

《心理学报》审稿意见与作者回应

题目：天气和空气污染对诚信行为的影响：一项校园丢钱包的现场实验

作者：赵玉杰；高扬；周欣悦

第一轮

审稿人 1 意见：

作者以丢钱包为例，探讨了天气和空气污染对人们诚信行为的影响。选题新颖而有意义，方法得当，结论可信。以下为具体审稿意见。

意见 1：

作者在研究一中使用了已发表研究的数据，又补充了中国和美国的天气状况。这里有一个问题跟作者商榷：在已经发表的研究中，中国和美国的数据是什么时段收集的？而作者的天气状况又是什么时段收集的？假如两个时段不匹配，事实上的检验也就没意义了。因为我们都知道，天气本身是不稳定的，时刻在变化，因此作者有必要披露这方面的信息，回应读者在这方面的疑虑。

回应：感谢您的建议。

我们同意您的观点，天气状况的不稳定性可能会对我们的研究结果产生挑战。事实上，Cohn et al (2019)发表的全球诚信行为的研究并没有记录准确的时间点，只记录了实验当天的日期，美国的实验时间集中在 2015 年 8 月 17 日到 2015 年 10 月 17 日，中国实验数据主要集中在 2015 年 7 月 7 日到 2015 年 7 月 24 日，根据日期的记录，我们通过气象网站记录了当天的天气状况以及空气污染状况。所以我们的天气状况是跟丢钱包当天进行匹配的。

当然这种匹配只精确到天，并没有精确到具体的几点几分。由于现场实验的局限性、人力的局限性和技术的局限性，大多数关于天气状况的研究都没有精确到几点几分的天气记录，而是统一采用了当天的天气状况。例如下表列举的研究以及实验方法：

表 1 相关研究总结

引用	期刊	说明
Flynn & Greenberg, 2012	<i>Journal of Applied Social Psychology</i>	位置：实验 1 气象因素记录原因：验证天气因素对帮助行为的影响 方法： <u>回溯方式记录当天平均气温等气象因素</u>
Keller et al., 2005	<i>Psychological Science</i>	位置：实验 1 气象因素记录原因：验证天气状况对被试心情以及认知的影响 方法： <u>回溯方式记录当天平均气温等气象因素</u>
Fetterman, Wilkowski, & Robinson, 2018	<i>Social Psychological and Personality Science</i>	位置：实验 1 气象因素记录原因：排除天气因素的影响 方法： <u>回溯方式记录当天平均气温等气象因素</u>
Rotton & Cohn, 2000	<i>Journal of Personality and Social Psychology</i>	位置：研究 1 气象因素记录原因：验证天气因素对犯罪率的影响 方法： <u>在有准确时间的情况下，以三小时为单位回溯当天气温等气象因素</u>

另外，根据您的问题，我们还考虑到，虽然我们记录了丢钱包当天的天气状况以及空气污染，但是被试做决策的时间很可能不是实验当天进行的，在研究中 Cohn 等也记录丢钱包之后 100 天的邮箱信息作为诚信行为的指标。为了排除决策时间和实验当天天气以及空气污染不匹配的问题。我们进一步分析了大部分被试做出决策的时间：只关注那些主动归还钱包的被试的归还时间，数据显示在全球 40 多个国家当中，人们归还时间的中位数是 26 分钟，而且 88% 的被试是在一天之内邮件联系失主的，且各个国家在收到邮件的时间上并不存在显著的差异（Cohn et al., 2019）。从中国和美国的数据上看，被试主动联系失主的时间与全球基本保持一致的水平，超过 85% 的被试会在第一天就主动邮件联系失主。超过 90% 的参与者是在钱包丢失之后两之内邮件联系失主的，且两个国家在归还时间上不存在显著的差异。因此，我们有理由推测，实验当天就是被试主要的决策时间，大部分的决策一定是在三天之内进行的。为此我们不仅检验了当天的天气状况与空气污染对于诚信行为的影响，还补充了后面三天的天气状况以及空气污染，检验结果与假设相符。（具体改动见：正文 p6 第 9 行开始；p7 第 18 行开始）

意见 2:

我们留意到，在研究二中，大多数学生其实都做出了诚信行为。哪怕面临天气和空气状况不利的条件。作者研究二中少数人的不诚信行为可以用天气和空气状况来解释，但对于绝大多数（超过 90%）学生都没有做出不诚信行为该如何解释？另外，天气和空气状况的解释是否充分，也就是说，作者虽然通过实验（严格意义来说不是实验，因为研究中的天气状况是自然记录，而非认为操纵）得出了跟预期接近的结果，但如何证明这些结果就来自于天气和空气状况？作者在解释时自然可以推测，说结果证实了假设，但因为研究本身是调查而非实验，只能表明相关而无法证明因果。因此，作者有必要在假设中讨论这些限制。

回应:

感谢您的建议。在研究 2 中做出不诚信行为的学生占比较低，综合考虑，我们认为可能还存在以下的原因影响了学生的诚信行为：（1）周围其他人的影响。与生活中的其它情景不同，课堂上身边往往会有其他人的存在，他人很可能起到了监督的作用，限制了不诚信行为的产生（Doleac & Sanders, 2015; Zhong, Bohns, & Gino, 2010）；（2）我们选用的金额数量比较少（二十块人民币），无法形成有效的诱惑；（3）三所学校都是我国顶尖的大学，学生的道德水平可能相对较高，因此出现非诚信行为的可能性相对较低。关于这一点我们在研究 2 的分析部分进行了补充。（**具体请见：正文 P10, 12 行-15 行**）

感谢您提出解释天气和空气污染对诚信行为解释充分性的问题，我们完全赞同您提出的可能存在其他解释的可能性！为了排除其它因素的影响，在实验过程中我们还记录了包括课程、天气状况的其它变量、位置等信息。为了更好的对天气状况以及空气污染对于诚信行为的影响，在原有分析的基础上，我们进行了逻辑回归，并添加了包括学校、课程类型、校园卡放置的位置等因素作为控制变量，在检验天气状况和空气污染对诚信行为影响稳健性的同时，也对其它因素对诚信行为的影响进行了探索。（**相关改动及结果请见：正文 p12 第 5 行开始**）。检验结果如下：

