要多个评价步骤。

平衡区组设计后，建议至少重复3次。

#### F.7.1.5 统计分析和结果解释

采用方差分析（ANOVA）进行结果分析。

QDA数据结果以多种方式呈现，最常见的有蛛网图或雷达图，以及平均值和显著性差异。

### F.7.2 Spectrum<sup>TM4</sup> 法

#### F.7.2.1 原则

c法是描述性剖面法基础上，同时参照文献中方法进行定性属性和强度标度点测定。Spectrum<sup>TM4</sup>法起源于风味剖面法和质地剖面法，通过扩大标度、记录更多参考属性等将方法扩展到食品描述性评价之外，包括肤感和织物感。该方法从评价员筛选、小组领导、小组培训、验证以及培训后维护等各个阶段都有明确的方法和步骤。通过实践操作形成一个描述性评价小组，多环节和多类别评价时获取强大、可再现的统计数据。

#### F.7.2.2 感官属性

Spectrum<sup>TM4</sup>法采用一个经过多种产品类型（食品、香水、肤感、织物和纸张等）Spectrum<sup>TM</sup>培训的评价小组获取项目或产品的感官属性。评价员按照小组组长给的评价步骤和指导来建立属性或词汇。这些属性通过外部物理参比样和语义进行定义，建立的描述词能够描述和区分产品类别中的样品。所选的属性应可感知、主要的、可正交和单一的。

---

<sup>4</sup> QDA<sup>®</sup>和 Spectrum<sup>TM</sup> 是目前适合定量描述剖面分析的商业程序实例。这些信息是为了方便本文件的使用者而提供的，并不表示 ISO 对这些程序的认可。

### F.7.2.3 标度

Spectrum<sup>TM4</sup>标度是一个0~15分的强度标度，每一分又分成10等份，共计150个点。该标度是对前期标度进一步细化，使评价员能够用更小的差异点进行样品区分。Spectrum<sup>TM4</sup>标度属于通用标度，涵盖了食品或肤感所能体验到的所有强度范围。采用绝对强度进行标度，15分并非真正终点，可以根据需要延长。大多数产品强度均在0~15分范围内，如若需要，具有更高强度属性的分值可以超过15分（例如，极端的酸味，更高的油腻感）。为比较样品内和样品间强度，培训中开展强度值跨属性的关联设计。例如，强度为5的甜味与强度为5的苦味或强度为5的薄荷味具有相同强度。Spectrum<sup>TM4</sup>法通用标度是基于大样本规模的食物种类，涵盖的强度范围包括刚能检测（阈值）到轻微、到强、到非常强。Spectrum通用标度参比样的标度值是10个来自不同行业评价小组重复评价的结果。定期审查标准样的标度值变化（常规产品的变化），并在参考文献[4]中公布。

### F.7.2.4 评价员

#### F.7.2.4.1 资格

##### F.7.2.4.1.1 招募与选拔

通过多种资格认证措施从招募人员中筛选评价员。招聘期间，对潜在评价员进行筛查，针对健康问题、适用性和对工作内容/时间的兴趣度、基础视力以及描述产品或情境的口头表达能力。通常而言，根据项目特点，招募人数为所需评价员数量的3倍。通过预筛的人员随后进行面试及相关敏锐度和技能的选拔。达到预筛和筛选标准，且敏锐度和组内互动优秀的人员选拔为评价员（建议18~20人）。

##### F.7.2.4.1.2 培训

培训过程主要包括一系列高强度的训练周以及每个训练周后紧接着的若干练习，增强对研究材料的熟悉度。1个培训日程示例可能是28小时训练加45小时练习。培训首先介绍Spectrum<sup>TM4</sup>的整体理念，再采用相关领域的多个类别展开培训。培训和练习初期主要介绍总体方法、强度标度和如何为产品选择属性词汇/投票流程。后期的培训和练习可能集中在评价小组的待测项目等特定领域。

### F.7.2.4.1.3 表现评价

培训结束时应采用验证实验确定小组的可靠性和一致性。验证实验可以评估评价小组整体和单个评价员的表现，并帮助确定潜在的再培训领域。经过培训的评价小组数据应具有再现性，且能有效区分产品。验证实验设计采用样品和重复性评价测试这两方面的表现。建议不仅是培训结束时，还应该评价功能性工作小组的工作表现（例如，每年一次）。小组验证的概念类似于其他科学仪器的校准。

### F.7.2.4.1.4 评价员数量

Spectrum<sup>TM4</sup>评价小组通常由10~12位评价员组成。也可能根据评价员的工作年限、验证表现和数据采集方法（例如，一致数据）出现例外，使用较少数量的评价员（推荐最少6位评价员）。

### F.7.2.5 步骤

该方法使用5个规范步骤来构建产品空间属性的词典。

- a) 标准参比样——评价员可以体验多种产品，这些产品已定义和扩展感兴趣的领域。产品可能包括市场产品和用于建立理想术语的样品。
- b) 术语生成/分组——评价员根据自己经验对产品进行描述。评价组组长根据相似性将这些术语进行分组，并与全组协作初步拟出一个词典。
- c) 属性参比样——评价员接触已产生术语的参比样。
- d) 属性改进——小组组长引导讨论参比样的适用性和术语的冗余性。从词典中删除冗余、有或合成的术语。通过和讨论建立最终词典。
- e) 词典验证——采用成对样品测试建立的词典。词典应有效区分样品。该步骤也可用于词典或参比样改进。

### F.7.2.6 统计分析和结果解释

可使用参数统计进行Spectrum<sup>TM4</sup>小组数据分析。强度标度上有151个差异点，可用的标准统计量包括平均值、标准偏差、方差分析、平均分离步骤和相关性等。频次属于非参数数据，例如存在或不存在的属性。无论是个人数据还是一致性评价的平均值，均可采用多元分析技术进行产品分类（聚类分析，主成分分析/因子分析等）。

