

《心理学报》审稿意见与作者回应

题目：双相障碍患者的风险决策偏好：来自三水平元分析的证据

作者：陆嘉琦, 李雨斯, 何贵兵

第一轮

非常感谢审稿专家提出的宝贵意见和建议。在认真学习审稿意见后，我们按照专家的审稿意见对原论文做了较大幅度的修改，以期进一步提升稿件质量。修改工作主要包括两方面：一是根据专家意见增加了日常风险行为相关的搜索关键词（如物质使用、性风险行为等，详见正文附表1），并将检索文献的时间范围扩大至2024年4月15日。相较此前的检索结果，本次增加了7212篇文献，最终纳入了73篇文献，比上次增加了15篇。在此基础上，我们对纳入文献进行了重新分析，发现主要研究结果仍与之前一致，即双相患者比健康个体更偏好风险寻求（详见正文的结果部分）。二是针对专家提出的其他意见对论文做了相应修改。

下面是我们根据审稿专家提出的问题和建议进行的逐一回复说明，文中修改部分都标记为红色。

审稿人1意见：该论文通过元分析方法系统探讨了双相障碍患者与正常健康群体的风险决策偏好差异，并检验了测量特征、心境阶段和样本特征等因素在其中的调节作用。该元分析在前人研究的基础上纳入了风险偏好的不同测量方式，元分析的步骤和流程总体上较为合理，论文流畅性较高，但是论文还存在很多问题，这些问题的存在大大降低了论文的整体质量和研究结论的可信度。以下是我的一些意见，供作者参考。

意见1：作者在引言中写道“研究者发现风险寻求增加是双相患者常见的行为特征，并认为该风险决策特征或许能成为双相障碍潜在的内表型（即核心症状）”，这一论述是前人公认的吗？有何支持性证据？

回应：感谢您的问题。这一论述主要参考了过往研究中的表述。例如，Hidroğlu 等人(2013)

写到"Risk-taking behavior and impulsivity are core features of bipolar disorder."和"Increased risk-taking behavior is one of the more recognized symptoms of bipolar disorder."; Ram íez-Mart ín 等人(2020)写到"potentially dangerous risky behaviours appear to be frequent components in the course of BD in its various phases and have been proposed as promising endophenotypes (core features) of the disease."。

双相患者风险寻求增加的特征有不同方面的支持性证据。首先，在DSM-V中躁狂或轻躁狂的诊断标准中有提及“过度地参与那些很可能产生痛苦后果的高风险活动”，和双相患者常与破坏性、冲动控制与品行障碍、物质使用障碍等与风险行为有关的精神障碍共病(APA, 2013)。其次，临床观察发现，双相患者确实容易做出一些风险行为，如风险性行为(e.g., Krantz et al., 2018)和物质滥用(e.g., Williams et al., 2017)等。其中，双相患者还有极高的自杀倾向，其自杀率是所有精神障碍中最高的(Plans et al., 2019)。最后，过往双相障碍与风险决策的行为实证研究与神经影像学研究发现了两者的关系(e.g., Edge et al., 2013; Lapomarda et al., 2021)。这些证据在之前的正文中虽有提及（见引言，正文 P2），但确实不够完善。

为了完善这一论述的提出和相关证据的阐述，我们对这一部分的正文进行了修改，具体见引言部分(P2)。为了便于审稿，我们将相关文字粘贴至下方。

——引言，正文第 2 页——

近年来，有研究者根据躁狂或轻躁狂诊断中“过度地参与那些很可能产生痛苦后果的高风险活动”的标准，结合双相障碍常与风险行为相关的精神障碍共病（如，物质使用障碍和破坏性、冲动控制与品行障碍）(APA, 2013)，推测双相患者比健康个体更偏好风险寻求(e.g., Ernst et al., 2004; Holmes et al., 2009)。实际观察也发现，患者在生活中常做出风险行为，如物质滥用(e.g., Williams et al., 2017)和风险性行为(e.g., Krantz et al., 2018)。因此，越来越多研究者认为，风险寻求增加是双相患者常见的行为特征，并可能成为双相障碍潜在的核心症状之一(Hıdıroğlu et al., 2013; Ramírez-Mart ín et al., 2020)。

参考文献

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>.
- Edge, M. D., Johnson, S. L., Ng, T., & Carver, C. S. (2013). Iowa gambling task performance in euthymic bipolar I disorder: A meta-analysis and empirical study. *Journal of Affective Disorders*, 150(1), 115–122. <https://doi.org/10.1016/j.jad.2012.11.027>
- Ernst, M., Dickstein, D. P., Munson, S., Eshel, N., Pradella, A., Jazbec, S., Pine, D. S., & Leibenluft, E. (2004).

- Reward-related processes in pediatric bipolar disorder: a pilot study. *Journal of Affective Disorders*, 82, S89–S101. <https://doi.org/10.1016/j.jad.2004.05.022>
- Hıdıroğlu, C., Esen, Ö. D., Tunca, Z., Yalçın, Ş. N. G., Lombardo, L., Glahn, D. C., & Özerdem, A. (2013). Can risk-taking be an endophenotype for bipolar disorder? A study on patients with bipolar disorder type I and their first-degree relatives. *Journal of the International Neuropsychological Society*, 19(4), 474–482. <https://doi.org/10.1017/S1355617713000015>.
- Holmes, M. K., Bearden, C. E., Barguil, M., Fonseca, M., Monkul, E. S., Nery, F. G., ... Glahn, D. C. (2009). Conceptualizing impulsivity and risk taking in bipolar disorder: Importance of history of alcohol abuse. *Bipolar Disorders*, 11(1), 33–40. <https://doi.org/10.1111/j.1399-5618.2008.00657.x>
- Krantz, M., Goldstein, T., Rooks, B., Merranko, J., Liao, F., Gill, M. K., Diler, R., Hafeman, D., Ryan, N., Goldstein, B., Yen, S., Hower, H., Hunt, J., Keller, M., Strober, M., Axelson, D., & Birmaher, B. (2018). Sexual Risk Behavior Among Youth With Bipolar Disorder: Identifying Demographic and Clinical Risk Factors. *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(2), 118–124. <https://doi.org/10.1016/j.jaac.2017.11.015>
- Lapomarda, G., Pappaiani, E., Siugzdaite, R., Sanfey, A. G., Rumiat, R. I., & Grecucci, A. (2021). Out of control: An altered parieto-occipital-cerebellar network for impulsivity in bipolar disorder. *Behavioural Brain Research*, 406, 113228. <https://doi.org/10.1016/j.bbr.2021.113228>
- Plans, L., Barrot, C., Nieto, E., Rios, J., Schulze, T. G., Papiol, S., Mitjans, M., Vieta, E., & Benabarre, A. (2019). Association between completed suicide and bipolar disorder: A systematic review of the literature. *Journal of Affective Disorders*, 242, 111–122. <https://doi.org/10.1016/j.jad.2018.08.054>
- Ramírez-Martín, A., Ramos-Martín, J., Mayoral-Cleries, F., Moreno-Küster, B., & Guzmán-Parra, J. (2020). Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. *Psychological Medicine*, 50(13), 2141–2153. <https://doi.org/10.1017/S0033291720003086>
- Williams, S. C., Davey-Rothwell, M. A., Tobin, K. E., & Latkin, C. (2017). People Who Inject Drugs and Have Mood Disorders—A Brief Assessment of Health Risk Behaviors. *Substance Use and Misuse*, 52(9), 1181–1190. <https://doi.org/10.1080/10826084.2017.1302954>

意见 2：作者指出前人研究结果不一致可能是由于心境阶段、风险偏好的测量方式和样本的人口学特征差异等调节变量的存在，但这三个变量的具体类别和影响没有厘清：双相患者的心境阶段只包括<轻>躁狂期、重性抑郁期和缓解期吗？是否存在其他类别的心境阶段？本研究中为何只关注这三种？这些问题需文中说明；

回应：感谢您的问题和建议。根据 DSM-V 诊断标准，双相患者有时表现为心境高涨、精力旺盛或活动增多，即<轻>躁狂期，包括躁狂发作和轻躁狂发作；有时表现为心境低落、疲惫乏力或对活动丧失愉悦感，即重性抑郁期，对应重性抑郁发作；而发作间隙症状通常得到缓解，情绪趋于稳定，即缓解期。因此，根据诊断标准，双相患者的心境阶段确实只包括这 3 个阶段。实际上，过往也有研究采用三阶段划分(Adida et al., 2011; Thomas et al., 2007)。为了便于读者理解，我们在引言部分增加了有关双相患者心境阶段的说明（见正文 P3）。为了便于审稿，我们将相关文字粘贴至下方。

双相患者的风险偏好可能随心境阶段变化而变化。根据美国精神疾病诊断与统计手册第五版(Diagnostic and Statistical Manual of Mental Disorder Fifth Edition, DSM-V; APA, 2013), 双相患者有时会表现出心境膨胀、极度自信, 目标活动增多, 易过度参与高风险活动, 即<轻>躁狂期, 包括躁狂发作和轻躁狂发作; 而有时会表现出情绪低落、自我贬低, 思考能力下降或犹豫不决, 几乎对所有活动的兴趣或愉悦感都明显减少, 即重性抑郁期(简称抑郁期), 对应重性抑郁发作。两者常统称为发病期。而发病间隙患者的症状得到缓解, 情绪趋于稳定, 即缓解期。因此, 患者存在 3 种不同的心境阶段, 且在不同心境阶段下他们的认知、动机和行为表现都有明显差异。

参考文献

- Adida, M., Jollant, F., Clark, L., Besnier, N., Guillaume, S., Kaladjian, A., Mazzola-Pomietto, P., Jeanningros, R., Goodwin, G. M., Azorin, J. M., & Courtet, P. (2011). Trait-related decision-making impairment in the three phases of bipolar disorder. *Biological Psychiatry*, 70(4), 357–365. <https://doi.org/10.1016/j.biopsych.2011.01.018>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>.
- Thomas, J., Knowles, R., Tai, S., & Bentall, R. P. (2007). Response styles to depressed mood in bipolar affective disorder. *Journal of Affective Disorders*, 100(1–3), 249–252. <https://doi.org/10.1016/j.jad.2006.10.017>

意见 3: 作者指出前人用了自陈量表、实验室行为任务和日常风险行为这三种方式评估风险偏好, 并指出不同类型的测量特征可能调节风险偏好的差异。这里罗列的日常风险行为的范围太过于宽泛, 包括抽烟、酗酒、危险驾驶、赌博、过度消费、工作成瘾、网络成瘾、性成瘾、攻击性等。这些划分的依据是什么? 本研究在文献筛选过程中把哪些作为目标风险行为纳入? 如何保证本研究纳入的日常风险行为全面且没有遗漏? 在我看来, 自陈量表和日常风险行为并非互斥类别, 例如, 当采用自陈量表测量日常攻击行为时, 这种测量方式应属于哪类? 此外, 自陈量表既可以测量风险行为意向, 也可以测量风险行为, 风险行为意向和风险行为的结果是否有差异?

