

Supplementary Evaluation of Two-Word Thesaurus in Test Anxiety

Yiran Kong, Yutong Luo, Jianxun Li

School of Psychology, Southwest University, Chongqing
Email: kongyiran07@163.com

Received: Aug. 29th, 2017; accepted: Sep. 12th, 2017; published: Sep. 19th, 2017

Abstract

Test anxiety is one of the important factors affecting the quality of student learning. The aim of the study was to supplement the dimensions of assessment test anxiety double word libraries, because of fewer dimensions of the already existing lexical database. The degrees of pleasure, arousal and image were assessed. Test Anxiety Inventory (TAI) and State-Trait Anxiety Inventory (STAI) were used to distinguish three groups, including high, moderate and low test anxiety. 45 undergraduates from Northwest University for Nationalities were participated this research to assess 406 words at the degrees of valence, arousal, image. The results showed that the coefficients of the internal consistency were all above 0.8; for gender there were significant differences in the degrees of arousal and pleasant, but not for the degree of image; for different anxious groups, there were no significant differences in three dimensions. The reliability and validity of the test anxiety words system are acceptable and this system can provide normalized stimulant materials for further research in text anxiety.

Keywords

Test Anxiety, Lexical Database, Assessment

考试焦虑双字词库的补充评定

孔仪然, 罗宇彤, 李建勋

西南大学心理学部, 重庆
Email: kongyiran07@163.com

收稿日期: 2017年8月29日; 录用日期: 2017年9月12日; 发布日期: 2017年9月19日

摘要

考试焦虑是影响学生学习质量的重要原因之一, 由于之前的词库维度较少, 意在之前双字词评定的基础

上增加愉悦度、唤醒度和形象性三个维度, 补充评定考试焦虑双字词库。研究通过测验焦虑量表(Test Anxiety Inventory, TAI)、状态-特质焦虑问卷(State-Trait Anxiety Inventory, STAI)两个问卷区分高中低考试焦虑组, 选取西北民族大学在校本科生45人, 对406个经过筛选的双字词进行愉悦度、唤醒度、形象性三维度的评定。结果显示, 词汇库内部一致性信度均在0.8以上; 被试在唤醒度、形象性上不存在性别差异, 但愉悦度上存在性别差异; 通过聚类分析发现, 愉悦度低唤醒度低的词汇较少。本研究所评定的词汇满足标准化词库的要求, 评分结果稳定、可靠, 可作为相关考试焦虑研究的实验材料, 但有待进一步改善。

关键词

考试焦虑, 词库, 评定

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 前言

考试焦虑(Test Anxiety)是个体在考试或评价情境中产生的一种消极的情绪反应(董军强, 陈建勇, 2013), 被界定为一种情境化的特质焦虑(Keogh, & French, 2001; 于靓, 陈睿, 张小聪, 等, 2011), 包括对考试或考试情境过度的生理和心理反应, 在考试情境中, 此特质个体的焦虑会上升。研究表明, 严重的考试焦虑不仅影响学生的学业成绩和学习活动, 还影响到学生的心理健康水平(Lowe, Lee, Witteborg, et al., 2008), 导致焦虑、抑郁、无助感及自杀的增加。为了降低考试焦虑的不良影响, 有必要进一步研究。而在以往的考试焦虑研究中, 采用的实验刺激材料包括图片和文字(Lawson, 2006; 陈睿, 刘潇楠, 周仁来, 2011), 然而, 由于缺乏一套标准化的刺激系统, 使得研究材料的选择和研究结果的可比性出现困难。

考试的典型形式是笔试, 文字形式的考试内容本身往往引发考试焦虑。因此, 采用文字材料作为刺激可以更好地研究考试焦虑。然而已有的考试焦虑文字材料即考试焦虑词汇库维度过少, 根据 Osgood 等的理论, 主要考虑以下三个维度理论, 考试焦虑作为一种情绪反应其中唤醒度、愉悦度可作为测评的维度, 而作为一种认知评价, 词汇的形象性也具有评定的必要(于靓, 陈睿, 张小聪, 2011)。因此, 有必要对建立的标准化考试焦虑词库进行补充评定, 既能使研究者更便捷地选择刺激材料, 又增强了不同研究结果间的可比性(Bradley & Lang, 2017)。因此, 本研究以期通过评定, 对建立的标准化考试焦虑词库进行补充评定。

