

第 4 章 Factors Influencing Sensory Verdicts
影響感官評定的因素

4.1 前言

好的感官測定需要將品嚐者(taster)視為測量儀器，但這些人會隨著時間以及彼此之間都有些的變動，也很容易出現**偏見/偏差(bias)**。為使**變異度(variability)**及**偏見**降為最低程度，實驗者就必須瞭解影響感官知覺(sensory perception)之基本的生理與心理因素。

Gregson (1963)指出：現實世界的感知(覺)並非一個被動的過程，而是積極、有選擇性的過程。

一個觀察者只記錄複雜的情況中，他能**容易看到及關聯為有意義的那些元素**。我們必須讓觀察者處於這樣的心境去了解要他(們)去測量的特性。這是透過訓練、並避免在樣品呈現、詢問卷文案及參與者的處置等會暗(內)藏的許多陷阱。

4.2.1 Adaptation 順應

順應：連續暴露於給予的刺激物或類似物，而導致該刺激物的敏感性減弱或改變。於感官試驗，此效應是閾值與強度評級的**變異度(variability)**之主要來源。

下例中，條件 B 的觀察者很可能感知試驗樣品的甜度較低，因為先測試蔗糖會降低甜味的靈敏度。條件 A 採用的水無味，不會造成疲勞(或引起甜味感知上的順應)。

交叉順應(“cross-adaptation”)的範例

Adapting stimulus		Test stimulus
Condition A	H ₂ O	Aspartame
Condition B	Sucrose	Aspartame

4.2 Physiological Factors 生理因素

感官試驗較受人被詬病的是：**心理的誤差影響個體的回應**。

通常，參加者如果是「**naive** 天真爛漫的」，未充分了解試驗方法及/或特別的產品或產品類別(**product category**)，尤須注意這些誤差。明顯地，採用**有經驗和訓練型**的人並依照建議的步驟與練習，可大為降低這些誤差的影響。

然而，仍有些心理的誤差直接關連特定的測試產品，例如 **contrast (對比/反差)**、**convergence (趨同)**效應與其它的 **context (文章的前後關係/語境)**效應，可透過適當的實驗設計的考慮而降至最低，但無法全部移除之。因此，感官專家的主要工作是要瞭解這些誤差和構思每項研究時的重點，然後確認這樣的效應能以相同的機率發生在所有產品。

交叉增強(“cross-potential”)或**促進(facilitation)**也很可能發生，在條件 B 的觀察者感知試驗樣品為更苦，因為先測試蔗糖會提升對奎寧的靈敏度。

交叉增強(“cross-potential”)或促進(facilitation)的範例

Adapting stimulus		Test stimulus
Condition A	H ₂ O	Quinine
Condition B	Sucrose	Quinine

4.2.2 Enhancement or suppression 增強(效)或抑制

增強或抑制涉及刺激物同時混合存在時的交互作用。

- **增強 Enhancement**：一物質的存在而提高第二種物質的知覺強度。
- **協同 Synergy**：一物質的存在而提高兩種物質所知覺的混合強度，混合物的知覺強度大於組成分的強度總和。
- **抑制 Suppression**：一物質的存在而減弱二種或以上物質混合物的知覺強度。

範例：

1. 混合物的總感知強度：

情況	效應名稱
Mix < A + B (each alone)	Mixture suppression (taste)
Mix > A + B (each alone)	Synergy

2. 可分析混合物的組成分：

情況	效應名稱
A' < A	Mixture suppression
A' > A	Enhancement

Mix = 混合物的感知強度

A = 組成分 A 原本的感知強度

A' = 在混合物中，組成分 A 的感知強度

A + B = A 和 B 組合的感知強度

理論上，訓練良好的品評員讓自己不會受到關於樣品的一些偶發知識(accidental knowledge about a sample)而影響；但實際上，這樣的訊息會影響受試者的評級；因此對他或她，更好的方式是忽視樣品的履歷。

[補充] Error of Expectation：

此期望誤差起因於受試者對產品已有一定的認識，由於瞭解(認知)而對特定的屬性或差異性有所期望。

例如在辨識(discriminative)試驗，受試者很可能所報告有差異的濃度是更低於預期，與此相反，受試者卻忽略那些不希望在此類產品中存在的差異。事實上，他可能過於了解樣品，而不應該挑選參加試驗。

