

## 《心理学报》审稿意见与作者回应

题目：道德推理与道德行为关系的元分析

作者：吴鹏，刘华山

### 第一轮

#### 审稿人1意见：

该文对道德推理与道德行为的关系进行了元分析，选题前沿，文献阐述逻辑基本清楚，研究过程清晰，研究结论基本合理。几个需要注意地方：

#### 意见 1：

在第 3.1 中，内容使用了异质性检验，但标题却使用了同质性检验，请一致；

#### 回应：

我们按照专家的意见，将标题“3.1 同质性检验”改成“3.1 异质性检验”（见 3.1）。

#### 意见 2：

在第 3.4 中，图 3 与图 4 非常不清晰；

#### 回应：

在我们使用的元分析软件 CMA 中，对于图的输出有一定的缺陷，以致无法非常清晰地保存效应量分布图。我们尝试通过截图、直接输出文本、图片等多种方式来得到清晰的图片，但效果都不是很好。其中，通过直接输出的图片是其中最清晰的，我们也就通过这种方式修改了原来的效果分布图，见图 3 和图 4。此外，为了更加清晰地表明原始研究的效应量，我们增加了“表 1 元分析中纳入的原始研究”（见 2.3 文献编码）。

#### 意见 3：

在第 4 的第四段中，结论是“采用标准化工具来探讨道德推理与道德的行为间的关系会比较差”，还是“采用标准化工具来探讨道德推理与不道德的行为间的关系会比较差”？请正确描述；

#### 回应：

根据专家的意见，我们将文中“采用标准化工具来探讨道德推理与道德的行为间的关系会比较差”改成“采用标准化工具来探讨道德推理与不道德的行为间的关系会比较差”（见 4 中第四段，第 4-5 行）。

#### 意见 4：

在第 4 中的最后两段中，被试年龄对道德推理与道德行为的关系的调节作用解释不清楚，请清楚解释；

#### 回应：

我们的研究发现，被试的年龄阶段能调节道德推理与不道德的行为间的关系，但不能调节道德推理与道德的行为间的关系。对于前一个显著的调节作用，研究表明，相比于成人，儿童、青少年的道德推理与不道德的行为之间有更强的联系。其原因可能是，儿童与青少年正处在道德发展的重要阶段，学校与家庭都会教导他们做出符合社会文化要求的、正确道德推理（Bar-Tal & Nissim, 1984; Cushman, Young, & Hauser, 2006）。同时，对于他们的不道德的行为可能会有更多的关注。即相比于促进道德的行为，父母与老师更加注重对不道德的行为的抑制。于是，在提升儿童、青少年道德推理水平的同时，家庭与学校也在极力

减少其不道德的行为。这两方面的作用就使儿童、青少年越能进行正确的、高水平的道德推理（即道德推理得分越高），也越能抑制自己的不道德行为。因此，儿童、青少年的道德推理与其道德行为之间有一定的负相关（Gasser & Malti, 2012; Guzman, 2006）。而对成人来说，其道德推理已经趋于稳定，个体间道德推理得分差异较小（Eisenberg et al., 2002; Reynolds & Ceranic, 2007）。但是成人间不道德的行为差异会比较大，有些成人会经常做出不道德的行为（Eisenberg et al., 2002）。因为不道德的行为可以为其带来一定的“好处”（比如，金钱、权势等），成人有强烈的动机去获取这些“好处”。此外，成人更能为自己的不道德行为寻找理由、做出辩解（Detert, Trevino, & Sweitzer, 2008），以减少不道德行为可能带来的自责、羞愧等负性影响。于是，成人间的道德推理并没有很大的差异，而其不道德的行为却有很大的差异，这两者变异间的不平衡就导致了其相关程度不大（见 4 中第五段，第 7-24 行）。

