

汉语课成绩测试问题初探 ——以汉语综合课为例

侯 红 玉

语言测试包括学能测试、成绩测试、诊断测试和水平测试。期中和期末考试具有诊断意义，可以考查学生的学习效果，教师还可以通过关注学生在测试中出现的问题，调整教学方法。不过，为了和平时小测试那种诊断性测试区分开，一般把期中和期末测试归于成绩测试(即阶段性测试)，旨在检查学生是否掌握了阶段性的学习内容，进而给学生评出分数等级，同时也将测试作为学生集中学习、积累知识的一次时机。

成绩测试作为汉语教学的一个重要环节，在实际操作中存在哪些问题呢？测试方式又有哪些呢？笔者汇总了从事留学生汉语教学的几十位教师的近百份期中和期末试卷，并进行了统计，对上述两个问题进行了初步探讨。

1. 成绩测试的原则

为了体现成绩测试的综合性 and 系统性，一般试题抽取范围比较广泛，而且有代表性。试题既覆盖大纲要求的基础知识、基本技能及重点、难点，也设计一些应用型试题，对学生的实际汉语能力进行考查，同时控制好试题的难度，使学生成绩拉开差距，有明显的区分度。另外，在测试中一般会吧单项测试和综合测试结合起来。单项测试主要考查语音、词汇、语法、汉字等单向知识，而综合测试包括完型填空、作文等，可以考查学生的综合的语言应用能力。中高级阶段的测试还包括对语篇能力（比如，段落的连贯性、修辞手段的使用等）、语感及对跨文化交际能力的考查。

2. 成绩测试中存在的问题

成绩测试中存在的问题主要是试题区分度低、指导语不统一、分类不细致、缺乏交际性等，下面分别进行说明。

2.1 试题区分度低

试题区分度低的一个明显表现是考试成绩没有正态分布。比如，题目偏难，成绩基本都在80分以下；或者题目偏容易，基本都在90分以上，好、中、差学生分数差距不大。造成这个问题的原因是试题难度不当或者评分标准有问题。

2.2 试题指导语不统一

例如，把下列词语连成完整的句子：开 玩笑 了 一 他 个 跟 我

对于上面这道题，有的试卷指导语是“连词成句”，有的是“把下列词语连成一句话”，有的是“组句”等。指导语不统一可能给学生带来一些困惑，有时还会给教师成绩评定带来麻烦，比如：“选合适的词语填空：（1）顺利（2）埋怨（3）轻视（4）冷漠（5）顺便……”，对于上面这道题，有的指导语要求每个词只能使用一次，而有的没有提出这个要求，所以有的词语学生使用了两次，而有的词没用。在完型填空中，有的指导语只写“完型填空”四个字，没有具体的词语数目或者字数要求，结果有的学生填了一个字，有的填了两个字、还有的填了一个词组，所以指导语应尽量提出明确的要求或提供答题范例。

2.3 试题分类不细致

比如，下面的一道词语搭配题：

昂贵的_____ 满意的_____ 过分_____ 批评_____ 开展_____

上面这道题包括填写名词中心语、动词中心语和宾语，但是因为题目没有详细分类，显得比较混杂，发散性强，增加了试题的难度。

2.4 试题缺乏交际性

试题缺乏交际性指的是试卷上只有知识点方面的试题，学生只要死记硬背即可以应试，缺少应用能力方面的试题（比如，作文、完成会话等）。

3. 成绩测试的规范化

3.1 试卷抬头格式的规范化

试卷抬头名称统一，相关信息完整，有利于建立题库和进行试卷分析。
试卷抬头标准化格式举例：

2011级本科留学生第五学期汉语课期末考试试卷（A卷）
（2013年7月）

姓名_____ 班级_____ 成绩_____

题号	一	二	三	四	五	六	七	八	总分
题型	写拼音 写汉字	选词 填空	词语 搭配	连词 成句	完成 会话	改错	完型 填空	作文	100
分值	15	15	10	10	10	10	10	20	
得分									

