

短时记忆容量的重新思考¹⁾

崔剑霞 吴艳红

刘艳芳

北京大学心理学系 北京 100870

摩托罗拉实验室 北京 100085

摘 要 通过研究发现,短时记忆 (STM) 的容量大小为 7 个单位。但在随后的 30 多年,该理论受到了质疑。本文通过总结大量的研究,提出了纯净的短时记忆容量和混合的短时记忆容量概念,并认为以前的研究结果代表的仅仅是混合的短时记忆容量。

关键词 纯净的短时记忆容量? 混合的短时记忆容量? 块? 注意焦点? 时间限制
中图分类号 B77

记忆是人类高级心理功能,是心理学研究中最活跃的领域之一。记忆过程包括编码、存储、提取三个过程。其中,特别是短时记忆存储容量限制的研究一直受到记忆研究者的普遍关注。对短时记忆存储容量的研究是进行人机界面研究的基础;只有更准确地了解短时记忆存储容量后,才能设计出恰当的人机界面系统,以保证界面布局合理、操作方便。提供的信息数量既不会多得让人一时无所适从,又可以充分利用人的注意和记忆资源,实现设计的最优方案。

1 短时记忆的容量

通过总结大量的对线性刺激的绝对判断、速度-准确性、速知、以及即时回忆广度等实验研究,发现被试的感觉通道容量或者回忆项目的数量,也就是记忆的容量在一个很小范围内波动,大概是 7 个。但需要注意在不同的实验中,7 所代表的项目是不同的,因此,使用了 Dm 这个单位。一个 Dm 的信息是指在 7 个几乎相同的项目之间进行判断时所需要的信息量。例如要求被试区分 7 个不同频率的纯音信号,显然区分 7 个纯音信号时,每个项目包含的 Dm 信息较少;而对于单词的区分,每个项目所包含的 Dm 信息就要多很多了。即使两种情况下被试记住的都是 7 个项目。认为再编码对于每个项目中包含 Dm 信息的数目的增加具有重要意义。虽然不同任务的机制可能不同,但却得到相同的容量范围 7;说明短时记忆接受、加工、存储的容量限制为 7 个块。但没有对块做出明确的定义。的理论提出后,长时间被人们广泛接受,而且写入教科书^[1]。

1) 摩托罗拉实验室、国家自然科学基金 10071、国家攀登计划项目 234 专 4 2 资助项目

通讯联系人: 526 784 7048, 37283 94#16; /; < =* - 2 - /z

收稿日期 2004 年 4 月 27 日; 修回日期 2004 年 8 月 1 日

2 C 理论框架

2.1 理论框架

A. H. 认为 Atkinson 的理论提出了 20 多年,但现在仍不清楚存储容量限制的性质,甚至不知道是否真的存在这样一个容量,因此不但要研究短时记忆 (Miller, 1956) 容量的大小,更重要的是研究它是否存在。因此 A. H. 于 1971 年总结了自己和他人的研究,重新考察了短时记忆的存储容量,提出了自己的理论。

先说明一下注意焦点的含义。注意焦点的概念由冯特在他 1907 年出版的《情感和注意心理学》中提出。冯特认为“注意是伴随着一种心理内容的清晰领会的状态”,并指出注意的范围和作用“意识也与视野一样,是一个以一定的阈限为境界的有限领域,任何的心理内容只有进入这个领域,才有被领会的可能。在该领域内有一个范围狭小的中心区域,任何心理内容只有进入这个中心区域,才会获得最大的清晰性和鲜明性。这个中心区域被称为‘注意焦点’。”^[1]

A. H. 的理论共有以下 4 条基本假设:第一,“注意焦点是容量限制的”;第二,“正常成年人注意焦点限制的平均值为 7 个块”;第三,“不存在其他心理功能的容量限制”;第四,“任何有意回忆的信息,不论来自当前刺激还是长时记忆 (Jensen, 1973) 都受到注意焦点的限制。第四条假设的前提是“只有在注意焦点中的信息才能够被意识到和报告出来”^[1]。