4.3 Psychological Factors 心理因素

4.3.1 Expectation Error 期望誤差

提供樣品的訊息都可能引發先入為主的想法(成見)，人們通常去找他們希望找到的東西。在測試時，例如：

- 閾值(threshold)的典型試驗，採用一系列濃度漸昇的(ascending)樣品，受試者(透過自我暗示 autosuggestion)會預期(下一個樣品會出現)知覺，而在實際上確可感知之前，就報告回應。
- 品評員如聽到過期產品被退貨回工廠，當日品評的樣品會傾向要去測出陳味(aged flavor)。
- 啤酒品嚐員如獲知啤酒花的用量，對苦味的評決就易生偏見。

期望誤差會折損試驗的有效性，透過保密樣品的來源、品評員不能給予試驗之外的任何訊息，以及樣品要編碼、呈現順序要平衡、隨機分配給參加者等做法來避免之。

4.3.2 Error of Habituation 慣性誤差

人有習慣性，在感官世界裡也的確存在，因而導致發生一種誤差：慣性誤差。起因：當面臨一系列緩慢增強或減弱的刺激時，連續給予同樣回應之傾向，例如日復一日的品質管制。品評員易於重複給同樣的分數，因而錯失任何形成中的趨勢或甚至接受偶發的缺點樣品。

慣性是普遍的，解決之道：透過變化產品的型式或呈現偽裝/做手脚的樣品(doctored samples)。

[補充] Error of Habituation and of Anticipation：慣性誤差與預期誤差

來自於受試者給予一系列產品或問題而給同樣的回應。

例如在閾值試驗，都以刺激濃度漸增或漸減之少量變化的一系列樣品，在超過確可感知刺激的濃度時，受試者會報告“沒有感知”，即是慣性(habitation)。

預期誤差(anticipation error)：當受試者在實際的閾值濃度達到之前要報告刺激物的感知情形，任何時候都會有所預期，因已回應好幾個“未感知”。某種程度上，這些的誤差等同於期望誤差。

4.3.3 Stimulus Error 刺激誤差

此誤差的起因：當不相關的標準 criteria (譬如容器樣式或顏色)影響了觀察者，如有標準建議是有差異，品評員會去找出即使並不存在的。

例如：

- 品嚐者在當時知道，比起軟木塞瓶裝，旋蓋瓶裝紅酒通常較為便宜，這會造成其所提供的樣品得到較低的評分。
- 緊急召集的品評小組品評會可能觸發已知製造缺陷之報告。
- 在試驗比較後面的時段所提供的樣品會被評級更受喜歡，因品評員知道為使感官疲勞的問題降至最低程度，品評團負責人會首先提供淡味的樣品。

這些誤差的解決對策：

- 避免留下不相關(以及相關的)的線索
- 規則地安排品評小組品評會期
- 經常且不規則地變化任何通常的呈現的順序或方式

4.3.4 Logical Error 邏輯誤差

邏輯誤差發生於當評定者心中將樣品的二種或更多的特性聯想在一起。例如：認知色澤較黑的啤酒會是較濃風味或較深色的沙拉醬傾向於已放置較久，引起觀察者修改其評決，而忽略自身的知覺。

降至最小的影響：保持樣品均一性、藉助有色玻璃或有色光等遮蔽(masking)差異。

有些邏輯誤差無法遮蔽之，但可用其它方式避免。例如：較苦味的啤酒總是傾向於得到較高的啤酒花風味評分，除採用訓練型品評員，品評小組領導為打斷這樣的邏輯關聯，透過偶而改以奎寧的加入來產生帶淡啤酒花香氣的高苦味偽裝樣品(doctoring a sample with quinine in order to produce high bitterness combined with low hop aroma)。

[補充] Stimulus Error 刺激誤差：

試驗中，當受試者對產品或刺激物持有（或認為他們有）先前知識(prior knowledge)，他們會基於自己對該物理刺激物的了解，而非以刺激物實際產生的感知做評分、或者想要發現一些非預期的差異時，就會發生刺激誤差。後者的情況尤難以解決，因為正常的生產循環或非經常用產品的製造方法，在正常的製造週期都會發生非產品變數(none-product variables)。

