

《心理学报》审稿意见与作者回应

题目：我国大中学生道德推脱水平的变迁及宏观成因

作者：王祥坤 辛自强 侯友

第一轮

审稿人 1 意见：

本研究选题新颖，兼具学术和现实关切，值得称赞！然而，目前仍有以下问题与作者商榷。

回应：

感谢审稿专家对本研究选题的认可。下面对审稿专家指出的问题逐一说明修改情况。

1、概念界定：就字面而言“道德推脱”似在描述某种策略或行动，但实际上作者将之定义为一种个体特质或倾向（“道德推脱水平”）。我建议作者的概念界定应更清晰（比如区分特质和具体情境下的策略），并在前言 1.1 节更多介绍与道德推脱特质有关的理论和研究。

回应：

感谢审稿专家的建议。我们的研究将“道德推脱”定义为一种认知倾向（稳定的人格特质），并从社会变迁的角度来考查道德推脱水平的变迁趋势。我们按照审稿专家的建议重新梳理了引言 1.1 节的内容，重点强调了道德推脱理论与（基于道德推脱特质的）操作性定义等内容。具体修改内容请见第 13 页第 2 段至第 14 页第 1 段的蓝色字体部分。

2、概念操作化：作者对“社会个体化水平”以及“社会正义水平”的操作化界定值得商榷。例如，作者首先说“个体化是个体脱离传统家庭和集群社会进入市场社会的过程……个体化的核心是个体对自身行为做出规划的能力”（p.7），这是很微观（心理层面）的概念。但后来却用市场经济发展水平、社会城镇化水平和离婚率来度量社会的个体化水平，存在巨大的分析水平跳跃（从微观到宏观）。个体化怎么就成了一个宏观的、社会结构层面的概念？作者需做更多解释。就“社会正义水平”而言，作者将“现行有效立法量作为社会正义水平的代理指标”（p.8）。然而，正如“形式民主”不一定等于“实质民主”，法律条文的数量就等于法制的质量？

回应：

感谢审稿专家指出概念操作上的问题，原稿的写作确实表述得不清楚。我们重新修改完善了论文以理顺概念的操作化定义。这里大致说明修改的思路。

第一个问题，个体化的微观与宏观构念之间的关系问题。二者的逻辑关系是，宏观社会结构层面的个体化会影响人的心理与行为个体化，前者是后者的重要推动力量。同时，我们想要重点强调的是社会层面的个体化。对此，我们在文中 1.4.1 部分做了全面的修改和调整，具体修改内容请见第 17 至 18 页的蓝色字体部分。

第二个问题，我们认为，若基于中国当下法治社会建设的现实情况来看，现行有效立法量作为社会正义水平的代理指标应该是合理的。立法是法治之先导，良法是善治之前提。若没有完善的法律条文，显然更可能存在社会正义水平低下问题，因此，从纵向来看，历年来立法数量的增加能在相当程度上指示社会正义水平的变化。对此，我们在文中补充了更详细的论证分析。具体内容请见第 19 页第 1 段的蓝色字体部分。

3、前言存在一些似是而非的逻辑推导。例如，“责任心强的个体具有更高的道德水平(Cohen et al., 2014)，也更加重视公平公正、强调规则意识、遵守道德规范(Ogunfowora et al., 2022)；而个体的责任心越弱则对道德规范的敏感性越低(De Angelis et al., 2016)。这说明责任心越强的个体道德责任意识也越强，并通过激活道德认知调节机制来抑制道德推脱。”(p.6) 我不清楚前文怎么就能支撑“这说明.....”之后的结论？再例如，“已有研究发现，在组织情境中，组织正义(员工对公正事件的公平感)会抑制员工的道德推脱(Lee et al., 2016; Moore & Gino, 2013)。”(p.8) 作者讨论的是“社会正义”这一“宏观社会因素”，怎么又用一个微观概念(“员工对公正事件的公平感”)来论证？第三个例子，“以往研究均发现道德推脱预示着更多不道德行为(Chen et al., 2016; Kavussanu & Stanger, 2017; Kowalski et al., 2014; Ma et al., 2019; Machackova, 2020)。.....新冠肺炎疫情期间的追踪分析发现，更高的道德推脱水平与更少的疫情防御行为有关(Bazzoli & Probst, 2022)。”(p.4) “更少的疫情防御行为”怎么就算作“更多不道德行为”了？

回应：

感谢审稿专家指出的问题。我们将以上这些不准确或不恰当的表达逻辑做了修改或删除(见第 16 页第 2 段与第 18 页最后一段中的蓝色字体部分)。此外，我们通读全稿，对类似逻辑推理不恰当之处一并做了修改完善。针对审稿专家指出的以上三点问题，具体修改内容如下：

