

Heuristic Walkthrough 在網站人機界面可用性評估之應用

許銀雄 詹榮昌

銘傳大學資訊管理研究所

桃園縣龜山鄉大同村德明路五號

hsong@mcu.edu.tw

r6750081@ss21.mcu.edu.tw

摘要

由於使用人口激增，在網際網路(Internet)上建立全球資訊網(WWW)網站已漸漸成為一個機構或組織重要的工作事項之一。而由於技術的進步使得網站的建立變的非常容易。然而，網站的容易建立並不保證網站的品質，所以我們才會在 Internet 上看到許多品質不良的網站。網站的品質因素主要包括內容、技術以及界面三者：其中內容以及技術比較廣為設計者所重視也熟知，而網站的界面則比較容易被忽略，國內大部分的網站設計者對人機界面的觀點尤為薄弱。這樣的結果就導致國內有許多可用性(Usability)不佳的網站。

為確保人機界面的可用性，其設計必須採用以使用者為中心(User-centered)的設計方法，而可用性評估是這類方法最核心的部分。Heuristic Walkthrough 則是一個簡單而又有效的人機界面可用性評估方法。本研究的主要目的是透過實驗的方式來瞭解這個方法在網站界面可用性評估時的一些特性以及瞭解國內網站可用性問題的實際情形。而實驗的進行則是選定國內具代表性的網站為評估對象，並由具備初級人機界面知識的大學學生做為評估人員來運用這個方法進行網站的可用性評估。

關鍵詞：Heuristic Walkthrough, Usability Evaluation, World Wide Web

一、簡介

由於採用超連結(hyperlink)簡易的呈現及操作方式，使得全球資訊網的使用不再侷限於技術人員，而結合多媒體的表現方式，使得 WWW 得以多采多姿的方式呈現，再加上政府的大力推動使得網路快速普及，種種因素促成了 WWW 的使用者數目以極快的速度成長。行政院三年三百萬人上網的目標便在這樣的機緣下提早於一九九九年一月達成。使用者數目的大量增加代表著 WWW 將走入家庭、進入一般人的生活之中，這也意謂著 WWW 上充滿顧客、商機、以及利潤。各企業機構，不管大小，紛紛建立網站以提供服務進而拓展商機，而政府各單位也為因應電子化政府的來臨而紛紛將服務搬上網路以達便民之效果。設立網站已漸漸成為機構或組織中一項重要的任務。

然而，由於網站及網頁設置的技術及工具不斷更新，再加上設計者不一定具有資訊相關背景，設計者往往將重點放在網站或網頁設置的技術以及工具的使用上，而對於網站及網頁的界面及可用性則大多僅憑自己的想法

加以考量甚或予以忽略。這樣子的結果導致有許多網站的人機界面設計不良，再加上網站的使用者是一般大眾、初次使用者、或是初學者，這些設計不良的人機界面，很容易造成使用者不知如何操作網站、不容易正確得到所要的資料、以及浪費許多額外的時間在已經非常擁塞的網路上等等[1]。而如果這些網站是許多人都必須使用的，那整體所產生的浪費就非常可觀。所以，在進行網站的設計時，其人機界面的考量比在設計傳統軟體時更為重要。

人機界面的設計方法大多強調以使用者為中心的觀念，所以這些設計方法被通稱為使用者中心設計方法(User-Centered Design, UCD)[2,3]，而可用性(Usability)評估是 UCD 的重要元素[2]。Heuristic Walkthrough[4]是結合 Heuristic Evaluation[5]以及 Cognitive Walkthrough[5]兩種方法的一種人機界面可用性評估方法，同時兼具這兩種方法的優點，是一種簡易、花費少而又有效的可用性評估方法。之前關於這類方法的研究以及數據，都是在於傳統軟體界面的可用性上[4,6,7]。然而，網站的人機界面的特性

與傳統軟體的人機界面卻不盡相同，例如：網站的使用者是一般大眾無法準確定義、網站大多以資訊提供為導向、網站開發的方式與工具不同等等。所以對於網站界面的設計者而言，目前並沒有相關的數據可供作為其界面評估設計之參考。

