

The International Affective Picture System (IAPS)

—Comparison of Evaluating Method in Young Adults Sample

Wenchi Huang, Shihkuang Chiang

Department of Counseling and Clinical Psychology, National Dong Hwa University, Hualien City
Email: skchiang@mail.ndhu.edu.tw

Received: Jan. 13th, 2014; revised: Jan. 18th, 2014; accepted: Jan. 24th, 2014

Copyright © 2014 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

Abstract

The International Affective Picture System (IAPS) was developed to provide a large set of standardized, emotionally evocative, internationally accessible photographs from a wide range of semantic categories. The researcher used the Self-Assessment Manikin (SAM), a non-verbal pictorial assessment technique, to measure the pleasure, arousal and dominance associated with the effects of IAPS on the participants. The aim of the present study was to check the difference and any correlation between SAM and a nine-point Likert scale. Twenty-six university students (aged 18-30 years) participated and evaluated 72 pictures twice over four weeks from IAPS using the SAM scale and a nine-point Likert. There were three results. First, the reliability of the internal consistency of the nine-point Likert scale was good. Second, the nine-point Likert scale was significantly correlated with SAM scale. Third, there were no significant differences between the means of the three dimensions of the nine-point Likert scale and the norm of IAPS. In conclusion, the two evaluation methods of IAPS are suitable to be used on young adults. The present data have provided the author with the opportunity to discuss its applications and limitations, as well as relevant recommendations for future research.

Keywords

IAPS; SAM Scale; Likert Scale; Young Adult

国际情绪图片系统的评估方式研究

—以年轻成人为例

黄玟绮, 蒋世光

国立东华大学咨商与临床心理系，花莲市
Email: skchiang@mail.ndhu.edu.tw

收稿日期：2014年1月13日；修回日期：2014年1月18日；录用日期：2014年1月24日

摘要

国际情绪图片系统的发展是为了提供大量标准化，能激发情绪且可以跨国使用的包含广泛语意类别的照片。国际情绪图片系统的研发者使用自我评定模型“一种非语文的图形评估技术”来测量参与者在图片的愉悦度、唤醒度和控制度上的反应。本研究的目的是检验自我评定模型与九点量表间的差异与相关。研究对象：26位受试者平均分为两组，每组针对72张IAPS照片在间隔4周以九点量表与SAM量表分别进行评估，但两组施测顺序采对抗平衡设计。研究结果：对于年轻成人两种评估方式在愉悦度、唤醒度和控制度皆有良好的内部一致性，且两种评估方式的结果没有明显差异。结论与建议：两种方式在年轻成人评估IAPS时都适用，但本研究也存在受试者控制度评估变异较大，未能对年轻成人以外的群体进行两种方式比较的限制，建议欲应用本研究结果时应注意前述限制。

关键词

国际情绪图片系统；SAM量表；九点量表；年轻成人

1. 引言

国际情绪图片系统(International Affective Picture System, IAPS)是由美国国家心理卫生院(National Institute of Mental Health, NIMH)针对情绪与注意力研究所需而编制的一套标准化情绪刺激图片系统(Lang, Bradley, & Cuthbert, 1997)。IAPS (Lang, Bradley, & Cuthbert, 2005)目前已有超过 900 张以上的图片，这些照片大多都是由报章杂志等媒介收集而来的，都有明确的内容与分辨率，包含人类情绪的喜、悲、恐怖、愤怒、威胁、魅力、丑陋、或各种物品、裸体、家、艺术作品、日用品、住宅、性接触、死亡、公害、污染、都市风景、海景、运动场面、战争灾害、新闻杂志、医学治疗、病人、切割的身体、动物幼儿、凶暴动物、昆虫、温暖的家庭、孩子的异想世界等内容。这些照片，有良好的内部一致性与再测信度，且在西方文化中也有相似的特性(Bradley & Lang, 2007)。

Wundt 是最先提出情绪有共同基础特性的学者(Wundt, 1896)；他擅长使用精确的实验来计量各种心理特性，同时他也对当时众多的情绪研究素材进行整理与归纳，发现情绪的促发因子包含文字、对象、事件；而这些刺激可以利用三种基础向度解释，即为愉悦度(lust, pleasure)、紧张度(spannung, tension)、抑制度(beruhigung, inhibition)。而后 Osgood 与他的同事延续 Wundt 的想法，他们设计出一种有 50 个词汇且具相反极性的语意量表(如冷与热、黑与白、快与慢)，并招募受试者，要求参与者去评价各种语言刺激；在这当中，他们对实验数据做因素分析后，发现 50%的变项可以被三个评估要素所解释，即为评价(evaluation)、活动(activity)、效力(potency)，这三个因子也在跨文化的研究中得到证实，上述实验开展并逐步奠定情绪的基本构面(Osgood, Suci, & Tanenbaum, 1957)。但是以语意量表作为研究工具，还是具有一些缺点，例如：1) 测量实验情境中每个刺激的各个不同语意量尺，费力又费时，而且在因素分析上统计过于庞大，容易对结果产生误差。2) 建立的语意量尺多为英文，且有些词汇较难被理解，尤其是母语非英文国家之受试者，常会因文化、语文理解上而有不同的判读。3) 若考虑跨文化或不同教育程度之受

