

《心理学报》审稿意见与作者回应

题目：大学生自伤行为的强化敏感性基础

作者：应梦婷 江光荣 于丽霞 鲁婷

第一轮

首先在此感谢各位审稿专家认真严谨的审稿，为我们的稿件提供了许多宝贵建议，并指出了其中存在的问题与不足，对我们进一步改进稿件提供了很大帮助。下面是我们针对各位审稿专家的意见作出的回复以及修改情况说明，同时，在稿件正文中也用蓝色对相应修改内容进行了标记。

审稿人 1 意见：

文章写作规范，概念准确，有一定的创新，行文清晰简明，使用的范式具科学性，对无关变量的控制也很细致。但是本文仍有一些值得商榷的地方。建议修改后再审。

意见 1： 研究一没有报告描述性结果，不知道两个研究的被试是怎样的被试，也会影响对文章结果的理解，比如：研究二的对照组与实验组之间在其他人口学变量上是否有差异？

回应： 感谢审稿专家的意见。已补充研究一的描述性统计结果，并对研究二的被试组成作了更细致的说明。见 2.2.2 描述统计结果和 3.1.1（2）被试的组成，蓝色标记部分。

意见 2： 研究二的入组标准中，有“入组的自伤被试必须同时满足 2 个条件：（1）自伤问卷总分大于 9；或自伤问卷总分大于 6 且至少有一种自伤行为的发生次数在 5 次及以上”，为何制定这样的标准？这一标准是否有文献支持？

回应： 针对审稿专家的意见，对此标准进行说明和修改完善如下：

本研究所用自伤问卷是国内研究者在质性研究的基础上参考国外自伤量表编制而成，目前以此问卷为研究工具的研究表明，该自伤问卷上得分越高，情绪表达不能、情绪调节困难、冲动性等核心病理特征的程度越严重（冯玉，2008；于丽霞，2013；于丽霞等，2013）。然而，目前尚缺乏对问卷划界分的相关界定，且由于是中国版问卷，缺乏国外研究证据的支持。本研究起初制定筛查标准时，主要是结合 DSM-5 对自伤行为的界定，采用经验划分法选取的划界点。DSM-5 中指出，自伤行为是在过去一年内有 5 天或以上出现自伤行为；结合前人研究结果，自伤问卷得分越高，其核心病理特征越严重，病理性程度越高，因此我们针对本问卷取 9 分（相当于自伤频次 ≥ 5 次、伤害程度重度以上）作为人为制定的经验标准，以尽量保证入组者自伤行为的典型性。然而，经您的提醒，我们也发现这样的标准难免有些主观，缺乏研究支持。因此，在本次修改过程中，研究者根据前人研究结果，重新制定了自伤组的入组标准，将条件之一的自伤总分 9 分标准进行了适当提高。因为最近以此问卷为研究工具的研究表明，问卷得分 10 分以上的自伤者在情绪调节困难和冲动性等核心病理特征上均显著高于 10 分以下的自伤者及无自伤者（于丽霞等，2013）。因此，为增加区分效应，选取自伤得分在 10 分以上者作为自伤组被试。

另一方面，由于部分自伤被试的自伤方式比较单一，仅在一、两类自伤行为上发生频次

≥5 次，这样，尽管问卷总分未达到 10 分，其病理程度并非一定低于 10 分以上的被试。事实上，重复自伤本就反映了自伤的病理性程度（Klonsky & Olin, 2008; 于丽霞, 2013），≥5 次也只是其自伤频次的一个最低值，实际自伤发生情况可能远高于分数代表的意义。参考于丽霞（2013）的研究结果可知，青春期后、进入大学阶段还存在高频次的重复自伤行为指示其自伤行为并非由发展阶段所致，而更可能是一种病理性自伤的表现。因此，我们将自伤问卷总分大于 6 分（相当于自伤频次≥5 次、伤害程度中度以上）且至少有一种自伤行为的发生次数在 5 次及以上的被试也纳入自伤组。这不仅符合 DSM-5 对自伤行为的界定标准，也突出了长期重复自伤者的自伤频次特征。

此外，在本研究一的实施过程中，研究者还收集了另一自伤特征变量“体验回避”的相关数据。本次修改过程中，我们对重新制定标准后不符合要求的被试进行剔除。并利用研究一的数据对分组标准进行检验：将研究一的 717 名被试分为三组，即①满足研究二自伤组标准的自伤者；②未达到入组标准的自伤者；③无自伤行为者。进行单因素 ANOVA，结果表明三组被试在体验回避的得分上差异显著， $F_{(2,714)}=42.65, p<0.001$ 。

最终，完善标准并增补被试后，研究二 87 名实验组被试中，符合 10 分标准的占 75.86%（66/87），符合 6 分及频次标准的占 24.14%（21/87）。针对此部分，文中已作出相应修改，见 3.1.1（1）被试的筛查与分组，蓝色标记部分。

意见 3：研究二中这样写道：“归入自我管理组或刺激寻求组的被试必须满足以下 2 个条件之一：（1）在某一功能维度上得分前 50%并在另一维度上得分后 50%，且在自我管理和刺激寻求两个维度上共同负荷较低；（2）在某一功能维度上得分为低分 50%但仅在这一维度上有得分，另一维度得分为 0，也就是说，其所有自伤行为仅仅出于一种目的。对于部分较难分组的被试，在实验开始前结合简短的访谈进行确认。”请问如何操作“自我管理和刺激寻求两个维度上共同负荷较低”？如何进行简短访谈而确认“部分较难分组的被试”？这些含糊的表达使得研究的可重复性大打折扣。是否需要交代：有多少符合（1），多少符合（2），又有多少是访谈确认的。

