

明尼蘇塔多相人格測驗的修訂

路 君 約

明尼蘇塔多相人格測驗 (Minnesota Multi-phasic Personality Inventory, 簡稱 MMPI), 原由美國明尼蘇塔大學精神病學系臨床心理學組主任 (Hathaway, S.R.) 教授及精神病學者麥肯勒 (Mckinley, J. C.) 教授所編製的。開始於1930年, 於1940年刊布其研究報告, 1943年正式由美國明尼蘇塔大學出版部出版, 1945年改由美國心理學公司 (Psychological Corporation) 承擔發行。旋即成爲應用最廣而接受研究最多的一種人格測驗⁽¹⁾。

最先介紹明尼蘇塔多相人格測驗到我國來的, 是民國四十三年國立臺灣大學心理學系應屆畢業生黃碧月和包重齡兩女士⁽²⁾。她們把它譯成中文, 刪除了十五個測驗題; 並從臺灣大學附屬病院、松山錫口療養院及臺北監獄, 取樣五十八人, 作初步的嘗試; 以此作爲畢業論文。

民國四十八年我國海軍在決定修訂是項測驗之後, 爲了保存原測驗的內容、題數和形式, 將每一個測驗題, 重新譯過。對於不適於我國文化背景的測驗題, 未加刪除, 僅根據原題的涵義, 換用相當的中國材料。同時, 力求合乎中國語法, 字斟句酌, 費時甚多。

修訂的工作已於今年完成, 現在將其修訂過程及結論, 報告於後以請教正。

一、取 樣

我國海軍修訂這個測驗的目的, 主要期望它能成爲檢定下列各種人員的初步工具; 適任海上戰鬥任務及艙面或艙內工作人員的人格組型, 情緒不穩定及有情緒困難的人員, 異常人格或異常人格傾向的人員。因此, 測驗的母體 (Population) 自然以軍事人員爲主。標準化樣本 (Standardized sample) 中有一部分爲大學學生及公務人員, 原在作比較之用。關於女性樣本, 由於軍中選配女性樣本困難; 就軍中說, 女性測驗的需要, 亦遠不及男性爲迫切, 故在作計劃時即未計及女性樣本。至於所取大學生及公務人員樣本中雖有少數女性樣本, 但數量太少, 未作統計處理。

所取標準化樣本, 其分佈如表一:

表一 標準化樣本的分佈

年 齡	類 別 人 數	軍事單位	大 專 學 生		公 務 人 員		總 數	
			男	女	男	女	男	女
16—25		627	486	261	41	42	1,154	303
26—43		1,077	197	4	212	43	1,486	47
44—54		20	8	—	2	5	30	5
55—		—	—	—	4	—	4	—
其 他		258	83	—	9	—	350	—
總 數		1,982	774	265	268	90	3,024	355

註: 詳細資料, 見海軍報告。

正常人員組，計共取得男性樣本3,024，女性樣本355。但後來於評閱過程中，凡發現作答不全，或有任意作答之可能者，盡行剔除。故實際用作標準化根據者，僅2,420人，其中計軍事人員部分有1,716人，佔71%；大學生部分有443人，佔18%；公務人員部分有261人，佔11%。

至於效標組，依據海軍修訂本測驗的目的，計取得心理異常的人員、監所受刑人員及問題行為等三類，共計979人，如表二、表三及表四。

表二 心理異常的樣本

年 齡	類 別 人 數	Anxiety		Paranoia	Schiz.		Hypocho.		Schiz(pa)	Obse- Comp.	De-Ma	Hysteria	
		男	女		男	女	男	女				男	女
16-25		3	1		5	1			1	1	2	6	
25-43		13		15	18	3	7		21	3	3	10	1
44-54				1	1		1	1	2		3		1
55-													
其 他					3					1	1		

N=130

* 本項樣本全部由精神病學醫生代為收集

表三 監所受刑人員的樣本 *

年 齡	類 別 人 數	搶	奪	傷	害	偷	竊	風	化	毒	品	其	他
		16-25		18		40		51		4		9	
26-43		46		36		36		11		18		39	
44-54		2		1								3	
55-													
其 他												10	

* 含新竹、高雄感化院及軍人看守所

表四 問題行為的樣本分布

年 齡	類 別 人 數	想	想	工	無	情	心	孤	憂	難	複	身	神	多	嚴	其	總
		家	退	作	工	緒	理	獨	鬱	於	補	體	經	病	重	他	和
16-25		4	5	10	2	2	3	2	1	4	8	5	4	1	7	22	80
26-43		10	66	31	29	16	18	24	6	39	42	41	20	14	8	15	379
44-54		1	2	1	1			2		1	2	2	1	1			14
55-					1												1
其 他			5	7	2	5	1		1	3	6	5	3	3	1	2	44

註：類別名稱及問題行為，均為由單位提出者。

心理異常人員的樣本，不但選擇困難，而對病人實施測驗尤為不易。這些樣本都是商請北投精神病醫生代做的，首先經院中診斷確定為某種心理疾病之後，由醫生或護理人員取得患者合作，進

行測驗。所取樣本 130 人，費時達半年之久。

監所受刑人員之樣本，係依照判決之罪行而選取的。取得監所之充分合作，派員進行測驗。至於問題行為之樣本，則先請求各有關單位調查製冊，然後由海軍總部定期派員前往舉行測驗。

正常人員樣本及效標組樣本，自選定至全部測驗完畢，前後費時凡一年半。

二、分析程序

正常人員組因為事先不能確定其特點，故在分析過程中除男女兩性外，復分為軍人、學生、公務人員及全體樣本四大類。⁽³⁾

1. 利用原來四個效度量表及臨床量表的標準，評閱正常人員組及效標組的全部測驗卷，分別計算其平均數及標準差等。藉以與美國原測驗結果相比較。

2. 分析作答人數百分比；依照“是”、“否”及“疑問”（未作答）三種作答方式，分別統計各組受試者對每一題的作答百分比。但因正常人員組有四類，每類又有年齡區分，所以在分析歷程中極為紛繁，並特別為了便於這種分析而設計了一種登記表格。