对于后一个不显著的调节作用，即年龄阶段不会影响道德推理与道德的行为之间的关系。我们认为原因可能有：①儿童和青少年会经常被要求或奖励去做出道德的行为，学校、家庭与社会也会教育他们进行正确的道德推理，这两方面的作用就使儿童和青少年的道德推理与道德的行为之间存在一定的正相关（Malti & Gutzwiller-Helfenfinger, 2010; Eisenberg et al., 2002）。而相比于儿童，成人对自己做出道德的行为会有更强烈的要求（Carlo, Crockett, Randall, & Roesch, 2007; Maeda, Thoma, & Bebeau, 2009）。成人又具有更强的体力、更好的能力来完成道德的行为，因此成人会有更多的道德行为。同时，随着年龄的增长，个体的道德推理水平会越来越高（Eisenberg et al., 2002; Narvaez & Gleason, 2007）。于是，道德推理与道德的行为随着年龄同时增长，成人的道德推理与道德的行为会都强于儿童与青少年。但相关关系只是关注这两者之间的联系，随时间共同的增长可能不会显著影响这两者之间的相关系数，于是年龄阶段不会影响道德推理与道德的行为之间的关系。②目前的道德推理测量工具主要针对于道德的行为，因此道德推理与道德的行为关系可能更密切。而一些道德推理的测量工具针对不同年龄群体有相应的版本，比如 PROM 就有儿童版、青少年版与成人版。测量工具针对各个年龄阶段进行了修订，确保各个年龄阶段被试都能准确作答，这也就保证了不同年龄阶段的道德推理与道德的行为之间都能展现出紧密的联系。③本元分析涉及的研究中探讨青少年的道德推理与道德的行为之间关系的研究数量（21）远多于针对儿童或成人的研究（7），三种年龄群体的研究数量的不均衡可能也会影响元分析中调节效应的分析结果（Borenstein et al., 2009）（见 4 中第六段，第 3-19 行）。

#### 审稿人2意见：

采用元分析技术探讨道德推理与道德行为的关系是非常必要的。然而，这篇文章在创新性和写作上存在明显的问题。

#### 意见 1：

研究缺少独特而有价值的新发现，只是对以往关于两变量关系的一个综合计量。

#### 回应：

我们的元分析针对当前道德心理学研究中的热点问题（即道德推理是否为重要的因素）而展开。在整个道德心理学研究历史中，皮亚杰与科尔伯格提出的道德认知理论一直都是研究的唯一视角。但从上个世纪末起，Haidt, Greene 等学者认为非认知因素（道德情绪、道德直觉等）才是重要的道德因素，甚至认为道德推理不会影响个体的道德行为。于是，当前道德心理学研究领域对于道德推理的作用存在一定的争论，本研究希望通过整合以往的相关研究结果，从更大范围、更久时间来检验这一争论。诚然，元分析就是对两变量关系的一个综合计量，但它不是一种简单的综合。在计算整体效应量的过程中，需要考虑各个原始研究所占的权重，也要考虑各原始研究背后的真效应量是否一致。对这些问题的考虑，保证了整体效应量可以更加准确可靠。本元分析通过对涉及 35 个独立效应量、6663 名被试的相关研究

进行效应量综合发现道德推理与道德的行为之间是存在一定的关系，而通过对涉及 48 个独立效应量、10065 名被试的相关研究进行效应量综合发现道德推理与不道德的行为之间也是存在一定的关系。综合来看，我们的元分析结果表明在国内外 40 多年来的心理学相关研究中道德推理与道德行为之间确实存在一定的关系，这是本研究的价值之一。此外，我们的元分析也考察了道德推理测量工具类型的调节作用。结果发现，采用标准化道德推理测量工具的研究中道德推理与道德的行为之间的关系更为紧密，这一结果意味着在探讨道德推理与道德的行为之间的关系时，研究者应该采用标准化测量工具。而本研究也发现道德推理与不道德的行为之间的关系并不会受道德推理测量工具类型的影响，这一结果表明道德心理学家今后需要开发专门针对不道德行为的道德推理测量工具，这是本研究的一个新发现，也是具有独特价值的研究结论。最后，本元分析探讨了被试年龄阶段的调节作用。结果表明，被试的年龄阶段只能影响道德推理与不道德行为之间的关系。依据这一结果，在儿童与青少年的道德教育中，教育者还是应该要加强学生的道德推理，通过增强道德推理来减少不道德行为。而要抑制成人的不道德行为，更多需要从道德情感、道德意志等方面进行干预。总体来说，我们的元分析虽然只是针对道德研究领域的基本问题而展开，但我们的研究结果既增进了道德心理理论的内容，质疑了当前排它性关注非认知因素的研究取向，也为道德教育实践提供了一定的实证研究贡献。

#### 意见 2:

研究的内容非常简单，然而论文的篇幅却很长，过度铺垫了外围文献，而且有过多的方法学的一般介绍。若能围绕要研究的问题以更直接的方式写作，文章会好些。目前该文章不适合在《心理学报》发表。

