

《心理学报》审稿意见与作者回应

题目：团队人口统计特征多元化与绩效关系的元分析

作者：卫旭华、刘咏梅、陈思璇

第一轮

总体修改说明：非常感谢两位外审专家的意见，我们认真思考了这些意见，阅读了国内外相关文献，对论文的不足进行了修正和完善。原论文基于两个维度将团队多元化划分为六种类型，专家一（意见 6）认为这种分类过细，建议仅考虑单一维度，即仅考虑分离、多样和不平等三种类型，而不区分任务导向和关系导向的性质差异。此外，专家一（意见 1 和意见 2）建议补充检验团队类型的调节作用，这势必会增加文章的长度，而专家二（意见 6）建议对论文的文字进行精炼。为了兼顾论文的质量及简洁性，我们对原论文的框架进行了调整，即遵循专家一的建议，仅考虑三种类型的团队多元化（分离、多样和不平等），并遵循专家二的建议对理论基础进行了加强，从而使论文的理论贡献更加明确（专家二意见 2）。同时，我们还对论文论据以及研究方法的描述进行了加强（专家一意见 3、4、5、7 以及专家二意见 1、3、4、5），以提升论文的严谨性。为了更好地总结团队多元化与团队绩效的关系，本次修改还补充了团队多元化领域 2014 年新发表的 14 篇中英文实证论文，并纳入到元分析中。具体修改意见及修改说明见下文。

审稿人 1 意见：

意见 1: 是否应该将团队类型作为调节变量。不同类型的团队，其特征存在很大差异。比如，职能团队、研发团队、高管团队、创业团队等等，各自的任务与使命、团队构成、团队过程都存在巨大差异。从元分析所包括的实证研究的对象也可以看到，不同研究可能关注不同类型的团队。对于不同类型的团队而言，各种团队异质性指标是否会有不同影响？我个人认为这个问题值得探讨。

回应: 我们遵循专家一的意见，将团队类型纳入调节变量分析。根据元分析样本的特点，并遵循 Bell 等(2011)对团队类型的划分，我们将团队划分为三类：高管团队、研发团队、混合工作团队。在修订版本中，我们增加了相关描述，包括理论假设、变量编码、结果分析与讨论。

首先，我们在“理论基础与研究假设”部分提出了我们关于团队类型调节作用的研究假设及其理由（正文第 4 页）：

“团队类型可能会调节多元化与绩效之间的关系(Horwitz & Horwitz, 2007; Joshi & Roh, 2009)。常见的团队类型包括高管团队、研发团队和一般工作团队等(Bell, et al., 2011; de Wit, et al., 2012)。从团队的工作内容来看，高管团队通常面临大量复杂模糊的决策任务，研发团队通常需要创造和开发新的产品与服务，而一般工作团队则以常规任务为主(Bell, et al., 2011)。

从分离型多元化角度来看，分离往往形成团队断层和子群体，并诱发子群体间的冲突(Thatcher & Patel, 2011, 2012)，而团队冲突领域的研究者则发现团队冲突的负面作用在高管团队以外的团队类型中表现的更为明显(de Wit, et al., 2012)。因此，本研究预期高管团队样本中的分离型多元化与团队绩效的负相关程度低于研发团队和一般工作团队。

从多样型多元化角度来看，高管团队和研发团队面临的任务具有一定的非常规性和不确

定性,经常需要一些独到的观点和新颖的方案,因而更加需要团队成员多样的观点来弥补单个高管决策的不足(de Wit, et al., 2012; 任兵, 魏立群, & 周思贤, 2011)。而对于一般工作团队而言,由于其所面临的任务相对较为常规,解决起来相对比较容易,因而对团队多样的需求较低(Horwitz & Horwitz, 2007)。这说明与一般工作团队相比,高管团队和研发团队更能从团队多样中获益。

从团队不平等型多元化角度来看,不平等可能会在权力更高的团队(如高管团队)中发挥更大的负面作用。研究表明,当团队成员的权力都很低的时候(如研发团队和一般工作团队),团队等级和不平等有助于解决冲突,促进团队和谐;而当团队成员的权力都很高的时候(如高管团队),团队成员更倾向于权力争斗,此时的不平等不利于团队冲突的解决,进而引发低绩效(Greer & van Kleef, 2010)。这说明团队不平等对于高管团队更加有害。基于此,提出如下假设:

假设 3: 团队类型会调节团队多元化与团队绩效的关系,即与其他类型团队相比,高管团队中的分离型多元化与绩效的负向关系更低,一般工作团队中的多样型多元化与绩效的正相关程度更低,高管团队中的不平等型多元化与绩效的负向关系更强。”

其次,我们在研究方法部分交代了团队类型的编码过程,详见正文第 5 页:“……团队类型依据各论文研究方法部分的样本描述进行识别……本研究借鉴 Bell 等(2011)关于团队的分类,将团队类型划分为高管团队、研发团队和其他团队(即混合工作团队)……为了保证多元化类型、效应值数据以及各情境调节变量编码的准确性,由两名博士生对所有实证研究进行独立编码,随后进行比对。各变量(……团队类型)的评定者间信度从 0.83 到 1.00,均值为 0.92,说明变量编码存在较高的一致性”。