2. 研究方法

2.1. 研究对象

来自西北民族大学的二、三年级的学生, 通过方便取样发放包含 TAI、STAI-S、STAI-T 三份问卷 1041 份, 以 $M \pm 1SD$ 筛选标准, 区分高、中、低三个焦虑组共 50 人参加实验。其中 5 位同学因评定数据不全被剔除, 最后有 45 名大学生的评定数据进入了分析(平均年龄为 20.85 ± 0.82 岁)。其中, 女生 19 名(20.86 ± 0.75 岁), 男生 26 名(20.85 ± 0.86 岁)。被试视力或矫正视力正常, 无任何精神疾病。实验前要求被试签署知情同意书, 并告知其实验内容及性质, 实验后给予少量报酬。

被试需要完成美国临床心理学家斯皮尔伯格编制、修订的测验焦虑量表(Test Anxiety Inventory, TAI)和状态-特质焦虑问卷(State-Trait Anxiety Inventory, STAI)中的特质量表(STAI-T)、状态量表(STAI-S),其中 TAI 中国修订本,包括忧虑性(TAI/W)和情绪性(TAI/E)两个分量表(Lang, Cuthbert, & Bradley, 1997)。共 20 题,按 1~4 级评分,“从不”为 1 分,“总是”为 4 分,总分最低为 20 分,最高为 80 分。国际上流行的考试焦虑测评量表多达数十种,但最有效的且应用最广泛的莫过于 TAI (Spielberger, 1980)。此外,状态-特质焦虑问卷采用 Likert4 点量表形式,每一项目以 1~4 计分,得分越高代表焦虑程度越高。本文选取其中的 STAI-T 用作效标。中文版 STAI 具有良好的信度和效度(王才康, 2003)。

2.2. 实验材料的收集制作

在原有词库所筛选出来的词汇的基础上,对 406 个双字词进一步进行添加维度地评定,将双字词统一处理为图片,并用 E-Prime 软件编制程序并呈现在计算机屏幕上,显示屏分辨率为 72 像素/英寸,像素为 1024 × 768,双字词图片大小为 20 mm × 40 mm,黑底白字、宋体加粗。

2.3. 维度的评定

在借鉴考试焦虑相关研究的基础上(Lawson, 2006; 陈睿, 刘潇楠, 周仁来, 2011; 于靓, 陈睿, 张小聪, 2011),对词汇进行评定时,从情绪研究的维度观点出发,根据奥斯古德等的理论,主要考虑以下三个维度:唤醒度,即阅读该词后感觉其是否能引起自己的警觉关注,如感到眼前一亮、令人敏感、刺激或迟顿、昏昏欲睡等;愉悦度,即阅读词汇后感受到的愉快与否、相关程度;形象性,即阅读该词时感觉具体或抽象的程度(李文利, 钱铭怡, 1995; 王一牛, ZHOU Li-Ming, 罗跃嘉, 2008; Kemp, Gray, & Eide, 2002; Aftanas, Varlamov, & Pavlov, 2002)。采用自我报告模式(Tucker, Hartry-Speiser, & Mcdougal, 1999),让被试分别在三个维度上对词汇进行 7 点量表评分,评分从 1 到 7,1 表示低值(低唤醒度、低形象性、低愉悦度),而 7 则代表高值(高唤醒度、高形象性、高愉悦度)。

2.4. 实验程序

实验前向被试解释实验内容,并签署被试知情同意书。然后让被试坐于距计算机屏幕 70 cm 的椅子上,熟悉实验流程后,主试输入被试的基本信息后进入正式实验。首先在屏幕中央出现 500 的注视点,然后出现词汇,要求被试根据自己的所体会到的感受,按计算机数字键 1~7 对词汇进行评分,评价时间由被试自己控制,但不能做长时间思考,按键评分后会呈现 500ms 的空屏,接着进入下一词汇评定(参见图 1)。评定完所有词汇的一个维度后休息一次,再进入另一维度的评定。词汇在每个维度内随机出现,且三个维度的顺序被试间在随机。

