

目 录

第一章 绪论	1
1.1 研究背景	1
1.2 研究目的	4
1.3 研究问题	5
1.4 小结	5
第二章 写作测试中的提示研究	7
2.1 写作测试中的提示固有特征研究	7
2.1.1 提示固有特征与写作得分	9
2.1.2 提示固有特征与文本特征	12
2.2 写作测试中提示与考生的交互研究	15
2.3 写作测试中提示与评分员的交互研究	19
2.4 小结	22
第三章 写作测试的效度验证框架与方法	25
3.1 效度的概念	25
3.2 效度验证框架	28
3.3 综合写作测试的构念	40
3.4 综合写作测试的效度验证框架	42
3.5 效度验证方法	44
3.5.1 Coh-Metrix 文本分析工具	45

3.5.2 多层线性模型	49
3.5.3 结构方程模型	53
3.5.4 多层面 Rasch 模型	56
3.6 小结	59
第四章 研究设计	61
4.1 综合写作任务简介	61
4.2 数据收集	62
4.2.1 提示、作文、参赛学生与评分员数据	62
4.2.2 认知维度的提示特征测量	66
4.3 研究变量	67
4.3.1 提示固有特征	67
4.3.2 基于考生认知的提示特征	68
4.3.3 基于评分员认知的提示特征与评分标准	69
4.3.4 综合写作表现	70
4.3.5 语言水平	71
4.3.6 评分行为	71
4.4 数据分析	72
4.5 小结	74
第五章 综合写作测试的效度验证：提示固有特征维度	75
5.1 提示固有特征	75
5.2 提示固有特征对综合写作表现的影响估算	76
5.2.1 提示话题域与文本特征	76
5.2.2 提示任务说明与文本特征	81
5.3 提示固有特征对综合写作表现的影响	83
5.3.1 提示话题域对文本特征的影响	83
5.3.2 提示任务说明对文本特征的影响	89
5.4 基于“评估使用论证”框架的讨论	92

5.5 小结	95
第六章 综合写作测试的效度验证：考生认知维度	97
6.1 基于考生认知的提示特征	97
6.2 基于考生认知的提示特征对综合写作表现的影响估算 ...	100
6.2.1 基于多层线性模型的提示认知影响估算	100
6.2.2 基于结构方程模型的提示认知影响估算	109
6.3 基于考生认知的提示特征对综合写作表现的影响	114
6.3.1 基于多层线性模型的提示影响	114
6.3.2 基于结构方程模型的提示影响	116
6.4 基于“评估使用论证”框架的讨论	118
6.5 小结	120
第七章 综合写作测试的效度验证：评分员认知维度	121
7.1 基于评分员认知的提示特征与评分标准	121
7.2 基于评分员认知的提示特征对综合写作评分的影响估算	122
7.2.1 基于评分员严厉度的提示影响估算	122
7.2.2 基于评分难易度的提示影响估算	128
7.2.3 基于评分员与考生偏差交互的提示影响估算	134
7.3 基于评分员认知的提示特征对综合写作评分的影响	139
7.3.1 基于评分员严厉度的提示影响	139
7.3.2 基于评分难易度的提示影响	141
7.3.3 基于评分员与考生偏差交互的提示影响	143
7.4 基于“评估使用论证”框架的讨论	145
7.5 小结	146
第八章 结语	149
8.1 研究的主要发现	149
8.2 研究的理论价值与实践意义	153

8.3 未来研究方向	155
参考文献.....	157
附录.....	187
附录 1 “外研社杯”全国英语写作大赛简介	187
附录 2 综合写作评分标准	190
附录 3 综合写作任务样题	191
附录 4 考生问卷调查	193
附录 5 评分员问卷调查	197
附录 6 各地区评分员与考生显著偏差交互统计	204