降低該風險：感官專家必須在試驗之前很徹底地檢視產品，確定有足量的產品和產品是具代表性的。

[補充] Logical Error and Leniency Error 邏輯誤差與容忍誤差：

這些誤差比較可能發生在那些就要做的測試工作但沒有給予明確的指導、且又缺乏經驗的測試者身上。

邏輯誤差：當人員不熟悉特定的工作，而依照自訂的邏輯程序來決定產品如何地做評級。

寬大(仁慈)(容忍性)誤差的產生：是因為測試者只根據他們對實驗人員的感覺來對產品評價而忽視產品的差異。

降低這些誤差至最低程度：給予測試者專門的試驗指導和採用經驗型的人；分析型試驗使用取得資格及訓練的人員。

4.3.5 Halo Effect 光環效應

當同時品評樣品的一項以上屬性，其評級會傾向於互相地影響(halo effect)。各項風味屬性連同整體接受性(overall acceptability)一起評分，會不同於單獨品評每項特性之結果。品評員常對產品作成總體印象(general impression)，故要求同時品評定其氣味、質地、顏色及滋味，和每個因子個別品評之結果就會有所不同。例如：

- 在柳橙汁消費者試驗，要求不僅整體喜歡度，也同時評級特定的屬性。
- 當產品頗受喜歡時，其它全部的項目如甜味、酸度、剛榨汁感、風味強度、口感等，傾向於也是喜歡的評級。反之，若產品不被喜歡，則大多數屬性也被評不受喜歡。

對策：當任何的特定變數都是重要時，呈現至各單獨的樣品組別品評那些特性項目。

[補充] Proximity Error 近似誤差：

此近似誤差，比起不相近/性質差異較大者，相近的特性傾向於被評級較為類似，因此，相近的特性彼此間的相關性顯然較高，比起將它們用其他屬性隔開、再做評級。

降至最小程度的作法：類似的屬性要分開、產品組重複之評級屬性、或將屬性項目隨機分配於評分卷上。很不幸的，在商業環境中的實用性與有用性，這些建議的每一項仍產生問題。

[補充] Error of Central Tendency 趨中誤差：

其特徵為受試者評分產品時都選擇尺度的中間範圍，回避兩端的高/低值（或極大值）。造成的影響：產品似更相同。當人員不熟悉試驗方法或產品、或者對試驗程序有些我行我素的，就容易發生這種評分誤差。更多經驗(及有資格認定)受試者比較不會表現評分誤差。

[補充] Halo Effect 光環效應：

在測試當中，消費者先回應了一系列針對各種產品屬性的問題之後才再回應產品嗜好性。在這種情形下，猶如是由先前的回應設定基調而接下來的回應則是在替這個基調進行辯護，換言之，在最後提問嗜好性將不能改變或消除這種的效應。

在涉及消費者、並特別要回答許多問題的測試，光環效應更常出現。由於要回答許多問題，就需要對產品重複測試，而重複測試會引起生理疲勞，因而使得情況更形惡化並且導致趨中誤差(errors of central tendency)的產生。為深入了解此問題，可以檢查評分卷內與各卷之間的回應樣式與變異情形。

為克服或避開光環效應，有些研究者會要求測試者從分數卡的不同地方開始進行評估或者改變問題提出的順序。實際上，這些努力只是表面功夫，因為效應仍然存在，只是改為影響其它問題的回應而已。

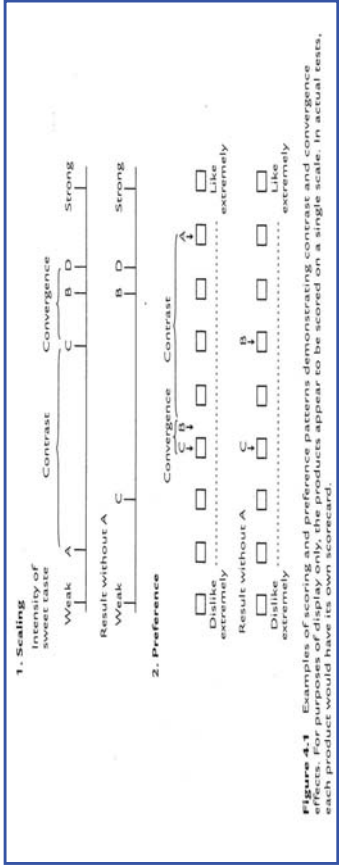
4.3.6 Order of Presentation of Samples 樣品呈現的順序

