

行政院國家科學委員會專題研究計畫成果報告

員工績效評估之研究：從顧客評估之觀點

The Employees' Performance Appraisal:
From the Customer Perspective

計畫編號：NSC 88-2416-H-034-001

執行期限：87年8月1日至88年7月31日

主持人：謝安田 執行機構及單位名稱：文化大學企管系

一、中文摘要

既有文獻中，不論以實驗或實地之調查，或者從直屬主管、同事、下屬、自我等不同績效評估者切入之研究，所探討之員工工作行為表現，其範圍皆在管理者可觀察之領域內。然而，在某些服務業中，有些員工非常重要之工作行為表現是不易被觀察到，針對此一績效評估問題，除文獻上尚失之闕如並未釐清外，在相關產業上亦具實務研究價值。因此，從認識論觀點，本研究即在嘗試找尋適合此特性員工之績效評估者，以補足績效評估理論尚有不足之處，同時解決實務問題。

本研究在設計上嘗試結合定性與定量方法，研究首先以電話進行實務訪問，樣本為60家綜合旅行社主管，訪談之目的主要在於瞭解績效評估實行現況，以進一步釐清研究問題。其次，透過特殊事件法(Critical Incident Technique, CIT)經由顧客與員工觀點，歸納出主要25個服務因素及11個比較準則；最後，運用層級分析法(Alytic Hierarchy Process, AHP)在11個比較準則下，進行各種不同績效評估者成對比較，經由不同層級權重運算後，研究結果發現：針對服務業員工不易被觀察到之重要工作行為表現，以顧客為績效評估者較傳統評估者更為正確與適合。

實證資料以旅行業為範圍，樣本則主要涵蓋下列兩部份：其一為定性方法部份，共包含302位顧客及84位員工樣本；其二為定量方法部份，共包含產、官(含消保團體)及學界專家樣本31位。

關鍵詞：績效評估、顧客評估、旅行社、團體旅遊、專題計畫、報告格式、國科會

Abstract

Traditional rating sources can not provide an unbiased solution to the problem of evaluating employees' performance for which the job performance behaviors cannot be observed by the management. From the epistemology perspective, this study is attempting to find an appropriate rating source for such employees' job performance behaviors.

Both qualitative and quantitative methods were utilized. The study first incorporated critical incident technique (CIT) method and categorized the observations of both the customers and employees. From this exercise then, twenty-five service features and eleven criteria were emerged. On the basis of these results, the analytical hierarchy process (AHP) method was then employed to conduct a pairwise comparison among alternative rating sources.

The results showed that using the customer as the rating source is more objective than the conventional sources in appraising the performance of the unseen employees. The empirical study focused on a service industry using a sample of 302 customers and 84 employees for the CIT, and 31 experts for AHP.

Keywords: Performance Appraisal, Customer Appraisal, Travel Agency, Group Package Tour, Research Project, Report Style, National Science Council

二、緣由與目的

依據 Murphy and Cleveland (1995) 指出，績效評估(performance appraisal)要能真正落實並進一步為管理者與員工所採信，

其前提在於績效評估是否具正確性。故在文獻回顧中與績效評估相關之研究，早期學者，如：Campbell and Fiske (1959) 及 Smith and Kendall (1963) 等，其研究之重點在於評估尺度及方法之發展，並深入探討如何降低評估者在評估時之偏見與繆誤；晚近學者，如：Harris and Schaubroeck (1988)、Bernardin, Dahmus, and Redmon (1993) 及 Bozeman (1997) 等，則著重於更正確之評估技巧發展、強調成本效益觀念及對不同評估來源者之差異進行探討。

此外，亦有學者 Ilgen, Barnes-Farrell, and McKellin (1993) 認為 Landy and Farr 於 1980 年在 *Psychological Bulletin* 所發表之文章，似乎可視為績效評估研究中一個重要與明顯之分野：兩位作者將績效評估研究之方向從以往著重於評估錯誤(rating errors)、評估者訓練(rater training)及評估尺度發展(rating scale development)，導引入一新的研究方向-瞭解評估過程(rating process)。