表 2 天气晴朗状况、空气污染对诚信行为影响的稳健性分析

变量	b	SE	Wald	p	95% Exp(B)
截距	-0.553	1.721	0.103	0.748	
天气是否晴朗	1.625	0.697	5.436	0.02	[1.296, 19.915]
空气是否存在污染	-1.934	0.784	6.085	0.014	[0.031, 0.672]
云量	0.003	0.013	0.071	0.79	[0.979, 1.029]
湿度	-0.021	0.029	0.535	0.465	[0.924, 1.037]
最大风力	-0.017	0.222	0.006	0.938	[0.636, 1.518]
温度	-0.008	0.031	0.067	0.795	[0.934, 1.054]
课程类型	0.851	0.369	5.325	0.021	[1.137, 4.823]
相对位置	-1.009	0.405	6.198	0.013	[0.165, 0.807]
教室大小	0.295	0.124	5.611	0.018	[1.052, 1.713]
学校	0.022	0.59	0.001	0.97	[0.322, 3.246]

首先，**在控制了其它变量的情况下，天气是否晴朗以及控制污染对于人们诚信行为的影响仍然是符合假设的，在天气非晴朗的条件下，人们更有可能做出不主动归还卡或金钱的不诚信行为，空气污染的情况下人们也更有可能出现不诚信的行为；**值得注意的是确实存在其它变量会影响学生的诚信行为，例如：学生在教室的位置会影响学生的诚信行为，校园卡放在教室后排丢失的概率显著的高于放在前排丢失的概率（ $b=-1.01$, $SE=0.41$, $Wald=6.2$, $p=0.013$ ），针对这一点，我们认为可能是多种因素共同造成的，例如：作在后排的学生感受到的来自老师和周围其它同学的监督更弱，而坐在前排的同学周围的人远多于后排的同学，这就导致了后面的学生更可能做出非诚信的行为（Doleac & Sanders, 2015; Zhong, Bohns, & Gino, 2010）。当然，虽然这些因素也会对学生的诚信行为产生影响，但是这并无法排除天气和空气污染对于诚信行为的影响。

最后，再次感谢您的真知灼见与批判性意见！我们在您的指引下对整个论文的假设推理与理论贡献进行了重新创作，整个数据报告过程也更加的细致与完善。真诚感谢！也希望我们的回复能够满足您的要求！

审稿人 2 意见：

该研究通过 2 个平行研究检验了天气与空气污染对个体诚信行为的影响，选题较为新颖，但研究规范性和创新性较弱，仅在未控制诸多额外因素（如性别、一卡通证件照、年龄、人格特质等）干扰前提下探究了天气晴朗与否、空气污染水平与归还钱包/一卡通的相关关系，很遗憾并未看到作者揭示天气与空气污染对诚信行为的影响机制，这对诚信行为研究领域的贡献不大。因此，建议退稿。

意见 1：

问题提出部分，作者基于诚信缺失所带来危害而引出从环境因素出发探究诚信行为的发生原因，一方面较缺乏科学的文献研究基础，及并未看出作者选此主题进行研究的缘由是基于何种考虑？另一方面，作者在 1.1 部分回顾诚信行为的影响因素中，较多揭示诚信行为研究方法或范式，并没有为问题提出进行详实的文献研究。因此，本研究选题较缺乏扎实、充分的理论基础。

回应：

感谢您的建议。我们对您给出的意见进行了认真回应并进行了改进，希望可以满足您的要求。

关于一卡通证件照的问题，在实验中三所学校我们用的全部是一个人的证件照。完整版的实验材料我们也重新添加到附录当中。另外，三个学校的校园卡照片使用的是同一版本，为了适应三所学校校园卡的颜色，对照片的尺寸和背景色进行了调整，因此一卡通证件照是经过严格控制的（详细材料见 22 页到 23 页）。

关于您提出的性别、年龄、人格特质控制变量的问题，我们在您第二条意见中进行了解释说明，并对原文进行了修改。

关于天气与空气污染对诚信行为的影响机制，我们之所以没有考察空气污染对于诚信行为的影响机制，是因为：（1）关于天气和空气污染对于人们道德行为和亲社会行为的影响目前在国际上已经展开了部分研究，例如：Lu (2018) 通过犯罪的数据发现空气污染与犯罪呈现正相关关系，在实验研究中他们通过给被试看空气污染的图片，之后通过经济游戏的方式发现看到空气污染的图片会让被试感到焦虑，进而做出更多的不道德行为。因此在前人研究中，已经考察过了空气污染对不道德行为的影响机制。因此在本次研究的设计阶段，影响机制并非重点内容。我们主要希望通过在现实生活中，做现场实验的方式来检验天气和空气污染对人们诚信行为的影响。当然，我们同意您说的，本文的一个限制是没有讨论影响机

制，因此在文章的限制部分我们进行了讨论（见正文 **4.4 研究局限和未来研究方向第一段;P15,第 10 行开始**）。

根据您的建议，我们首先对问题提出部分进行了逻辑上的重新梳理和文章的撰写，在学报关于提出问题、以及讨论部分允许的字数范围内（3500 字）对理论进行了补充说明。改动集中在问题提出部分的第二段、第三段，以及 1.1 诚信行为的影响因素的第一段、第三段。

（**修改见正文 p1-p3**）

意见 2:

研究 1 采用二手数据，作者仅选择 Cohn 等（2019）文章中 40 多个国家的 2 个国家（中国和美国）被试数据，且 2 个国家国土面积较大、区域环境及空气质量存在较大差异，作者该如何规避这些额外因素的干扰？此外，非常值得关注的的一个问题是，诚信既可能是一种受状态因素影响的变量，也是一种较为稳定的品质或特质，因此它可能受个体层面因素影响较大，而作者在 2 个研究中均忽视对这些额外因素的干扰。

回应:

感谢您的建议。如您所言，在 Cohn 等（2019）Science 发表的文章中在 40 多个国家中进行了实验，我们选择了中国和美国两个国家，一共包括 1400 个样本。这主要是因为客观原因造成的，各个国家的空气质量标准并不一致，选择中国、美国恰恰时因为两个国家的空气质量标准很接近，在空气污染上具有可比性；另一方面，大部分国家的空气质量指数我们无法获得。

两个国家确实存在国土面积大，区域环境存在较大差异的问题。这些因素也是为什么我们选择这两个国家的一个理由。在每个国家实验的时间跨度很短，例如，在中国的 8 个城市实验的时间集中在 7 月 7 日到 7 月 24 日，如果不是因为中国国土面积大，实验城市分布广，气象数据很可能存在很强的同质性。这一点在空气污染上表现得更加明显，如果这个国家本身面积很小，例如瑞士，数据收集时间很短的情况下，不同城市之间得空气污染水平几乎没有差异的，无法让我们用来考察空气污染的差异导致的影响。

