

(一)總加量表

1、形式與沿革

總加量表係由一組問題結合而成，每一個問題之下均附有一個標尺，來測量受訪者對這個問題贊成或反對的程度(範例6-2)。

這種量表由Likert [234]所創設，故又稱「李克量表」，每個項目的標尺最早由五個刻度(Steps或Points)構成，但日後無論刻度或格式均有不同的變化。總加量表發展迄今，已經成爲行爲學界最常利用的測量工具，不過，Oskamp也發現它是最容易「使用」的量表，也是最容易「誤用」的量表。

範例6-1 總加量表的各種格式

問：我們應該加入聯合國，不問任何理由？

答：_____ 非常贊成

1.

_____ 贊成

_____ 沒意見

_____ 反對

_____ 非常反對

非常反對 _____ : _____ : _____ : _____ : _____ 非常贊成

2.

成

非常反對 _____ : _____ : _____ : _____ : _____ 非常贊成

3.

非常反對 非常贊成

4.

常贊成

非常贊成

5.

贊成

沒意見

反對

非常反對

*第一種是Likert原來設計的量尺，第二至第四種是日後的變形；第五種是本資訊系統提供的視窗標準單選式介面。

2、測量理論

假設受訪者對量表中任一項目所抱的態度，可以根據由贊成到反對的程度，配合一個由高而低的分數，那麼把所有項目的分數加起來，就應該是他對整個問題的態度分數，此即Shaw[310]又將這種量表名爲「總加」的原因。

Likert創設這個量表之初，只是爲了實用上的目的，並未深入闡述其理論上的意義。

本研究者建議，以下兩個簡易的公式，可以充分表達其理論上的概念：

$$f (X \text{ 態度的強弱}) \rightarrow R (\text{實數})$$

即每一種態度（X）都可以找到一個實數（R）來對應，但為了簡便計，通常是以整數來對應。

$$\sum X_i (\text{每一個項目的分數}) = X_t (\text{態度分數})$$

即各項目分數相加，是為受訪者的態度分數。

根據上述測量理論，編製一套總加量表，應具備以下必要條件：

a. 應有相當數量的項目。譬如我們發現一個人喜歡吃甘蔗、蜂蜜、巧克力、冰糖蓮子、千層糕，應該可以推斷他偏好甜食；如果我們只知道他喜歡吃甘蔗，便很難作相同的推測，因為他很可能只是愛好經得起咀嚼的食品，並不特別偏好甜食。項目如果太少，誤差就會很大，方法學上特稱為「信度」（reliability）低。

b. 項目必需對所要探討的問題具備「區別力」（discriminatory power）。Oskamp[263]曾舉例說，一組意欲區別「國際主義者或是孤立主義者」的量表，如果包含一個項目：「追求和平是人類主要的目標」，無論國際主義者或是孤立主義者都很可能給予很高的評價，這種項目就是低區別力的項目，也是不好的項目。

c. 每一位受訪者對每一個項目的評分，其間需有「相關關係」（correlation），尤其每一個項目分數應與總分相關，也就是說，甲項目的分數偏高，乙、丙、丁...及總分均應比例偏高，反之亦然。

再舉上述「甜食」為例，假設一位食品工廠經理，打算了解甜食市場，於是他設計了一個量表，包含九個項目：「甘蔗、蜂蜜、巧克力、冰糖蓮子、千層糕、魷魚絲、牛肉乾、鹽酥蠶豆、乾饅饅」，偏好前面五種的，給高分，偏好後面四種的，給低分，他認為：總分愈高的人愈多，就是甜食市場愈大，反之則愈小。結果資料回收後，他卻發現，有時甘蔗的分數高，果然前面五項分數及總分也高；但有時甘蔗分數很高，前面其他四項分數都低，造成總分不高不低、甚至偏低的情形。

產生這種現象，因為甘蔗不僅具有「甜食」的屬性，也具有「利咀嚼」的屬性，前一個屬性和千層糕前面的四個項目結成一個集團，後一個屬性又和魷魚絲後面的四個項目結成一個集團，這兩個集團強大的力量，使得甘蔗忽而與前面的項目相關，忽而與後面的項目相關，忽而與總分相關，忽而不相關、甚至負相關，使總分的高低跟著不隱定，無法根據總分作恰當的決策。

