

博士研究生英语入学 考试的个案分析

王 鸥

【提 要】从难易程度、考试题型与测试方式等方面来看,博士研究生英语入学考试在校际间存在着较大的差异,因此我院这一考试无疑具有其特殊性。笔者把中国社会科学院研究生院博士研究生英语入学考试作为个案,在回顾历史演化和现状发展的基础上,通过对考试难度、难度变化与难度合理性等因素的统计分析,提出未来在考试命题和测试方式等方面的调整建议,从而使这一考试能够更加科学与公正。

【关键词】博士学位英语入学考试 考试难度 难度的合理性 统计分析 正态分布

〔中图分类号〕G **〔文献标识码〕**A **〔文章编号〕**1000-2952(2012)06-0016-07

英语作为国际语言(International Language),是研究生的必修课程,也是从事研究过程中重要的外语语言交流工具。作为进入社会科学研究领域激烈竞争的第一道门槛,周密准备和顺利通过我院博士研究生英语入学考试,必然要耗费考生大量的精力与资源。更重要的是,考试中的试题的难度与题型,评判方式与标准,对考生英语学习方法和外语能力提高将起到十分重要的导向作用。

一、现状与问题

与硕士研究生英语入学考试的全国统一考试不同,我国博士研究生英语入学考试的一个重要特点是命题与考务工作一直是由各学校自行组织并实施完成的,国家也并没有制定统一的考试大纲与相关细则。这一制度为博士研究生招生单位的不同专业特点提供了一定的自由度,例如我院在英语测试中一直突出社会科学

领域的研究范围与特点。而从另一方面看,这一制度又给此项重要考试带来了一定的随意性和非规范性。“由于考卷的命题、施考、阅卷工作都是各自为政,自成一体,因此,关于此类考试试题的设置、难度的掌握的校际间交流甚少。”^①总体考察:各个高校博士研究生英语入学考试的考试程序和相关制度基本相同,而难易程度、考试题型与测试方式却存在着较大的差异。

从中国社科院研究生院招收博士研究生30年的发展过程来看,博士研究生英语入学考试从命题、施考、评判到分数统计等一系列工作,主要都是由外语教研室与相关部门配合协调完成的。经过了前期的波折和调整,近十年的考务工作已逐渐步入正轨。试卷设置命题,考试施考程

^① 张波:《博士研究生英语入学考试质量解析——案例分析》,《中国社会科学院研究生院学报》2006年第6期。

序等各个环节已经形成了一套行之有效的程序与规范，保证了历年该项工作的顺利完成。与此同时，每年一度的博士研究生英语入学考试也逐渐演化成为一种例行公事：首先，保密工作为重中之重；其次，从完成命题的数月前期准备到考前卷面落实与印刷装订，再到主观题与客观题的判卷与登分；但是，对考题难度是否适中，卷面题型是否合理，评判标准是否公正等重要环节，缺乏事后的统计分析和科学评估，由此，更谈不上进一步的调整与改进。

作为我国高等学历教育中最高教育等级，博士研究生培养无疑有其特殊性，正是基于这一特点，预计国家不会在短期内推出类似于硕士研究生英语入学全国统一考试的政策安排。因此在未来相当长的一段时间内，博士研究生入学考试依然会由各高等院校自行组织考试、分别命题与评判。鉴于此，有必要在发展演变和现状研究的基础上，对本院博士研究生英语入学考试难度与合理性进行统计分析和深入探讨，进一步优化改进。

二、程序与设置

表 1 试卷的基本题型与分值分布

	题型	题量 (题号)	分值	总分	
PART I	词汇	Sec. A 替换	10 (1~10)	1	10
		Sec. B 填空	10 (11~20)	1	10
PART II	语法	Sec. A 替换	10 (21~30)	1	10
		Sec. B 改错	10 (31~40)	1	10
PART III	阅读理解	20 (41~60)	1.5	30	
PART IV	翻译	Sec. A 英译中	5 (1~5)	3	15
		Sec. B 中译英	3 (1~3)	5	15
总计		68		100	

上表显示我院博士研究生英语入学考试考卷的题型，分别为词汇、语法、阅读理解和翻译四大部分，同时按照考试的答题方式的不同，又分为客观题和主观题两种类型。在题型和分值分布中，词汇、语法和阅读理解前三部分为客观题，一共为 60 道小题，分值一共为 70 分；第四部分为主观题，其中包括英译中、中译英两大部分，

