

《心理学报》审稿意见与作者回应

题目：影响动作图片命名的因素

作者：陈永香 朱莉琪

第一轮

审稿人 1 意见：本研究考察了动作图片命名的因素，研究发现命名一致性、熟悉性和视觉复杂性是影响动作图片命名的重要变量，未发现词频、习得年龄、词长等变量的作用。选题有一定的意义，存在问题如下：

意见 1：从题目来看，本文的研究的重点影响动作图片命名的因素，但在问题提出部分研究者首先提出的问题是“目前国内还没有标准化的动词图片库，本研究拟对国际上常用的一个动词图片库进行标准化研究”，而且文中有不少相关内容的论述。如果本文的目的之一是建立标准化的图片库，那么论文题目不能涵盖本文研究内容；另外，如果是对图片库进行标准化研究，并建立相应的常模，那么本研究的选取的样本量（被试量）是否足够？

回应：本研究的题目参考张清芳和杨玉芳（2003）发表在《心理学报》的论文《影响图画命名时间的因素》，该研究对名词图片库做了标准化，同时也考察了影响名词图片命名的因素。本研究同样涵盖两方面的内容，一是对动词图片库的标准化，二是也考察了影响动词图片命名的因素。为了和该研究相区别，本研究暂时命名为《影响动作图片命名的因素》。

在被试量方面，本研究参考了张清芳和杨玉芳（2003）的研究，以及 Liu, Hao, Li, & Shu(2011)最近对名词图片库所做的标准化研究，每组有效被试 34-41 人不等。

意见 2：词频数据来自《现代汉语通用词表，2003》，需要介绍该词表，并给出图片名称对应的频率信息，例如平均数、标准差等，并给出该变量与其他变量的相关系数等信息。

回应：这是目前考察汉语加工的研究者经常引用的一个资料，其引用方式为(已列在参考文献中)：Chinese Linguistic Data Consortium. (2003). 现代汉语通用词表 [Chinese Lexicon](CLDC-LAC-2003-01). Beijing, China: State Key Laboratory of Intelligent Technology and Systems, Tsinghua University, and Institute of Automation, Chinese Academy of Sciences.

该资料为清华大学、中国科学院自动化所联合开发的一个数据库，规模约 10 万词，包括词、拼音和词频等信息。因该资料已被研究者较广泛地应用，故本研究直接引用，未在文中详细介绍。其他已发表论文对该资料的引用如：Qu, Q., Damian, M. F., Zhang, Q., & Zhu, X. (2011). Phonology Contributes to Writing Evidence From Written Word Production in a Nonalphabetic Script. *Psychological science*, 22(9), 1107-1112.

图片名称对应的词频信息已在文中说明。“本研究中的词频来自现代汉语通用词表（2003）。本研究所有动词的词频均值为 2265.3 次/百万词，标准差为 7847.1 次/百万词，范围是（0.91-81700）次/百万词，偏度为 6.688。为纠正偏态分布，本研究参考 Stadthagen-Gonzalez & Davis（2006）的研究等，采用词频对数来考察词频和其他变量之间

的关系（如表2）。”（见 P16。）在表2中，提供了词频对数和其他变量之间的相关系数。

意见3：研究结果：大量研究发现，习得年龄是影响图片命名的重要因素，无论命名名词图片，还是命名动词图片，习得年龄都是影响图片命名的重要因素。这一研究结果稳定地表现在目前几乎所有的研究文献中，因此不知本研究为什么没有发现习得年龄效应。对于频率，多数研究发现频率是影响图片命名的因素（虽然有少数研究未发现），本研究亦未发现该效应。

回应：本研究遵循了有关图片命名和评定研究的程序，结果发现习得年龄和频率对动词图片命名的影响不显著，这和名词研究结果不太一致。在讨论部分，我们对这两个结果进行了探讨。对此，我们尝试做出的解释是：

“在动词图片命名任务中，由于图片特征到语义通达过程吸收了较多的变异”，而习得年龄和词频效应主要作用于“语义-语音”通达阶段，故在逐步回归分析中，习得年龄和词频效应不显著。当然，这种解释仍需要更多研究来验证。（见 P18。）

意见4：讨论部分：根据图1，作者认为 AoA 的作用体现在语义通达后的语音加工阶段，这有待商榷。大量研究发现，习得年龄效应出现认知加工的多个阶段，例如图片识别，语义提取、Lemma 选择和语音加工等多个阶段。如果 AoA 的作用体现在语义通达后的语音加工阶段，那么如何解释前人发现的图片识别、图片语义分类等不需要语音产出任务中的习得年龄效应。