2.5. 统计处理

用 SPSS17.0 对实验数据进行描述性统计分析、差异性分析、聚类分析及散点图分析等。

3. 结果

3.1. 被试信息

被试在 TAI、STAI-S、STAI-T 三个量表上的描述性统计结果见表 1。经独立组 T 检验表明,被试在两个量表上的得分在 0.05 水平上性别不存在显著性差异。

根据被试的 TAI 得分分为三组,见表 2。经方差分析表明,三组被试在 TAI 和 STAI-T/S 上的得分存在显著差异($F_{(2,44)} = 52.60, p < 0.001$; $F_{(2,44)} = 37.33, p < 0.001$; ($F_{(2,43)} = 35.77, p < 0.001$)。

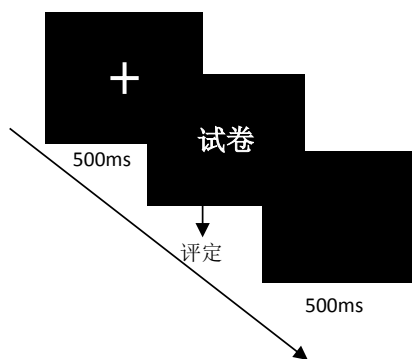


Figure 1. Experiment process

图 1. 实验流程图

Table 1. TAI, STAI-S and STAI-T scores among male and female groups: M(SD)

表 1. 男女大学生的 TAI 和 STAI-S/STAI-T 量表得分的描述统计结果: M(SD)

量表	性别	人数	M ± SD	t	P
TAI	男	26	33.70 ± 10.83	-1.41	0.17
	女	21	38.38 ± 11.53		
STAI-S	男	26	39.69 ± 12.08	-1.97	0.06
	女	18	47.06 ± 14.42		
STAI-T	男	26	42.77 ± 12.52	-0.05	0.96
	女	20	42.95 ± 10.55		

Table 2. TAI scores in each condition by level of test anxiety: M(SD)

表 2. 根据被试 TAI 得分的分组情况: M(SD)

评定维度	人数	M ± SD		
		STAI-S	STAI-T	TAI
低考试焦虑	15	32.67 ± 7.71	31.66 ± 6.03	25.76 ± 5.47
中考试焦虑	15	39.00 ± 7.49	43.93 ± 7.55	34.53 ± 7.59
高考试焦虑	15	56.20 ± 7.79	53.67 ± 8.10	48.47 ± 5.45

3.2. 考试焦虑词汇在三维度上的属性得分情况

分别对 45 名被试评定的 406 个词汇, 在愉悦度、唤醒度和形象性上的得分, 进行描述性统计分析, 结果见表 3。进一步分析表明, 男女大学生在唤醒度、形象性上的评价差异都不显著, 但在愉悦度上差异显著(见表 4), 不同考试焦虑组别被试在唤醒度、形象性上的评价不存在显著差异, 在愉悦度上差异显著(见表 5)。

3.3. 评分一致性分析

采用内部一致性信度中的克隆巴赫系数(Cronbach 的 α 系数), 对 45 名被试判断的 406 个词汇的评分结果, 进行信度的分析, 以评定测验结果的一致性、可靠性及稳定性。信度系数越高即表示该测验的结果越一致、稳定可靠。结果表明, 在愉悦度、唤醒度和形象性的评分上, 被试评定的内部一致性信度系数分别为 0.92、0.81、0.94。这表明, 被试在词汇评定中的评分一致性很高, 结果可信。

Table 3. 406 words scores among three dimensions: M(SD)**表 3.** 406 个词汇在三个维度上评分的描述统计结果: M(SD)

评定维度	人数	Min	Max	M ± SD
愉悦度	45	1.80	6.87	4.74 ± 1.01
唤醒度	45	2.40	6.14	4.72 ± 0.74
形象性	45	3.07	6.57	4.71 ± 0.67

Table 4. Means and standard deviations in three dimensions for the male and female groups**表 4.** 男女大学生在三个维度上评分结果比较

评定维度	性别	人数	M ± SD	t	P
愉悦度	女	19	4.33 ± 1.18	2.44	0.02
	男	26	5.03 ± 0.75		
唤醒度	女	19	4.52 ± 0.88	1.62	0.11
	男	26	4.87 ± 0.60		
形象性	女	19	4.53 ± 0.76	1.57	0.12
	男	26	4.84 ± 0.59		