本研究之主要目的有二：一是希望找出 Heuristic Walkthrough 應用在網站界面可用性評估的相關數據，以提供網站設計者作為其網站界面設計規劃之參考；二是呈現國內部分具代表性網站人機界面的情形，以便提供作為網站相關人員之參考。本研究將採用實驗的方式來進行。

本篇論文的架構為：第二節介紹 Heuristic Walkthrough 方法；第三節介紹實驗之設計；第四節介紹實驗之結果與分析；最後則是結論與探討。

二、Heuristic Walkthrough 評估方法

可用性檢視方法(Usability Inspection Method)是指由評估者依據某種程序與規範來檢視人機界面，以發現其中可用性問題的一種方式 [8]。Heuristic Evaluation (以下簡稱 HE) 以及 Cognitive Walkthrough (以下簡稱 CW) 是兩個很常用的可用性檢視方法，而 Heuristic Walkthrough (以下簡稱 HW) 則是將這兩種方法結合，藉此來達到兩者互補的效果。以下將先介紹 HE 以及 CW，然後再對 HW 做一介紹。

Heuristic Evaluation

HE 這個方法是由評估者個別自由地檢視受評估之人機界面，評估者依據十個主要原則(Guidelines)，檢視此人機界面中是否有違背這些原則之處，藉由此一方法找出人機界面中的可能問題。圖一所列的即是 Jacob Nielsen 所使用的十個原則 [1]。當所有評估者列出所找到之問題之後，再由主持人整理這些問題，然後在將這些問題交由所有評估者評估這些問題的嚴重度(Severity)，最後再由主持人整理這些問題並將各個評估者所評之嚴重度加以平均以作為此問題之嚴重度。

這個方法的優點是它簡單、快速、成本低、而且評估人員不一定要是人機界面專家。以人機界面專家來擔任評估者的效率當然會比以非人機界面專家來的好。根據 Nielsen 的實驗 [5]，五個專家可以找出超過 75% 的可用性

問題；而如果是五個一般的評估者，則可以找出大約 50% 的可用性問題。而評估者的人數以三至五人的經濟效益最高。

1. 系統狀態可以明白表示出來。
2. 系統與真實狀態相契合。
3. 使用者能自由地操控界面。
4. 整體符合一致性或標準。
5. 可以預防錯誤的產生。
6. 使用辨識而非使用記憶回想。
7. 使用上的彈性與效率。
8. 美觀且精簡的設計。
9. 幫助使用者確認及診斷錯誤，並在錯誤發生時，幫助使用者復原。
10. 線上輔助及文件。

圖一、用於 Heuristic Evaluation 的十個原則

Cognitive Walkthrough

Cognitive Walkthrough 主要是源自於認知模型 [9]，它把認知模型轉換成數個簡單的問題，來作為評估者進行評估時之依據。在評估進行之前，界面設計者首先必須先定義受測系統的使用者、使用環境、主要的工作情節(Task scenario)、以及依目前設計完成各工作情節的動作順序(Sequence of actions)等。而在評估過程中，再由評估者依據使用者特性以及使用環境來就每個工作情節的每個動作或步驟詢問那些由模型轉換成的問題，來找出目前設計的可用性問題。圖二所列即是那些被詢問的問題 [8]。

因為這個方法是以工作 (Task) 為中心，分別對主要工作作較深入的分析以及評估，所以它發現的問題比較有深度；然而，也正因为如此，它通常無法發現整體的問題，也就是廣度不夠。

Heuristic Walkthrough

Heuristic Walkthrough 主要是結合 Heuristic Evaluation 以及 Cognitive Walkthrough 兩種方法，希望透過這樣的結合來改進 Heuristic Evaluation 深度不夠以及 Cognitive Walkthrough 廣度不夠的缺點，來達到