感谢您提出的关于个人特质以及其它控制变量方面的意见。这不但是本次研究的弱点，这一类的真实现场研究都普遍具有这个弱点。包括之前发表在 SCIENCE、PNAS 上的几篇文章（例如：Gneezy, Keenan, & Gneezy, 2014; Piff, Stancato, Côté Mendoza-Denton, & Keltner, 2012），几乎没有控制个人特质变量的。我们查阅了之前一些关于天气影响的真实场景研究，也并未考虑人格特质的影响，例如下面的这些研究：

表 3 相关研究总结

引用	期刊	介绍
Cunningham, 1979	<i>Journal of Personality and Social Psychology</i>	<p>研究内容：光照条件对于帮助行为的影响</p> <p>研究 1：记录被试愿意帮助研究人员回答问题的数量，记录光照条件；</p> <p>研究 2：记录餐厅服务人员收到的小费，记录光照条件；</p> <p>是否控制影响帮助行为的人格特质：否</p>
Rotton & Cohn, 2000	<i>Journal of Personality and Social Psychology</i>	<p>研究内容：温度对于不道德行为的影响</p> <p>研究：利用警局接到报警的数据，补充气象数据</p> <p>是否控制影响不道德行为的人格特质：否</p>
Cohn, Marechal, Tannenbaum, & Zund, 2019	<i>Science</i>	<p>研究内容：全球的诚信状况；金钱数量对于诚信行为的影响</p> <p>研究：将不同金额的钱包交给被试，记录被试是否主动归还</p> <p>是否控制影响诚信行为的人格特质：否</p>

意见 3：

研究 2 的样本来自中国三所名校（北大、浙大和清华），其研究结果能否代表更为广泛的大学生群体？此外，研究设计较为简单，数据分析方法较为单一。

回应：

感谢您的建议。我们同意您的观点，三所学校确实不能够代表我们国家的大学生群体，因此在讨论部分的实践价值中将本次研究对大学生诚信方面的意义进行了删除。这当然不仅是我们这个研究的问题，在几乎所有的心理学实验室研究中可能都普遍存在这个问题，研究局限于一个学校或者几个学校的大学生被试。

学生并非本次研究的目标对象，本次研究也并非定位在大学生诚信行为，而是立足于更加广泛的人群。这也就是为什么我们重新分析了 Cohn 等（2019）的研究数据，在他们的研究中，广泛的采取了更有真实性的社会样本。

最后，再次感谢您的洞见与细心！您的系列意见让我们对整个论文的理论基础与理论贡献进行了重新思考与写作，整个数据报告过程也更加的细致与完善。真诚感谢！也希望我们

的回复能够满足您的要求！

.....
编委意见：

作者们好，两位审稿人的意见一位比较负面，另一位比较正面。你们的文章我自己也读了，觉得研究内容比较贴合实际，有些意思，希望你们能在审稿人意见的基础上对文章进行修改后再投。我这里也有两个问题，和审稿人 1 的有些类似：

意见 1：

据我对 Science 那篇文章的理解，研究一中实验当天的天气并不一定是被试做决定那天的天气。如果是这样的话，会对你们的结果有什么影响？

回应：

感谢您的建议。为了排除决策时间和实验当天天气以及空气污染不匹配的问题。我们进一步分析了大部分被试做出决策的时间：如果我们只关注那些主动归还钱包的被试的归还时间，在全球 40 多个国家当中，人们归还时间的中位数是 26 分钟，而且 88% 的被试是在一天之内邮件联系失主的，且各个国家在受到邮件的时间上并不存在显著的差异（Cohn et al., 2019）。从中国和美国的数据上看，被试主动联系失主的时间与全球基本保持一致的水平，超过 85% 的被试会在第一天就主动邮件联系失主。超过 90% 的参与者是在三天之内选择邮件联系失主的，且两个国家在归还时间上不存在显著的差异。因此，我们有理有推测，实验当天就是被试主要的决策时间，大部分的决策一定是在三天之内进行的。为此我们不仅检验了当天的天气状况与空气污染对于诚信行为的影响，还补充了后面两天的天气状况以及空气污染。纳入共三天的天气状况后的分析结果仍然和之前相同。（**具体改动见：正文 p6 第 9 行开始；p7 第 18 行开始**）

意见 2：

研究二中丢失钱包的比例太少，这对统计检验会产生挑战。建议除了用平常的卡方检验，也用其他的方法，比如说逻辑回归。在逻辑回归中，还可以加入学校和课程等变量。

回应：

感谢您宝贵的建议，按照您的意见，我们进行了逻辑回归的分析，结果仍然跟之前一致。另外，我们还添加了包括学校、课程类型、校园卡放置的位置等因素作为控制变量，在检验天气状况和空气污染对诚信行为影响稳健性的同时，也对我们记录的其它因素对诚信行为的

影响进行了探索。在控制了其它变量的情况下，天气是否晴朗以及污染对于人们诚信行为的影响仍然是符合假设的。（相关改动及结果请见：正文 p12 第 5 行开始）

最后，再次感谢您宝贵的意见并给我们机会进一步完善文章的机会。希望我们的修改和回复能够满足您的要求！

参考文献：

- Cohn, A., Mar échal, M. A., & Zünd, C. (2019). Civic honesty across the globe. *Science* (pp. 24-26).
- Cunningham, M. R. (1979). Weather, mood, and helping behavior: Quasi experiments with the sunshine samaritan. *Journal of Personality and Social Psychology*, 37(11), 1947.
- Doleac, J. L., & Sanders, N. J. (2015). Under the cover of darkness: How ambient light influences criminal activity. *Review of Economics and Statistics*, 97(5), 1093-1103.
- Gneezy, U., Keenan, E. A., & Gneezy, A. (2014). Avoiding overhead aversion in charity. *Science*, 346(6209), 632-635. doi:10.1126/science.1253932
- Grant, A. M., & Mayer, D. M. (2009). Good soldiers and good actors: Prosocial and impression management motives as interactive predictors of affiliative citizenship behaviors. *Journal of Applied Psychology*, 94(4), 900-912. doi:10.1037/a0013770
- Kouchaki, M., & Desai, S. D. (2015). Anxious, threatened, and also unethical: How anxiety makes individuals feel threatened and commit unethical acts. *Journal of Applied Psychology*, 100, 360–375. doi:10.1037/a0037796
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.
- Lu, J. G., Lee, J. J., Gino, F., & Galinsky, A. D. (2018). Polluted morality: Air pollution predicts criminal activity and unethical behavior. *Psychological science*, 29(3), 340-355.
- Piff, P. K., Stancato, D. M., Côté S., Mendoza-Denton, R., & Keltner, D. (2012). Higher social class predicts increased unethical behavior. *Proceedings of the National Academy of Sciences*, 109(11), 4086-4091.
- Pun, V. C., Manjourides, J., & Suh, H. (2016). Association of ambient air pollution with depressive and anxiety symptoms in older adults: results from the NSHAP study. *Environmental health perspectives*, 125(3), 342-348.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163.
- Rotton, J., & Cohn, E. G. (2000). Violence is a curvilinear function of temperature in Dallas: a replication. *Journal*

of personality and social psychology, 78(6), 1074.