## 附件 G

## (资料性附录)

## 动态主导型感官属性测试 (TDS)

**G.1 原则**

动态主导型感官属性测试法是一种描述性感官方法,包括持续评价哪种感觉占主导地位。对于评价者来说,“主导感觉”被定义为在给定时间引起注意的感觉,这并不意味着这种感觉必须是产品中非常强烈或最激烈的感觉。与时间强度相比,这种方法考虑到随着时间的推移的多维性。对于每次评价,该方法允许在整个品尝过程中的不同时间收集引用的一系列感官属性。评价员可以在一张纸上提供他的答案,只要他有一个秒表。但对于评价员来说,使用计算机更舒适,这样评价员不必担心时间,可能更专注于识别主导属性。收集的数据可以进行统计处理,以计算每个属性的主导曲线随着时间的推移。

**G.2 属性**

提供给评价员一个属性列表;他/她必须在该列表中选择在整个品尝过程中占主导的属性。因此,属性列表中的定义是确定评价员回答的关键要素。TDS属性列表可以从属性生成步骤开始构建,如定量描述性分析,然后通过一致性或统计分析进行删减。

建议该列表中最多包含10到12个属性。列表太长,部分评价员可能无法对应整个列表。

属性在列表中的顺序通常会影响属性的选择顺序:列表顶部的属性可能比底部属性被更早选择,建议平衡各评价员的属性顺序。但对任一评价员来说,属性的顺序在整个实验过程中应该始终保持一致(无论产品和重复过程如何变)。

可以在同一列表中组合多个属性类型(例如,风味属性和质地属性)。由于强度级别与属性类型间存在较大差异,建议至少针对每种类型属性单独评价,也可选择将几种不同类型属性放在1个轮次中评价,以获取补充信息。

**G.3 标度**

该方法最简便的形式就是不使用标度。评价时仅需在属性表中选择占主导的属性。



## G.4 评价员

### G.4.1 培训

TDS的培训与定量剖面有所不同：不需要对强度进行培训，更多是针对不同感官属性进行训练，提升主导感觉的判断。优选评价员需为每个TDS属性开发一致定义。使用参比样帮助他们识别每个属性。

培训期间（次数根据评价员而定，1到4次）应着重引导评价员熟悉TDS的特性（例如，使用计算机系统）。

一旦评价员熟悉了TDS法，该方法所需的培训课程数量一般少于定量描述分析。通常2到4个轮次就可以适应属性列表。

### G.4.2 评价员数量/重复次数

需要经过培训的评价员，建议TDS使用比定量剖面更多些的评价员，如约16位评价员，及重复2次或更多，保证每个产品至少收集30个评价。

## G.5 步骤

——将一口量大小的产品一次性提供给评价员，要求他们整个放在嘴里，同时开始评价（通常通过点击“开始”按钮）。

——当评价员认为主导感觉知觉已经改变时，他/她必须在属性列表中指出新的主导感觉。评价员重复该过程直到感知结束。点击“停止”按钮，停止计时。

——允许评价员自由地给相同产品选择相同属性，或者完全不选择列表中的一个或多个特定属性。

——测试过程应明确指示何时开始和停止评价，确保收集可靠的数据。

## G.6 统计分析和结果解释

### G.6.1 原始数据处理

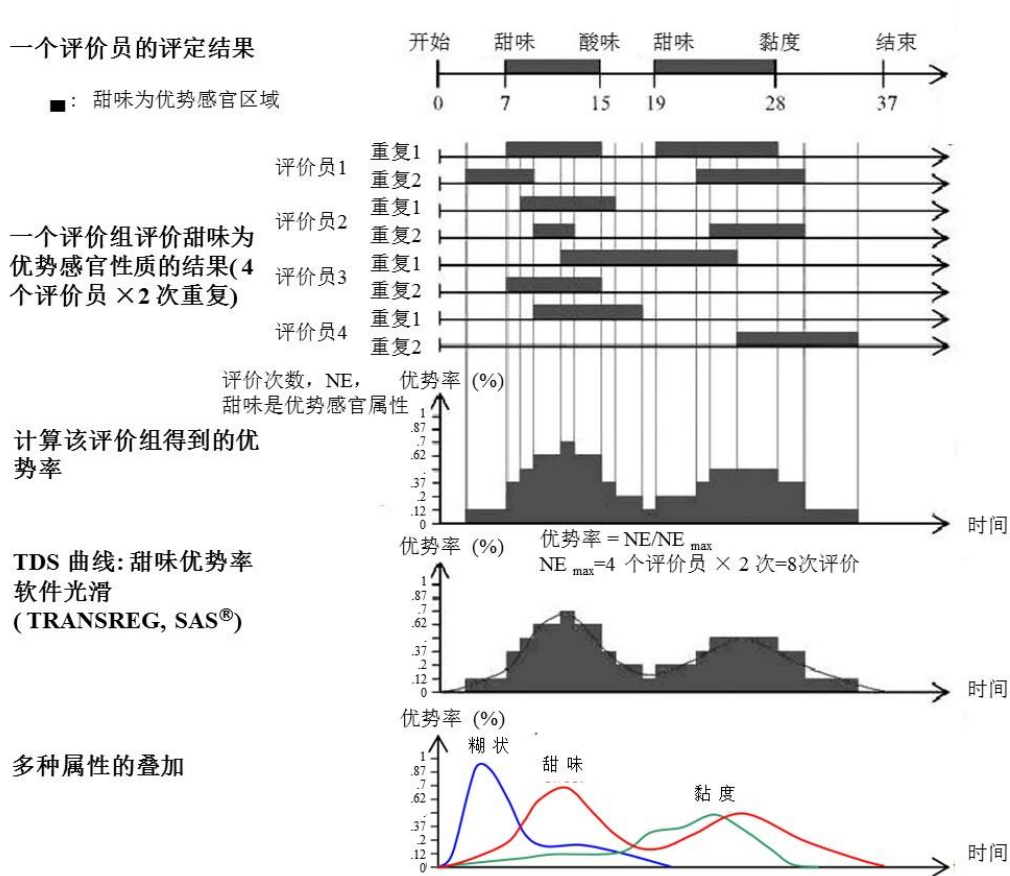
感官分析师必须确定每次评价期间占主导的属性名称和计时开始后所需时间。占主导的属性是指从被选中一直到选择其他属性前的属性。

