

汉字键盘输入人的认知模型

张 倪

(中国科学院心理所)

陈一凡

(北京信息工程学院)

【摘要】 本文从汉字键盘输入方法评测的重要意义和目前遇到的困难出发,说明以工程心理学的方法建立的汉字键盘输入认知模型是设计完整的评测系统和优化输入方案的有效手段。

一、DVORAK 键盘引起的思索

汉字键盘输入技术是在我国推广计算机应用的关键技术之一。十多年来,通过科技工作者不懈的努力,该项技术已经进入实用阶段。在已经提出的五百多种汉字编码方案中有五十多种已经上机运行,其中有一定知名度的约十余种[1,2]。尽管有些系统的输入速度已超过西文输入,但是“易学易用、高效、符合汉语规范”的输入系统还有待作深层的研究。因此新的方案和系统仍不断地出现。

面对如此众多的输入方法,用户面临选用哪一种最适合自己的工作需要,国家主管部门更有一个判定各个方案的优劣问题。在解决这个问题前,我们应该回顾目前为各国所采用的通用小键盘对推广计算机应用所起的作用。通用小键盘的英文字母排列沿用100多年前的英文机械打字机方式,这种键盘称 DVORAK 键盘或 OWERTY 键盘。作为计算输入设备增加、调整功能键后形成目前的 ASCII 键盘。这种键位分布似乎已纳入西方文化的范畴,为几代人所习惯使用,从而成为工业标准。当有人从人一机工程学的角度发现这种键盘的不科学性而设计出合理的键位分布时,用户不愿改变从小就学会了的指法,厂商也不会冒险修改标准。这样,全世界使用英语的人只好继续在目前这种通用方式上浪费精力。

汉语不是拼音文字,不能像英语那样,一个字母一个字母的对应键入。在用小键盘键入时,必须用脑力进行字、词的特征信息编码,然后击打与码元对应的键位。这就使得汉语输入计算机的技术,远比输入英语复杂;选择合适方案或最佳方案的决策过程也更复杂。如果我们不能在目前汉语输入方法还在发展的过程中就对各种输入系统的素质和性能进行科学测试,为用户提供选择依据,为汉字输入方式的进一步发展指明方向,一旦等到某种并非最好的方法为广大用户和生产厂家所接受,我们也要面临像英语键盘一样的历史性的失误。这显

本文1991年7月1日收到

然是任何使用汉语的人都不愿意见到的结果。

为了对民族的长远利益负责，我国汉字键盘输入方法评测界自七十年代末以来做了大量工作，与输入方法发展同步，不断研究测试技术和评测规则，并组织了多次全国性的和个案性的评测活动[3,4,5]，为评价各种方案的优劣提供了一定的科学依据。但是，迄今为止汉字键盘的输入评测技术还很不完善，很多难点有待于解决。

二、统选操作员心理素质对评测结果的干扰

人通过键盘将汉字键入计算机。因此，使用统选操作员在同样控制条件下对不同输入方法进行学习后，再依他们的操作水平来评价各个汉字输入系统的优劣很自然地成为首先考虑的评测方法[3, 5]。这种评测方法的最大优点是符合实际使用条件。汉字键盘输入实际上是由人和计算机组成的人—机系统来完成的。现代人—机系统的理论认为，人是系统的核心，只有系统的操作过程符合人的信息加工特点，才可能获得高效率的人—机系统。从这个意义上讲，以统选操作员的学习过程和成绩来评价汉字输入方法应该具有最大的逼真性。然而，从另一个角度，正是这种人的因素的介入，又给评测工作带来了极大的困难。困难的集中点在于，使用统选操作员进行评价的一个必要的前提假设是：分配给各种不同输入方法的统选操作员的各项素质必须是一致的。由此产生了以下三个方面的问题。

(一) “一致的”三个字可以包含无穷维的参数。目前，最常考虑到的是：

1. 年龄、性别、一般文化水平
2. 对英文键盘的熟悉程度
3. 对汉字输入技术的经验
4. 手指的灵活性
5. 学习能力
6. 一般认识能力，如：记忆、反应灵敏性等