第一点：“责任心强的个体具有更高的道德水平(Cohen et al., 2014)，也更加重视公平公

正，强调规则意识，遵守道德规范(Ogunfowora et al., 2022)；而个体的责任心越弱则对道德规范的敏感性越低(De Angelis et al., 2016)。这意味着责任心可能会通过不断强化人们的道德责任意识来抑制其自身的道德推脱倾向。”

第二点：“与社会正义对应的微观概念是组织正义，分析它与道德推脱的关系有助于揭示社会正义的约束机制。例如，已有研究发现，在组织情境中，组织正义会抑制员工的道德推脱行为(Lee et al., 2016; Moore & Gino, 2013)。”

第三点：因该例子不恰当，我们将其与文后的参考文献一并删除。

4、统计分析问题：横断历史元分析的数据点可能存在自相关（autocorrelation）或者不独立（同一年数据更相近），导致相关和回归分析的结果可能不准确（假阳性）。不知作者有没有做相应的稳健性检验或统计参数校正？比如采用 ARIMA 算法或者稳健的标准误（Robust SE）估计？

回应：

基于上述建议，我们采用稳健标准误的方法补充分析了道德推脱对年代的回归分析以排除假阳性的问题。我们以 R 语言 `fixest` 包计算怀特稳健标准误，并在原文中补充了上述分析结果，具体内容见第 23 页第 1 段和第 26 页第 1 段的蓝色字体部分。

次要问题：

1、本文的一些行文断句还需润色。例如，“道德是社会关系的重要调节力量，随着我国经济社会的持续发展，社会道德水平发生了怎样的变迁，有人认为是道德滑坡(孙立平, 2007)，但随着社会持续变迁道德水平也可能发生积极变化。”(p.3) 这里有的逗号该标成句号或问号。

回应：

感谢专家的建议。我们将这句话重新组织断句，确保语言表达措辞更准确，并在文稿的正文部分做了修改，见第 13 页引言首段的蓝色字体部分。此外，我们对照审稿专家的建议通读全稿，对类似行文断句表达不明确的语句一并做了修改完善，以增强文稿的可读性。

2、元分析不充分。例如，“(2) 对文献的其他信息进行编码，包括文献类型（1=核心刊物，2=一般刊物，3=学位论文或论文集，4=外文期刊）、数据收集地区（1=西南地区，2=西北地区，3=东北地区，4=华北地区，5=华中地区，6=华东地区，7=华南地区，8=未报告或者跨

地区)和被试生源地(0=无明确城乡信息,1=城市,2=农村,3=既有城市又有农村)等信息。”(p.10)后文中有利用这些信息进行分析吗?

回应:

感谢审稿专家的提醒。我们参考以往研究的做法,将文献类型、数据收集地区、被试生源地(转化成虚拟变量)作为协变量纳入回归模型,检验控制以上这些变量后年代对道德推脱水平的预测效果。补充的分析结果见第23页第1段(研究1)和第26页第1段的蓝色字体部分(研究2)。

3、研究一 Bandura 的量表有 32 个题目。所有题目都从属一个维度吗?没有分量表?为什么研究一不像研究二那样多用几个量表?

回应:

研究1中,我们采用了实证研究中使用较多的32题版的道德推脱量表。该量表共有8个分维度构成,这些分维度的项目可以合成一个总的维度(即8个分维度的一阶因子模型可以合成得到1个二阶因子模型)。鉴于已有实证研究文献中绝大多数只提供了总维度的数据,所以我们的分析中只用了总维度得分。

此外,研究1没有像研究2那样用多个量表的数据是出于两点考虑。首先,我们在做数据收集时,汇总发现在相同的年代跨度范围内,其他量表获得的数据量有限(有效的数据量均小,且这些数据集中在个别年代),所以未将这些量表的数据纳入分析。例如,Bandura的26题版量表的有效数据只有9个;Caprara等32题版的公民道德推脱量表有效数据只有3个。其次,目前已有75篇文献共81组数据,这符合横断历史元分析的文献数量要求。而添加少量其他量表的数据会引入新的变异源,又无法实现有效的统计控制,所以保持这一个量表数据结果会更合理。

4、时间滞后分析中3年、1年和当年的预测项是同时回归吗?

