《心理学报》审稿意见与作者回应

题目：《二语情绪概念理解中的空间隐喻》

作者：沈曼琼 谢久书 张昆 李莹 曾楚轩 王瑞明

第一轮

审稿人 1 意见：

意见 1：2.3 结果与分析部分中“根据以往数据处理的方法(Meier, Robinson, 2004)，删去词语判断准确率低于 75%的被试 10 名(王穗苹, 迟毓凯, 王瑞明, & 吴岩, 2005)”，但是，Meier 和 Robinson（2004）的研究中，并没有删去准确率低于 75%的被试，而是“Latencies were then log-transformed to normalize their distribution (Ratcliff, 1993). Next, we replaced trials that were 2.5 SDs below or above the grand latency mean”。另外，值得思考的是，在只有 30 个被试参加的实验中删去 10 个被试是否合适？Meier 和 Robinson（2004）的研究中，只是“Inaccurate trials were dropped from the analysis. ”；在 Outlet 等人的研究中，是“Trials with errors in any of the two tasks (8.79%, 743 trials) were excluded from the latency analysis.”。是否这两个研究中用的方法会更为合理、可行？

回应：非常感谢审稿人的宝贵意见，为了验证实验结果是否稳定并增加有效被试量，同时针对审稿人提出的补充相关变量实验结果的要求，我们重新收集了实验数据。先前实验中的被试只是通过大学英语四级，所以对二语情绪词汇的判断正确率低，导致删除的被试较多。新实验中使用的被试均通过了大学英语六级考试。实验结果中有效被试人数显著增多，最主要的是，实验结果与之前版本的结果相同，证明本实验结果稳定，具有可重复性。同时，在按照不同数据标准处理时，实验结果基本保持稳定。本研究报告中最终选择的数据处理标准和依据具体如下：

（1）删除准确率低于75%的被试。Meier and Robinson (2004) 在数据处理时没有删除被试，主要因为被试在实验中是使用母语进行反应。所以，被试的正确率都很高，无需删除被试。然而，本研究是一项跨语言的研究，被试使用二语进行反应。因此，有些被试的正确率太低。根据先前的有关研究，错误率高于20%的被试一般需要删除(Gozli, Chasteen, & Pratt, 2013)。Connell（2007）在语言理解的颜色表征研究中指出，被试的错误率高于25%的必须删除；而在知觉加工影响概念加工的研究中，Dantzig, Pecher, Zeelenberga, & Barsalou (2008)

删除了准确率低于70%的被试。所以，本研究参考前人的方法和本研究的具体情况，删除了准确率低于75%的被试。根据这一标准，实验1使用32名被试，最后有效被试27名（删除5名），实验2使用40名被试，最后有效被试36名（删除4名）。

（2）删去两类实验任务反应错误的数据。本研究的因变量是任务一（空间位置判断任务）中的反应时和准确率。任务二（情绪概念词一致性判断）的作用主要有两个，一个是掩蔽实验的真正目的，预防被试效应对实验结果的影响；另一个是辨别被试对前面的情绪词是否进行了认真的加工。根据审稿人的意见，并参考以往研究（Ouellet, Santiago, Funes, & Lupi áñez, 2010），我们在修改稿中采用了这一标准，即删去两类实验任务反应错误的数据，然后再进行统计分析，从而使实验结果更加的纯净。

（3）剔除3个标准差以外的极端数据。首先，舒华教授的《心理学研究方法》中指出，“原则上，我们尽量使用所有的有效数据来进行统计分析。但是根据样本分布，在进行统计检验分析之前，往往需要先将那些被认为不是来自于该样本总体的数据剔除掉”。在经验操作上，可以按照平均值加减三个标准差的原则来处理。在这个范围之外的数据，我们称之为极端数据，统计上有理由认为它不是来自抽样样本所代表的总体，需要对这部分数据进行特殊处理。其次，在前人的研究中，也多采用剔除3个标准差以外的极端数据的标准(Vega, 2013; Pecher, Dantzig, Boot, Zanolie, & Huber, 2010; Zwaan & yaxley, 2003; 莫雷，伍丽梅，王瑞明，2007)。本研究中，我们已经删除了错误率较高的被试，并删除了两类实验任务反应错误的数据，因此，相对于剔除2.5或者2.0个标准差的极端数据的处理方法，剔除3.0个标准差可以保证更多的有效数据参与分析，因此本研究最终采用平均值加减3个标准差的原则来剔除极端数据。