回应: 感谢您的问题和启发性建议。您的问题让我们意识到确实有必要对风险决策测量方式的划分, 尤其是日常风险行为, 做进一步澄清。我们之前在文献搜索时采用 "risk behavior" 及相近词汇 (如 risk taking、risk seeking 等) 查找文献, 上述具体风险行为 (如抽烟、酗酒、危险驾驶、赌博、过度消费、工作成瘾、网络成瘾、性成瘾、攻击性等) 均来自检索的文献。

为了使得纳入的风险行为尽可能合理、全面，我们主要参考了美国疾控中心开发的青年风险行为检测系统(YRBSS; Brener et al., 2004)和我国疾控中心在此基础上调整的健康危险行为调查问卷(季成叶, 2007)，并结合过往风险行为元分析(Roberts et al., 2021; Tinner et al., 2018)，**最终聚焦他们都关注且无争议的风险行为**，包括物质使用、性风险行为、不健康饮食、缺乏锻炼和危险驾驶等。由于明确了日常风险行为的分类，我们增加了具体风险行为的关键词并再次进行了文献搜索，**具体关键词见正文附表 1**。

为了便于读者理解，本次修改时我们在正文中增加了日常风险行为的分类描述，**具体见引言部分(正文 P4)**。为了便于审稿，我们将相关文字粘贴至下方。

关于专家提及的测量方式的划分问题，原文表述确实不够准确，尤其是未在文字上清晰区分自陈量表和日常风险行为。我们基于 Frey 等人(2017)对风险决策测量方式的划分，再次明确了这 3 种测量方式，并调整了测量方式的分类表达，即**风险态度量表、行为实验任务和日常风险行为**（原文中的自陈量表实际是指风险态度量表）。其中，风险态度量表通过一些抽象问题（如“你是一个风险寻求者还是风险规避者”）或更具体的情境问题（如“你多大可能性将一天的收入用于赌马”）测量个体一般性风险态度或对特定领域的风险态度；行为实验任务则通过决策对任务、赌博任务或更具模拟现实性的实验任务测量个体的风险偏好；而日常风险行为主要在临床和流行病学研究中使用，测量的是个体真实发生的行为表现（如，“你一天抽多少支烟”）。

风险态度和日常风险行为的相关较高，但并不完全一致。正如专家所言，风险态度关乎风险行为意向，而日常风险行为则是生活中实际发生的可观测的真实风险行为。例如，对于抽烟，前者评估个体抽烟的意愿或可能性，而后者评估个体抽烟的真实状况（如频率）。因此，本次修改时我们进一步明确了测量方式的归类（见 **P11 方法部分**），以分析测量方式的调节作用。**为了便于审稿，我们将相关文字粘贴至下方。**

——引言，正文第 4 页——

而第三种测量方式则常见于临床和流行病学研究，直接观察和记录个体真实发生的日常风险行为（如，“你一天抽多少支烟”）。目前研究多关注健康领域的日常风险行为。如，美国疾控中心的青年风险行为检测系统(Youth Risk Behavior Surveillance System, YRBSS)涵盖了非故意伤害行为、烟草使用、酒精和其他物质使用、与意外怀孕和性传播感染有关的性行为、不健康饮食和缺乏锻炼等 6 种风险行为(Brener et al., 2004)。我国疾控中心也有类似调查，其保留了 YRBSS 的后 3 种行为，同时将前 3 种行为调整为 4 种行为，即导致非故意伤害的行为、导致故意伤害的行为、物质成瘾行为和精神成瘾行为(含网络成瘾、游戏成

瘾、赌博成瘾、消费成瘾等) (季成叶, 2007)。此外, 过往对风险行为的元分析也多关注物质使用、性风险行为、危险驾驶、赌博等行为(e.g., Roberts et al., 2021; Tinner et al., 2018)。据此, 本研究将重点关注上述研究涉及的典型日常风险行为, 包括物质使用、性风险行为、不健康饮食、缺乏锻炼以及非故意伤害行为中的危险驾驶行为等。

——方法, 正文第 11 页——

基于过往研究对风险决策测量方式的分类(Frey et al., 2017), 本研究在纳入的文献中区分了 3 种测量方式, 以考察双相障碍对风险决策的影响是否因测量方式的不同而不同。具体编码为: 1 = 风险态度量表; 2 = 行为实验任务; 3 = 日常风险行为。值得注意的是, 风险态度量表聚焦个体的一般性风险态度或在特定领域的风险态度, 而日常风险行为则聚焦个体真实发生的风险行为。

参考文献

- 季成叶. (2007). 青少年健康危险行为. *中国学校卫生*, 28(4), 289–291.
- Brener, N. D., Kann, L., Kinchen, S. A., Grunbaum, J. A., Whalen, L., Eaton, D., Hawkins, J., & Ross, J. G. (2004). Methodology of the youth risk behavior surveillance system. *MMWR. Recommendations and reports : Morbidity and mortality weekly report. Recommendations and Reports*, 53(RR-12), 1–13.
- Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, 3(10), e1701381. <https://doi.org/10.1126/sciadv.1701381>
- Roberts, D. K., Alderson, R. M., Betancourt, J. L., & Bullard, C. C. (2021). Attention-deficit/hyperactivity disorder and risk-taking: A three-level meta-analytic review of behavioral, self-report, and virtual reality metrics. *Clinical Psychology Review*, 87, 102039. <https://doi.org/10.1016/j.cpr.2021.102039>
- Tinner, L., Caldwell, D., Hickman, M., MacArthur, G. J., Gottfredson, D., Lana Perez, A., Moberg, D. P., Wolfe, D., & Campbell, R. (2018). Examining subgroup effects by socioeconomic status of public health interventions targeting multiple risk behaviour in adolescence. *BMC Public Health*, 18(1), 1180. <https://doi.org/10.1186/s12889-018-6042-0>

意见 4: 关于样本的人口学特征差异, 性别差异是否适用于双相患者可理解为男性和女性双相患者是否存在风险偏好差异, 和本文的核心问题没有实质性关系, 年龄和风险偏好负相关并不能直接推导出年龄能够调节双相障碍(vs. 正常人群)与风险偏好的关系。建议作者进一步梳理其中的研究问题和理论逻辑。

回应: 感谢您的建议。原文关于性别和年龄的调节作用的表述确实不够清晰, 为此本次修改时我们完善了引言中有关性别和年龄调节作用的论述, 具体见引言部分(正文 P5–P6)。为了便于审稿, 我们将相关文字粘贴至下方。

——引言, 正文第 5–6 页——

值得关注的是，样本的人口学特征差异也可能导致双相患者风险偏好的研究结果不一致。Byrnes 等人(1999)的元分析揭示，在 14 种风险决策测量（共 16 种）中，男性均比女性更偏好风险寻求。日常观察亦发现，相较于男性，女性对风险事件更敏感(Paluckaitė & Žardeckaitė-Matulaitienė, 2017)，更厌恶风险(Croson & Gneezy, 2009)，更少做出风险行为(Charness & Gneezy, 2012; Dir et al., 2014)。双相患者中，性别与风险偏好的关系同样值得探讨。研究表明，男女双相患者在临床症状上存在差异(Benazzi et al., 2003)。如，女性更易抑郁，自杀意图高，容易出现进食、焦虑等问题；而男性更易躁狂，伴随酗酒和其他物质滥用问题(Azorin et al., 2013)。然而，这些症状差异是否意味着风险偏好的不同尚不清晰，因此不同性别的患者与健康个体的风险偏好差异是否不同仍需探究。本研究旨在探讨女性占比对双相患者与健康个体风险偏好差异的影响，以检验性别的调节作用。如若双相患者存在与健康个体相似的性别差异，那么女性占比的增加可能会减少两者的风险偏好差异；反之，如若这一性别差异消失或逆转，则两者的风险偏好差异可能会随女性占比增加而扩大。

年龄对风险偏好的影响也值得关注。研究表明，相较青少年，成年人常表现出更低的风险寻求(Defoe et al., 2015; Josef et al., 2016)。然而，这一趋势在双相患者中是否存在值得探究。双相患者常伴有认知功能损害(Bora et al., 2009)，且随着年龄增长，这种损害可能加剧(John et al., 2019; Lewandowski et al., 2014)，因而可能导致老年患者做出更多非理性的风险决策。另一方面，双相障碍的发病年龄较早(Bolton et al., 2021)，超过六成患者在 21 岁前发病(Duffy et al., 2009)。青少年和成年早期患者常面临混合症状、病程长和共病障碍多等问题，这可能对其正常发育和社会心理功能造成较大影响，继而增加了他们自杀、物质滥用、学业问题等风险(Birmaher et al., 2013)。据此推测，年轻患者也可能更偏好风险寻求。可见，目前研究对双相患者的年龄如何影响其风险偏好尚存争议，因而双相患者与健康个体的风险偏好差异如何受年龄影响也不得而知。如若年龄对双相患者和健康个体的风险偏好的影响趋势相似，那么两者的风险偏好差异就可能随年龄增大而减小；反之，如若这一影响趋势减弱甚至方向相反，那么两者的风险偏好差异则可能随年龄增大而扩大。因此，我们将年龄作为潜在调节变量，探究其是否影响双相患者与健康个体在风险偏好上的差异。此外，根据元分析需要，本研究还将纳入样本受教育程度和所属地区作为潜在调节变量。

参考文献

Azorin, J. M., Belzeaux, R., Kaladjian, A., Adida, M., Hantouche, E., Lancrenon, S., & Fakra, E. (2013). Risks

- associated with gender differences in bipolar I disorder. *Journal of Affective Disorders*, 151(3), 1033–1040. <https://doi.org/10.1016/j.jad.2013.08.031>
- Benazzi F. (2003). The role of gender in depressive mixed state. *Psychopathology*, 36(4), 213–217. <https://doi.org/10.1159/000072792>
- Birmaher B. (2013). Bipolar disorder in children and adolescents. *Child and Adolescent Mental Health*, 18(3), 140–148. <https://doi.org/10.1111/camh.12021>
- Bolton, S., Warner, J., Harriss, E., Geddes, J., & Saunders, K. E. A. (2021). Bipolar disorder: Trimodal age-at-onset distribution. *Bipolar Disorders*, 23(4), 341–356. <https://doi.org/10.1111/bdi.13016>
- Bora, E., Yucel, M., & Pantelis, C. (2009). Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *Journal of Affective Disorders*, 113(1-2), 1–20. <https://doi.org/10.1016/j.jad.2008.06.009>
- Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125(3), 367–383. <https://doi.org/10.1037/0033-2909.125.3.367>
- Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior and Organization*, 83(1), 50–58. <https://doi.org/10.1016/j.jebo.2011.06.007>
- Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2), 448–474. <https://doi.org/10.1257/jel.47.2.448>
- Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. (2015). A meta-analysis on age differences in risky decision making: adolescents versus children and adults. *Psychological Bulletin*, 141(1), 48–84. <https://doi.org/10.1037/a0038088>
- Dir, A. L., Coskunpinar, A., & Cyders, M. A. (2014). A meta-analytic review of the relationship between adolescent risky sexual behavior and impulsivity across gender, age, and race. *Clinical Psychology Review*, 34(7), 551–562. <https://doi.org/10.1016/j.cpr.2014.08.004>
- Duffy A. (2009). The early course of bipolar disorder in youth at familial risk. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 18(3), 200–205.
- John, A., Patel, U., Rusted, J., Richards, M., & Gaysina, D. (2019). Affective problems and decline in cognitive state in older adults: a systematic review and meta-analysis. *Psychological medicine*, 49(3), 353–365. <https://doi.org/10.1017/S0033291718001137>
- Josef, A. K., Richter, D., Samanez-Larkin, G. R., Wagner, G. G., Hertwig, R., & Mata, R. (2016). Stability and change in risk-taking propensity across the adult life span. *Journal of Personality and Social Psychology*, 111(3), 430–450. <https://doi.org/10.1037/pspp0000090>
- Lewandowski, K. E., Sperry, S. H., Malloy, M. C., & Forester, B. P. (2014). Age as a predictor of cognitive decline in bipolar disorder. *The American Journal of Geriatric Psychiatry*, 22(12), 1462–1468. <https://doi.org/10.1016/j.jagp.2013.10.002>
- Paluckaitė, U., & Žardeckaitė-Matulaitienė, K. (2017). Adolescents' perception of risky behaviour on the Internet. In Z. Bekirogullari, M. Y. Minas, & R. X. Thambusamy (Eds.), *Health and Health Psychology - icH&Hpsy 2017*, vol 30. European Proceedings of Social and Behavioural Sciences (pp. 284–292). Future Academy. <https://doi.org/10.15405/epsbs.2017.09.27>