互補的效果[7]。它的實施分成兩個階段：先進行 Cognitive Walkthrough 再進行 Heuristic Evaluation。這樣安排的另一個目的是：評估者在第一階段做完，比較正確地瞭解系統主要的工作之後，可以在第二階段的 Heuristic Evaluation 中找到較確切的問題。

1. 使用者是否清楚要達成這個效果？
 2. 使用者是否清楚系統提供某個正確的动作來達成這個效果？
 3. 使用者是否可以利用系統提供的動作正確地達成這個效果？
 4. 界面是否有提供適當的回應，讓使用者感覺工作進行順利？
- 圖二、Cognitive Walkthrough 所使用的問題

三、實驗設計

本節將分別就實驗分組、網站的選擇、實驗之流程以及嚴重度分別進行說明。

實驗分組

本實驗的人員是由大學資訊管理研究所以及資訊管理學系選修人機界面課程的兩班同學組成。總共分為八組，每組由一位研究生擔任主持人負責統籌實驗的準備以及進行，另外每組有四至五位的大學部同學擔任評估人員。

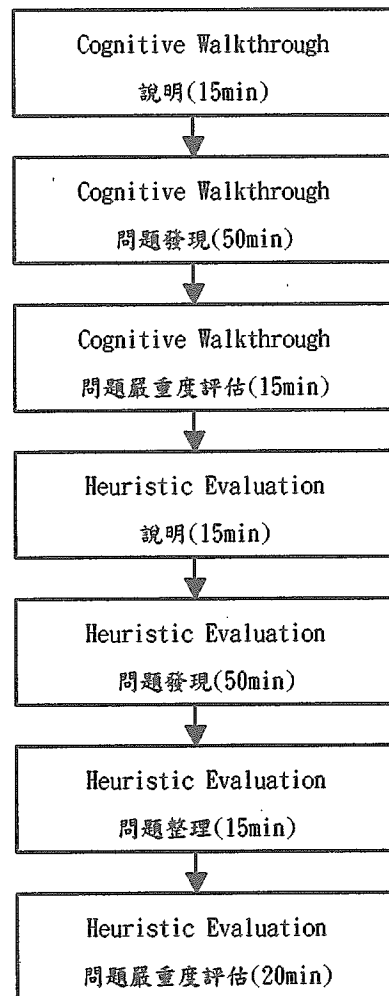
網站的選擇

為了讓本實驗的結果可以呈現國內網站人機界面的現況，我們希望選擇一些具有代表性的網站作為實驗對象，我們選擇了四個民間網站以及四個政府網站，分別是一個一九九八年第一屆網際金像獎的得獎網站、兩個被 PC Home 選為 100 中文好站的網站（分別屬新聞以及線上購物網站）、以及四個具代表性的政府網站（分別屬中央級、部會級、縣市政府、整合性政府網站）。

實驗的流程

實驗的流程是依照 Heuristic Walkthrough 的進行程序來做，其程序如圖三所示。其中先進行的 Cognitive Walkthrough 是採用同組人員一起討論的方式進行，而之後的 Heuristic Evaluation 則是評估人員個別進行評估。而評估的時間，兩種方法分別有五十分鐘，主要是考量到評

估者的注意力以及體力之負荷。而由於這樣子的時間限制，所以在 Cognitive Walkthrough 評估中，每組都是使用了十個工作情節來進行評估，而在 Heuristic Evaluation 評估時，每組則選定大約 100 個網頁進行評估。



圖三、實驗進行程序

嚴重度評估

嚴重度的評估主要的目的是讓評估者評定問題的嚴重程度，以作為設計者在進行問題改進規劃時的參考[6]。本實驗要評估者依據問題發生所造成的影響程度以及問題發生的頻率兩項因素來給予 0-3 等級的嚴重度，等級越高表示越嚴重。等級的給定是讓評估者參照表一來評定。