回应:

两项研究中,时间滞后分析的预测项均是分别单独做的回归分析,不是同时回归分析。我们也在文中补充说明了具体操作方法。

.....
审稿人2意见:

《近十年来我国大中学生道德推脱水平的变迁》审稿意见 该文通过横断历史元分析的

方法,探讨了中学生和大学生群体道德推脱水平的变迁趋势及其与反映社会个体化水平和正义水平的宏观指标的关系。但是该文在结果计算方面存在较大问题。坦白而言,本人未曾进行过横断历史元分析的研究,也未曾以横断历史元分析为主要方法发表过学术论文(但是发表过一般元分析的文章)。拜读贵稿件之后,本人对文章的结果产生了较大的疑惑,但是也说不清楚问题出在哪里。之后利用一段时间认真阅读了国内有关横断历史元分析的文章,逐渐明白了该稿件的问题所作,现将我的思考撰写如下。

回应:

感谢审稿专家花费宝贵的时间来审阅我们的论文以及所提的问题。这为我们提高论文质量提供了宝贵的机会。下面对审稿专家提出的问题逐一说明修改情况。

第一,研究结果与我的常识存在很大冲突。该文发现:“11年来中学生道德推脱下降了1.76个标准差”,“12年来(2008至2020年)大学生道德推脱水平下降了2.76个标准差”。一方面,这与一般统计学知识相违背。本人很难想象一个心理变量的平均值会下降2.76个标准差(更何况本文采用了Z分数计算,标准差是1)。另一方面,大学生和中学生道德推脱的均分一般为2.00左右,标准差为0.50左右(我的主观认识)。如果下降1.76个标准差或2.76个标准差,意味着这些群体就不存道德推脱了。因为 $2.00 - 0.50 \times 2.76 = 0.62$,小于1,但是道德推脱是1-5计分。举个例子,杨继平等(2013)一文显示,大学生道德推脱平均分是2.13,标准差为0.56,如果按照该文的平均下降速度,大学生的道德推脱得分应该是 $2.13 - 0.56 \times 2.76 = 0.5844$,这明显不对(道德推脱是1-5计分)。[参考文献:杨继平 王兴超 杨力. 观点采择对大学生网络偏差行为的影响:道德推脱的中介作用[J]. 心理科学, 2014, 37(3): 633-638.]

回应:

感谢审稿专家的提示。经过反复核查数据和推算效应量 d 的计算过程,我们比较确信的是按照已往的经验做法,本研究中计算得到的效应量结果并无错误。经过认真分析和推算,我们认为这里出现Z分数计算得到效应量 d_z 的数值出现系统性增大的原因是它与原始分数计算得到的效应量 d 之间因标准差的不同单位造成的。因此,Z分数的效应量 d_z 在判定效应大小时就不能用以以往有关原始分数计算效应量大小的经验标准。若我们继续采用以往的Cohen's d 的经验判定,就会觉得不合理。下面我们给出Z分数效应量 d_z 与原始分数效应量 d 之间的数量关系以说明上述问题,并以研究1的数据进行举例分析。

(1) 推导Z分数与原始分数分别计算变化量 d_z 与 d 值之间的数量关系如下:

首先假定：起始年的原始分数预测值为 X_1 ；终止年的原始分数预测值为 X_2 。而由 Z 分数计算公式：

$$Z = \frac{X - M}{SD} \quad (\text{公式 1})$$

可知， X_1 和 X_2 的预测值所对应的标准分 Z 分数分别为：

$$Z_1 = \frac{X_1 - M}{SD} \quad (\text{公式 1a}) \quad Z_2 = \frac{X_2 - M}{SD} \quad (\text{公式 1b})$$

已知原始分数计算的效应量值的计算公式为：

$$d = \frac{X_2 - X_1}{SD_M} \quad (\text{公式 2})$$

则套用公式 2 的计算方法，求得 Z 分数变化量的计算公式为：

$$d_z = \frac{Z_2 - Z_1}{SD_Z} \quad (\text{公式 3})$$

接下来，我们将公式 1a 和 1b 带入公式 3 替换标准分 Z_1 和 Z_2 得到：

$$d_z = \frac{\frac{X_2 - M}{SD} - \frac{X_1 - M}{SD}}{SD_Z}$$

整理化简后得到：

$$d_z = \frac{\frac{X_2 - X_1}{SD}}{SD_Z} \quad (\text{公式 4})$$

由公式 2 可得到其变式：

$$X_2 - X_1 = d \times SD_M \quad (\text{公式 2a})$$

再将公式 2a 带入公式 4 可得：

$$d_z = \frac{d \times SD_M}{SD_Z} \quad (\text{公式 5})$$

已知标准分 Z 分数的标准差为 1；则 SD_Z 为标准分 Z 分数的标准差的平均值也为 1。因此，公式 5 化简为：

$$d_z = \frac{SD_M}{SD} \times d \quad (\text{公式 6})$$

由公式 6 可知，转换为 Z 分数后计算的效果量 d_z 与原始分数计算的效果量 d 之间相差倍数关系为 SD_M / SD ，其中，原始得分的平均标准差 SD_M 要比计算 Z 分数时均值的标准差 SD 要大。因此， Z 分数计算的效果量（变化量 d_z ）要比使用原始分数所得的效果量（变化量 d ）要大。

(2) 以研究 1 为例带入数据计算变化量 d_z 与 d 值之间的数量关系如下：

研究 1 中原始分数计算中学生道德推脱水平下降趋势的变化量 d 为 0.62；而采用 Z 分数计算其变化量 d_z 为 1.54。我们已知中学生道德推脱水平的原始得分的平均标准差（标准