再次,我们在“结果分析”部分交代了分析结果,详见正文第 6 页:“表 4 列出了团队类型作为潜在调节变量的分析结果。可以看到,团队类型对多样型多元化与团队绩效关系的调节作用显著($Q_B=24.34, p<0.001$),多样型多元化与研发团队绩效($\rho=0.23, p<0.001$)和高管团队绩效($\rho=0.07, p<0.001$)显著正相关,但多样型多元化与混合工作团队绩效关系并不显著。团队类型对其他类型多元化与绩效关系的调节作用不显著,因此假设 3 得到部分验证”。

最后,我们在“讨论”部分讨论了我们的研究发现,详见正文第 8 页:“元分析结果验证了团队类型对多元化和绩效关系的调节作用。对于决策型的高管团队和智力型的研发团队而言,由于成员彼此之间的相互依赖程度较高,多样型多元化能够较好地促进知识和经验的整合,因而其积极作用更大;而对于一般的工作团队而言,由于工作任务的相互依赖程度较低,多样型多元化并不能发挥信息整合的优势。同时,通过进一步的元分析,我们发现研发团队、高管团队的多元化与绩效的关系也存在地域差异。在东方国家中,研发团队的多样型多元化表现出了更强的正面作用,我们认为这主要是得益于东方国家人们在研发任务中展示出较高的合作水平(Kim, et al., 2007);而在西方国家中,研发团队的多样型多元化与团队绩效的关系并不显著,其可能的原因在于西方个体主义国家的人们在研发任务中更加倾向于说服其他人,展示出较多的竞争行为(O'Neill, et al., 2013),进而削弱了研发团队的绩效。此外,在西方国家中,高管团队不平等型多元化表现出了较强的负面作用,这与西方的高管团队等级研究结果较为一致;而在东方国家中,高管团队不平等型多元化与团队绩效的关系并不显著,可能的原因在于中国是高权力距离社会,对于不平等的接纳程度更高(Hofstede, 1984),从而在一定程度上削弱了团队不平等的消极作用”。

意见 2: 从所引用的实证研究来看,东方背景(主要是中国情境)下的研究,大多数是针对高管团队,如何区分文化因素和团队类型因素?此外,表 3 的结果表明,文化背景的调节效应不是很显著,是否与此有关?

回应: 非常感谢专家一的新颖视角, 我们赞同这一观点, 并认为文化因素和团队类型之间可能会存在一定的交互作用, 因而我们分析了特定团队类型中多元化与绩效关系的文化差异。由于专家二认为“文化”一词过于宽泛, 建议我们将文化改为地域, 因此在正文的描述中, 我们用“地域”替换了“文化”。基于这一观点的数据分析描述为(第7页, 数据表格详见正文): “为了更进一步理解团队和地域的调节作用, 我们对团队类型和地域进行了交叉, 以比较特定团队类型中地域的调节作用, 结果如表5所示。可以看到, 东、西方国家研发团队的多类型多元化与绩效关系存在显著差异 ($Q_B=4.98, p<0.05$), 与西方国家研发团队相比 ($\rho=0.06, p=0.52$), 东方国家研发团队多类型多元化与绩效的正相关程度更强 ($\rho=0.31, p<0.001$)。此外, 与东方国家高管团队相比, 西方高管团队的不平等型多元化表现出显著的负面作用 ($\rho=-0.03, p<0.05$), 说明地域对于高管团队中的不平等型多元化与绩效关系的调节作用显著 ($Q_B=5.93, p<0.05$)”。

同时, 我们在结果讨论部分也对这一发现进行了深入的探讨, 分析了其蕴含的理论和逻辑, 具体描述为: “同时, 通过进一步的元分析, 我们发现研发团队、高管团队的多元化与绩效的关系也存在地域差异。在东方国家中, 研发团队的多样型多元化表现出了更强的正面作用, 我们认为这主要是得益于东方国家人们在研发任务中展示出较高的合作水平(Kim, et al., 2007); 而在西方国家中, 研发团队的多样型多元化与团队绩效的关系并不显著, 其可能的原因在于西方个体主义国家的人们在研发任务中更加倾向于说服其他人, 展示出较多的竞争行为(O'Neill, et al., 2013), 进而削弱了研发团队的绩效。此外, 在西方国家中, 高管团队不平等型多元化表现出了较强的负面作用, 这与西方的高管团队等级研究结果较为一致; 而在东方国家中, 高管团队不平等型多元化与团队绩效的关系并不显著, 可能的原因在于中国是高权力距离社会, 对于不平等的接纳程度更高(Hofstede, 1984), 从而在一定程度上削弱了团队不平等的消极作用”。

意见3: 实证研究的变量编码, 第一阶段独立编码的一致性如何? 是否有相应的统计指标?