3. 修訂原來的效度量表、臨床量表及建立新量表。檢驗各種正常人員組對每一測驗題“是”與“否”作答百分比，與效標組作答百分比差數的可信度。依據可信度的水準為各效標組選定測驗題及其作答方式。

檢定兩個百分比的差數的可信度，係求這兩個百分比差數的標準誤，用下列的公式。

$$\sigma_{D\%} = \sigma_{p_1 - p_2} = \sqrt{PQ \left[\frac{1}{N_1} + \frac{1}{N_2} \right]}$$

$$P = \frac{N_1 P_1 + N_2 P_2}{N_1 + N_2}$$

$$Q = 1 - P, \quad \text{然後再求CR}$$

$$CR = \frac{(P_1 - P_2) - 0}{\sigma_{p_1 - p_2}}$$

CR 等於 2.58，則可信水準達 0.01；因此，為各效標組選題，應以 2.58 為最低標準。美國原來為各效標組選題，其兩個百分比之差；至少應為其標準誤的二倍。而大多數效標組的題則在二倍以上⁽⁴⁾。其次選題時尚須注意究竟應如何選定“是”或“否”的作答方式；這須看效標組係以那種作答方式，則其百分比高於正常組為定。

4. 選定各效標組的測驗題之後，重新評閱正常人員組之測驗卷，分別求其平均數及標準差等，以便制訂常模。

5. 根據各正常人員組在量表上的成績，製訂標準分數及側面圖。

三、結果

依據前項分析程序，所得結果如下：

1. 我國樣本在美國原效度量表及臨床量表上之結果：

(1) 正常人員組在美國原量表之平均數、標準差，以及各樣本平均成績之差的可信水準，如表五。

表五 正常組各種樣本在美國原量表上之統計量數及平均數之差的可信水準

項 分 類 別	N	L	F	K	?	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	
全 體 樣 本	2,389	M	11.35	14.05	7.93	10.35	27.53	23.26	17.32	28.57	12.72	18.97	22.09	18.91	
		σ	6.85	3.92	16.21	6.71	5.89	6.28	6.02	6.02	8.08	3.72	8.66	10.94	3.24
		σ_m σ_a	0.14 0.04	0.08 0.06	0.33 0.23	0.14 0.10	0.12 0.09	0.19	0.09	0.12	0.17 0.12	0.18 0.13	0.18 0.13	0.22 0.16	0.07 0.05
軍 人	1,701	M	8.51	13.94	7.40	11.24	28.10	23.91	17.63	28.22	13.02	19.33	22.82	18.84	
		σ	2.96	4.19	13.06	6.68	5.92	6.52	5.41	5.41	6.75	3.52	8.74	10.93	4.41
		σ_m^2 σ_a^2	0.0052	0.0104	0.1101	0.0265	0.0207	0.0251	0.0173	0.0270	0.0450	0.0073	0.0450	0.0706	0.0115
學 生	443	M	6.56	13.88	11.01	8.67	26.52	21.78	17.21	30.16	12.56	20.33	22.67	19.72	
		σ	2.88	7.14	14.25	5.24	6.16	5.50	8.58	5.41	5.41	4.59	7.98	11.01	5.06
		σ_m^2 σ_a^2	0.0187	0.1149	0.4903	0.2150	0.0836	0.0684	0.1665	0.0663	0.0663	0.0372	0.1439	0.2817	0.0579
公 務 人 員	259	M	8.50	15.04	6.12	7.37	25.49	21.45	15.47	28.12	10.86	14.24	16.50	18.00	
		σ	3.17	2.52	12.06	5.41	4.49	5.30	4.60	4.60	6.86	4.25	7.55	8.96	4.70
		σ_m^2 σ_a^2	0.0392	0.0248	0.6232	0.1164	0.0781	0.1094	0.0825	0.1835	0.1835	0.0703	0.2221	0.3130	0.086
軍人與學生平均數之差的 可信度		$\sigma_{m1}^2 + \sigma_{m2}^2$	0.0239	0.1253	0.6013	0.2415	0.1064	0.0935	0.1839	0.0933	0.0445	0.1891	0.3523	0.0694	
		σ_D	0.154	0.354	0.775	0.491	0.326	0.306	0.429	0.325	0.211	0.435	0.594	0.263	
		CR	1.93 12.00	0.06	3.61 4.60	2.57 5.00	1.58 4.00	2.13 6.80	0.42	1.94 6.00	0.46 2.18	1.00 2.29	0.15	0.15	0.88 3.34
軍人與公務人員平均數 之差的可信度		$\sigma_{m1}^2 + \sigma_{m2}^2$	0.0444	0.0352	0.7333	0.1429	0.0993	0.1346	0.0998	0.2105	0.0777	0.2672	0.3836	0.0976	
		σ_D	0.211	0.187	0.856	0.378	0.315	0.367	0.316	0.458	0.278	0.517	0.619	0.312	
		CR	0.01	1.10	1.28	3.87 10.00	2.61 8.00	2.46 6.00	2.16 6.90	0.10	2.16 7.00	5.09 9.00	6.32 10.00	0.84 2.69	
學生與公務人員平均數 之差的可信度		$\sigma_{m1}^2 + \sigma_{m2}^2$	0.0578	0.1397	1.1145	0.3314	0.1642	0.1778	0.2489	0.2494	0.1076	0.3661	0.5907	0.1441	
		σ_D	0.241	0.374	1.055	0.576	0.405	0.422	0.498	0.499	0.328	0.605	0.768	0.379	
		CR	1.92 7.9	1.16 3.10	4.89 4.60	1.30 2.25	1.03 2.54	0.33	1.74 3.00	2.04 4.00	1.70 5.00	6.09 10.00	6.17 7.00	1.72 4.00	