差的平均值) SD_M 为 0.53; 而中学生道德推脱水平的原始得分均值 M 的标准差 SD 为 0.17。由此可知, 道德推脱原始分数获得的变化量 d 通过套用公式 6 推算得到 Z 分数的变量 $d_z = 0.53/0.17 * 0.62$, 结果约为 1.93。此结果比用 Z 分数计算所得变化量值 1.54 比较接近(数值不一致是因每个计算步骤保留两位有效数字引起的误差), 并且这与原始得分计算的变化量相差近 2.50 至 3.00 倍左右。此时, 若再使用原始的 Cohen's d 值效应量的经验标准来判定 Z 分数计算所得效应量, 就会发现 d_z 效果量过大, 这会引起误解。

因此, 为了避免这个问题, 我们不再使用 Cohen's d 的经验标准来判断 Z 分数计算所得效应量, 而是直接使用标准分数来描述效果量即可, 即中学生道德推脱水平下降了 1.54 个标准分。同理, 研究 2 中效果量也应以标准分来描述效果量, 即大学生道德推脱水平下降了 1.92 个标准分。修改后的效应量计算过程与结果见第 23 页的第 2.2.2 小节和第 26 页的第 3.2.2 小节中蓝色字体部分。

此外, 审稿专家认为: “另一方面, 大学生和中学生道德推脱的均分一般为 2.00 左右, 标准差为 0.50 左右(我的主观认识)。如果下降 1.76 个标准差或 2.76 个标准差, 意味着这些群体就不存道德推脱了……”。我们认为审稿专家这里所做的分析是有待商榷的。第一, 这里计算的效应值是基于 Z 分数获得的 d , 其含义依然表达的是相对量数而不是像原始分数那样基于李克特量尺记分所具有的实质性含义, 因此并不能直接拿来与原始得分进行对照分析。第二, “因为 $2.00 - 0.50 * 2.76$ 等于 0.62, 小于 1, 但是道德推脱是 1-5 计分。”这里所做的计算可能并不准确(在统计计算中并无直接将 d 值带入计算的方式)。因此, 由此得出道德推脱就不存在了的结论可能并不正确, 这一点仍需与审稿专家商榷。

第二, 我发现该文要么在 Z 分数计算出问题了, 要么在平均标准差出问题了。具体而言, 基于第一条件的疑惑, 我分析了该文的效应值计算问题—— d 值。其过程如下: “首先, 以中学生道德推脱 Z 分数为因变量(y), 以年代为自变量(x), 对样本量进行加权, 从而建立回归方程: $y = Bx + C$ (具体方程结果为 $y = -0.16x + 314.85$)。其次, 分别将年代 2010 和 2021 代入回归方程获得中学生道德推脱的 Z 分数 Z_{2010} 和 Z_{2021} 。最后, 计算 Z_{2010} 和 Z_{2021} 之差, 再除以 11 年间的 Z 分数的平均标准差 1 (因为标准分的标准差为 1, 所以 11 年间的平均标准差也为 1), 即可得到 d 值 ($d = (Z_{2010} - Z_{2021})/1$)。结果表明, 11 年来中学生道德推脱下降了 1.76 个标准差。”我根据回归方程 $y = -0.16x + 314.85$, 分别将年代 2010 和 2021 代入回归方程获得中学生道德推脱的 Z 分数。其中 Z_{2010} 为 8.51, Z_{2021} 为 6.75。我不太明白作者的计算过程, 但是根据统计学知识: 如果这是中学生道德推脱的 Z 分数, 那

么 2010 年中学生道德推脱的 Z 分数无论如何也不能是 8.51 吧？同样 2021 中学生道德推脱的 Z 分数也不能是 6.75 吧？如果 2010 年和 2021 年中学生道德推脱的 Z 分数计算没有错误，那么“11 年间的 Z 分数的平均标准差为 1（因为标准分的标准差为 1，所以 11 年间的平均标准差也为 1）”可能是错误的？该文没有呈现大学生 d 值的计算过程，也没有呈现出大学生道德推脱 Z 分数的回归方程。但是我武断地推测，作者大学生的计算过程和中学生一样，也应该是存在错误的。

回应：

由上一条回复意见中的推导结果可知，转换为 Z 分数后计算的效果量 d_z 与原始分数计算的效果量 d 之间相差倍数关系为 SD_M / SD 。这里 Z 分数得分按照回归方程的求解过程在数值上没有错误；而且根据标准分的性质可知，标准分的标准差为 1 其平均标准差也是 1 也无不当之处。大学生 d 值的计算过程与研究 1 中中学生的计算过程相同。我们的分析说明，使用 Cohen's d 的经验标准来判断 Z 分数计算所得效应量不再合适，而更合理的做法是直接使用标准分数来描述效果量。