#### 回应:

根据专家的意见，我们对文章的内容进行一定的修改及删除，以更直接明了的方式描述我们的研究（见 1.4 中第三段，第 6-7 行；2.4.1 中第一段，第 1-2 行；3.1 中第二段，第 1-2 行）。而专家指出的我们的文章中有过多方法学的一般介绍，是因为我们考虑到元分析技术还是属于一个较为新的研究方法，对于其中的一些重要技术内容需要做一定的阐释，以便读者可以准确、清晰地了解我们的研究过程。在我们的文章中，主要是针对模型选择与发表偏差两部分进行了详细的介绍。这是因为我们认为在元分析技术中，这两个问题对于最终结果的准确性有着重要的影响。于是，我们用了一定的篇幅来比较两种模型的差异，也简单介绍了发表偏差的危害及检验方法。通过对这两个问题的介绍，读者就可以准确理解我们的整个研究结果，也可以判断出我们的研究结果是可靠的。

#### 审稿人3意见:

讨论了一个很有理论价值的话题，即道德推理与道德行为表现之间的关系。而且采用元分析的方法来回答这个问题也是合适的。但文章存在如下问题:

#### 意见 1:

作者混淆了显著性与相关大小的区别，以致作者可能得出了错误的结论。相关显著只是说明此系数受偶然因素影响的可能性很小，小于 5%；而没有说明相关的大小。元分析结果显示，道德推理与行为之间的相关为 0.238，显示二者之间仅有 5.8% 的共同变异。也就是说，这个数字显示，道德推理和道德行为之间关系确实不大。

#### 回应:

在我们的元分析中，以相关系数作为效应量分析了道德推理与道德行为之间的关系。确实如专家所指出的，我们的研究结论可能夸大了道德推理与道德行为之间的关系程度。相关系数的显著性只是表示了最终得到的系数不等于零，受偶然因素影响的可能性很小。变量间

关系程度的强弱应该与其相关系数的绝对值大小有关,本元分析最后得到的平均效应量分别为 0.238、-0.188。这两个相关系数的数值大小只能表明道德推理与道德行为之间有一定的联系。

**意见 2:**

依照前一个意见,文章的引言部分要做调整,讨论部分需要重写。

**回应:**

根据专家的意见,我们修改了讨论部分的表述,将“这两个总效应量均说明道德推理与道德行为之间还是存在比较强烈的联系的”改为“这两个总效应量均说明道德推理与道德行为之间还是存在联系的”(见 4 中第二段,第 2-3 行)。在本元分析研究中,我们主要关注道德推理与道德行为之间是否确有关系。因此,我们关心的是相关系数是否显著,相关系数的数值大小不是我们研究的重点。我们对引言部分稍作修改,就将“道德推理与道德的行为有显著的正相关,道德推理与不道德的行为有显著负相关”改为“道德推理与道德的行为之间有正相关,道德推理与不道德的行为之间有负相关”(见 1.4 中第三段,第 11-12 行)。我们也修改了结果部分的表述,将“本元分析发现道德推理与道德行为之间还是存在比较强烈的联系的”改成“本元分析发现道德推理与道德行为之间存在联系”(见 5 中,第 1 行)。此外,我们对“道德推理与道德行为的联系”进一步展开了讨论,补充探讨了道德行为与道德推理相关程度不强的原因。从道德双加工理论与道德动机社会直觉模型的角度,分析了道德行为是受多种道德因素共同影响的结果,而道德推理只是其中很小的一部分(见 4 中第二段,第 16-25 行)。

**意见 3:**

50 篇文章的基本特征,包括原始研究报告的各个维度间相关系数,需要整理出来,附于文中。

**回应:**

根据专家的意见,我们将原始研究的信息进行了整理,以表格的形式呈现(见 2.3 中,表 1)。

**意见 4:**

引用参考文献格式可能有问题。Borenstein, Hedges, Higgins, & Rothstein, 2009 多次出现,却是同样的模样。第二次及其以后要简写。

**回应:**

按照专家的意见,我们修改了有问题的参考文献格式(见 2.4.2 中第一段,第 8 行、第 14 行; 3.1 中第二段,第 2 行; 3.2 中第三段,第 8 行)。

**意见 5:**

1.5.2 节涉及的不同阶段是哪些阶段?