意见 5：既然作者将攻击行为考虑为 risk-taking behavior 的一种，为何没有纳入其他考察了双相障碍与攻击行为关系的文献？比如 Ballester, J., Goldstein, T., Goldstein, B., Obreja, M., Axelson, D., Monk, K., ... & Birmaher, B. (2012). Is bipolar disorder specifically associated with

aggression?. *Bipolar Disorders*, 14(3), 283-290.

回应：感谢您的宝贵意见和建议。我们在最初文献检索时采用的是“risk behavior”、“risk taking”、“risk seeking”等相近关键词，未具体设置与攻击行为相关的关键词，因而导致您提及的这篇文献未被纳入。值得重视的是，您在审稿意见中还提到“攻击行为在多大程度上属于风险行为的概念范畴”这一重要问题，这启发我们重新审视攻击行为与风险行为的区别。Allen 和 Anderson(2017)以及 Anderson 和 Bushman(2002)认为，攻击行为一般具有故意伤害的目的、会造成他人身心伤害，同时受害者还有回避动机。而 Leigh(1999)认为，风险行为是指能使行为人获得某种形式奖赏（而非故意伤害他人）但同时又伴随潜在自我危险的行为。美国疾控中心开发的青年风险行为检测系统也未将故意伤害行为（如攻击）纳入其中。基于攻击行为和风险行为在定义上的差别，本研究在分析双相障碍与风险决策的关系时暂不纳入攻击行为。

参考文献

- Allen, J. J., & Anderson, C. A. (2017) Aggression and violence: definitions and distinctions. In: P Sturme, editor. *The Wiley Handbook of Violence and Aggression*. West Sussex, UK.
- Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, 53, 27–51. <https://doi.org/10.1146/annurev.psych.53.100901.135231>
- Leigh, B. C. (1999). Peril, chance, adventure: Concepts of risk, alcohol use and risky behavior in young adults. *Addiction*, 94(3), 371–383. <https://doi.org/10.1046/j.1360-0443.1999.9433717.x>

意见 6：建议在“2.3 数据提取和质量评价”部分具体说明每个提取的变量（包括调节变量）的具体分类和编码方式。例如，患者所处的心境状态分成几类？当心境状态的信息不明确时如何进行分类处理？

回应：感谢您的建议。我们将提取变量的情况进行了说明，见方法部分(正文 P8–P9)。为了便于审稿，我们将相关文字粘贴至下方。

——方法，正文第 8–9 页——

本研究需要提取的数据变量如下：

- (1) 基本信息：文献信息（第一作者和发表年份）和效应量数量
- (2) 样本特征
 - a) 样本量
 - b) 年龄：提取文献中各个组别的平均年龄（标准差）。
 - c) 性别比：提取文献中各个组别的男女人数或性别比，最终都转换成女性占

比。

- d) 受教育程度：考虑到纳入文献大多采用受教育年限来反映受教育程度，所以本研究提取文献中各个组别的受教育年限。若文献未报告该指标，本研究将其标为缺失值。
- e) 所属国家及地区：提取文献中样本所属国家，并根据国家所在地区进行标注（如，欧洲、北美洲、亚洲等）。
- f) 心境阶段：提取文献中样本当时所处的心境阶段（缓解期、<轻>躁狂期和抑郁期）。若文献未报告心境阶段或者将不同心境阶段的患者混合为一组进行比较，本研究将其标为缺失值。
- g) 诊断标准：提取文献中用于判断样本患病与否的诊断标准。
- h) 抑郁/躁狂程度：提取文献中样本的抑郁/躁狂程度（均值和标准差）及其量表名称。

（3）测量特征

- a) 测量方式：提取文献中所涉及的风险决策测量方式名称（风险态度量表、日常风险行为、行为实验任务）和所有反映风险偏好的指标名称。
- b) 风险决策：提取文献中样本在所有风险决策指标上的表现（均值和标准差）。

意见 7：在统计分析时，为何剔除大于或小于所有效应量均值或中位数 3 个标准差的效应量？其中，中位数上下 3 个标准差的标准有何依据？

回应：感谢您的意见。Kepes 和 Thomas(2018)认为，在元分析中，纳入的研究应具有代表性，才能帮助研究者得出一般性的结论。在多数情况下，极端值会影响研究的代表性，以致干扰总体效应量大小甚至得出有偏的结论(Lipsey & Wilson, 2001)。因此，本研究在正式分析前剔除极端值，选用标准为“大于或小于所有效应量均值或中位数 3 个标准差的效应量”。这一标准在以往元分析研究中常被采用(e.g., Fernandes & Garcia-Marques, 2020; Kathawalla & Syed, 2021)。此外，Leys 等人(2013)提出，均值本身容易受极端值的干扰，应考虑采用中位数。为了避免极端值对均值的可能影响，本研究同时采用均值和中位数上下 3 个标准差作为极端值排除标准，符合其一者将被排除。我们在正文中也增加了极端值剔除标准的依据（正文 P9）。为了便于审稿，我们将相关文字粘贴至下方。

——方法，正文第 9 页——

为了避免极端值对研究结论的影响(Kepes & Thomas, 2018)，本研究参考过往的元分析

研究(e.g., Fernandes & Garcia-Marques, 2020; Kathawalla & Syed, 2021), 选择剔除大于和小于所有效应量的均值或中位数 3 个标准差的单个效应量。

参考文献

- Fernandes, A. C., & Garcia-Marques, T. (2020). A meta-analytical review of the familiarity temporal effect: Testing assumptions of the attentional and the fluency-attributional accounts. *Psychological Bulletin*, 146(3), 187–217. <https://doi.org/10.1037/bul0000222>
- Kathawalla, U., & Syed, M. (2021). Discrimination, Life Stress, and Mental Health among Muslims: A Preregistered Systematic Review and Meta-Analysis. *Collabra: Psychology*, 7(1), 28248. <https://doi.org/10.1525/collabra.28248>
- Kepes, S., & Thomas, M. A. (2018). Assessing the robustness of meta-analytic results in information systems: Publication bias and outliers. *European Journal of Information Systems*, 27(1), 90–123. <https://doi.org/10.1080/0960085X.2017.1390188>
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49, 764–766. <https://doi.org/10.1016/J.JESP.2013.03.013>
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

意见 8：为何要将南美洲和大洋洲合并？鉴于大洋洲的效应量只有 1 个，在调节效应分析时不将其纳入分析即可。

回应：感谢您的建议。根据您的建议，本次修改中我们对效应量只有 1 个的大洋洲样本不纳入调节效应的分析，具体见结果部分(正文 P19)。

意见 9：为何要将自陈量表和日常风险行为的研究合并？既然测量方式也是调节变量之一，建议将所有数据放在一个数据文件中，对不同变量的相应类别都进行区分和编码。

回应：感谢您的意见和建议。我们采纳您的建议，先将所有效应量放在一个数据文件进行了总体模型的分析，并将测量方式（3 种）作为调节变量进行分析。本次修改时，我们在方法部分增加了对测量方式这一调节变量的描述(正文 P11)，为了便于审稿，我们将相关文字粘贴至下方。

结果发现，测量方式的调节作用不显著($p = .070$)，但风险态度量表和行为实验任务的差异显著，其他两两间比较差异不显著。因此我们根据测量方式区分出行为实验任务模型和日常风险态度和行为模型进行分析。其中，将风险态度量表和日常风险行为合并原因有：① 这 2 种测量主要以量表或问卷形式进行，测量形式上较接近。而且，研究确实发现这两种测量方式的相关较高(Frey et al., 2017)；② 在本研究中，这 2 种测量方式的效应量无

显著差异；③ 风险态度量表的文献数较少（5 篇），若再予以区分，可能会影响研究结论的稳健性。本次修改时我们对合并原因进行了简要阐述，见结果部分(正文 P20)。为了便于审稿，我们将相关文字粘贴至下方。

——方法，正文第 11 页——

基于过往研究对风险决策测量方式的分类(Frey et al., 2017)，本研究在纳入的文献中区分了 3 种测量方式，以考察双相障碍对风险决策的影响是否因测量方式的不同而不同。具体编码为：1 = 风险态度量表；2 = 行为实验任务；3 = 日常风险行为。值得注意的是，风险态度量表聚焦个体的一般性风险态度或在特定领域的风险态度，而日常风险行为则聚焦个体真实发生的风险行为。

——结果，正文第 20 页——

尽管本研究发现患者在所有测量方式上均比健康个体更偏好风险寻求，但每种测量方式的效应量大小不同，尤其是行为实验任务和风险态度量表的差异显著($\beta = 0.371, p = .022$)。因此，本研究区分测量方式形成 2 个子模型再次检验，包括行为实验任务模型和日常态度和行为模型，其中后者包括风险态度量表和日常风险行为的研究。本研究将这两种方式进行合并的原因，一是这 2 种测量主要以量表或问卷形式进行，测量形式上较接近，两者效应量在本研究中也无显著差异，且以往研究表明这两种测量方式的相关较高(Frey et al., 2017)；二是本研究中纳入风险态度量表的文献仅 5 篇，将其单独分为一个子模型，可能会影响结论的稳健性。

参考文献

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, 3(10), e1701381. <https://doi.org/10.1126/sciadv.1701381>

意见 10：虽然我没有一一核对不同研究的效应量编码，但我发现 Clark et al. (2001)这篇文献中根据 IGT 计算的效应量(Hedges' g)存在问题。这篇文献中 manic patients($M = 48.3, SD = 13.3, n = 15$)的风险偏好水平比 comparison subjects ($M = 37.2, SD = 14.0, n = 30$)更高，计算的 Hedges' g 应为正值，但作者计算的结果是负的。个别研究效应量计算错误会导致元分析的整体结果不可信。与此相对应，作者还尝试通过其他途径（如查找相关元分析）获得相关数据，这种间接获取数据的方式的可靠性存疑。