這種情形，在方法學上稱為這組量表具「多向性」（multidimensional）根據國內外研究經驗顯示，總加量表經常具有多向性，而且往往不如上述例子那麼明顯清楚，因此研究者必

需把量表中的各個「向面」(dimensions)分析出來，才能尋求合理的解釋與推斷。爲了符合總加量表的測量理論，以及滿足上述三個必要的條件，編製總加量表必須經過「淨化測量」(purify measurement)的手續。

3、編製程序

a.設計題庫 (item pool)

由於在測量之前，我們無法知道那些項目的區別力如何，所以應該廣泛收集各種有關的、可能的項目，Nunnally認爲通常在二十條至八十條之間組成題庫，再從中挑選出恰當的項目。

項目的「區別力」不是「重要性」。本文作者在1983年在臺灣推動第一個政治態度調查與選舉預測時，鑑於毫無前例，所以在事前請教了許多的先進。其中包括一位在學術與政界兩方面，都有崇高聲譽的知名前輩。他非常認真的看過我擬的題庫後，極嚴肅又善意的對我說：「統雄，你竟然犯了一個大疏漏，沒有問選民『清廉』重不重要！」

『清廉』是一個不證自明的絕對價值，沒有人會認爲政治人物可以不『清廉』，這種項目就缺乏區別力與測量意義。量表設計是一種科學，不是設定一種標準來考試。

b.淨化測量-信度分析

抽選一個小樣本，利用題庫做一次「前測」(pilot study)借以鑑別項目的好壞。淨化總加量表常用的方法是「項目分析」(Item Analysis)傳統的項目分析，常經過計算「區別力」等三道手續[39][171]。但在電腦系統的協助之下，則宜採用Nunnally[258]建議的「各項目與總分相關係數分析法」(r_{it} method)分析，取捨的效果可能更佳，這個方法其實就是量表的「[內在一致信度分析](#)」。

吳聰賢[39]曾以對「家庭計畫的態度」爲例，說明傳統項目分析的步驟；本研究者則以相同資料，用「 [\$r_{it}\$ 相關係數法](#)」分析，相互比較，比較中可以發現：

(a)兩種分析方法，最後取捨的項目大致差不多。

(b)傳統的方法，程序比較繁雜。(當然，如果沒有電腦協助，計算相關係數反而比較麻煩)。

(c)傳統的方法有幾個步驟均訴諸主觀，但相關係數法則可借客觀的統計標準，作更合理的取捨。

c、決定量表的項目

量表最後實際採用的項目，數目如果太少，很容易發生歪曲，項目若是太多了，又很容易令受訪者厭倦。根據統計推算及國外研究經驗，項目如果超過五十，可信性已近飽和，增加可信的程度會漸趨緩慢，所以一般量表的項目上限可以定爲五十題。

有些學者建議用反復調查的方式，逐次減少一些項目，等到項目最少，而仍能維持最低可信程度時，就以這個數目為準。不過，這種方法除了對有強大經費、長期性的研究之外，顯得相當複雜。

Nunnally[258]以心理度量的立場認為項目的下限宜訂為三十條。

本研究者依據研究經驗，建議兩個原則：

(a) 對於抽象程度很高的問題、完全不了解的問題、精密的心理測驗，或是學業成就測驗，項目不宜少於三十。

(b) 對於很具體的問題、已有其他相關研究資料的問題、或是其他研究者掌握充分證據的問題，項目數量可以酌量減少。最少似可至五條，但如果再少，就可能失去「總加」意義的量表性質，而是直接詢問法的問題了。

4、刻度

總加量表中每一個項目都附著一個標尺，再依受訪者「贊成或反對」的程度分成若干個「刻度」(Steps或Points)，訪問後每一個刻度都給予一個分數以便後續分析。

在以往文獻中，討論刻度的「度數」方面，包括三個--且都偏向形式上的問題：

--最適刻度的範圍大約是多少？

--度數是奇數還是偶數好？

--要不要「沒意見」或是「不知道」的「中間分數」(middle point)？

但本文作者要特別提出2個文化與方法論的問題：

--刻度的「相對客觀性」與「文化適域性」為何？處理方法為何？

--刻度反映的到底是「點」測量，還是「區間」測量？

(1) 形式上的問題

Likert最早使用總加量表時採用5個度數，日後發展由3刻度到100刻度的都有。

從區別受訪者而言，刻度當然是愈多愈好，受訪者之間的差異才容易反映出來，研究者才能有所決策的依據。

但是刻度太多時，也容易發生兩種爭論：

一個受訪者假設面對一個一百刻度的標尺時，他如何分辨「96的贊成」和「97的贊成」有何不同？其間1%的差別何在？

其次，兩位受訪者，一位可能喜歡以強烈的意見反映自己強烈的態度；另一位卻可能習慣以溫和的意見反映自己強烈的態度。那麼，這兩者在刻度很多的標尺上可能相去很遠，但其間的差距實際上並不存在。