回应：感谢审稿专家的意见，您的意见帮助我们在表述上做到更严谨、明确，以利于读者的理解。事实上，在条件（1）中存在这样一部分被试，他们在某一功能维度上得到在高分 50%，在另一功能维度上得分为 0。这部分被试其实可以与条件（2）合并，作为功能单一的自伤被试，即仅在某一功能维度上有得分，其所有自伤行为仅仅出于一种目的，在本研究中，这类被试占实验组的 86.21%（75/87）。

另一类被试则是在两类功能维度上均有得分，但存在明显的功能倾向。这类被试在原稿件中的操作性定义为：在某一功能维度上得分前 50%而另一维度上得分后 50%，且两个维度上共同负荷较低。实际上，所谓“共同负荷低”旨在保证两个维度的百分位数有一定距离，避免 51 百分位和 49 百分位的情况出现。但经您的提醒，我们也发现这样的表述不够规范。在本次修改稿中，为增加两类自伤被试的区分效应，我们对这类混合功能被试的分组标准进行了提高，选取高分 27%和低分 27%作为极端组划分标准，重新对被试进行了筛查，对不符合新标准的被试数据进行剔除。对于所有混合功能的自伤被试，均在实验开始前由主试通过简短的半结构式访谈对其自伤功能类型进行再次确认。访谈问题主要如下：（1）你近一年内发生的自伤行为，主要动机是什么；（2）你最近一次采取自伤行为的动机是什么；（3）过去的自伤行为主要给你带来什么功能。以此为筛查标准对这部分被试做进一步分组。在本研究中，这类被试共占实验组的 13.79%（12/87）。

针对您的意见，本部分已在表述上进行修改，见 3.1.1 (1) 被试的筛查与分组，蓝色标记部分。

意见 4: 研究二：(1) 作为一个组间比较，样本量较小，且观测值标准差较大。是否需要采用相关统计方法提前计算所需样本量？(2) 文中“根据多元方差分析的数据要求和统计前提，实验数据经检验满足多元方差分析的使用条件。”是否需要报告对独立假设、正态假设、方差和协方差齐性假设的检验结果？

回应: 感谢审稿专家的意见。(1) 原样本量较小的确是本研究的不足，本次已在修改时限内采用同样的问卷、同样的程序、同样的主试重新筛查补充被试，现三组被试共 127 人(N1=51, N2=36, N3=40)。使用 G Power 软件对研究的统计检验力进行后效检验，Power 接近 1。尽管按中等效应量 $f=0.25$ 对所需样本量进行事先估计，显示总被试量达到 159 人更为适宜，但考虑到本研究所用自伤被试属于特殊样本，加之本次增加样本后研究结果的相对稳定性，以及较好的统计检验力。我们有理由认为本次修改后的研究结果是科学可靠的。增补被试后的实验结果，见稿件 3.2 实验结果，蓝色标记部分。

(2) 经您的提醒，我们已在文中补充报告多元方差分析各假设条件的检验结果。见 3.2.5 行为实验结果，蓝色标记部分。

意见 5: 表 3 的表头为“自伤行为对强化敏感性各维度的回归结果”，自伤行为作为因变量，这样的表述是否准确？

回应: 对于该表头，我们认为此处的表述应该是准确的。一般而言，我们说自变量 X 对因变量 Y 的预测作用，或自变量 X 对因变量 Y 的影响，但若以“回归”进行表述时，应是因变量 Y 对自变量 X 的回归，此处参考郭志刚主编的《社会统计分析方法》第一版，如 p31 “用粗死亡率 y 对老年人口比例 x_1 和人均国民生产总值 x_2 做回归”，p57 “妇女曾生子女数对年龄的回归”，p71 “食品支出对人均收入和价格回归”。此外，本刊已发表文章中也有类似的表述方式，如于丽霞等人 (2013) 发表在《心理学报》2013 年第 3 期的文章中表 5 的表头为“自伤行为对情绪调节困难、冲动性人格的 logistic 回归分析”。综上，我们认为这样的表述是符合规范的。

意见 6: 该研究没有阐述该研究存在的不足，实际上存在样本量较小，观测值标准差较大，BIS/BAS 的内部一致性不理想以及文章中的一些细节等不足。

回应: 感谢审稿专家的意见，已在文章讨论部分增加“研究局限与未来展望”，对研究中存在的不足、改进建议和未来研究方向进行了补充和说明。见稿件 4.4 研究局限与未来展望。

相关参考文献:

Bijttebier, P., Beck, I., Claes, L., & Vandereycken, W. (2009). Gray's reinforcement sensitivity theory as a framework for research on personality–psychopathology associations. *Clinical Psychology Review*, 29(5), 421–430.

Klonsky, E. D. & Olino, T. M. (2008). Identifying clinically distinct subgroups of self-injurers among young adults: A latent class analysis. *Journal of Consulting and Clinical Psychology*, 76(1), 22-27.