回应：感谢您的宝贵意见。确实如您所说，通过 Clark et al. (2001)的数据计算的效应量为正

值。我们此前计算并未出错，但在后期编码时出现了偏差。不同行为实验任务测量指标的含义不同，如 IGT 任务选用的净分数，分数越高表示个体越偏好风险规避；而 BART 任务选用的调整后平均充气次数，次数越多表示个体越偏好风险寻求。那么在这两种任务中正效应量的含义就不一致。我们在对效应量进行编码时，正效应量表示双相患者比健康个体更偏好风险寻求，负效应量表示患者比健康个体更偏好风险规避。Clark et al. (2001)的文章中采用了 IGT 任务但没有采用经典指标净分数，而是风险卡牌的选择数量（越多越偏好风险寻求），之前我们未注意到这一点，所以未能正确编码。为了保证结果的可靠性，我们此次对纳入文献的指标含义、计算与编码进行了核对。为了便于读者理解，我们完善了介绍编码的文字，具体见方法部分(正文 P9)。为了便于审稿，我们将相关文字粘贴至下方。

为了减少因数据缺失无法计算效应量而剔除的文献，我们尝试在同一主题已发表的元分析中获取数据，最终 Yechiam 等人(2008)的效应量数据来自 Edge 等人(2013)元分析。我们完善了这部分的文字描述，具体见方法部分(正文 P7)。为了便于审稿，我们将相关文字粘贴至下方。

——方法，正文第 9 页——

g 值采用双相患者在风险决策测量指标中的表现减去健康个体的表现，但在某些风险决策测量方式中不同指标的大小对应的风险偏好不同。例如，IGT 任务的测量指标为净分数，个体净分数越大表明其越偏好风险规避，此时 g 值为正表明双相患者比健康个体更偏好风险规避；而 BART 任务的测量指标为调整后平均充气次数，个体充气次数越多则表明其越偏好风险寻求，此时 g 值为正表明双相患者比健康个体更偏好风险寻求。因此，不同测量方式中 g 值为正便有 2 种截然相反的含义，所以本研究对效应量进行统一编码，使 g 值为正均表示双相患者比健康个体更偏好风险寻求， g 值为负则表示更偏好风险规避。

——方法，正文第 7 页——

如果发现未报告可转换指标的研究，我们尝试通过在已发表的双相障碍与风险决策的元分析(Edge et al., 2013; Richard-Devantoy et al., 2016; Ramírez-Martín et al., 2020)中查找获得相关数据，若仍无法获得则排除该研究。最终 Yechiam 等人(2008)的效应量数据可从 Edge 等人(2013)的元分析中获取。

参考文献

Clark, L., Iversen, S. D., & Goodwin, G. M. (2001). A neuropsychological investigation of prefrontal cortex involvement in acute mania. *American Journal of Psychiatry*, 158(10), 1605–1611.

<https://doi.org/10.1176/appi.ajp.158.10.1605>

Edge, M. D., Johnson, S. L., Ng, T., & Carver, C. S. (2013). Iowa gambling task performance in euthymic bipolar I disorder: A meta-analysis and empirical study. *Journal of Affective Disorders*, 150(1), 115–122.

<https://doi.org/10.1016/j.jad.2012.11.027>

Ramírez-Martín, A., Ramos-Martín, J., Mayoral-Cleries, F., Moreno-Küster, B., & Guzmán-Parra, J. (2020). Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. *Psychological Medicine*, 50(13), 2141–2153. <https://doi.org/10.1017/S0033291720003086>

Richard-Devantoy, S., Olié, E., Guillaume, S., & Courtet, P. (2016). Decision-making in unipolar or bipolar suicide attempters. *Journal of Affective Disorders*, 190, 128–136. <https://doi.org/10.1016/j.jad.2015.10.001>

Yechiam, E., Hayden, E. P., Bodkins, M., O'Donnell, B. F., & Hetrick, W. P. (2008). Decision making in bipolar disorder: a cognitive modeling approach. *Psychiatry Research*, 161(2), 142–152. <https://doi.org/10.1016/j.psychres.2007.07.001>

意见 11：作者将风险态度和行为所属领域划分为健康、经济、安全、社会、其他以及总体态度，这一划分方式有何依据？不同领域划分的编码标准是什么？在表 2 中，为何作者将身体攻击和言语攻击分别作为安全和社会领域的风险行为？这两者均反映了反社会倾向，在多大程度上属于风险行为的概念范畴？如果属于风险行为，为何两者属于不同领域？

回应：感谢您的意见。我们主要参考了 Blais 和 Weber(2006)开发的风险决策领域特异性量表和相应修订版本中对决策领域的分类，即健康、经济、安全、社会、娱乐等。考虑到实际纳入的文献并未涉及上述所有领域，因此在进行元分析时基于上述分类结合文献内容确定所涉领域。

正如专家指出的，我们在表 2 中将身体攻击和言语攻击分别作为安全和社会领域的风险行为确有不妥，具体原因见对问题 3 的回答。因此，本次修改将不纳入攻击行为。

参考文献

Blais, A.-R., & Weber, E. U. (2006). A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. *Judgment and Decision Making*, 1(1), 33–47. <https://doi.org/10.1017/s1930297500000334>

意见 12：既然测量特征是作者要考察的调节变量之一，为何在报告结果时不先报告总效应量，再分析测量特征的调节效应？研究方法部分关于不同测量方式之间的效应量差异检验应该放在结果部分报告。

回应：感谢您的建议。根据您的建议，我们先将所有效应量放在一个数据文件进行总体模型的分析，报告总体效应量大小，并对测量方式进行调节效应检验，然后进一步根据测量方式区分模型再次分析报告结果。我们已将关于不同测量方式之间的效应量差异检验放至结果部分(正文 P18)。

意见 13：关于异质性检验，为何不对所有效应量只进行一次异质性检验，再进行调节效应检验？

回应：感谢您的建议。确实如您所说，应先对所有效应量进行分析。因此，我们首先对所有效应量进行异质性检验(正文 P16)，然后区分分子模型后再次进行检验(正文 P20 和 P23)。其中，考虑到测量方式不同可能会影响异质性，因此区分分子模型后再次进行检验可以更好地认识效应量间的异同。

意见 14：作者在文中报告结果时多次使用“(边缘)显著”的表述，容易混淆对结果的解读，强烈建议作者不要报告边缘显著的结果，统一描述为不显著，只报告统计上显著($p < .05$)的结果。对于类似 $p = .09$ 这样的结果没有必要花大量篇幅再解释这类结果。

回应：感谢您的建议。根据您的建议，我们将文中所有 $0.05 < p < 0.1$ 的结果统一描述为不显著(由于在文中较为分散，具体见结果部分)。除了重点解释显著结果外，本文主要对有启发但可能因效应量太少使结果不稳健的不显著结果进行了一定的解释和分析。

意见 15：作者指出本研究发现的双相患者比健康个体更偏好风险寻求的结果与 Ramírez-Martín 等人(2020)的元分析结果一致，但是 Ramírez-Martín 等人(2020)并没有发现双相障碍与正常人群在风险行为上的显著差异，何来一致的结果？

回应：感谢您的意见。Ramírez-Martín 等人(2020)中有 3 种风险决策行为实验任务即爱荷华赌博任务、剑桥赌博任务和气球模拟风险任务。其中，爱荷华赌博任务的效应量显著(Hedges' $g = -0.43, p < .001$) (该研究中负效应量表示患者更偏好风险寻求)，这与本研究结果一致。而在剑桥赌博任务(Hedges' $g = 0.41, p = .199$)和气球模拟风险任务(Hedges' $g = 0.43, p = .280$)中的效应量不显著(该研究中正效应量表示双相患者更偏好风险寻求)，这与本研究区分任务后的结果部分一致。为了更准确表达避免歧义，我们完善了文字表达，见讨论部分(正文 P24)。为了便于审稿，我们将相关文字粘贴至下方。

——讨论，正文第 24 页——

研究发现，双相患者比健康个体更偏好风险寻求(Hedges' $g = 0.301$)，该结果与 Edge 等人(2013)、Richard-Devantoy 等人(2016)以及 Ramírez-Martín 等人(2020)中 IGT 任务的元分析结果一致，也与大多实证研究结果相同(e.g., Malloy-Diniz et al., 2011; Mason et al., 2014)，这回答了本研究第一个问题，即“双相患者与健康个体的风险偏好有无差异”。

参考文献

- Edge, M. D., Johnson, S. L., Ng, T., & Carver, C. S. (2013). Iowa gambling task performance in euthymic bipolar i disorder: A meta-analysis and empirical study. *Journal of Affective Disorders*, 150(1), 115–122. <https://doi.org/10.1016/j.jad.2012.11.027>
- Malloy-Diniz, L. F., Neves, F. S., de Moraes, P. H. P., De Marco, L. A., Romano-Silva, M. A., Krebs, M. O., & Corrêa, H. (2011). The 5-HTTLPR polymorphism, impulsivity and suicide behavior in euthymic bipolar patients. *Journal of Affective Disorders*, 133(1-2), 221–226. <https://doi.org/10.1016/j.jad.2011.03.051>.
- Mason, L., O'Sullivan, N., Montaldi, D., Bentall, R. P., & El-Deredy, W. (2014). Decision-making and trait impulsivity in bipolar disorder are associated with reduced prefrontal regulation of striatal reward valuation. *Brain*, 137(8), 2346–2355. <https://doi.org/10.1093/brain/awu152>.
- Ramírez-Martín, A., Ramos-Martín, J., Mayoral-Cleries, F., Moreno-Küster, B., & Guzmán-Parra, J. (2020). Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. *Psychological Medicine*, 50(13), 2141–2153. <https://doi.org/10.1017/S0033291720003086>
- Richard-Devantoy, S., Olié, E., Guillaume, S., & Courtet, P. (2016). Decision-making in unipolar or bipolar suicide attempters. *Journal of Affective Disorders*, 190, 128–136. <https://doi.org/10.1016/j.jad.2015.10.001>

意见 16：很多边缘显著的结果在多大程度上具有统计学意义？在讨论这些结果时要谨慎，不要过分解读或夸大这些结果。

回应：感谢您的建议。确实如您所问的，边缘显著结果的统计学意义不大，但是考虑到部分结果可能因目前研究量少而难以探测显著差异，因此我们对不显著的原因做了一些推测和分析。本次修改时，我们完善了边缘显著结果的表述，由于比较散落，在此不一一列举，见讨论中红色修改的部分。

意见 17：作者在讨论部分放一个全英文的图有何必要？显得很突兀。如果需要，直接文字表述核心观点即可。

回应：感谢您的建议。确实如您所说，不应直接放入全英文的图，本次修改时我们对图文进行翻译，重新绘制成中文图 5（见正文 P29）。为了便于审稿，我们将图粘贴至下方。