表目录

表 2.1	写作评估中写作任务的各个维度	8
表 3.1	效度渐进矩阵	26
表 3.2	二语词汇能力测试中使用完形测试的理由分析	28
表 3.3	测试公平性框架	30
表 3.4	以证据为中心的评估论证模型	35
表 4.1	“外研社杯”全国英语写作大赛省级复赛的综合 写作任务提示信息	63
表 4.2	综合写作提示下作文的描述性统计	64
表 4.3	参赛学生问卷的描述性统计	65
表 4.4	评分员问卷的描述性统计	65
表 4.5	考生问卷与评分员问卷的信度统计	67
表 4.6	提示固有特征分类编码	68
表 4.7	基于考生认知的提示特征描述	68
表 4.8	基于评分员认知的提示特征描述	69
表 4.9	基于评分员认知的评分标准描述	70
表 5.1	提示固有特征分类	76
表 5.2	社会话题作文中与写作得分显著相关的文本特征指标	77
表 5.3	教育话题作文中与写作得分显著相关的文本特征指标	77
表 5.4	商业话题作文中与写作得分显著相关的文本特征指标	78
表 5.5	个人话题作文中与写作得分显著相关的文本特征指标	78
表 5.6	社会话题提示作文中具有写作得分预测力的文本 特征指标	80

表 5.7	教育话题提示作文中具有写作得分预测力的文本特征指标	80
表 5.8	商业话题提示作文中具有写作得分预测力的文本特征指标	80
表 5.9	个人话题提示作文中具有写作得分预测力的文本特征指标	80
表 5.10	显性对立观点提示作文中与写作得分显著相关的文本特征指标	81
表 5.11	隐性对立观点提示作文中与写作得分显著相关的文本特征指标	82
表 5.12	显性对立观点提示作文中具有写作得分预测力的文本特征指标	83
表 5.13	隐性对立观点提示作文中具有写作得分预测力的文本特征指标	83
表 6.1	基于考生认知的提示特征的探索性因子分析结果	98
表 6.2	基于考生认知的提示特征的探索性因子分析结果： 方差解释量	98
表 6.3	基于考生认知的提示特征的因子描述表	99
表 6.4	多层线性模型中各变量的描述性统计	101
表 6.5	综合写作得分零模型的估计结果	102
表 6.6	综合写作得分多层线性模型的统计结果	103
表 6.7	内容得分零模型的统计结果	104
表 6.8	内容得分多层线性模型的统计结果	104
表 6.9	结构得分零模型的统计结果	105
表 6.10	结构得分多层线性模型的统计结果	106
表 6.11	语言得分零模型的统计结果	107
表 6.12	语言得分多层线性模型的统计结果	107
表 6.13	多层线性模型中提示因子变量的描述性统计	108

表 6.14	综合写作得分多层线性模型的统计结果 (提示因子变量)	109
表 7.1	多层面 Rasch 模型下评分员严厉度的统计结果	122
表 7.2	评分员严厉度分组结果	125
表 7.3	二元逻辑回归分析的 Hosmer-Lemeshow 检验	126
表 7.4	评分员严厉度与其提示认知的二元逻辑回归分析结果	126
表 7.5	多层面 Rasch 模型下评分难易度的统计结果	128
表 7.6	评分员对评分标准重要性认知的统计结果	130
表 7.7	评分难度与评分标准认知的对比统计	132
表 7.8	基于评分员认知的提示特征与评分难度的对比统计	133
表 7.9	多层面 Rasch 模型下评分员与考生的偏差交互模式	136

图目录

图 3.1	效度验证的社会认知框架	29
图 3.2	解释性论证	32
图 3.3	Toulmin 推断模型	33
图 3.4	托福考试效度论证模型	34
图 3.5	以论据为中心的评估论证模型	36
图 3.6	“评估使用论证”框架的推理链	38
图 3.7	Toulmin 论证模型	39
图 3.8	“评估使用论证”框架的主张	42
图 6.1	初始模型：考生对提示特征的认知与其综合写作表现之间的关系模型	110
图 6.2	假设模型：考生对提示特征的认知与其综合写作表现之间的关系模型	111
图 6.3	考生对提示特征的认知与其综合写作表现之间的关系模型	112
图 6.4	考生对提示特征的认知与其综合写作表现之间的关系简图	114