分别为 15 分，共 8 道小题，分值为 30 分。客观题的特点是答案具有唯一性，试卷采用的客观题全部是四个备选答案，其中只有一个为正确答案的选择题，答案必须填入有 [A]、[B]、[C]、[D] 四个选项的答题卡上。主观题的特点是考生必须动手落笔，最终将答题内容写在特别印制的答题纸上。主观题的另一个特点是答题内容呈多样化，答案缺乏唯一性。在主观题中值得注意的是，英译中考题所提供的英文内容，是包括带有下划线的五个翻译段落的整篇文章。虽然考试仅仅要求翻译下划线部分，但是整篇文章提供的主题、中心思想和逻辑结构等上下文关系对答题部分的理解非常重要。中译英部分为主题相互独立的三个中文段落，一般与政治时事和社会科学的内容相关。^①

在考试的程序方面，针对博士研究生考试的特点，保密工作无疑是重中之重。我院博士研究生英语入学考试一方面基于参与出题老师的职业道德，同时也不断完善相关措施和保密制度，近十年来没有在这一环节出现任何问题，保证了考试的公平公正。在此基础上，考试的公平公正还有赖于考试程序的建立与完善。多年来在出题的分工协作方面，在客观题的计算机判卷程序的调整，答题卡的机读识别与最后分数的统计，主观题判分标准与评定方式，直到最终分数的计算、登记与核查等方面，均形成了一套有效的工作程序和制度，中外老师分工协作，基本保证了这一考试工作的顺利完成。

在这项考务工作中，最需要研究与改进的就是考试的命题环节，这也是本文统计分析的重点部分。

三、统计与分析

从考试命题环节来看，难度与题型及其合理性是测试方式和考核标准最重要的基本要素，因此也是本文统计分析的要点。各高等院校博士研究生英语入学考试的差异，主要表现在难度与题型上，两者在校际间差异也是我院未来

① 详见《中国社会科学院研究生院博士研究生入学考试英语试题及解析》，中国商务出版社 2012 年版。

调整的重要参照与借鉴。以下根据可获得的资料,对我院博士研究生入学考试两个时间段的分数及其它数据进行初步的统计与对比。

表2 博士研究生考试两个时间段分数的统计与对比
(2005年~2007年与2010年~2012年)

年份	第一时段			平均值	第二时段			平均值
	2005	2006	2007		2010	2011	2012	
最高分	82.00	83.00	87.00	84	74	79	86	80
平均分	49.46	47.90	50.88	49.41	43.61	46.09	49.50	46.40
分数线	57	55	59	57	51	53	58	54
过线比例%	25.7	25.9	27.2	26.27	25.88	29.57	25.90	27.11
高分段比例%	1.9	2.1	4.8	2.9	0.00	0.02	0.07	0.03
实考人数	1094	1032	1136	1087	1257	1037	1165	1153

1. 难度的分析

表2综合了可获得的基本数据,以最高分、平均分反映我院博士研究生英语入学考试分数的基本状况及其变化,以分数线和超过分数线的比例、高分段比例^①等指标反映这一考试的基本难度与难度变化。同时表2又反映出了2005年至2007年和2010年至2012年两个时间段分数、难度及变化的对比,这一考试题型结构和分值分布在这两个时间段没有任何差别,具有完全的可比性。

试卷题量是考题难度的一个重要衡量指标。在这两个时间段中,历次考试的题型和分值是完全相同的,题量也保持基本相同。回到表1可以清楚地看到,共四大类7部分题型,总考题数共68题。虽然略少于一般水平考试所要求的80~100题量,但客观题部分的每一道题,特别是词汇和语法部分,其长度和难度都远远超过一般水平的考试;同时第四大类主观题的翻译部分共8题,特别是英译中有一定阅读难度,所有的主观题又都需要有一个从理解构思到翻译书写的过程,相当耗费时间,至少要用考试时间的三分之一方可完成。同时,历年试题卷面总量都保持在15~16页左右。^②在小小

时时间的严格约束下,再考虑到博士研究生考试的心理压力,先不谈答题质量,能够最终完成全部主客观题的答题已属不易。由此不难看出,我院博士研究生英语入学考试单从考卷总题量上看,就已经显示出了相当高的难度。