- 冯玉. (2008). *青少年自我伤害行为与个体情绪因素和家庭环境因素的关系*(硕士学位论文). 华中师范大学, 武汉.
- 郭志刚主编. (1999). *社会统计分析方法: SPSS 软件应用*. 北京: 中国人民大学出版社.
- 于丽霞. (2013). *一样自伤两样人: 自伤青少年的分类研究*(博士学位论文). 华中师范大学, 武汉.
- 于丽霞, 凌霄, 江光荣. (2013). 自伤青少年的冲动性. *心理学报*, 45(3), 320 - 335.
-

审稿人 2 意见:

该文选题有一定新意, 行文规范, 为研究自伤提供新的视角, 有一定的理论和实践意义。但是, 该论文在理论概念和实验过程(实验二)存在明显缺陷, 影响了该研究结果的科学性。

意见 1: 关键变量“自伤”界定不明确。正文中的定义为非自伤自伤(NSSI), 英文摘要中为蓄意自伤(DSH), 两者的内涵和外延有一定的区别, 测量方式也会有区别; 问卷中并无与自杀意图相关的内容, 那么本文中所指的自伤究竟是什么, 文中对被试的筛选与分组是否值得推敲?

回应: 感谢审稿专家的意见。本文所指自伤是沿用非自杀性自伤(NSSI)的术语。经您的提醒, 我们也发现原稿件在中英术语及定义上存在一些不妥, 需进一步明确以利于读者的理解。

首先, 目前学术界对自伤的术语使用一直未达成共识, 自伤(self-harm, SH)、蓄意自伤(deliberate self-harm, DSH)、非自杀性自伤(Non-Suicidal Self-Injury, NSSI)、自残(self-mutilation, SM)等相关术语多达 33 种(江光荣等, 2011)。这些术语指代的行为范围不尽相同, 它们在不同程度上被研究者所使用。其中, NSSI 的使用频率最高, 且为一些专业组织机构(如“国际自伤研究学会”, International Society for the Study Self-Injury, ISSS)所采用(江光荣等, 2011), 本研究中的自伤亦沿用此术语。本次修改已针对您的意见, 在文中进行补充说明, 并且在定义上, 按业界对非自杀性自伤的默认共识进行了完善。见 1.2 强化敏感性与自伤行为, 蓝色标记部分。

其次, 稿件正文与英文标题、摘要中所用术语不一致, 的确是我们的疏忽。起初撰写英文摘要时, 考虑到“自伤者”的英文表述用 self-harmer 似更通顺, 故将自伤的术语也写作 deliberate self harm, 却忽视了两个概念之间的本质区别。对此, 我们已对英文题名及摘要作出修改, 见稿件英文摘要, 蓝色标记部分。

最后, 关于研究所用自伤问卷问题, 经我们检查发现, 原稿件中所附问卷版本有误。我们在实际研究中特意在指导语部分强调了“在没有自杀动机的情况下”, 因此基于该问卷的被试筛选与分组也都是在排除自杀意图的前提下根据被试自伤行为的水平及其功能进行的。相应修改见附录 2 青少年自我伤害问卷, 蓝色标记部分。

意见 2: 文献较为陈旧, 需补充近几年的研究;

回应: 感谢审稿人提供的意见。已补充近 5 年的最新参考文献, 见稿件参考文献, 蓝色标记部分。

意见 3: 取样方面: (1) 在研究一中, 取样过程过于随意, 样本的人口学资料未见报告,

不能判断是否存在偏差；根据文献报道，大学生自伤（NSSI）的报告率差异较大，但通常为 10% 以上，本文中未见相关结果，因研究二的取样基于研究一的筛查，因此考虑是否有必要提供相关结果。

回应：感谢审稿专家的意见。研究一采用的是分层整群取样的方法，在武汉市 5 所综合性普通高等院校，对大一至大四 4 个年级的学生以班级为单位随机抽样。相关描述性统计结果已在稿件中进行补充，见 2.2.2 描述统计结果，蓝色标记部分。

(2) 研究二未对分组被试进行操作性定义（如，何为自我管理组）；

回应：感谢审稿专家的意见。已在研究二部分对三组被试的操作性定义进行补充，并对分组标准作了进一步详细说明。并且，经您的提醒，我们发现原稿件在自伤功能问卷所用项目上未交待清楚。原自伤功能问卷的自我管理维度除包括情绪管理的项目外，还有部分涉及日常事务管理或与他人相关的项目，而这部分功能并非我们想研究的、对应“个人负强化”的功能类型。从内容上看，只有情绪管理的 5 个条目对应我们希望研究的个人负强化功能，同时无论从 DSM-5 还是前人的大量研究看，它也是最主要的自伤功能。原研究者也发现，自伤被试更多地通过自伤达到自我管理中情绪管理的目的，得分最高的五种自伤原因中四种与情绪有关（郑莹，2006）。研究过程中，为了保证问卷两个维度的结构，我们发放的问卷中仍包含了自我管理和刺激寻求两个维度的全部项目（附录中即发放版问卷），但实际对自伤被试进行分组时，只将“情绪管理”部分的条目得分纳入统计，以保证入组的是以情绪管理（即管理负性情绪）为行为目的的自伤者，最大限度地保证其“个人负强化”的功能，与对应“个人正强化”功能的刺激寻求组达到区分效果。然而，这部分在原稿件中未作交待，容易造成读者误解。本次修改已在文中进行补充说明。见 3.1.1（1）被试的筛查与分组，蓝色标记部分。