而有關績效評估之功能，早期在於為員工薪資調整、職務升遷時之測量工具；近年來，其功能則漸漸轉化為顯示員工訓練是否需要、檢視個人與組織目標是否達成及為組織內溝通、發展之有效工具(Murphy and Cleveland, 1995)。

在績效評估中有關績效評估者之重要性事實在科學知識之本體論(ontology)上已獲得確認(Landy and Farr, 1980; Murphy and Cleveland, 1995)，而就認識論(epistemology)言之，不同之學者針對不同的績效評估者議題進行探討：有的學者從單一績效評估者切入，探討其信度、效度及可能產生之各種偏誤(biases)，如：Farh and Werbel (1986)、Farh, Werbel, and Bedeian (1988)、Fox, Ben-Nahum, and Yinon (1989) 及 Fox and Dinur (1988) 等；有些學者則從不同績效評估者間(inter-rater)進行深入研究，如：Vance, MacCallum, Coovert, and Hedge (1988) 及 Schrader and Steiner (1996) 或者利用 meta-analysis 方法嘗試瞭解不同研究間之相關性，學者

如：Harris and Schaubroeck (1988) 及 Mabe and West (1982) 等；晚近學者則提出了所謂之 360° 績效評估，試圖從更周延、更宏觀角度(facets)來進一步擴充對此領域知識之瞭解，如：Pollack and Pollack (1996)、Bozeman (1997) 及 Borman (1974) 等。

從上述之文獻進一步整理與分析可發現，在績效評估者相關研究中，研究所探討之主軸主要環繞在直屬主管(immediate supervisor)、自我(self)、同事(peer)及下屬(subordinate)等績效評估者。學者 Campbell (1982) 及 Bozeman (1997) 分別指出：如果從所有績效評估者所蒐集之資訊皆一致，則管理者沒有必要從多重績效評估者(multiple sources)獲取績效評估相關訊息。顯然，每一個績效評估者針對被評估人所觀察到的現象有所不同，而由不同績效評估者所獲得之資訊對管理者而言亦有不同之用途，如：Lawler (1967)、Blood (1974)、Murphy and Cleveland (1995) 及 Cleveland, Murphy, and Williams (1989) 等學者所指出。此外，亦有學者認為：績效評估若僅建構在單一之績效評估者上，則顯然在績效評估相關資訊之取得必有不足之處(Murphy and Cleveland, 1995)。

文獻中對於與顧客評估(customer as rating source)相關之研究探討較少。Bretz, Milkovich, and Read (1992) 曾指出：在某些情境下(specific situations)，以顧客為主軸的績效評估也許可以提供績效評估所需之相關資訊，不過作者亦認為，以顧客為績效評估者也可能產生某些研究上問題，如：顧客對於員工工作之瞭解程度有限(limited job knowledge)以及可能產生所謂之選擇性偏誤(selection bias)現象。此外，Pollack and Pollack (1996) 研究中則明白指出：在少數論及以顧客為績效評估者文獻中，對於以顧客為績效評估者之研究，尚僅止於概念探討階段。

但依據 Borman (1997) 及 Wherry and Bartlett (1982) 表示，評估者與被評估者若有較密切之關係(close proximity)，則可獲得較正確的績效評估結果；而 Murphy and Cleveland (1995) 亦曾明確指出，組織

在發展績效評估體系時，必須把顧客的想法納入。

因此，綜觀上述學者研究所獲致之論點，似乎隱含顧客在「某些現象」下，可嘗試引入為績效評估者，而透過顧客之角度，似乎也可提供管理者有異於其他績效評估者，如：直屬主管、同事、下屬及自我評估等所提供之有效管理資訊。總之，過往學者主要致力於發展或補足績效評估者此領域知識認識之深度與廣度，並進一步擴充績效評估者此知識之本體。本研究即在此理論背景下，從各種不同績效評估者理論探討中，進一步嘗試發展以顧客為績效評估者此一知識。