以上所列举的六个项目只是一般概念中“一致的”所包含的极小部分的内容。然而，即使依这些项目选择统选操作员也要花费大量的人力、物力。全国性评测工作每举行一次需要几十万元经费，其中相当大的一部分是花费在统选操作员选择上。

与此相关的一个问题是，这种用统选操作员的评测方法是一种相对比较的方法，只能在有若干测试方法同时参加评测的场合应用。对个别的新的输入方法或是改进后的系统，难以进行个案性的评测。

(二) 严格地说，“一致的”所包含的内容远非以上几个项目。实际生活中，没有二个完全一样的人。为了某一特定的目的，往往把“一致的”规定在某些指标范围之内。而为了尽可能逼近理论上的“一致”，又必然导致这些指标数目的增加。为了评价某种操作的目的，一般需要考虑控制以下一些因素^[6]：

1. 认知能力
一般智力能力；语言能力；数学能力；推理能力；知觉速度；记忆能力；空间和机械能力。
2. 动作能力
控制精度；四肢协同运动能力；反应朝向能力；反应时；手臂运动速度；节奏控制；手

的灵巧性；手指的灵活性；手臂和手的稳定性；腕指快速运动；手动准确性。

3. 与当前任务有关的工作经验
4. 学习的能力
5. 情绪状态及其对操作的影响
6. 动机因素
7. 小组内人员之间的关系对操作的影响。

以上所列举的仅仅是应考虑到并应该加以控制的因素，还不是具体的测试项目。往往每一因素要由多项测试来决定。由此可见，选择统选操作员是一项很复杂的工作。在时间和经费受到限制的条件下，很难达到选出同样水平的统选操作员的目标。特别是与情绪和动机有关的因素更是难以控制。因此，要对每一参加评测的输入系统都配给上述各项都一致的统选操作员，实际上是不可能的。这就容易使人对评测结果的公正性、科学性产生疑虑，进而影响评测结论的权威性。

(三) 退一步考虑，即使仅以目前已被考虑到的几项测试内容来选择统选操作员，也存在许多问题。早在1987年，中华杯汉字输入竞赛的组织者已经考虑到统选操作员心理选择的重要性，委托中科院心理所和北京信息工程学院对待选的204人做了五个方面的心理测试。并以测试结果作为选择和分配统选操作员的参考。当时所测的项目和统计数据见表-1，各指标间的相关分析见表-2。

表-1 204名统选操作员心理素质测试数据统计表

测试项目	平均值	标准差	最大值	最小值
反应速度	27.43	3.44	32	13
推算速度	43.73	7.13	75.68	0
联想反应速度	60.97	8.71	96.34	0
联想学习进步率	41.34	35.37	128.19	-77.44
认知能力	98.98	59.04	301.72	-17.41

* 表内数值均为相对值

表-2 心理素质指标的相关分析

	反应速度	推算速度	联想学习进步率	联想反应速度
推算速度	-0.197			
联想学习进步率	0.257	-0.265		
联想反应速度	-0.145	0.175	0.124	
认知能力	0.118	-0.046	0.208	-0.143

测试结果表明，这些指标之间的关系非常复杂，难以用简单加权的方法确定个体的心理素质。五项指标中，除速度指标外，其它两项的标准差都很大。联想学习进步率的平均数为

41.34, 标准差为35.37。认知能力的平均数为98.98, 标准差为59.04。这些数据表明受试者之间个体差异很大。所选用的测量指标在没有经过效度和信度检验的情况下很可能难以量出可靠的结果。五项指标之间, 没有任何两项之间的线性相关数在0.3以上。特别是反应速度和推算速度之间是负相关, 即反应越快的, 推算速度越慢。反应速度和联想反应速度之间也呈负相关。以选入的140人和未选入的64人比较, 唯有联想反应速度的差异达到了统计学显著性水平($P < 0.001$)。其它四项指标均不能作为确定受试者能否被选为统选操作员的指标。被选入的人员, 最后只有以随机方式归为五类。再将五类人员平均地分配给各组, 以达到各组之间“水平一致”的目标。尽管如此, 中华杯组织还是做了有益的探索, 暴露出心理选拔是一个非常困难的任务。