而刻度的後兩個問題，其實二而為一，因為奇數的標尺就是多了一項「中間分數」的緣故。

學者研究後發現，含「中間分數」的標尺，容易產生以下四個問題：

(a) Bogart [107], Presser [277]認為中間分數常成為「遁詞」。許多受訪者並不是沒有意見，只是不肯多思索，就隨便以中間分數為答案。

(b) Converse [134], Faulkenberry et al. [152]發現「沒意見」(nonopinion)和「不知道」(no opinion)本質上並不相同，但均列為「中間分數」(在英文量表中，都用“don't know”一詞代表)容易引起混淆。「沒意見」是一種中立態度(ambivalent attitudes)，應該計分；「不知道」是對研究者的問題不知所云(nonexistent attitudes)應在資料中剔除，不予計分。如果沒有分別處理，統計量就容易偏頗，使推論不正確。而在一般奇數刻度的標尺中，並無法區別這兩者的不同。

(c) 喜歡選中間分數的人，常具有某些特徵，容易造成統計上系統性偏差。譬如在美國，Ferber [154], Francis et al. [162]發現社會階層較低(Peripheral positions)的人，如：女性、有色人種、低教育程度、低收入、參與政治程度低的人，喜歡用「中間分數」搪塞，其實並不表示那是他的態度。

(d) Ryan[298]指出當資料中「中間分數」很多時，容易帶給統計分析困擾，也容易發生統計上的爭議。

從一般文獻中看出，國外使用的刻度常在5個左右。

(2) 文化上的問題

但根據國內研究經驗，本研究者建議以使用百分刻度與百分評分法頗為理想，其優點是基於以下的觀察、分析與討論：

1. 相對客觀性與文化適域性

Likert測量的是一種「虛擬量」(當時的術語稱為「模糊量」-vague quantifier)，不具物理上的客觀性；但如果高比例的人，認同其數字的高低與實質的高低，觀念上相當接近，就可視為具備某種程度的「相對客觀性」。而這種相對客觀性，可能並非是放諸四海而皆準的，

「文化適域性」係指一種觀念或技術--在此即蔚為測量工具--在某地區適用，但移往其他地區卻可能產生不同效果，所以必須根據每一個地區的特色作「適域性」調整。

Likert最早使用5個度數，可能就是歐美採用5刻度作人類行為測量(尤其在學習成就評量上)早已有長遠的歷史與傳統，大家都很熟悉，自然而然(自然就是文化)的接受這種形式。

本研究者在1983年推動臺灣第一次的選舉調查，當時還是所謂「威權時代」，對公共事務--尤其選舉--作調查，還存在著心理上的禁忌，所以絕大多數的人民之前不曾接觸過調查，甚至不知道什麼是調查。當我們在作預訪籌備階段，發現如果採用美國常用的5點量表，受訪者在生活經驗中過去並沒有5點制評分法的經驗，很難回答。即使回答，我們分析其中的一致性也很低。

本文作者觀察到這個現象，認為不能硬生生移植這種量表，必須考慮尋求臺灣文化中具備相對客觀性的測量方法。

而我國從小學開始，就實施百分評鑑法，而不同於國外的5等評鑑法。因此國人對百分法的內涵意義均十分清楚而敏銳：60就是普通（及格），80就是很好（或很贊成），90以上就是非常好了，60以下則是不好（或不贊成），50以下很不好，30以下更是非常不好了。

在這種文化情境下，較容易使受訪者態度與評分刻度呈現自然、與相互呼應的結果。

正式實施之後，果然也突破了受訪者不了解的障礙，順利的在臺灣開創了公共調查與選舉預測。

2.容易評分與回答

傳統的評分法，在設計格式時係以視覺為主，但在電話訪問中，如果反復詢問「您非常不贊成？很不贊成？不贊成 … … 非常贊成？」，問不了三、四題，雙方頭都昏了。因此，宜改用適合聽覺的百分評分。

