
Title	Evaluating Chinese comprehension principles and practice (in Chinese)
Author(s)	Soh Kay Cheng
Source	<i>Seminar on "The Teaching of Reading Skills in Chinese Language", Ngee Ann Polytechnic, Singapore, 27-28 May 1988</i>

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.

EVALUATING CHINESE COMPREHENSION

PRINCIPLES AND PRACTICE

(in Chinese)

Soh Kay Cheng

Paper presented at the Seminar on

"The Teaching of Reading Skills in Chinese Language"

Ngee Ann Polytechnic, 27 - 28 May 1988

华文阅读评测的原则与方法

苏启祯

要谈阅读的评测，我们必须先了解阅读的过程，也就是阅读牵涉到那些心理现象，而在什么情况下一个读者可以说是已经理解了一段话语。

让我们先默读一段话语，看看我们理解它到什么程度：

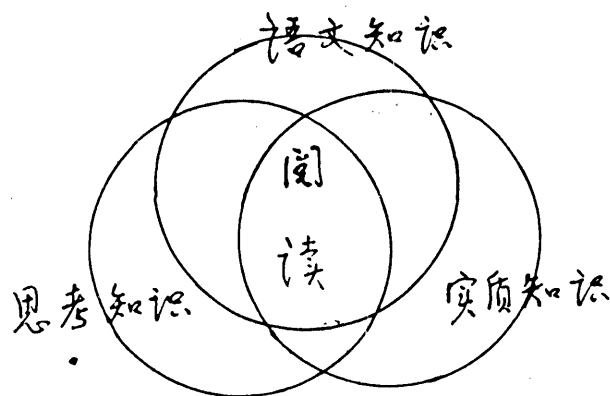
(例 1)

「他用典当了珠宝的钱作资金。我们的英雄对人们的嗤笑充耳不闻。他说：你们被眼睛欺骗了。它不能用桌子来代表，而应该用鸡蛋来代表。他率领了三个姐妹去寻求证据。他们勇往直前，有时要越过巨大无比的寂静。但大多数时间里，他们要越过险恶的高山峻岭。日子变得很长。那些有疑问的人们担心会掉到边缘之外。最后，不知道从哪里出现了有翅膀的动物，这意味着无可匹敌的成功。」

当我们默读这段话时，每个字每个词我们都懂，而且每句子也都合乎语法。但是，我们理解了吗？我们得到了什么讯息？为什么我们读不出意思来呢？

作者写一段话语，目的是要在读者的脑中引发出一个代表他们所要描述的实况的模型。要达到这个目的，作者必须运用读者熟悉的词汇，用读者熟悉的语法规则排列。更重要的是，读者对作者所要描述的情况和所要表达的感受，已经有直接的或间接的经验（即实质知识）。同时，读者必须有能力将被引发的经验加以组织而进行思考。缺少了这三个条件的任何一项，一段话语便不能理解而显得莫明其妙，甚至语无伦次。这三个条件和阅读的关系可以用下图表明：

(图 1)



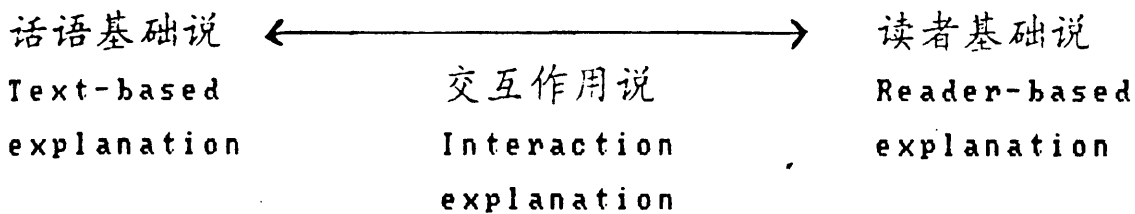
刚才，我们默读后，脑中呈现许多影象，这是和作者已经成功地用词语适当地排列，但我们却不知所云。为什什呢？这是因为我们不能把这些影象合理地组织成符合我们的经验的实况模型。这当然不是因为我们没有思考能力，而是作者用了许多譬喻，而不直接描述。这情况下，作者和读者之间思想的交际便中断了。作者表情达意的企图也失败了，而读者吸取经验的目的也达不到了。谁之过？这要看我们赞同阅读心理的哪一派说法。

「话语基础说」(如gough, 1986)认为文字的意义大部分存在于文字中。阅读时，读者把符号变成声音，用声音引发有关的知识，从而理解一段话语所含有的意义。例如学生看到「蛙」这个符号，想到它的发音，而这个声音wa使他联想到一种青色的，会跳的小动物等等。然后，将这个影象和其他同样被引发的影象组织，构成实况模型，便理解了话语。