试卷结尾注明考试依据的教材、第几册、从第几课到第几课，比如：《博雅汉语》（冲刺篇Ⅱ）第6课～第10课。

3.2 试题指示语的规范化

试题的指导语可以使用教材课后练习中的语言，因为学生已经习惯了，不会在答题时产生困惑。如果是新的题型，应对试题进行举例说明，即提供答题的范例，以免造成学生对题目不理解而不能正确答题。另外，指导语中应标出每道大题中每小题的分值及该题总分，作文要提供标有字数的稿纸，以方便学生计算分值和作文字数。

3.3 把握好试题难度，提高试题区分度

一般来说，中等难度的试题（即难度指数在0.30—0.70）具有较好的区分度。也有专家认为，理想的难度是0.5左右（即50%的人可以答对）。在一份试卷中，中等难度的试题一般应占70%，比较难的和比较容易的题各占15%，这样测试出来的成绩会呈现正态分布，学生的成绩会拉开差距。

3.4 尽量减少试题造成的投机因素

测试中的投机因素是指学生不是凭水平而是凭猜测答题。为了减少试题造成的投机因素，首先指导语要有明确的要求，比如，在“填写量词”时，

有的指导语提示“不能用‘个’”，而有的则没有提示，所以有的学生把不会写的都写上了“个”，给评分带来了不便。其次，成绩测试中的考试范围要广泛，这样才能全面考查学生对知识的掌握情况，能在一定程度上保证测试的可信度。

3.5 应尽量提供测试的语境

相对于平时测验而言，中期末测试范围比较广，为了减少学生在考试中的焦虑情绪，降低考试难度，应尽量提供测试的语境。比如，根据拼音写汉字这类题，可以让学生直接根据拼音写出汉字，也可以把生词放在句子中完成测试任务：

(1) lúnkuò () (2) 那座建筑在月光下只能看见 lúnkuò ()。

显然，(2) 比 (1) 更能引导学生正确答题。

3.6 通过试题的调整，保证同一个班级内测试的公平性

在“自我介绍”和“朗读”这两类测试中，由于外向学生比较乐于表达，“自我介绍”成绩可能比较好，而“朗读”测试外向学生不一定有优势，为了保证测试的公平性，可以科学地设计试题，保证同一个班级内测试的公平性。

3.7 合理确定分值

分值能否合理分配关系到成绩测试的可信度，据蔡云凌（2010）统计，汉语综合课试卷中，“根据拼音写出汉字”类试题的总分不同试卷分值从5分到30分不等，题目数量从5个到20个不等。“阅读理解”类的总分不同试卷分值从5分到20分不等。此外，还存在分值不尽合理的地方，同样是完成句子，不同试卷确定的分值有1分、1.5分、2分、3分、5分等不同级别。题目的分值应与难易度成正比。不同种类的题型难度不同，比如，提供了选项的选择题，难度上比较低，而让学生自己填写答案的填空题难度要高，相应地，后者分值也应高于前者。同样生词听写难度要低于完成句子或者造句，前者分值自然要比后者低。

3.8 题型尽量丰富，包含比较大的信息量

成绩测试的目的在于考查学生的综合能力，而综合能力的考查依赖于丰

富的题型。一般，本科1~2年级的成绩测试除了作文以外，出十道大题左右比较合适。

3.9 要详细列出评分标准

测试题要详细列出评分标准，尤其是作文。这样便于学生正确答题和了解得分的依据，也便于教师评定成绩。

3.10 平行班要统一标准

一般来说，期中考试往往只测上半学期学习的内容，期末考试只测下半学期的内容，期中、期末考试测试的大概都是七、八周内学习的内容，一般考试前只告诉学生考试范围、题型，不需要细划出重点。为了公正，平行班要统一标准。有的教师把以往考试的样题发到学生邮箱，学生可以预测自己大概的成绩，熟悉考试题型，这样有利于调动学生学习的热情。

此外，试题顺序的安排一般是由易到难，造句类、作文类试题由于难度大，排在后面。试题的题量要合适，一场1.5小时的考试，中等水平的学生答题时间应为1小时左右。在期末考试前，应把学生平时成绩（比：占30%）和期中（比如：占30%）汇总以后告知学生，让学生心里有数，有利于从容应对期末考试。