我们在修改稿中对最终使用的数据处理标准进行了明确说明。（请见 P5-8）

意见 2：对于每一个 trial，被试需要有两个反应，一是判断白点出现的位置，二是判断情绪词的正负性。显然，结果中词语判断的正确率是有关第二个判断的结果，而反应时和正确率是第一个判断的结果。但是，由于在现在对结果的描述中，没有很明确地对二者进行区分，很容易让读者误解。

回应：感谢审稿人的宝贵意见。在修改稿中，我们进一步在实验设计、结果与分析部分中明确指出本研究的因变量是空间位置判断任务中的反应时和准确率。同时，我们也在实验程序中进一步明确了实验中的两个任务。本实验中，每个被试有两个反应，任务一是判断白点出现的位置，任务二是情绪词一致性的判断。任务二中，被试的任务是判断屏幕上出现的“？

积极？”或者“？消极？”与之前出现的情绪词表达的情绪是否一致。具体我们在修改稿中做了修改（见 P6、8）。

意见3：建议补充词语正负性判断的结果，以表明该结果确实是由情绪词的正负性引起的。

回应：感谢审稿人的意见。正如问题 2 中所提到的，本实验中，每个被试有两个反应任务，任务一是判断白点出现的位置（如上或下），任务二是情绪词一致性的判断。本研究的主要目的是考查情绪概念理解中空间信息的激活情况，所以因变量是空间位置判断任务中的反应时和准确率，最后统计分析时主要呈现的也是该实验结果。在任务二中，被试的任务是判断屏幕上出现的“？积极？”或者“？消极？”与之前出现的情绪词表达的情绪是否一致，主要是保证被试能够认真理解前面出现的情绪词。该任务的判断结果主要用来筛选被试和有效数据。为了避免漏掉一些重要信息，我们在重新实验时，结合两位审稿人的意见，增加了情绪概念词这个因素，该因素有两个水平：积极情绪概念词和消极情绪概念词。因此，实验一的实验设计变为 2 (情绪概念词：积极和消极)×2 (空间位置方向：水平和垂直)×2 (隐喻一致性：一致和不一致)的三因素被试内设计。实验二的实验设计变为 2 (情绪概念词：积极和消极)×2 (隐喻一致性：一致和不一致)的两因素被试内设计。最后结果中更加详细地报告了两个实验的实验结果。最主要的实验结果与之前的结果完全一致，证明了本实验的假设以及实验结果的可重复性和稳定性。具体请见我们在修改稿中做的修改（见 P6-8）。

意见4：以往的研究都会具体报告上和下或者左和右分别引起的反应时和正确率变化，并且，上下或者左右的结果往往有差异。而在本研究中，则只是报告了一致和不一致的效应，似乎遗漏了较为重要的数据，建议进行添加。

回应：感谢审稿人的意见。在本研究中，我们关注的核心问题是在垂直和水平这两个不同方向上，二语情绪概念词是否存在空间隐喻性，以及垂直方向的空间隐喻和水平方向的空间隐喻的强度差异。因此，比较垂直与水平方向上的隐喻，是本研究的重点。目前专门研究上下（即垂直方向）或左右（即水平方向）的研究已经很多，研究结果也比较一致。在上下方向上，研究发现人们倾向于分别利用上和下的空间来表达积极和消极效价，比如使用“high on life”形容一个人很开心，而使用“down in the dumps”形容一个人心情很差(Lakoff & Johnson, 1980, 1999)。Chasteen (2010)等人先在注视点位置出现一个单词，之后出现目标刺激 X/O，这个刺激可能出现在注视点上面，也可能出现在注视点下面，被试的任务就是对目标刺激做

