

《心理学报》审稿意见与作者回应

题目：青少年抑郁情绪的发展轨迹：界定亚群组及其影响因素

作者：侯金芹 陈祉妍

第一轮

审稿人 1 意见：通过追踪研究模拟出青少年在青春期抑郁情绪的发展轨迹，并同时界定出抑郁情绪发展的亚类型。该论文的设计和统计方法都有新意。总体来说，这篇文章有科学价值。但是，我也有几点建议供作者参考：

意见 1：作为一个纵向研究，流失率分析是保证最终样本代表性的重要步骤，不能省略。

回应：谢谢审稿专家的建议。相关信息已补充至文中，详见正文第 3 页。

意见 2：没有 3 个时间点各变量均值，标准差的数据

回应：谢谢审稿专家的建议，男女生抑郁情绪在三个时间点的均值和标准差详见正文第 5 页，表 2。青春期发育等预测变量的均值和标准差详见方法部分，正文第 4-5 页

意见 3：新近的，权威的参考文献明显缺少，可能表明作者没有充分掌握该领域的前沿研究成果。

回应：谢谢审稿专家的建议。本研究新近的参考文献确实不多，原因如下：

西方对青少年抑郁情绪发生发展的研究从 90 年代就开始，2000 年初研究逐渐达成统一，得出了很稳健的研究结果。如，青春期女生的抑郁情绪显著高于男生；女生抑郁情绪迅速增长，呈现倒 U 型发展曲线；抑郁情绪的发生发展存在着个体差异性，分为若干个亚群组等；既然已经掌握了青少年抑郁情绪发生发展的大框架，因此该领域新近的、前沿的研究方向已经再次细化：如，从神经生理与分子遗传等角度探讨抑郁情绪的发生与发展(e.g.Chen, Li, & McGue, 2013; Chen, Li, Natsuaki, Leve, & Harold, 2014; Heller et al., 2013);从临床样本对照的角度来探讨抑郁的影响因素，(e.g.Boyes, Girgis, D’Este, Zucca, & Lecathelinais, 2013; Br édart et al., 2015; Kuchibhatla & Fillenbaum, 2011; M’Bailara, Cosnefroy, Vieta, Scott, & Henry, 2013); 从心身发展的角度探讨生理对抑郁的影响，(e.g.Kuo, Chen, & Tzeng, 2014; Sabiston et al., 2013); 将研究的年龄范围扩大至老年人(e.g.Huang et al., 2011; Kuchibhatla, Fillenbaum, Hybels, & Blazer, 2012)和婴幼儿(e.g.McManus & Poehlmann, 2012)等等。

然而国内从发展的角度对青少年抑郁情绪发生发展的动态研究还处于初步阶段，只有先明了国内青少年抑郁情绪发展的脉络才能开展后续的研究，在西方研究的基础上先探讨国内青少年抑郁发生发展的特点，比较分析异同是我们迈出的第一步。因此，本研究对青少年抑郁情绪所引用的参考文献虽不新近但都很权威。

意见 4：统计分析的思路比较混乱。建议作者把 3 个分析放到一个综合的 MSEM 中，将生理发育、学业压力和人际关系纳入发展模型中。

回应：谢谢审稿专家的建议。本文之所以分成 3 个分析而未放到一个模型之中，是从统计方法与解决的问题所需两方面进行考虑的。理由如下：

首先，从统计方法的角度来讲，是分为以变量为中心的方法(variable-centered)和以人为中心的方法(person-centered)两种。其中，前者包括回归、因素分析和结构方程模型等，