Table 5. Means and standard deviations in three dimensions for the level of test anxiety groups**表 5.** 不同考试焦虑组别被试在三个维度上的评分结果比较

	评定维度	人数	M ± SD	F	P
低考试焦虑	愉悦度	15	4.82 ± 0.77	0.16	0.86
	唤醒度	15	4.94 ± 0.38		
	形象性	15	4.84 ± 0.67		
中考试焦虑	愉悦度	15	4.90 ± 0.87	0.23	0.79
	唤醒度	15	4.74 ± 0.53		
	形象性	15	4.79 ± 0.45		
高考试焦虑	愉悦度	15	4.49 ± 1.31	0.001	0.99
	唤醒度	15	4.48 ± 1.09		
	形象性	15	4.50 ± 0.83		

3.4. 聚类分析

用快速聚类法对词汇的唤醒度进行聚类, 将词汇分为两类, 其中唤醒程度高的词汇: $n = 221$, $M = 5.2800$, $SD = 0.35$; 低相关的词汇: $n = 185$, $M = 3.88$, $SD = 0.406$, 两类词汇之间唤醒程度差异显著, $t = -36.60$, $p < 0.0001$ 。且两类词汇在形象性上的差异比较表明显著, $t = -5.31$, $p < 0.0001$ 。为了保证词汇在形象性上无差异, 且都保持在较高形象性上, 将词汇按照形象性评分从高到低排序, 逐个删除分值较低的词汇, 并检验形象性差异, 最终排除 152 个词汇。最后确定 254 个词汇入选考试焦虑词汇库(见表 6)。再次对入选的两类词汇进行唤醒度和形象性检验, 结果表明, 两类词汇在考试唤醒度上差异显著, $t = -27.64$, $p < 0.001$, 在形象性上的差异不显著, $t = 0.41$, $p > 0.05$ 。

用愉悦度对两类词汇进行聚类分析, 最后聚类出四类词汇(见表 7)。对四类词汇的唤醒、形象性进行

方差分析, 结果表明, 四类词汇在唤醒度上的差异均显著, $F_{(3,250)} = 480.80, p < 0.0001$, 这表明聚类分析有一定的区分能力, 有助于后续相关研究对词汇的选择; 但在形象性上差异不显著, $F_{(3,250)} = 1.69, p > 0.05$ 。

3.5. 唤醒度与愉悦度之间的评分分布散点图

为进一步分析唤醒度与愉悦度之间的关系, 对愉悦度和唤醒度评分平均分之间做了评分分布散点图。从图 2 中可以看出, 被试对词库中的词汇在愉悦度和唤醒度两维度评分的平均分分布状况。整个评分分布几乎呈现线型分布, 有个别极端相关和威胁的词汇出现。且唤醒度评分和愉悦度评分都处于较高水平时, 词汇相对较多; 而唤醒度的评分相对较低时, 愉悦度评分处于较低水平, 词汇较少。

Table 6. 254 words scores among imagery and valence dimensions: M(SD)

表 6. 254 个词汇在形象性和愉悦度上评分的描述性统计结果: M(SD)

评定维度	分类	n	Min	Max	M±SD
形象性	高唤醒	158	4.78	6.11	5.21 ± 0.33
	低唤醒	96	4.78	5.96	5.23 ± 0.32
愉悦度	高唤醒	158	5.00	6.00	5.54 ± 0.50
	低唤醒	96	2.00	5.00	4.19 ± 0.55

Table 7. Lexical clustering results

表 7. 词汇聚类结果

评定维度		n	唤醒度	形象性
高唤醒度	高愉悦度	86	5.55 ± 0.14	5.23 ± 0.27
	低愉悦度	72	5.02 ± 0.23	5.19 ± 0.27
低唤醒度	高愉悦度	91	4.02 ± 0.30	5.24 ± 0.28
	低愉悦度	5	3.23 ± 0.13	4.93 ± 0.07