(2) 效標組在美國原量表上的平均數及標準差，如表六。

表六 效標組在美國原量表之統計量數

類別	年齡	人數	統計量數	L	F	M	?	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma
Paranoia	26-43	4	M	9.50	9.75	15.75	0.00	3.75	21.00	17.50	15.55	24.25	12.25	10.35	14.75	20.50
			σ	1.00	1.50	1.71	0.00	1.50	1.41	2.08	3.70	2.87	2.25	5.50	9.21	5.20
			σ_M	0.50	0.75	0.85	0.00	0.75	0.70	1.04	1.85	1.43	1.13	2.75	4.60	2.60
			σ_σ	0.35	0.53	0.60	0.00	0.53	0.50	0.73	1.31	1.01	0.79	1.94	3.25	1.84
Paranoia	26-43	12	M	8.50	15.75	15.17	1.92	9.08	26.00	22.33	19.50	27.92	16.00	22.33	25.75	21.42
			含 F>15 L>10 σ	2.15	6.94	3.43	6.64	6.02	5.75	7.69	5.76	5.43	4.86	9.85	13.82	6.50
			σ_M	0.62	2.00	0.99	1.92	1.74	1.66	2.22	1.66	1.57	1.40	2.85	3.99	1.88
			σ_σ	0.44	1.42	0.70	1.35	1.23	1.17	1.57	1.17	1.11	0.99	2.01	2.82	1.33
Schizophrenia	26-43	7	M	8.00	9.71	12.29	15.29	11.29	27.43	23.86	17.43	26.71	14.29	19.43	21.43	19.71
			σ	2.00	3.47	2.93	20.97	3.90	7.70	6.72	3.10	2.69	6.12	7.48	9.03	4.07
			σ_M	0.76	1.31	1.11	7.94	1.48	2.92	2.54	1.17	1.02	2.32	2.83	3.42	1.54
			σ_σ	0.58	0.93	0.78	5.61	1.04	2.06	1.80	0.83	0.72	1.64	2.00	2.41	1.09
Schizophrenia	26-43	46	M	9.65	14.43	15.02	8.19	11.54	27.67	25.61	19.52	27.52	15.41	16.83	23.06	19.69
			含 F>15 L>10 σ	3.87	7.04	4.30	18.42	6.03	6.17	6.43	5.03	4.35	4.48	8.61	11.34	4.93
			σ_M	0.42	1.04	0.63	2.72	0.89	0.91	0.95	0.74	0.64	0.65	1.27	1.67	0.73
			σ_σ	0.30	0.73	0.45	1.92	0.63	0.64	0.67	0.52	0.45	0.47	0.90	1.18	0.51

(3) 臨床量表間的相關：為檢查臨床量表間的相關，曾抽取100人，所得結果如表七。

表七 臨床量表間的相關

	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma
Hs		0.48	0.61	0.31	0.17	0.37	0.52	0.55	0.07
D			0.45	0.22	0.25	0.19	0.36	0.25	-0.35
Hy				0.29	0.12	0.16	0.06	0.11	-0.07
Pd					0.28	0.55	0.44	0.46	0.40
Mf						0.34	0.31	0.33	-0.16
Pa							0.61	0.68	0.49
Pt								0.88	0.40
Sc									0.26

N=100

2. 分析我國正常樣本作答人數百分比：⁽⁶⁾

3. 修訂原來的效度量表及臨床量表：

(1) 效度量表：

(甲) L量表：明尼蘇塔男性正常人員在 MMPI 的 L 是表上的原始分數之平均數及標準

差為 $M=4.05$, $\sigma=2.89^{(6)}$; 而我國正常樣本在該量表上之平均數為 $M=8.14$, $\sigma=2.93$ (參看表五)。平均數超出一倍以上。美國原量表規定 T 分數超過 70, 應視為無效, 即 $M+2\sigma$, 原始分數為 10 分。依據我國樣本, 則原始分數 14, 始為 T 分數 70。故原標準應不適於我國。

根據 Dr. Hathaway 所選 L 量表的內容, 主要有四項, 為有關侵犯情感、不好的思想、誘惑及缺乏控制等。選這四項題目的原則為: ①此種事實在大部分正常人員中都普遍存在或經驗過的, 而受試者通常敢於承認此種事實的存在。參看表八, 可知美國正常樣本在 L 量表上對各題的反應百分比, 均超過 50%, (僅 135, 165 兩題低於 50%)

表八 中美兩國樣本在量表上的反應百分比

題號	15	30	45	60	75	90	105	120	135	150	165	195	225	255	285
美%	52	81	66	88	94	85	77	71	27	86	43	78	73	58	90
中%	38	25	42	35	77	56	48	21	19	78	37	40	43	29	59

②此種反應或情感, 意義明顯, 而通常不為社會所接受。

由上述兩項原則, 可知每題的內容, 雖然代表“不為社會所贊許”的行為或情感; 但在個人方面來說, 即使承認也無傷大雅。同時, 由於它們的普遍存在性, 如果個人予以否認, 便表示有過分將自己向好處想的趨向了。

依據建立 L 量表的原則, 再根據我國正常人員的反應, 選取適當的題目, 編製適合我國的 L 量表。經選定的題目為 15、45、59、90、105、147、150、208、225、280、282、285、411、418、438、共十六題。其中與美國原 L 量表重複之題計有八題。我國正常樣本在此新 L 量表上的平均數為 7.09, 標準差為 3.40 (參看表十四)。就這項成績看, 如 T 分數 70, 其原始分數應為 14, 即受試者在本量表上得 14 分, 應視為無效。

(乙) F 量表: 美國原 F 量表為 64 題, 平均數為 3.88, 標準差為 4.24。我國樣本則為 $M=11.34$, $\sigma=6.85$, 相差近乎三倍了。根據臨床經驗, F 分數不能超出原分數十六分即 $M+3\sigma$, T 分數為 80 了。其選題標準原為正常人員認為“是”或“否”, 不得高於 10%, 凡和 10% 的正常人員作答方向相反者即得分。但就原量表之 64 題看, 我國正常人員組之反應分配如表九。