Zhong, C. B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological science*, 21(3), 311-314.

第二轮

审稿人 3 意见:

该研究采用了 2 个研究探究天气情况以及空气污染指数对大学生个体诚信行为的影响。在阅读了修改后的研究以及两位审稿人以及编辑的意见之后。我仍有以下问题需要作者回答。

意见 1:

两个研究都是对归还他人钱财行为的研究,但是两个研究所呈现的数据展现的诚信行为有很大差异。研究一在对 Cohn (2019) 的中美数据进行二次分析的时候,选取的是美国 25 个城市 2015 年 8 月-10 月两个月的天气数据;对比的是中国 8 个城市 2015 年 7 月一周内的天气数据。为什么不用同一时段的数据呢? 由于气候、温度、光照都有可能影响诚信行为,这些因素在做元分析的时候希望能够得到更好的平衡和控制。

回应:

感谢您的建议。我们赞同您的观点,两国的数据在同一时间段确实更加严谨。然而,本文之所以没有用同一时段的数据主要是因为 Cohn 的研究中,中国和美国的数据就是在这两个时间段进行采集的,本文并未对两个国家的数据进行选择,所有研究 1 的数据都是 Cohn 研究中原始的中美两国的数据。Cohn 的研究在世界范围内进行,在研究中并未保证每个国家在数据采集的时间段是一致的 (Cohn et al., 2019)。

另外,两个国家的数据本身并不存在对比的关系,之所以选取两个国家共 1400 个数据,是为了在气象、空气污染数据可取得性的条件下,尽可能保证最大的样本量来检验天气、空气污染对于诚信行为的影响。

最后您提到了气候、温度对于诚信行为也存在影响。在气候上面,两国之间确实存在很大的差异,这一点即使在相同的时间段我们也没有办法完全避免,我们也已经将这一点添加到我们研究 1 的不足当中 (详细请见正文 18 页 17 行)。

由于数据采集时间的不同,不同城市之间的温度确实存在很大的差异,为了排除温度对

于天气晴朗状况、空气污染对于诚信行为的影响，我们补充了实验当天的平均温度数据，并进行了逻辑回归分析，最终的结果表明，即使在控制了平均温度的情况下，天气是否晴朗以及空气污染仍然会影响人们的诚信行为，即：天气晴朗的情况下人们更多的做出诚信行为 ($b = 0.49$, $Wald = 9.51$, $p = 0.002$, 95% Exp(B) [1.20, 2.23])，而空气污染越严重，人们越有可能做出不归还钱包的不诚信行为 ($b = -0.02$, $Wald = 57.26$, $p < 0.001$, 95% Exp(B) [0.98, 0.99])。不仅如此，结合上一轮审稿人对被试决策时间的建议，我们对实验后面两天的数据也进行了逻辑回归分析，结果表明在控制温度的情况下，天气是否晴朗与空气污染对诚信行为的影响与前面的结果保持一致（详细修改请见正文第 7 页第 8 行-第 8 页第 9 行）。

表 1 天气晴朗状况、空气污染对诚信行为的影响

变量	实验当天			实验后第一天			实验后第二天		
	b	Wald	p	b	Wald	p	b	Wald	p
截距	1.90 (0.43)	19.17	<0.001	2.37 (0.41)	34.32	<0.001	1.89 (0.14)	37.65	<0.001
天气是否晴朗	0.49 (0.16)	9.51	0.002	0.45 (0.13)	11.53	0.001	0.87 (0.14)	37.65	<0.001
空气污染	-0.02 (0.002)	57.26	<0.001	-0.014 (0.002)	45.47	<0.001	-0.01 (0.002)	41.56	<0.001
当天平均气温	-0.07 (0.02)	16.78	<0.001	-0.10 (0.02)	31.98	<0.001	-0.08 (0.02)	24.22	<0.001

意见 2:

研究一，中国和美国的 1400 个现场实验数据中，归还钱包数量为 585，占比 41.8%，未进行归还的数量为 815，占比 58.2%。研究二，大学生的校园卡归还校园卡的比例都很高，大多数都是在 90% 以上。对于本土收集的数据，为什么会出现如此大的差异？作者从被试本身的角度做了一部分解释（例如，都是 985 高校，都在教室有监控，卡套内的钱只有 20 元等）。希望做成能从两个研究的研究方法上面也对如此大的差异做一个讨论。为什么 Cohen 的研究和作者的研究在中国人诚信行为上的数据有这么巨大的差别。

回应:

感谢您的建议。因为丢失的金额，丢失的位置，丢失的方式，归还的方式等在两个研究中都很不一样。不太容易进行直接比较。我们同意您的建议，在研究 2 中丢失的钱或者校

园卡与研究 1 中 Cohn 的研究确实存在很大的差异。实际上，我们认为无论是研究 2 的结果还是 Cohn 等人的研究结果都没有办法代表中国人实际的诚信水平，“丢钱包”的研究范式已经在众多经典的研究中得到应用，但是针对相关文献的梳理我们发现，针对同一国家的诚信行为在不同的研究中往往呈现的结果并不相同，例如：Stoop (2012)在荷兰进行的“丢钱包”实验，最终发现荷兰人平均的归还率为 50%，这一结果在 Cohn 等人的研究中上升为 75%左右；同样的，在 Cohn 的研究中美国人的邮件归还钱包的比率大概为 30%，但是在 Mullen 和 Nadler (2008)的研究中，美国学生平均的归还比例为 87.5%，不同群体之间得诚信水平确实存在比较大得差异。

根据您的建议，我们增加了讨论部分。在研究方法上，我们认为可能是因为下列因素的不同导致了两个研究的诚信行为结果出现较大差异（**详细修改请见正文：第 10 页 16 行-第 11 页 22 行**）：