G.6.2 TDS 曲线

图G.1显示了甜味属性TDS曲线的计算。该步骤分别考虑了每个属性。对于每个时间点，计算选择某一占主导属性的评价员（评价员×重复）比例。将这些平滑（非迭代平滑样条）比例与时间进行对应作图，计算TDS曲线。1个产品所有属性的TDS曲线均在同一图中展示。

为了更深入展示TDS图形，可以绘制两条线。一条是“偶然水平”线，是某一属性可以被偶然感知的主导率。它的值 $P_0$ 等于 $1/p$ ，其中 $p$ 是属性的数量。另一条是“显著性水平”，最小值。这个比例应被认为是显著高于 $P_0$ ，通过二项式分布计算获得。

注：由于进入口腔的持续时间不同，评价员之间感觉的时间尺度也不同。计算 TDS 曲线时需考虑这一现象，可以对每个评价员的数据进行标准化，即将开始评分到吞咽视为一个咀嚼期，使 x 轴显示从 x=0（占主导的属性第一次提出）到 x=100（吞咽），计算各个主导属性持续时间占整个咀嚼期的比例。非标准化的曲线很有趣，可以看出产品之间持续时间的差异，但由于所有评价并非同时结束，难以看懂，而标准化后的曲线提升了产品性能（更高的主导率，特别是在刚开始和结束前）。



图G.1 - 甜味属性计算示例（4个评价员，2次重复）

### G.6.3 TDS 差异曲线

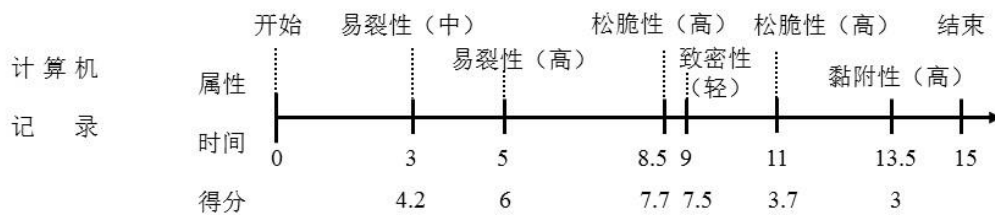
TDS图在同一图上显示了 $p$ 个属性的曲线。为了比较2个产品，可以将主导率差异的 $p$ 曲线进行叠加。这些差异仅在与零显著不同时绘制，凸显了产品随时间的差异。使用通常的测试来比较2个二项式比例，获得每个时间差异曲线的显著性范围。

### G.6.4 其他统计分析

也可使用其他几种统计方法。例如：研究总体产品长度的“停止”次数的方差分析，随时间变化的主导持续时间的PCA或典型变化分析（CVA），以概述随时间变化的平均产品差异，PCA随时间变化的产品轨迹，突出其随时间演变的产品差异。

### G.7 相似和方法变化

TDS法通常由筛选过的经过培训的评价员进行。也可由“当地消费者”通过TDS计算机系统开展。属性列表必须足够简单，易于被准消费者理解。刚开始的几分钟，需要花点时间向消费者介绍感觉的时间概念，让他们对属性定义清晰，并指导他们使用TDS计算机系统。有必要给消费者几个热身样品，使其熟悉方法和属性。对于本地消费者的TDS，通常不会重复。因此，建议至少使用30个受访者。根据参考文献[28]，可要求评价员对所选属性的强度进行评价。这个要求使任务难度显著增加，需要开展相应的培训。为了不混淆2个不同的认知过程：占主导的属性（定性任务）选择和强度评价（定量任务），建议使用按钮而不是强度标度来指明强度。图G.2给出一个饼干响应的例子。



图G.2 - TDS的计算机记录示例

酥脆的属性占了主导地位两次，第一次的持续时间是3.5秒（在5-8.5秒之间，得分：4.2），第二次持续了2.5秒（在11和13.5秒之间，得分：3.7）。TDS中，酥脆这个属性占主导的总持续时间是 $3.5\text{ s} + 2.5\text{ s} = 6\text{ s}$ 。脆性强度的评分通过2个评分上的平均加权计算（持续时间）： $4.2 * 3.5/6 + 3.7 * 2.5/6 = 3.99 \approx 4$ 。

## 附件 H

## (资料性附录)

## 评价小组量化一个属性时的单变量分析

**H.1 总则**

该方法尤其适用于定量描述剖面、Spectrum™和QDA<sup>5</sup>法，提供一种对产品间相对重要属性进行比较的方法。

**H.2 双因素方差分析（产品和评价员因素），无重复****H.2.1 ANOVA 模型**

通常，基本模型是

$$y_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} \quad (\text{H.1})$$

其中，

$y_{ij}$ 是第*i*个评价员（*i*: 1到 $\alpha$ ）对第*j*个产品（*j*: 1到*p*）的评分值；

$\mu$ 是评分值的平均值；

$\alpha_i$ 是评价员影响；被假定为随机的；

$\beta_j$ 是产品影响；被假定为固定的；

$\alpha\beta_{ij}$ 是评价员\*产品的交互作用。

实验设计通常称为随机完全区组设计（1个评价员是1个区组），也称为内部评价员设计或重复测量设计。

**H.2.2 数例**

10名评价员对3种产品展开评价，结果如表H.1所示。对H.1中数据进行双因素方差分析，结果如表H.2所示。

<sup>5</sup> QDA®和 Spectrum™ 是目前适合定量描述剖面分析的商业程序实例。这些信息是为了方便本文件的使用者而提供的，并不表示 ISO 对这些程序的认可。