(3) 样本量远远达不到统计学要求。作者采用单变量 ANOVA 方法，对 3 组被试的 Q-任务得分差异进行了统计学分析，但是 3 组被试人数只有 52 人（N1=21, N2=15, N3=16），我用 G Power 分析软件对其统计效力进行了检验，其 Power 为 0.32， $\beta=0.68$ ，假阳性率 $>60\%$ ，这一结果完全没有科学性。按作者单变量 ANOVA 统计方法，效果量确定为中等程度， $f=0.25, Power=0.8, Groups=3$ ，我进行了样本量估计，结果总样本应大于 159 人，平均每组至少应达 53 人。

回应：感谢审稿专家的意见。研究二的样本量较小的确是本研究的不足之处，我们已进行被试的增补。但我们认为您在原研究结果的统计检验力进行后效检验时存在偏颇。我们在研究二的 MANOVA 和 ANOVA 中报告的效应量均是 η^2 ，因为 η^2 的含义是“实验处理之后各組间平方和在总体平方和中所占的比重”，用此作为方差分析后效果大小的指标意义比较好理解，同时这一指标也为 SPSS 所采用（胡竹菁，戴海琦，2011；郑昊敏，温忠麟，吴艳，2011）。然而，在对方差分析的统计检验力进行检验时，应先将 η^2 通过公式 $f = (\eta^2 / (1 - \eta^2))^{1/2}$ 换算成效果大小的指标 f ，这个 f 也是国外研究最常用的 effect size，其值小于 0.1 为小的效应量，0.25 左右为中等，大于 0.4 为大的效应量；然后再通过进一步换算成 Φ 值查表（胡竹菁，戴海琦，2011）或直接用 G Power 软件来评估其统计检验力的大小。G Power 在对方差分析的统计检验力进行计算检验时正是使用“effect size f ”，而非 η^2 ， f 值则可在 determine 里计算得出。回到本研究中，原 Q-任务和 CARROT 的单因素 ANOVA 得出的 η^2 分别为 0.194 和 0.229，换算成效应量指标 f 分别是 0.491 和 0.545，均属于大的效应量，进一步对统计检验

力进行检验，得出 Power 分别为 0.88 和 0.94，均在一般标准 0.8 以上。

另一方面，国外有学者对后验的检验效能提出质疑，认为后验统计检验力并不可靠（Carstensen, 2002），不提倡使用。并且正如您所指出的，按单因素 ANOVA 统计方法，效果量确定为中等程度， $f=0.25$ ， $Power=0.8$ ， $Groups=3$ 对样本量进行事先估计，总样本量应达到 159 人为宜。并且，研究二原本在样本上也存在一些不足，如两类功能被试的交叉问题，Q 任务上两类自伤被试的差异边缘显著，即便减去最后 5 个 Q 刺激，显著性改善也不大。鉴于以上考虑，我们仍在本次修改时限内采用同样的问卷、同样的程序、同样的主试尽力增补了被试。并在原基础上，对研究二的筛查和分组标准作了一定改进，对不符合新标准的原被试数据进行了剔除。最终，三组被试共 127 人（ $N_1=51$ ， $N_2=36$ ， $N_3=40$ ）。补充数据后，多元方差分析的 η^2 为 0.258， $f=0.590$ ；Q-任务的单因素 ANOVA 的 η^2 为 0.271， $f=0.610$ ；CARROT 的单因素 ANOVA 的 η^2 为 0.260， $f=0.593$ 。采用 G Power 对统计效力进行检验，三者的 Power 均大于 0.99，接近 1。

综上所述，尽管使用软件按 $f=0.25$ 的中等效应量计算，总被试量达到 159 人更为适宜，但考虑到本研究所用自伤被试属于特殊样本，加之本次增加样本后研究结果的相对稳定性，以及较好的统计检验力。我们有理由认为本次修改后的研究结果是科学可靠的。已在文中补充报告效应量 f 值。修改后的实验结果见 3.2 实验结果，蓝色标记部分。

意见 4:（1）统计方法表述过于简单。（2）回归分析表述不规范，（3）为何未做单因素线性回归分析？

回应: 感谢审稿专家提供的宝贵意见。（1）已对统计方法的表述进行补充说明，见 2.1.4 和 3.1.5，蓝色标记部分。

（2）关于回归分析的表述，我们认为文中表头的表述方式应该是准确的，即“自伤行为对强化敏感性各维度的回归结果”。一般而言，我们说自变量 X 对因变量 Y 的预测作用，或自变量 X 对因变量 Y 的影响，但若以“回归”进行表述时，应是因变量 Y 对自变量 X 的回归，此处参考郭志刚主编的《社会统计分析方法》第一版，如 p31 “用粗死亡率 y 对老年人口比例 x_1 和人均国民生产总值 x_2 做回归”，p57 “妇女曾生子女数对年龄的回归”，p71 “食品支出对人均收入和价格回归”。此外，本刊已发表文章中也有类似的表述方式，如于丽霞等人（2013）发表在《心理学报》2013 年第 3 期的文章中表 5 的表头为“自伤行为对情绪调节困难、冲动性人格的 logistic 回归分析”。综上，我们认为这样的表述是符合规范的。