2. 难度变化的分析

表2统计结果显示:从考试平均分数和分数线两个指标来看,平均分除了2007年略高于50分,其他年份均低于这一数字,2010年仅为43.61分,为此不得不将录取分数线降低到近50分的水平。统计期内各年的录取分数线从来没有达到60分,最高在2007年达到了59分,6年来平均分数仅为55.5分。考生能达到分数线的比例常年徘徊在25%,在较高的年份也没有超过30%。尽管有高分的出现,但是70分以上为高分段的比例甚少,并且此类高分考生多为从事英语工作,特别是高等院校的外语教师。以上分析初步表明:我院博士研究生英语入学考试考题难度过高,即使是在试卷相对简单、平均分较高的年份,难度也是偏高的。

从两个时间段考试的各项统计数据对比还可以清楚地看到:尽管我院一直在试图调整与降低考试难度,但考试分数和达标比例不但没有发生明显的提升,反而几乎所有的分数数据都有所下降,平均分在两个时间段的平均值从49.41下降到了46.40,录取分数线的平均值也由第一时间段的57分调整到了第二时间段的54分。这说明只有通过大幅度降低录取分数线才能维持考生过线比例的稳定,保持足够的生源。这一状况从最高分和高分段比例来看更加明显,最高分3年的平均值从第一阶段的84分下降到第二阶段的80分,而高分段比率在第二阶段趋向于零,几乎失去了统计意义。这一状况反映了考试难度依然居高不下,更加重要的是这种状况说明了对考试难度调整的无效。

难度调整失灵需要更深入地进一步探讨:试

① 由于我院博士研究生英语入学考试难度较高,高分段的统计以70分划分,高分段为70分以上的考生的数量占英语入学考试考生总数的比例。

② A4纸,小四号字。中英约4万个字符左右。

卷的难度惯性,考生结构的变化,加之考生对于难度的心理预期等因素都值得进一步的研究。但从命题角度来看,考题自身难度及考题题型的合理性更为重要,这些是考试命题的内在因素。

3. 难度的合理性分析

考试难度可以由质和量两个方面分析,并且两者具有一定的替换关系。考试的总量在前面已经有所分析,考试的质主要表现在考试难度的合理性安排上。下面从客观题和主观题两个方面分别进行定量和定性的分析。

(1) 客观题合理性的分析

对考题合理性的定量分析在我院外语入学考试中一直是个空白,造成这一问题的原因是此类定量分析需要繁杂的计算,花费大量时间和精力,个人研究十分困难。由于涉及到了大量的数据录入和统计计算工作,作为一个尝试,我们仅选用了最近两年即2011年和2012年的考试为样本,范围也限定在客观题内,在此基础上进行计算整理和统计分析。^①

总体来说,首先经过博士研究生考试初期的磨合与变动,近十几年来试题的题型和难度已经相当稳定,因此,统计样本的选定既有一定的历史代表性,也能反映这一考试的现实状况。其次,从表2可以看出经过几年的难度调整,2011年度到2012年度的考试难度调整的效果具有一定的典型性,两年间平均分与分数线差值分别为3.41分与5分,这两项数据的增加值的变化在本文所选定的统计年间差值都是最大的,说明这两次考试难度的调整起到了一定的作用。

表3 客观题答题正确率比例分布的对比

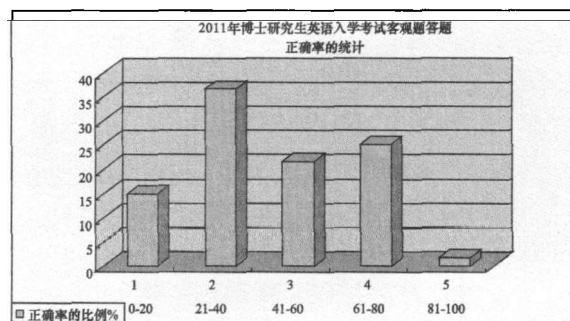
答题正确率的划分区间(%)	2011年		2012年	
	题数(道)	比例%	题数(道)	比例%
0~20	9	15.00	1	1.67
21~40	22	36.67	19	31.67
41~60	13	21.66	23	38.33
61~80	15	25.00	13	21.66
81~100	1	1.67	4	6.67
总题量/百分比	60	100	60	100