回应: 我们根据先前的编码记录计算了评定者间信度, 并在正文中进行了补充说明, 详见正文第页: “为了保证多元化类型、效应值数据以及各情境调节变量编码的准确性, 由两名博士生对所有实证研究进行独立编码, 随后进行比对。各变量(样本量、团队规模、效应值、多元化类型、绩效信度系数、国家地域、团队类型)的评定者间信度介于0.83到1.00之间, 均值为0.92, 说明变量编码存在较高的一致性。对于有差异的数据, 两名博士生再次检查核对论文中的原始数据, 对于仍存在差异的数据在讨论后达成一致”。

意见4: 元分析的统计指标问题。目前仅报告了根据样本量进行加权得到的相关系数(r)。一般的元分析还会报告根据抽样误差和测量误差修正之后的估计值(ρ), 并且基于 ρ 值及其置信度区间来下结论。目前完全是基于 r 的数值来下统计结论, 难以完全保证结果的可靠性。

回应: 感谢专家一的建议, 我们此前的确忽略了修正问题, 当前的版本已经考虑了这一因素。在正文中我们交代了相关的描述, 并在数据报告中列出了修正前的 r 和修正后的 ρ 。

在正文第5页“3.2 变量编码”部分, 我们交代了“除了收集各变量的效应值之外, 还收集了通过量表测量的团队主观绩效的内部一致性信度系数, 以用于后续元分析的测量误差修正。对于某些未报告信度的主观绩效变量, 通过元分析样本中的信度平均值替代。对于团队多元化和客观绩效变量, 由于都是客观数据, 因而假定其不存在测量误差(Hülshager, Anderson, & Salgado, 2009)”。

在正文第5页“3.3 元分析过程”部分, 我们交代了“元分析软件采用 Comprehensive Meta Analysis 2.0, 首先使用 Hunter 和 Schmidt(2004)的元分析程序估计未修正的基于样本加权的平均相关系数 r , 其次估计根据测量误差(信度)修正后的总体相关系数 ρ , 并基于此计算

修正后的总体相关系数的 95% 置信区间。对于元分析计算模型的选择，由于研究样本涉及范围广，所有研究样本和情境均等同的假定很难成立，因而本研究遵循 Borenstein, Hedges, Higgins 和 Rothstein (2011) 的建议，在分析过程中全部采用随机效应模型”。

结果报告详见正文第 6-7 页的表 2、表 3、表 4、表 5，表中既报告了修正前的 r ，也报告了修正后的 ρ ，95% 的置信区间是修正后的置信区间。

意见 5: 从表 2 的结果看，任务取向多样性总是正向影响，而分离指标（无论是（任务还是关系取向）几乎总是负性影响。是否意味着某些类型的特征（如任务取向的人口学特征）应该扩大多元化程度，而另外一些则应该避免（如导致分离的人口学特征）？后者是否印证了新近关于团队断裂带(team faultlines)的研究结果？

回应: 感谢专家一的建议。参考了专家一的建议 6（详见后面的修改说明），当前版本仅仅区分了三种类型多元化。当前结果表明，尽管团队分离表现出了一定的负面作用，但这一作用并没有达到统计显著（即置信区间包含 0），因此，我们尚不能断定分离型多元化一定会形成团队断裂带，进而造成消极的影响。同时，近期团队断裂带的相关研究也并没有显示其一定会产生负面作用。

我们认为团队断裂带理论是团队多元化理论的拓展，而团队子群体理论（subgroup theory，详见 Carton 等 2012 年发表在 AMR 上的综述）则是团队断裂带理论的进一步拓展。因为团队断裂带（正文称之为“团队断层”）衡量的是团队划分为若干个子群体的可能性，因而原来提交的版本仅提及了子群体，没有涉及团队断裂带。事实上，Carton 等(2012)年关于子群体的理论也是建立在团队断裂带理论和 Harrison 等(2007)的多元化分类基础之上的，更进一步，他们依据 Harrison 等(2007)的多元化分类提出如下子群体理论：分离型多元化→基于分离的断层→基于身份的子群体，多样型多元化→基于多样的断层→基于知识的子群体，不平等型多元化→基于不平等的断层会→基于资源的子群体。

考虑到多元化、断层、子群体之间的潜在关系，因此，我们根据专家一的建议，对理论基础和研究假设部分进行了加强，详见第 3-4 页：“根据社会分类理论，当团队分离型多元化程度较高的时候，分离往往会诱发团队断层（faultlines），并促进团队内部基于身份的子群体的产生(Carton & Cummings, 2012, 2013; Cooper, Patel, & Thatcher, 2014)。此时子群体成员往往会更关心子群体内部成员的利益，而将子群体以外成员当做威胁，这会降低子群体之间的合作意愿和凝聚力。在这种消极氛围的影响下，分离型多元化可能会诱发更高的子群体间冲突(Carton & Cummings, 2013; Thatcher & Patel, 2011)”。

同时，我们在“讨论”部分阐述了未来展望，详见第 9 页：“此外，团队多元化不仅仅局限于个体成员之间的差异，也可以表现为团队断层和团队子群体之间的差异(Carton & Cummings, 2012)。例如，分离型多元化会形成基于分离的断层，进而产生基于身份的子群体；多样型多元化会形成基于多样的断层，进而产生基于知识的子群体；不平等型多元化会形成基于不平等的断层，进而产生基于资源的子群体(Carton & Cummings, 2012)。因而，未来研究者也可以关注团队断层、团队子群体多元化对团队交互过程和结果的影响”。

意见 6: 我觉得目前对于团队异质性的三类指标（分类、多样性和不平等）的探讨还没有达成共识，仍有研究的必要。因而我建议表 2 中，不区分人口学变量的性质差异（即不区分关系导向还是任务导向），报告这三类指标的结果，也有意义。