综上所述, 在实际进行的大规模评测工作中, 不可能选出所谓心理素质一致的统选操作员; 统选操作员分配到组后, 仍有许多难以控制的情绪、动机和小组环境因素。因此, 使用统选操作员的评测方法难以保证评测的客观性、公正性和科学性。加上以统选操作员的操作作为评测指标的方法不能用于个案性的评测工作。这就使得汉字键盘输入评测不得不考虑其它更可靠、客观和实用的评测手段。

三、易学性速度素质和心理负荷测量

统选操作员所带来的困难都是由于人的因素所带来的。为了获得更加可靠、客观、实用的评测手段, 一个很自然的想法是, 排除人在评测指标中的干扰, 用完全客观的指标来评测汉字输入方案的优劣。

为此, 汉字键盘输入评测界近年来做了很多严密的工作, 建立了一些以计算机程序化为基础的评测工具。以求解决在排除人的因素的前提下, 对汉字输入方案进行客观评测的问题。如“七五”期间完成的科研项目“键位分布合理指数与动态平均码长综合指标的自动测定”^[7, 8]。这种评测法的最大优点在于其客观性、系统性、绝对性和经济性。

客观性, 用这种方法对汉字输入系统进行测定时, 无需选择特别的输入人员, 只要由任何了解被测方案输入规则的人, 将从标准题库中^[9]选出的测试文本按该系统的输入规则(包括使用联想等策略)键入计算机, 测试系统即可给出键位分布合理指数、动态平均码长和速度码长为该系统速度素质的测试结果。也可将输入规则置于测试系统后用“模拟操作员”即由测试系统扫描码本和标准题库的方式来得到上述速度素质参数^[10]。

系统性, 该系统所给出的动态平均码长值不依赖特定的被测文本, 而是结合以大量统计工作为依据的标准题库, 求得被测方案在一般情况下的码长。键位分布合理指数也是依据于工程心理学数据所获得的结果。因此, 所得评测结果具有很高的稳定性。

绝对性, 因为无需选择相对一致的统选操作员, 使得该系统可以在任意时间, 任意地点对任意输入方法进行个案评测。不同时间、不同地点所做的输入方法的评测结果之间可以直接比较。从这个意义上讲, 这种自动测试系统是一种绝对测量方法, 而不是象使用统选操作员那样的相对测量方法。

经济性, 显而易见, 使用这种自动测量方法可以节约大量的人力、物力。使经常性的评测成为可能。

上述自动评测方法的最大优点在于排除了人为因素的干扰, 然而我们应看到上述测试系

统由于未能解决汉字输入过程中人的因素的测试而使测试结果具有局限性。

如前所述，汉字键盘输入技术是一门相当复杂的学科，涉及到包括汉字语言文字学、计算机科学、工程心理学在内的多门学科的内容。这就决定了汉字键盘输入方法的评测也是一项复杂的工作。目前评测界所面临的重大问题自然是如何以人一机系统的角度来更有效的评测汉字输入系统的三个主要指标的问题。具体可以归于以下三个主要方面：

(一) 易学性的测量。易学性是就人而言的。即接受某一输入方法的人需要多少训练方可达到一定的操作水平，在间隔相当长的时间不使用该输入方案后，需要多少训练方可恢复到原有的操作水平。需要建立“训练量”的计量体系。无论对专职操作员和非专职操作人员，训练量与学会使用后实际操作时的心理负荷是相关的。

(二) 速度素质的精确的测量。从人一机系统的角度看，汉字键盘输入包括人对输入文本的字、词进行特征信息提取、编码、按提示选择和击键两个过程。速度码长不包含前一个过程，因此它只能表示输入系统速度素质的一个重要指标，而不能代表速度素质的全部内容。目前，我们尚不明确各种编码规则对人的编码速度的影响。但是，可以肯定地说，速度码长相等的输入方法并不保证人的编码速度也是一致的。为了获得速度素质的更精确测量，有必要进一步考虑影响人的编码速度的因素。