（3）在回归分析方法的选择上，我们做了较大的修改。下面先对您的意见进行回复：我们之所以选择多因素回归，而未做单因素回归，首先也是最重要的，是因为惩罚敏感性与奖励敏感性是通过联合作用共同影响个体行为的（Corr, 2002）。对于自伤行为同样如此，自伤行为的发生受到环境中多种惩罚或奖励刺激的影响，与强化敏感性各维度相联系，因此将多个变量的最优组合一起放入回归方程考察其对自伤行为的共同影响，比用单变量进行预测更有效，也更符合实际。从研究目的上来讲，我们希望找出对自伤行为有预测作用的强化敏感性因素，即最优组合，以最大限度地预测在哪几种因素存在时，自伤的风险最大，所以采用多因素回归的方式。这样也避免出现单因素回归时某个自变量的回归系数受另一自变量影响的情况，可以得到各自变量对自伤行为的净作用。其次，我们在做多元回归分析之前已经进行了各变量与自伤行为的相关分析，而单因素回归系数与相关系数本就存在一定联系， $b=r*s_y/s_x$ ，可见 b 与 r 呈正相关。结合前述研究目的上的考虑，我们认为可以不必再单独做单因素回归分析。

在回归分析的方法上，本次修改过程中我们发现此前的多元线性回归分析存在不妥，由于自伤问卷上得分为 0 的被试较多，分数呈偏态分布，并不适合进行线性回归。因此，我们已对数据分析方法做出修改，改为有序 Logistic 回归分析（Ordinal）方法。首先将自伤水平转换为分类变量，按前人研究证实有效的分类标准将自伤总分划分为 0 分、1 分、2-10 分和 10 分以上四个等级（于丽霞等，2013），然后以自伤等级为因变量，强化敏感性各维度为自变量进行 Logistic 回归分析，研究结论与前保持稳定。修改后的回归分析见 2.2.3 强化敏感性对自伤行为的预测作用，蓝色标记部分。

意见 5：部分结果缺少基本的统计数据（平均数、标准差）。且部分对结果的解释内容与讨论部分混杂。

回应：感谢审稿专家的意见。已对研究一和研究二的部分描述性统计结果进行补充，并对讨论部分做出修改。具体见 2.2.2 描述统计结果，3.1.1（2）被试的组成和 4 讨论，蓝色标记部分。

意见 6：（1）讨论的第一部分是否存在过度推测，或偏离主题现象；（2）第二部分对本文主题有喧宾夺主之嫌，作者本意可能是想表述从强化敏感性角度对自伤进行分类很有必要，但文中表述的观点范围过于宽泛，与主题不符。

回应：感谢审稿专家的意见。针对您的意见，我们分别回复如下：

（1）由于本研究是对强化敏感性与自伤关系的探索性研究，此前并无研究对两者关系进行直接探讨。因此，对于研究一得出强化敏感性对自伤行为的预测作用，研究者试图在其他间接证据的基础上对其作出合理的解释。从现有文献来看，情绪管理障碍、应对技能等都可能两者关系中起到中介作用，但就目前阶段而言，我们只能基于已经明确的证据进行分析讨论，并在此基础上尝试对两者的联系进行解释，而对于强化敏感性影响自伤行为的具体机制，我们还不能给出明确的解释，这也正是未来的研究方向。然而，由于我们在讨论部分对参考文献的细化工作有所欠缺，部分研究证据未详细标注出处，可能给读者带来疑问和误解。对此，我们参考您的意见，在修改稿中对讨论部分的研究证据进行明确细化，对部分主观性表述进行删改，并补充了一些新的研究证据，以提高分析讨论的可靠性。此外，我们也增加了未来研究展望的部分。具体见稿件 4.1 和 4.4，蓝色标记部分。

（2）研究二是在研究一所得结果的基础上，参考与自伤联系密切的相关领域研究证据（Bijttebier et al., 2009），希望进一步探讨不同功能自伤行为的强化敏感性基础。两个研究在逻辑上层层递进，逐渐深入，且始终围绕自伤者的强化敏感性基础这一核心主题。从研究结果上看，研究二也的确验证了假设，发现不同功能的自伤行为存在不同的强化敏感性模式。从某种意义上说，研究二对目前自伤研究的意义是可以与研究一齐平，甚至超越研究一的，因为它对目前研究中存在的问题及未来研究的方式提供了一些启示。在研究二表明不同功能类型的自伤行为存在本质差异的情况下，提示未来应考虑对自伤行为进行分类研究，我们认为这样的思考是有意义，也是合情合理的。但与此同时，在您的提醒下，我们也意识到，或许因为这一研究的发现令我们太兴奋，以致在讨论的表述上的确存在一些不当之处，尽管我们实事求是地反映了研究二的结果，但对其讨论不够。对此我们已进行修改，力求更谨慎、客观地表达我们的观点。具体见稿件 4.2，蓝色标记部分。

相关参考文献:

- Bijttebier, P., Beck, I., Claes, L., & Vandereycken, W. (2009). Gray's reinforcement sensitivity theory as a framework for research on personality–psychopathology associations. *Clinical Psychology Review*, 29(5), 421–430.
- Carstensen. (2002). Post-hoc power: Don't do it. Retrieved March 2, 2008, from <http://www.biostat.ku.dk/~bxc/SDC-courses>
- Corr, P. J.(2002). J. A. Gray's reinforcement sensitivity theory: tests of the joint subsystems hypothesis of anxiety and impulsivity. *Personality and Individual Differences*, 33, 511–532.
- 郭志刚主编. (1999). *社会统计分析方法: SPSS 软件应用*. 北京: 中国人民大学出版社.
- 胡竹菁, 戴海琦. (2011). 方差分析的统计检验力和效果大小的常用方法比较. *心理学探新*, 31(3), 254–259.
- 江光荣, 于丽霞, 郑莺, 冯玉, 凌霄. (2011). 自伤行为研究: 现状, 问题与建议. *心理科学进展*, 19(6), 861–873.
- 于丽霞, 凌霄, 江光荣. (2013). 自伤青少年的冲动性. *心理学报*, 45(3), 320–335.
- 郑吴敏, 温忠麟, 吴艳. (2011). 心理学常用效应量的选用与分析. *心理科学进展*, 19(12), 1868–1878.
- 郑莺. (2006). *武汉市中学生自我伤害行为流行病学调查及其功能模型*(硕士学位论文). 华中师范大学, 武汉.

审稿人 3 意见:

《大学生自伤行为的强化敏感性基础》通过两个研究发现, 强化敏感性可以有效预测自伤行为, 且自我管理和刺激寻求两种自伤行为有不同的强化敏感性基础。本研究对理解强化敏感性在大学生自伤行为发生和发展中的作用具有开拓性意义, 研究设计和研究范式规范, 研究结果可靠, 尚可完善之处如下:

意见 1: 个别专业词的缩写虽在专业内很常见, 建议英文简称第一次出现时, 还需注明中英文全称, 如: 边缘性人格障碍 (Borderline Personality Disorder, BPD)。

回应: 感谢审稿专家的建议。已在文中进行检查与补充。

意见 2: 个别处参考文献引用不详, 建议补上, 尽量将文献出处细化, 并引用一手文献, 如: BAS 对欲求刺激进行反应, 伴随产生积极情绪体验, 促进正在进行的行为; BIS 和 FFFS 对厌恶刺激及目标冲突进行反应, 伴随产生焦虑、恐惧等消极情绪, 导致抑制或逃避行为。根据 RST 的最新观点, 惩罚敏感性体现了 BIS 和 FFFS 两个系统的反应性, 奖励敏感性体现了 BAS 的反应性。(没标出处)

回应: 感谢审稿专家的意见。此处已补充文献出处, 见 1.1 强化敏感性概述, 蓝色标记部分。并再次进行通篇阅读, 仔细检查参考文献引用及标注情况, 将文献出处细化。

意见 3: 自伤行为的操作性定义的具体用词和文献出处还有待精准。例如: 文章中有时使用“自我伤害行为”, 有时用“自伤行为”, 建议统一使用“自伤行为”; 文章中对自伤行为的界定是“自我伤害行为是一种复杂的心理病理行为, 表现为个体在无自杀意图的前提下, 故

意且直接地伤害自己的身体组织,如用刀割伤、以头撞墙、阻止伤口愈合等,具体的伤害形式可达数十种(江光荣等,2011)。”。核对应文献时,发现原文献中对自伤行为的两处描述都与此处描述有些出入,一处为“自我伤害是一种复杂且危险的心理病理行为,表现为个体蓄意以不同方式破坏自己的身体组织,如用利器割伤/划伤、打火机烧伤、以头撞墙等,具体伤害形式可能超过十多种(Ross & Heath, 2002; Whitlock, Eckenrode, & Silverman,2006; 郑莺,2006)。”。另一处是“在定义上,业界基本将自伤默认为:在没有明确自杀意图的情况下,个体故意、重复地改变或伤害自己的身体组织,这种行为不为社会所认可,且不具致死性或致死性较低。”。建议使用本领域公认的或本课题组认可的自伤行为的操作性定义,并准确注明文献出处。

回应:感谢审稿专家提供的建议,帮助我们在术语及定义上做到更准确、规范,更有利于读者的理解。有关原文献中的两处定义,前者“自我伤害是一种复杂且危险的心理病理行为,表现为个体蓄意以不同方式破坏自己的身体组织,如用利器割伤/划伤、打火机烧伤、以头撞墙等,具体伤害形式可能超过十多种(Ross & Heath, 2002; Whitlock, Eckenrode, & Silverman, 2006; 郑莺, 2006)。”实则是自伤行为的一个初步介绍,并非严格意义上的定义,由于研究中的自伤是指非自杀性自伤(Non-Suicidal Self-Injury, NSSI),而该定义并未强调“无自杀意图”这一核心要素。因此,根据您的建议,我们在修改稿中使用原文献的后一处定义,即业界公认的自伤定义为准,并对文中自伤行为的用词进行了统一。具体修改见稿件 1.2 强化敏感性与自伤行为,蓝色标记部分。

意见 4:在研究一中,既然是“以班级为单位集体施测”,被试取样是如何随机选取的,建议再详述。

回应:感谢审稿专家的意见。此处在原文中的确未交待清楚,容易引起读者猜测或不解。实际上,研究一采用的是分层整群取样的方法,在武汉市 5 所综合性普通高等院校,对大一至大四 4 个年级的学生以班级为单位进行随机抽样。文中已对取样方法进行补充说明,见稿件 2.1.1 被试,蓝色标记部分。

意见 5:文中提到“在研究二中,被试分组时,对部分较难分组的被试,在实验前结合简短的访谈进行确认”,采用的访谈提纲是怎样的,谁进行的访谈,访谈的目的是什么?如可能建议适当交待下。