回应: 我们遵循了专家一的建议，仅仅区分了分离、多样和不平等三种类型。首先，Harrison 等(2007)的这一分类在西方有很大的影响力（已被引 1028 次），且从理论层面和操作层面而言，更加合理（近两年每年都有 200 次以上的引用）。其次，与西方研究形成鲜明对比，中国学者对这一分类的重视程度不高，这可能造成理论方面的偏差。最后，合并后元分析的效

应值数量更多，所得到的结论也更加稳健。此外，合并后的元分析一方面验证了 Harrison 等(2007)多元化分类的必要性，另一方面也能够呼吁中国研究者重视这一分类的理论基础和测量方式。

意见 7: 表 4 中，关系取向分离与任务绩效以及创新绩效的相差系数 r 差别很大，但统计检验结果不显著，是否与创新绩效的研究太少有关？如果根据抽样误差进行修正，是否会得到不同结果？

回应: 由于专家二认为篇幅需要压缩，且认为绩效类型的调节作用显得有些奇怪，而专家一提出需要增加团队类型为潜在的调节变量，且这一操作具有更大的现实意义。在权衡研究意义和篇幅之后，现在的版本不再包含“绩效类型”作为潜在调节变量的任何描述。

至于专家一指出的关系取向分离与任务绩效以及创新绩效两个相关系数相差很大，调节效应检验却不显著。这与样本量有很大关系，在该情形下，创新绩效的研究只有一篇（仅包含 11 个团队），因此其研究结果可能并不具备代表性。我们在讨论中也曾指出这一不足：“部分元分析的效应值数量偏少，可能会限制其有效性”。对于这么小样本的分析，即使通过修正，其参考意义也不大。当前版本通过合并一些过细的分类，也在一定程度上提升了各类多元化效应值的数量和用于分析的样本量，并提升结果的有效性。

审稿人 2 意见:

意见 1: “三人行必有吾师”跟团队多元化没有什么关系，放在这里不是很适合。

回应: 我们已经删除了这一句话，并对文章其他部分进行了检查，尽量避免不合适的语句。

意见 2: 关键的变量，如“关系取向分离、任务取向分离、关系取向多样、任务取向多样、关系取向不平等和任务取向不平等”是如何定义的，或者这样定义有什么意义？如果只是由于操作方法不同而进行这样的区分，对理论的贡献不是很大。

回应: 感谢专家二对于理论基础方面的建议。同时，参考专家一的建议 6，当前版本仅仅考虑了分离、多样和不平等三个关键变量。Harrison 等(2007)是基于以往经典理论提出了多元化分类，然后才强调每一种多元化则应采取与之对应的操作方法。我们首先在原文中对三个关键变量的理论界定进行了加强，具体如下（正文第 2 页）：

“分离型多元化指的是团队成员在立场和观点上的差异程度，通常是信念、价值观或态度上的对立或者不一致(Harrison & Klein, 2007)。例如，当团队内年轻成员支持改革方案，而年长成员由于保守而反对改革的时候，就会形成团队年龄分离型多元化。从表 1 可以看到，当团队内所有成员拥有相同观点的时候，分离型多元化程度最低；而当团队内形成两个势均力敌的极端派系的时候，分离型多元化程度最高(Harrison & Sin, 2006; Solanas, Selvam, Navarro, & Leiva, 2012)。

多样型多元化通常指的是与知识和经历相关的种类上的差异，如专长、职能背景和行业经历方面的多元化(Harrison & Klein, 2007)。从表 1 可以看到，当团队内所有成员处于同一分类的时候，多样型多元化程度最低；而当团队内所有成员所处的分类都不同的时候，多样型多元化程度最高(Harrison & Sin, 2006; Solanas, et al., 2012)。

不平等型多元化通常指的是团队成员在有价值的资源或社会资产方面的差异(Harrison & Klein, 2007; Magee & Galinsky, 2008)。例如，任期较长的成员在占有组织各项资源上具有明显优势，进而会形成团队任期不平等型多元化。从表 1 可以看到，当团队内所有成员拥有相同资源的时候，不平等型多元化程度最低；当团队内一人高高在上，占有了所有资源，而其他成员均处于最底层，没有任何资源的时候，不平等型多元化程度最高(Harrison & Sin,

2006; Solanas, et al., 2012)。”

同时，我们在研究假设部分明确了这三种分类的理论基础，如下（正文第 3 页）：

“分离型多元化的理论基础为相似性吸引理论（similarity attraction）和社会分类理论（social categorization）。相似性吸引理论认为团队成员的相似性能够降低分离型多元化，此时所有成员的态度和观点较为类似，有助于产生更高的合作水平、信任和社会整合程度（Guillaume, et al., 2013; Williams & O'Reilly, 1998）。

多样型多元化的理论基础为信息加工理论（information processing）和变异-选择-保留理论（variation, selection, and retention），这些理论通常认为拥有不同人口统计特征的成员可能拥有不同的经验和看问题的方式，会为团队带来不同的决策信息（Han, Han, & Brass, 2014），能够拓展团队视野并促进团队知识的整合，从而促进更高的团队创造性和创新能力、更高的决策质量以及更高的团队灵活性（Guillaume, et al., 2013; van Dijk & van Engen, 2013）。

不平等型多元化的理论基础为公平理论（equity theory）和社会等级理论（social hierarchy），这些理论通常认为不平等能够增加团队成员之间的竞争，降低成员之间的沟通意愿，并提升成员的不公平感（C. Anderson & Brown, 2010; Connelly, Tihanyi, Crook, & Gangloff, 2014）。”