2.2 纯净的/混合的 容量

A. H. 的理论基础有一个非常重要的前提,就是他对“容量”的定义。A. H. 认为虽然很多人承认短时记忆的存储容量是有限的,而且可以用“块”来估计它的容量,但有的时候“块”并不容易鉴别。因为人们可能利用已有的知识经验将单个项目或者较小的“块”组成较大的“块”。除此之外还存在一些受时间限制的过程,它们不是容量限制的,但可能与容量限制的存储过程混合在一起共同回忆信息,这种情况发生时,无法评估其他非容量限制机制的作用大小,所以 A. H. 把此时的估计结果称为“混合的短时记忆容量估计”,而把不受这些因素影响的称为“纯净的短时记忆容量估计”。如果能够找到一些方法,从中得出纯净的短时记忆容量限制,将十分有利于研究的深入。^[1]

刺激呈现后, LTM 中就会有相应的记忆痕迹,而这些痕迹保持激活的时间是有限的,如果在这有限的时间内对痕迹进行复述等加工,那么这些痕迹保持激活的时间就会延长,从而被记住并在提取过程中表现出来,否则这些痕迹就会消退。因此 Bower & Tulving 等认为,“在没有复述的情况下,‘短时记忆中信息保持激活状态的时间’称为‘时间限制’。它是一种无需注意而在短时间内保持激活的存储机制。”^[1:31]

A. H. 理论的提出具有重要意义,他从 3 个角度细化了 LTM 容量的概念。一方面, A. H. 将短时记忆容量概念细分为纯净的和混合的 2 类,这是以前的研究者不曾涉及到的;另一方面, A. H. 强调了 LTM 容量与注意焦点之间的关系,也就是明确指出,研究的 LTM 容量针对被试能够报告出来的记忆,而不包括那些也许存在但不能被报告出来的记忆,因为有了更为详细和准确的概念,因此对 LTM 容量的研究也会更明确和清晰。

2. 能观察到纯净存储容量限制的 3 种情况

为了能够根据提出的理论研究纯净的 LTM 容量, A. H. 深入研究了哪些实验情景可以观察到纯净的 LTM 容量,大致可分为 3 种情况。

第一：信息超载限制新组块的形成。呈现刺激时使加工系统信息超载。由于信息过多，被试为了尽可能多的获得信息，因而不可能在刺激呈现的很短时间内对刺激进行复述或编码。既然在刺激呈现的时候，被试没有对信息进行再加工，那么刺激呈现之后，所存储的信息仍就保持刺激呈现时的状态。

第二：采取措施阻止对刺激项目再编码以形成较大的块。例如：在每个组块中给被试呈现相同的几个刺激，但要求被试回忆刺激呈现的序列位置，这样被试设法记住所有刺激呈现顺序的能力就受到了限制。另外，在刺激呈现的过程中，通过要求被试不断重复一个单词，也可以阻止被试对所呈现刺激进行复述。

第三：检测由容量限制引起的操作的不连续性。实验中，如果被试仅仅依靠容量限制的存储就能完成当前的任务，对这个任务的操作水平就比较高；但如果不能单独依靠容量限制的存储完成任务，而是还需要长时记忆和时间限制的存储的支持，或者说操作过程中卷入了更多的存储机制，此时被试的操作水平就会下降，即准确率降低、反应时变长。以一次呈现刺激的项目数量为横坐标，反应时或正确率为纵坐标，可以发现在容量限制处，操作曲线的斜率会发生较大的变化，也就是发生了操作的不连续性。

第四：检测由容量限制引起的其他方面的效果。除了第 1 种情况，还可以观察到由容量限制引起的其他效果。例如：人们倾向于将 n 个项目组成一个块来回忆， n 个词之间插入的项目为 1 个或少于 1 个的时候，前一个词对后一个词的启动效应会强于多于 1 个项目的情况。