3.可以區分「沒意見」和「不知道」

如果研究者想觀察真正的「中立態度」時，就是60至69分區間的受訪者。

另外，在答案中增加「不知道」的選項，反映真正「不知道」的受訪者。

在訪問過程中，形式上既然沒有中間分數，就不會發生以中間分數搪塞的系統性偏差，使不想動腦筋的人，無隙可遁，不致使中間分數太多造成統計困擾的情形。

(3) 新挑戰：「點」測量與「區間」測量

A. 5點制的盛行

在1983年的選舉調查之後，民意調查如雨後春筍般盛行。許多機構引入各式各樣的調查，但似乎多沒有注意到「相對客觀性」與「文化適域性」的問題，而直接採用5點量表。近30年來，多數中年以下的人民對5點評分制應該已經不陌生，5點制亦應具備了相當程度的相對客觀性。

B. 「點」測量與「區間」測量

「虛擬量」的特性，就是測量值的差異不一定是真實差異。百分制有一個優點，就是可以作「事後歸類」，把80至100視為1類、60-80視為1類…等，不要再在類別中分高下，避免假性差異。也就是對受訪者而言，是作「點測量」，而最後可得到「整體」的「區間測量」結果，可能較符合人類行為的特質。

而5點制在執行上，等於是讓受訪者在事前就作歸類了，其實有相當微妙的不同。

但臺灣目前是百分制和5點制並行，本文作者在臺、美教學期間，就發現兩地似乎對「5點」代表的意義有所不同：（美國的5點，在許多地方化為「分」時，是0-4分，但為比較之方便，本文均依臺灣習慣，作為1-5「分」來討論。）

5點制有沒有滿分？是「點測量」的100分？還是區間測量的80分？

5點的「及格」是「4」？還是「3」？

5點的「1」是「0分」嗎？

美國人民在文化上，似不常打最低分，臺灣呢？

5點制用「1~5」的形式，或用「-2，-1，0，1，2」的形式，在資料收集的結果會不會有所不同？

本文作者在2011年所指導的小型研究，發現臺灣人民對5點制已否具備基礎「[相對客觀性](#)」，其「區間性」的客觀程度與百分制相當接近；但在重要切點：如及格、優秀點上，仍不如百分制明確。本文作者建議：若有恰當的資源，應作更廣泛與深入的研究，

5、優點與限制

總加量表體製簡明，容易設計、容易作答，諒必仍為將來民意調查的主要工具之一。

然而，總加量表仍有其限制。

(1) 過去文獻發現的問題

整理過去文獻，可歸納為兩項主要的限制：

a.編製總加量表，必須淨化測量（譬如事前作項目分析）。沒有淨化的總加量表，是蕪雜而很可能沒有意義的。

b.總加量表假設人類的態度是可以用「等距尺度」(interval scales)畫分開來的。人類的態度可以「等距」畫分，還是只能「等序」(ordinal)畫分？仍然是方法學者如Campbell [116], Stevens [332]等認為是見仁見智的問題。

(2) 從實務經驗與基礎理論發掘的問題

本研究者從事人類行為測量研究數十年後，發現總加量表可能還有2項深入的問題：

a. 總加量表對許多主題有先天效度不足-尤其是吳統雄所發現的「[取用效度](#)」-的現象。

b. 總加量表在許多主題上，只是儀式作用，有宗教效果，可能沒有科學效果。

以上2個問題，其實只是一體的兩面。

測量工具的「[信度與效度](#)」，在行為研究上過去主要強調的是「[構念效度](#)」，亦即某種抽象的特質，如：重力、健康、心理健康、學習滿意度…，是否存在，是否能夠被測量到。這種效度可以用MTMM法與相關方法來衡量。

但在許多[人類取用行為](#)的情況中，並不是「[構念](#)」是否存在的問題，而是人類取用行為作用，會扭曲測量的過程，使總加量表這種工具，測量不到目的變項。

吳統雄發現人類「[TX取用模式](#)」是反應行為函數，和社會氛圍作用函數，交互作用形成的取用行為「類 \tanh 積函數」。

簡單的說：人類的認知反應是依據對認知對象的：興趣、理性抉擇、社會相信的交互作用，而最經常反應的態度，是依據社會相信而作集體行為反應，而不是對事實作反應。

先舉一個簡單的例子，譬如以直接詢問法測量「在網路使用色情資訊」，歷年來都不到百分之1。但以流量流向法測量，許多數據都指出在百分之50以上。就是「直接詢問法」這項工具，在測量「在網路使用色情資訊」這類問題，是效度是很低的。而且，這不是工具本身的問題，而是「人類行為」的問題，當受訪者預期「社會相信」認為「在網路使用色情資訊」是一件不好的事，他不會回答事實，而是依據集體行為作否認。