此外，Harrison 等(2007)的这一分类在学术界的影响力很大，截止 2015 年 2 月底，他们关于团队多元化分类的论文在谷歌学术被引用 1028 次，近两年平均每年都保持 200 次以上的引用率，也在一定程度上说明多元化分类理论是得到了学者们的认可的。谷歌学术截图如下：



意见 3：样本选取的标准应该更清楚地说明。如从那年到那年发表的文章，发表在什么杂志上的文章？元分析的样本应该越全越好。

回应：已补充相关说明，如下（详见正文第 4-5 页）：

“.....查找的时间范围限定于 1984-2014 年间公开发表的学术论文。中文文献为发表在心理学 CSSCI 期刊（如心理学报、心理科学）以及国家自然科学基金委认定的 30 种管理科学重要期刊上（如管理世界、南开管理评论）的论文.....最终得到 33 篇中文文献。英文文献在 Web of Knowledge 中检索公开发表的 SSCI 期刊（如 Academy of Management Journal、Strategic Management Journal）.....最终得到 104 篇相关英文文献。这 137 篇中英文文献共涉及 49 个期刊（期刊目录详见附件）.....”。

除了通过正文的五条标准在如上数据库中检索之外，我们还通过以往发表的元分析文献进行查漏，正文中描述如下：“为了减少发表偏差（publication bias），本研究也通过参考以往元分析的文献部分进行查漏补缺”。同时我们也通过相关检验进行了发表偏差的检验，详见第 5 页“元分析过程”说明部分。

最后，我们还增加了期刊统计，放至论文的附录部分，可以看到元分析文献涵盖了国内外组织行为和工业与组织心理学的主流期刊，如下：

期刊名	论文数	期刊名	论文数
Academy of Management Journal	16	管理科学	2
Group & Organization Management	9	科学学研究	2
Journal of Organizational Behavior	8	中国软科学	2
Administrative Science Quarterly	6	心理学报	1
Strategic Management Journal	6	Academy of Management Learning & Education	1
Journal of Management	5	Applied Psychology	1
管理评论	5	Canadian Journal of Administrative Sciences	1
International Journal of Human Resource Management	4	Corporate Governance: An International Review	1
Organization Science	4	European Management Journal	1
Small Group Research	4	Human Relations	1
管理世界	4	Intercultural Education	1
管理学报	4	International Journal of Conflict Management	1
科学学与科学技术管理	4	Innovation: Management, Policy & Practice	1
南开管理评论	4	Information Systems Research	1
Creativity and Innovation Management	3	Journal of Applied Social Psychology	1
Journal of Applied Psychology	3	Journal of Business & Management	1
Journal of Business Ethics	3	Journal of Educational Administration	1
Journal of Management Studies	3	Journal of Personnel Psychology	1
科研管理	3	Organizational Behavior and Human Decision Processes	1
British Journal of Management	2	Research Policy	1
Group Dynamics: Theory, Research, and Practice	2	Systems Research and Behavioral Science	1
Journal of Business Research	2	The Leadership Quarterly	1
Journal of Business Venturing	2	研究与发展管理	1
Journal of Managerial Psychology	2	预测	1
Management Science	2	合计	137

意见 4: 两个调节变量的定义和描述不是很清楚。文化这个概念太笼统，根据你的研究，还不如直接定义成地域（东、西方），这与你后续的测量也很好契合。另一个调节变量更是不好理解，本来探讨多元化与绩效的关系，为什么还要探讨绩效类型的调节作用？很有 data driven 的嫌疑。

回应: 感谢专家二的建议。当前版本对文章涉及的所有调节变量进行了说明，并且按照专家二的建议，将“文化”替换为“地域”。具体描述如下：

“团队多元化的作用机理可能会随着国家和地域的不同而不同(Johnson, Schnatterly, & Hill, 2013)。借鉴先前组织管理领域的元分析，本研究将国家地域分为东方国家和西方国家两类(Anderson, et al., 2010)，处于同一地域的国家往往拥有相似的价值观和文化(Gupta, Hanges, & Dorfman, 2002)。

团队类型可能会调节多元化与绩效之间的关系(Horwitz & Horwitz, 2007; Joshi & Roh, 2009)。常见的团队类型包括高管团队、研发团队和一般工作团队等(Bell, et al., 2011; de Wit, et al., 2012)。从团队的工作内容来看，高管团队通常面临大量复杂模糊的决策任务，研发团队通常需要创造和开发新的产品与服务，而一般工作团队则以常规任务为主(Bell, et al., 2011)”。

至于专家二所提出的绩效类型可能是出于数据驱动的嫌疑，我们是基于以往团队领域元分析的调节变量选择了绩效类型（如 de Wit, Greer, & Jehn, 2012, JAP）。最初版本所包含

的调节变量还包括地域、团队规模、团队绩效类型、团队类型、团队绩效测量方式等，但由于篇幅限制，一些调节变量没有包含进来。虽然区分不同的绩效类型是很有必要的，但是专家一（意见 1）提出需要增加团队类型为潜在的调节变量，且这一操作具有更大的现实意义。在权衡研究意义和篇幅之后，现在的版本不再包含“绩效类型”作为潜在调节变量的任何描述。