（1）两个研究在“诚信行为”测量上并不完全相同。在 Cohn 的研究当中，他们将被试是否会主动邮件联系失主座位衡量诚信行为的主要标准，而且在将钱包交给被试之后，研究人员并未回到实验地点检查自己是否还能够要回“丢失”的钱包，只是将未收到邮件直接定义为“不诚信行为”；而在研究 2 的本土研究中，我们将金钱或者卡的直接丢失作为衡量诚信行为的标准，只有在研究人员确认放卡地点无法找回校园卡或者金钱的时候猜认定被试做出了“不诚信行为”。实际上，我们认为虽然 Cohn 的方法也可以在一定程度上反应诚信行为，但是用研究 2 的方法用来定义诚信行为更为合适。在 Cohn 的研究当中，即使被试没有通过邮件联系研究人员归还钱包，我们也不能完全肯定被试将钱包占为己有，而主动归还的钱包当中，由于研究人员并没有真的将钱包取回，研究人员同样无法确认是否存在金钱数额减少的不诚信行为。诚然，考虑到 Cohn 等人研究在世界范围内进行，数据量巨大，这些问题并不能很好的规避也在情理之中；但是研究 2 中，我们很好的解决了这样的问题。

（2）另一方面，在 Cohn 的研究中钱包的“丢失”概率明显较研究 2 多的原因也在于“邮箱”的使用，这一点在中国的数据当中表现的尤其明显。在研究中，Cohn 的数据发现在所有的实验国家当中，中国归还钱包的行为是最少的，仅为 14.2%，而根据实验流程，这是在研究人员主动将钱包交给被试并请求被试联系失主的情况下。我们有理由相信，如果钱包完全是被试自己捡到的，那么主动归还的概率会更低。但是这样低的概率多大程度上是由于我国邮箱的使用不普及导致的呢？在研究 2 中由于我们是预留了电话号码的，即使在完全没有主动要求被试的情况下，仍然有超过 24%的被试会主动打电话归还丢失的校园卡。因此，在 Cohn 的研究中，邮箱的使用可能也是导致中国不诚信行为增多的原因之一。

意见 3:

在研究二中，作者为了在视觉上夸大天气对诚信行为的影响，把图 2 和图 3 的 Y 轴起点标为 88% 和 82%。这种做法太随意，希望作者统一改为 0%。

回应:

感谢您宝贵的建议，按照您的意见，我们将图 2 和图 3 的 Y 轴起点统一改为了 0%。(详细修改请见正文：第 12 页图 2、第 13 页图 3)。

意见 4:

作者在文献综述中强调过光照对诚信行为的影响，并且作者在研究假设中着重检验了是否晴天对诚信行为的影响。作者还发现研究二的数据中，上经济类课程的同学的归还率比思政课以及历史课要多。原因之一是选中的课程大多是公共选修课。作为高校老师，我们知道公共选修课经常被统一安排晚上进行，所以白天上课与晚上上课教室内的光照程度是很不一样的。作者在审稿人 2 以及编辑的意见下，做了逻辑回归。回归中也添加了不少需要控制的因素，建议作者从原始数据中尽量恢复一下实验数据采集时的光照情况，并统计一下光照情况对诚信行为的影响。环境会对诚信行为产生影响，在研究二中，室内的光照环境相对于室外的天气以及污染程度是不可忽视的影响因素。

回应:

感谢您宝贵的建议，关于室内的光照环境。在数据采集的过程中，我们虽然并未记录实验教室的光照环境，但是在记录中，我们记录了课程是在上午、下午、还是晚上进行的，在您的建议下，我们重新分析了我们的数据，将课程是在哪一个时间段进行的进行了控制（编码：上午：0；下午：1；晚上：2）。

首先，关于课程与上课时间的关系，结果正如您的建议，数据证明，由于经济学课存在较多的公共选修课，经济学课与其它课程相比，确实存在很多在晚上上课的情况。

表 2 上课时间段与课程类型的关系

课程类型	上课时间		
	上午	下午	晚上
思政课	50 (45.9%)	76 (45.5%)	12 (9.2%)
历史课	37 (33.9%)	55 (32.9%)	48 (36.6%)
经济学课	22 (20.2%)	36 (21.6%)	71 (54.2%)

但是数据采集的时间对于诚信行为并不存在显著的影响 ($\chi^2(2, N=407)=0.491, p = 0.782$)。我们认为, 虽然之前的研究表明, 在黑暗的条件下人们往往会做出更多的不诚信、不道德的行为 (Hartley, 1974; Page & Moss, 1976), 但是能见度下降才是背后的根本原因, 而在研究 2 中, 实验的场景主要设置在教室, 即使是在晚上, 课堂环境下往往具备很好的照明条件, 这也是我们并未发现在不同的时间段会影响人们诚信行为的原因。

表 3 时间段与诚信行为的关系

诚信行为	上课时间		
	上午	下午	晚上
诚信	102 (93.6%)	159 (95.2%)	125 (95.4%)
非诚信	7 (6.4%)	8 (4.8%)	6 (4.6%)

为了排除室内的光照环境的影响, 我们在研究 2 的回归分析中加入了课程的时间段的变量。最终的结果表明, 在控制时间段的情况下, 天气的晴朗状况以及空气污染对于诚信行为的影响仍然是存在的, 天气晴朗的情况下, 人们更少的做出拿走卡或者钱的不诚信行为 ($b= 1.633, Wald = 5.491, p = 0.019, 95\% \text{Exp(B)} [1.306, 20.078]$), 而在空气污染的情况下, 人们的不诚信行为则会表现除更多的不诚信行为 ($b= -1.911, Wald = 5.865, p = 0.015, 95\% \text{Exp(B)} [0.032, 0.695]$)。

表 4 天气晴朗状况、空气污染对诚信行为影响的稳健性分析

变量	b	SE	Wald	p	95% Exp(B)
截距	-0.617	1.73	0.127	0.721	
天气是否晴朗	1.633	0.697	5.491	0.019	[1.306, 20.078]
空气是否存在污染	-1.911	0.789	5.865	0.015	[0.032, 0.695]
云量	0.002	0.013	0.035	0.851	[0.977, 1.028]
湿度	-0.019	0.03	0.415	0.519	[0.926, 1.04]
最大风力	0.004	0.224	0.000	0.987	[0.647, 1.558]
体感温度	-0.014	0.032	0.185	0.667	[0.926, 1.051]
课程类型	0.782	0.386	4.101	0.043	[1.025, 4.656]
相对位置	-1.013	0.407	6.196	0.013	[0.163, 0.806]
教室大小	0.299	0.125	5.739	0.017	[1.056, 1.721]
学校	-0.037	0.6	0.004	0.95	[0.297, 3.12]
时间段 (上午、下午、晚上)	0.215	0.351	0.376	0.54	[0.624, 2.466]