表H.1 无重复研究结果实例

评价员	产品 1	产品 2	产品 3
1	2	3	3
2	8	6	7
3	5	4	5
4	4	3	5
5	5	3	5
6	4	4	8
7	4	5	5
8	6	2	4
9	4	3	8
10	5	6	5
平均值	4.7	3.9	5.5

表H.2 表H.1的双因素方差分析

方差来源	因素状态	自由度 ( <i>df</i> )	平方和 ( <i>SS</i> )	均方 ( <i>MS</i> )	<i>F</i> 值	<i>p</i> 值
评价员	随机	$\alpha - 1 = 9$	34.300	$MS_{\alpha} = 3.811$	$MS_p / MS_{\alpha * p} = 3.95$	0.038
产品	固定	$p - 1 = 2$	12.800	$MS_p = 6.400$		
评价员*产品	随机	$(\alpha - 1) * (p - 1) = 18$	29.200	$MS_{\alpha * p} = 1.622$		

通过评价员\*产品的交互作用对产品因素进行验证。表H.1的描述词并不显著。

注：如果评价员因素是固定的，例如为研究评价员在小组培训期间表现，除非假定评价员\*产品的交互作用进行实验误差估算，否则 2 个固定因素均不能被检验。有了这个假设，就可以针对误差对 2 个固定因素进行检验，即评价员\*产品的相互作用。

### H.3 双因素方差分析（产品和评价员因素），有重复

#### H.3.1 ANOVA 模型

通常，基本模型是：

$$y_{ijr} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijr} \quad (\text{H.2})$$

其中：

$y_{ijr}$ 是第*i*个评价员（*i*：1到 $\alpha$ ）对第*j*个产品（*j*：1到*p*）在第*r*次重复中（*r*：1到*r*）的评分值；假定所有产品的都是相同的重复次数；

$\mu$ 是评分值的平均值；

$\alpha_i$ 是评价员影响；被假定为随机的；

$\beta_j$ 是产品影响；被假定为固定的；

$\alpha\beta_{ij}$ 是评价员\*产品的交互作用；

$\varepsilon_{ijr}$ 是随机重复误差；是对实验误差的估计。

这种设计通常称为随机完全区组设计，每个实验单元不止一个观测值，由产品和评价员的交互作用形成。该模型中重复是随机，但也具有相关性，例如，在同一批次产品、同一轮次中进行重复性研究。事实上，随机完全区组设计中，无需将重复1中产品P1的分值与重复1中产品P2或P3的分值进行关联。一个例子是在单个品尝轮次中，10位评价员均需评价来自4个不同品种的2根胡萝卜样品。每次重复提供给每个评价员1根胡萝卜，因此每个品种需要20根胡萝卜，每个评价员需评价 $4 * 2 = 8$ 根胡萝卜。

### H.3.2 数例

表H.1的评价员对产品进行了2次评价，表H.1中只给出了重复1的结果。表H.3给出了2次重复的结果，双因素方差分析结果如表H.4所示。

表H.3 研究结果示例（两次重复）

评价员	产品 1		产品 2		产品 3	
	重复	重复	重复	重复	重复	重复
	1	2	1	2	1	2
1	2	3	3	3	3	5
2	8	7	6	6	7	7
3	5	6	4	6	5	7
4	4	5	3	5	5	8
5	5	7	3	5	5	6
6	4	6	4	2	8	8
7	4	5	5	8	5	7
8	6	8	2	5	4	6
9	4	3	3	2	8	7
10	5	7	6	4	5	8

表H.4 表H.3双因素方差分析

方差来源	因素状态	自由度 ( <i>df</i> )	平方和 ( <i>SS</i> )	均方 ( <i>MS</i> )	<i>F</i> 值	<i>p</i> 值
评价员	随机	$r - 1 = 9$	48.350	$MS_{\alpha} = 5.372$	$MS_{\alpha}/MS_e = 3.322$	0.0063
产品	固定	$p - 1 = 2$	38.033	$MS_p = 19.017$	$MS_p/MS_{\alpha*p} = 6.673$	0.0068
评价员*产品	随机	$(r - 1) * (p - 1) = 18$	51.300	$MS_{\alpha*p} = 2.850$	$MS_{\alpha*p}/MS_e = 1.763$	0.0824
误差		$(r - 1) * a * p = 30$	48.500	$MS_e = 1.617$		
总计		$a * p * r - 1 = 59$	186.183			

通过评价员\*产品的交互作用对产品因素进行验证。 $\alpha=0.05$ 时，产品因素显著。

评价员因素和产品\*评价员的交互作用是针对误差进行检验。评价员因素显著，产品\*评价员的交互作用不显著。

注：如果评价员是固定的，例如为研究评价员在小组培训期间的表现，则表H.4的3个因素均是针对误差进行检验。则产品的*F*值等于11.763 ( $p$ 值=0.00017)；另外2个*F*值不变。

## H.4 三因素方差分析（产品，评价员和重复因素）

### H.4.1 重复因素与其他 2 个因素交互

#### H.4.1.1 ANOVA 模型

含有8个部分：

$$y_{ijr} = \mu + \alpha_i + \beta_j + \delta_r + \alpha\beta_{ij} + \alpha\delta_{ir} + \beta\delta_{jr} + \alpha\beta\delta_{ijr} \quad (\text{H.3})$$

其中

$y_{ijr}$ 是第*i*个评价员 ( $i: 1$ 到 $\alpha$ ) 对第*j*个产品 ( $j: 1$ 到 $p$ ) 在第*r*次重复中 ( $r: 1$ 到 $r$ ) 的评分值；