表九 我國正常樣本在美國原 F 量表上反應分布情形

題號	○%	?%	題號	○%	?%	題號	○%	?%
14	12	2	*113	90	1	209	7	3
*17	92	1	115	27	6	210	6	1
*20	30	32	121	9	2	211	16	2
27	15	2	123	8	1	215	30	2
31	22	2	139	22	2	218	21	2

34	26	3	146	57	2	*220	95	1
35	26	6	151	7	2	227	2	1
40	30	2	156	28	2	245	9	3
42	20	10	*164	86	2	246	5	1
48	38	3	168	29	2	247	9	4
49	13	2	*169	51	2	252	32	3
50	31	4	*177	98	0	256	36	2
53	7	2	184	20	4	*257	90	1
*54	68	4	*185	83	2	*258	38	4
56	9	1	*196	84	2	269	13	2
* 65	90	2	197	8	2	*272	82	2
66	17	6	*199	47	5	275	15	2
* 75	77	2	200	16	4	*276	87	1
* 83	88	1	202	8	2	286	31	3
85	7	1	205	9	1	288	12	2
*112	68	2	206	22	8	291	9	2
						293	22	2

* 表明記分方向應為“否”

根據表九之分析，就我國正常人員組說，合乎原F量表標準者僅有二十二題。另行選出145、187、224、360、391、512六題，共合28題。我國正常人員組在此28題上的平均分數為1.94， $\sigma=2.43$ 。因此，我國受試者在本項新量表上，其原始分數不應超過10分。因為這類題目均為正常人員很少經驗的，或很少有這種事實存在的。如果有人在這類題目上得高分，便有下列幾種可能了：①受試者不了解題意或不細心，②異於正常人員的性格及觀念，③對自己往壞的方面去想。至於究竟屬於那種可能情形，應參考其他資料或臨床量表，予以推論。也就因為效標複雜，所以得分也很紛歧，而平均數為1.94，標準差為2.43的特殊結果。

(丙) K量表：原來K量表是依據經驗程序選出30題，其正常組平均數為13.45， σ 為5.66。在我國正常樣本上為M為14.05 $\sigma=3.92$ 。這是兩國正常樣本在量表上相差最少者。

這個量表一部分功能，原是用來校正某幾個臨床量表的。不過近來已經很少再做這種校正之用，所以此次並未再作深入的分析和修訂。

(丁) ?分數：根據編製者稱，其標準並不是依經驗程序而建立的，完全根據臨床的經驗。在我國根據實施的結果，未作答超出30題者為數極少。平均數為7.93， $\sigma=16.21$ 。故仍可以30題為標準，凡未作答的題，超出30題者，應該探究其原因，以決定該項實施結果是否有效。

(2) 臨床量表：美國原有臨床量表九個。此次修訂時，對臨床量表僅在作試探性的檢查，是否適於我國。又因限於樣本之選擇，故僅對 Hysteria, Paranoia 及 Schizophrenia 三種量表予以修訂。在取樣過程中，附帶又取到 Anxiety Reactions 的樣本若干，因此，在臨床量表方面有 Psychoses 及 Neuroses 各兩個量表了。同時，為 Schizophrenic 效標組選題，係根據正常人員組全部樣本及其中之大專學生組，分別選取，藉以探討正常組如單純為學生是否有顯著差異。所以

自各單位及監所中，原選有十七類樣本。在為各類樣本選定題目以後，復求這十七類的量表的內部相關（依據正常組 675 人在各該量表上的成績），發現其中有些相關高達 0.9 以上。足證原來各單位對於問題行為之分類過於主觀，有些所謂之問題行為，名稱雖不同，實質上並無區別。將相關高達 0.9 以上之各量表合併或淘汰；最後，所發展之新量表則如表十二所載，為想家、孤獨、憂鬱、暈船、多病、工作態度、反社會及強暴八種。

(4) 各量表間的重疊題數，及由此種重疊所產生的內部相關，如表十三。

表十三 各量表間重疊題，及因此而產生的內部相關

量 表	L	F	K	1	2	3	4	5	6	7	8	9	10	11	12	13
L	16	0	0	0.03	0.07	0	0.02	0.04	0.05	0	0.06	0	0.03	0.04	0.12	0.11
F	0	28	0	0.26	0.05	0.11	0.11	0.03	0.04	0.13	0.11	0.12	0.07	0.03	0.06	0.15
K	0	0	29	0	0.05	-0.04	-0.02	0.03	0.02	0	0.02	-0.04	-0.03	0.05	-0.02	0.05
1	1	15	0	60	0.15	0.11	0.02	0.02	0.01	0.02	0.12	0.01	0.11	0.06	0.11	0.02
2	2	2	2	9	41	0.11	0.02	0.02	0.02	0.01	0.01	0.13	0.12	0.11	0.08	0.09
3	0	3	-1	5	4	25	0.02	0.16	0.08	0.14	0.12	0.10	0.12	0	0.02	0.07
4	1	7	-1	22	18	10	94	0.02	0.25	0.26	0.02	0.22	0.02	0.07	0.08	0.11
5	1	1	1	10	10	5	14	31	0.01	0.14	0.12	0.10	0.26	0.21	0.07	0.12
6	2	2	1	10	11	4	30	10	78	0.11	0.15	0.22	0.28	0.07	0.09	0.07
7	0	8	0	19	10	8	29	9	10	88	0.22	0.09	0.16	0.13	0.13	0.16
8	2	5	1	8	7	5	15	6	12	19	52	0.12	0.11	0.04	0.12	0.15
9	0	7	-2	12	9	5	25	6	11	9	9	77	0.15	0.14	0.06	0.11
10	1	3	-1	7	6	4	19	11	24	13	6	11	44	0.11	0.15	0.16
11	1	1	2	4	5	0	5	9	5	8	2	9	5	36	0.01	0.29
12	5	3	-1	8	5	1	8	4	8	12	9	5	10	6	72	0.28
13	4	6	2	12	5	3	9	6	12	14	10	8	10	19	26	58