意见 5: “假设 1c: 关系取向多样与团队绩效弱正相关或不相关”不是一个好的假设，因为你的 null hypothesis 就是“没有相关”。同时，“弱正相关”显然也没有必要，有正向相关即可。

回应: 感谢专家二指出这一不严谨的假设。当前版本所有的假设都有明确的方向性，即强调是“显著正相关”还是“显著负相关”。

意见 6: 整篇论文太长，建议精简到 1 万字左右。例如，2.3 一节完全可以简化很多。

回应: 我们已对文章架构进行了调整，并对文章字数进行了精炼（包括原文的 2.3 节）。

当前版本由于包含了修改说明以及所有的元分析文献，所以页数较多，但正文实际字数约为 9700 字。

除了以上修改外，我们也对中英文摘要进行了调整，并增删了部分参考文献。新增文献为：

- Ali, M., Ng, Y. L., & Kulik, C. T. (2014). Board age and gender diversity: A test of competing linear and curvilinear predictions. *Journal of Business Ethics, 125*, 497–512.
- Andrevski, G., Richard, O. C., Shaw, J. D., & Ferrier, W. J. (2014). Racial diversity and firm performance: The mediating role of competitive intensity. *Journal of Management, 40*, 820–844.
- Connelly, B. L., Tihanyi, L., Crook, T. R., & Gangloff, K. A. (2014). Tournament theory: Thirty years of contests and competitions. *Journal of Management, 40*, 16–47.
- Daniel, S., Agarwal, R., & Stewart, K. J. (2013). The effects of diversity in global, distributed collectives: A study of open source project success. *Information Systems Research, 24*, 312–333.
- de Poel, F. M., Stoker, J. I., & Van der Zee, K. I. (2014). Leadership and organizational tenure diversity as determinants of project team effectiveness. *Group & Organization Management, 39*, 532–560.
- Du, X. (in press). Does confucianism reduce board gender diversity? Firm-level evidence from China. *Journal of Business Ethics*.
- Hoch, J. E. (2014). Shared leadership, diversity, and information sharing in teams. *Journal of Managerial Psychology, 29*, 541–564.
- Hofstede, G. (1984). Cultural dimensions in management and planning. *Asia Pacific Journal of Management, 1*, 81–99.
- Li, W., Li, D., Meng, D., & Han, X. (2014). An empirical study on TMT heterogeneity and firm performance in Chinese manufacturing companies with declined performance. *Chinese Journal of Management, 11*, 819–827.
- Li, W., & Wu, D. (2014). Study on the relationship between demographic characteristics of new CEO and improvement of corporate performance: TMT heterogeneity moderation effect. *Chinese Journal of Management, 11*, 1158–1167.
- Lin, X., Bai, X., & Lin, L. (2014). Effects of similarity and accuracy indices of shared mental models on team creativity. *Acta Psychologica Sinica, 46*, 1734–1747.

- Lin, X., Lin, L., Wang, Y., & Bai, X. (2014). Empowering leadership and team performance: The transitive memory system as a medium variable. *Management Review*, 26(1), 78–87.
- Perry-Smith, J. E., & Shalley, C. E. (2014). A social composition view of team creativity: The role of member nationality-heterogeneous ties outside of the team. *Organization Science*, 25, 1434–1452.
- Ruiz-Jiménez, J. M., del Mar Fuentes-Fuentes, M., & Ruiz-Arroyo, M. (in press). Knowledge combination capability and innovation: The effects of gender diversity on top management teams in technology-based firms. *Journal of Business Ethics*.
- Wei, X., & Liu, Y. (2014). A mediated moderation analysis of the relations among past performance, team efficacy and team conflict. *Science of Science and Management of S. & T.*, 35(9), 152–160.
- Zona, F. (2014). Board leadership structure and diversity over CEO time in office: A test of the evolutionary perspective on Italian firms. *European Management Journal*, 32, 672–681.

第二轮

审稿人 1 意见：

意见 1：我认为经过修改，论文显著完善。在修改稿中，作者出于篇幅角度考虑，删去了对“绩效类型”这一调节变量的分析。我个人觉得，讨论绩效类型——常规绩效或创新绩效——还是有意义的。而且将两者合并为单一指标，其含义不清。建议保留，当然最终还得由作者决定是否保留。

回应：我们遵循专家一的建议，在正文“4 结果”部分保留了“绩效类型”调节变量的分析，作为假设 1 的延展分析，并在正文“5 讨论”部分对这一调节变量的启示进行了详细讨论。这样可以在不大幅增加篇幅的情形下，让读者理解尽可能多的情境调节变量。具体如下：

正文第 6 页：“由于本研究的团队绩效既包含一般任务绩效（即一般任务过程的效率和产出），也包含创新绩效（即创新过程的效率和产出）(Bell et al., 2011; Joshi & Roh, 2009)，因此我们也比较了团队多元化与不同类型绩效之间关系的差异，即绩效类型的潜在调节作用（见表 3）。可以看到，多样型多元化与团队创新绩效（ $\rho=0.12$, $p<0.001$ ）和任务绩效（ $\rho=0.05$, $p<0.01$ ）的正相关程度存在显著差异（ $Q_B=6.43$, $p<0.001$ ），而不平等型多元化与团队创新绩效（ $\rho=0.05$, $p<0.05$ ）和任务绩效（ $\rho=-0.02$, $p=0.18$ ）的关系也存在显著差异（ $Q_B=5.87$, $p<0.05$ ），说明绩效类型能够显著调节多元化与团队绩效之间的关系”。