一般用于界定因变量的影响因素，论述自变量和因变量之间的关系。后者包括聚类分析、潜类别分析等，关注于人与人之间的不同，目标在于将人群分为不同的亚群组，使得群组内部的个体更为相似，个体差异更小。本研究分析 1 采用的是以变量为中心的分析方法，意在探讨男女生抑郁情绪随时间的变化趋势。分析 2 采用的是以人为中心的方法，考虑到个体差异性，将抑郁情绪的发展轨迹区分为不同的亚群组。基于此，分析 1 和分析 2 不能合并至一个模型中。既然群组内部更为相似，群组之间的差异很大，是哪些因素造成了这些差异？这是研究三意欲解决的问题。因此，研究 3 是在研究 2 的结论基础上展开的，这是研究 2 和研究 3 之间的逻辑关系。

再次，研究 1 和研究 2 之间的逻辑关系是什么？这个问题要从意欲研究的问题的角度上来说。美国比较稳健的研究结果都发现：（1）从静态的平均水平来看，女生的抑郁情绪显著高于男生(Ge et al., 2003; Hankin et al., 1998)；（2）从动态的发展轨迹来看，青春期女生的抑郁情绪呈现倒 U 型曲线发展，进入青春期急剧升高，历经平台期，然后下降。而男生的抑郁情绪在整个青春期保持平稳(Adkins, Wang, Dupre, van den Oord, & Elder Jr, 2009; Ge, Natsuaki, & Conger, 2006)。但是国内目前关于此方面的研究还仅仅停留在静态的平均水平，并且现有的研究与美国的结果并不一致。国内有的学者发现男女生的抑郁情绪一样高(Li, McGue, & Gottesman, 2012)，还有的研究者发现男生的抑郁情绪甚至比女生还高(Sun, Tao, Hao, & Wan, 2010)。为什么会出现这种差异？想要解决这个问题，我们仅仅停留在静态的平均数比较层面是不够的，因此动态的发展轨迹可以从发展的角度为我们提供一个视角来解决这种中西方的差异。是不是因为发展轨迹的不同而导致了男女生抑郁情绪的均值比较与西方不同？这是研究 1 所想要解决的问题。

从宏观的发展轨迹方面我们发现中西方青少年抑郁情绪的发生发展确实会存在差异，那么这种差异是怎么造成的？现有学者普遍支持抑郁是连续体，即持有“量的差异”的学者们认为轻度、中度和重度抑郁是线性相关的，处于同一个连续体的不同点(Brown & Barlow, 2005; Flett, Vredenburg, & Krames, 1997)，那么中西方青少年抑郁情绪发生发展总趋势的异同会不会是因为其中各类型亚群体所占的比重不同所致？（即，轻度、中度和重度的比例不同所致）。基于此，研究 2 着重分析亚群体的特点，这是研究 2 想要解决的问题。

总起来说，研究 1 和研究 2 是层层推进，分别从宏观总体和微观亚群体两个角度来探讨中西方目前研究结果的差异。

综上所述，3 个分析之间相互推进又相互联系，但分析方法有所差异，因此，暂未合并。

意见 5: 对于女生群体，结果也是显示 2 组最佳且最节省，这从 LMR-LRT 检验和熵的值可以看出。

回应: 谢谢审稿专家的建议。

确定亚群组的组数是潜类别分析中的重要一环，受到研究问题、模型拟合指数、模型的节俭性、理论可证以及可解释性多方面的影响。Nylund 等学者（2007）采用模拟运算的方法发现在众多的指数中，尤其是在初期的探索性分组中 LMR-LRT 的效果最好，其次是 BIC。其中 BIC 越低越好，LMR-LRT 达到显著性水平就说明 K 组的分类比 K-1 组的分类，有显著性差异。因此，现有的研究中一般首先采用这两个指标进行分组。然后，再根据一些模型成功聚敛的指标如熵（entropy），每组人数不少于总人数的 1%，以及后验概率较高等指标对分组进行优化。即，BIC 和 LRT 在确定分组权重上要大于熵值。综上所述，第一步先根据 BIC 和 LMR-LRT 进行选择，2 组和 4 组的 LMR-LRT 均显著，但 4 组的 BIC 小于 2 组，因此，选择 4 组的分类结果。虽然 4 组的熵值小于 2 组，但尚可接受。