第三，下降标准差的计算结果与图 1 和图 2 散点图的直观感觉相冲突。例如，图 2 中 2008 年和 2020 年在 Y 轴上对应的 Z 分数值，两者相差怎么看也不到 2.76 个标准差？以上三点是我的直观思考，并无十分严格的统计计算过程，但我认为我的逻辑推理和分析是没有问题的。如果我的逻辑推理和分析有错误，愿意接受作者反诉，并向作者致歉。

回应：

诚如审稿专家所言，图 2 直观上来看，确实没有 2.76 个标准差那么大的变化。这里是因研究 2 中采用了样本量加权回归方程来计算效应量 d_z 的数值有关。修改后的论文中，我们对道德推脱水平（Z 分数）随年代的变化量所使用的估计方程不再采用加权样本量回归分析，从而使得 d_z 不受样本量加权回归的干扰。以上是我们对审稿专家指出的问题所做的回应。不妥之处，恳请审稿专家进一步批评指正。

其它问题：鉴于结果计算方面存在较大问题，我简单说一下文章存在的其它问题。

1. 文章的逻辑基础存在较大问题。从哲学上讲，不道德的对立面并不一定是道德。同样，道德推脱的降低，并不一定意味着社会道德水平的上升。

回应：

感谢审稿专家的提示。我们在文中补充了 1.2 节专门讨论道德推脱与社会道德水平的关

系。具体请见第 14 页第 1.2 小节中的蓝色字体部分。

本文论证道德推脱水平降低与社会道德水平上升逻辑关系的推导直观上来看不那么直接，但可以从间接的逻辑推理得出。我们做此逻辑推导是基于以下两方面原因。第一，这里将道德推脱作为反映社会道德水平的指标，其内在的逻辑关联是道德推脱反映不道德行为，不道德行为反映社会道德水平。道德推脱反映（或解释）不道德行为，是道德推脱理论阐述的核心观点；而且已有大量的实证研究也揭示了道德推脱作为不道德行为的前置因素的证据。文中已做了讨论，这里不再赘述。而不道德行为反映社会道德水平，可以从三方面来理解。首先，道德哲学上，道德与不道德是基于价值判断的。道德给予我们与他人一起生活的规则，为我们的欲望和行动设定了界限，它告诉我们什么是允许的，什么是不被允许的；道德赋予我们做决定的指导原则，告诉我们什么应该做，什么不应该做(所罗门, 希金斯, 马丁, 2019)。由此可见，道德全然关注的是对错，是价值判断。我们由此可做如下判断：对道德规则的遵从就是道德，而对道德规则的违反就是不道德。其次，心理学的实证研究中（尤其是在社会心理学研究中），常常将道德行为描述为亲社会性行为，而将不道德行为描述为反社会性行为。我们也往往将二者看作是一体两面（如同一个硬币的两面）。最后，也是最重要的是，在社会现实层面，社会道德水平的高低是由一个社会的道德底线决定的；而道德底线取决于社会中人们的不道德行为的反社会性程度。反社会性程度越大意味着社会道德水平越低；反之，则意味着社会道德水平越高。可见，不道德行为是社会道德水平的指示器。第二，本研究中道德推脱的社会变迁分析是建立在宏观分析水平上。基于群体层面的横断历史元分析，可以概括出群体（宏观）层面（或进一步解释）社会变迁现象(蔡华俭 等, 2023; 辛素飞 等, 2022; 辛自强, 池丽萍, 2008)。基于群体层面（或宏观水平）的道德推脱水平变迁趋势可以作为揭示社会道德水平的代理指标。因此，综合以上两点理由，我们认为建立在社会宏观层面的道德推脱水平的降低趋势在一定程度上是可以反映社会道德水平的上升。对此，恳请审稿专家批评指正。

参考文献：

- 蔡华俭, 张明杨, 包寒吴霜, 朱慧珺, 杨紫嫣, 程曦, 黄梓航, 王梓西. (2023). 探索社会变迁过程中的心理与行为变化: 研究设计与分析方法. *心理科学进展*, 31(1), 在线出版.
<https://doi.org/10.3724/SP.J.1042.2023.001>
- 所罗门, 希金斯, 马丁. (2019). *哲学导论: 综合原典教程*(第 11 版)(陈高华译). 天津: 天津人民出版社.
- 辛素飞, 梁鑫, 盛靓, 赵智睿. (2021). 我国内地教师主观幸福感的变迁(2002~2019): 横断历史研究的视角. *心理学报*, 53(8), 875–889. <https://doi.org/10.3724/sp.J.1041.2021.00875>
- 辛自强, 池丽萍. (2008). 横断历史研究: 以元分析考察社会变迁中的心理发展. *华东师范大学学报(教育科学版)*(2), 44–51. <https://doi.org/10.16382/j.cnki.1000-5560.2008.02.001>