上面 n 种情况都基于一个假设：被试不能将项目组成更高水平的块。除此之外，A. J. Flus 还做了一些其他限制：实验中每个项目本身就是记忆中的一个块，块内要有较强的相互关系，而不同块之间的关系较弱或没有关系^[1]。

容量限制的实验证据

1 信息过载时的容量限制估计

A. J. Flus [22] 曾做过忽略语音的整体报告的研究，他让被试匹配计算机屏幕上的几幅图画的同时听一系列数字。实验组要求被试忽略声音信号，最后一个听觉数字呈现后（视觉任务突然停止）要求被试回忆数字系列；控制组被试则在做视觉任务的同时注意听觉信息。结果发现，不论声音系列长短，实验组被试成人的回忆成绩保持在 1.3 个左右，儿童成绩更低一些，还出现了个体差异。不论是成人还是儿童，实验组与控制组被试回忆数字数量的差异都在 0.8 之间，说明注意条件下可能加入了某种过程，并假定它独立于纯容量限制，且能反应形成大组块时注意的使用。^[1]

需要说明的是，这个结果与 A. J. Flus 的理论并不矛盾。A. J. Flus 假设由注意焦点的容量限制产生了 L5 容量限制，是指在提取的时候，处于注意焦点中的所有激活信息的一部分信息，注意决定哪些激活的项目处于注意焦点中，因此上述实验的实验组被试测量的正是这部分信息。

2 通过阻碍长时记忆再编码、被动存储和复述估计容量限制

在 @L, Z [23] 的 n 个再认实验中，给被试呈现一系列 n 个字母的真词或假词，如 @L, Z，真词又分为高频词、低频词 n 类。每个实验中都有 0.0~1.0 个词，分为 n 个 D%/* 呈现 6 1~3, 3~1。其中有一部分词只呈现过一次，而另一部分词是对原来呈现过项目的重复，被重复的词和重复的词之间的间隔有 2 种 6~0, 1~3, 2, 1~0~3，要求被试判断当前呈

现的项目是否在前面呈现过。测量被试的反应时间和正确率。这些实验中，由于呈现的刺激序列非常长，被试很难对项目进行复述，因此符合 A. H. 观察到纯净存储容量限制的要求。结果发现不论真词还是假词，也不管高频词还是低频词，基本都在插入项目为 1 个的时候，被试的操作曲线出现不连续现象。结果支持 A. H. 短时记忆的容量限制是 7 的理论^[1]。

M. J. (1974) 对简单项目进行了变化检测实验，以检验人们每一次可以注意到多少个项目。他首先给被试呈现一幅真实景物的原始图像，间隔很短一段后再呈现一幅改变的图像，要求被试报告图像中哪里发生了变化。第二幅图画呈现到被试报告出变化为止。在此研究中，图片材料是难于复述的，符合 A. H. 的要求。结果发现在没有变化的干扰项目中检测出变化的靶子项目所需的时间，与所呈现的项目数量成正比。对于方向变化的检测，当项目的呈现时间大于 800 ms 的时候，操作曲线的斜率与图像的变化率成正比，说明只有有限的信息能够进入视觉短时记忆，呈现时间的延长并不能增加进入短时记忆的项目的数量。另一方面，当项目间隔时间 $t_{int} = 0$ 时，视觉短时记忆容量为 3K 8 个项目，但后来的实验表明这是一种混合的容量，通过排除短时成分 (即 $t_{int} > 0$) 影响后，容量下降到了 1K 1 个项目^[1]。

具有操作不连续性时的容量限制估计

Officer (1972) 做过多客体追踪实验。实验的过程是：首先在计算机屏幕上呈现一些光点或者小的客体，在这些光点或小客体运动之前，其中有几个闪烁几下之后停止闪烁，随后所有客体开始随机运动，运动停止后要求被试报告哪些客体曾经闪烁过。Officer 的实验结果是：当追踪的客体数量从 1 个增加到 3 个的时候，被试的操作水平明显下降，也就是操作在 1K 3 个项目之间发生了不连续^[2]。