意见 6: 如果可能，建议用 bootstrapped LMR-LRT。

回应：谢谢审稿专家的建议。Nylund 等人(2007)建议，在初期的探索性分组中，还是建议采用 LMR-LRT，因为 bootstrap LMR-LRT 明显增加了运算时间。确定分组时可依据 BIC 和 LMR 的 p 值这两个指标进行，得到一些初步的模型之后再通过 BLRT 指标进行优化。结合审稿专家和 Nylund 等学者的建议，我们再次对目前的模型进行了运算，发现男生两组，女生四组的分类的 BLRT 值依然显著，具体结果如下：

(1) 男生 2 组的 bootstrap 运算结果：

PARAMETRIC BOOTSTRAPPED LIKELIHOOD RATIO TEST FOR 1 (H0) VERSUS 2 CLASSES

| | |
|----------------------------------------|---------------|
| H0 Loglikelihood Value | -7548.290 |
| 2 Times the Loglikelihood Difference | 458.267 |
| Difference in the Number of Parameters | 3 |
| Approximate P-Value | 0.0000 |
| Successful Bootstrap Draws | 5 |

(2) 女生 4 组的 bootstrap 运算结果：

PARAMETRIC BOOTSTRAPPED LIKELIHOOD RATIO TEST FOR 3 (H0) VERSUS 4 CLASSES

| | |
|----------------------------------------|---------------|
| H0 Loglikelihood Value | -8213.517 |
| 2 Times the Loglikelihood Difference | 88.953 |
| Difference in the Number of Parameters | 3 |
| Approximate P-Value | 0.0000 |
| Successful Bootstrap Draws | 5 |

审稿人 2 意见：该研究考察了抑郁情绪的发展模式，研究的优点是采用了追踪数据，做了以人为中心的统计分析，并考察了不同类发展模式的预测因素。研究有一定新意，然而，也存在明显问题：

意见 1：研究按照自然班来取样，应该充分报告每一次调查时被试的年级分布信息、年龄信息等。每次调查的准确时点要报告出来。每次调查时，每个年龄组和年级组的样本量是多少？

回应：谢谢审稿专家的建议。此研究的一轮来自于全国青少年心理健康状况调查项目，采用多阶段分层整群随机抽样的方法，从东部、中部、西部和东北部四大区域选取了九个城市，每个城市按照重点和非重点分层抽取中小学，每所学校分层抽取自然班，调查以班级为单位进行。

追踪研究是建立在一轮全国大调研的基础之上的，但由于各方面条件的限制，追踪研究的样本只是一轮大调研中的一部分样本，以学校为单位的追踪调研主要是对其中 7 个城市所选取学校的非毕业班同学进行追踪研究（相关信息请详见正文第 3 页的样本部分描述）。由于涉及到升级和重新分班，所以后续的追踪研究并不能实现以自然班的形式来取样，更多地是依据是否参加了第一轮调研来进行的。由于年龄和年级信息具有很高的重叠性，本研究就以年龄段来表示，各年龄在各轮的分布情况见下表（见正文第 4 页，表 1）。

出生日期为 1992-1998 组

| | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 总计 |
|------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 一轮 (2008 年底 09 年初) | 418 | 665 | 699 | 771 | 748 | 833 | 918 | | | | 5052 |
| 二轮 (2009.12-2010.7) | | 371 | 399 | 415 | 643 | 436 | 375 | 502 | 50 | | 3192 |
| 三轮 (2011.5-2011.9) | | | | 210 | 219 | 151 | 344 | 352 | 190 | 178 | 1644 |