2. 缺乏问题意识。文章没有说清楚为何要进行道德推脱的横断历史元分析。

回应:

感谢审稿人的建设性意见,我们进一步补充强调了本研究的必要性,并将其补充在引言的 1.5 小节部分(第 19 页第 2 段的蓝色字体部分)。

3. 中学阶段是青少年道德认知快速发展的关键时期,年龄是一个非常重要的干扰因素。在进行中学生道德推脱水平变化分析时,必须考虑年龄对结果的干扰和影响。

回应:

谢谢审稿专家的提示。我们结合审稿专家的意见,将年龄作为控制变量纳入道德推脱水平对年代的加权回归模型来检验年龄对中学生道德推脱水平变迁是否有干扰。原始研究的 79 组数据中,有 47 组数据报告了年龄信息(这些数据在 2010 至 2021 年均有分布)。我们以这 47 组有效数据进行层次回归分析,以此来考查控制年龄后年代是否还能显著预测中学生道德推脱水平的下降趋势。上述分析结果请见第 23 页的第 1 段蓝色字体部分。

4. 道德推脱存在明显的性别差异,需要考虑性别的调节作用。

回应:

诚如审稿专家所言,道德推脱确实存在性别差异,分析时应该考虑这一因素的影响。然而,本研究中只收集到提供男女生道德推脱得分这一信息的少量原始文献。具体而言,研究 1 中,中学生组只有 30 组提供了男女生道德推脱得分(共有 79 组数据);研究 2 中,大学生组只有 18 组提供了这一得分(共有 54 组数据)。而且,这些得分数据的年代分布多数集中在近期的某几年内,使得其不能从整体上考查性别对近十多年来大中学生道德推脱变迁的调节作用。因此,本研究未能在研究结果部分考查性别的调节作用。但是我们吸收了审稿人的这一建议,并在讨论部分的研究不足中补充了这一点并提出未来研究建议,期待未来的研究通过积累更多数据来解答这一问题。具体内容见第 30 页蓝色字体部分。

5. 英文摘要质量有待提高。整体而言,建议退稿。

回应:

感谢审稿专家的评阅。我们认真阅读修改了英文摘要,并请英文写作水平较好的专业同行润色修改。修改后的英文摘要的语言质量有了提升。以上问题,我们均做了修改完善,恳

请审稿专家进一步批评指正。

第二轮

审稿人 1 意见：

非常感谢作者的仔细修改，较好地回应了我之前的大部分疑问。但仍有以下问题需作者进一步考虑。

回应：

感谢审稿专家对论文修改工作的肯定。下面进一步对审稿专家所提问题逐一说明修改情况。

1、关于社会个体化。我认同作者的以下回应：“宏观社会结构层面的个体化会影响人的心理与行为个体化，前者是后者的重要推动力量”。但在 1.4.1 节中，作者仍然没有把这个关系讲清楚。在我看来，作者首先要给出“社会个体化”的定义，其次介绍宏观社会结构层面个体化的表现，最后介绍宏观变化在心理层面（包括道德推脱）的影响。

回应：

感谢审稿专家的建议。我们遵照专家所提的写作逻辑对这部分内容做了修改，具体内容请见第 22 页至 23 页的红色字体部分。

2、关于道德推脱的维度。作者在 1.1 节要明确道德推脱分维度之间的关系（比如：高度共变还是相对独立？）。进一步而言，八个分维度的变迁模式和机制是否相似？作者在前言部分要更小心（如 16 页，作者推测共情会抑制非人化，但能否据此推广为共情会降低道德推脱的整体水平？），在讨论部分似乎也应增加理论探讨（鉴于作者说无法做实证分析）。

回应：

感谢审稿专家的提示。从以往多方面的研究证据来看，道德推脱各维度之间应该是高度共变的。Bandura 等人(1996)认为，道德推脱各维度理论上具有一致的心理过程机制；实证研究中，道德推脱分维度之间以及与总维度都具有高相关性(Caprara et al., 2009)。例如，以青少年为研究对象的相关研究表明，道德推脱各维度的相关系数在 0.35 至 0.59 之间，各分维度与总维度相关系数在 0.67 至 0.76 之间(滕召军, 2018)。这说明道德推脱各维度之间具有

高共变性。若道德推脱各维度之间以及与总维度之间高共变性成立，非人化是道德推脱的子维度，研究表明共情会抑制非人化；那么据此可以推测共情也会降低整体的道德推脱水平。我们在引言 1.1 节中做了适当补充，具体内容见第 19 页红色字体部分。

此外，根据审稿专家的建议，我们在本研究的讨论部分适当补充了道德推脱各维度的变迁趋势的讨论（具体见第 36 页讨论部分 4.3 节的红色字体部分）。然而，这一问题更值得未来的研究工作进行实证分析检验，提供明确的研究证据。

参考文献：

- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71(2), 364–374. <https://doi.org/10.1037/0022-3514.71.2.364>
- Caprara, G. V., Fida, R., Vecchione, M., Tramontano, C., & Barbaranelli, C. (2009). Assessing civic moral disengagement: Dimensionality and construct validity. *Personality and Individual Differences*, 47(5), 504–509. <https://doi.org/10.1016/j.paid.2009.04.027>
- 滕召军. (2018). *暴力电子游戏对青少年道德认知的影响*(博士学位论文). 西南大学.