表3是对2011年与2012年相同数据内容的比较,其背后需要大量的统计。统计首先针对每年60道客观考题,逐题统计出每道题考生答对的数量,再计算每道题的答题正确率,即答题正确的人数与考生总数的比值。表3按照答题正确率比值的大小划分了五个区间,每一区间以20%为单位,由此统计出落入这一区间中答题正确率的题数和总题数的百分比。例如,考生答题正确率极低的0到20%区间段,即答题的正确率仅在20%以下,2011年的题数为9题,到2012年减少到仅为1题。在这一区间的题数与总题数的比例也从15%下降到了1.67%。在81%到100%的正确率的空间,也就是考题过于容易的区间,题数从2011年的1题增加到了2012年的4题,落入区间的题数占总题数的比例也有较大幅度的上升。通过两年间两个典型区间的对比不难看出,2012年的考题难度有所降低。

难度降低仅是一个侧面,考试难度中的合理性则是更加值得研究探讨的。一般水平考试的难易度,即答题正确率应该控制在30%~70%之间,考虑到博士研究生入学考试的特殊性,正确率低于20%或高于80%无疑是难易度失常。这一失常的原因需要逐题个案的分析,但此类题量过多无疑会影响到考卷整体的合理性。

根据表3按年度数据分别绘制了下面的图1与图2,便于更直观地对测试的合理性进行分析和考察。

图1 2011年客观题答题正确率的分布



① 王方文同学对此作了大量计算统计工作,对本文有所贡献。

图2 2012年客观题答题正确率的分布

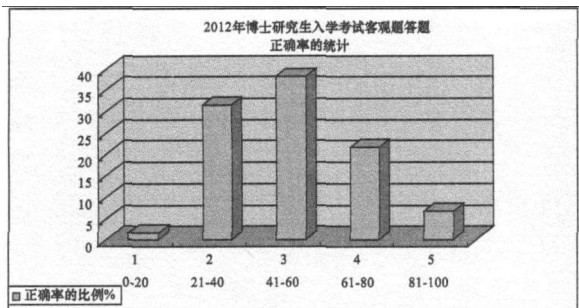


图1与图2分别直观展示了2011年与2012年我院博士研究生英语入学考试客观题正确率的分布状况。与表3统计方式相同,两图分别划分了五个区间,每一区间以20%为单位,方柱形显示了答题正确率的分布状况。

2011年分布则显示,方柱形分布趋向于左端,即向答题正确率低的一端偏斜。说明当年答题正确率过低比率区间(0%~20%)和正确率较低比率区间(21%~40%)比例过高,具体的题数分别达到了9题和22题,所占比率分别为15%和36.67%,这两个区间相加共为31题,占总量的51.67%,超过了总体量的一半。更加具体的统计显示:其中有过多的考题过难且不尽合理,还有的考题答对率仅为10.51%或11.76%,也就是说仅仅有略高于1/10的考生答对题。分析表明:2011年客观题的难度过高且不尽合理。

这一情况在2012年的考试中有了较大的改善。从图2看,2012年正确率的分布更趋向于标准的正态分布,说明了出题难度合理性的改善。“标准化考试,结果分数分布在多大程度上符合正态分布的常规,是考试质量检查的第一标准。”^①具体统计数据更能说明问题,落入过难和较难两个区间的总题数降到了20题,为客观题总数的1/3。特别是图2的最左端,即从0%到20%这一考题过难的区间仅有一道考题,并且该题的答对率为19.65%,已经非常接近于20%的答题正确率。这说明考题过难,整体答卷不尽合理的状况有了较大的改观。

与此同时,表3的统计结果也表明了另一种倾向:考生答对率超过81%以上的区间,题数从2011年的1题上升到了2012年的4题,其中题号为46的考题(阅读理解题型)的答对率

高达87.12%,即有接近90%的考生都能答对,这表明有些考题过于容易。虽然这一比例尚在可接受的范围,但这种状况与趋势也是同样值得注意和尽量避免的。

考题偏难与不尽合理的状况值得研究推敲,初步分析表明:这其中有的是考生知识结构不尽合理和语言运用能力不足的问题,而命题自身与题型结构是问题产生的主要因素,在此由于篇幅的限制不尽详述。总体来讲,客观题合理性的问题有所改善,但考题依然偏难,部分考题仍然值得推敲与研究。

(2) 主观题合理性的分析

由于主观题的特点与性质,本文对此部分仅作定性分析,而这一分析基本能够反映出其难度及变化与合理性。与客观题难度调整不同,主观题对调整的反应较为敏感,考生的答题水平有了较大的改善,其变化不但表现在高分段的考生比例有较大的增加,而且反映在考生整体应答水平的提高上。