$\alpha_i$ 是评价员作用；被认为是随机的；

$\beta_j$ 是产品作用；被认为是固定的；

$\delta_r$ 是重复作用；被认为是固定或随机的因素；可能最相关的是随机因素；

$\alpha\beta_{ij}$ 是评价员\*产品的交互作用；

$\alpha\delta_{ir}$ 是产品\*重复的交互作用；

$\beta\delta_{jr}$ 是产品\*重复的交互作用；

$\alpha\beta\delta_{ijr}$ 是评价员\*产品\*重复的交互作用。

这种设计通常被称为双因素完全重复测量设计或三因素交叉设计。

当每个实验单元的观测值多于1个时，如评价员-产品-重复3个1组，则该设计有第9个组分，可对实验误差进行估计。

通常，剖面研究中，重复因素是评价轮次。例如，H.3中的重复1在星期二进行，重复2在同一周的星期四进行，则重复1为轮次1，重复2为轮次2。

**H.4.1.2 数例**

当评价员因素随机，产品和轮次因素固定时，表H.3数据进行三因素方差分析，结果如表H.5所示。

表H.5 表H.3的三因素方差分析

方差来源	因素状态	自由度 (df)	平方和 (SS)	均方 (MS)	F值	p值
评价员	随机	$\alpha - 1 = 9$	48.350	$MS_{\alpha} = 5.372$	$MS_p/MS_{\alpha*p} = 6.673$ $MS_s/MS_{\alpha*s} = 8.405$	0.0067 0.0176
产品	固定	$p - 1 = 2$	38.033	$MS_p = 19.017$		
轮次	固定	$r - 1 = 1$	16.017	$MS_s = 16.017$		
评价员*产品	随机	$(\alpha - 1) * (p - 1) = 18$	51.300	$MS_{\alpha*p} = 2.850$	$MS_{p*s}/MS_{\alpha*p*s} = 0.787$	0.47
产品*轮次	固定	$(p - 1) * (s - 1) = 2$	1.233	$MS_{p*s} = 0.617$		
评价员*轮次	随机	$(\alpha - 1) * (s - 1) = 9$	17.150	$MS_{\alpha*s} = 1.906$		
评价员*产品*轮次	随机	$(\alpha - 1) * (p - 1) * (s - 1) = 18$	14.100	$MS_{\alpha*p*s} = 0.783$		
总计		$\alpha * p * s - 1 = 59$				

针对评价员\*产品的交互作用检验产品因素，结果显著。根据评价员\*轮次的交互作用检验轮次因素，结果显著。针对评价员\*产品\*轮次的交互作用检验产品\*轮次的交互作用，结果不显著（参考文献[10]，361-367）。

其他所有因素不能被检验。但是，如果评价员\*产品\*轮次的相互作用被假定为可忽略的，则该部分可被认为是对实验误差的估计。这种假设下，评价员因素、评价员\*产品的交互作用和评价员\*轮次的交互作用均可以针对评价员\*产品\*轮次的交互作用进行检验。

注1：当轮次因素随机时，该设计具有2个随机因素（评价员和轮次因素）和1个固定因素（产品因素）。

——针对评价员\*轮次的交互作用检验评价员因素；F值等于 $5.372 / 1.906 = 2.818$ （p值= 0.069）。

——针对评价员\*轮次的交互作用检验轮次因素；F值等于 $16.017 / 1.906 = 8.403$ （p值= 0.017）。

——针对评价员\*产品\*轮次的交互作用检验评价员\*产品的交互作用，F值等于 $2.850/0.783 = 3.640$ （p值= 0.0044）。

——针对评价员\*产品\*轮次的交互作用检验产品\*轮次的交互作用，如表H.5所示（p值= 0.47）。



——产品因素和评价员\*轮次交互作用不能被检验。对于产品因素而言，这非常困扰。但是，可通过计算类似F的F'克服这个缺陷（参考文献[10]，368-370）：

$$F' = MS_p / (MS_{\alpha*p} + MS_{p*s} - MS_{\alpha*p*s}) = 19.017 / (2.850 + 0.617 - 0.783) = 19.017 / 2.684 = 7.083$$

F'分母的自由度由Satterthwaite的近似值得出：

$$\frac{(MS_{\alpha*p} + MS_{p*s} + MS_{\alpha*p*s})^2}{\frac{MS_{\alpha*p}^2}{df_{\alpha*p}} + \frac{MS_{p*s}^2}{df_{p*s}} + \frac{MS_{\alpha*p*s}^2}{df_{\alpha*p*s}}}$$

对于这个例子：

$$df = (2.6842)/(0.451 + 0.190 + 0.034) = 7.204 / 0.675 = 10.672$$

分子自由度为2和分母自由度为11（最接近10.67的整数值）的F'=7.085，概率等于0.0027。产品因素显著。

注2：当3个因素固定时，除非假定评价员\*产品\*轮次的交互作用可忽略，否则不能检验因素效应。这种假设下，所有6个因素（3个主要因素和3个相互作用）的检验均是针对评价员\*产品\*轮次的交互作用进行的。

#### H.4.2 重复因素嵌套在产品因素中

这种类型设计需仔细分析实验条件。以下有3个例子，例1：肉烤熟后的感官分析，重点关注猪肉品种的影响。共有5个品种，每个品种有4头猪。这4头猪（重复）嵌套在每个品种（主要因素）中。例2：啤酒生产商正在生产一种啤酒新产品，有5种替代配方。每个配方有3个批次。批次因素嵌套在配方因素内。例3：研究6种马铃薯的差异。每个品种种3行，随机分配一块田中。马铃薯收获后，将每行马铃薯放在不同容器中。行因素嵌套在品种因素中。

这种设计通常被称为嵌套因子设计。嵌套因子名称因方法而异，本节中被称为为批次。

##### H.4.2.1 方差分析模型

模型有5个组成：

$$y_{ijr} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \delta_r + \beta\delta_{jr} \quad (\text{H.4})$$

其中

$y_{ijr}$ 是第*i*个评价员（*i*: 1到 $\alpha$ ）对第*r*次批次中（*r*: 1到*r*，假设所有产品都是相同数量的批次）的第*j*个产品（*j*: 1到*p*）的评分值；