3、关于道德推脱与社会道德水平之间的关系。道德的面向很多，道德推脱只是一方面。建议作者通过道德推脱水平变迁推断社会道德水平变迁时应更谨慎，要明确可能存在的边界。比如，作者多次提到阎云翔的研究，并认为自己的研究能补充其质性研究的不足。但阎氏涉猎的问题很广泛，作者能否以自己的“点”来回应阎氏的“面”？

回应：

诚如审稿专家所言，阎氏的质性分析和理论描述是基于更宏大的历史叙事逻辑来对道德变迁现象进行阐述；而本研究从道德心理现象的某一个侧面来描述人们的道德心理的时代变迁问题。若单就这一点来说，我们的研究工作确实不能对其所研究的内容进行补充。事实上，我们的工作只是基于特定道德心理现象的来说的，并不能直接回应阎氏的研究问题。对此，我们在文中修改了相关表述。

尽管如此，我们依然要强调的是，本研究借助横断历史元分析方法从定量研究的角度来对道德心理变迁进行量化研究并建立预测模型；而以往的研究并未就道德心理变迁问题提供过类似的量化分析证据。总之，本研究为考查我国道德心理变迁趋势首次提供了清晰的量化研究证据，同时引入横断历史元分析方法，可能在某种程度上也具有一定的方法论意义。

4、感谢作者在回归分析中使用了怀特稳健标准误，但正文也应介绍使用该标准误能避免什么统计问题（“自相关”）。

回应:

感谢提醒，我们在文中相应位置补充了采用该校正方法的目的。

5、建议研究二做一下不同量表效应量的比较。

回应:

基于审稿专家的建议，我们在研究 2 中补充了三个量表各自的效应量结果。具体内容见 3.2.2 节中的红色字体部分。

.....
审稿人 3 意见:

该文章采用横断历史研究方法揭示了近 10 余年以来我国大中学生道德推脱水平的变化，并从我国社会个体化水平和社会正义水平这两类社会宏观因素来解释该群体道德推脱水平变化的宏观原因。该论文关注的大中学生道德推脱水平是当前社会和教育的热点问题和急需解决的问题，立德树人根本任务的核心就是立德，立德的前提是掌握大中学生道德水平现状与发展趋势。因此该文研究结果为大中学生立德教育工作提供了一项重要依据与支撑，应该说是一项满足国家重大急需的研究工作。总体上来说，该文研究问题具有较强的现实意义，研究范式合理，数据分析正确，研究结论为该领域提供了新的知识贡献。同时，也请作者考虑如下问题以进一步提升论文质量：

回应:

感谢审稿专家对论文选题现实意义和知识贡献的认可。下面对审稿专家提出的问题逐一说明修改情况。

1. 引言第一段作者从社会道德水平发生变迁设问引入，但是本文聚焦的道德推脱水平的变迁，道德推脱水平与社会道德水平之间并不是必然的一一对应的关系，因此这样引入在逻辑上不够清晰或直接，建议直接从道德推脱这一现象本身引入研究问题。

回应:

感谢审稿专家的建议。我们借鉴审稿专家的指导意见，对引言的第一段做了修改，使得文章从关注道德推脱本身来引入研究问题。具体内容见第 18 页的红色字体部分。

2. 前言第四段“在社会现实层面，道德水平的高低由一个社会的道德底线决定，而道德底线

取决于社会中不道德行为的发生数量。”这个有实证研究或理论文献支持吗？

回应：

感谢提示，我们已在文中补充了相关的参考依据。

3. 前言部分“1.4.2 社会正义水平的影响”第二段作者论证组织正义，似乎也没有采纳组织正义的指标来进行预测，作者试图用组织正义过渡到社会正义水平，但是这二者并不一定是直接的或直线水平关系。建议作者聚焦在社会正义宏观水平论证。

回应：

感谢审稿专家的建议。我们将组织正义这部分内容（及其文后的参考文献）做了删减，使 1.4.2 节中的内容都聚焦到社会正义宏观水平来论证。

4. 研究一结果部分“Z 分数计算的效应量与原始分数计算的效应量之间会因标准差的单位不同而出现系统性改变（效应值变大），因此，不能继续采用原有的 Cohen' d 值的经验标准来评估 Z 分数效应量的大小程度。”读者若是没有看到作者回应审稿专家的回复与分析，单从文章角度看到这个表述容易产生疑惑。建议作者换一种表述或者删除该句话。