反应。实验结果表明，当前面出现的单词与后面的目标刺激有隐喻一致时，被试的反应更快。类似的，Meier and Robinson（2004）的实验中，情绪词呈现在注视点上面或者下面，被试直接对单词的效价进行反应。研究结果表明，积极情绪词出现在屏幕上方时会判断的更快些，消极情绪词出现在屏幕下方会判断的更快些；并且，积极情绪词会促进上方空间刺激的识别，消极情绪词会促进下方空间刺激的识别。在道德概念的研究中，抽象的汉语道德概念可以激活垂直空间的具体概念(上或下)，形成汉语道德概念的垂直空间隐喻，即“道德是上，不道德是下”的感觉-动作经验(王锃, 鲁忠义, 2013)。在左右方向上，研究发现与身体右半边有关的动作与积极情感有关，比如收缩右手的肌肉、观看右视野刺激，而一样的动作，用左边身体来做，则与消极情感有关(Davidson,1992; Natale, Gur, & Gur, 1983)。由于比较上和下、左和右的差异的研究比较多，结果也比较一致，本研究最主要的的目的不在于比较上和下、左和右的差异，而是从跨语言和跨通道的角度上，探讨情绪概念词是否存在垂直或者水平方向上的隐喻。因此，在修改稿中，我们没有直接报告上和下、左和右的差异。实际上，本研究中，在水平位置方向中，隐喻一致指的是积极情绪词后面紧跟右边的白点或者是消极情绪词后面紧跟左边的白点，隐喻不一致指的是积极情绪词后面紧跟左边的白点或者是消极情绪词后面紧跟右边的白点；在垂直方向中，隐喻一致指的是是积极情绪词后面紧跟上边的白点或者是消极情绪词后面紧跟下边的白点，隐喻不一致指的是积极情绪词后面紧跟下边的白点或者是消极情绪词后面紧跟上边的白点。所以，不同空间位置上情绪概念词（积极和消极)2×隐喻一致性（一致和不一致)2的检验结果也能在一定程度上体现出左右和上下的差异。先前研究是在不同实验中分别比较水平位置和垂直位置的空间隐喻情况，所以可以用上下、或左右直接作为研究变量，但本研究中位置作为被试内变量在同一个实验中，并且研究主要关注二语情绪概念词是否存在空间隐喻性，以及垂直方向的空间隐喻强度和水平方向的空间隐喻的差异，所以用隐喻一致性作为研究变量更为妥当。我们在修改稿对此进行了进一步的说明（请见 P4-8）。

意见 5：本文的摘要和行文中有个别多字或漏字的现象。例如，正文第 2 也页第一段中“根据知觉符号理论所认为的内在符号与外在物体原型间具有类似性关系”，似乎可以去掉“所认为的”。

回应：感谢审稿人细心的建议，为了更准确的表达，按照您的建议，我们将该部分修改为“知觉符号理论认为内在符号与外在物体原型间具有类似性关系，Stanfield 和 Zwaan (2001)根据该理论开创性地使用句图匹配范式对知觉符号理论进行了实验验证”。此外，我们通过大声

朗读再次对文章的行文表述进行了修改和润色。

审稿人 2 意见：

意见1：材料匹配的具体信息应提供，如字母数，频率（熟悉性）、效价、唤醒度和控制性等。特别是这80个词似乎又进一步进行了评价，是否所有材料都入选，究竟多少个材料进行了实验不清楚。