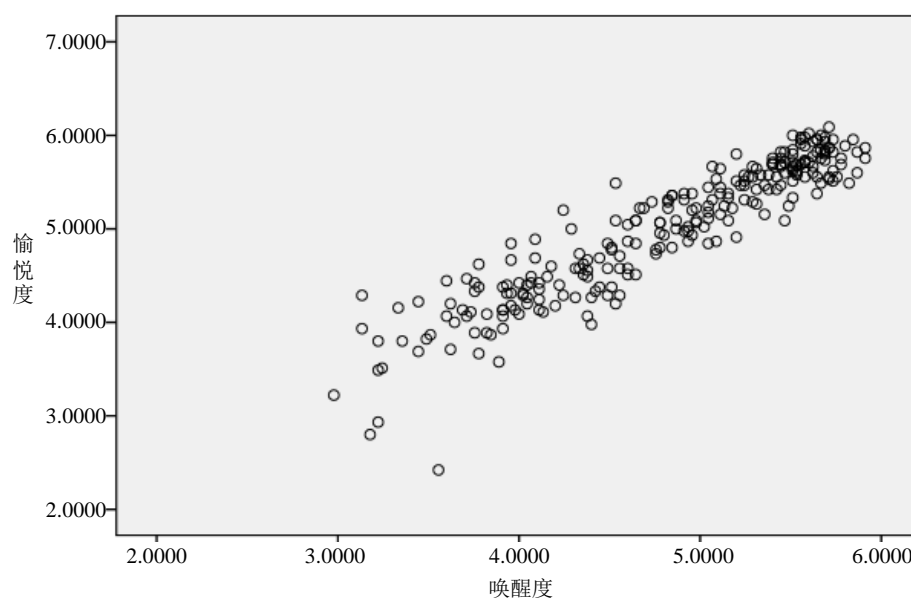


Figure 2. Scatter score between arousal and valence

图 2. 唤醒度和愉悦度之间评分散点图

4. 讨论

该研究采用自我报告模式,在借鉴汉语情感词系统及考试焦虑图片库系统的评定基础上,编制了考试焦虑词库系统,在词汇材料的选择上,力求广泛丰富。其次,除评定了唤醒度和愉悦度外,还增加了形象性的维度。因为考试焦虑作为一种特殊情境下的情绪反应具有维度性的特点,刺激与个体的主观唤醒程度、唤醒的情感体验、唤醒事件的形象性的不同,对个体产生的影响也可能会有差异。形象性作为对语言共性,是语言认知评价的重要因素之一,具有评价意义。内部一致性信度分析结果表明,被试在对愉悦度、唤醒度和形象性的评分上,评定的内部一致性很高,这说明在同一维度下被试的评分一致性和同质性较好,不同被试对同一词汇的评定较为一致、可靠。

由评分分布散点图可以看出,考试焦虑词库中的词汇的唤醒度和愉悦度的平均值成线性地分布在七点量表上,整个评分分布范围较广,有个别极端词汇出现。由于考试焦虑词库评定主要针对学生群体关于考试焦虑的认知评定,而这一群体对考试相关词汇的唤醒度要求较高,考虑到要尽量保持高唤醒度和低唤醒度词汇在个维度上的平衡,因此在词汇的选择上多选取的是形象性相对较高的词汇,使形象性的评分集中于较高水平。同时在词库选择过程中,根据逐个排除形象性低的词汇,并依次检验形象性差异,最终保证入选词汇库的高、低唤醒度词汇间无差异。

以唤醒度进行聚类,与考试高及低相关的词汇的分布较均匀,类别间相关度差异显著;而结合愉悦度进行聚类的结果表明,四类词汇间在唤醒度差异显著。这表明聚类分析在这些维度上有一定的区分能力,且被试对词汇的评定较为一致,但在形象性上差异不显著,表明形象性在愉悦度上不具备区分能力。

此外,在0.05水平上高中低考试焦虑组别的被试,在词汇三个维度的评定上的差异都不显著。被试对词汇的评定较为一致,即在对词汇的评定阶段,被试组间差异不显著。不同性别的被试在唤醒度和形象性上差异不显著,但在愉悦度上差异显著,原因可能是由于女性的积极情绪与消极情绪的相关程度比男性高原因导致的,比如女性更倾向于以辩证方式体验和表达情绪,评定出相对更接近于平均数的分数(伏春宇,杨炳钧,2006)。