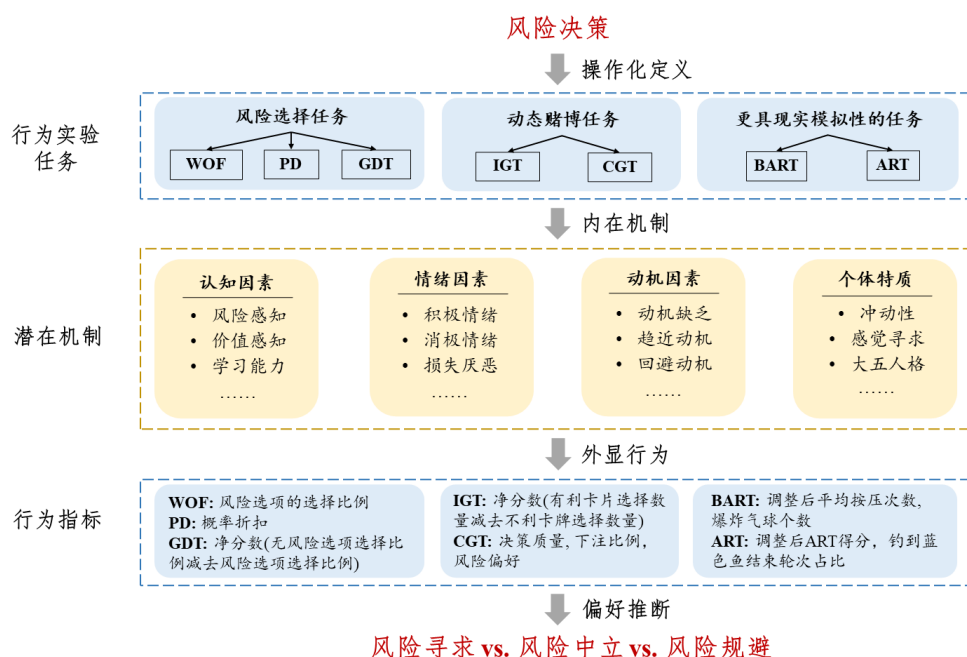


图 5 不同行为实验任务中反映不同风险偏好的潜在框架（摘自“因匿名审稿隐去该参考文献”）

注：WOF = 幸运轮盘任务；PD = 概率折扣任务；GDT = 骰子任务；IGT = 爱德华赌博任务；CGT = 剑桥赌博任务；BART = 气球模拟风险任务；ART = 钓鱼风险任务。

意见 18：作者在讨论部分指出未来应该重视纵向研究，但是这类比较双相障碍患者和健康人群的研究本质上是准实验研究，虽然无法随机分配到实验组和控制组，但只要匹配不同组被试的其他特征，也能在一定程度上排除其他变量的影响，从而推断因果关系。纵向研究在操作层面上如何实现？是否确实能够增强因果推断的效力？

回应：感谢您的意见和建议。确实如您所说，比较双相患者和健康个体的研究是准实验研究，只要通过匹配不同组被试的其他特征，能一定程度推断因果关系。然而，这种因果推断的效力仍然会受到多种因素的制约，例如样本的代表性、测量的准确性、控制变量的选择等。尤其是考虑到患者收集难度和个体差异，因此开展纵向研究是增强推断效力的方法之一。具体地，纵向研究能够揭示变量之间的时间顺序，观察到变量之间的动态关系，可以进一步支持因果关系的推断。

在操作层面上，可通过对双相患者和健康个体进行多次测量，包括风险决策、症状表现、生活质量等，了解各变量的变化趋势，以及两者的差异随时间的变化，进而通过交叉滞后模型或格兰杰因果检验等分析方法推测因果关系。对双相患者的长期追踪，不仅能从被试内角度检验心境阶段的作用，而且还能捕捉风险决策变化与发病与否的关系。本次修改时，我们完善了有关重视纵向研究的内容，具体见讨论部分(正文 P30)。为了便于审稿，我们将相关文字粘贴至下方。

最后，未来应重视纵向追踪研究，以弥补当前双相障碍与风险决策关系研究中依赖横断研究在因果推断上的局限。横断研究虽能揭示双相患者与健康个体在风险偏好上的差异，但受限于个体间差异，当样本量较小时尤其如此。而纵向追踪研究不仅能消除部分个体间差异的影响，追踪患者风险决策随时间的变化，构建患者风险决策的动态发展模型，还能从被试内角度比较同一患者随心境阶段转换的风险偏好变化，更深入地理解心境阶段的作用。因此，未来研究应重视对双相患者队列的长期随访。

意见 19：预注册文档不可见，无法了解与正文中报告的方法是否一致。

回应：感谢您的意见。非常抱歉此前由于我们的疏忽，误删了最后一位数字造成预注册文档不可见。目前已修改，可通过 <https://www.crd.york.ac.uk/PROSPERO/>，搜索预注册编号 CRD42022346204。

意见 20：“双相患者的风险偏好相比健康个体是否有改变以及变化方向和程度需要进行系统性梳理。”这一表述不当。本研究关注是否有差异以及差异的方向和程度，没有涉及时间上的变化或改变，只是对不同组别的被试进行了比较。

回应：感谢您的建议。根据您的建议，我们已将上述表述修改为“双相患者的风险偏好与健康个体是否存在差异以及差异的方向和程度需要进行系统性梳理”（见正文 P6）。同时，我们通读全文并修改了与此类表述相似的表达，由于比较散落，在此不一一列举，可具体见文中红色修改的部分。

意见 21：图 2 中的 HC 和 BD 分别代表什么需要说明。

回应：感谢您的建议。我们在图 2 中增加了 HC 和 BD 的图注（见正文 P16），便于读者理解。

意见 22：图 4 的漏斗图中的横坐标 Observed Outcome 指的是什么？是 Hedges' g 吗？

回应：感谢您的意见。图 4 横坐标 Observed Outcome 表示观察到的效应量即 Hedges' g。为了便于读者理解，文中所有漏斗图的横纵坐标均替换为中文，见图 3（见正文 P17）和图 4（见正文 P21）。

.....

审稿人 2 意见：

论文“双相障碍患者的风险决策偏好：来自三水平元分析的证据”具有重要的理论和现实意义；元分析程序规范；研究结果可靠，语言表达清晰。有以下几个问题供作者参考。

意见 1：建议作者不止报告论文篇数，也同时分类报告日常态度和行为、实验任务等类别中的具体研究的数量，以及具体的效应量。总体感觉日常态度和行为方面的研究总数有点少，在三水平元分析中可能会影响结果的稳定性。

回应：感谢您的意见。本次修改时，我们在结果中报告了不同测量方式的文献数量（见正文 P12），并在测量方式的调节作用时报告了各个测量方式的效应量大小（见正文 P18）。为了便于审稿，我们将相关文字粘贴至下方。

本次修改时，我们根据日常风险行为分类增加了对应的关键词进行补充搜索，并将文献的检索时间扩大至 2024 年 4 月 15 日。此次补充检索新纳入 10 篇日常态度和行为文献，但患者比健康个体更偏好风险寻求的结果并未发生改变，可见结果具有一定的稳健性。

——结果，正文第 12 页——

在风险决策测量方式上，有 48 篇文献采用行为实验任务，其中 23 篇采用 IGT 任务，8 篇采用 BART 任务，12 篇采用 CGT 任务，其余采用 RC tasks 任务。有 5 篇文献采用风险态度量表，报告被试对各个决策领域或总体的风险态度，以及有 22 篇文献报告被试真实的日常风险行为状况，包括物质使用、缺乏锻炼、风险性行为等方面。

——结果，正文第 18 页——

就测量特征而言，本研究并未发现测量方式显著的调节作用($p = .070$)。不论采用风险态度量表(Hedges' $g = 0.624, p < .001$)、日常风险行为(Hedges' $g = 0.312, p < .001$)还是行为实验任务(Hedges' $g = 0.252, p < .001$)的效应量都显著。其中，行为实验任务与风险态度量表的效应量间存在显著差异($\beta = 0.371, p = .022$)，其余效应量间差异不显著。因此，采用不同风险决策测量方式探究双相障碍与风险决策的关系时，研究结果方向一致，都表现为双相患者比健康个体更偏好风险寻求，但具体差异程度有所不同。风险态度量表效应量最大，而行为实验任务效应量最小。

意见 2：文中“图 5”的内容均是英文，应该用中文语言表示较为合适。

回应：感谢您的建议。我们将图 5 进行翻译（见正文 P29）。为了便于审稿，我们将图粘贴至下方。

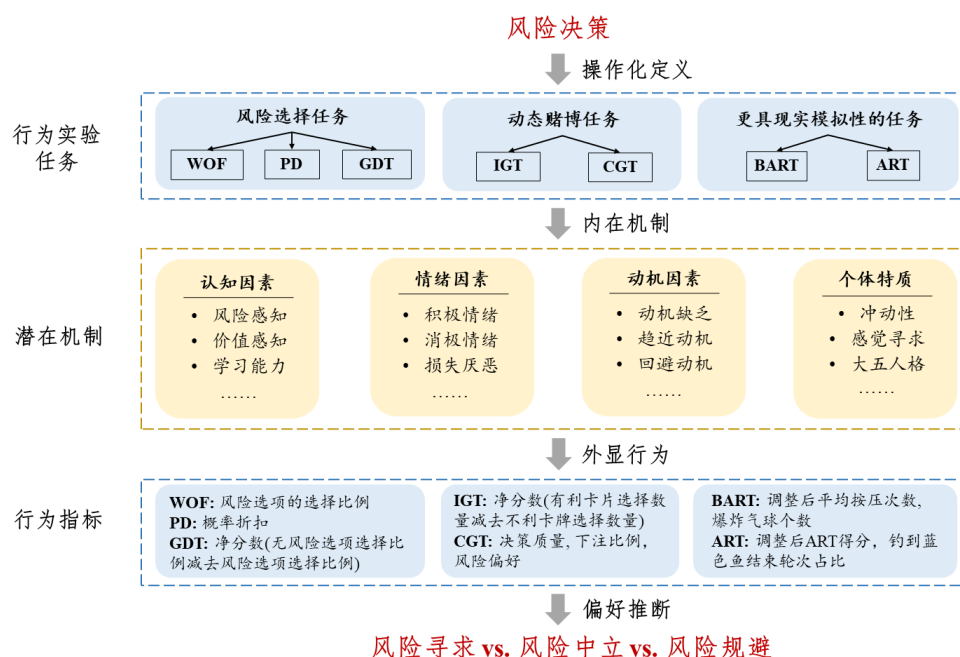


图 5 不同行为实验任务中反映不同风险偏好的潜在框架（摘自“因匿名审稿隐去该参考文献”）

注：WOF = 幸运轮盘任务；PD = 概率折扣任务；GDT = 骰子任务；IGT = 爱德华赌博任务；

CGT = 剑桥赌博任务；BART = 气球模拟风险任务；ART = 钓鱼风险任务。

意见 3：关于“地域”方面，建议作者有更深入的思考和讨论，例如不只是简单停留在地域层面，可以进一步考虑影响风险偏好行为的文化层面（如个体主义、集体主义；松紧文化等）、社会经济发展层面的因素（如经济不平等、社会流动水平等）。