「读者基础说」(如good man, 1976)却认为文字的意义大部份存在于读者的背景经验当中。阅读时，读者凭既有的经验，根据前文却去推测后文，再以真正出现的后文和前文互相印证。必要时，再假设、再印证。因此，阅读被说是一种「语文心理上的猜谜游戏」。

这两派说都试图解释阅读过程，但出发点和观点都相反。「话语基础说」采取微观的态度，以单独的字词为出发点。至于「读者基础说」却采取宏观的态度，以文章整体作出发点。就象许多事情一样，折衷的看法时常比较合理。因此，「交互作用说」(如Rumelhar, 1976)认为在阅读过程当中，文字的意义一部份存在于文字中，一部分存在于读者中。阅读时，读者有时以文字为起点，当经验和文字能配合时，一切顺利理解。同时，读者也不断推测后文。如果一切顺利，话语得到理解。如果预测的后文不出现，阅读有困难，读者便会停下来搜索文字所代表的意义，甚至根据前文和后文去推测意义。换句话说，从「交互作用说」的立场看来，阅读牵涉到「话语基础说」和「读者基础说」所个别描述的活动过程。对阅读过程的这三种解释，其间的关系可以用下图来表明：

(图 2)



综合以上各点，阅读并不是一种刻板的、机械式的单向直译，由形到音，由音到义。阅读时，读者必须根据既有的知识，在被文字引发后，加以组织，从而得到新的理解，新的知识，并且经验新的感受。以这个观点来看，阅读既不是「从文字中汲取信息」，也不是「将已有的经验投射在文字上」，而是「将被文字引发的经验重新排列，创造新经验」。据此，阅读是话语的再创作。

* * *

前面已经提到，阅读牵涉到三个因素，也就是语文知识、实质知识和思考能力。这意味阅读的评测必定也牵涉到这三个因素。换句话说，一个阅读测题并不只测量学生有多少语文知识，它同时也测量学生有没有和它有关的经验，而且也测量学生思考能力的高低。请看下例：

- (例 2) —— 达·芬奇的时代“仅能看到”固有色，
—— 达·芬奇在自己的时代“创造”了固有色。
- (a) 因为…所以…
 - (b) 虽然…不过…
 - (c) 与其说…不如说…
 - (d) 不是…就是

这是个「高层次」的测题，它测量学生能不能看出两个单句之间的关系。要正确回答，学生必须对美术（固有色）和美术史（「达·芬奇」）有所认识，并且将两个单句所提供的信息逻辑地排列。很明显的，缺乏有关的实质经验和思考能力，这个测题便难以回答。

如果将这测题稍为改一改，变成：

- (例 3) —— 与其说达·芬奇的时代“仅能看到”固有色，
—— 达·芬奇在自己的时代“创造”了固有色。
- (a) 虽然说
 - (b) 所以说
 - (c) 不如说
 - (d) 不是说

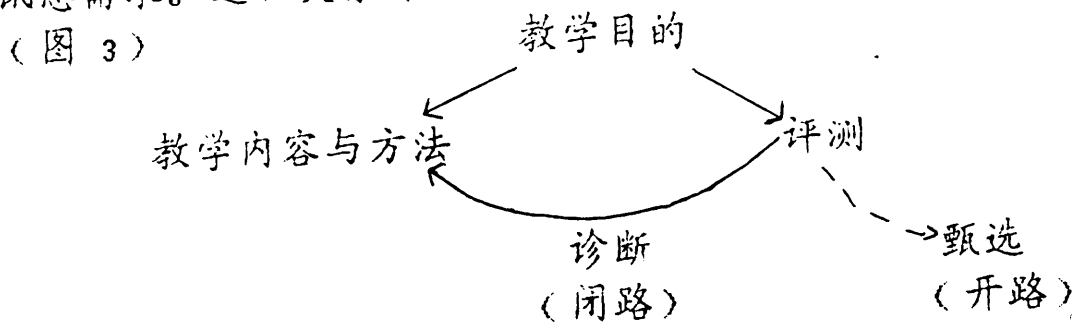
象一样的测题，学生只要学过“与其说…不如说”的搭配，即使不懂美术，不能作高层次思考，也能答对。可是，这样的测题并不测量阅读理解，而且测量关联词语的联想（记忆），是低层次的测试。