回应：

感谢审稿专家的提示。我们删除了此句，在文中仅强调使用标准分数作为评估效应大小的指标。

5. 研究二的结果部分 3.2.3 大学生道德推脱水平与社会指标的关系中“各项指标的预测作用强于中学生组（研究 1）。”这两个群体数据回归系数应该是不能直接比较大得出该结论的，建议作者谨慎，或者进行统计上比较给出统计结果也可以的。

回应：

感谢提示。我们在文中的结果和讨论部分均删除了此类表述。

6. 讨论部分，“这两个年龄组被试的结果揭示了一个积极的变化：近十多年来大中学生群体的道德推脱水平降低了，这可能表明大中学生群体的道德水平有所上升，符合道德提升模型。”“这可能表明……”这个结论需要谨慎，虽然作者也表示是“可能”，但是毕竟道德推脱水平降低了，并不必然导致或反映道德水平上升了。

回应：

感谢审稿专家的严谨把关。我们已将此句做了删减处理。

7. 讨论部分可以进一步加强该研究的现实意义，尤其是从我国社会个体化水平和社会正义水平这两类社会宏观因素对道德推脱的预测作用来进一步阐述现实意义，更能突出本文的创新与贡献。

回应：

感谢审稿专家的建议。在讨论的 4.3 节中，我们之前从宏观层面讨论了本研究的政策意义。结合专家的建议，我们进一步从微观层面讨论研究的社会现实意义，具体而言，重点围绕立德树人的教育实践工作来讨论本研究的现实意义（请见第 35 页的红色字体部分）。

第三轮

审稿人 1 意见：

除以下两个问题外，我对作者的其他修改无异议，建议修改后发表。

回应：

感谢审稿专家对上一轮修改工作的肯定，下面我们对审稿专家指出的两个问题说明修改情况。

1、第 19 页：“……实证研究确实表明道德推脱各维度之间以及它们与量表总分之间具有高相关性(Caprara et al., 2009)，例如，这两种相关系数分别在 0.35 至 0.59 之间和 0.67 至 0.76 之间(滕召军, 2018)。道德推脱量表各维度之间较高的共变关系意味着如果能揭示量表总分的变迁趋势，则该趋势也很可能适用于具体维度，反之亦然。”首先，不知作者“高相关性”的标准是什么？一般认为 $r=.10$ 效应量小， $r=.30$ 效应量中等， $r=.50$ 效应量大。其次，.67-.76 的相关也不过是 40%-60% 的共变，分量表是可能存在特异性的。这里的推理还是不够严密。

回应：

感谢审稿专家的提醒，此句表述可能存在不当之处，对此，我们做了修改。具体修改内容，参见第 20 页的红色字体部分。

2、第 22 页对“社会个体化”的解释较晦涩难懂。例如：“个体的制度化抽离和重新嵌入”以及

“……造成系统风险的内在化”。这两句话都太过抽象，能不能更通俗些或举例说明？既然作者说社会个体化涉及“社会关系”以及“社会制度”在结构上的现代化转变，逻辑上是不是应先明确“社会关系”以及“社会制度”发生了哪些转化？目前 1.4.1 节的第一段还是没能从“结构”、“组织”或者“制度”等社会学层面把“社会个体化”的内涵讲清楚、讲透彻。

回应：

感谢审稿专家的建议，我们重新阅读学习贝克、阎云翔等有关风险社会、个体化和中国社会的个体化等理论著作后，系统梳理了个体化的理论内涵和我国社会个体化水平的变迁过程，并对 1.4.1 节进行整体的调整以提升内容的逻辑性和可读性。具体修改内容，参见 1.4.1 节红色字体部分。

第四轮

编委意见：

作者已按照外审专家意见作了修改，建议发表。

主编意见：

该文经过专家的仔细评审和作者们的认真修改，已经达到学报发表水平。建议做一些小的修改后发表。具体地，文章 2.2.3 和 3.2.3 都说和“社会指标的关系”。这里不妥，其实考察的是道德推脱和个体化和社会正义的关系。指标是用来标识个体化和社会正义的，有意义的是个体化和社会正义，而不是指标。建议文中标题改一下。相应下面的文字表述也改一下。

回应：

诚挚感谢各位审稿专家和编委专家为帮助我们提高论文质量贡献力量和智慧，也感谢专家们对我们论文修改工作的肯定。我们接受主编所提的修改意见，文章中 2.2.3 节和 3.2.3 节使用“社会指标”这一表述确实不妥。对此，我们将标题和文中内容统一做了修改。具体修改内容参见 2.2.3 节和 3.2.3 节中的红色字体部分。此外，我们也对文中存在类似的表述欠妥之处做了修改。敬请主编批评指正。