综上所述,考试焦虑双字补充词汇库可以作为一套作为补充修订词库,补充了标准化的刺激系统,为今后考试焦虑的研究提供标准化的实验刺激材料。然而,考试焦虑词库仍需继续完善。首先,词汇库中只选择了双字词作为材料,且涉及的数量较为有限,每个维度分段上的词汇较少,尤其是低唤醒、低愉悦词汇不多,为了能给今后的相关研究提供更多的选择空间,有必要在后续研究中进一步扩充词汇库;此外,词库是在有限范围内的被试中进行评定的,对于同一个词汇,不同评定者也许有不同感受,因此,该词库只能提供一个相对稳定的标准。同时,在今后的研究中需要进一步探讨词库在其他样本群体中的适用性。

5. 结论

考试焦虑双字补充词汇库可为今后考试焦虑的研究提供标准化的实验刺激材料。

参考文献 (References)

- 陈睿,刘潇楠,周仁来(2011). 不同程度考试焦虑个体对威胁性刺激注意机制的差异. *心理科学*, 1, 151-154.
- 董军强,陈建勇(2013). 核心自我评价、应对方式对大学生考试焦虑的影响. *科技通报*, 29(1), 209-214.
- 伏春宇,杨炳钧(2006). 英汉语言形象性对比研究. *外国语文*, 22(1), 126-131.
- 李文利,钱铭怡(1995). 状态特质焦虑量表中国大学生常模修订. *北京大学学报:自然科学版*, 31(1), 108-114.
- 王才康(2003). 考试焦虑量表(TAI)的信度和效度研究. *中国临床心理学杂志*, 11(1), 69-70.
- 王一牛, Zhou, L.-M., 罗跃嘉(2008). 汉语情感词系统的初步编制及评定. *中国心理卫生杂志*, 22(8), 608-612.

- 于靓, 陈睿, 张小聪, 等(2011). 考试焦虑图片库的编制:在中国大学生群体中的试用. *中国临床心理学杂志*, 19(1), 38-41.
- Aftanas, L. I., Varlamov, A. A., Pavlov, S. V. et al. (2002). Time-Dependent Cortical Asymmetries Induced by Emotional Arousal: EEG Analysis of Event-Related Synchronization and Desynchronization in Individually Defined Frequency Bands. *International Journal of Psychophysiology Official Journal of the International Organization of Psychophysiology*, 44, 67-82.
- Bradley, M. M., & Lang, P. J. (2017). International Affective Picture System. In *Encyclopedia of Personality and Individual Differences* (pp. 1-4).
- Kemp, A. H., Gray, M. A., Eide, P. et al. (2002). Steady-State Visually Evoked Potential Topography during Processing of Emotional Valence in Healthy Subjects. *Neuroimage*, 17, 1684-1692. <https://doi.org/10.1006/nimg.2002.1298>
- Keogh, E., & French, C. C. (2001). Test Anxiety, Evaluative Stress, and Susceptibility to Distraction from Threat. *European Journal of Personality*, 15, 123-141. <https://doi.org/10.1002/per.400>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International Affective Picture System (IAPS): Instruction Manual and Affective Ratings. In *FL: NIMH Center for the Study of Emotion and Attention*. Gainesville: University of Florida.
- Lawson, D. J. (2006). *Test Anxiety: A Test of Attentional Bias*. An Abstract of the Thesis Presented in Partial fulfillment of the Requirements for the Degree of Doctor of Philosophy.
- Lowe, P. A., Lee, S. W., Witteborg, K. M., et al. (2008). The Test Anxiety Inventory for Children and Adolescents (TAICA). *Journal of Psychoeducational Assessment*, 26, 215-230. <https://doi.org/10.1177/0734282907303760>
- Spielberger, C. D. (1980). *Test Anxiety Inventory: Preliminary Professional Manual*. Palo Alto, CA: Consulting Psychology Press.
- Tucker, D. M., Hartry-Speiser, A., Mcdougal, L. et al. (1999). Mood and Spatial Memory: Emotion and Right Hemisphere Contribution to Spatial Cognition. *Biological Psychology*, 50, 103-125.

附 录

评定词汇表举例

钢笔 升学 容貌 失眠 实话 倾听 倾诉 团聚 单词 获奖

Hans 汉斯

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2160-7273, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: ap@hanspub.org