不仅如此, 为了进一步确认晴朗状况以及空气污染的影响不会受到时间段的影响, 我们还用另外两种方式进行了进一步的数据分析:

(1) 首先, 我们在逻辑回归中验证了时间段与空气污染以及时间段与天气晴朗状况是否存在对于诚信行为的交互影响, 最终的结果表明, 上课时间段与空气污染和晴朗状况并

不存在显著的交互作用 ($p_{\text{晴朗} \times \text{时间段}}=0.695$; $p_{\text{空气污染} \times \text{时间段}}=0.229$):

(2) 在数据分析中, **如果我们只保留白天采集的数据(上午、下午), 检验天气状况和空气污染对诚信行为的影响。最终的结果与前面的数据分析结果基本一致** ($b_{\text{晴朗}}=1.963$, $Wald=4.521$, $p=0.033$, 95% Exp(B) [1.166, 43.522]; $b_{\text{空气污染}}=-1.754$, $Wald=3.231$, $p=0.072$, 95% Exp(B) [0.026, 1.172])。

表 5 研究 2 天气晴朗状况、空气污染对诚信行为的影响 (只保留白天收集数据)

变量	b	SE	Wald	p	95% Exp(B)
截距	-0.865	2.322	0.139	0.421	
天气是否晴朗	1.963	0.923	4.521	0.033	[1.166, 43.522]
空气是否存在污染	-1.754	0.976	3.231	0.072	[0.026, 1.172]
平均云量	-0.007	0.017	0.153	0.695	[0.977, 1.028]
平均湿度	-0.001	0.036	0.001	0.982	[0.931, 1.073]
最大风力	0.038	0.284	0.017	0.895	[0.595, 1.813]
体感温度	-0.056	0.038	2.169	0.141	[0.877, 1.019]
课程类型	0.351	0.498	0.499	0.48	[0.536, 3.769]
相对位置	-0.973	0.469	4.306	0.038	[0.151, 0.947]
教室大小	0.434	0.174	6.205	0.013	[1.097, 2.171]
学校	-0.347	0.848	0.168	0.682	[0.134, 3.721]

从上面的结果我们发现, 如果只看白天收集的数据, 我们无法得到“经济学课上的诚信行为更多”的结论, 当然这可能是因为此时经济学课程的数据量很少 ($N=58$), 但是考虑到折本身并不是本次研究重点讨论的问题, 我们将原文中针对课程对诚信行为影响的分析进行了删除, 更加突出天气晴朗状况、空气污染对于诚信行为的影响。(详细改动请见正文: 13 页第 8 行开始)

最后, 再次感谢您宝贵的意见并给我们机会进一步完善文章的机会。希望我们的修改和回复能够满足您的要求!

参考文献:

- Cohn, A., Maréchal, M. A., & Zünd, C. (2019). Civic honesty across the globe. *Science* (pp. 24-26).
- Hartley, J. E. (1974). Lighting reinforces crime fight. *Pittsfield, MA: Buttenheim*.
- Mullen, E., & Nadler, J. (2008). Moral spillovers: The effect of moral violations on deviant behavior. *Journal of Experimental Social Psychology*, 44, 1239–1245.
- Page, R. A., & Moss, M. K. (1976). Environmental influences on aggression: The effects of darkness and proximity of victim. *Journal of Applied Social Psychology*, 6(2), 126-133.
- Stoop, J. (2012). From the lab to the field: envelopes, dictators and manners. Munich Personal RePEc Archive, MPRA Paper No. 370
-

第三轮

审稿人 3 意见:

作者根据两位审稿人以及编委的意见已经做了详尽的回复。对提出的相关质疑做了非常认真的回应。虽然本研究的原创性和理论性都不是很高,并且在数据处理方面面临诸多挑战。但是作为结合 Cohen 文章后续的一项研究,还是有价值和贡献的。通过讨论与修改,可以看出对现有实验设计以及数据的处理是得当的。建议发表,也希望能够有更多的实验能够填补这两篇文章所发现的不一致留下的空白。

回应:

感谢您的肯定以及宝贵意见

编委意见:

现已收到新的审稿人(审稿人 3)对你们文章的意见和你们的回复。我自己仔细看过了新版本的文章,你们没有对理论和文献部分做什么修改,主要是加了一些新的数据分析,其结果大多支持之前的假设。下面是我认为文章现在还存在的问题,有些可能是无法用新的分析方法改变的:

意见 1:

效应量(effect size)普遍较小。“天气是否晴朗”在你们研究一中的效应量是 $\phi=0.053$, 研究二中是 $\phi=0.10$;“空气污染”是二元变量时,在研究一中不显著,也没有汇报效应量,

在研究二中是 $\phi=0.11$ 。这些数值都很小。这两个因素之所以还能在回归分析中出现统计显著，很可能是因为你们样本量比较大。也就是说，今后的研究如果样本不够大的话，很可能观察不到这两个因素的作用... 说这些，想让你们 a) 在文章讨论中指出这点；b) 看一下之前研究的效应量，和你们的相比是更大还是更小（估计不太可能），差别的原因是什么。心理学的重复危机想必你们也知道，假阳性的结果是大家都不愿意见到的。

回应：

感谢您的宝贵建议。我们同意您的观点，在本次研究中无论是天气的晴朗状况还是空气污染，效应量都是比较小的，也正如您所说，天气以及空气污染对于诚信行为的影响确实可能在大样本研究的基础上才能实现。我们认真分析了背后可能存在的原因。

首先，在您的建议下，我们对比了前人关于天气以及空气污染影响的相关研究，并重点关注了相关研究的效应量。为了便于比较，我们将卡方检验的效应量转化为相关系数 r ，研究 2 中空气污染的效应量 $r=0.1078$ ，天气晴朗状况与诚信行为影响的效应量在两个研究中分别是： $r=0.053$ ； $r=0.1$ 。通过对比发现，很多相关研究的效应量都较小：例如：在空气污染对道德行为的影响研究中，Lu 等（2018）通过研究发现空气污染条件下人们更容易做出不道德甚至是违法行为。在研究 1 中，他们通过二手数据分析发现了空气污染和违法行为的正向关系（ $N=9360$ ），但是实际上从相关系数看效应量可以发现，空气污染对于各种违法行为影响的效应量并不大。其中，污染综合指数与各种犯罪行为（包括抢劫、盗窃等）的相关性范围是： $r=0.07-0.1$ （ $p<0.05$ ）； PM_{10} 与犯罪行为的相关性范围为： $r=0.04-0.08$ （ $p<0.05$ ）；不仅如此，在研究 2 中，作者通过给被试看图片的方式操纵了被试空气污染的感知（空气污染图片 vs. 普通图片， $N=256$ ），并验证了空气污染与不道德行为之间的因果关系，这项研究的效应量是 $Cohen'd=0.23, 95\%CI[0.015, 0.476]$ ，效应量在 0.2-0.5 之间，说明效应量较小（Cohn, 1988），为了便于比较，我们将 $Cohen'd$ 转化为相关系数， $r=0.114$ 。