$\alpha_i$ 是评价员作用；通常是一个随机因素；

$\beta_j$ 是产品作用；通常是一个固定因素；

$\alpha\beta_{ij}$ 是评价员\*产品的交互作用；

$\delta_r$ 是批次作用；通常是一个随机因素，但也可能是一个固定因素；

$\beta\delta_{jr}$ 是产品\*批次的交互作用；

当产品和批次都固定时，唯一的随机因素是评价员。表H.3分析很简单，结果如表H.6所示。

表H.6 表H.3的嵌套方差分析（产品和批次因素固定）

方差来源	因素状态	自由度 (df)	平方和 (SS)	均方 (MS)	F值	p值
评价员	随机	$\alpha - 1 = 9$	48.350	$MS_{\alpha} = 5.372$	$MS_p/MS_{\alpha*p} = 6.673$	0.0068
产品	固定	$p - 1 = 2$	38.033	$MS_p = 19.017$		
评价员*产品	随机	$(\alpha - 1) * (p - 1) = 18$	51.300	$MS_{\alpha*p} = 2.850$	$MS_b/MS_{\alpha*b} = 4.970$	0.0071
批次（产品）	固定	$(b - 1) * p = 3$	17.250	$MS_b = 5.750$		
评价员*批次（产品）	随机	$(\alpha - 1) * (b - 1) * p = 27$	14.100	$MS_{\alpha*b} = 1.157$	$MS_{p*s}/MS_{\alpha*p*s} = 0.787$	
总计		$a * p * b - 1 = 59$				

针对评价员\*产品检验产品因素，针对嵌套在产品因素中的评价员\*批次交互作用检验批次因素，这2个因素均显著。

注 1：当批次因素随机时，此时有 2 个随机因素（评价员和轮次因素）和 1 个固定因素（产品因素）。评价员因素、评价员\*产品交互作用和批次因素是针对评价员\*批次（产品）的交互作用进行检验。评价员F值等于  $5.372/1.157 = 4.643$ （p值= 0.096），评价员\*产品的交互作用F值等于  $2.850/1.157 = 2.463$ （p值= 0.0166）。批次F值已经在表 H.6 中计算出。

产品因素用准F或F'进行检验。

第一步：计算均方，通过以下计算获得：

$$MS_{test.P} = MS_{b(p)} + MS_{\alpha*p} - MS_{\alpha*b} = 5.750 + 2.850 - 1.157 = 7.443$$

第二步：找到 $MS_{test.P}$ 的df数，通过以下计算获得：

$$Den = \frac{MS_{test.P}^2}{\frac{MS_{b(p)}^2}{df_{b(p)}} + \frac{MS_{\alpha*p}^2}{df_{\alpha*p}} + \frac{MS_{\alpha*b}^2}{df_{\alpha*b}}} = \frac{7.443^2}{\frac{5.75^2}{3} + \frac{2.85^2}{18} + \frac{1.157^2}{27}} = 4.808$$

第三步：计算F'：

$$F' = \frac{MS_b}{MS_{test.P}} = \frac{19.017}{7.443} = 2.555$$

最后一步：在H0下找到F'的相关概率，分子df= 2、分母df= 5（最接近4.808的整数）时值为0.172。产品因素不显著。

注 2: 当 3 个因素固定, 所有的检验均针对评价员\*批次 (产品) 的交互作用进行。评价员  $F$  值等于 4.64 ( $p$  值=0.0009); 产品  $F$  值等于 16.43 ( $p$  值=0.0002), 评价员\*产品的交互作用  $F$  值等于 2.46 ( $p$  值=0.017)。表 H.6 中计算已计算出批次  $F$  值。

## H.5 随时间推移的评价数据分析

### H.5.1 总则

表H.4中共有3个因素: 评价员, 产品和轮次 (或批次)。本节中一样, 只是用时间因素替代了轮次因素。尽管与产品因素相比, 轮次因素是次要因素, 但本节中, 时间因素与产品因素一样显著, 因为产品随时间而变化。表H.7给出1个例子; 借鉴了1项研究数据, 10位评价员以0~20分对4种牙膏的清新特性展开评价。总和如表H.8所示。

表H.7 随时间推移4次评价4种牙膏的清新特性强度

P1				P2				P3				P4			
T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
8	12	9	8	11	10	2	2	9	10	7	7	6	3	3	1
14	17	15	15	18	17	14	14	16	18	14	15	15	18	17	14
16	17	12	12	15	14	10	9	14	10	8	9	13	15	9	6
15	19	14	14	13	17	10	10	5	15	10	10	10	13	8	7
9	8	6	6	7	9	5	5	7	9	3	4	6	7	4	2
20	20	16	16	19	19	16	16	19	17	15	15	19	17	15	12
11	14	11	10	10	10	7	8	11	14	9	8	10	15	9	8
15	19	13	13	15	17	7	6	12	15	9	10	13	14	10	7
12	12	5	4	10	10	4	3	10	10	2	2	9	9	3	8
10	10	7	7	13	8	3	2	13	11	5	6	11	9	5	2

表H.8 表H.7的总和

	时间 1	时间 2	时间 3	时间 4	
产品 1	130	148	108	105	491
产品 2	131	131	78	75	415
产品 3	116	129	82	86	416
产品 4	112	120	83	67	382
	489	528	351	333	1701

### H.5.2 所有评价员对 4 种产品进行评价: 产品、时间和评价员三因素交互

表H.7中, 第1行给出了评价员1的数据; 第二行给出评价员2的数据, 以此类推。

**H.5.2.1 三因素方差分析**

结果如表H.9中所示，其显示方式与表H.5略有不同。检验步骤按表H.5进行。产品因素（固定因素）是针对产品\*评价员的交互作用进行检验。时间因素（固定因素）是针对时间\*评价员的交互作用进行检验，而时间\*产品的交互作用是针对误差（即产品\*时间\*评价员的交互作用）进行检验。