回应：感谢审稿人的意见。材料匹配的具体信息已经在文章中进行了补充。我们首先从英语情绪词库(Affective Norms for English Words, ANEW)(Bradley, Lang, 1999)，并参照汉语情感词系统 (Chinese Affective Words System, CAWS)(王一牛, 周立明, 罗跃嘉, 2008)，综合熟悉性、效价、唤醒度和控制性指标选择了80个英语情绪概念词汇作为实验材料，其中积极情绪词40个，消极情绪词40个。在前人的关于情绪概念的研究中，实验材料主要报告效价和熟悉度两个维度（郑志伟, 2013; 郭晶晶, 2011; de la Vega, Dudschig, De Filippis, Lachmair, & Kaup, 2013），因此实验前为了保证实验材料的适用性，另外20名在校大学生对这些词汇的熟悉性和效价又进行了评定。根据 Bradley(1999)和张钦(2013)等人的操作方法，我们对80个词汇进行9级评定。熟悉性指的是对词汇的熟悉程度,1表述非常不熟悉，9表示非常熟悉。效价指情绪词是积极还是消极，当被试觉得该词汇表达非常快乐时，评定等级最低为1，当被试觉得该词汇表达非常不快乐、愤怒时，评定等级最大为9。评定结果表明，积极情绪词和消极情绪词熟悉性差异不显著、情绪效价差异显著。此外，对积极词和消极词的笔画数进行 *t* 检验，结果发现积极情绪词和消极情绪词字母数差异不显著。在修改稿中，我们进行了详细的补充（见 P5）。

意见2：建议结果的报告进一步分积极词、消极词，上下及左右来报告。合并为一致，不一致，有些笼统。

回应：感谢审稿人的意见。该意见跟第一位审稿人的意见基本类似，前面的回复中已经有所说明，在此再进一步说明如下：

一方面，我们在修改稿中增加了情绪概念词这个因素。该因素有两个水平：积极和消极。因此，实验一的实验设计变为2 (情绪概念词：积极和消极)×2 (空间位置方向：水平和垂直)×2 (隐喻一致性：一致和不一致)的三因素被试内设计。实验二的实验设计变为2 (情绪概念词：积极和消极)×2 (隐喻一致性：一致和不一致)的两因素被试内设计。实验最主要的结果与之前的结果一致，证明了本实验的假设以及实验结果的可重复性和稳定性。具体为，实验一的

跨语言研究结果发现，在单通道条件下，二语情绪概念理解中上下空间信息可以激活，即存在上下空间隐喻，而在水平空间位置条件下没有发现空间隐喻现象。在实验二的跨通道研究中，同样发现了垂直空间位置条件下，二语情绪概念存在空间隐喻。另外，我们也发现了情绪概念词类型跟隐喻一致性的交互作用，具体我们在修改稿中做了修改。（见 P6-8）

另一方面，在本研究中，我们关注的核心问题是在垂直和水平这两个不同方向上，二语情绪概念词是否存在空间隐喻性，以及垂直方向的空间隐喻和水平方向的空间隐喻的强度差异。因此，比较垂直与水平方向上的隐喻，是本次研究的重点。目前专门研究上下（即垂直方向）或左右（即水平方向）的研究已经很多，研究结果也比较一致。在上下方向上，研究发现人们倾向于分别利用上和下的空间来表达积极和消极效价，比如使用“high on life”形容一个人很开心，而使用“down in the dumps”形容一个人心情很差(Lakoff & Johnson, 1980, 1999)。Chasteen (2010)等人先在注视点位置出现一个单词，之后出现目标刺激 X/O，这个刺激可能出现在注视点上面，也可能出现在注视点下面，被试的任务就是对目标刺激做反应。实验结果表明，当前面出现的单词与后面的目标刺激有隐喻一致时，被试的反应更快。类似的，Meier and Robinson（2004）的实验中，情绪词呈现在注视点上面或者下面，被试直接对单词的效价进行反应。研究结果表明，积极情绪词出现在屏幕上方时会判断的更快些，消极情绪词出现在屏幕下方会判断的更快些；并且，积极情绪词会促进上方空间刺激的识别，消极情绪词会促进下方空间刺激的识别。在道德概念的研究中，抽象的汉语道德概念可以激活垂直空间的具体概念（上或下），形成汉语道德概念的垂直空间隐喻，即“道德是上，不道德是下”的感觉-动作经验(王锃，鲁忠义, 2013)。在左右方向上，研究发现与身体右半边有关的动作与积极情感有关，比如收缩右手的肌肉、观看右视野刺激，而一样的动作，用左边身体来做，则与消极情感有关(Davidson,1992; Natale, Gur, & Gur, 1983)。由于比较上和下、左和右的差异的研究比较多，结果也比较一致，本研究最主要的的目的不在于比较上和下、左和右的差异，而是从跨语言和跨通道的角度上，探讨情绪概念词是否存在垂直或者水平方向上的隐喻。因此，在修改稿中，我们没有直接报告上和下、左和右的差异。实际上，本研究中，在水平位置方向中，隐喻一致指的是积极情绪词后面紧跟右边的白点或者是消极情绪词后面紧跟左边的白点，隐喻不一致指的是积极情绪词后面紧跟左边的白点或者是消极情绪词后面紧跟右边的白点；在垂直方向中，隐喻一致指的是是积极情绪词后面紧跟上边的白点或者是消极情绪词后面紧跟下边的白点，隐喻不一致指的是积极情绪词后面紧跟下边的白点或者是消极情绪词后面紧跟上边的白点。所以，不同空间位置上情绪概念词（积极和消极)2×隐喻一致性（一致和不一致)2的检验结果也能在一定程度上体现出左