		問題發生所造成的影響程度		
		高	中	低
問題發生的頻率	高	3	3	2
	中	3	2	1

表一、嚴重度評定表

四、實驗結果整理與分析

本節將分別就實驗的結果加以說明。

首先，表二所示的是八個實驗網站的基本資料：其中評估人數是該小組的評估者人數；而網站大小則是指該網站包含的網頁數目。原先各組都選定了大約一百個網頁要評估者在五十分鐘內評估完畢，後來發現網頁長短以及網站品質都不一致等因素，導致各組無法在五十分鐘內評估完原先預設之網頁，也因此產生了不小的差異。表中的評估頁數是該小組五十分鐘內所評估的網頁數目。

網站編號	評估人數	網站大小	評估頁數
1	5	約250頁	101
2	5	約110頁	76
3	5	>1500頁	42
4	5	79頁	79
5	5	>1000頁	69
6	4	>3500頁	40
7	5	約250頁	74
8	4	>2000頁	65

表二、網站基本資料

表三所示的是各個網站在實驗中所發現的問題數。其中，CW與HE問題數分別是Cognitive Walkthrough以及Heuristic Evaluation評估所發現的問題，而總問題數則是把CW與HE問題扣除掉相同的問題之後的問題數。由這個表我們可以發現以下幾個現象。首先是CW以及HE所共同發現的問題並不多，這與之前在傳統軟體的人機界面上的實驗很不一樣[4,6]。其原因應該是傳統軟體的人機界面較為單純，所以兩種方法注意的情事較為接近，而網站的內容則較為複雜導致兩種方法在進行時專注到其各自的重點有關。從這點來看，我們可以知道網站的界面確實與傳統軟體的界面有不一樣的性質，以及網站在進行可用性評估時必須要同時使用兩種以上不同性質的評估方法。第二個現象是Heuristic Evaluation所發現的問題明顯比Cognitive Walkthrough所發現的問題多許多。這個現象的主要原因是本實驗所選擇的網站大部分為得獎網站，所以在架構上沒有比較嚴重的問題所致。另外，則是要來看看在表的最下方的每組發現問題數的平均值以及標準差。這標準差蠻大的，會導致這樣的結果的主要原因有二：一是網站之間的差異，如一號以及四號網站的架構較為簡明導致CW問題數較少；二是因為第二組人員

的年級較其它組高一個年級，素質特殊所致。

從表三我們可以看出發現總問題數的平均是48.5個問題，這代表著一個由四至五位初具人機界面知識的評估小組可以利用Heuristic Walkthrough評估方法在三個小時內找出48.5個可用性問題。另外，表二的評估網頁數平均值為68.2頁，把48.5除以68.2代表著在這次實驗中，每個網頁大約可以找出0.71個問題。所以，假設是一個具有一百個網頁的小型網站，依此數據則大約至少會有71個可用性問題。而如果再考量Nielsen[5]的數據，這四至五個一般評估者大約只可以找到50%的問題，則這個小型網站所隱藏的問題數可能超過一百個。

網站編號	CW 問題數	HE 問題數	總 問題數	共同 問題數
1	8	23	27	4
2	24	58	78	4
3	19	40	58	1
4	8	48	51	5
5	18	23	41	0
6	18	28	46	0
7	12	39	46	5
8	15	26	41	0
平均	15.3	35.6	48.5	
標準差	5.6	12.8	14.9	

表三、所發現之問題數目

		Cognitive Walkthrough問題編號			
		1	2	3	4
網站 編號	1	2	2	3	1
	2	5	6	7	6
	3	11	8	0	0
	4	7	1	0	0
	5	6	2	5	5
	6	11	2	4	1
	7	3	1	7	1
	8	6	1	1	7
總計		51	23	27	21

表四、Cognitive Walkthrough所問題歸類

表四所示是將Cognitive Walkthrough階段所發現的問題依四個詢問問題作歸類。其中以違反第一個問題（使用者是否清楚要達成這個效果？）的問題數最多，這個問題所反應的是網站架構設計不當而導致使用者不清楚要完成他的某項工作時必須要進入某一個奇怪的網頁，例如：要察看“近期新商品”卻必須要進入“線上購物”網頁。違反第二個問題較多的是二號及三號網站，它的問題主要在於使用新奇的技術導致使用者無法知道如