意见 2: 调查的样本流失率太高，影响结果可靠性。“大约 50%的青少年继续参加了第三轮的追踪研究”，这个百分数相对谁来说的？不论如何，前后样本的代表性是不可比的。是否抑郁程度与此有关？作者说“第三轮调查流失的样本与三轮均参加的青少年的抑郁情绪得分差异不显著”，这是就第一次的数据来说的吗？ $t=1.83$ 虽然不显著，但也是一个接近显著的水平了。样本问题，大大限制了结论的可信性。

回应: 谢谢审稿专家的建议。追踪率是对意欲追踪的样本来说的，追踪样本是指一轮中七个城市（除去北京和庆阳）中所调查学校的非毕业班同学。

修改稿中已详细添加了关于追踪样本的介绍，对追踪方式的介绍、对流失率的分析以及对流失样本和追踪样本的抑郁情绪比较分析等，详见正文第 3 页研究方法部分。

意见 3: 如何分离出年龄和年级的影响？如何分离出年龄变化和样本组个体差异的影响？群组序列设计部分的结果实际上杂糅了 cohort（出生年龄）组的个体差异、发展性（自然年龄）差异等，在这种情况下拟合出的增长曲线代表的是什么？每个年龄组样本量大小的影响是否被考虑进拟合模型了？

回应: 谢谢审稿专家的意见。这是一个很好的问题。

1. 由于年龄和年级具有很高的重叠效应，同一个年级中个体的年龄差异不会太大，在此情况下，分离年龄和年级的效应的研究意义不大。因此，在实际研究中一般会从“年龄”和“年级”这两个变量中二取一。在本研究中我们选取了年龄作为主要的变量，理由如下：

(1) 本文是一个追踪研究，除去第一轮中以自然班级为单位收集数据，在后续的追踪研究中，由于原有的班级已被打乱，追踪研究更多的是以学生个体为单位进行的，班级的信息被削弱。

(2) 其次，本文的主要研究目的是探讨青少年抑郁情绪随年龄的变化趋势，因此年龄是核心变量。

综上所述，本研究主要探讨了年龄的效应，未考虑班级的效应，但再次感谢审稿专家的意见，在后续的研究中我们会逐步细化班级的信息。

2. 第二个问题是关于年龄效应（不同 cohort 组）和发展效应（同一 cohort 组随时间的变化）这之间的关系如何分离？

回应: 最理想的追踪研究是同一个年龄组的个体被追踪多次得到的发展曲线，这样的研究从设计上就有效地分离了年龄效应和发展效应。但是本文的研究设计是对多个年龄段的个体进行了有限次数（3 次）的追踪。从设计上不能规避年龄效应和发展效应，但是这个问题可以通过统计方法来解决。

群组序列设计（cohort sequential design）(Nesselrode & Baltes, 1979)，又称为加速追踪

设计 (an accelerated longitudinal design) (Bell, 1953)就是专门用于解决对多个年龄段个体进行了有限次数追踪的数据,并且已有研究表明这种研究方法对发展曲线拟合的正确度非常接近传统的追踪研究(对 cohort 追踪多次的研究)(Duncan, Duncan, & Hops, 1996)。

群组序列设计摒弃了传统的以追踪轮数 (wave) 为单位的分析方法,而采用的是以年龄为分析单位 (age as the unit of analysis), 每个年龄组的样本均被考虑进模型中, 这种方法有以下几点优势:

(1) 适合年龄跨度大的样本设计, 可以充分考虑到年龄变异。

(2) 样本年龄本身的跨度以及个体年龄的增长使得两次追踪之间出现了年龄重叠(例如, 本研究中一轮的年龄跨度是 10-16 岁, 二轮变为 11-17 岁, 一轮到二轮之间 11-16 岁即为重叠年龄), 群组序列设计将这种年龄差异以及个体发展差异考虑在内, 使得在有限追踪次数下模拟出更长的发展曲线成为可能。即, 群组序列效应同时分析了年龄效应和发展效应。

(3) 大规模的追踪研究往往不能保证所有地区的数据在同一个时间点上收集上来, 在同一轮的追踪研究中, 数据的收集时间会略有差异。群组序列设计通过以年龄为单位的方法也很好控制了影响。