(三) 心理负荷即易用性的测量。随着汉字键盘输入方式越来越多地应用于人一机系统，操作人员的一个普遍反映是，输入汉字比输入英文要累。究竟累到什么程度？不同方法引起累的差别怎样？现代认知心理学提示我们，人的心理负荷不一定和动作的次数呈线性关系。在绝大多数情况下，操作者宁可以多次击一个键的简单动作来完成一定的作业，而不愿在多个键中选择一个，击一次键来完成同样的作业。对于汉字键盘输入，类似这种关系的原則究竟应该怎样？所有这些问题的解决，都有待于对输入方法的心理负荷的测量^[11]。

四、汉字键盘输入的认知模型

汉字键盘输入是在人一机界面完成的一种操作。如果我们能够按照人的视听感觉、认知和动作对整个过程进行分解和说明，我们就可以获得一个汉字键盘输入的认知模型。这个模型的最初形式还仅仅是一种解释性模型。再以工程心理学的实验方法测出模型中各个参数和变量间的函数关系，我们就可以获得具有定量说明能力的汉字键盘输入的数学模型。数学模型的计算机程序化将为我们提供可用于实际评测的测试软件。该软件可在汉字输入评测界目前已经获得的成果的基础上，进一步包括人的认知和决策因素在输入过程中的作用，对汉字输入系统作出更精确评测。并为设计和改进输入系统进行技术导向。有必要指出的是，我们这里所说的人是广义的人或称之为人的群体特性，是完全不同于使用统选操作员方式中的具体的人。具体的人所固有的人的因素会给评测工作带来干扰和困惑。而对广义人的认知和决策过程的了解，则可以使我们将人的因素包括在评价系统之内，从而更精确地评测各个汉字键盘输入系统。例如，在现实生活中，如想找出两个简单反应时间完全一样的人是很困难的。但是，如果要知道是听觉反应时快，还是视觉反应时快，却是很简单的事情。实验心理学早已获得公认的结果，即：听觉反应时约 130 毫秒，而视觉反应时为 170 毫秒^[12]。130 毫秒和 170 毫秒均对人的群体而言，即对广义而言的统计数据。与此相类似的是，我们很难说哪两个人的记忆能力一样，但我们已明确地知道，人的工作记忆容量是 7 ± 2 个单元。象

这样可以用于建立描述人的认知过程模型的心理学的研究成果还有很多。同时，用于解释英文输入方式的认知模型也可以作为我们建立汉字键盘输入的认知模型的借鉴^[13]。我们的任务是，合理地将这些知识应用到汉字输入方案的评测工作中去。