4. 成绩测试的信息反馈

4.1 及时发放试卷

一般在考试后，学生希望尽快知道自己的成绩，因此，期中试卷应及时发还学生，因为考试后学生对试题及成绩的关注度比较高，易于及时修改考试中的问题。

4.2 注意公布成绩的方式

出于自尊，很多学生不想让别人知道自己的成绩，所以不希望教师当众公布成绩。同时学生想知道自己本次考试在全班中的位置，所以教师不妨公布成绩的梯度，即公布每个分数段学生的人数。

例如：

分数	90-100	80-90	70-80	60-70	60 以下
人数	5	5	3	2	1

另外，如果能公布考试成绩的平均分的话，学生也可以借此知道自己在全班中的学习情况。为了鼓励学生，可以公布 90 分以上和进步最大的学生的名字，这也是激励学生学习的一种方式。

4.3 做好讲评和错误修改

期中考试以后教师对全班具有共性的错误及时进行讲评，至于每个学生出现的不同的错误，因为很难面面俱到，所以不妨引导学生分析自己知识和学习方法上的不足，有效地解决问题。教师自己也可以通过对学生的测试发现教学中、试卷设计中的不足，并及时改进。一般来说，试卷中的共性错误反映了教师教学中的问题，²而个别学生的错误则说明了学生本人学习的问题，而这两方面的问题都可以通过测试后的信息反馈得到改进。比如：有的教师在讲评后布置作业，让学生把自己试卷中的错误改正过来并以正确的形式再写一遍，这样就给学生一个反思的机会，避免知道分数就草草了事，造成错误不断重复出现。期末测试中的问题如果新学期教师还继续上同一个班的课的话，也可以把测试中的问题反馈给学生。

5. 汉语综合课测试题型汇总

汉语综合课在初级阶段承担着听、说、读、写这四大技能的训练任务，语音、语法、词汇这几个语言要素是测试的重点内容，在中级阶段则侧重对语篇能力的考查，高级阶段重点测试语言应用能力。笔者统计了一下，汉语综合课常见测试题型主要包括以下内容：

5.1 看拼音写汉字

(1) 生词式：yóujú_____ (2) 短语式：cānguān xuéxiào_____

(3) 句式式：Qǐng bǎ yīzi bān guòlai. _____

下面的试题比较有新意³：

(4) 妹妹说要去买一些_____票，还要买一瓶酱_____。(yóu)

(5) 把____上的行李____给我。(di)

5.2 选字填空

样题:我有一____狗,他的名字____阿黄,他喜欢____去跑。(叫、出、只)

5.3 选词填空(可选的词语数目可以多于句子的数目,从而形成对选择的干扰)

样题:选词填空:①而且 ②勉强 ③热烈 ④坚持 ⑤由

(1) 这个访问团()他担任团长。

(2) 他不愿意参加这个俱乐部,我们劝了半天,他才()同意了。

(3) ~ (5) 略

5.4 选择题

(1) 两选一类

样题:留学还是不留学我()一下再说吧。(A打算 B考虑)

两选一类试题平时测验可以使用,但是中期末这类重要的测试不宜太多,以免造成凭猜测答题的投机行为,最好是以多项选择题(在ABCD四个选项选择一个正确的)来代替两选一类试题。

(2) 多项选择类

样题:我说什么你都当成(),以后准有你吃亏的时候。

A. 追时髦 B. 听不进去 C. 耳旁风 D. 闲不住

多项选择题是被广泛使用的一种题型,其优点是评分客观,阅卷速度快,有利于试题的标准化,但是也存在以下三个问题:一是多项选择题只考辨认而忽略语言的生成,而能辨认句子的正误不等于能说出或者写出正确的句子,不能反映学生实际的语言输出能力。二是找到四个合适的选项对教师来说是有难度的,设计不好的话,会出现一些模棱两可的选项。多项选择题和填空题、问答题等测试方式相比,填空题和问答题效度高于多项选择题。也就是说,如果测试中的多项选择题比例较大,学生的测试成绩将不能有效地反映其实际语言能力。在大规模考试中(如:HSK),可以通过设计大量多项选择题以减少阅卷所需要的人力,快速进行机器阅卷,而小规模的教学测试