三、結果與討論

研究之資料搜集主要分為顧客與員工兩部份，其樣本則來自一家位於台北之大型綜合旅行社所提供之顧客名單及該公司員工。該公司主要業務涵蓋所有商務與休閒旅遊服務，並有 4 家分公司分設於北、南、東、西三地，目前並通過 ISO-9002 認證，員工人數共有 165 人。考量到時間落差，對評估之正確性影響。故請該公司提供最近 6 個月內參加過其團體旅遊之顧客名單資料。此外，該公司並提供小型會議室 1 間，供研究人員及訪員在研究期間之資料搜集並整理及開會討論時所用。

經由嚴謹的資料搜集過程後，研究結果發現，顧客為績效評估者，在 11 者績效評估者中，有 7 者其權重值為最高(經由 AHP 方法之計算)，包含有含溝通能力並力藥基本常識能力並解說能力並個人之耐心並親切感並責任感及服務熱忱，其整體權重值為所有績效評估者中，包含直屬主管並下屬並同事並自我並高層主管及顧客中最高者。此結果顯示以顧客為績效評估者是合適的，本研究之假設獲得支持。

持者 Borman(1974) 曾指出，組織內位處不同職位層級的績效評估者，能顯著觀察到被評估者各種不同層面之工作表現。本研究認同這樣之看法，但是否績效評估者一定要為組織內之員工呢？此外，兩位持者 Pollack and Pollack(1996) 曾

指出，相較於評估員工個人，顧客較適宜評估產品或服務。

服務業員工之特性及其角色之重要，過往文獻已有許多探討。而某些服務業並與顧客有密切接觸之員工，譬如含旅行業並之領隊並壽險業之保險經紀人... 紀，其扮演之角色可能更形重要，但往往其工作之執行卻在管理者可以觀察範圍之外，團體言之即，顧客很可能在消費或購買服務產品過程，並不會接觸到提供服務之公司，但卻一定會接觸到這些實際提供服務之員工。顯際，這些員工在此部份服務實際之工作表現與績效，對於服務業管理者而言是固有重要意義。

針對上述有此特性之服務業員工績效評估問題，文獻上缺乏探討，而乏本研究所進行之實務電話訪問調查，查可發現實務管理者針對此一不看查員工績效評估問題，則仍際依循一般較際統之績效評估者進行員工工作考評，查即，以直屬主管及高層主管為主要績效評估者。

統過本研究之實證，此一績效評估問題獲得進一步之釐清，績效評估者不一定要清處組織內，針對不看查員工之績效評估問題，組織外之績效評估者-顧客，較其清組織內之績效評估者來之合適。而此一研究結果，不僅強化了對於以顧客為績效評估者有關知識之瞭解，而乏認識論觀瞭，更可補足績效評估者相關理論尚有不足之處。

四、計畫成果自評

研究之最終結果與原計畫相當符合，且在有關於團體旅遊理論(group travel theory)部份，本研究之發現與既有理論有諸多差異之處，其結果可顯著擴充既有理論尚有不足之處。研究部份內容於日前投稿於 Tourism Management (SSCI)，目前已獲該國際學術期刊之接受，擬於年底或明年初出版。