试者，应在指导语与研究使用前先进行严谨的信效度检验，若无法有标准化的施测环境，其结果难以外推。

为了改进上述缺点，Lang(1980)设计一种图片导向型(*picture-oriented*)的评估工具，称为自我评测模型(*Self-Assessment Manikin, SAM*)，受试者透过这些拟真的图形，在评估 IAPS 图片的愉悦度、唤起度、控制度的主观感受时，可以更直接、快速的反应个人的情绪经验。除了图片外，SAM 也可以有效的运用在评估情绪刺激的多种情境，如影像、声音、广告、痛觉反应或更多情境上的情绪自我评估，也可以用在小孩、焦虑症患者、特定恐惧症、心理病态、和其他临床族群(Bradley & Lang, 1994)。

回顾文献，使用 IAPS 的研究大多利用其作为情绪促发、评估、电生理反应之工具或评估各种情绪刺激媒介方式，但较少对 IAPS 的评估方式做研究；仅 Bradley 与 Lang(1994)曾尝试比较 SAM 与 Mehrabian 与 Russell(1974)的 18 种双极性语意量表的差别，结果显示 SAM 比起冗长的语意量表，反应情绪较具直觉性、反应时间较有效率；Lang, Greenwald, Bradley 与 Hamm(1993)发现 SAM 在评定情绪反应时，与生理讯号呈高相关，因此他们认为 SAM 是一种易于使用、且能替代生理测量的研究工具。虽然使用 SAM 具有前述优点，但使用 SAM 时需要大量的纸张、批阅人力和时间；蒋世光，谭伟象，花茂琴，陈皖兰与张兆贤(2012)的研究在实际操作观察中也发现 SAM 的标准化指导语解说复杂且耗时、倘若受试者分心而注意力涣散则会忽略讯息，例如忽略 SAM 人形图示之间的空格也可选择，造成受试者仅勾选图案、因此降低答案之变异性，或作答时因作答时间较短暂，而错看人物图形错选完全相反的答案。因此本研究的主要目的是探讨使用 SAM 与九点量表来评估 IAPS 是否存在差异。

2. 研究方法

2.1. 受试者

本研究招募 30 位自愿参与之研究生与大学生者为受试者，当中男女比例需小于 1:2 以符合 IAPS (Lang, et al., 2005)之建议。并于当中排除严重肢体障碍、精神疾病、视力不良未矫正、色盲等影响参与者接受标准化刺激之潜在可能因素。本研究将 30 位受试者分为两组，A 组与 B 组各有 15 人，每组分别接受九点量表与 SAM 量表两种评估方式，两种不同评估方式间隔 4 周，但两组施测顺序采对抗平衡设计。在排除 4 位无法完成前测或后测的参与者后，最后 A 组包括男性 4 名、女性 9 名；B 组包括男性 5 名、女性 8 名；两组受试者在年龄上无明显差异($t = 0.61, p = 0.55$)，整体受试者年龄平均为 23.65 岁，标准偏差 2.86 岁，以卡方检定样本同构型，可发现两组成员在性别、学历上皆无明显差异(表 1)。

2.2. 研究工具

2.2.1. IAPS

本研究使用蒋世光等人(2012)研究中所建议之包括 18 类主题共 72 张的 IAPS 照片作为实验材料(表 2)。

2.2.2. SAM 作答纸

SAM 是将愉悦度、唤起度、控制度各以不同的人型漫画图片呈现，让参与者能快速理解与判断、减少参与者对语言的负荷量之非语文量尺。量尺中皆有五个人形图，依序代表该向度的不同程度，而这些入型皆是等比例的呈现该向度之程度变化，每个图案程度下方皆有数字，研究者需要求参与者将贴近自身情绪经验的数值圈起；若参与者较难判断反应，量尺位于人形之间的白色小框意指邻近人形的平均值，可帮助参与者更细微分辨情绪(Bradley & Lang, 1994)(图 1)。