意见 4: 负性生活事件的测量工具是如何来的, 信度和效度如何确定。

回应: 负性生活事件采用的是中学生生活事件调查 (Junior High Life Experiences Survey) (Swearingen & Cohen, 1985)的修正版。中文版本的负性生活事件包括家庭、同伴学校环境三大方面, 本研究主要涉及同伴以及学校环境两大方面, 这些指标已被广泛应用于以往的研究之中, 信效度良好(Chen, Yu, Wu, & Zhang, 2015; Ge, Conger, & Elder Jr, 2001)。相关信息已补充进正文, 详见正文第 4 页。

意见 5: 结果的报告不规范。结果“男生的抑郁情绪随时间稳步增长, 而女生的抑郁情绪在一二两轮间大幅增长, 二、三轮间增长速度放缓”, 缺乏相应的推论统计的数据支持。

回应: 谢谢审稿专家的意见。这部分表述确实有失严谨, 已在正文中相应部分删除

意见 6: 表 2 模型拟合指数是就哪个模型来说的? 不清楚。线性模型和二次曲线究竟哪个拟合好? 似乎要选出一个最好的, 来报告具体的结果, 以此来说明增长趋势问题。说女性是倒 U 型曲线似乎很牵强。

回应: 表 3 (原来的表 2) 模型是男生、女生分别进行的, 分别对 $y=ax^2+bx+c$ 这个曲线进行拟合, 如果二次项的系数显著 (即 a 显著), 则说明是非线性; 若 a 不显著, b 显著, 则说明是线性的。(并未进行线性和非线性的直接比较)

男生的数据结果显示轨迹是线性的 (二次项系数不显著), 模型拟合良好。

女生的数据结果显示二次项和一次项都显著, 说明是非线性的。方程是 $y=-3.45x^2+6.27x+10.30$, 开口向下的抛物线, 其中顶点坐标是 (0.91,13.8), 由于模型将 15 岁设置为坐标原点, 0.1 代表 1 岁, 因此抛物线的拐点大约是 24 岁。从抛物线函数的角度来看, 10 岁-24 岁是单调递增函数, 而 24 岁之后是单调递减函数。因此, 女生的数据确实模拟出以 24 岁为拐点的开口向下的倒 U 型曲线。

意见 7: 女生在第二轮的抑郁情绪显著高于男生, 为什么如此? 第一和第三轮为什么不如此?

回应: 这是单纯从平均数的比较上得出的结论。正如目前国内的研究有的发现男生抑郁情绪高于女生(Sun et al., 2010), 有的发现两者没有差异(Li et al., 2012)一样。本文提出的一个新的思路去解决这种已有差异的不一致即是也许静态的平均数比较掩盖了太多的信息, 比如年

龄效应，比如发展效应等，而应该从动态的发展角度来比较。

首先，青春期的抑郁情绪随着年龄的发展有很大的变化。描述性统计部分单纯比较每一轮的性别差异，未考虑到年龄的影响。

其次，诸如研究 2 对亚群组的分析，男生和女生的抑郁情绪发展轨迹还分为不同的亚群组，在每一轮的收集点，每个亚群组的比重不同，也会造成平均数上的差异。

据上分析，单纯地比较平均数的差异会掩盖掉很多信息而出现假性的结果。为了消除误解，正文中已删掉了男女生在一轮、二轮和三轮的性别差异比较分析，详见表 2（正文第 5 页）

意见 8：虽然该研究在设计和统计方法上有一定新意，但存在较为严重的问题，特别是问题 2 和 3。综合考虑，论文的质量和创新性似乎还达不到心理学报的要求，建议暂不发表或转投他刊。