五、參考文獻

1. Bernardin, H. J., "An Analytic Framework for Customer-based Performance Content Development and Appraisal," Human Resource Management Review, Vol. 2, No. 1, (1992), p. 83.
2. Bernardin, H. J., S. A. Dahmus, and G. Redmon, "Attitudes of First-line Supervisors toward Subordinate Appraisals," Human Resource Management, Vol. 32, No. 2 & 3, (1993), pp. 315-324.
3. Blood, M. R., "Spin-offs from Behavioral Expectation Scale Procedures," Journal of Applied Psychology, Vol. 59, No. 4, (1974), pp. 513-515.
4. Borman, W. C., "360° Ratings: An Analysis of Assumptions and a Research Agenda for Evaluating Their Validity," Human Resource Management Review, Vol. 7, No. 3, (1997), pp. 299-315.
5. Borman, W. C., "The Rating of Individuals in Organizations: An Alternate Approach," Organizational Behavior and Human Performance, Vol. 12, No. 1, (1974), pp. 105-124.
6. Bozeman, D. P., "Interrater Agreement in Multi-source Performance Appraisal: A Commentary," Journal of Organizational Behavior, Vol. 18, No. 4, (1997), pp. 313-316.
7. Bretz Jr., R. D., G. T. Milkovich, and W. Read, "The Current State of Performance Appraisal Research and Practice: Concerns, Directions, and Implications," Journal of Management, Vol. 18, No. 2, (1992), pp. 321-352.
8. Campbell, D. J., "Editorial: Some Remarks from the Outgoing Editor," Journal of Applied Psychology, Vol. 67, No. 6, (1982), pp. 671-700.
9. Campbell, D. T. and D. W. Fiske, "Convergent and Discriminant Validation by the Multitrait-multimethod Matrix," Psychological Bulletin, Vol. 56, No. 2, (1959), pp. 81-105.
10. Cleveland, J. N., K. R. Murphy, and R. E. Williams, "Multiple Uses of Performance Appraisal: Prevalence and Correlates," Journal of Applied Psychology, Vol. 74, No. 1, (1989), pp. 130-135.
11. Farh, J. L. and J. D. Werbel, "Effects of Purpose of the Appraisal and Expectation of Validation on Self-appraisal Leniency," Journal of Applied Psychology, Vol. 71, No. 3, (1986), pp. 527-529.
12. Farh, J. L., J. D. Werbel, and A. G. Bedeian, "An Empirical Investigation of Self-appraisal-based Performance Evaluation," Personnel Psychology, Vol. 41, No. 1, (1988), pp. 141-156.
13. Fox, S. and Y. Dinur, "Validity of Self-assessment: A Field Evaluation," Personnel Psychology, Vol. 41, No. 3, (1988), pp. 581-592.
14. Fox, S., Z. Ben-Nahum, and Y. Yinon, "Perceived Similarity and Accuracy of Peer Ratings," Journal of Applied Psychology, Vol. 74, No. 5, (1989), pp. 781-786.
15. Harris, M. M. and J. Schaubroeck, "A Meta-analysis of Self-supervisor, Self-peer, and Peer-supervisor Ratings," Personnel Psychology, Vol. 41, No. 1, (1988), pp. 43-62.
16. Ilgen, D. R., J. L. Barnes-Farrell, and D. B. McKellin, "Performance Appraisal Process Research in the 1980s: What Has It Contributed to Appraisals in Use?" Organizational Behavior and Human Decision Processes, Vol. 54, No. 3, (1993), p. 362.
17. Landy, F. J. and J. L. Farr, "Performance Rating," Psychological Bulletin, Vol. 87, No. 1, (1980), pp. 72-107.
18. Lawler III, E. E., "The Multitrait-multirater Approach to Measuring Managerial Job Performance," Journal of Applied Psychology, Vol. 51, No. 5, (1967), pp. 369-381.
19. Mabe III, P. A. and S. G. West, "Validity of Self-evaluation of Ability: A Review and Meta-analysis," Journal of Applied Psychology, Vol. 67, No. 3, (1982), pp. 280-296.
20. Murphy, K. R. and J. N. Cleveland, Understanding Performance Appraisal: Social, Organizational, and Goal-based Perspectives, (California: SAGE Publications, 1995).
21. Pollack, D. M. and L. J. Pollack, "Using 360° Feedback in Performance Appraisal," Public Personnel Management, Vol. 25, No. 4, (1996), pp. 507-528.
22. Schrader, B. W. and D. D. Steiner, "Common Comparison Standards: An Approach to Improving Agreement Between Self and Supervisory Performance Ratings," Journal of Applied Psychology, Vol. 81, No. 6, (1996), pp. 813-820.
23. Smith, P. C. and L. M. Kendall, "Retranslation of Expectations: An Approach to the Construction of Unambiguous Anchors for Rating Scales," Journal of Applied Psychology, Vol. 47, No. 2, (1963), pp. 149-155.
24. Vance, R. J., R. C. MacCallum, M. D. Coovet, and J. W. Hedge, "Construct Validity of Multiple Job Performance Measures Using Confirmatory Factor Analysis," Journal of Applied Psychology, Vol. 73, No. 1, (1988), pp. 74-80.
25. Wherry, R. J. and C. J. Bartlett, "The Control of Bias in Ratings: A Theory of Rating," Personnel Psychology, Vol. 35, No. 3, (1982), pp. 521-551.