另外，Officer (和 ...) 针对注意优先机制问题进行了研究，旨在探讨同时呈现多个项目的视觉情景中，如果有另外一个项目突然出现，它是如何抓住人们的注意，从而使自己得到优先加工的。结果发现被试的注意转移后，一次可以抓住 7 个项目^[1]。

非直接效果估计容量限制

S. J. (1972) 曾给被试声音呈现 3 个系列的项目，每系列包括 3 个项目，每个项目呈现 300 ms 之后让被试做 1000 ms 分散注意任务。测验时要求被试将项目放回的正确系列和位置中。结果发现当一个项目被放置到正确系列之后，它的位置混淆最多发生在相邻的第三个位置。延迟时间不同，操作的效果也会不同，但位置混淆的范围没有变化。这一现象也许能用短时记忆容量限制为 7 的理论解释^[1]。

J. C. (1977) 设计了一种任务，让被试检验 BP 1Q 9 这样的等式，B 后第 1 个字母时 9，因此这个等式是对的，加数可以是 1、2、3。发现当加数为 K 时，实验结果都很好，但当加数 3 的时候操错结果就比较差了，也就是说有一个学习曲线的不连续性发生。这个不连续性与被试报告的策略转换有关，被试报告加数是 3 的时候相对比较困难，因此会有意学习某种策略帮助操作^[1]。

其他支持证据

1 来自其他理论或模型的推导

M. J. 等用一个最近发现的神经生理学模型——皮层网络模型来解释短时记忆的容量限制问题并证实了 A. H. 短时记忆容量是 7 的理论。该模型认为在额叶 (PFC) 和视觉

'1(" ,4fž#=#, !f% 例如 1žFž!, žž#=#, !f% /, !fžS' N5~ 区域之间存在视觉工作记忆的大脑皮层环路。外

呈现不等式的条件下，被试的某种判断任务都完成得较好？而在分散呈现不等式的条件下，被试的位置判断任务操作成绩显著下降。某种条件下，被试完成位置判断任务的过程中都插入了分散注意的其他任务，差别仅在于中间呈现只需要被试记住一个位置就能完成任务，而分散呈现则需要被试在3个位置中区分出最初的那个位置后，才能完成任务。实验结果说明记忆0个位置和记忆3个位置的操作水平有明显差异。这个结果支持短时容量为7的说法。

对 C 理论的质疑

A. J. 通过深入细致的研究，提出短时记忆容量为7的理论，而且有大量的研究结果支持他的理论，但同时也有很多人对 A. J. 的研究提出了疑问，还有一些人的研究结果与他的理论并不一致。

1 C 对块的定义问题

如果不能真正的鉴别块，就无法根据块来量化容量限制。那么，来自不同实验报告的结果之间则不具有可比性，所得到的结论的可信度就会受到威胁。A. J. 对块的定义是6块是一个概念的集合，其中的各概念之间具有很强的联系，而与同时使用的其他块间的联系很弱^[1]。这个定义没有限制块的特征、数量和数量^[7]、不具有操作性^[10]，而且有循环论证的意味^[11]。

2 C 低估了短时记忆容量的大小

有人提出疑问，提取过程的瓶颈，如果存在，以及序列关系的记忆是否会影响测量的容量大小^[10]？还有人认为存在与注意焦点以外、不能报告的信息应该算做容量内^[11]。

C 理论与已有的实验结果不符

认为 A. J. 的理论不能解释经典的序列记忆实验。因为按照 A. J. 的理论所得的结果与实际的实验结果的趋势是相反的。因此 A. J. 短时记忆容量为7的理论如要成立，他必须能解释这个矛盾^[11]。