如果要使用多项选择题的话，一般要注意以下几点：

一是题干要简练明确，减少学生阅读题干的时间，而且尽量少用否定式或者双重否定式提问，如：“请问，下面哪件事今天没发生？”可以用“请问，下面哪件事今天发生了？”来代替，以免造成学生因为粗心而误答。

二是为了规范化，多个选项的长度尽量应该一致。例如：

请问：张华不高兴的原因是：

- A. 得知自己的发明没有获得认可 B. 没有获奖
C. 受到了嘲笑 D. 家被盗了

为了规范化，可以将A项改为“发明未被认可”，这样每个选项的字数就大体一致了。

三是不要在题目中给予提示，如：

请问，设计这所房子的是：

- A 设计师 B 水暖工 C 建筑工人 D 装修工

由于题干和选项含有共同的词“设计”，有的学生会倾向于选“A”，所以题干中的关键词不应在选项中重复出现。另外，选项中尽量不使用“从不”、“总是”、“所有”等表示范围太广的词语，有经验的学生都知道这样的选项一般不是正确答案。

四是题干相对于各个选项来说，应具有概括性。例如：

请问，前门是_____。

- A. 北京的一条商业街 B. 重庆的一条商业街
C. 上海的一条商业街 D. 天津的一条商业街

上面的题应改为：前门是哪个城市的一条商业街？

- A. 北京 B. 重庆 C. 上海 D. 天津

五是每道题最好只有一个考点，以免使学生产生困惑。

六是答案应均匀分布，不要让学生找到规律。比如，不要把答案都集中到A选项上。

七是为了保证效度，应尽量设计四个选项，以减少猜测的命中率。

5.5 量词填空

样题：在空格内填入适当的量词，不能用“个”

一_____雨伞 一_____楼

5.6 词语搭配

样题：填名词：打扫_____ 填动词：_____足球

填形容词：_____的鲜花 填补语：兴奋得_____

5.7 词语填空

样题：成语填空：温文尔（ ） 填近义词：支持（ ）

填反义词（不能使用“不”）：喜欢（ ）

5.8 解释句中加点部分的意思

样题：这件事八字还没一撇呢。（ ）

5.9 根据下面句子的意思写出相应的成语

样题：事情不因刮风或者下雨而受阻，照常进行。（成语：_____）

5.10 写出部首相同的汉字（适用于初级阶段）

样题：打——扌（ ） 边——辶（ ）

5.11 组词

(1) 用所给的语素分别组成一个动词，每空只写一个字

样题：_____论 表_____

(2) 用所给的多音字组词，每空只写一个字

样题：干(gān)_____ 干(gàn)_____

(3) 写出派生词

样题：请参考下面句子中“化”的用法，写出两个不同的派生词

王府井是北京有名的商业化的购物中心。

派生词：_____化 _____化

5.12 连线题

样题：

马大哈	太过分了，让人难以忍受，不能再继续看
看不下去	技术或者水平不高的人

二把刀	粗心大意的人
砸锅卖铁	重视、认真对待
当回事儿	比喻把自己所有的钱或东西都拿出来

5.13 对句中画线部分进行提问（适用于初级阶段）

样题：我们是在街上看见他的。

提问：_____？

5.14 造句

样题：只有……才……：_____。

5.15 改写句子

样题：在牛街能买到清真食品，作为穆斯林，我每周都去那儿，考试前很忙的话，也去。

（用“即使……也……”改写）_____。

5.16 为指定词语选择适当的位置

样题：A 你们班 B 同学 C 是 D 从日本来的吗？（都）

这类题适合考查虚词和某些特殊用法的表达。

5.17 句式转换（包括肯定句、否定句和疑问句之间的相互转换及“把字句”和“被动句”之间的转换等）

样题：我把面包吃了。（请改成被动句）→

这类题应该慎重使用，因为“把字句”和“被动句”在意义上是不同的，如果不给予意义上的提示，机械地进行替换，学生会觉得两个句子意思一样。

5.18 改错题

改错题对学生来说比较难，因为母语不同，学习中出现的负迁移现象也不同，比如：有一道改错题是：“我的发烧终于好了”，以英语为母语的学生容易受英语“my fever”的影响，出现“我的发烧”这样的偏误，而其他国家的学生不一定会出现类似问题，同一道题对不同国籍的学生来说，难度不同，所以这道题作为国际班的试题就不合适。此外，在教学和测试中，教师应尽可能不断重复正确的句子，而不要为了练习而重复学生的偏误，从不强化偏误的角度出发，我们认为改错题不应占太大比例。