回应：感谢您的建议。由于本次修改增加了新文献，地区调节作用发生了改变，仅在行为实验任务模型中显著。我们根据最新结果讨论了地区的作用，并在日常态度和行为模型结果的讨论中尝试结合了集体主义、松紧文化等文化层面的因素，具体见讨论部分（正文 P25–26）。为了便于审稿，我们将相关文字粘贴至下方。

——讨论，正文第 25–26 页——

本研究在总体模型中并未发现地区显著的调节作用，但无论地区如何，双相患者都表现出比健康个体更高风险寻求的趋势(欧洲: Hedges' $g = 0.333$; 北美洲: Hedges' $g = 0.189$; 南美洲: Hedges' $g = 0.415$; 亚洲: Hedges' $g = 0.220, p = .087$)。然而，细分测量方式后结果呈现差异。在行为实验任务中，地区的调节作用显著，其中欧洲和南美洲的效应量显著，且显著大于北美洲，而北美洲和亚洲的效应量不显著。对于亚洲研究，或许因研究数量不足(48 篇行为任务研究中仅 7 篇为亚洲研究)削弱了统计检验力(42.7%)，掩盖了亚洲患者与亚洲健康个体在风险偏好上的差异。随着未来亚洲研究的积累，其统计检验力有望提升，结果可能有所变化。而对于北美洲研究，尽管统计检验力也较低(8.18%)，但现有趋势表明，增

加研究可能不足以改变结果，因此我们推测这更可能与行为实验任务选用有关。不同地区选用行为实验任务类型有所差异($\chi^2(9, N = 48) = 25.40, p = .003$)。相较欧洲研究多采用 IGT 和 CGT 任务以及南美洲研究都采用 IGT 任务，北美洲研究选用的任务更为平衡。加之，不同行为实验任务的效应量不同。据此，北美洲效应量不显著或许与不同任务融合有关，但这也表明未来需积累不同地区的研究，并均衡选用行为实验任务，才能更准确分析地区的作用。

在日常态度和行为中，地区的调节作用并不显著，但欧美的效应量显著而亚洲的效应量不显著。亚洲研究数量不足可能是效应量不显著的原因之一，未来需积累更多研究再次检验这一调节作用。同时，文化差异也是值得关注的因素。目前研究多围绕风险性行为和物质滥用展开。基于文化松紧程度(Gelfand et al., 2011)，相较欧美国家，亚洲国家往往处于“紧”文化，通常拥有较严格的社会规范，对越轨行为的容忍度低、惩罚度高。那么风险性行为、物质滥用（尤其是大麻等毒品）自然在欧美文化中的接受度和常见度都比亚洲文化更高(Arria et al., 2017; Bragazzi et al., 2021; Jia et al., 2018; Kandel et al., 1981)。即便是双相患者也可能存在这一文化差异。另一方面，上述结果也可能和亚洲与欧美个体在风险偏好上的固有差异有关。亚洲个体通常比欧美个体更偏好风险寻求(Chen et al., 2020; Du et al., 2002)，这可能与集体主义文化有关。因为集体主义文化能给个体提供强大的社会支持，即使冒险失败个体也可能因他人的帮助而得到“缓冲”(Hsee & Weber, 1999)。那么，不同地区健康个体的基线差异可能导致亚洲患者与同一地区的健康个体间的差异变得不明显。未来需积累更多这方面的研究，更好地检验地区的作用，并探索文化差异与地区效应的关系。

参考文献

- Arria, A. M., Caldeira, K. M., Allen, H. K., Bugbee, B. A., Vincent, K. B., & O'Grady, K. E. (2017). Prevalence and incidence of drug use among college students: An 8-year longitudinal analysis. *The American Journal of Drug and Alcohol Abuse*, 43(6), 711–718. <https://doi.org/10.1080/00952990.2017.1310219>.
- Bragazzi, N. L., Beamish, D., Kong, J. D., & Wu, J. (2021). Illicit drug use in Canada and implications for suicidal behaviors, and household food insecurity: findings from a large, nationally representative survey. *International Journal of Environmental Research and Public Health*, 18(12), 6425. <https://doi.org/10.3390/ijerph18126425>.
- Chen, X. J., Ba, L., & Kwak, Y. (2020). Neurocognitive underpinnings of cross-cultural differences in risky decision making. *Social Cognitive and Affective Neuroscience*, 15(6), 671–680. <https://doi.org/10.1093/scan/nsaa078>
- Du, W., Green, L., & Myerson, J. (2002). Cross-cultural comparisons of discounting delayed and probabilistic rewards. *Psychological Record*, 52(4), 479–492. <https://doi.org/10.1007/BF03395199>
- Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliaich, A., Ang, S.,

- Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Subirats Ferrer, M., Fischlmayr, I. C., Fischer, R., ... Yamaguchi, S. (2011). *Differences between tight and loose cultures: a 33-nation study*. *Science*, 332(6033), 1100–1104. <https://doi.org/10.1126/science.1197754>
- Hsee, C., & Weber, E. U. (1999). Cross-national differences in risk preference and lay predictions. *Journal of Behavioral Decision Making*, 12(2), 165–179.
- Jia, Z., Jin, Y., Zhang, L., Wang, Z., & Lu, Z. (2018). Prevalence of drug use among students in mainland China: A systematic review and meta-analysis for 2003-2013. *Drug and Alcohol Dependence*, 186, 201–206. <https://doi.org/10.1016/j.drugalcdep.2017.12.047>
- Kandel, D. B., Adler, I., & Sudit, M. (1981). The epidemiology of adolescent drug use in France and Israel. *American Journal of Public Health*, 71(3), 256-265. <https://doi.org/10.2105/ajph.71.3.256>.

意见 4：一些参考文献格式问题，如“Biological psychiatry”，期刊首字母大写。

回应：感谢您的建议。我们对参考文献的格式进行了仔细核对，对您提及的期刊首字母大写等格式问题进行了修改。

第二轮

审稿人 2 意见：

作者比较好地对意见进行了回应和论文修改，无其他意见。

审稿人 3 意见：

本文借助三水平元分析，详细地考察了双相障碍患者的风险决策偏好，具有较强的研究意义与一定的创新性。但文章目前仍然存在以下问题，建议进一步修改完善。

意见 1：在引出本文的研究话题时，由于目前的逻辑是“双相障碍存在普遍性、严重性和复杂性，而探究其认知和行为特征，有助于有效识别、评估和干预”，那么接下来应当强调的是：为什么在双相障碍的识别和评估中，风险偏好这一特点尤为重要（即相较于其他特点，其重要性十分突出）？但目前的引言只是提到了双相患者可能会有更多的风险行为，并没有强调风险行为在双相患者众多行为表现中的重要性或特殊性。只有阐明了风险偏好的重要性或特殊性，才能进一步提升本文的研究价值。

回应：感谢您的宝贵建议。我们在引言部分增加了有关风险偏好的重要性和特殊性的文字阐述（见正文第 2 页）。为了便于审稿，我们将相关文字粘贴至下方。

——引言，第 2 页——

决策(decision-making)作为高级认知功能，又关乎个体日常行为，其如何受双相障碍的影响备受关注。有研究发现双相患者的决策功能受损，容易做出非理性、不适宜的决策

(e.g., Amlung et al., 2019; Ramírez-Martín et al., 2020), 进而影响其行为。双相患者的风险决策尤为重要, 它可能产生潜在的严重后果, 不仅关乎患者的社会、职业功能, 甚至可能危及身体健康和生命安全, 如自杀行为等(Miller & Black, 2020)。因此, 有必要深入了解患者的风险偏好, 及时识别并干预其风险行为。

风险决策(risky decision-making)指个体在风险或不确定情境中的决策。个体的风险决策偏好(简称风险偏好)常分为风险寻求和风险规避。风险寻求(risk-seeking)是指个体在面对两个及以上风险选项时选择高风险选项的倾向(Meertens & Lion, 2008; Lejuez et al., 2002)。与之相反的则是风险规避(risk-aversion)。双相障碍的诊断标准指出, 患者在躁狂或轻躁狂发作时容易过度参与那些很可能产生痛苦后果的高风险活动(APA, 2013), 这意味着双相患者常伴有风险决策异常。对患者的临床观察(Krantz et al., 2018; Williams et al., 2017)、行为实证研究(Edge et al., 2013; Richard-Devantoy et al., 2016)和神经影像学研究(Blankenstein et al., 2017; Miklowitz & Johnson, 2006)也表明双相障碍会影响患者的风险偏好。

参考文献

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing. <https://doi.org/10.1176/appi.books.9780890425596>.
- Amlung, M., Marsden, E., Holshausen, K., Morris, V., Patel, H., Vedelago, L., Naish, K. R., Reed, D. D., & McCabe, R. E. (2019). Delay Discounting as a Transdiagnostic Process in Psychiatric Disorders: A Meta-analysis. *JAMA Psychiatry*, 76(11), 1176–1186. <https://doi.org/10.1001/jamapsychiatry.2019.2102>
- Blankenstein, N. E., Peper, J. S., Crone, E. A., & van Duijvenvoorde, A. C. K. (2017). Neural Mechanisms Underlying Risk and Ambiguity Attitudes. *Journal of Cognitive Neuroscience*, 29(11), 1845–1859. https://doi.org/10.1162/jocn_a_01162
- Edge, M. D., Johnson, S. L., Ng, T., & Carver, C. S. (2013). Iowa gambling task performance in euthymic bipolar I disorder: A meta-analysis and empirical study. *Journal of Affective Disorders*, 150(1), 115–122. <https://doi.org/10.1016/j.jad.2012.11.027>
- Fischhoff, B., & Broomell, S. B. (2020). Judgment and decision making. *Annual Review of Psychology*, 71(1), 331–355. <https://doi.org/10.1146/annurev-psych-010419-050747>
- Krantz, M., Goldstein, T., Rooks, B., Merranko, J., Liao, F., Gill, M. K., Diler, R., Hafeman, D., Ryan, N., Goldstein, B., Yen, S., Hower, H., Hunt, J., Keller, M., Strober, M., Axelson, D., & Birmaher, B. (2018). Sexual Risk Behavior Among Youth With Bipolar Disorder: Identifying Demographic and Clinical Risk Factors. *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(2), 118–124. <https://doi.org/10.1016/j.jaac.2017.11.015>
- Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, 8(2), 75–84. <https://doi.org/10.1037/1076-898X.8.2.75>