回应：谢谢审稿专家的建议。对于问题 2 追踪样本的流失问题，在审稿专家的指导下，正文中详细阐述了追踪的流程、追踪方式以及对流失样本的分析（详见第 3 页方法部分）。由于全国七个城市的追踪样本难度大，追踪率保持在 60% 以上，虽然追踪率不高，但尚可接受。

对流失样本的分析发现，在一轮中小学五年级、初二和高二即将升入毕业班的同学中，参加了追踪研究和未参加追踪研究的青少年在抑郁情绪上有显著差异，其它的流失样本与未流失的均无显著差异。即将升入毕业班同学流失掉的样本抑郁情绪显著高于未流失掉的样本存在两种可能性：

（1）样本存在偏差。

（2）目前各学校中都存在一些隐形的提升升学率的方法。例如，将那些升学无望，不认真学习的后进青年劝退以减少升学率的分母。而这些后进青年的问题行为和抑郁情绪都显著高于一般同学。从这个角度来说，本文的取样倒是有效代表了总体样本。

但从现有的数据来看，我们不能明确说明到底是哪一种情况，因此这也是本研究的一个局限之处，已在正文中予以说明（详见正文 12 页）。

关于问题 3，本文的研究设计确实不能有效区分年龄效应和发展效应，但是本文所采用的研究方法正是为了解决此问题而产生的，年龄效应和发展效应通过统计方法得以统合，详情请见对问题 3 的解答。

青少年的抑郁情绪在青春期的发生发展是一个重要的研究课题，它影响到青少年发展的方方面面，甚至能有效预测成年期抑郁的发病率。西方对青少年抑郁情绪的研究在 90 年代就已经开始了，关注于青少年抑郁情绪的发展轨迹、性别差异、亚群组的界定，亚群组不同发展模式的影响因素等等。截至目前，关于青少年抑郁情绪的发展已经得到了很多稳健的结论，例如，女生的抑郁情绪在青春期呈现倒 U 型曲线的发展，女生的抑郁情绪比男生严重等。但是国内对这方面的研究基本还处于未开荒的阶段：

（1）因为国内大规模的追踪研究很少，大型的追踪数据库基本没有。

（2）目前国内已有的零星的关注青少年抑郁情绪的研究来看，现有的研究本身存在不一致，并且与西方稳健的结论存在不一致性。这些均提示我们国内青少年抑郁情绪的发生发展可能会呈现出独特的发展模式。

（3）对抑郁的影响因素中，我们有一些特异性的因素，如学业压力，如人际关系等。

因此，开展国内青少年抑郁情绪的发展性研究以及对文化特异性的影响因素进行研究是非常必要的。本文也初步证实了中西方确实存在一些差异，但究竟是哪些因素造成了这些差异还需要进一步探究，还需要广大学者一起努力。本文虽然存在不少的局限性，但也算是迈出了青少年抑郁情绪发生发展研究的第一步，为后续研究奠定了基础，希望审稿专家可以予以考虑。

审稿人 3 意见：该研究基于大型数据库，对中国青少年抑郁情绪的发展特点进行了分析和讨论，采用多种方法，对结果的分析层层深入，是一篇思路清晰的研究报告。下面我提出几点看法和建议：

意见 1：关于抑郁量表评定标准和计分说明，文中缺乏基本的介绍：

回应：在抑郁的研究领域，“抑郁是否是连续的？”一直以来都是一个具有争议的话题。即，轻度或者中度的抑郁个体 (mild and moderate depressive symptoms) 与被临床诊断为抑郁症的个体 (major depressive disorder as a syndrome) 是否具有本质区别 (qualitative difference) 还是只是存在量的差异 (quantitative difference)? 持有“质的差异”的学者认为，轻度抑郁个体只是局限于轻微的烦躁不安的情绪状态而重度抑郁患者则受到各种负面情绪的综合困扰，如快感缺失，饮食和睡眠习惯改变，自杀意念或者想法以及无望和无助感等。抑郁的测量主要采用临床诊断的方法，如 DSM-IV。持有“量的差异”的学者们认为轻度、中度和重度抑郁是线性相关的，处于同一个连续体的不同点，抑郁的测量可以采用量表，如流调中心抑郁量表 (CES-D) 和贝克抑郁量表 (BDI) 等 (Flett, Vredenburg, & Krames, 1997)。近些年来，越来越多的学者们倾向于认为抑郁是连续的，因此采用量表的方法来测量抑郁的研究也越来越多 (Brown & Barlow, 2005)。