改错题的样题：不知怎么了，他最近常常发脾气我。

为了降低难度，还可以把错误的地方划出来。例如：不知怎么了，他最近常常发脾气我。

5.19 连词成句

样题：

(1) 我 食堂 去 吃饭 (5个词)

(2) 脸谱 的 京剧 通过 可以 体现 性格 人物 (8个词)

这类题要注意控制句子的长度(即词语数量)，试题顺序可以逐步从短句扩展到长句。

5.20 给句子排序

A所以我们说电视广告对儿童的影响很大。

B据一项调查表明，在中国城市儿童中，约有80%的孩子喜欢看电视广告。

C心理学家认为人物形象对儿童形成男人、女人的角色意识会起到重要作用。

D特别是广告中的男女人物往往会给儿童留下深刻的印象。

句子的排列顺序为：_____

5.21 用所给的词语完成句子或者会话

样题：

(1) 完成句子：我一定要努力学习，_____ (报答)。

(2) A: _____? (印象)

B: 我根本没去过那儿。

5.22 完型填空

完型填空有两种形式，一种是选择式的，即提供多个选项，让学生选择合适的填写；另一种是自由填写式的，即不提供选项，让学生自己填写。当然，后一种难度大一些。完型填空的考点应该是学生熟悉的，为了方便学生阅读和推理，考点应尽量等距离出现，而不要过于集中。

样题：

(1) 选择式的完型填空

一些医学专家认为, 人们在生活中有一种 (I) 的疾病——“时间病”。“时间病”是指人为时间所迫而感到 (II) 紧张, 甚至因此而 (III) 旧病复发, 如冠心病、中风等。

(I) A、常规 B、常见 C、日常 D、常常

(II) A、健康 B、高兴 C、烦躁 D、麻烦

(III) A、发现 B、引申 C、发起 D、引起

(2) 自由填写式的完型填空

美国人重视中国当代文学, 翻译 () 很多中国当代作家的小说, 比如: 余华的《兄弟》、姜戎的《狼的图腾》、莫言的《生死疲劳》等; 而德国人 () 不重视中国当代文学, 认为中国当代文学平庸、没有价值。德国人喜欢 () 是中国当代诗人的诗歌。

5.23 作文

作文属于综合性测试, 能集中体现学生语法、词汇的实际运用情况和成段表达能力, 初中高级阶段都可作为测试题。

(1) 作文题型

作文测试题在初级阶段可以采用看图写故事、写信、回电子邮件、写留言条等形式。

样题: 看图说话



注: 本图选自《亨利》148 页, 卡·托·安德生等作, 译林出版社, 1992 年。

在中高级阶段, 为了体现阶段性考试的特点, 一般选择课文中出现过的话题或者是贴近学生实际生活的话题, 以免学生看到题目不知从何谈起。也可以列出一些词语让学生选择部分或者全部写出一篇文章。

命题作文举例 1: 介绍一处名胜古迹

命题作文举例2：阅读下面的短文，然后按照要求写作文

爸爸是个足球迷

我爸爸是个足球迷。

大家也许还没忘记在新加坡举行的世界杯足球赛亚洲区预选赛中国队对阿联酋队的那场比赛。那天，大家都睁大眼睛注视着电视屏幕。上半场0:0结束，下半场在离比赛结束还有16分钟的时候，中国队进了一球，“1:0”！全场沸腾了。