**回应:**

根据专家的意见,我们补充了年龄阶段的具体信息(见 1.5.2 中,第 4 行)。

**意见 6:**

修改文字。有些地方过于口语化,不够精炼。见文稿。

**回应:**

根据专家的意见,我们修改了文中的部分文字表述(见 1.4 中第三段,第 6-7 行; 2.4.1

中第一段，第 1-2 行；3.1 中第二段，第 1-2 行）。

## 第二轮

### 审稿人1意见：

修改或考虑如下问题

#### 意见 1：

道德推理是否与道德行为有显著相关？这是本文的核心问题。元分析方法的作用是综合多个效果量的方法，并寻找相关原因。之所以要综合它们，是因为它们受到抽样误差的影响。在有些研究里表现出了显著的相关，在另一些研究里则没有表现出相关。效果量的这种变异有可能是在“皮亚杰等人的理性取向”下的，也可能是“Haidt 非理性取向”下。但本研究所处理的仅仅是理性取向下的。因此，没有办法考察理性与非理性之间的差异。也就是说，过多地铺陈非理性取向的文献也许是触动作者做本研究的起点，但会让论文的主题不够鲜明，显得虎头蛇尾。因此，建议前言部分的中心更明确些。

#### 回应：

正如专家所指出的，本研究的出发点就是想探讨道德行为与道德推理之间的关系。我们的主要研究目的是想证明道德心理的“理性取向”应该还是有其合理性的，不应该如 Haidt 等学者所指出的道德研究可以不考虑认知因素（特别是道德推理）。在“Haidt 非理性取向”下的研究不会关注于道德推理，因此本元分析中纳入的原始研究多是在“皮亚杰等人的理性取向”下的，我们的研究确实不能探讨理性与非理性的差异。我们按照专家的意见，修改了前言部分的论述以明确论文主题。具体来说，我们删减了部分关于非理性取向的论述（见 1.1 中第三段，第 4-5 行；1.4 中第三段，第 3-6 行）。

#### 意见 2：

表 1 全是理性取向的？表 1 的标题行用汉字即可。

#### 回应：

“理性取向”与“非理性取向”的差异就在探讨的道德因素不一样，道德推理完全是“理性取向”下的研究主题，因此很少有“非理性取向”下的学者会讨论道德推理与道德行为之间的关系。本元分析中纳入的原始研究多是“理性取向的”，但也有一些研究可能是从“道德双加工”角度同时考察了“理性因素”与“非理性因素”，如（洪丽, 2005; 毛静思, 2012; Miller, Eisenberg, Fabes, & Shell, 1996）。此外我们按照专家的意见，将表 1 的标题行改为汉字（见表 1）。

#### 意见 3：

图 3 和图 4 并非必须有的。如果不清楚，可以用统计表描述同样的。

#### 回应：

我们在文章中添加图 3 与图 4 是想直观表明各研究的效应量之间存在很大的变异，这一明显的差异也是后面进行调节效应分析的原因之一。由于统计表不如图直观，因此我们保留图 3 与图 4。但为了进一步增加图表的清楚度，我们对图 3 和图 4 进行了小的修改，考虑到结果部分前面的内容，图中只保留重要的统计信息（即各个原始研究的效应量及总效应量的分布森林图），删除了重复的统计信息（即基本统计信息）。感谢专家的宝贵建议。

#### 意见 4：

讨论部分需要做小的修改。不必重复前言部分铺垫性的内容。直接陈述对前言部分提出

的问题是如何回答的。与以前的研究有哪些相同之处，有哪些不同之处？为什么？调节效应分析结果如何？可能的原因是什么？未来研究应该关注哪些问题？本研究的不足有哪些？

**回应：**

我们按照专家的意见，对讨论部分进行了一定的修改。首先，删除了重复前言部分的铺垫性文字并修改了部分内容（见 4 中第一段）。其次，为了明确、直接陈述对前言部分所提出问题的回答，我们在讨论部分开头就直接介绍了研究结果（见 4 中第一段，第 3-5 行），并增加了二级与三级标题以突出展现各个具体结果及其可能原因的讨论（见 4.1、4.2、4.2.1、4.2.2 与 4.3）。最后，我们增加了研究不足与展望部分，探讨了本研究的不足之处与未来研究方向（见 4.3）。

**审稿人2意见：**

同意发表。