如前分析,我院主观题命题仅体现在试卷的第四部分,即翻译题型中的英译汉与汉译英两个部分。翻译的第一部分英译汉分数的提高主要表现在考题难度降低和合理性的改善上:英文翻译文章的选用主体更加突出、结构更加鲜明、词汇难度降低、文章更加流畅。^②这一变化主要考虑到由于题量已经很大,主观题从阅读构思到翻译落笔需要消耗较多的时间。从改进的效果来看,考生的答题更加完整和顺畅,以往考生没有时间完成或者只能完成部分答题的比例有所减少,更能真实地反映学生英语阅读理解能力与翻译水平,平均分数也有所提高。

翻译的第二部分汉译英的难度没有明显的改变,变化主要表现在学生的英语写作能力上。由于近年来博士研究生考生趋于年轻化,英语写作的水平和能力均有了明显的提高。特别是考生中有较大的比例是在我院获得了硕士学位,

① 张波:《博士研究生英语入学考试质量解析——案例分析》,《中国社会科学院研究生院学报》2006年第6期。

② 参见《中国社会科学院研究生院博士研究生入学考试英语试题及解析》,中国商务出版社2012年版。

硕士研修期间参加了英语学术写作课程，英语写作能力已经达到了相当高的水准。从卷面上也可以看出：过去那种放弃汉译英，答题纸最后部分呈空白状的现象已大为减少，反之，许多考生表现出了相当高的英文写作能力。这也是未来改变题型，增加写作部分的条件和基础。

综上所述，针对我院博士研究生英语入学考试的变化，客观题的难度有待进一步地调整，其合理性需要进一步地改善；主观题宜保持现有难度，增加新的测试方式和这部分的分值比例。因此，这一分析不但涉及到了卷面难度调整和合理性改善，而且也未来题型的重构提供了依据。笔者认为：在现有题型的框架下，对难度和合理性的调整是必要的，但这些变化和改善是有限的；必须在未来对整体结构进行调整，引入新的考试题型和更加科学的测试方式。

四、调整与建议

1. 客观题的调整

题型调整的一个基本原则是突出语言技能的测试和实际应用能力的提高，改变历史和传统制度上形成的以考试自身为导向的测试模式。我院博士研究生英语入学题型已经沿用了十几年，这一结构中客观选择题的分值比例占总成绩的70%之多，其中的题型包括词汇、语法和阅读理解三大类。词汇和语法部分为传统测试方式，主要用于测试考生语言知识的掌握程度，这两种测试方式侧重于语言的用法(Usages)，而忽略了语言技能的实际应用(Use)。由此，此类题型难以很好地考察报考者的语言交流和应用能力，而考生往往通过死记硬背，应用所谓的“应试技巧”获取高分。国外典型英语水平托福考试(TOEFL)相继取消了语法专项考试部分，并将词汇专项考试并入了阅读理解部分。雅思考试(IELTS)也正是由于其基于语言基本交流能力的测试方式，为越来越多的教育机构所认可。国内考试的发展趋势也基本如此，全国硕士研究生入学统一英语考试(NETEM)就完全取消了词汇与语法的专项测试，增加写作这种更加注重实际能力和应用的考评方式。

针对我院博士研究生入学考试，在客观题中

只有阅读理解部分题型是对考生语言能力的测试，可以考虑逐步减少以至完全取消词汇专项客观题，将其并入阅读理解部分。在阅读理解中，文章提供的上下文(Context)关系为备考词汇提供了明确和生动的语义，使得测试更加科学。在客观题中，语法测试题型的调整也成为必然，调整可以采取难度与结构并举的原则，先降低难度，在时机成熟时改变题型结构，以更加科学的测试方式，如写作来测试考生语言结构的能力。