至少五種類型的偏見因呈現的順序而引起。

- Contrast effect 對比/反差效應：品質較好的樣品出現在品質較差的樣品之前，會造成第二個樣品比單獨被評級時(單一樣品)得到較低的評級。反之緊接在特別差樣品後面的樣品會傾向於較高的評級。

[補充] Contrast (對比/反差誤差) 和 Convergence (趨同誤差) Errors：

可能是最難處理的誤差。這些誤差與產品相關，且常一起發生。
對比誤差的特徵：兩產品被評分差異很大，差異幅度超過預期。當“較差”品質產品出現在“較高”品質產品的後面時可能發生，相差程度被擴大；比起和品質相近的產品一起進行，較高品質者被評得更高分數。
趨同誤差：對比/反差誤差的相反效應，通常是因兩個(或更多個)產品之間的對比而導致屏蔽或掩蓋了其中某個產品與其它產品之間的較小差異所引起的。



短期試驗(sip and evaluate)時，會對呈現的第一個樣品產生偏見/偏差，而長期間試驗(放在家庭一星期)會對呈現的後最樣品產生偏見/偏差。實驗組中第一配對樣品的辨識(Discrimination) 大於後續的配對。

[補充] Time-Order Error 時序誤差：

是一項次序效應(order effect)、第一樣品(first-sample)誤差或位置誤差(position error)。在於評分的誤差，導致第一個樣品（一組樣品中）被給予比預期更高的評分，而忽視該產品本身。如果樣品的呈現順序不平衡，以及結果的討論若不考慮提供順序，此誤差造成結果解釋上的困擾。

- **Group effect 群集效應**：一個好的樣品出現在一群差的樣品之中，比單獨出現時會傾向於被評級較差。對比效應的相反。
- **Error of central tendency 中間傾向誤差**：一組樣品中靠近中間者傾向於比在兩短者較為更被喜歡。在三角試驗法，孤單樣品(odd sample)如放在中間位置有較多的機會被檢出。也發現於採用尺度及類別(categories)時。
- **Pattern effect 樣式效應**：品評員使用所有可用的線索(當然，這在他們的部份是何合法的)，在呈現的順序上很快找出樣式(pattern)。
- **Time error/positional bias 時間誤差/位置偏見**：在一系列的試驗期間，一個人的態度會產生微妙的變化，從預期參與第一個樣品，以至於疲頓或最後已是無所差異。因此，第一個樣品常會是例外地被喜歡或拒絕。

Table 4-1 Examples of scoring patterns from a series of two-product tests displayed by serving order as a demonstration of order effects

Example	Serving order	Product	
		X	Y
A	First	6.4 ^a	6.3
	Second	6.1	6.0
	Pooled	6.3	6.1
B	First	6.5	6.4
	Second	6.7	6.1
	Pooled	6.6	6.2
C	First	6.3	6.5
	Second	6.2	5.9
	Pooled	6.3	6.2
D	First	6.4	6.5
	Second	7.0	6.5
	Pooled	6.7	6.5

^a Entries are mean values based on the use of the nine-point hedonic scale (1, dislike extremely; 9, like extremely), N = 50 in each case except for pooled values, with a monadic-sequential serving order. See text for further explanation.

Example	Serving order	Product
A		X Y
	First	6.4 ^a 6.3
	Second	6.1 6.0
	Pooled	6.3 6.1

範例 A：顯示第一個樣品(first-sample)呈現之典型效應，兩樣品都被評分比平均值高出 0.3 單位，其結論為兩樣品同樣程度被喜歡，時序誤差同等地影響樣品兩者。

Example	Serving order	Product
B		X Y
	First	6.5 6.4
	Second	6.7 6.1
	Pooled	6.6 6.2

範例 B：更複雜的可能性，先評樣品 Y 則樣品 X 的喜歡程度更高；先評 X 之分數 6.5，位居第二品評則提高為 6.7；反之對樣品 Y 亦然，先評 Y 的分數為 6.4，第二品評則降至 6.1。純粹看結果會建議 X 比 Y 更被喜歡，可能 6.6 對 6.2 pooled result 達到 95%信賴區間的統計顯著性。然而，也出現爭論即都是第二品評時兩樣品的差異 0.6，如自己的樣品是 Y，但會有興趣於樣品 X 吸引消費者，這些結果讓我們了解若從產品的角度這不是很成功的。