表十三的編號順序為 1. 嚴重想家，2. 孤獨，3. 憂鬱，4. 工作態度，5. 暈船，6. 多病，7. 反社會，8. 強暴，9. Anxiety，10. Hysteria，11. Paranoia，12. Sch-I，13. Sch-II。對角線中的數字，表明僅屬於本量表的題數，左下方的數字表明與其他量表真正重疊的題數。右上方數字，表示由於此種重疊題數，而使該兩個量表間的相關所增高的數值。

4. 標準化樣本的平均數及標準差：

依據我國樣本，修訂成之效度量表 L 及 F，臨床量表 Anxiety、Hysteria、Paranoia Schiz-I、及 Schiz-II 等五個量表，以及發展之八個新量表後；根據這十五個量表，重新評閱標準化樣本之測驗卷。

不過，原有標準化樣本 2,389 人，為了運用 Computer 作統計分析，僅自其中取出四分之一強共 675 人。其統計結果如表十四：

表十四 標準化樣本的平均數及標準差

樣本	量				Anxiety (1)	Hysteria (2)	Paranoia (3)	Schiz-I (4)	Schiz-II (5)	想家 (6)	孤獨 (7)	憂鬱 (8)	工作態度 (9)	暈船 (10)	多病 (11)	反社會 (12)	強暴 (13)
	新修訂 L	新修訂 F	M	σ													
軍事人員 N=489	M	1.79	19.69	16.52	14.36	29.55	19.52	14.56	11.84	6.94	28.31	10.06	24.66	29.03	17.57		
	σ	3.28	0.83	3.26	3.15	8.06	5.56	6.22	4.43	2.75	13.54	3.32	8.36	9.81	5.67		
	σ_m	0.15	0.44	0.15	0.15	0.34	0.25	0.28	0.20	0.12	0.61	0.15	0.38	0.44	0.26		
	σ_s	0.11	0.08	0.11	0.10	0.24	0.18	0.20	0.12	0.09	0.42	0.11	0.27	0.31	0.18		
大專學生 N=120	M	1.94	18.13	13.95	11.55	21.56	18.96	13.70	9.66	6.83	23.06	7.68	20.60	26.76	16.04		
	σ	2.43	8.54	1.22	2.12	6.36	3.78	6.47	2.04	2.25	11.33	2.75	7.95	7.85	2.99		
	σ_m	0.27	0.78	0.11	0.19	0.58	0.34	0.59	0.19	0.21	1.03	0.25	0.73	0.70	0.07		
	σ_s	0.19	0.16	0.08	0.14	0.41	0.24	0.42	0.13	0.15	0.73	0.18	0.51	0.91	0.19		
公務人員 N=66	M	1.12	15.69	14.88	12.44	29.92	21.05	12.84	9.76	6.12	19.77	8.49	20.61	24.29	15.68		
	σ	4.17	6.15	2.38	4.71	6.23	4.46	3.98	2.92	2.48	8.48	2.53	6.75	7.26	3.03		
	σ_m	0.51	0.76	0.29	0.58	0.77	0.55	0.49	0.36	0.31	1.05	0.31	0.83	0.86	0.37		
	σ_s	0.36	0.16	0.21	0.41	0.54	0.39	0.35	0.25	0.22	0.75	0.22	0.59	0.63	0.26		
總人數 N=675	M	1.75	18.98	15.90	13.67	28.17	19.57	14.08	11.25	6.84	26.54	9.48	23.54	28.16	17.03		
	σ	2.40	9.55	3.10	3.44	7.85	4.50	6.14	4.20	2.65	13.10	3.30	8.35	9.36	5.15		
	σ_m	0.13	0.37	0.12	0.13	0.30	0.17	0.24	0.16	0.10	0.50	0.13	0.32	0.36	0.20		
	σ_s	0.09	0.09	0.08	0.09	0.21	0.12	0.17	0.12	0.07	0.37	0.09	0.23	0.25	0.14		

根據表十四，即可製訂各量表的標準分數換算表並製訂側面圖。標準分數係採用T分數，標準分數70，即M以上兩個標準差，標準分數30，即M以下兩個標準差。

四、信度研究

以臺灣省立師範大學教育系同學為樣本 (N=41)，兩次測驗相距三個月，其穩定係數如表十五：

表十五 臨床量表的穩定係數

量 表	1	2	3	4	5	6	7	8	9	10	11	12	13
穩定係數	0.73	0.78	0.58	0.61	0.67	0.74	0.66	0.57	0.82	0.71	0.57	0.77	0.71

五、討 論

1. 關於中譯本的問題：筆者於五十三年春在美國明尼蘇塔大學時，始知 Dr. Hathaway 初擬在美國如遇到中國受試者，即直接採用中譯本。但他為了慎重起見，特請由一位中國醫生將中譯本譯成英文，藉以檢查譯文的符合程度，經發現約有三十餘題似有較嚴重的錯誤。但經一再審核，其所以認為有較嚴重錯誤者，多係原題與我國文化情形不相符合而特予改變的。例如，原題295，I liked "Alice in Wonderland" by Lewis Carroll, 中譯本則為“我喜歡「夢遊仙境」一類的事”。原題513：I think Lincoln was greater than Washington, 中譯本則為：“我認為謀求別人的福利，比爭取本身的自由更偉大”。所謂嚴重與原文不符者，多係此類題目。同時，我們知道明尼蘇塔多相人格測驗原為一由效度標準決定的測驗 (Criterion-oriented test) ⁽⁹⁾。因此，縱然譯文或有出入，那麼，依據特殊的效度標準，選取區別力 (discriminative power) 高的題，組成特殊的量表，應該也沒有不良影響了。

2. 取樣問題：我國修訂明尼蘇塔多相人格測驗，原在適於軍中要求；所以取樣亦以軍人為主。其所以取樣大專學生及公務人員，則在互相比較。其次，以軍中女性人員較少，而對該項測驗之需要，亦遠不及男性的迫切。故雖在大專學生及公務人員中取得少數女性樣本，亦未作進一步的分析。