回应：首先，已有图片识别、图片语义分类研究主要考察的是名词图片；其次，即便是对于动词研究，图片命名任务也相对特殊，因为其加工难度更大。在动词图片命名任务中，被试需要从一个静态的、较抽象的图片中提取和想象动态的动作信息，进而通达某一动词的语义。这一加工过程变异较大、且需要耗费较长的加工时间，因而可能吸收了主要的变异。

本研究认为 AoA 效应可能主要体现在“语义-语音”通达的阶段，这确实是一种推测。推测理由是“在词汇习得过程中，口语习得主要是在语义和语音之间建立联结关系”。还有一个理由是，在动词图片命名中，图片中的场景和被试童年时学习的动作场景可能是有出入的，因此即便在图片特征-语义通达阶段有习得年龄效应，在动词图片命名任务中，该线索的可利用度也会很低。相比之下，被试对图片的“熟悉性”则显著影响图片命名时间。

审稿人2意见：本研究基于 IPNP 图片库考察了汉语大学生对动词图片的命名及其诸多因素的影响。结果发现了动词图片命名区别于名词图片命名的特征，为后续的动词图片加工研究提供了材料。这一研究具有十分重要的应用推广价值。但尚存在一些问题值得商榷。

意见1：本研究中的“词汇获得年龄”指的是动词图片的获得年龄还是动词图片所对应词汇的获得年龄？这两个指标有无区别？作者在文中提出，“图片习得年龄评定。在图片命名的基础上，将图片名称和图片同时呈现给被试，请被试在7点量表中评定该动词的口语习得年龄”。在作者在附录中是这样表述的：“我们所说的“认识”是指您听到这个词后知道它的意义，在生活中或书籍、报纸、电脑等地方看到与该图片类似的动作时能认出这个动词，知道它是什么意思（不是识字，而是口语的理解与表达）。”这两处指的显然不是相同的评定对象。

回应：本研究中的词汇习得年龄是指“动词”的习得年龄。同时，由于词汇语义具有一定的模糊性，为了方便被试了解具体是怎样的动作，我们在评定词汇年龄的同时提供了动词图片供被试参考。比如，汉语动词的“吹”既可以表示“吹头发”（dry）又可以表示“吹蜡烛”（blow），而在英语中这是两个不一样的动词。因此，为了方便以后和有关的英语研究作对照，我们在评定词汇习得年龄的时候不仅提供了词汇、还提供了相应的图片作为参考。

意见 2：作者在文中提到“被试表达了图片中的动作，或者说出的短语中包含恰当的核心动词（此时反应时无效）”。此处作者对“命名正确标准”和“有效反应时”的界定都不明确，有效反应时应具备怎样的标准？

回应：已删除“(此时反应时无效)”字样。如何删除反应时的数据，在方法部分有很清楚的界定，如“删除反应时数据的标准：1）删除命名为动词词组的反应时数据”（见 P13。）

意见 3：根据该标准未删除一名被试（所有被试的反应时平均值为 1925.5ms，标准差为 1691.7ms）标准差为什么会如此偏大？

回应：感谢审稿人的提醒！作者检查了一下，发现原来计算的是所有被试、所有未经删除的 trials 的整体反应时和标准差，因而标准差较大。作者和张清芳老师课题组的博士生讨论了一下，澄清了两个问题。首先，这个计算应该是先算每个被试的平均值，然后计算所有被试的平均值和标准差（这样计算出来的标准差为 241.1ms）；其次，作者和小同行讨论后发现，现在很多研究者倾向于认为，不能单纯因为某个被试反应太快或太慢而删除数据（前提是受试遵从指导语且有一定的正确率）。因此，本研究去掉了这一删除数据的标准，在正文中删除了“（所有被试的反应时平均值为 1925.5ms，标准差为 1691.7ms）”等字样（见 P14）。

意见 4：删除反应时数据的标准：1）删除命名为动词词组的反应时数据。请进行举例说明。以便于读者更好地理解作者的删除标准。

回应：已增加举例，如“吹蜡烛”。

意见 5：删除错误命名的反应时。何为错误命名？正如作者在附录中给出的一样，“逮捕”和“搜查”均为正确命名，但这两个词汇似乎并不是表示同一个意义。对于这样的情况，如何处理？

回应：同一动作场景对应的是哪个动词，本身具有一定的模糊性，因而动词的命名一致性相对比较低。本研究的研究者母语为汉语，正确命名的判定标准是认为被试提供的动词可以适当描述图中的动作场景。因此，对正确命名的判定确实有一定的主观成分。为了让读者对此有自己的评判，我们在附录中同时提供了命名次数最高的词汇、以及我们认为的概念一致词。