总之，很多研究者认为，短时记忆容量限制为7的理论，其条件限制合理、结果解释适当，支持证据有力，所以相对块限制的理论，块限制的说法更可信。但这并不是说，块限制就是最终的答案。科学研究处于不断发展进步的过程中，人们总是会通过不断发现问题，把对事物的研究推进更深的层次。在 1971 年的理论观点提出近 30 年后，A. J. 等人对他进行反驳，即使此观点不完全正确，但他对原有理论的冲击力是不容置疑的，其中提出的各种问题正是得出与块理论相同结果的研究者应该考虑和解释的。同时，还应重视7这个神奇数字的价值，现在称它为混合容量限制，可为什么混合容量限制也具有相当的一致性呢？一致性的存在可能与短时记忆组块过程的机制有关系，这同样需要进一步研究。

参 考 文 献

1. 朱滢主编，实验心理学，北京：北京大学出版社，1998。

8 A : flz S` S"Čřfi J V` 9% , fi 9 @` ři fl% 5; Ž M, ů, F1 fiřřfi, ž lž fi Ž VŽ` Ž%=#Žřfi, F L; , fi45Ž!# @Ž#, /6 1 ČŽ Vfi fiřřfi Ž. Ž (lž fi Ž +Ž!Dfi%L=fiž, F 1 ==lž; Žž(ł, ž- A; % VŽ` Ž%=#Žřfi 0222` ~60` 7, K 0` 2`
 ~ @/L, ŽŽ 9- L; , fi45Ž!# N#=#fi @Ž#, /! F! U, !>(fiž> S, ž, !>(- ., " lžfi%, F ŽS=Žl#Žřřfi%=(//; , %C/6 JŽfi žlžč @Ž#, /! fiž> A, Čřfi, ž` 0223` _0 3` 6 0 0` 7K 0 0, 8

7 MŽž(ž* M 1 - R, " fiłC, fi lž Zlž>(C; fi 5; Ž 9S (fižž Ž, F @fiC/fi% S" #DŽl (lž +l(, ž- BŽ: fi l, fi% fiž> B!fiž L A Žž Ž (_ _ _ _ _` °6 0` 0

2 Ofiřřfi (L- @` %fi 9% #Žřfi +l (" fi% 5! fi / * l ž C 6 1 fiřřfi, ž fiž>) Žl / Ž=fi fi% Y! CfiłGfi, ž- A, Čřfi fi Ž) (//; , % C /` 022 _` °6` 23K 1` ~

0` Ofiřřfi (L` ., ; ž(, ž V S- @Ž /; fił (# (, F1 fiřřfi, ž fi%) fi, fi /- ., " lžfi%, F 9S=Ž!# Žřřfi%) (//; , % C /6 Z" #fiž) Žl / Ž=fi, ž fiž>) Ž fi ! #fiž Ž` 022` ~ 08` °6 70, K 7, 3

00 Sfi lž Ž. L`), (fi, ž fi% & ž Žl fił fi lž J, ž C 45 Ž!# @Ž#, /- @Ž#, /! fiž> A, Čřfi, ž` 0220` 026 11, K 1` ~

0 J, Cfiž T V- 5; : fi !> 1 ž fiřřfi Ž 5; Ž, /! F1 " fi #fiřřfi, ž-) (//; , % C /fi% MŽ lž :` 0277` 236` 3, K 3, ~

01 Mfiřřfi, ž Ž 1` U, %fiž (T` @` " lž . @- 1 SŽ" !, =; / (l, l, G fi% / /, " ři, F U, !* lž C @Ž#, /! J! #fi (6 B Ž fi Ž Ž řř fi # L Ž C! Ž Cfi fi, ž fiž> U fi lž A; " ž* fiřřfi fi, ž- BŽ: fi l, fi% fiž> B!fiž L / lž ž Ž (_ _ _ _ _` °6 012