2.2.3. 九点量表作答纸

本研究参考 SAM 量表的三个向度，以愉快程度、唤醒程度、控制程度请参与者回应，如表 3 之九点量表范例设计，在最左边标号题项，采李克特式量尺，其程度包含 1~9，当中 1、3、5、7、9 有特别写

Table 1. Chi-square test and *t* test for characteristics of sample
表 1. 本研究样本特性的检定

	A 组	B 组	<i>t</i>	<i>p</i>
人数(男/女)	13(4/9)	13(5/8)		
大学生	5	4		
研究生	8	9		
χ^2	0.44	3.26		
<i>p</i>	0.51	0.07		
平均年龄(标准偏差)	24.0(3.34)	23.3(2.39)	0.61	0.55

Table 2. The numbers of IAPS used in this study
表 2. 本研究选用之 IAPS 编号

图片编号									
8033	9921	1302	9432	9008	9911	7284	4659	2661	4002
7002	2691	4607	7235	4004	4534	7286	4608	3053	2703
6243	1930	6311	5520	7175	8050	5611	8341	3215	4235
7340	2395	9500	9341	9925	2800	6821	1932	5530	5811
8490	4574	4255	7475	4669	2360	3069	8220	9290	2700
7052	1525	5510	8040	1947	8179	3000	1603	8178	2005
2299	3063	2154	9300	3016	2455	2683	3180	2491	5500
6415	4542								

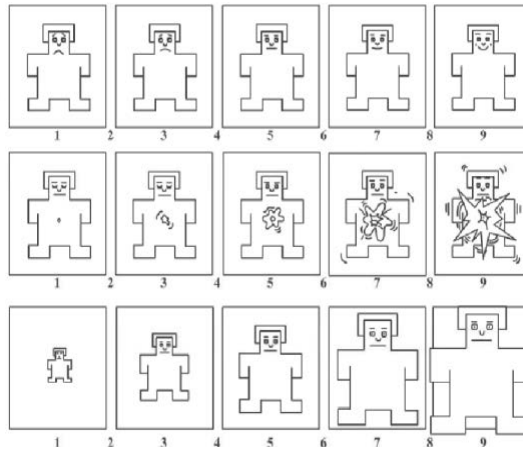


Figure 1. Manikin figures in SAM

图 1. SAM 人形

上中文文字描述。以愉快程度为例，如 1 = 非常不愉快、3 = 有点不愉快、5 = 中等、7 = 有点愉快、9 = 非常愉快，当中的 2、4、6、8 提供参与者在介于程度于程度之间作更细微的确定与选择。

2.3. 施测程序

受试者正式实验前，皆签署研究同意书。本研究依据 IAPS 的标准化评估程序进行，完成一张照片的评估共需 26 秒，其中先呈现 5 秒的标示题号、接着 6 秒呈现 IAPS 刺激照片、随后 15 秒为作答提醒和受试者评估时间(图 2)，完成 72 张图片的时间约需四十分钟。进行正式施测前每位受试者皆接受研究者清

Table 3. Sample of nine points Likert scale
表 3. 九点量表设计范例

愉快程度									
	非		有					有	非
	常		点					点	常
	不		不		中			愉	愉
	愉		愉		等			快	快
	快		快						
1	1	2	3	4	5	6	7	8	9
2	1	2	3	4	5	6	7	8	9
3	1	2	3	4	5	6	7	8	9