同样，天气相关的研究中也存在较小的效应量。例如：Keller 等（2005）研究了天气状况和人们心情之间的关系，他们发现天气状况越好人们的心情越好。但是在两项相关研究中，无论是温度还是晴朗状况与心情的相关性都在 $r=0.2$ 左右，效应量并不大。因此，本次研究效应量较小并非特例，相关研究也存在效应量较小的情况。

关于效应量较小的原因，我们认为可能是下列因素造成的：

1. 对于行为的影响因素很多，天气只是其中的一个因素。我们的准实验研究在保证情景尽可能自然真实的条件下也包含了很多混淆因素在里面，复杂的环境因素很可能导致了

效应量的下降。

2. 两个研究的场景都在室内也可能是造成效应量较低的原因。研究 1 和研究 2 的实验场景一个在各种机构的前台，一个在教室内，室内的场景可能削弱了天气以及空气污染对于人们诚信行为的影响。这一点在之前的研究中也有表现，在研究天气状况与帮助行为的关系时，Keller 等（2005）在研究 1 中采用户外实验的形式，而研究二则在一家餐厅内部进行，结果发现晴朗状况与帮助行的相关性从 $r=0.36$ 下降到 $r=0.2$ ，研究人员在文中也谈到，效应量的下降很可能是受到室内的环境影响（“This correlation was smaller than in Experiment 1, which might be expected, since participants were indoors and were thus partially screened from effects of the sun”）。

综上，我们对前人研究和我们研究的效应量进行了对比，同时还分析了本次研究效应量较小的可能原因。当然，我们认为较小的效应量并不足以影响本次研究的意义。一方面，不少学者认为效应量的大小不应该局限于相关的统计分析，也应该关注研究的“长远影响”（Funder & Ozer, 2019），虽然本次研究发现的天气状况以及空气污染对诚信行为的影响效应量较小，但是考虑到诚信行为作为影响社会经济发展、社会和谐的重要因素，从长远来看，较小的效应量也会产生非常大的影响；另一方面，心理学实验的可重复性危机确实引起了广泛的关注，但是部分学者认为大的效应量不应该是解决这一问题的唯一方式，他们认为：大样本小效应量的研究在重复实验中反而是更加值得信任的（Funder & Ozer, 2019）（“In our view, enough experience has already accumulated to make one suspect that small effect sizes from large-N studies are the most likely to reflect the true state of nature”）。

相关内容我们已经在讨论部分中进行了添加，详细请见讨论部分 P19 第 2 行-第 18 行

意见 2：

对比表一，二和三，一个让人比较困惑的结果是：随着天数的靠后，“天气是否晴朗”对诚信行为的影响好像更加重了。为什么？如果“绝大多数参与者进行决策的时间是在捡到钱包的当天”，这个结果是不通的。另外，“当天平均气温”在这三个表中也都是显著的。我欣赏你们如实的汇报，但汇报后的解释呢？为什么会出现这个结果？它和你们的假设有什么关系？

回应：

您的这个建议非常中肯且有建设性，对我们的帮助很大。从目前的数据来看确实存在后面两天的天气状况显著性甚至高于实验当天的情况，这样的结果确实值得深入分析。后面两

天的显著性“看似”高于第一天，我们进一步分析了这个结果：

1. 在短时间内，以天为单位的天气变量存在一定的连续性，例如，我们进行了第一天天气晴朗状况与第二天天气晴朗状况的逻辑回归分析，结果发现，从结果上看，第一天和第二天的天气晴朗状况存在显著的正向相关关系 ($b = 3.34$, $Wald = 283.66$, $p < 0.001$, $95\% \text{ Exp}(B) [19.13, 41.61]$)，同理，从结果上看第二天的天气晴朗状况对第三天的天气也存在显著的影响 ($b = 2.01$, $Wald = 195.95$, $p < 0.001$, $95\% \text{ Exp}(B) [5.639, 9.905]$)。因此从数据上看，三天的天气晴朗状况本身就存在明显的相关性，这样的相关性可能是导致后面两天的天气晴朗状况与诚信行为也是显著相关的原因。这一点在空气污染的数据当中也存在类似的情况。

2. 另外，虽然第三天天气影响“看起来”显著性比较高，但是是不是意味着第三天的天气的影响比第一天天气的影响要显著的大呢？我们用第一天的天气和第三天的天气作为重复测量的自变量，发现交互效应并不显著 ($p=0.356$)。这意味着第一天的天气影响和第三天的天气影响没有显著的差异，虽然有着肉眼上的差异。

在回应第一轮审稿意见的时候，我们在说明绝大多数参与者都是在实验当天 (>85%) 进行决策的同时还补充了后面两天的数据分析，实际上，我们认为无论后面两天的天气状况与空气污染对于诚信行为是否存在影响都无法排除我们的假设，即：天气晴朗、空气污染较轻的情况下，人们更容易做出诚信行为。诚然，研究 1 确实无法彻底解决“是否参与者全部是在实验当天进行决策的”这样的问题，因为这样的原因，我们进行了研究 2 作为重要的补充和支持。（详细修改请见：P9 第 1 行-14 行）

另外，根据您的建议，我们也增加了关于温度与诚信行为的影响的讨论。之前的部分研究表明，温度对于人的行为存在明显的影响。例如，在温度适宜的情况下，人们会做出更多的助人行为 (Fetterman, Wilkowski, & Robinson, 2018)。之后的相关研究又发现，温度对于行为的影响并非线性的，舒适的温度才会对人们的行为存在积极的影响（例如：助人等），而过高或者过低的温度往往会产生明显的负面作用 (Anderson, 2001; McCrae, Terracciano, Realo, & Allik, 2007)，例如：炎热的天气下暴力行为会明显增多。结合研究 1 中国、美国两国的实验数据收集时间，中国数据收集时间在 7 月，而美国的数据收集在 8-10 月，数据收集时间是两国平均温度较高的月份。我们认为，在炎热的夏天，平均温度较高可能导致被试感觉不舒服，从而对诚信行为存在显著的负面影响。