表H.9 表H.7的三因素方差分析

方差来源	因素状态	自由度 (df)	平方和 (SS)	均方 (MS)	F值	p值
评价员	随机	$\alpha - 1 = 9$	2056.18	$MS_{\alpha} = 228.46$		
产品	固定	$p - 1 = 3$	161.22	$MS_p = 53.74$	$MS_p/MS_{\alpha*p} = 8.842$	0.0003
评价员*产品	随机	$(\alpha - 1) * (p - 1) = 27$	164.09	$MS_{\alpha*p} = 6.08$		
时间	固定	$(t - 1) = 3$	716.12	$MS_t = 238.71$	$MS_t/MS_{t*\alpha} = 38.85$	0.0000
时间*产品	固定	$(t - 1) * (p - 1) = 9$	45.21	$MS_{t*p} = 5.02$		
时间*评价员	随机	$(t - 1) * (\alpha - 1) = 27$	167.19	$MS_{t*\alpha} = 6.19$	$MS_{t*p}/MS_{error} = 2.43$	0.0168
误差		$(\alpha - 1) * (p - 1) * (t - 1) = 81$	167.23	$MS_{error} = 2.06$		
总计		159	3477.24			

**H.5.2.2 重复测量的多元方差分析**

三因素方差分析需要关于方差-协方差（称为球度）的假设。如果用时间因素（T1, T2, T3, T4）作为因变量进行多元方差分析（MANOVA），则这个假设的限制性较小。关于表H.10的解释，请查阅统计类书籍。

表H.10 所有评价员评价四种产品的表H.7的多元方差分析

	测试	值	F	df	误差df	p值
时间	Wilk	0.08710	87.34527	3	25.000 00	0.000000
	Pillai	0.91290	87.34527	3	25.000 00	0.000000
	Hotelling	10.48143	87.34527	3	25.000 00	0.000000
	Roy	10.48143	87.34527	3	25.000 00	0.000000
时间*产品	Wilk	0.49116	2.29936	9	60.994 04	0.026935
	Pillai	0.59736	2.23762	9	81.000 00	0.027499
	Hotelling	0.85590	2.25070	9	71.000 00	0.028152
	Roy	0.48699	4.38287	3	27.000 00	0.012273
时间*评价员	Wilk	0.17562	2.22086	27	73.655 21	0.003738
	Pillai	1.20213	2.00592	27	81.000 00	0.008874
	Hotelling	2.69818	2.36507	27	71.000 00	0.002065
	Roy	1.6219	4.86657	9	27.000 00	0.000640

多元方差分析结果表明，4个测试的时间\*产品的交互作用显著。

注：可创建一个新变量代表时间的变化量，然后执行另一个MANOVA：Diff<sub>1</sub> = 时间 2-时间 1，Diff<sub>2</sub> = 时间 3-时间 2，Diff<sub>3</sub> = 时间 4-时间 3。这些新变化或差异变量可衡量每个时间间隔内发生的变化（见参考文献[8]）。

**H.5.3 每个评价员评价 4 种产品中 1 种：评价员因素嵌套在与时间因素交互的产品因素中**

因此，表H.7中是40位评价员；10位评价产品1，10位评价产品2，10位评价产品3和10位评价产品4。

由表H.7数据得到 表H.11和表H.12。

表H.11 表H.7的嵌套方差分析

方差来源	因素状态	自由度 (df)	平方和 (SS)	均方 (MS)	F值	p值
产品 评价员 (产品)	固定	$p - 1 = 3$	161.22	$MS_p = 53.74$	$MS_p/MS_{\alpha*p} = 8.84$	0.0003
	随机	$(\alpha - 1) * p = 36$	2220.28	$MS_{\alpha*p} = 61.67$		
时间 时间*产品	固定	$(t - 1) = 3$	716.12	$MS_t = 238.71$	$MS_t/MS_{error} = 77.09$	0.00000
	固定	$(t - 1) * (p - 1) = 9$	45.21	$MS_{t*p} = 5.02$	$MS_{t*p}/MS_{error} = 1.62$	
误差		108	334.43	$MS_{error} = 3.10$		

表H.6中，批次因素是嵌套产品因素中的一个固定因素，而本节中嵌套在产品因素中的评价员因素是一个随机因素，时间因素和时间\*产品交互作用的检验并不相同。

表H.12 当评价员嵌套在产品中时表H.8的多元方差分析

	测试	值	F	df	误差df	p值
时间	Wilk	0.149451	64.49977	3	34.0000	0.000000
	Pillai	0.850549	64.49977	3	34.0000	0.000000
	Hotelling	5.691156	64.49977	3	34.0000	0.000000
	Roy	5.691156	64.49977	3	34.0000	0.000000
时间*产品	Wilk	0.636200	1.88093	9	82.8977	0.066100
	Pillai	0.401348	1.85334	9	108.0000	0.066810
	Hotelling	0.512838	1.86141	9	98.0000	0.066763
	Roy	0.339206	4.07047	3	36.0000	0.013742

注：MANOVA 是以时间因子 (T1, T2, T3, T4) 作为因变量执行。

对Roy检验而言，时间\*产品的交互作用是显著。p值接近其他3个检验的显著值。

## 参考文献

### 一般性

- [1] ISO 704, Terminology work — Principles and methods
- [2] ISO 4121, Sensory analysis — Guidelines for the use of quantitative response scales
- [3] ISO 13300-1, Sensory analysis — General guidance for the staff of a sensory evaluation laboratory —Part 1: Staff responsibilities
- [4] ISO 13300-2, Sensory analysis — General guidance for the staff of a sensory evaluation laboratory —Part 2: Recruitment and training of panel leaders
- [5] GACULA Jr. M.C., 1997. Descriptive Sensory Analysis in Practice. Food & Nutrition Press, 712 pp.
- [6] LAWLESS H.T.. Quantitative Sensory Analysis. Psychophysics, Models and Intelligent Design. Wiley & Sons, 2013, p.
- [7] LAWLESS H.T.. HEYMANN H., 2010, Sensory Evaluation of Food. Principles and Practices. 2nd ed., Springer, 596 pp.
- [8] MEILGAARD M.C.. CIVILLE G.V., CARR B.T., 2007. Sensory Evaluation Techniques. 4th ed., CRC Press, 448 pp.
- [9] STONE H.. BLEIBAUM R.N., THOMAS H.T., 2012. Sensory Evaluation Practices, 4th ed., Academic Press, 438 pp.