何使用。違反第三個問題較多的主要是二號及七號網站，它的問題分別是使用英文讓使用者無法正確操作以及字形太小讓使用者查詢不易等。而違反第四個問題較多的是二號及八號網站，它的問題主要是錯誤的鏈結導致網頁找不到。

表五所示是將 Heuristic Evaluation 階段所發現的問題依十個原則作歸類。其中以違反第八個原則（美觀且精簡的設計）的問題數最多，大致的問題包括：版面配置凌亂、圖片文字編排不當、以及網頁太長等等問題。由此可以感覺的出來，一般的評估者還是對於界面的美觀及精簡有較高的注意。而以違反第一個原則（系統狀態可以明白顯示出來）的問題數最少。

		Heuristic Evaluation 的原則編號									
		1	2	3	4	5	6	7	8	9	10
網站編號	1	1	0	5	2	1	0	5	4	5	0
	2	4	10	5	4	5	7	8	8	4	3
	3	4	5	4	2	6	4	4	9	1	1
	4	0	9	1	3	6	4	5	16	0	4
	5	0	2	2	2	1	1	3	6	2	4
	6	1	4	2	1	2	0	3	8	3	4
	7	3	1	8	3	4	0	9	9	1	1
	8	0	2	1	2	5	2	4	6	2	2
總計		13	33	28	19	30	18	41	66	18	19

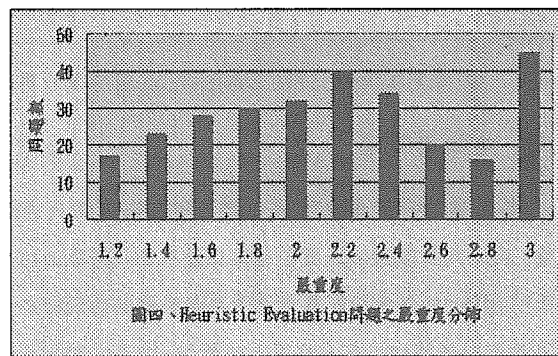
表五、Heuristic Evaluation 所發現問題歸類

Heuristic Evaluation 所發現的問題，有部分單獨存在，亦即屬於某個網頁的個別問題；有部分問題則重複出現在不同的網頁之中，我們稱這樣子的問題為整體性問題。整體性的問題通常所代表的是問題的產生並非不小心所造成，而是網站設計者的認知與使用者的預期之間有較大的差異。從表六中可以看出有蠻多這種整體性的問題，這某種程度顯示出國內網站設計者對於使用者以及人機界面的一些不正確的想法。

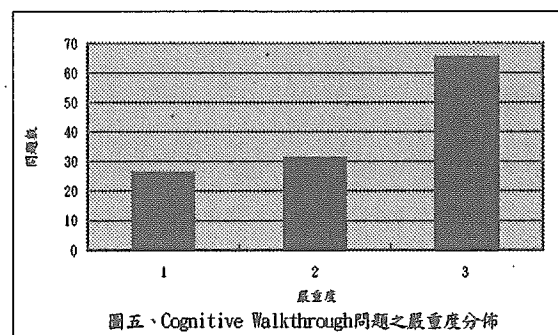
網站編號	整體性問題	個別問題
1	5	17
2	7	52
3	9	30
4	15	32
5	6	17
6	3	25
7	11	27
8	5	21

表六、Heuristic Evaluation 問題之分類

圖四所示是 Heuristic Evaluation 所發現的問題的嚴重度分佈圖，由於每個問題的嚴重度是由同組的每一位評估者評定之後再平均，所以會有小數點。圖中是統計介於區間的問題數。其中比較特殊的是介於 2.8 與 3 區間的問題數特別高，它的原因在於其中包含了不少錯誤鏈結的問題，而那類問題都被評定為最高嚴重度 3。