根据前人的研究，天气状况之所以会影响人们的行为主要是通过情绪，在本次研究中我们也倾向于认为无论是晴朗状况还是温度对诚信的影响很大程度上是因为情绪。但是相比于天气的晴朗状况，温度的变化更加明显，温度究竟对人们的行为存在怎样的影响在前人的研

究中也并未得到一个统一的结论。并且，我们也并未在研究 2 中发现温度（体感温度）对于诚信行为存在显著的影响。关于这一点，根据您的建议，我们在研究 1 和研究 2 的数据分析中讨论了温度对于诚信行为的影响（详细修改请见 P7 第 14 行到 P8 结束；P14 第 7 行-第 9 行；P15 页 15 行-P16 页第 6 行）。

意见 3:

研究二也有类似的问题。“相对位置”和“教室大小”在统计上是显著的，其效应和那两个主要分析的因素差不多。这些不在假设范围内的显著效应应该仔细考虑其原因，但也不要过分地给 post-hoc 的解释，尺度不太好掌握。

回应:

感谢您的意见。根据您的意见，我们减少了这部分的解释和讨论。在进行研究 2 的数据收集时，为了排除其它因素对于诚信行为的影响，我们对一些可能影响诚信行为的变量进行了记录。因此，我们在收数据之前就已经判断相对位置和教室大小会对诚信行为存在影响。最终的结果也支持了我们的假设。最终的结果表明，相对位置和教室大小对于学生的诚信行为存在显著的影响，具体而言：放置卡的位置越靠教室后排，学生越有可能做出拿走卡或钱的非诚信行为；相反教室越大，学生做出不诚信行为的可能性越低。

我们认为两个变量之所以会对诚信行为存在显著的影响是因为老师和其他学生的监督，具体而言：

1. 关于前后位置对诚信行为的影响：坐在前排的学生一方面距离老师更近，另一方面，由于教室的座位排布，前排的学生往往也会更多的受到来自身后同学的监督；相比之下，坐在后排的学生感受到的来自老师和周围其它同学的监督更弱。事实上，当人们感受到外界的监督时，就会做出更少的非诚信行为（Doleac & Sanders, 2015; Zhong, Bohns, & Gino, 2010）。

2. 关于教室大小对于诚信行为的影响：在实验收集过程中我们发现，在较大的教室内学生数量更多。这也就意味着，在较大的教室内，学生感受到的来自身边学生的监督也更加明显。因此，在较大的教室内，学生的诚信行为反而较多。

根据您的建议，我们也在研究 2 的讨论中进行了修改，具体修改请见 P16 第 7 行-19 行。

意见 4:

你们的研究整体还是比较有意思和意义的，这也是审稿人 3 同意发表的主要原因。但我

希望你们能在我这次评论的基础上，再仔细地改文章，中间不要以应付审稿人和编辑的心态来做，而是想怎么把文章写好，让读者们信服。最后，那些表和图不能一味地往文章里放，想想更好的呈现和表达的方式。

回应：

感谢您的肯定和建议。根据您的建议我们已经对文章进行了对应的修改，尤其是在数据的结果分析和最终的讨论部分，使得文章的结果更加清晰，数据分析和讨论更具有说服力。另外，在您的建议下，我们对于文中的图标进行了重新的整理和删减。**删除了可以通过文字表达清楚的无关紧要的表格，并将部分表格进行合并**。除此之外，在您的建议下，我们重新梳理了整篇文章的逻辑，尤其是数据分析和讨论部分，并进行了详细的修改，使得文章的逻辑更加清晰，分析更加准确细致。（**详细请见整篇文章黄色标注部分**）

最后，再次感谢您宝贵的意见并给我们机会进一步完善文章的机会。根据您的意见我们对文章的结构以及分析进行了调整和深化，希望我们的修改和回复能够满足您的要求！

参考文献：

- Anderson, C. A. (2001). Heat and violence. *Current directions in psychological science*, 10(1), 33-38.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Doleac, J. L., & Sanders, N. J. (2015). Under the cover of darkness: How ambient light influences criminal activity. *Review of Economics and Statistics*, 97(5), 1093-1103.
- Fetterman, A. K., Wilkowski, B. M., & Robinson, M. D. (2018). On feeling warm and being warm: Daily perceptions of physical warmth fluctuate with interpersonal warmth. *Social Psychological and Personality Science*, 9(5), 560-567.
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156-168.
- Keller, M. C., Fredrickson, B. L., Ybarra, O., Côté S., Johnson, K., Mikels, J., . . . Wager, T. (2005). A warm heart and a clear head: The contingent effects of weather on mood and cognition. *Psychological Science*, 16(9), 724-731.
- Lu, J. G., Lee, J. J., Gino, F., & Galinsky, A. D. (2018). Polluted morality: Air pollution predicts criminal activity and unethical behavior. *Psychological Science*, 29(3), 340-355.
- McCrae, R. R., Terracciano, A., Realo, A., & Allik, J. (2007). Climatic warmth and national wealth: some

culture-level determinants of national character stereotypes. *European Journal of Personality: Published for the European Association of Personality Psychology*, 21(8), 953-976.

Meyvis, T., & Van Osselaer, Stijn M J. (2018). Increasing the power of your study by increasing the effect size. *Journal of Consumer Research*, 44(5), 1157-1173.

Zhong, C. B., Bohns, V. K., & Gino, F. (2010). Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological science*, 21(3), 311-314.

第四轮

编委意见:

感谢作者们对我上次意见的认真回复。我总体比较满意。鉴于文章所要阐述的观点其实没有很多，也没有很复杂。建议对文章整体进行删减，有些图表能放到 Appendix 中就不用放在正文里，把之前为了应付审稿人意见的一些过分细致的内容适当进行精简，reference 也是。

回应:

感谢您的肯定和宝贵建议。根据您的建议，我们对全文进行了删减，主要包括：理论部分我们进行了精简，删除了与文本研究不太相关的文献总结；数据部分我们将过分细致的分析进行了适度的删减；对于相关图表，我们将两个图标放在了附录。

再次感谢您宝贵的建议。

第五轮

主编意见:

一个会吸引大众兴趣的研究，目前字数约为 1 万 5 千字，在不影响文章旨趣的情况下，请尽量精简。

回应:

感谢您的肯定和宝贵建议。根据您的建议，我们对全文进行了进一步的删减。

再次感谢您宝贵的建议。