显然，要设计“纯粹的”阅读测题，丝毫不牵涉实质知识和思考能力的阅读测试，虽然不是不可能，却非常困难。这是因为文字只是代表既有或自创的经验的符号，文字本身没有意义，文字的意义来自使用者的经验。照此类推，语文本身没有内容，语文的内容必向须人类的其他活动（音乐、美术、经济、科学、政治、文学等等）借用。其实，「井蛙之见」、「夏虫不可与语冰」、「之乎者也」，正说明经验在阅读中的重要，也说明文字的空洞虚无。

以上的论点也意味编作阅读测试时，实质经验和思考能力的份量应该在合理的程度上尽量轻。要不，测量的到底是理解能力的强弱，实质经验的有无，或是思考能力的高低，便无从确定。以测试理论来说，这样的测题，效度(Validity)很低，不能真正反映阅读能力。

* * *

教学目的决定教学内容与方法及评测方式，这是众所周知的。但值得提醒的是，评测有不同的方式，以配合不同的讯息需求。这种关系可以用下图来表明：



教师在指导学生阅读后，需要可靠有效的信息，以便确定学生已掌握哪些阅读技巧，而在哪些方面需要更进一步的指导。要做到这点，教师需要有关个别学生阅读能力相当细微的信息。这种信息，并不是一般作为甄选评测的课室测试所能提供的。教师需要一套有系统的阅读微技测试，作为搜集信息的工具。以下是阅读微技的一些可能类别。

(例 4) 玫瑰园中所栽种的玫瑰种类，可说是_____，
数也数不完。

- (a) 美不胜收 (c) 鲜艳夺目
(b) 盈千累万 (d) 五花八门

(从语境中出现的同义词语去推断相应的词语)

(例 5) 老板对工人说：“你想不想吃饭！_____。”

- (a) 厨房里还有些剩饭
(b) 想吃，就赶快去吧
(c) 吃饭是件重要的事
(d) 再不好好地干，我就把你辞退

(理解语句的交际意义)

(例 6) 青年人有着旺盛的精力，有着美好的理想，有着火样的热情，更有坚强的奋斗精神。_____，才能使自己的志向成为现实。

- (a) 总而言之 (c) 换句话说
(b) 那就是说 (d) 只有这样

(通过语法组合标志理解语段各部份的关系)

虽然，编制一套完整阅读微技测试，不是一朝一夕由个别教师在百忙中可以完成的工作。但是，通过集思广益，通力合作，并且假以时日，这样的一套测试是可以达成的。编制后，广泛应用再略加修订，似乎可以暂时高枕无忧，虽然不能一劳永逸。

话说回来，在这样的一套测试问世之前，教师需要自己个别或小组合作设计测试，以便日常考查用。这项阅读教学上的重要工作，牵涉到好几个基本问题。

第一个问题是：要测量什么？这必须从一篇文章或一段话语为什么产生谈起。一般说来，一篇文章产生的目的不外是传讯（如：说明文、通告、启事、新闻），说服（如：议论文、广告、评论、社论）或消遣（如：笑话、散文、诗歌、小说）。当然，一篇话语可能兼有一种以上的目的。同时，作者写作的意图不一定也就是读者阅读的意图。为达到目的，作者必须选用某些词汇，运用某些语法，采取某种语调。目的能不能达到（读者理解、误解或不解），要看作者所用的

词汇、语法、语调是不是读者熟悉的，要看读者能重新创造作者心目中的影象和实况模型。因此，阅读的评测可以从下列四方面着手：

- a. 测量读者对语境中词汇的了解。
- b. 测量读者对语段的内容的了解。
- c. 测量读者对作者意图的了解。
- d. 测量读者对语段的情绪感受。

前两项的测量，可通过教师熟悉常用的词汇和语法测题及理解问答，这里不再赘述。至于后两项，不妨以刚才用过的一个例子来说明。

(例 7) 老板对工人说：“你想不想吃饭！再不好好地干，我就把你辞退。”

老板说这话的目的是：

- (a) 要工人吃饱饭，有力做工。
- (b) 要工人认真工作，不偷懒。
- (c) 要工人自动辞职。
- (d) 要辞退工人。

(作者意图的理解)

(例 8) 老板_____对工人说：“你想不想吃饭！再不好好地干，我就把你辞退。”

- (a) 高兴地
- (b) 生气地
- (c) 关心地
- (d) 和气地

(读者对语段情绪的感受)