圖五所示則是 Cognitive Walkthrough 所發現的問題的嚴重度分佈圖，由於問題的嚴重度是由全組的評估者討論後得知，所以只有三個值。另外與圖四不同的是 Cognitive Walkthrough 所發現的問題中，有最多的問題是嚴重度 3，而嚴重度是 1 的問題數則最少。嚴重度為 3 的問題中有一大部分是來自於第一個詢問問題，亦即是網站架構的問題。另外，由於 Cognitive Walkthrough 所評的工作項目都是該網站主要的工作項目，而且是依據設定的程序來進行，所以如果發生問題通常都是主要的問題，問題的嚴重度通常也會比較高。



表七所示是 Heuristic Evaluation 評估中各評估人員所發現的問題數目。其中除了第三、五、六組的二號評估人員較為突出外，變異係數大約維持在 30% 以下。若以此發現問題數的平均而言，每人在 Heuristic Evaluation 過程中所發現的問題的平均數目為 10.8 個問題。

網站 編號	評估 人數	各組評估人員編號					平均	標準 差	變異 係數
		1	2	3	4	5			
1	5	5	8	7	9	5	6.8	1.8	26.3
2	5	13	15	20	14	14	15.2	2.8	18.2
3	5	13	17	9	10	8	11.4	3.6	32.0
4	5	16	9	12	10	18	13.0	3.9	29.8
5	5	5	11	6	5	3	6.02	3.0	50.0
6	4	7	15	11	7		10.0	3.8	38.3
7	5	13	16	18	12	15	14.8	2.4	16.1
8	4	8	7	7	11		8.3	1.9	22.9

表七、評估者在 Heuristic Evaluation 時發現的問題數

五、結論與探討

由上述的分析與說明，我們大概可以有下列幾點結論。首先是國內網站仍然普遍存在著不少人機界面的問題而不自知，而這些問題嚴重性主要在於網站的設計者與使用者之間的認知仍有一段不小的差距。而這之間的差距仍有賴使用者以及設計者的覺醒以及網站人機界面設計方法的研究來加以改善。其次是網站人機界面與傳統軟體人機界面的特性並不完全相同，而這之間的差異會導致評估方法運用時的性質轉換，所以有需要針對網站的人機界面設計與評估方法進行更進一步的研究以符合網站的需求。最後則是本研究雖然是採用了大學部剛修習人機界面課程而且是第一次使用這方法的學生當作評估人員，仍能在有限的時間內找到不少的網站人機界面問題，我們相信只要經過幾次的磨練，這些評估人員的效率應當可以再大幅的提昇，這表示 Heuristic Walkthrough 對於網站人機界面的評估是有效的。

參考文獻

- [1] Shum, S. B. and McKnight, C., "World Wide Web Usability: Introduction to This Special Issue," *Int. J. Human-Computer Studies*, 47, 1997, pp.1-4.
- [2] Gould, J. D. and Lewis, C. "Designing for Usability: Key Principles and What Designers Think," *CACM*, 2(3), March 1985, pp.300-311.
- [3] Karat, J., "Evolving the Scope of User -Centered Design," *CACM*, 40(7), 1997, pp.33-38.
- [4] Karat, K., Campbell R., and Fiegel T., "Comparison of

Empirical Testing and Walkthrough Methods in User Interface Evaluation," proceeding of CHI'92, May 1992, pp.3-7.

- [5] Nielsen, J. and Molich, R., "Heuristic Evaluation of User Interfaces," *ACM CHI'90*, pp.249-256, 1990.
- [6] Jeffries, R., Miller, J. R., et al., "User interface evaluation in the real world: a comparison of four techniques," in proceedings of ACM CHI'91, April 1991, pp.119-124.
- [7] Sears, A., "Heuristic Walkthroughs: Finding the Problems Without the Noise", *International Journal of Human-Computer Interaction*, 1997, pp.213-234.
- [8] Nielsen, J. and Mack, R. L. ed., *Usability Inspection Methods*, John Wiley & Sons, 1994.
- [9] Polson, P., Lewis, C., et al., "Cognitive walkthroughs: A method for theory-based evaluation of user interface," *International Journal of Man-Machine Studies*, 1992.