2. 主观题调整

主观题比重的增加和英语写作的引入已经是博士研究生英语入学考试的基本趋势。“近几年来，一些高校对于考生是否具备英语运用能力的问题越来越重视，因此，此类试卷中主观题部分的分值有所增加。”^①各学校考试题型结构上的现状与变化也为我们提供了参照和借鉴，其主要表现在：1)取消和压缩词汇、语法等对语言知识的专项测试；2)保留扩充阅读理解、完型填空等的语言能力测试的客观题型；3)增加更加能够合理考查外语综合能力，如写作等主观题型。^②“从全国重点院校的博士研究生入学考试英语试卷来看，向主观题型发展是今后研究生试卷的趋势，如北京航空航天大学、西安交通大学考博试卷中全是主观题型，仅考汉译英与英译汉和写作；又如北京大学考博试卷中有篇章改错，还曾采用没有阅读客观题，只有主观问答和阅后写总结。现在增加了英文解释和选择。因此，考生要想取得好成绩，必须重视主观题型的备考和能力的提高。”^③我院试卷中主观题分值仅占总成绩的30%，其中题型也只是英汉两种语言的互译，仅在汉译英部分考察与评定考生对英语语言结构的把握，无疑是不够全面和不尽合理的。去掉语法专项考查，增加写作题型，一方面能使测试手段更加趋于合理，另一方面也使得考试内容更加接近于实际

① 张波：《博士研究生英语入学考试质量解析——案例分析》，《中国社会科学院研究生院学报》2006年第6期。

② 博士研究生入学考试辅导用书编审委员会：《考博英语——全国名校真题精解》，机械工业出版社2012年版。

③ 《教育部博士研究生英语考试大纲解析》，http://www.b2cedu.com/bulletin/hq_bulletin_view.asp?d_id=74&zt=kaoboyingyu。

应用。

3. 听力题型和口语测试的引入

各高等院校博士研究生英语入学考试的试卷内容基本上分为客观判断题和主观笔试题；考试程序大致分为初试和复试两个阶段，复试包括以面试的形式测试英语口语和与专业相关的外语表达能力。外语教学中的听、说、读、写四种形式并举，口语和书面语言紧密联系，相辅相成，缺一不可，早已成为外语教学的共识。传统外语教学违反语言自然发展这一客观法则，重视语言的读写功能，忽视其最基本的听说手段已经是贻害无穷。因此，对博士研究生入学考试的口语能力的测试不仅是必要的，而且应对此提出更高的要求。在被调查的全国高校中，愈来愈多的学校把听力和口语的测试引入到了博士研究生的入学考试中来。^①

我院曾经在博士研究生英语入学考试中，把听力作为正式考试的一部分，并将其测试结果计入总分成绩，后因种种条件的限制和由此引发的问题而取消了听力测试。新校园的先进设备为这一测试提供了客观条件，考生结构的变化使得听力测试更能反映外语的基础水平。因此，建议重新逐步引入听力测试，同时，有必要在复试中增加英语口语表达能力的测试。为了更好地过渡和防范风险，一开始可以不计总成绩，作为复试或者免修考试的参考，条件成熟后转为外语入学考试的正式测试项目。

五、原则与预期

多年来难度调整的实践证明，要想达到预期效果并非易事，未来对考题的合理性和题型结构的改变将面临更大的困难。这其中涉及到种种外在的因素，诸如考生的结构与英语水平的变化，考生对于考试题型的适应程度以及对考试难度的心理预期等。但作为标准化水平考试的博士研究生入学英语考试，调整与改变的基本原则应该是使考试难度更加适中，题型与程序更加合理，测试更加能够符合语言自然规律和反映应试者的真实水平。调整与改变的重要前提是保持各项指标相对稳定，避免出现大的波动。可以预期，主观题的增加对判卷难度和老师素质会提出更高的要求，听力测试和口语复试的引入也会受限于种种客观条件的约束。然而必须清醒地认识到，改变与调整是未来发展的必然趋势，早改比晚改更加主动，及时调整比被迫调整将更加合理。

本文作者：中国社会科学院研究生院外语教研室副教授、博士
责任编辑：马光

^① 会议调研：北京高等院校研究生英语教学协会2012年青岛会议。

A Case Study of English Entrance Examination for PhD Candidates

Wang Ou

Abstract: As to the difficulty level, question types and test modalities, there exists significant dissimilarity in the English Entrance Examinations for PhD candidates among the different schools. Therefore, the English Entrance Examination for PhD candidates of CASS undoubtedly possesses its particularity and characteristics. As a case study of this examination, based on the historical evolution and the current situation, this paper makes a statistic analysis on difficulty level, the adjustment of difficulty level and rationality of difficulty level, raises the modified proposal on the test modalities and the test setting and looks forward to more scientific and equitable examination in the future.

Key words: PhD English Entrance Examination; difficulty level of examination; rationality of difficulty level; statistic analysis; normal distribution