表 3 绩效类型对团队多元化与团队绩效关系的调节作用

变量	地域	k	N	r	ρ	95%置信区间	Q_W	Q_B
分离型多元化	创新绩效	9	595	-0.06	-0.06	[-0.21, 0.08]	11.16	0.17
	任务绩效	24	2024	-0.03	-0.03	[-0.11, 0.06]	87.13***	
多样型多元化	创新绩效	50	14834	0.11	0.12***	[0.07, 0.16]	483.55***	6.43***
	任务绩效	140	33434	0.04	0.05**	[0.02, 0.07]	567.97***	
不平等型多元化	创新绩效	20	32011	0.04	0.05*	[0.00, 0.09]	93.48***	5.87*
	任务绩效	102	81102	-0.02	-0.02	[-0.04, 0.01]	498.62***	

注：[†] $p<0.100$ ；* $p<0.050$ ；** $p<0.010$ ；*** $p<0.001$ ； k 表示效应值个数； N 表示团队个数； r 表示未修正的平均相关系数； ρ 表示修正的总体相关系数；置信区间为基于修正的总体相关系数的 95%置信区间； Q_W 为组内异质性检验统计量； Q_B 为组间异质性检验统计量

正文第 8 页：“需要指出的是，团队多元化与不同类型绩效的关系存在显著差异。多样

型多元化与创新绩效的相关性高于与一般任务绩效的相关性,这说明任务取向多样有利于促进新想法的产生,但在转化为最终团队输出的时候,其效力有所下降。令我们感到意外的是,不平等型多元化能够显著提升团队创新绩效,但与任务绩效关系并不显著。这可能得益于社会等级带来的沟通效率的提升(Magee & Galinsky, 2008),进而改善团队的创新绩效水平,但与此同时,不平等可能对团队成员态度造成消极影响(Harrison & Klein, 2007),进而削弱了创新向任务结果的转化。因此,我们建议研究者在进行团队多元化的相关研究时,需要明确区分团队结果变量类型以及多元化类型,提升研究结论的针对性和可靠性”。

由于增加了一个表格,我们对正文中的表号也进行了相应调整。

审稿人 2 意见:

意见 1: 篇幅还是过长建议简化。语言上可以更精简。

回应: 本文是定量元分析,涉及的主效应和调节变量较多,为了保证理论基础及机理解释的充分性,篇幅相对较长。我们遵循专家二的建议,尽量在不改变原文内容的基础上,对文字进行了精简(正文删除了大约 800 字)。遵循的原则如下:

- ①用更简洁的语言表达文中的一些描述;
- ②对表格注释内容进行了简化;
- ③删除部分仅引用一次的参考文献。

由于遵循专家一的建议,增加了一些内容(约 1000 字),调整后正文字数约为 9900 字(不包含图表、摘要和参考文献),符合《心理学报》对正文字数的要求。

意见 2: 两个调节变量的选取有些没有系统性,尤其是国家地域的调节作用,建议可以删除。

回应: 元分析的一个重要目的就是通过已有的大量实证研究,找出导致研究结论不一致的调节变量,这是近期团队多元化研究者所倡导的思想(如 Guillaume 等, 2013, *Journal of Occupational and Organizational Psychology*)。这也是为什么专家一认为应当保留“绩效类型”这一调节变量的重要原因。目前尚未有元分析对不同国家地域情境下的多元化-绩效关系进行比较,因此,国家地域的调节作用是一个非常重要的理论贡献。这一分析结果无论是对方研究者还是西方研究者都具有很重要的理论参考意义。考虑到心理学报的国际影响力,我们认为国家地域的分析是有助于国际同行了解东方的管理理论和实践。

专家二的质疑说明我们对调节变量选取的描述不够,我们在引言部分增加了关于这一考虑的描述,详见正文第 1 页:“从研究情境的角度来看,各个研究都有其特有的情境,如不同的研究可能采用不同类型的团队为样本。同时,近年来,东方国家研究者对团队多元化的关注与日俱增,其中一些研究发现基于西方国家的多元化理论和实践可能并不适用于东方国家。这说明不同情境下的多元化对团队结果的影响可能会有所差异……”。

我们在讨论部分也对这一考虑进行了说明,详见正文第 8 页第 3 段第 1 句话。同时,英文摘要对于调节变量选取的理由和结果也进行了充分的说明。

意见 3: 语言要更规范化,(截止 2015 年 2 月底,谷歌学术引用率为 1028 次)这样的表达没有必要。

回应: 已经删除除了这一不规范的表达,详见正文第 2 页第 1 段。

同时,作者们也认真诵读了全文,对其他不规范的表达、错别字进行了修改。

第三轮 主编终审意见

意见 1: 作者在考虑研究的概念模型时, 只考虑到“团队多元化”有 3 种分类(水平), 而“团队绩效”并没有做相应的分类。待分析结果不理想后, 才想到“团队绩效既包含一般任务绩效(即一般任务过程的效率和产出), 也包含创新绩效(即创新过程的效率和产出)(Bell et al., 2011; Joshi & Roh, 2009)”。照说, “团队绩效”亦是林林总总, 这也是常识。我的意见是, 要么从头开始就将分类型的“团队绩效”纳入模型, 要么很醒目地“acknowledge this limitation”。