Example	Serving order	Product
C		X Y
	First	6.3 6.5
	Second	6.2 5.9
	Pooled	6.3 6.2

範例 C：另一種可能結果，如同範例 A 所發生的一兩樣品(居第二品評)都呈現降低，但是 Y 樣品的變化幅度較大，暗示樣品 Y 可能有缺點或 X 可能有機會。

Example	Serving order	Product
D		X Y
	First	6.4 6.5
	Second	7.0 6.5
	Pooled	6.7 6.5

範例 D：另一可能的結果，X 增高，而 Y 幾乎未變化。居第二位品評時會增加喜歡 X 甚於 Y (7.0 相對於 6.5)的一項可能解釋，Y 樣品具有只在第一位（然後是 X）品評時才明顯易見之官能的缺點。

所有這些的影響必須透過採用平衡且隨機順序的呈現(the use of a balanced, randomized order of presentation)而降為最低。

平衡的(“**Balanced**”)意指每種可能的組合都呈現相同的次數，每次的品評小組品評會每個樣品在第一、第二及第三…及第 n 位置都要出現同樣的次數，如果要呈現樣品的數量大，可採用平衡不完全區集設計(a balanced incomplete block design)。

隨機的(“**Randomized**”)意指根據機會法則(the laws of chance)所選擇組合出現之順序。

實際上，隨機化(randomization)是透過抽取袋裡的樣品卡來達成，或者計畫藉助於亂數碼的編集。

4.3.9 Capriciousness (任性) 對比 Timidity (膽怯)

有些人傾向使用尺度的兩端，導致影響結果的比重過大。另有人傾向挑尺度的中間，導致樣品間的差異變小。

要取得有再現、有意義的結果，品評團領導必須每日監控新成員的評分，需要時採用偽裝樣品 doctored samples 作為例證。

4.4 Poor Physical Conditions 身體狀況差

品評員(1)如患有感冒(做品嘗員時)、皮膚病症、神經系統障礙(觸覺型品評小組)；(2)如有牙齒衛生不佳、齒肉炎；(3)情緒低落、過度工作壓力而使他們難以集中；(4)如目前有服用藥物而會干擾對滋味或嗅覺之能力。

4.3.7 Mutual Suggestion 相互建議

一品評員的回應會受到其他品評員所影響，由於這，品評員分隔於品評小間，避免判斷從反應至臉部的表情而引起他人的揣度，品評時不允許任何出聲，試驗區域要無外界的聲音干擾，與製備區隔離。

4.3.8 Lack of Motivation 動機欠缺

品評員的努力程度會做出辨識微小的變化，或為所給予的印象找出適切的話詞，或一致性的給分，這些對結果都有決定性的重要性。

品評團領導的責任是要創造讓大家感覺舒適及做有意義工作之氛圍。

保持興致的品評員總是效率高。激勵最好是在一個善解、明確定義的試驗情況。

維持品評員的興趣可透過試驗完後給予他們的結果報告，讓品評員覺得品評團是項重要的活動，這可透過控管的、有效的方式來運作試驗而巧妙達成。

再者，老化(aging)會影響感官靈敏度，老年人對各種滋味及嗅覺的靈敏度較低，吸菸者可能是好品嚐者，但避免品評之前 30 至 60 分鐘抽菸。濃咖啡會影響感知達一小時以上，測試應該不要在飯後 2 小時喝咖啡，因此，品評工作的適合時段介於早上 10 時至中餐，通常，個別品評員的最佳時間取決於他或她的生理時鐘：當一個人最清醒與一個人的精神力量處高峰之這一天的時候。

Mattes (1986; J. Sensory Stud., 1 (3-4): 225, 1986)綜述健康或營養失調有各種方式影響感官功能。滋味與嗅覺功能障礙與慢性疾病關聯，對某些滋味受器靈敏度的順應可能誘發與健康問題有關的情況譬如肥胖、糖尿病及其他等，因此，身體狀況對感官品評是非常重要的，身體狀況差的品評員不可能得到正確的試驗結果。