- Meertens, R. M., & Lion, R. (2008). Measuring an individual's tendency to take risks: The risk propensity scale. *Journal of Applied Social Psychology*, 38(6), 1506–1520. <https://doi.org/10.1111/j.1559-1816.2008.00357.x>
- Miklowitz, D. J., & Johnson, S. L. (2006). The psychopathology and treatment of bipolar disorder. *Annual Review of Clinical Psychology*, 2, 199–235. <https://doi.org/10.1146/annurev.clinpsy.2.022305.095332>
- Miller, J. N., & Black, D. W. (2020). Bipolar Disorder and Suicide: a Review. *Current Psychiatry Reports*, 22(2), 6. <https://doi.org/10.1007/s11920-020-1130-0>
- Ramírez-Martín, A., Ramos-Martín, J., Mayoral-Cleries, F., Moreno-Küster, B., & Guzmán-Parra, J. (2020). Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. *Psychological Medicine*, 50(13), 2141–2153. <https://doi.org/10.1017/S0033291720003086>
- Richard-Devantoy, S., Olié, E., Guillaume, S., & Courtet, P. (2016). Decision-making in unipolar or bipolar suicide attempters. *Journal of Affective Disorders*, 190, 128–136. <https://doi.org/10.1016/j.jad.2015.10.001>
- Williams, S. C., Davey-Rothwell, M. A., Tobin, K. E., & Latkin, C. (2017). People Who Inject Drugs and Have Mood Disorders—A Brief Assessment of Health Risk Behaviors. *Substance Use and Misuse*, 52(9), 1181–1190. <https://doi.org/10.1080/10826084.2017.1302954>

意见 2：在分析“测量方式的调节作用”时有些问题。举例来说，在风险选择任务中偏好“高风险高收益选项”，与在日常生活中做出一些诸如“吸毒、酗酒等高风险行为”，不止是测量的方式差异。前者应当是中性的，因为我们不能认为在经济任务中偏好风险就是有错的；后者则是有害的，因为这些行为要么伤害身体健康，要么违规违法。因此将这种差异称为测量方式的调节作用有所不妥，建议作者对此进行进一步的分析。

回应：感谢您极具价值的意见和建议。在风险决策研究中，确实存在以不同风险偏好测量类型(risk preference measure categories, Bagañi et al., 2023)揭示个体风险倾向的情况。Frey 等人(2017)将风险决策测量类型分为风险态度量表、行为实验任务和日常风险行为 3 种。诚如您所言，这 3 种类型不仅在测量形式上不同，在测量内容上往往也有差异。本质上可以理解为研究者对风险偏好的操作定义不同。因此，关于个体冒险或是保守的结论通常是基于特定的操作定义和测量类型（包括测量形式和内容）得出的。这也可能是不同研究关于个体冒险性的结论不完全一致的原因之一。然而，人的风险行为确实体现在众多领域，不同领域风险行为的测量类型也有一定程度的异同。虽然人们默认不同测量类型所揭示的人的冒险性在本质上都是内在风险偏好，但相关结论是否具有跨测量类型的一致性仍是值得探究的问题。本研究关注双相患者在不同测量类型下的风险偏好是否稳定，试图得出更为稳健的结论，因此将上述 3 种风险偏好测量类型作为调节变量进行分析。在专家意见的启发下，我们认为“风险偏好测量类型”比之前的“测量方式”能更准确地体现研究意图。因此，我们在正文中将“测量方式的调节作用”修改为“风险偏好测量类型的调节作用”。同时，我们还在讨论部分进一步分析了探讨不同测量类型背后的潜在心理机制的必要性（见正文第

27 页)。为了便于审稿,我们将相关文字粘贴至下方。

——讨论,第 27 页——

本研究还考察了双相障碍与风险决策偏好关系是否因测量特征不同而有所差异。首先确实发现不同**风险偏好测量类型**间的结果不同,主要是风险态度量表和行为实验任务间的结果有显著差异。但稳定的是,不论在哪种测量**类型**中,双相患者均比健康个体更偏好风险寻求(风险态度量表: Hedges' $g = 0.624$; 行为实验任务: Hedges' $g = 0.252$; 日常风险行为: Hedges' $g = 0.312$), 仅仅在程度上有所差异。**这种程度差异或许不仅源于测量形式不同,还和测量内容有关。风险态度量表将风险偏好定义为一般性风险态度或对特定领域的风险态度,行为实验任务将风险偏好抽象为选择高风险高收益选项的倾向,而日常风险行为则将风险偏好具象为具体风险行为的频率。这些操作定义差异及其具体选取的指标差异都可能使得个体对备选选项或风险行为的感知发生变化(如,风险大小、后果性质等),最终影响其风险偏好。因此,探索双相患者在不同测量类型上表现出差异背后的原因,即探究双相患者风险寻求增加的潜在心理机制是未来研究的重要方向之一。**

参考文献

- Bagañi, A., Liu, Y., Kapoor, M., Son, G., Bürkner, P., Tisdall, L., & Mata, R. (2023). *Meta-Analyses of the Temporal Stability and Convergent Validity of Risk Preference Measures*. PsyArXiv. <https://doi.org/10.31234/osf.io/d7nuj>
- Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, 3(10), e1701381. <https://doi.org/10.1126/sciadv.1701381>

意见 3: 在讨论的第一部分,作者仅用三行半的文字交代了本文的核心结论,其余内容均在探讨双相患者拥有更高风险偏好的潜在原因。建议在此处缩短这部分内容的篇幅,因为考察潜在的心理机制并非本文的研究重点;但考虑到这一话题的重要性,建议将其放到本文的研究不足与对未来研究展望中。

回应: 感谢您的宝贵建议。我们在讨论的第一部分中精简了探讨双相患者有更高风险偏好潜在原因的内容(见讨论,正文 P24–P25),并在研究不足与未来研究展望中增加了关于潜在心理机制探讨的必要性(见讨论,正文 P30–P31)。为了便于审稿,我们将相关文字粘贴至下方。

——讨论,正文第 24–25 页——

研究发现,双相患者比健康个体更偏好风险寻求(Hedges' $g = 0.301$),该结果与 Edge 等人(2013)、Richard-Devantoy 等人(2016)以及 Ramírez-Martín 等人(2020)中 IGT 任务的元分

析结果一致，也与大多实证研究结果相同(e.g., Malloy-Diniz et al., 2011; Mason et al., 2014)，这回答了本研究第一个问题，即“双相患者与健康个体的风险偏好有无差异”。研究者基于模糊痕迹理论(Fuzzy-Trace Theory, FTT; Brainerd & Reyna, 1990)推测双相患者风险寻求增加可能有 2 种途径(Lukacs et al., 2021)：一是患者在强烈情绪影响下可能改变其对信息的加工方式(Rivers et al., 2008)，变得不再像健康个体那样倾向于要义加工(Gist processing)，而是过度依赖字面加工(Verbatim processing)。这可能使得他们对风险和价值的感知发生变化，最终改变其风险偏好(Reyna et al., 2015)；二是患者依然依赖要义加工，但他们可能赋予事件不恰当的要义，潜在创造一个“预加载反应”(Pre-load response)，使其未来面对同一决策时再次做出不恰当反应(如，躁狂期患者因亢奋情绪易赋予风险行为积极要义和“风险寻求”反应)。然而，目前仅 Sicilia 等(2020)发现双相患者的要义加工偏好下降，但尚无研究直接验证上述推论是否成立，因此仍需未来进一步研究来检验。

——讨论，正文第 30 页——

③ 鉴于目前双相障碍影响风险决策的研究侧重于揭示现象，鲜有探究现象背后的心理机制，因此本元分析着重梳理了主效应而未能探究其心理机制。

——讨论，正文第 31 页——

第四，未来研究应重视探究双相障碍影响风险决策的心理机制。目前双相障碍与风险决策间关系的研究多停留在现象揭示，直接探究现象背后心理机制的研究相对匮乏。有研究者基于 FTT 理论提出，双相患者在强烈情绪状态下既可能过度依赖字面加工，改变其风险感知或价值感知，进而影响其风险偏好；也可能赋予事件不恰当要义，使其未来面对同一决策时仍做出不恰当行为(Lukacs et al., 2021)。也有研究者认为，<轻>躁狂期患者因对行为趋近系统超敏继而比健康个体更偏好风险寻求(Katz et al., 2021)。还有研究者从动机强度(Hershenberg et al., 2016)、损失厌恶(Lasagna et al., 2022)等角度提出原因推测。但目前只有一些间接证据(e.g., Collett et al., 2016; Sicilia et al., 2020)而非直接证据支持上述推测。因此，未来研究有必要检验这些推测以揭示双相障碍影响风险决策的心理机制。这不仅有助于描绘患者风险决策的心理过程，还能为干预其风险决策提供理论依据和实践指导。

参考文献

- Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review, 10*(1), 3–47. [https://doi.org/10.1016/0273-2297\(90\)90003-M](https://doi.org/10.1016/0273-2297(90)90003-M)
- Collett, J. (2016). *It's not all about that bas: trait bipolar disorder vulnerability weakly correlated with trait bas and not predictive of risky decision-making*. (Doctoral dissertation). Swinburne University of Technology.

- Edge, M. D., Johnson, S. L., Ng, T., & Carver, C. S. (2013). Iowa gambling task performance in euthymic bipolar I disorder: A meta-analysis and empirical study. *Journal of Affective Disorders*, 150(1), 115–122. <https://doi.org/10.1016/j.jad.2012.11.027>
- Hershenberg, R., Satterthwaite, T. D., Daldal, A., Katchmar, N., Moore, T. M., Kable, J. W., & Wolf, D. H. (2016). Diminished effort on a progressive ratio task in both unipolar and bipolar depression. *Journal of Affective Disorders*, 196, 97–100. <https://doi.org/10.1016/j.jad.2016.02.003>
- Katz, B. A., Naftalovich, H., Matanky, K., & Yovel, I. (2021). The dual-system theory of bipolar spectrum disorders: A meta-analysis. *Clinical Psychology Review*, 83, 101945. <https://doi.org/10.1016/j.cpr.2020.101945>.
- Lasagna, C. A., Pleskac, T. J., Burton, C. Z., McInnis, M. G., Taylor, S. F., & Tso, I. F. (2022). Mathematical modeling of risk-taking in bipolar disorder: Evidence of reduced behavioral consistency, with altered loss aversion specific to those with history of substance use disorder. *Computational Psychiatry*, 6(1), 96–116. <https://doi.org/10.5334/cpsy.61>
- Lukacs, J. N., Sicilia, A. C., Jones, S., & Algorta, G. P. (2021). Interactions and implications of Fuzzy-Trace theory for risk taking behaviors in bipolar disorder. *Journal of Affective Disorders*, 293, 305–313. <https://doi.org/10.1016/j.jad.2021.06.035>
- Malloy-Diniz, L. F., Neves, F. S., de Moraes, P. H. P., De Marco, L. A., Romano-Silva, M. A., Krebs, M. O., & Corrêa, H. (2011). The 5-HTTLPR polymorphism, impulsivity and suicide behavior in euthymic bipolar patients. *Journal of Affective Disorders*, 133(1-2), 221–226. <https://doi.org/10.1016/j.jad.2011.03.051>.
- Mason, L., O'Sullivan, N., Montaldi, D., Bentall, R. P., & El-Deredy, W. (2014). Decision-making and trait impulsivity in bipolar disorder are associated with reduced prefrontal regulation of striatal reward valuation. *Brain*, 137(8), 2346–2355. <https://doi.org/10.1093/brain/awu152>.
- Ramírez-Martín, A., Ramos-Martín, J., Mayoral-Cleries, F., Moreno-Küster, B., & Guzmán-Parra, J. (2020). Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. *Psychological Medicine*, 50(13), 2141–2153. <https://doi.org/10.1017/S0033291720003086>
- Reyna, V. F., Weldon, R. B., & McCormick, M. (2015). Educating Intuition: Reducing Risky Decisions Using Fuzzy-Trace Theory. *Current Directions in Psychological Science*, 24(5), 392–398. <https://doi.org/10.1177/0963721415588081>
- Richard-Devantoy, S., Olié, E., Guillaume, S., & Courtet, P. (2016). Decision-making in unipolar or bipolar suicide attempters. *Journal of Affective Disorders*, 190, 128–136. <https://doi.org/10.1016/j.jad.2015.10.001>
- Rivers, S. E., Reyna, V. F., & Mills, B. (2008). Risk taking under the influence: A fuzzy-trace theory of emotion in adolescence. *Developmental Review*, 28, 107–144. <https://doi.org/10.1016/j.dr.2007.11.002>
- Sicilia, A. C., Lukacs, J. N., Jones, S., & Perez Algorta, G. (2020). Decision-making and risk in bipolar disorder: A quantitative study using fuzzy trace theory. *Psychology and Psychotherapy*, 93(1), 105–121. <https://doi.org/10.1111/papt.12215>