意见 6：表 1 中的 Skew 代表什么？

回应：Skew 应为“偏度”，已改为中文。在 Liu 等人（2011）的研究中也同样报告了偏度。

意见 7：“词频效应在汉语名词图片命名中并不稳定地出现，而本研究发现，动词图片命名中的词频效应不显著”。这一结论怎么得出的？

回应：该描述不太准确，已修改为“已有研究发现，词频效应在汉语名词图片命名中并不稳定（如，张清芳，杨玉芳，2003；Liu et al., 2011）。而本研究发现，汉语动词图片命名中的词频效应不显著。”（见 P18。）

意见 8：“拳击”是名词还是动词？在汉语中存在很多兼类词，对于这类词如何处理？它的命名一致性等指标是否会受到兼类的影响？在选入标准图片库时，要作何处理？

回应：本研究是对国外的一个动词图片库做的汉化研究，其中的所有动词都是“动作词汇”，对应于具体的动作。关于“兼类词”，本研究未考虑这一因素。事实上，有国内语言学家（如，沈家煊，2007）认为，所有的汉语动词都可以用做名词。汉语动词没有词形变化，如汉语中的“跑”在英语中可能对应于“run”“running”等不同的形式。而动名词形式是可以充当主语的，从语法角度来说它相当于名词。

意见 9：本文在文中注解参考文献时，格式有问题，如(张振军, 陈宝国, & 丁国盛, 2011)。

回应：根据《心理学报》文献著录要求，已对引文和参考文献格式进行了相应的修改。在中文文献的引文和参考文献中，删除了“&”符号。

审稿人 3 意见：本研究为汉语动词加工研究建立图片库，是一项奠基性的工作；研究的价值显而易见。研究的方法恰当，数据可靠，讨论深入。

意见 1：文章中几组被试来自于同一群体，保证了被试的同质。但是，由于命名任务与熟悉性评定任务由两组不同被试完成。计算相关时，如果以被试为随机变量时，则两组被试需要先进行匹配。如果以项目（图片）为随机变量，在论文中应说明。

回应：首先感谢审稿专家的肯定。“本研究参考张清芳和杨玉芳（2003）以及 Liu 等人（2011）的研究，以项目（即图片）为随机变量做相关分析和回归分析。”已在结果部分补充说明，见 P15。

另外，本次所做的其他修改包括：1) 按要求将图表的字体调为六号字宋体。2) 将所有小数点之前的零补齐。所有修改详见 word 文档修订模式。再次感谢所有审稿人及编辑部老师的宝贵意见！

第二轮

审稿人 1 意见：作者很好地回答了上轮提出的问题。还有以下问题建议考虑。

意见 1：被试组成的介绍不便于读者理解，建议介绍的更为清晰些。共有多少人参与实验，分成几组，每组多少人参与什么任务。

回应：已修改如下，同时见正文 P13。

“有效被试为来自北京地区高校、自愿参加实验的大学生共 148 名。所有被试的母语均为汉语普通话。其中 36 人（男 12 人）参加图片命名测试，40 人（男 16 人）参加熟悉性和复杂性的评定（一半被试先评熟悉性，一半被试先评复杂性，随机分配），36 人（男 13 人）参加表象一致性的评定，36 人（男 17 人）参加动词口语习得年龄的评定。另有 6 名被试因

为规律作答、或者不作区分地答题而被删除。所有被试在实验完毕获得少量报酬。”

意见 2: “图片习得年龄”表示“该动词的口语习得年龄”是否合适。

回应: 已将“图片习得年龄”修改为“动词的口语习得年龄”，见 P14。

意见 3: AOA 效应在言语产生中的发生阶段是否有更为充足的证据，或者本研究中的解释是否更倾向于一种假设或推论。建议作者在论述这一问题时，既要照顾到名词图片产生的相关研究，同时考虑动词言语产生的可能性影响因素。

回应: 就这一项研究而言，AOA 效应在言语产生中的发生阶段确实是本研究的推论，因此我们把它放在讨论部分，并未把它当成一项“研究结果”。讨论部分的图一是一个示意图，它可以较直观地解释我们的研究结果。事实上，名词图片命名也会有类似的过程（即先从图片特征提取语义、然后通达语音并输出）。不同的是，名词图片表示的是静态的客体，从图片特征到语义通达可能相对更加容易（体现在名词图片命名时间更短）。

值得一提的是，我们手头还有一些正在整理、尚未发表的数据表明，在动词词汇命名任务中，口语习得年龄（AoA）效应的确可能发生于“语义-语音”阶段。相比之下，书面语习得年龄（AoL）效应则更可能发生于“字形-语音”这一直接通路，口语和书面语两种不同的习得年龄效应出现了分离的现象。这些证据表明我们在本研究中的推论可能是合理的。