图-1是一个汉字键盘输入的认知模型的框图。

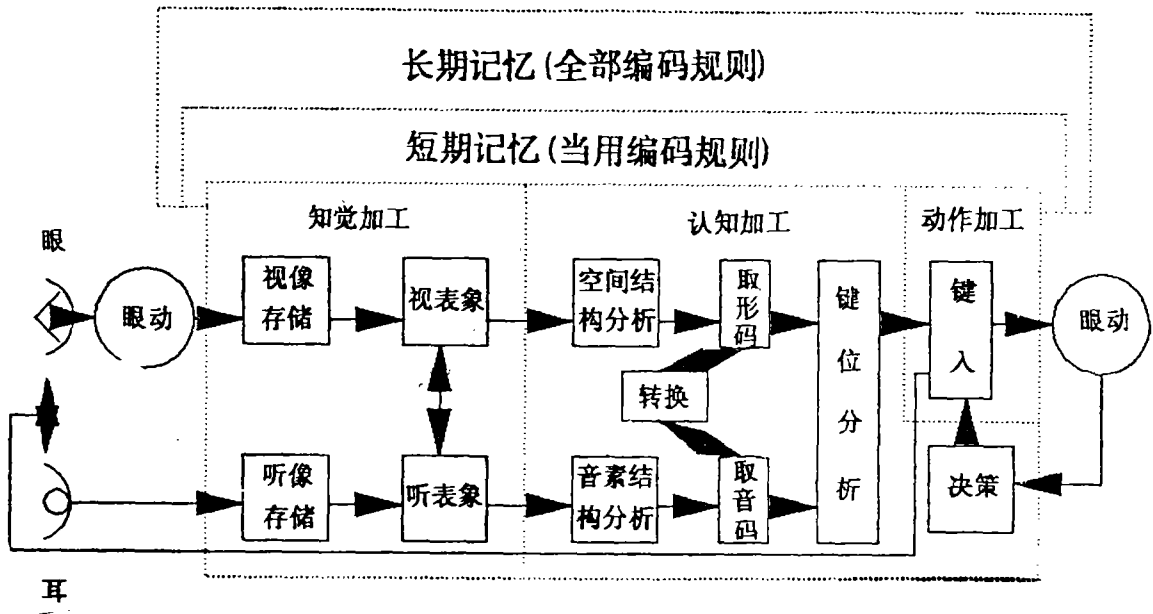


图-1 汉字键盘输入的认知模型框图

这个框图是以信息加工的阶段为基础。基本的原则是：每一信息加工阶段都要耗费时间和资源；而总的资源量受到人的认知特征的局限；这样，操作者在大多数情况下是一个串行加工者。图中的大多数阶段的各项参数和信息加工特征已为工程心理学所发现和确定。结合一套操作原则，就可使该模型有很强的可操作性。这里，将我们已经明确的操作原则简列如下：

1. 认知加工器的活动周期律
2. 知觉加工器的活动周期律
3. 编码特性律
4. 认知加工器活动周期律
5. 费茨律
6. 练习指数律
7. 分辨律
8. 不确定律
9. 目标律
10. 问题空间律

模型的三个不同层次的加工过程和容量限制将分别为评测汉字输入方法的三个主要素质提供理论依据：即，长期记忆量与易学性；短期记忆量与心理负荷和易用性；认知、动作加工与易用性和输入速度。

针对汉语与英语的差别，我国认知心理学界已经完成了一些基础理论性的工作。如，我

们已经知道, 中国人对以印刷形式呈现的汉字的形的加工时间为 857 毫秒, 而对音的加工为 1075 毫秒^[11]。框图中的另一些项目的参数, 如不同类形编码的解码时间, 还有待进一步实验室研究来确定。

参 考 文 献

- [1] 苏东庄、袁琦: 中文信息处理在中国的发展, 1990年中文与东方语言计算机处理国际会议专题报告, 1990
- [2] 陈一凡: 汉字键盘输入技术的现状与发展方向, 国际电子报, 1991年第12期
- [3] 全国汉字键盘输入评测办公室: 汉字键盘输入方法评测规则, 1990
- [4] 胡宣华、谢克中等: 汉字编码输入技术评估, 1990年中文与东方语言计算机处理国际会议专题报告, 1990
- [5] 台湾资讯工业策进会: 中文电脑输入法与输入器调查评估总结报告, 1984
- [6] Gavriel Salvendy: Handbook of Human Factors
- [7] 陈一凡, 张鹿, 周志农: 键位相关速度当量的研究, 中文信息学报, 1990, 4
- [8] 陈一凡, 张鹿: 键位分布合理指数与动态平均码长综合指标的自动测定, 中文信息学报, 1991, 1
- [9] 何厚存、胡卫明等: 汉字键盘输入操作方法评测用试题库的研究与建立, 中文信息学会汉字编码专业委员会年会论文, 1990
- [10] 何厚存、金向阳: 汉字键盘输入操作员的模拟, 中文信息学会汉字编码专业委员会年会论文, 1990
- [11] Christopher Wickens: Engineering Psychology and Human Performance, 1984
- [12] Woodworth R.S. and Schlossberg, H. Experimental Psychology, 1956
- [13] Card S.K., Moran T.P. and Newell A.: The Psychology of Human-Computer Interaction, 1983
- [14] 张武用等, 汉字匹配的偏侧化效应和事件相关脑电位, 心理学报, 1988, 4

A Cognitive Model of Typing Chinese into Computers

Zhang Kan

Chen Yi Fan

(Institute of Psychology, CAS) (Beijing Information Technology Ins.)

Abstract

With discussions of importance of evaluating approaches of typing Chinese into computers and problems in such evaluations, a cognitive model of typing Chinese into computers is suggested and the potential significance of the model in such evaluations is discussed.