……

你瞧，我爸爸够不够迷的？

题目：请补充文中“……”处省略的内容，通过想象，写出“爸爸”当时的神态和动作。

要求：

A 格式规范；正确使用标点符号；用词准确；语言通顺；内容丰富、生动。

B 字数在100字左右，插在省略号处，注意上下文的连贯性。

(2) 作文测试要注意的问题：

A 在字数上应体现年级梯度，比如：一年级 200字左右、二年级 300-500字、三年级 500-1000字、四年级 1000-2000字。

B 列出详细的评分标准，使学生有所依据，也有利于教师客观评定成绩。

参考标准如下：

内容充实；中心突出；语言流畅；用词丰富，能正确使用比较复杂的句式和一定的修辞手段，增加语言的表达力；句子之间和段落之间要连贯；不应少于500字；格式正确；能正确书写汉字和使用标点符号。内容丰富10分，语法表达10分，连贯性3分，汉字书写2分，共计25分。

上述参考标准在初中高不同阶段可以根据学生实际水平做出相应的调整。

5.24 阅读下面的短文，然后回答问题

阅读理解测试要注意以下几点：

(1) 课内和课外材料相结合。比如：课文部分占60%，课外占40%。如果都出书上的题目，则体现不出学生实际的阅读水平，分数也难以拉开差距。

(2) 可以在报刊、杂志和网络上选择知识性、时效性、趣味性强的文章作为阅读材料，在篇幅、生词量、语言方面进行适当修改后采用。

课外阅读试题材料可以参考：HSK样题、《国际金融报》、《环球时报》、《新发现》、《中国中学生报》、《中国经济周刊》、《读者文摘》、《人民文摘》、《市场报》、《人民画报》、《人民日报海外版》及百度网、中国民族文化网、中国寓言网、中国百科网、中学生作文网、人民网、一品故事网等。

一段四五百字的语言材料，由于信息容量小，可能只能设计出一两个题目，而一段二百字的语言材料，由于信息容量比较大，却可以设计出四五道试题，所以选取信息量大的语言材料比较好。

常见的阅读理解题型有：判断正误、多项选择和问答题等。其中，“判断正误”猜测的可能性比较大，信度较低，不宜占较大比例。此外，其他题型还有：填同义词、填近义词、填表格、连线、综合填空、分析段落大意等。测试点包括阅读材料的细节、关键词语的理解、句子和段落的大意及文章的中心思想等。

阅读材料的文体一般要比较丰富，记叙文、说明文、议论文等尽量都涵盖到。同时还要结合学生的实际水平，尽量选择丰富的题材，比如，环境、家庭、情感、社会、政治、经济、民俗、历史、生物等。

(3) 把握好文章的长度和测试题数量的关系，并注意控制阅读材料中生词的数量。

结语

本文以汉语综合课为例介绍了成绩测试中存在的一些问题，分析了测试如何规范化，并介绍了常见的一些测试题型。其他课型，比如，听力、口语、阅读、写作、听说、视听说等课程的成绩测试各有特点，希望以后能进行这

方面的考察。仅就汉语综合课而言，学生的实际情况不同，选择的题型也多种多样。本文的阐述只基于个人的教学经历和有限的统计资料，分析的角度不可避免会带有一定局限性。

注

- 1 《对建立“对外汉语各课型题库”的思考——从汉语精读课的试卷分析谈起》蔡云凌，《云南师范大学》（对外汉语教学与研究版），2010年11月第8卷第6期
- 2 《谈成绩测试的科学化》陈若凡，《世界汉语教学》，2002年02期
- 3 《略谈对外汉语成绩考试的改进》崔颂人，《语言教学与研究》，2006年第4期

参考文献

- 1 《谈成绩测试的科学化》陈若凡，《世界汉语教学》，2002年02期
- 2 《大型对外汉语口语成绩测试的探索》陈昭玲，《汉语学习》，1998年04期
- 3 《对外汉语中级精读课成绩测试试题设置》李玲，《才智》，2012年17期
- 4 《略谈对外汉语成绩考试的改进》崔颂人，《语言教学与研究》，2006年第4期
- 5 《对外汉语教学测试与评估的历史演变与发展趋势》杨翼，《中国考试》，2008年第4期
- 6 《多项选择答案对Cloze Test的影响》喻爱菊，《外语教学与研究》，1990年第1期
- 7 《多项选择题出题技巧及应用》黄素华，《外语界》，1997年第3期