第二个问题是：要测量阅读学习的成绩或是测量阅读能力的掌握？这两种不同测试的差异相当微妙，单从测试的内容和形式看不出。前者测量学生对已教过的课文的内容和文字能记忆多少，能理解多少；这是成绩测试 (Achievement Test)。它能提供信息，让教师评定课文的教学效果。然而，这种测试并不能真正地反映学生的阅读能力，因为它的记忆成份很重。要测量学生的阅读能力有没有适当地发展，测试必须考查学生的学习转移 (transfer of learning) 或说应用 (application)；要看学生能不能将从课文中学到的语文知识转化为能力，进而理解从来不曾见过的话语。这种测试是

所谓学力测试(Proficiency Test)。显然,成绩测试要根据教过的课文编制,而学力测试则必须选用未教过的话语篇章来设计。

学习而不能转化是白费功夫、浪费精力和时间。因此,阅读的测试应以学力的考查为最终目的,而成绩的考查只是必须经过的一个教学阶段。从另一个角度来看,成绩测试是倒摄的,焦点在过去,而学力测试是前摄的,焦点是将来。教师既需要适当的信息去了解教学情况,也需要确定学生能不能学以致用。因此,阅读评测有必要兼顾成绩与学生的测量。

同时,应考虑的是,教学过程中的测试应用着重阅读微技的测量,以便教师能够及时补充、补救。但是,这些微技必须在适当的时机里加以统合应用。因此,学期、学年及学阶的评测应该着重整体的理解测量。

*

*

*

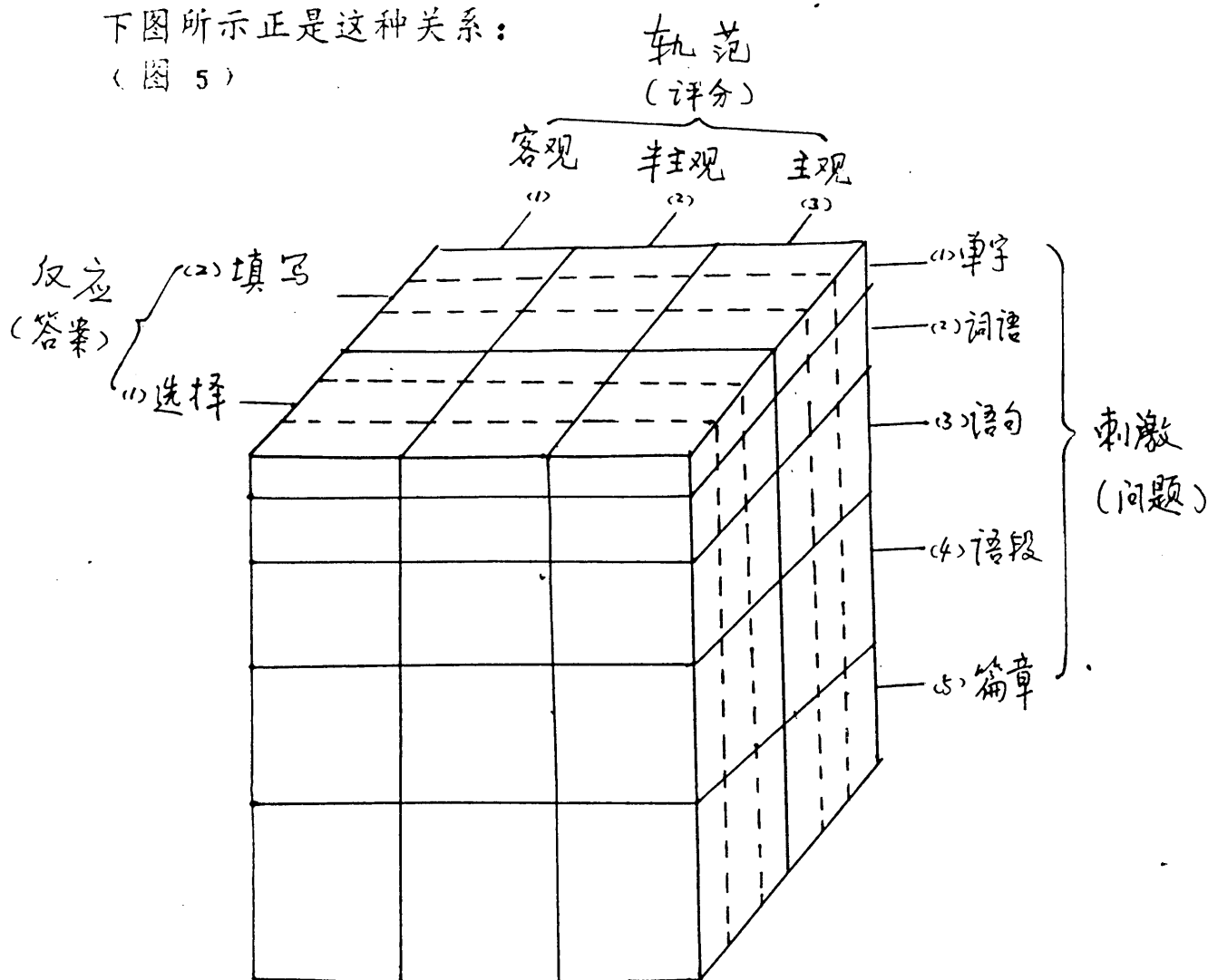
第三个问题:测题要采取怎样的形式?这是一般教师在设计测试时最先想到的问题(例如,要出多少题填充题?要出多少题选择题?要出是非题和问答题吗?)其实,这是应该最后才考虑的问题。测题形式是帮助教师达到测试目的的工具;形式本身不是目的,过份重视或给予优先考虑,似有舍本逐末之嫌。