0` M / = # fi B` T fił fi Ž V 9- R" ž fi, ž fi% SŽ" !, fi fił ž C , F L; , fi45 Ž!# @Ž#, /6 5; Ž SŽ" fi% @ Ž /; fił (# (, F @ Ž řř fi fi fi fi C Ž- B Ž: fi l, fi% fiž> B!fiž L A Žž Ž (_ _ _ _ _` °6 0` 1

03 5, > fi V- L fi fiřřfi (, F L 5 @!) , = Ž! fiřřfi (lž 1 ž fi% (@ fi / Ž Ž fi% & (B Ž fiřřfi & ž Ž! (fiž> fi Ž Sfi fi! Ž , F Y" ! Y: ž L fi ! fi C Ž J! # fi fi fi 4 , ž (6 5; Ž Afi Ž , F B! >(, ž C 1 / [" l fi, ž- B Ž: fi l, fi% fiž> B!fiž L A Žž Ž (_ _ _ _ _` °6 0` 2

08 U / žž L- 1 >> fi, ž fiž> L` Dfi fi / fi, ž D / Z" #fiž fiřřfi Sfi fi! Ž` 022 _` 1376` ~ 2K` ~ 3`

0` V Ž: fiřřfi L` A; Ž ž J- V i (/ fił fiřřfi @ Ž /; fił (# (, F L` Dfi fił ž C fiž> A, " ři ž C 6 S Ž" !, = (//; , % C / fi% 9! > ž ž Ž fi, # fi # " % fi řř fi 4 Cž (l / =) fiřřfi (- ., " lžfi%, F 9S=Ž!# Žřřfi%) (//; , % C /6 Z" #fiž) Žl / Ž=fi, ž fiž>) Ž fi ! #fiž Ž` 022` ~ 3` 6 237K 2` 3

07 L /; ž Ž l ž Ž! U \` V Ž" D Ž ž Z` U Ž Ž ž l / * @ B- A; fi l fi fiřřfi fił ž C A; " ž* (lž +l (fi% L; , ! fi 5 Ž!# @ Ž #, /6 S, fi @, lž fi fiž Y ž Ž R Ž fi fi` lž = Ž! V i # Ž ž (ł, ž X B Ž: fi l, fi% fiž> B!fiž L / lž ž Ž (_ _ _ _ _` °6 0` 0

02 L /; " D Ž fi fi 5`) lž ž (/;) 1 - Z, : & ž fił / l (fi Ž A fi = fi / fi 4 J i # fiž> 1 fiřřfi, ž fi% R, / (- B Ž: fi l, fi% fiž> B!fiž L A Žž Ž (_ _ _ _ _` °6 0` 8

~ B Ž fi # fiž A) - 5; Ž L fi C Ž fiž> Sfi fi! Ž , F 1 A; " ž* - B Ž: fi l, fi% fiž> B!fiž L A Žž Ž (_ _ _ _ _` °6 007

0 Sfi lž Ž. L` S Ž fiřřfi J- J, ž C 4 5 Ž!# @ Ž #, /! L = fiž- B Ž: fi l, fi% fiž> B!fiž L / lž ž Ž (_ _ _ _ _` °6 01`

.. Y fi > Ž (M V` . Ž # Ž % B- U; Ž! Ž fi Ž @ fi C / B Ž fi * (V, : ž 6 B, " ž fi fił Ž (fiž> fi Ž R, / (4 i fiřřfi, ž ž L /; l G, =; ! Ž ž fi fi- B Ž 4 ; fi l, fi% fiž> B!fiž L A Žž Ž (_ _ _ _ _` °6 013

.1 ., " - 5; Ž @ fi C / S" # D Ž l R, " !6 A fiž fi 9S = fił ž L fiřřfi ž D Ž C (l Ž! fi% @ Ž #, /! L / fiž V fi fi X B Ž: fi l, fi% fiž> B!fiž L A Žž / 4 Ž (_ _ _ _ _` °6 0, 8