MMPI 在美國原來的正常樣本，係在明尼蘇塔大學醫院，搜集陪病者的測驗成績而組成的。後來為了發展新量表，又組成一種新樣本⁽¹⁰⁾，復將原來的常模加以修正。計男性標準化樣本平均年齡33.1，標準差為10.9，教育水準為9.7（初中三至高中一年級的程度），社會經濟水準屬中等。但此項樣本能否代表美國全國成人，則似仍有問題⁽¹¹⁾。

關於效標組，其中心理異常的樣本，係請陸軍總醫院北投精神病醫院代為抽取的。樣本有限，一共僅有130人，以類別分，最少者不過十餘人。那麼，以其作為效標組，所選出的題目，自然難期十分穩定。後來對於 Paranoia 及 Schizophrenia 兩量表進行複效度研究 (cross validation) ⁽¹²⁾，發現所選之題有甚大之出入。至於所選樣本中，Paranoid schizophrenia 部分似不應與

Schizophrenia 合併為一起。

問題行爲的樣本，亦不甚多。而問題行爲的分類，係依據各單位所報出的。雖然並沒有客觀的標準，但後來於分析選定題目後，計算各量表間的內部相關，凡相關高的各量表均行合併，自其中選出一個作為代表。那麼，最先分類是否精確，便不是個重要問題了。

3. 統計的問題：根據表五我國正常樣本在美國原來各量表上的平均數及標準差，再和美國正常樣本相比較；我國正常樣本平均數均高於美國的，且有高達兩倍以上者。這顯然表示美國的常模，不適用於作為解釋我國樣本反應的依據。

根據德國的修訂⁽¹³⁾，發現德國男性正常樣本在各量表上的成績，均在標準分數50~60之間，均稍高於美國樣本的成績。Spreen 認為造成此種差異的可能原因有四：

- (1) 特殊量表的不正確：L及F量表的T分數，不能代表美國的正常人員的反應。
- (2) 樣本的差異：美國原來的正常樣本，其代表性頗有問題。
- (3) 受試者作答態度的原因：美國陪病人的樣本做測驗，比較來得防禦性大些。
- (4) 量表的效度：德國重效度考驗 (re-validation) 時，發現原量表的題目，有50%須予改變。其原因有二，一為美德兩國精神病學者所用診斷名詞不盡相同。在德國用“Schizophrenia”則限於嚴重的 Paranoid 或 Cotatonic schizophrenics，而必須住院者。而用“Psychopath”一詞則遠為廣泛，包括着若干異常的情況，如 Hyperthymic, depressive, Compulsive, sensitive, Psychasthenic, Pseudologicistic, Fanatic, Sexually abnormal Psychopath 等。二為某一測驗題對德國受試者，其意義確有差異。

日本修訂 MMPI，日本正常樣本在D量表上得分很高，顯然與美國不同。他們認為可能的原因是日本遭二次大戰的失敗，舉國陷於悲慘頹廢的境地，所以D量表上有極不尋常的表現。當然他們似乎沒有考慮到文化差異的一項因素。

從這些例子裏，都可說明，不論兩個地區的文化是否接近，要想把一個地區的人格測驗，移用到另一個文化地區，必須經過適當的修訂程序，是毫無問題的了。否則，如果直接採用另外一個地區的人格測驗，以它原來的標準來解釋測驗結果，很可能這一地區的人員都“人格異常”了。這次海軍的修訂，最初先依據美國原來的量表和標準，來檢查中國正常樣本和異常樣本的反應趨向，然後再修訂或發展適於我國的量表和標準，正是注意文化因素在人格測驗上可能產生的影響。

在這裏，尚須加以說明的，正常組的統計處理，初時，根據美國的效度量表，將有疑問的測驗卷（即L大於10，F大於15的）全部剔除。但後來懷疑美國的效度量表是否適於我國仍是一個問題（後來證明並不適於我國！），同時，發現如果將有疑問的測驗卷廢棄，不合併作統計處理，則所建立的常模含有選擇因素在內，似不能代表正常組的真正情況。所以最後在建立正常組常模時，除了未作完的測驗卷不計算外，餘者則全部計算在內。

正常組雖然在統計分析時，依據軍人、學生及公務人員，分別處理，同時並發現他們之間，在大多數量表上都有顯著的差異（參看表五及表十四）。但最後仍然合併起來。所根據的理由：每種樣本由於缺乏其母體（Population）的詳盡資料，因此，取樣不論多少都不是一個真正的樣本。但他們都是我們社會上所承認的正常人員。將若干不同的這類樣本合併起來，那麼，對於我國正常人員

的代表性，也許更大些。美國原來的正常組，所包括的年齡範圍，從16歲到65歲，而年齡組間的反應差異也十分顯著；他們最後則採取了各年齡組的平均數作為常模。這樣，如果一位受試者的測驗成績與常模有了極端的差異，例如，受試者在某一量表上的成績在平均數以上或以下兩個標準差的位置，便應該表示不同尋常了。

4. 效度量表的修訂：效度量表共含有四個分數，即L、F、K、 γ 。“ γ ”原來並不是一個真正的量表，它受著主試者實施測驗態度的嚴重影響。而K量表是由效標決定的，和L及F兩個量表的來源有異。這次的修訂，只限於L及F兩個量表。在分析過程中，發現中美兩國在L及F量表上有顯著的差異，也就是說，原來的這兩個量表不能直接移用於我國。根據原來的選題標準，為L量表選出的題，和美國原L量表相同者僅有八題，佔百分之五十。足證中美兩國對於所謂“優良”的標準並不一致。

就F量表說，原有64題，依據我國樣本的反應，而合於“等於”或“低於”百分之十的標準的，僅有22題。易言之，約有三分之二的題不適用於我國。這就是說，在美國認為“不良”的行為，在我國未必有相同的意義。依據原來的選題準則，僅能選取28題。也許由於題數太少而所含因素太多，使我國正常人員在這個量表上的測驗結果， $M=1.94$ ， $\sigma=2.43$ ，沒有達到理想的境界。如以原來的標準，平均數以上三個標準差作為異常的標準，則原始分數在9~10分之間。因此，我國受試者在這個量表上，其原始分數應訂為不得超出10分。