不论测题用什么形式,每个测题都有三个基本的、不可或缺的部份。从心理学的观点分析,测量学生的阅读能力时,教师最先提出一种刺激进行反应,然后,由教师将学生的反应和教师预期的反应加以比较,再以预定的规则评分。这个历程可用下图来表明:



例如教师用测量学生对一则笑话的理解,最实际的测量方法是讲这则笑话,再观察学生有没有大声笑,或只作会心的微笑、或是显得不知所云。当然,象这样最实际的测试也是最符合语文教学目标的。不过,在教室中,用这方式,非常不方便。

教师在经年累月的教学工作中，已经熟悉了许多不同形式的测题；没有必要作多余的介绍。但将这些经验系统化应该有用。既然如上面所提的，测题有三个构成部份，将它们的常用形式综合起来，可看出它们之间的关系及可能的变化。下图所示正是这种关系：



根据这个模式，教师可以设计许多种不同的测题。实际上，可能形式不只这些，因为选择式(Selection type)还可细分选择题、配合题、是非题等等，而填写式(Supply type)又可细分为填充题、短答案题、问答题等等。

教师面对这繁简不一的多种测题形式，要怎样决定取舍呢？以下几点是值得考虑的：

1. 问题的长短简繁要配合学生程度及教学目的。
2. 作答方式力求配合阅读能力教学的层次。
3. 评分力求客观。
4. 同一测试中，形式变化尽量少。
5. 读写兼顾。

以上五点，前四点都是不言而喻的。不过，「读写兼顾」一点有更进说明的必要。虽然阅读和书写是两种不同的活动，但是，书写能提供大脑以视觉之外的动觉信息，并且刺激右半球的活动，对注意及记忆都可能有帮助。因此，阅读评测不能只要求学生默读作答，而要在作答中用适量的书写活动。

和阅读测试有关的另一个问题是：测题和语段的关系。传统的理解测试中，学生先默读一段文字，然后回答一系列假设和该段文字有关的问题。换句话说，这种形式的阅读评测有个假设那就是学生能不能正确回答，要看学生理解或不理解先前的语段。这假设肯定成立吗？如果学生不理解甚至不默读选用的语段，便不能正确回答吗？假如不默读或不理解也能回答，那么，这种测试测量的又是什么？这的确是个值得研究的问题。

与这有关的是，近年来，教师对「完形测试」(Cloze test, 或译「定位填充」)日渐熟悉。这种测试在编作上非常方便，评改也迅速。完形测试虽然并不测量固定预定的某种语文知识(如量词、助词等)，但却测量学生能不能通过前文和后文去推测被删略的字、词。同时，在固定距离上留空待填，被删略的字、词构成文字的「随机样本」(random sample)，可说是合理的甚至理想的综合阅读测试。值得注意的是近年来以华文为材料的完形测试都显出，学生在完形测试和传统阅读测试的得分，有相当可观的相关程度。这种既合阅读教学目标、设计与评分又便利，而且能测定学生阅读能力的形式，是不是应该更广泛利用，以提高工作效率，也是个值得研究的问题。

* * *

华文的教学一向很传统，讲求多读、多写。多读和多写都需要多时。然而，时间是有限的资源，又能妥善分配，不能无中生有。幸亏时间只是学习中一个因素(当然是重要的因素)，其他因素如学生的动机和学习方法也能影响教与学。当前的问题是：在有限度的练习时间与应用机会的情况下，怎样改变阅读教学的策略，以求事半功倍。当阅读教学策略有所改变时，阅读评测方式当然也必须相应改变。这是华文教师有必要拨出宝贵时间共同切磋的课题。