第三轮 编委复审意见

论文选题有意义，作者也较好地回答了审稿人的问题。存在以下问题：

意见 1: 问题提出部分应重写，作者对动词与名词之间的结构、加工等方面的区别关注不够，国内已有文献引证方面严重不足，如：

- (1) 陈新葵等.中文身体动词的语义特异性加工.华南师范大学学报，2013 年第 4 期。
- (2) 陈新葵等.影响汉语动词名词识别因素的回归分析.心理科学，2010 年第 4 期。
- (3) 陈新葵等.汉语名词动词认知的影响因素比较.华南师范大学学报，2008 年第 4 期。
- (4) 陈新葵等.动、名词损伤的功能定位.应用心理学，2004 年第 1 期。
- (5) 陈新葵等.动、名词的认知及其脑机制.华南师范大学学报（社会科学版），2005 年第 6 期。
- (6) 韩在柱等.汉语名词特异性损伤的个案研究.心理科学，2005 年第 4 期。
- (7) 韩在柱等.动词特异性损伤的原发性进行性失语个案.中国临床康复，2005 年第 4 期。
- (8) 石东方等.动词信息在汉语句加工早期的作用.心理学报，1999 年第 1 期。
- (9) 汉语及物动词词典中宾语信息的表征.心理学报，2004 年第 4 期。

回应: 非常感谢编委专家的提醒！作者下载并阅读了这些文献的全文，其中和本文最相关的五篇实证研究文献（对应上述编号 1、2、6、7、9）已经加入了本文的问题提出和讨论部分，详见正文及参考文献修订模式。

意见 2: 大量研究表明，AOA 效应不仅仅发生在语义到语音的阶段，国内的研究也支持语义假设，请修改。

回应：感谢编委专家的意见。在讨论部分我们已做相应的修改，删除了“在词汇习得过程中，口语习得主要是在语义和语音之间建立联结关系，因而 AoA 效应可能主要体现在“语义-语音”通达的阶段”等字样，修改了图 1，并补充了对相应文献的讨论。

“在 Liu 等人（2011）的研究中，评定的名词口语习得年龄与图片命名反应时之间存在 0.475 的正相关 ($p < 0.01$)。在本研究中，评定的动词口语习得年龄和图片命名反应时之间仅存在 0.137 的正相关 ($p < 0.05$)；当习得年龄和其他因素共同预测命名反应时时，习得年龄的影响并不显著、未能进入逐步回归方程。陈新葵和张积家（2010）通过词汇判断任务发现，习得年龄对名词识别的影响大于对动词识别的影响。综合考虑以往研究和本研究的结果可以发现，习得年龄效应的大小随词类和实验任务不同而有差别，习得年龄对汉语名词加工的影响要大于对动词加工的影响。”（见 P21。）

本研究只是关于动词图片加工的初步研究，而动词图片命名包括词条选择、语音编码等复杂的过程。“关于习得年龄效应具体作用于词汇加工的哪一阶段，仍需更多的研究来验证”（见 P21）。

意见 3：还有一些重要的影响因素未涉及，如情境可获得性，是否是身体动词，如果是，身体部位有何影响，建议增加这方面的事后分析。

回应：本研究主要对一个国外的动词图片库做了标准化工作，其中评定的主要指标也主要参考了国内外进行图片库标准化研究的文献（如张清芳，杨玉芳，2003；Liu, Hao, Li, & Shu, 2011）。

词汇的心理语言学指标有很多都是高度相关的。关于“情境可获得性”这一指标，作者查阅了相关的文献，发现它常被用于书面词汇的评定，而本研究的主要研究材料是动词图片，图片本身即提供了一个特定的情境，如果在图片基础上再评定情境可获得性可能会有一定的污染。在今后进一步的研究中，我们可以尝试去评定图片所对应的动词名称的情境可获得性，这又是另一个值得研究的题目了。

由于我们的材料是一个动词图片库，都是一些方便以图示来表达的动词（见另外提供的附件），其中大部分都是身体动词。关于身体部位对动词加工有何影响，前人已有一些研究成果（如陈新葵，张积家，2013；该文献已引用）。今后我们可能考虑以此图片库为材料，进一步考察身体部位影响动词加工的具体机制。

意见 4：这项研究的意义就在于提供了一个图片库，并有各项指标，因此，请提供完整的图片和各项评定指标，便于今后的研究者采用。

回应：完整的图片和各项评定指标见作者另外提供的两个附件，供评审专家考察。今后如有研究者愿意采用，我们也十分乐意提供完整的资料。