5. 臨床量表的修訂：對於臨床量表，並不是這次海軍修訂的主要目標。最初的構想，是在這次修訂過程中，選出二、三個量表，依據我國的異常樣本予以“重效度考驗”看看結果如何，是否須對全部臨床量表加以修訂。不幸，重效度考驗的結果，發現 Hysteria, Paranoia 及 Schizophrenia 三個臨床量表，中美兩國間有出奇的差異。兩國間相同的題，在 Hysteria 上有八題，在 Paranoia 上有二題，在 Schizophrenia 上有六題。僅佔各該量表題數的六分之一，至十分之一。

由此可知，原來的九個臨床量表，要想應用於我國，均須修訂。紙筆式的人格測驗 (Paper and pencil personality inventory) 是提供若干問題情境，誘發受試者的反應；決定反應的因素，不是個人的能力，而是影響個人態度的經驗和文化背景；自然不可直接採用外國的標準了。

6. 新發展的量表：對問題行為的鑒別，這是海軍修訂的重點之一。自各有關單位選取問題行為樣本十四類，復自監所中另選取三類。於各量表完成之後，由各量表間的內部相關的特性，歸納為八種問題行為的量表如下：想家、工作態度、孤獨、憂鬱、暈船、多病、反社會及強暴。工作態度包含着由各單位提出的想退役、工作效率低、情緒極端起伏、心理緊張、難於管教、複補、身體多病及神經衰弱九類；反社會則包括偷竊及搶奪二類。

各量表間重疊的題（參看表十三），最高的有達30題者。各量表獨有的題，最多的有19題，如“多病”；最少者為三題如“暈船”。由於量表間的重疊而產生的內部相關，最高的有達0.28者。所以各量表的內部相關，雖然多在0.60左右，但事實上均應稍低於各該統計數字。

7. 標準化樣本的平均數及標準差：標準化樣本原有2,389人，為了運用 Computer 計算，節省用費，僅選出675人的測驗卷作為統計的依據。這675人係隨機取樣選出的，每隔三份或四份測驗卷

抽取一份。根據這675人的平均成績及標準差，轉換T分數。至於所製側面圖，在效度量表中，L及F則採用新修訂的量表，K量表及“？”則採用我國新常模。關於臨床量表有四個，Anxiety、Hysteria、Paranoia及Schizophrenia；不過爲了比較，Schizophrenia又分爲I及II，I爲依據學生樣本選的題，II爲依據全部樣本選的題。問題行爲則有八個，在前面已經分別說過了。

9. 側面圖組型的問題：自從因素測驗發展以後，無論能力測驗或人格測驗，對於測驗結果的解釋，已經不再重視某一個特殊分數的涵義，而逐漸重視側面圖組型的分析；MMPI自然亦不例外。MMPI側面圖組型的精細分析，有待於臨床經驗的累積，這不是一朝一夕可以完成的。

不過，在使用的初期，似可着重受試者在各量表上的反應一般趨勢。如果受試者在某類的量表上，均得高分或低分；而在其餘的量表上均在正常範圍以內，便有某種可能的趨勢；或有進一步檢查的需要。這樣的注重各量表上的相對關係，再參考其他的資料，則MMPI對受試者行爲特點或人格特質的了解，一定會有它的特殊貢獻了。

六、結論：根據修訂過程、統計結果、分析及討論、所得結論如下：

1. MMPI的中譯本，保有美國原測驗內容、形式及題數。每一題的語氣、措詞及含義，亦均力求和原來的測驗本相符合。其中僅有少數題目，由於屬於宗教等，或與我國文化背景顯然有異者，則予以修改，但仍盡量保持其原來的意義。

2. 標準化樣本，最後爲：以軍事人員爲主體有1,689人，佔71%，（這只是粗略的分類，事實上，軍事人員中，有一部分取自軍事學校，似可歸之於大專學生類中）大專學生有431人，佔11%；公務人員有253人，佔11%；共計2,389人。作爲分析“作答百分比”及修訂與發展量表的根據。

此外，尚應抽選我國女性樣本，建立我國女性的MMPI的常模。

3. 效標組共分爲兩類，一爲心理異常類，係由北投精神病院醫生代爲抽取的，合Anxiety、Hysteria、Paranoia及Schizophrenia各類患者。一爲問題行爲類，由海軍各有關單位自行選報的，分爲十四類共518人。另由軍人看守所、新竹及高雄感化院依據宣判罪行抽取344人。

4. 我國正常樣本在美國原效度量表及臨床量表上的測驗結果，均較美國原製標準爲高。就各量表的平均成績看，且有高達三倍以上者。顯示美國常模不能作爲解釋我國人員反應的依據。

5. 效度量表爲MMPI一種創發性的設計，目的在檢查受試者的作答態度；這在紙筆式的人格測驗上，是一種重要的發展。含有四個分數：

(1) ？分數：這是受試者沒有作答的總題數。在我國正常樣本上，平均未作答的題數爲7.93，標準差爲16.21。未作答的題超出30題者爲數不多。故仍可以30題，作爲標準；如未作答的題超出30題，應該深查其原因，以決定該項測驗結果是否有效。如將是項分數，視作受試者對某類問題情境持審慎態度或不肯表示態度的指標，則在中美兩國也大有不同。就未作答百分比高的十五題來看，中美正常樣本相同的僅有二題。我國正常樣本對於有關人際關係及性的問題，較持審慎態度；而美國正常的樣本，則偏重於宗教及自我知覺的問題。⁽¹⁴⁾

(2) L量表：根據測驗原編製的原則，爲我國選取的L量表，和美國的大有出入。所選取的十六題中僅有半數和美國原L量表相同。我國正常樣本在該新L量表上的平均數爲7.09，標準差爲