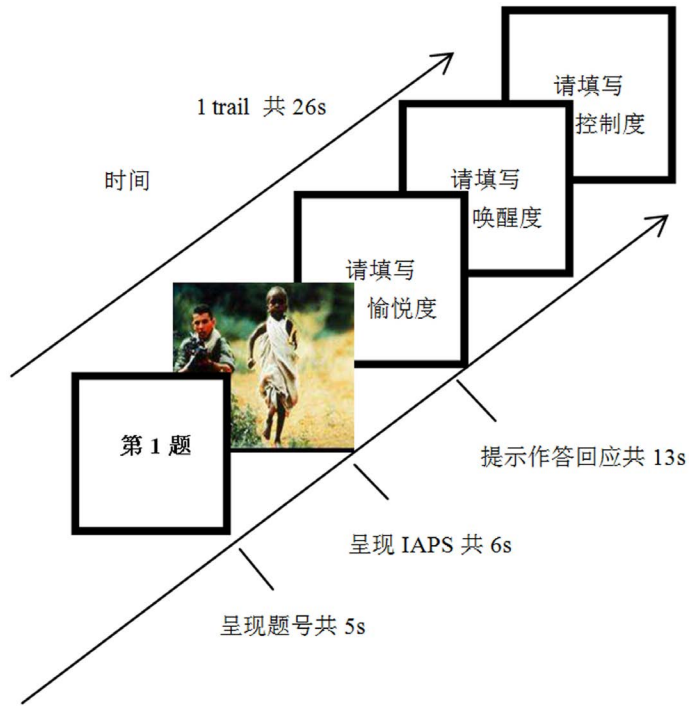


Figure 2. A demo of experiment procedure
图 2. IAPS 实验程序示意图

楚的指导语说明与三题 IAPS (2005 版) (Lang et al., 2005) 练习题(编号为 4200 正性照片、7010 中性照片、3100 负性照片)。受试者完成前后两次评估后则赠与参与研究的小礼物。

2.4. 资料分析

本研究以 SPSS 16.0 版分析收集到的资料。

3. 研究结果

3.1. 内部一致性

表 4 显示 26 位受试者以九点量表评估的愉悦度、唤起度、控制度之内部一致性信度系数为 0.75、0.98、

Table 4. Cronbach's α of two methods
表 4. 两种评估方式的内部一致性

	九点量表	SAM
愉悦度	0.75	0.68
唤起度	0.98	0.97
控制度	0.97	0.94

0.97; 以 SAM 量表评估的愉悦度、唤起度、控制度之内部一致性信度系数为 0.68、0.97、0.94。

3.2. 九点量表与 SAM 量表的差异与相关系数

表 5 显示 26 位受试者以九点量表与 SAM 量表分别评估的愉悦度($t = -2.51, p > 0.05$)、唤起度($t = 0.19, p > 0.05$)、控制度($t = 0.29, p > 0.05$)皆没有达到显著差异; 两种评估方法在愉悦度($r = 0.72$)、唤起度($r = 0.80$)、控制度($r = 0.39$)的相关系数皆达显著正相关。

4. 讨论

本研究发现九点量表与 SAM 量表的内部一致性系数除了愉悦度之外, 皆在 0.94 以上, 相对地, 九点量表评估的愉悦度的 α 值为 0.75、SAM 量表评估的愉悦度的 α 值为 0.68, 这样的结果与蒋世光等人(2012)研究中的年轻成人以 SAM 量表评估的愉悦度的 α 值为 0.93 的结果差异较大。本研究推测可能的原因之一是本研究样本数较小的问题, 参考其他以小样本进行之 IAPS 内部一致性研究, 本研究发现其结果与本研究相似, 例如在 Backs 等人的研究中, 21 位年轻人(平均 20.02 岁)的愉悦度 α 系数为 0.63, 仍低于本研究之结果(Backs, Silva, & Han, 2005)。影响愉悦度的整体 α 值的另一个可能原因为不同价性的图片张数不同, 本研究采用蒋世光等人研究的 72 张图片中, 正性和负性图片各有 32 张、而中性照片仅有 8 张, 本研究也发现若将图片按照其正性、负性和中性分别检验内部一致性, 结果发现正性和负性图片的愉悦度 α 系数将可分别提升至 0.88 和 0.93; 但以 SAM 量表评估的中性图片之愉悦度 α 系数为 0.48, 与九点量表评估的中性图片之愉悦度 α 系数则为 0.72, Kline(2000)认为一个好的测验其信度系数至少要大于 0.70, 本研究综合表 4 和表 6 的结果, 发现在样本数较小和中性图片较少的情形下, 以九点量表评估的 α 系数能符合 Kline 的要求, 但以 SAM 量表评估的结果则无法符合(表 6)。

表 5 显示虽然九点量表与 SAM 量表分别评估的愉悦度、唤起度、控制度没有显著差异, 两种评估方法在愉悦度、唤起度、控制度的相关系数也达显著正相关, 但两种方法评估的控制度之相关系数却明显较愉悦度与唤起度之相关系数低。Bradley, Codispoti, Cuthbert 与 Lang(2001)认为控制度对情绪的解释度较愉悦度与唤起度低, 本研究认为其原因可能是控制度的评估包含了个体对于自身能力、事件经验、环境等条件的考虑, 所以不同个体间的变异性上比愉悦度及唤起度大, 且从美国 NIMH 的常模中, 控制度的标准偏差也大多比愉悦度、唤起度的标准偏差高(Lang et al., 2005)。因此有些研究会不考虑控制度的评量(Grühn & Scheibe, 2008; 胡少华, 魏宁, 郭文滔, 胡健波, 谭云飞, 许毅, 2005)或不详细讨论(Backs et al., 2005; Silva, 2011; 王莹, 许晶, 张炳蔚, 冯霞, 2008)。