回应:感谢审稿专家的意见。由于研究二对自伤被试进行分组时,存在这样一部分被试,他们在两类功能维度上均有得分,但存在明显的功能倾向。对于这类被试,我们通过极端组划分法结合访谈法对其进行功能分组。被试首先需满足在某一功能维度上得分前 27%而在另一维度上得分后 27%,随后在实验开始前由主试人员通过简短的半结构式访谈对其自伤功能类型进行再次确认。访谈问题主要如下:(1)你近一年内发生的自伤行为,主要动机是什么;(2)你最近一次采取自伤行为的动机是什么;(3)过去的自伤行为主要给你带来了什么功能。

意见 6:在 1.2.4 强化敏感性与临床心理障碍部分,作者有提到“过高/过低的惩罚敏感性/奖励敏感性与抑郁症、焦虑障碍、强迫症、进食障碍、物质滥用、人格障碍等多种临床心理障碍和行为问题相联系(Bijttebier et al., 2009)。其中,BPD、抑郁症、强迫症、进食障碍和物质滥用均被认为与自伤存在不同程度的联系”,本研究在取样和处理数据的过程中,是如

何控制自伤与各种心理障碍（例如，BPD 和自杀）共病的情况，以避免或减弱同时存在的其他心理障碍对强化敏感性和自伤的关系的影响的？

回应：审稿专家考虑得十分周到。事实上，自伤行为存在于多类个体中，也的确与多种心理障碍相联系。广泛的临床评估发现，多数自伤者同样存在符合其他精神障碍诊断的症状，尤其是进食障碍和物质滥用（DSM-5）。自伤也曾被认为是 BPD 的症状之一，但在 DSM-5 中已对此进行修改，将自伤作为一种未定型的障碍。因此自伤到底是一种独立的障碍还是其他心理障碍的症状，目前不得而知。然而，我们认为这并不影响对自伤行为的研究。在这样的背景下，我们主要关注自伤这个行为本身，即探讨具备这类行为的个体具有怎样的强化敏感性基础。

通过对自伤行为本身的研究，发现自伤行为背后的独特特征，将更利于对这类行为的理解，进一步发现它与其他心理障碍间的区别和联系。如自伤与 BPD 的关系一直是业界关注的重点，长久以来，它一直被认为是 BPD 的一个核心症状，但近年的研究表明，BPD 并不存在于所有自伤者中，两者在某些方面存在本质差异。BPD 患者常表现出心理失常的敌意攻击行为，而自伤行为则更多地与亲密、合作行为及积极关系相联系（DSM-5），这些特征是所有自伤行为所共有的，正是这些证据使得自伤行为不再被完全涵盖于 BPD 之内。同理，我们也希望通过研究发现自伤者共有的稳定特征。假设样本中自伤者存在的共病类型是随机分布，那么如果所有自伤者都表现出稳定的强化敏感性特征，而这些特征又与 BPD 或其他心理障碍表现出的模式存在差异，那至少可以证明两类障碍不是完全重叠的，进而为区分自伤与其他障碍提供启示。在当今自伤归类尚无定论的背景下，这样的工作便显得更有意义。

此外，对于与自伤有着本质区别的自杀，本研究是进行了控制的。由于本研究中的自伤是指非自杀性自伤（NSSI），因此在取样时确保所有自伤问卷上的得分均是在无自杀动机情况下的自伤行为。此部分在原文中未交代清楚，且所附自伤问卷版本有误，实际研究中特意在问卷指导语部分强调了“在没有自杀动机的情况下”。相应修改见 2.1.2 研究工具和附录 2 青少年自我伤害问卷，蓝色标记部分。

最后，正如审稿专家的考虑，为了避免或减弱共病对研究效应的影响，我们在今后的研究中还是应该尽量尝试对该问题进行控制。本研究在这方面的不足也已在 4.4 研究局限与未来展望中进行补充说明。

第二轮

审稿人 1 意见：

我仔细查看了作者对三位审稿人的意见回复，很认真，并且基本上都进行了补充和修改，不能补充的部分也进行了较为合理的解释。可以发表。

回应：感谢审稿专家认真严谨的审稿，以及对修改稿的肯定。

审稿人 2 意见：

通读了《审稿意见和作者回应》和修改稿全文，感谢作者对审稿人的诸多建议和质疑做了深入思考和审慎探究，并做了必要的修改。以下建议供参考：

意见 1：关于自杀的控制还有待商榷。作者在回应中，明确提出本研究中的自伤行为是指非

自杀性自伤，其定义是“在没有明确自杀意图的情况下，个体故意、重复地改变或伤害自己的身体组织，这种行为不为社会所认可，且不具致死性或致死性较低；具体的伤害形式可能超过十多种，如用利器割伤/划伤、打火机烧伤、以头撞墙等（江光荣等，2011）。”，但作者仅在使用的自伤问卷指导语中强调了“在没有自杀动机的情况下”，并将之作为对自杀控制的手段，进而提到“基于该问卷的被试筛选与分组也都是在排除自杀意图的前提下根据被试自伤行为的水平及其功能进行的”，其他部分未见提到如何控制了自杀。如何来科学考量被试是“没有明确的自杀意图”和“没有自杀动机”的，又如何删除整群取样中必然包含的有自杀观念（在人群中比例大约 5-10%）、自杀计划、自杀准备、自杀尝试或自杀未遂等倾向的被试，是一个有待探究的问题，建议同时使用新效度好的自杀评估工具，以更有效地控制自杀。