在這種情境之下，不能改善工具，只能換工具-如換成流量法。總加量表的先天限制，也完全相同。

再舉一個具有國際管理意義的例子：富士康連續跳樓事件與其因應。富士康禮聘了多位頂級的管理大師，協助防患於未然，其中一項，是召募員工時，[新增心理測驗](#)。

這項測驗是總加量表，多達70題，信度的基本條件應該達到了。

項目包括：

- 如果周圍有喧嚷聲，不能馬上睡？
- 常常思考將來的事情並感到不安
- 一整天孤獨一人時常常心煩意亂
- 即使自己發生了重大的事情，也不會喪失理智
- ……等。

富士康希望此後招進來的員工具備「心理健康」，不要再跳樓。「心理健康」這項「構念」應該是存在的。

量表中的項目「不能睡」「不安」「心煩意亂」「喪失理智」…也應該可以測量「心理健康」。

但是報導上說「應徵民眾大排長龍，排出一公里長的隊伍」通過心理測驗「最後會進入復試的人員，只有求職民眾的百分之一」。

「[TX取用模式](#)」指出：人類有極大需求時，是不計代價要取得的，應徵者在這樣的低機會、高需求下，一定努力按照「富士康社會裡的相信」去回答，事實是次要、甚至是不需要的。自己即使是容易「不能睡」「不安」…，也會回答「沒有」。

這不是量表設計的問題，是人類行為基礎知識的問題。

富士康所聘請的頂級管理大師們，雖然按照常見教科書，編製了很長的量表、增加了很多工作量、也花費了很高的額外成本。但在行禮如儀外，是否能夠完全測量出「心理健康」的員工？很值得持續觀察。

再舉一個教育界中的常見例子：「學生意見反映」-這是本研究者定的名稱。在當前教育制度下，這個活動稱為「教學評鑑」。

「教學評鑑」要測量的是「教學品質」，應該是一個存在的構念。當前制式的量表，第一題常是「老師學識淵博」；最後一題常是「老師從不遲到早退」。

這些項目看起來，好像也能測量「教學品質」，但也發生了與人類取用行為基礎的扞格。

老師學識淵博？

[Galileo](#) 學識淵博嗎？[Pearson](#) 學識淵博嗎？

他們當代頂尖的學術大師：紅衣大主教、皇家院士都不知道，如何要求一位大學生知道？

老師從不遲到早退？

政治大學有一位教學卓越教師，獲邀到各校演講，如何提高「教學評鑑」分數的方法。

他非常誠摯的提出了一項秘訣：他和同學約定，他班上的上課鐘，比照學校的上課鐘晚10分鐘，學生都非常開心。

如果學生都按照事實回答，應該每人都回答「老師一直遲到早退」，那他也當選不了教學卓越教師了。

但學生都非常開心，所以每個人都作了不實的回答：老師從不遲到早退。

這個量表想要測的是「教學品質」，但能夠測到的分量是很少的。「教學品質」就是不易、也不能用總加量表來測量。

這種類型的總加量表，測量到的是什麼呢？其實是「人類的取用傾向」，換較通俗的名稱，就是「使用者滿意度」，在學校就是「學生意見反映」。

「使用者滿意度」量表，不論其中項目怎麼換來換去，核心自變項還是只有2個：「效益（高）」與「成本（低）」。

對「學生意見反映」就是：學生所感受到的實用性（高）、上課趣味性（高）、與課業壓力（低）。

總加量表由於人類基礎行為的影響，使其與某些欲測的構念，存在先天無法測量的問題，即使獲得了測量數據，也有可能反映的是另一種構念。

(3) 總加量表不是一定必要

本研究者自1980年代起，在投票行為的系列研究中，便發現了以下現象：

過去的研究，對候選人的評價都是採用由若干個項目組成的總加量表，請受訪者就項目一一評價。

胡佛與其同儕使用的項目包括：候選人的表現或成就、家庭背景、學識、品行、風度、能力、勇氣、經歷的特殊遭遇、聲望等。

陳義彥與其同儕使用的項目包括：候選人的口才、作人熱心、品德、學問、敢說話、保障人權、對整個社會的貢獻等。

如果受訪者對候選人有明確的認識，從研究方法來說，這樣的測量結果當然較好。

但是，本研究者在研究經驗中發現，絕大多數的受訪者對絕大多數的候選人，均無親身接觸、更無明確認識。如果一位受訪者對一項明確項目作了評價，再進一步請他以具體事例來印證他的評價，他極大可能說不出來。他可能是把他綜合的感受，再分割到他其實不見得清楚的項目上。