回应: 感谢主编的建议, 我们将“不同类型的团队绩效”纳入了我们的研究模型。在摘要、引言、假设、方法、结果、讨论、结论、英文摘要部分分别增加或修改了相关的描述。

由于增加了相关假设, 字数有所增加, 故对原文的文字进行了精简, 调整后正文字数约为 9995 字(不包含图表、摘要和参考文献), 符合《心理学报》对正文字数的要求。

相关修改如下:

(1) 摘要

正文第 1 页: “调节效应检验显示, 团队人口统计特征多元化与不同类型绩效的关系存在显著差异, 且……”。

(2) 引言

正文第 1 页: “其次, 从研究情境角度来看, 各个研究都有其特有的情境, 如不同研究可能采用不同类型的团队绩效作为结果变量, 也可能采用不同类型的团队作为研究样本……”。

(3) 假设

正文第 3-4 页, 增加了“2.3 绩效类型的调节作用”相关假设描述:

“团队多元化与绩效的关系可能会随着绩效类型的不同而不同。常见的团队绩效可以分为创新绩效(即创新过程的效率和产出)和一般任务绩效(既一般任务过程的效率和产出)两类(Bell et al., 2011; Joshi & Roh, 2009)。从不同类型的多元化来看, 多样型多元化能够整合团队成员不同的信息来源, 有利于提升团队创新绩效(Van Knippenberg & Schippers, 2007)。然而, 在创新过程中, 多样型多元化往往会引发消极的团队冲突(Qian et al., 2013), 进而降低团队任务绩效(de Wit et al., 2012)。因此, 多样型多元化与创新绩效的正向关系可能会强于其与一般任务绩效的关系。从分离型多元化和不平等型多元化来看, 分离和不平等会降低团队成员的合作意愿, 这种不合作既不利于团队创新, 也会进一步影响团队任务绩效表现。此外, 分离和不平等会产生消极的交互过程, 如较低的凝聚力(Thatcher & Patel, 2011)和较高的不公平感(Halevy, Chou, & Galinsky, 2011), 更加不利于任务绩效的提升。因此, 与创新绩效相比, 分离型多元化和不平等型多元化与一般任务绩效的负面关系更强。故提出如下假设:

假设 2: 绩效类型会调节团队多元化与团队绩效的关系, 即多样型多元化与创新绩效的正相关程度高于一般任务绩效, 而分离型多元化、不平等型多元化与团队一般任务绩效的负相关程度高于创新绩效。”

正文第 4 页: 研究模型也进行了相应的调整, 详见图 1。

由于增加了假设, 故原文中的假设序号均进行了相应的调整。

(4) 方法

正文第 5 页: “绩效类型、国家地域和团队类型依据各论文研究方法部分的样本描述进行识别。绩效类型主要包含一般任务绩效和创新绩效。其中, 一般任务绩效主要是指一般任务过程的效率和产出(Joshi & Roh, 2009), 而创新绩效主要是指创新过程的效率和产出(Bell et al., 2011)……”。

(5) 结果

正文第 6 页（上一版本已论述，本轮独立为一个新的段落）：“由于本研究的团队绩效既包含一般任务绩效，也包含创新绩效(Bell et al., 2011; Joshi & Roh, 2009)，因此我们也比较了团队多元化与不同类型绩效之间关系的差异，即绩效类型的潜在调节作用（见表 3）。可以看到，多样型多元化与创新绩效（ $\rho=0.12$, $p<0.001$ ）和任务绩效（ $\rho=0.05$, $p<0.01$ ）的正相关程度存在显著差异（ $Q_B=6.43$, $p<0.001$ ），而不平等型多元化与创新绩效（ $\rho=0.05$, $p<0.05$ ）和任务绩效（ $\rho=-0.02$, $p=0.18$ ）的关系也存在显著差异（ $Q_B=5.87$, $p<0.05$ ），说明绩效类型能够显著调节多元化与绩效之间的关系。因此，假设 2 得到部分验证。”

（6）讨论

正文第 8 页（上一版本已论述，本轮独立为一个小节）：

“5.2 绩效类型对多元化-绩效关系的调节作用

调节效应检验结果显示，团队多元化与不同类型绩效的关系存在显著差异。多样型多元化与创新绩效的相关性高于与一般任务绩效的相关性，这说明任务取向多样有利于促进新想法的产生，但在转化为最终团队输出的时候，其效力有所下降。令我们感到意外的是，不平等型多元化能够显著提升团队创新绩效，但与任务绩效关系并不显著。这可能得益于社会等级带来的沟通效率的提升(Magee & Galinsky, 2008)，进而改善团队创新绩效水平，但与此同时，不平等可能对团队成员态度造成消极影响(Harrison & Klein, 2007)，进而削弱了创新向任务结果的转化。因此，我们建议研究者在进行团队多元化的相关研究时，需要明确区分团队结果变量类型以及多元化类型，提升研究结论的针对性和可靠性。”

（7）结论

正文第 10 页：“（2）从绩效类型来看，多样型多元化、不平等型多元化与创新绩效的相关性显著高于其与一般任务绩效的相关性；（3）……”。

（8）英文摘要

增加了关于绩效类型作为调节变量的描述，详见正文第 17-18 页。