3.40。如以T分數70為標準，則在該新L量表上的原始分數，如超出14，便應視為無效。

(3) F量表：這個量表的功能有三：①檢查受試者作答是否細心或對題意是否了解，②鑒別異於正常人的性格及觀念，③受試者是否將自己往“壞處”想。這並不是“pure”測驗。因此，建立一個適當的新量表，便比較困難。經過最後的分析，僅選出二十八個題作為我國的F量表，平均數為1.94，標準差為2.43。受試者原始分數超出10分，便應視為測驗結果無效，或應參考臨床量表的結果，予以推論。

(4) K量表：這個量表和L及F量表的來源不同，是由效標決定的。效度標準建立困難，我國只能暫時建立新常模。好在這個量表的作用已經簡化；過去用它作為校正某些臨床量表的功能，已經廢棄了。

6. 臨床量表原來有九個。這次修訂，受了取樣的限制，只是對於 Hysteria、Paranoia 及 Schizophrenia 重予效度考驗，藉以檢查原來的臨床量表是否可以直接應用於我國。結果發現這三個量表，在中美兩國間有極顯著的差異。兩國在這三個臨床量表上相同的題目，僅有六分之一至十分之一。因此，證明其餘的六個臨床量表，均須修訂。

7. 新發展的量表：計有八個為：想家、工作態度、孤獨、憂鬱、多病、反社會及強暴。預期此新量表對問題行為之防止及海軍士氣之建立，可具有適當的作用。

8. 各量表間的重疊題目有高達30題者，各量表獨有的題最多的有19題，最少的有三題者。由於量表間题目的重疊，所增高的內部相關係數最高有達0.28者。

9. 標準化樣本的平均數及標準差之計算，則依據由原樣本2,389人中，所抽取之675人的測驗成績。

10. 測驗結果的處理：MMPI 修訂本的測驗結果，亦不應過分重視單獨量表的特殊分數。雖然側面圖的分析，須待累積臨床經驗的資料。但在初期使用時，應重視受試者在各量表上的一般趨勢。凡在各量表上的成績起伏過分劇烈，或在有關的某些量表上得分特別高，而在另一些有關量表上得分低，均應特別重視。易言之，應以各量表間的相對關係作為判斷或解釋的根據。在各量表上得高分，固應重視，而得分過低也不容予以忽視。

關於測驗的實施、記分及繪製側面圖等，一如美國原來測驗的程序，在這裏不再贅述。

11. 測驗結果的運用：MMPI 的修訂本，可以提供四種資料：①屬於效度量表的，檢查受試者的作答態度，同時，配合臨床量表作為診斷之參考。②心理異常的傾向(量表1~5)，例如抑鬱、癡慮或自我中心。③工作情緒或工作環境的適應(量表6~11)。④問題行為的傾向(量表12,13)。總之，由於在各量表上或在各類量表上的成績之相對關係，鑒別受試者的情緒穩定的程度或情緒適應的特點，藉以發現該受試者應否需要予以特別的照顧，或作進一步的檢查。

12. 紙筆式的人格測驗的作答方式，多分為“是”或“否”，而此種“是”或“否”全屬主觀的。測驗題本身在本質上並無“是”或“否”的絕對標準。因此，其信度及效度自難期達到能力測驗的水準。根據所作的一項信度研究，兩次測驗相距三個月，其穩定係數為0.57~0.82。

對測驗結果的解釋和運用，有如使用其他各種診斷工具，經驗是一項重要因素。具有充分的心理學或精神病學的知識，加上實際的不斷的印證，則每可充分發揮這種測驗應有的性能。

註 釋

- (1) 見路君約，修訂 MMPI 的初步報告，測驗年刊第十二輯。
- (2) 見包重齡，MMPI 是什麼？測驗年刊第四輯。
- (3) 同註一
- (4) 見 Welsh and Dahlstrom, *Basic Readings on the MMPI in Psychology and Medicine*, University of Minnesota Press, 1956, p. 56.
- (5) 詳見我國海軍刊行之 MMPI 修訂報告。
- (6) 見 Dahlstrom and Welsh, *An MMPI Handbook*, 1960, p. 48.
- (7) 此項分析係由海軍各單位主官或輔導人員所作之判斷。
- (8) 此項樣本分別取自海軍及陸軍看守所，高雄及新竹感化院，以其宣判之罪行為根據。
- (9) 見 Cronbach, *Essentials of Psychological Testing*. 2nd. ed., Harper, 1960, p. 468.
- (10) 見 Hathaway and Briggs, Some normative data on new MMPI scales. *J. of clin. Psychol.* vol. XIII, No. 4. 364-368. 1954
- (11) 見 Spreen, O. and Spreen, G., The MMPI in a German speaking population. *Acta Psychologica* 21 (1963) 265-273. North Holland Publishing Co.
- (12) 見王明美及李美耶，對 MMPI Pa 及 Sc 兩量表的複效度研究。臺大心理學系，畢業論文，未刊布。
- (13) 同註十一。
- (14) 見路君約，對 MMPI 效度量表的研究，長期科學發展委員會報告。

A CHINESE REVISION OF MMPI FOR THE CHINESE PEOLPE

CHUN-YO LU

ABSTRACT

The work to revise the MMPI for the Chinese people, beginning in 1959, was completed in 1966. The main purpose of this revision was to make it a useful means of screening and selecting personnel for Chinese Navy, as well as a diagnostic aid for Chinese civil psychiatric cases.

Test records of 2,420 normal cases and criterion groups including 130 pathological cases, 331 prisoners and 518 problem behavior cases have been carefully analysed. The means and standard deviations were computed; and for both the normal group and the criterion groups, the frequency of responses for each of the 550 items was determined. All these data has been compared with the American norms provided by Dr. Hathaway and others; and significant differences between the two cultures have been found.

Based on this research, 2 new validity scales, L and F, and 12 new clinical scales, namely Anxiety, Hysteric, Paranoid, Schizophrenic, Homesich, Isolation, Depression, Work attitude, Sea-sickness, Sickliness, Anti-social and Violence, have been developed.