意见 4：关于审稿人 2 的意见 3，作者目前的回应还不到位。目前本文仍然是根据地区进行了异质性分析，这一做法可能会削弱本文的社会意义。是否可以选取“文化层面（如个体主义、集体主义；松紧文化等）、社会经济发展层面的因素（如经济不平等、社会流动水平等）”其中的 1~2 个因素，重新对样本进行分类。举个简单的例子：将目前这些研究涉及的国家划分为“集体主义和个人主义”或“文化松紧度高和低”，从而分析在以不同文化为主导

的国家，双相患者表现出的风险偏好是否存在差异。

回应：感谢您的宝贵建议。根据您的建议，我们从文化层面选择了集体主义/个体主义和松紧文化这 2 个因素，将现有研究涉及的国家划分为“集体主义和个体主义”或“文化紧度高和低”（集体主义和个体主义划分参考 <https://www.hofstede-insights.com/country-comparison-tool> 中的个体主义分数，低于 50 分的为集体主义国家；文化紧度分数参考 Gelfand 等人(2011) 的研究，其中大于所有国家均分（6.5）为高文化紧度国家即处于紧文化，反之为低文化紧度国家即处于松文化）。然后比较不同文化背景下的双相患者的风险偏好差异。**结果表明（见下表）**，无论双相患者是处于集体主义还是个体主义国家，亦或是处于文化紧度高还是低的国家，他们几乎都比健康个体更偏好风险寻求。具体地，集体主义国家的效应量除了在日常行为和态度模型中不显著外，其余都大于个体主义国家的效应量；而高文化紧度国家的效应量在行为实验任务模型中比低文化紧度国家的效应量大，但在日常态度和行为模型中却相反。我们在讨论中简要阐述了上述结果（见正文第 26 页），**为了便于审稿，我们将相关文字粘贴至下方（限于篇幅，相关数据表格放在附录中）。**

表 1 集体/个体主义和文化紧度对双相障碍与风险决策偏好关系的调节效应结果表

模型	调节变量	No. of ESs	Q_E (df)	F (df1, df2)	Hedges' g/β	SE	95% CI	t	p
总体模型	a. 集体/个体主义	176	-	-	-	-	-	-	-
	集体主义 C	36	Q_E (174) = 684.26, $p < .001$	F (1, 174) = 2.34, $p = .128$	0.426	0.092	[0.245, 0.607]	4.65	<.001
	个体主义 I	140			0.268	0.047	[0.176, 0.361]	5.73	<.001
	C vs. I	-			-0.157	0.103	[-0.360, 0.046]	-1.53	.128
	b. 文化紧度	142	-	-	-	-	-	-	-
	文化紧度低 L	75	Q_E (140) = 577.28, $p < .001$	F (1, 140) = 0.00 $p = .990$	0.315	0.064	[0.187, 0.442]	4.89	<.001
	文化紧度高 G	67			0.316	0.070	[0.178, 0.454]	4.54	<.001
	L vs. G	-			0.001	0.095	[-0.186, 0.189]	0.01	.990
行为实验 任务模型	a. 集体/个体主义	88	-	-	-	-	-	-	-
	集体主义 C	18	Q_E (88) = 252.72, $p < .001$	F (1, 86) = 6.90, $p = .010$	0.522	0.114	[0.295, 0.749]	4.58	<.001
	个体主义 I	70			0.178	0.064	[0.051, 0.306]	2.78	.007
	C vs. I	-			-0.344	0.131	[-0.604, -0.084]	-2.63	.010
	b. 文化紧度	71	-	-	-	-	-	-	-
	文化紧度低 L	41	Q_E (69) = 240.07, $p < .001$	F (1, 69) = 1.34, $p = .252$	0.220	0.089	[0.042, 0.398]	2.46	.016
	文化紧度高 G	30			0.373	0.098	[0.178, 0.567]	3.82	<.001
	L vs. G	-			0.153	0.132	[-0.111, 0.417]	1.16	.252
日常态度 和行为模型	a. 集体/个体主义	88	-	-	-	-	-	-	-
	集体主义 C	18	Q_E (86) = 352.98, $p < .001$	F (1, 86) = 0.83, $p = .364$	0.222	0.165	[-0.107, 0.550]	1.34	.183
	个体主义 I	70			0.386	0.070	[0.246, 0.526]	5.47	<.001
	C vs. I	-			0.164	0.180	[-0.193, 0.521]	0.91	.364
	b. 文化紧度	71	-	-	-	-	-	-	-
	文化紧度低 L	34	Q_E (69) = 309.86, $p < .001$	F (1, 69) = 2.09, $p = .153$	0.462	0.104	[0.254, 0.669]	4.44	<.001
	文化紧度高 G	37			0.242	0.111	[0.021, 0.463]	2.19	.032
	L vs. G	-			-0.219	0.152	[-0.523, 0.084]	-1.45	.153

注：集体主义国家（个体主义分数低于 50 分）和个体主义国家划分参考 <https://www.hofstede-insights.com/country-comparison-tool>；文化紧度分数参考 Gelfand 等人(2011)，其中大于所有国家均分（6.5）为高文化紧度国家即处于紧文化，反之为低文化紧度国家即处于松文化。

——讨论，第 26 页——

上述地区差异深层次上可能与文化差异有关。一方面，当前研究涉及的风险行为主要包括风险性行为、物质滥用等。基于文化松紧度(Gelfand et al., 2011)， 相较欧美国家，亚洲国家往往处于“紧”文化，通常拥有较严格的社会规范，对越轨行为的容忍度低、惩罚度高。因此欧美文化对风险性行为、物质滥用（尤其是大麻等毒品）等行为的接受度和常见度比亚洲文化更高(Arria et al., 2017; Bragazzi et al., 2021; Jia et al., 2018; Kandel et al., 1981)。即便是双相患者也可能存在这一文化差异。我们尝试根据文化紧度将现有研究所涉国家分为“高”和“低”2 种类型，然后比较 2 种文化下患者与健康个体的风险偏好差异是否不同。结

果表明（见附表 4），虽然不论处于“紧”文化还是“松”文化的患者均比健康个体更偏好风险寻求，但我们确实发现在日常态度和行为测量中，“松”文化的效应量(Hedges' $g = 0.462$)大于“紧”文化的效应量(Hedges' $g = 0.242$)，这初步验证了我们的推测，但仍需更多研究检验。另一方面，亚洲与欧美个体在风险偏好上可能存在固有差异。亚洲个体通常比欧美个体更偏好风险寻求(Chen et al., 2020; Du et al., 2002)，这可能与集体主义文化有关。因为集体主义文化能给个体提供强大的社会支持，即使冒险失败也可能因他人的帮助而得到“缓冲”(Hsee & Weber, 1999)。因此，不同地区健康个体的基线差异可能导致亚洲患者与同一地区健康个体间的差异变得不明显。我们将现有研究所涉国家分为“集体主义国家”和“个体主义国家”进行比较。结果发现（见附表 4），在日常态度和行为中，仅个体主义国家的效应量显著(Hedges' $g = 0.386$)，这与前述研究结果相符。但在行为实验任务中，却发现集体主义(Hedges' $g = 0.522$)和个体主义国家(Hedges' $g = 0.178$)的效应量均显著，且集体主义国家的效应量显著更大，这与前述研究结果不符，值得未来进一步探索。总之，未来需积累更多关于双相患者风险决策的地区和文化差异的研究，更深入探究文化差异与地区效应的关系。

参考文献

- Arria, A. M., Caldeira, K. M., Allen, H. K., Bugbee, B. A., Vincent, K. B., & O'Grady, K. E. (2017). Prevalence and incidence of drug use among college students: An 8-year longitudinal analysis. *The American Journal of Drug and Alcohol Abuse*, 43(6), 711–718. <https://doi.org/10.1080/00952990.2017.1310219>.
- Bragazzi, N. L., Beamish, D., Kong, J. D., & Wu, J. (2021). Illicit drug use in Canada and implications for suicidal behaviors, and household food insecurity: findings from a large, nationally representative survey. *International Journal of Environmental Research and Public Health*, 18(12), 6425. <https://doi.org/10.3390/ijerph18126425>.
- Chen, X. J., Ba, L., & Kwak, Y. (2020). Neurocognitive underpinnings of cross-cultural differences in risky decision making. *Social Cognitive and Affective Neuroscience*, 15(6), 671–680. <https://doi.org/10.1093/scan/nsaa078>
- Du, W., Green, L., & Myerson, J. (2002). Cross-cultural comparisons of discounting delayed and probabilistic rewards. *Psychological Record*, 52(4), 479–492. <https://doi.org/10.1007/BF03395199>
- Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliaich, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Subirats Ferrer, M., Fischlmayr, I. C., Fischer, R., ... Yamaguchi, S. (2011). *Differences between tight and loose cultures: a 33-nation study*. *Science*, 332(6033), 1100–1104. <https://doi.org/10.1126/science.1197754>
- Hsee, C., & Weber, E. U. (1999). Cross-national differences in risk preference and lay predictions. *Journal of Behavioral Decision Making*, 12(2), 165–179.
- Jia, Z., Jin, Y., Zhang, L., Wang, Z., & Lu, Z. (2018). Prevalence of drug use among students in mainland China: A systematic review and meta-analysis for 2003–2013. *Drug and Alcohol Dependence*, 186, 201–206. <https://doi.org/10.1016/j.drugalcdep.2017.12.047>

Kandel, D. B., Adler, I., & Sudit, M. (1981). The epidemiology of adolescent drug use in France and Israel. *American Journal of Public Health*, 71(3), 256-265. <https://doi.org/10.2105/ajph.71.3.256>.

第三轮

审稿人 3 意见：

作者较好地回应了上一轮中提出的意见，文章质量有较大提升，目前没有其他意见了。

编委意见：

感谢作者们对审稿人意见细致的回复。文章经过几轮修改后，各个方面都有了显著的提高，达到了发表的水平。

主编意见：

文章经过多轮评审和修改，已经达到学报发表要求，同意发表。感谢审稿专家和作者们的辛勤付出！