右和上下的差异。先前研究是在不同实验中分别比较水平位置和垂直位置的空间隐喻情况，所以可以用上下、或左右直接作为研究变量，但本研究中位置作为被试内变量在同一个实验中，并且研究主要关注二语情绪概念词是否存在空间隐喻性，以及垂直方向的空间隐喻强度和水平方向的空间隐喻的差异，所以用隐喻一致性作为研究变量更为妥当。我们在修改稿对此进行了进一步的说明（请见 P4-8）。

意见3：研究者假设消极情绪与左侧一致，积极情绪与右侧一致，根据是什么？

回应：（1）以往研究发现，与身体右半边有关的动作与积极情感有关，比如收缩右手的肌肉、观看右视野刺激，而一样的动作，用左边身体来做，则与消极情感有关(Davidson,1992; Natale, Gur, & Gur, 1983)。Casasanto（2009）的研究表明，空间与效价之间潜在的联系是与身体的特异性有关的，当被试被问到买哪一个东西，选择哪一份工作，哪个外星物体更可信时，右利手的人倾向于选择右边的，而左利手的人倾向于选择左边的。此外，有研究使用 EEG 技术验证了左利手和右利手的情绪动机单侧化不同，右利手的左侧额叶与趋近动机有关，而右侧额叶与回避动机有关；而用 fMRI 技术发现了，让左利手和右利手的被试看一样的与手有关的动词，结果左利手的被试右半球激活，而右利手的被试左半球激活（Casasanto, 2011; Davidson, 1992）。Casasanto 的研究表明，这种潜在的空间与效价之前的联系是与身体的特异性有关，即右利手的被试认为右侧空间是积极的，而左侧空间是消极的；左利手的被试则表现出相反的模式。由于本研究的被试均为右利手，因此，根据先前的相关文献，本研究假设消极情绪与左侧为一致条件，积极情绪与右侧为一致条件。

（2）在英语表达中，“右”经常与“积极”相联系，而“左”经常与“消极”相联系。例如“the right answer（准确答案）”“my right-hand man（得力助手）”均表达积极的含义并与右边的空间相联系。相反，“out in left field（驴头不对马嘴）” 和“two left feet（笨手笨脚）”则表达了消极的含义且与左侧的空间相联系。同时，拉丁语的 right（右）和 left（左）对应 dexterand（幸运的，右侧的）和 sinister（阴险的，左侧的），而这些词分别又是英语单词“skillful” 和“evil”的词根（Casasanto, 2009）。

总之，关于情绪与左右空间方位的关系，前人已经有了一定的研究，并且研究结论也比较一致，因此我们的研究中，据此提出了有关假设。我们在前言部分对此做了进一步的说明（请见 P3-4）。