也就是說，受訪者很可能是在對明確的項目，作不明確的評價。雖然這些項目總加起來，很可能又接近原來綜合的評價。

受訪人在評價候選人時，並非在評價他「認識」的候選人，而是通過大眾媒介或人際傳說等「社會相信」作用，所捏塑出來的候選人「社會形象」。再經由受訪者的個人「感受」，評價候選人的「社會形象」。

故本研究嘗試採用受訪者對候選人綜合的感受，以符合事實上受訪者不易分割的「形象評價」，作為預測的指標，這也是本研究與過去其他相關研究不同的理論與方法之一。

選民這種不明確的選擇，或可從我國固有文化智慧中「彌子瑕與衛靈公」的故事來說明。

衛靈公寵愛貌美的彌子瑕。有一天，彌子瑕的母親生病了，便盜用了皇車趕回家看母親。

有人趕緊道衛靈公面前打小報告。不料，衛靈公聽了之後，欣然感慨道：「彌子瑕真是孝順啊！爲了探望母親，不惜觸犯殺頭王法。」

又有一次，衛靈公和彌子瑕在花園手牽手散步，彌子瑕隨手摘了一粒桃子，咬了一口，覺得十分甜美，就順手把剩下的部分塞進了衛靈公的嘴裡。衛靈公喜孜孜的說：「彌子瑕真是有愛心啊！有好吃的東西就會與我分享。」

後來彌子瑕年老色衰，衛靈公便叫人把彌子瑕推出去斬了。當然，也免不了有人來說情。衛靈公聽了卻勃然大怒，罵道：「盜用皇車，依法當斬；又把沾了口水的桃子給我吃，更是對我的汗辱！」

彌子瑕就是候選人。美貌就是他的形象；而他是善敬孝道或是大逆不道，是因存關愛還是心存汗辱，則是他的本質。

衛靈公就是選民。當他熱戀彌子瑕形象的時候，彌子瑕的一切本質都是好的；但他不再被彌子瑕的形象吸引的時候，彌子瑕的所有本質突然都是壞的。

而那些來打小報告的、來說情的，或者只是來忠諫的，則是各種媒體。在順著衛靈公意思的時候，或許會發生推波助瀾的效果；但如果和衛靈公意見不一致的時候，也起不了什麼作用。

衛靈公根本沒有愛過彌子瑕這個「人」，他愛的--也就是「TX取用模式」中的「需求反應」--是美貌。

選民根本不認識候選人，他選的是自己心中追求的形象。

本研究發現，極少數選民能夠說出候選人的具體政見，他所能舉出的優點往往是泛泛的條件--譬如「學歷好」，但是客觀條件上對手的學歷並不差，甚至更好。只是他愛上候選人的時候，候選人的一切條件都是優點；而對手一切相同的優點，通通不見了。

量表中的許多項目，並非沒有區別力，但經由人類的取用行爲，紛紛變質了。

譬如，在測量馬英九和陳水扁的差異時，會出現馬英九的鄉土親和性比陳水扁高，陳水扁的國際化經驗比馬英九好的資料。與其這樣分項測量總加，還不如直接打個形象分數，因爲選民的選擇本來就是很形象的。

這種情形不僅發生在大眾行爲，對菁英行爲也是一樣。

本研究曾經參加許多不同種類的評審活動，評審表也是一種總加量表，分成許多項目。但相當比例的評審，其實並沒有分項評審，還是打個總分，或是再將總分平均分派回各項目去。

亦即，還是有相當比例的評審，真正的需求只是領個出席費，並沒有對評審的標的極具興趣，所以就依據社會相信、打個社會形象分數，而且，會形成一種集體行爲。

總加量表固然流行，它真正的必要性，其所反映出來的訊息，應如何解釋，仍值得研究者縝密地思考。