综言之, 本研究为极少数比较 SAM 量表与九点量表评估 IAPS 的差异性的探索性实证研究, 研究结果显示对于年轻成人两种评估方式在愉悦度、唤醒度和控制度皆有良好的内部一致性且评估的结果没有明显差异, 支持两种方式都可以在年轻成人评估 IAPS 时使用, 增进 IAPS 使用上的选择性, 但因为控制度在个体间的变异较大, 故建议未来其他研究者可依照实验目的斟酌是否要加入控制度的评估。本研究未能对同样适用 SAM 量表评估的儿童与临床族群进行九点量表的评估, 故对于欲将本研究结果应用于上述对象前应进行验证。

Table 5. Differences and correlation between two methods
表 5. 九点量表与 SAM 量表的差异与相关

	评估方式	平均数(标准偏差)	<i>t</i>	<i>r</i>
愉悦度	九点量表	4.71(0.26)	-2.51	0.72**
	SAM 量表	4.80(0.22)		
唤起度	九点量表	5.62(0.87)	0.19	0.80**
	SAM 量表	5.60(0.85)		
控制度	九点量表	4.91(0.91)	0.29	0.39*
	SAM 量表	4.86(0.79)		

注: * $p < 0.05$, ** $p < 0.01$ 。

Table 6. Cronbach's α of pleasure by three valences of IAPS
表 6. 依价性分类后之愉悦度内部一致性系数

	九点量表	SAM
负性图片(32 张)	0.93	0.94
中性图片(8 张)	0.72	0.48
正性图片(32 张)	0.88	0.90

致 谢

感谢协助本研究之教授及匿名审查委员以及本研究所有协助之研究参与者。

参考文献 (References)

- 王莹, 许晶, 张炳蔚, 冯霞(2008). 国际情绪图片系统在 116 名中国老年人中的试用. *中国心理卫生杂志*, 22 期, 903-907.
- 胡少华, 魏宁, 郭文滔, 胡健波, 谭云飞, 许毅(2005). 中国和美国健康成人情感反应差异的跨文化研究. *中国临床心理学杂志*, 13 期, 265-267.
- 蒋世光, 谭伟象, 花茂琴, 陈毓兰, 张兆贤(2012). 国际情绪图片系统在台湾年轻成人的适用性与其分类方式探讨. *中华心理学期刊*, 54 期, 495-510.
- Backs, R.W., Silva, S.P.D., & Han, K. (2005). A comparison of younger and older adults' self-assessment manikin ratings of affective pictures. *Experimental Aging Research*, 31, 421-440.
- Bradley, M.M., Codispoti, M., Cuthbert, B.N., & Lang, P.J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. [Research Support, U.S. Gov't, P.H.S.]. *Emotion*, 1, 276-298.
- Bradley, M.M., & Lang, P.J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49-59.
- Bradley, M.M., & Lang, P.J. (2007b). The International Affective Picture System (IAPS) in the study of emotion and attention. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 29-46). New York: Oxford University Press.
- Grühn, D., & Scheibe, S. (2008). Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age? *Behavior Research Methods*, 40, 512-521.
- Kline, P. (2000). *The handbook of psychological testing* (2nd Edition). London: Routledge.
- Lang, P.J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in Mental Health Care Delivery Systems* (pp. 119-137). Norwood: Ablex.
- Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (1997) In P. J. Lang, R. F. Simons, & M. Balaban, Eds., *Attention and Orienting, Sensory and Motivational Processes*. Hillsdale: Erlbaum, 97-135.

- Lang, P.J., Greenwald, M.K., Bradley, M.M. & Hamm, A.O. (1993). Looking at pictures: Evaluative, facial, visceral, and behavioral responses. *Psychophysiology*, 30, 261-273.
- Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). International Affective Picture System (IAPS): Affective rating of pictures and instruction manual Technical Report A-6. Gainesville: University of Florida.
- Mehrabian, A., & Russell, J.A. (1974). The basic emotional impact of environments. *Perceptual & Motor Skills*, 38, 283-301.
- Osgood, C., Suci, G., & Tannenbaum, P. (1957). The measurement of meaning. Urbana: University of Illinois.
- Silva, J.R. (2011). International Affective Picture System (IAPS) in Chile: A cross-cultural adaptation and validation study. *Terapia Psicológica*, 29, 251-258.
- Wundt, W.M. (1894). Lectures on human and animal psychology. Whitefish: Kessinger Publishing.