意见4：实验2的反应时之差仅有7毫秒左右，但统计显著，能否进一步核实数据，看是否可

靠。

回应：感谢审稿人的意见。我们核实了以往的数据，没有发现数据分析和统计中的错误。在重新收集数据（具体说明请见对第一位审稿人的答复）后，差异量仍然比较小，但是统计检验确实得到了显著的差异。同时，以往很多研究中，隐喻的一致与不一致的差异量也比较小，但是，统计检验亦显著。例如，关于时间表征与空间位置的研究发现，左侧空间促进过去时间的加工，右侧空间促进将来时间的加工。当目标刺激在左边时，被试对与过去相关的词汇反应比对与将来相关词汇的反应快5毫秒，当目标刺激在右边时，被试对与将来相关的词汇反应比对与过去相关词汇的反应快3毫秒，尽管这个差异比较小，但是无论在被试分析还是项目分析中，都出现了比较高的 *F* 值，效应都非常显著(Ouellet, Santiago, Funes, & Lupiáñez, 2010)。根据 Meier 的研究结果，积极情绪词出现在屏幕上方时被试的反应更快，差异值约30毫秒，而消极情绪词出现在屏幕下方时被试的反应更快，差异值约10毫秒(Meier & Robinson, 2004)。Gozli 等人将实验任务分为线索-目标探测任务和线索-目标辨别任务，并在任务中分离线索-目标一致任务和线索-目标不一致任务，探讨实验任务对视觉空间机制的影响是发挥促进还是抑制作用。结果发现，一致任务的反应时(M=426)显著快于不一致任务(M=435)，差异值仅9毫秒（Gozli, Chasteen, & Pratt, 2013）。因此，关于隐喻的研究，存在差异值较小，但是效应很稳定的现象，而本研究结果在经过核实数据和重新实验之后，确定结果稳定，数据可靠。（见 P8）。

第二轮

审稿人 1 意见：

意见 1：摘要中“实验中被试首先看到一个英文情绪单词，接着进行视觉空间位置判断，最后再让被试判断之前呈现的情绪词是积极词还是消极词。”出现了两次，显得有些重复，建议进一步修改摘要，将实验 1 和 2 的方法、目的、结果结合起来写作。

回应：感谢审稿人的宝贵意见。在修改稿中，我们对摘要进一步进行了修改，删除了重复的内容，并把两个实验的目的、方法和结果结合到一起阐述。

意见2：引言中最后一段中建议将“本研究的逻辑是”改为“因此，基于前人的研究，我们认为”。

回应：感谢审稿人细心的建议，我们按照您的建议对文章进行了修改，并通过朗读全文，再次对文章的行文表述进行了修改和润色。具体见 P16。

意见3：建议在2.2.2 实验设计中只提自变量和因变量，其他说明性的文字提到2.1实验目的中。

回应：根据审稿人的建议将 2.2.2 实验设计中的其他说明性的文字提到 2.1 实验目的中，并在实验二中也做了同样的修改。具体见 P17、19、20。

意见 4：表一中说明“+”的是标准误还是标准差？

回应：表一中“+”是指标准差。为了使文章的表达更准确，按照您的建议，我们对文章中的表1和表2均做了修改，将表1中的“空间位置判断任务中的平均反应时(ms)和准确率(%)”改成“空间位置判断任务中反应时(ms)和准确率(%)的均值和标准差”，表2也做了同样的修改，具体见 P18、20。

意见 5：2.3 结果中建议将“删除任一任务反应错误数据 830 个，占总数据的 16.21%，其中位置判断错误数据 59 个，占总数据的 1.15%，词汇判断错误数据 786 个，占总数据 15.35%”，改为“反应时只包括两个任务反应都正确的反应时”。

回应：我们按照您的意见对2.3结果进行了修改，并对3.3结果也做了同样的修改。具体见 P18、20。

意见 6：注意统计数值的标准写法，如等号前后应有空格等。

回应：感谢审稿人细心的意见。修改稿中，按照您的建议，我们已经对文章的统计数值的标准写法，如等号前后应有空格、小于号前后应有空格等做了全文的检查，并进行了相应的修改。

审稿人 2 意见：

意见1：修改稿较好回应了我的疑虑，作者重新采集了数据，得到了一致的结果，说明结论可信，建议发表。

回应：感谢评审专家百忙之中对本文进行评审，感谢评审专家对本文的肯定！