### 统计分析

- [10] ABDI H.. EDELMAN B., VALENTIN D., DOWLING W.G., 2009. Experimental Design and Analysis for Psychology, Oxford University Press, 538 pp.
- [11] GACULA M.. SINGH J., BI J., ALTAN S. Statistical Methods in Food and Consumer Science, Second Edition, 2009, p.
- [12] HOWELL D.. Statistical Methods for Psychology. Duxbury, Fifth Edition, 2002, p.
- [13] NAES T.. BROCKHOFF P.B., TOMIC O., 2010. Statistics for Sensory and Consumer Science, Wiley & Sons, 287 pp.

### 一致性剖面

- [14] BRANDT M.A.. SKINNER E.Z., COLEMAN J.A. Texture Profile Method. J. Food Sci. 1963, 28 (4) pp. 404–409

[15] CAIRNCROSS S. E.. and SJOSTROM L. B. Flavor Profiles - A New Approach to Flavor Problems. Food Technol. 1950, 4 pp. 308–311

[16] CHAMBERS D.H.. ESTEVE E., RETIVEAU A. Effect of milk pasteurization on flavor properties of seven commercially available French cheese types. J. Sens. Stud. 2010, 25 (4) pp. 494–511

#### 偏离参比剖面

[17] LARSON-POWERS N.M.. PANGBORN R.M. Descriptive analysis of the sensory properties of beverage s and gelatine containing sucrose and synthetic sweeteners. J. Food Sci. 1978, 43 (11) pp. 47–51

[18] BOYLSTON T.D.. REITMEIER C., MOY J., MOSHER G.A., TALADRIZ L. Sensory quality and nutrient composition of three Hawaiian fruits treated by X-irradiation. J. Food Qual. 2002, 25 (5) pp. 419–433

#### 闪现剖面

[19] SIEFFERMANN J.-M.. 2000. Le profil Flash. Un outil rapide et innovant d'évaluation sensorielle descriptive, AGORAL 2000, XIIèmes rencontres, Montpellier, France In : Tec. & Doc., L'innovation: de l'idée au succès, 335-340.

[20] DELARUE J.. SIEFFERMANN J-M., 2004. Sensory mapping using Flash profile - Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. Food Quality and Preference, 15 (4), 383-392.

#### 自选剖面

[21] WILLIAMS A. A.. LANGRON S. P. The use of free-choice profiling for the evaluation of commercial ports. J. Sci. Food Agric. 1984, 35 (5) pp. 558–568

[22] GUÀRDIA M. D.. AGUIAR A. P. S., CLARET A., ARNAU J. and GUERRERO L. Sensory characterization of dry-cured ham using free-choice profiling. Food Qual. Prefer. 2010, 21 (1) pp. 148–155

[23] PIGGOT J.R.. WATSON M.P. A comparison of free-choice profiling and the repertory grid method in the flavor profiling of cider. J. Sens. Stud. 1992, 7 (1) pp. 133–145

#### 定性感官剖面

[24] CAMPO E., BALLESTER J., LANGLOIS J., DACREMONT C., VALENTIN D. Comparison of conventional descriptive analysis and a citation frequency-based descriptive method for odor profiling: An application to Burgundy Pinot noir wines. *Food Qual. Prefer.* 2010, 21 (1) pp. 44–55

#### 定量描述性剖面

[25] FINDLAY C.J., CASTURA J.C., and LESSCHAEVE I. Feedback calibration: A training method for descriptive panels. *Food Qual. Prefer.* 2007, 18 (2) pp. 321–328

[26] LAWLESS L. J.R., CIVILLE G.V., 2013. Developing Lexicons: Review A. *J. Sens. Stud.*, 28 (4) pp. 270–281

#### 定量描述性分析

[27] STONE H., SIDEL J., OLIVER R.S., WOOLSEY A. and SINGLETON R.C. Sensory evaluation by quantitative descriptive analysis. *Food Technol.* 1974, 28 (1) pp. 24–34

[28] SEE ALSO STONE ET AL. 2012, pp. 250-274 (op cit.)

#### 光谱

[29] CIVILLE G. V., DUS C. A. Development of terminology to describe the handfeel properties of paper and fabrics. *J. Sens. Stud.* 1990, 5 (1) pp. 19–32

[30] MUNOZ A. M. and CIVILLE G. V.. Universal, product and attribute specific scaling and the development of common lexicons in descriptive analysis. *J. Sens. Stud.* 1998, 13 (1) pp. 57–75

[31] SEE ALSO MEILGAARD ET AL., 2007, pp. 189-253 (op cit)

#### TDS

[32] PINEAU N., CORDELLE S., IMBERT A., ROGEAUX, M., SCHLICH P., 2003. Dominance temporelle des sensations – Codage et analyse d’un nouveau type de données sensorielles. XXXVèmes Journées de Statistiques, Lyon, 2–6 Juin, 777-780.

[33] PINEAU N., SCHLICH P., CORDELLE S., MATHONNIERE C., ISSANCHOU S., IMBERT A., ROGEAUX M., ETIEVANT P., KOSTER P. Temporal dominance of sensations. Construction of the TDS curves and comparison with time-intensity. *Food Qual. Prefer.* 2009, 20 (6) pp. 450–455