回应：感谢审稿专家的意见。在研究一中，我们通过自伤问卷对被试的自伤行为进行测量，指导语旨在保证被试的自伤行为是在无自杀意图的前提下进行的，而无自伤行为人群中的自杀意念则不在本研究考虑的范围内。在自杀领域，自杀意念、自杀计划、自杀准备、自杀尝试、自杀未遂等分处于层次不同的阶段，但都基于自杀意念之上。从自杀意念的角度讲，自杀意念在普通人群中的终生发生率较高，但只有以自杀意念为基础、为目的而表现出自伤的自杀行为才是对本研究有干扰、可能混淆研究结果进而需要我们控制的对象。一方面，现有研究表明，从自杀意念发展为自杀尝试的人群比例是很小的，因此我们认为单纯的自杀意念本身（即只有自杀意念而无自杀尝试）对自伤行为研究的影响并不大，没有必要完全控制；另一方面，对少数基于自杀动机发生的自伤行为，我们认为通过本问卷的指导语是可以得到有效控制的。本研究所采用的问卷以有无自杀动机为门槛，一开始便已排除自杀意图及以此为基础的进一步行为，其中包括对自伤行为有混淆作用的自杀尝试等。并且，以往国内外本领域内的相关研究也多采取此方法来控制自杀意图，本研究所用自伤问卷在编制之初也是借鉴国外自伤量表修订而成。因此，出于被试作答质量的考虑，我们没有再单独选用自杀评估工具对自杀意图进行测量。但我们认为审稿专家提出的建议十分有借鉴意义，在未来研究中，应该尝试使用简短的自杀评估工具对自杀意图进行更科学严谨的控制。作为本研究的局限之一，已在修改稿中进行补充说明，见 4.4 研究局限与未来展望，蓝色标记部分。

在研究二中，所有被试除完成自伤问卷，还完成了自伤功能问卷，后者本即是对被试自伤动机的一个考察。我们在对被试进行筛选和分组时，结合被试自伤问卷显示的自伤水平，自伤功能问卷显示的自伤动机，以及部分被试的半结构式访谈信息，已能较好地确定入组的自伤被试，在这一过程中，任何表现出自杀动机的被试均被排除在外。通过上述程序，我们基本已能保证所有入组被试并非出于自杀的意图进行自伤行为，而是出于调节负面情绪或寻求刺激的动机。我们认为这样的控制是可靠且符合自伤定义的。经您的提醒，我们也在研究二中对自杀意念的控制进行了补充说明，见 3.1.1（1）被试的筛查与分组，蓝色标记部分。

意见 2：关于无效问卷的情况还需要详细描述。文中提到“共发放问卷 778 份，回收有效问卷 717 份，回收有效率为 92.16%。”，但未详细交代无效问卷删除的标准。按照目前研究取样和研究目的，无效问卷可能包括漏答误答问卷、Marlowe—Crowne 社会称许性量表（MCSD）简版得分大于 5 分的问卷和“有明确自杀意图”的问卷三部分。建议补充无效问卷的具体信息和删选情况。

回应：感谢审稿专家的建议。此处补充说明如下：本研究共发放问卷 778 份，根据以下标准剔除无效问卷 61 份：（1）全问卷漏答率超过 15%，或指导语理解有误导导致问卷作答方式错

误，共 43 份；(2) Marlowe—Crowne 社会称许性量表 (MCSD) 简版得分大于 5 分，共 10 份；(3) 自伤问卷作答中明确表明存在自杀意图，共 8 份。最终回收有效问卷 717 份，回收有效率为 92.16%。文中亦作出相应修改，见 2.1.1 被试，蓝色标记部分。

第三轮 编委复审意见与作者回应

本文经过两次修改，已经比初稿提高及完善很多。

回应：感谢编委专家对修改稿的肯定。

意见 1：仍待澄清的问题：AB 问卷是两个等值问卷还是将量表题目按不同顺序排列的问卷？

回应：A、B 两个版本的问卷是指，在全问卷合成时，BIS/BAS 量表中文版、青少年自我伤害问卷、Marlowe—Crowne 社会称许性量表这三个量表的排列顺序不同。A 版问卷为青少年自我伤害问卷→BIS/BAS 量表中文版→社会称许性量表；B 版则为 BIS/BAS 量表中文版→青少年自我伤害问卷→社会称许性量表。如此操作的目的旨在平衡两个核心变量——自伤和强化敏感性的施测顺序，控制项目启动效应、项目语境诱发的情绪及其他与项目语境和项目嵌套有关的偏差，以在一定程度上降低共同方法偏差。文中已作出补充说明，见 2.1.3，蓝色标记部分。

意见 2：还需要对文章进行删节，目前文章字数超出了杂志的要求。

回应：根据编委专家的建议，我们在保证行文结构完整、条理清晰并保留各审稿专家修改意见的基础上，仔细对文章进行了再阅读并作出必要删节，已从上一稿的 15 页缩减至 13 页(包括中文摘要和图表)，参考文献由原来的 68 篇减至 58 篇。目前所保留的内容实属本研究之重点或有助读者理解的必要说明，并无冗余或赘述之文，若进一步删节恐破坏文章结构、影响可读性，故作此修改。望予以理解。