本研究采用抑郁具有“量的差异”的观点，采用流调中心抑郁量表 (CES-D) 来测量青少年及其父母的抑郁，得分越高说明抑郁情绪越高，并未采用 cut-off 划段。流调中心抑郁量表 (The Center for Epidemiological Studies Depression Scale, CES-D) 为美国国家心理健康中心 (NIMH) 的 Radloff 于 1977 年编制，在国际上被广泛用于对普通人群进行抑郁的调查，适用于青少年 (Radloff, 1977; 陈祉妍, 杨小冬, 李新影, 2009)。CES-D 共 20 题，要求被试使用 0-3 评定最近一周内症状出现的频度。CES-D 有多种简版，本次调查使用了 CES-D 中的 13 题，包括情绪低落、积极情绪缺乏、绝望、人际交往困难、做事困难、睡眠困扰等症状。

正文中概述了这部分内容，详见正文第 3-4 页。

意见 2：该研究的主要结果在于发现中国青少年抑郁情绪发展的亚类型及影响因素，但文中对于这些亚类型结果的分析略有些简单，不同类型的群体有什么具体的特点？以及为什么会存在多种亚类型人群？仅仅依赖于影响因素的逻辑回归分析是不够的。即使是逻辑回归分析，其结果的分析也可以再深入细致化。

回应：谢谢审稿专家的建议。

(1) “不同类型的群体有什么具体的特点？”已在正文中男女生个亚群组后面补充了有关发展轨迹的截距和斜率的信息 (详见正文 7 页)，进一步将类似“逐渐缓升组”，“逐年陡升组”等表述以数字的形式具体化，可以与图 3 相互补充，将信息表达地更为清晰与具体。

(2) “为什么会存在多种亚类型人群”？这正是研究 3 的目的所在，本文试图从影响因素的角度入手来分析亚类型人群的原因，但由于研究设计本身的局限性，本研究只能从人际、学业压力以及青春期发育等几个角度入手来探讨。但这些因素肯定远远不够，例如：

a. 本研究的追踪设计是从青春期开始的，有些影响因素可能从儿童期已经开始发挥作用了，本研究捕捉不到

b. 父母遗传方面的影响。

c. 遗传和环境之间的相互作用等等。

因此，我们很赞同审稿专家的“仅仅依赖于影响因素的逻辑回归分析是不够的。即使是逻辑回归分析，其结果的分析也可以再深入细致化”的建议，影响因素不但单独发挥作用，彼此之间还可能产生交互作用，共同影响抑郁情绪的发生与发展。但本研究为后续研究奠定了第一步，为以后细致化地收集与探讨影响因素的研究奠定了基础，这些都需要我们继续努力，在后续的研究中进一步深化。

研究的局限性也在正文中得以阐述，详见正文 12 页。

意见 3: 关于相关影响因素的研究，该研究主要关注负性因素，对结果的解释有限，讨论中也很少关注青少年情绪发展过程中的积极因素影响，如父母关爱、同伴友谊对抑郁情绪的调节作用

回应: 谢谢审稿专家的建议。仅仅关注于负性因素的影响确实是本研究的一个局限，已在正文中补充对积极因素的论述。详见正文 11 页。

参考文献:

- Adkins, D. E., Wang, V., Dupre, M. E., van den Oord, E. J. C. G., & Elder Jr, G. H. (2009). Structure and stress: Trajectories of depressive symptoms across adolescence and young adulthood. *Social Forces*, 88(1), 31-60. doi: 10.1353/sof.0.0238
- Bell, R. Q. (1953). Convergence: An accelerated longitudinal approach. *Child Development*, 145-152.
- Boyes, A. W., Girgis, A., D'Este, C. A., Zucca, A. C., & Lecathelinais, C. (2013). Prevalence and predictors of the short-term trajectory of anxiety and depression in the first year after a cancer diagnosis: A population-based longitudinal study. *Journal of Clinical Oncology*. doi: 10.1200/jco.2012.44.7540
- Br élard, A., Merdy, O., Sigal-Zafrani, B., Fiszler, C., Dolbeault, S., & Hardouin, J. (2015). Identifying trajectory clusters in breast cancer survivors' supportive care needs, psychosocial difficulties, and resources from the completion of primary treatment to 8 months later. *Supportive Care in Cancer*, 1-10.
- Brown, T. A., & Barlow, D. H. (2005). Dimensional versus categorical classification of mental disorders in the fifth edition of the Diagnostic and statistical manual of mental disorders and beyond: Comment on the special section. *Journal of Abnormal Psychology*, 114(4), 551-556.
- Chen, J., Li, X., & McGue, M. (2013). The interacting effect of the BDNF Val66Met polymorphism and stressful life events on adolescent depression is not an artifact of gene-environment correlation: Evidence from a longitudinal twin study. *Journal of Child Psychology and Psychiatry*, 54(10), 1066-1073.
- Chen, J., Li, X., Natsuaki, M. N., Leve, L. D., & Harold, G. T. (2014). Genetic and environmental influences on depressive symptoms in Chinese adolescents. *Behavior Genetics*, 44(1), 36-44.
- Chen, J., Yu, J., Wu, Y., & Zhang, J. (2015). The influence of pubertal timing and stressful life events on depression and delinquency among Chinese adolescents. *PsyCh Journal*.
- Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. *Psychological Methods*, 1(3), 236-248.
- Flett, G. L., Vredenburg, K., & Krames, L. (1997). The continuity of depression in clinical and nonclinical samples. *Psychological Bulletin*, 121(3), 395-416.
- Ge, X., Conger, R. D., & Elder Jr, G. H. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental Psychology*, 37(3), 404-417.
- Ge, X., Kim, I. J., Brody, G. H., Conger, R. D., Simons, R. L., Gibbons, F. X., & Cutrona, C. E. (2003). It's about timing and change: Pubertal transition effects on symptoms of major depression among African American youths. *Developmental Psychology*, 39(3), 430-439. doi: 10.1037/0012-1649.39.3.430
- Ge, X., Natsuaki, M. N., & Conger, R. D. (2006). Trajectories of depressive symptoms and stressful

- life events among male and female adolescents in divorced and nondivorced families. *Development and Psychopathology*, *18*(1), 253-273. doi: 10.1017/S0954579406060147
- Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology*, *107*(1), 128-140. doi: 10.1037/0021-843x.107.1.128
- Heller, A. S., Johnstone, T., Peterson, M. J., Kolden, G. G., Kalin, N. H., & Davidson, R. J. (2013). INcreased prefrontal cortex activity during negative emotion regulation as a predictor of depression symptom severity trajectory over 6 months. *JAMA Psychiatry*, *70*(11), 1181-1189. doi: 10.1001/jamapsychiatry.2013.2430
- Huang, J.-F., Wong, R.-H., Chen, C.-C., Mao, I.-F., Huang, C.-C., Chang, W.-H., & Wang, L. (2011). Trajectory of depression symptoms and related factors in later life—a population based study. *Journal of affective disorders*, *133*(3), 499-508.
- Kuchibhatla, M. N., & Fillenbaum, G. G. (2011). Trajectory classes of depression in a randomized depression trial of heart failure patients: A reanalysis of the SADHART-CHF trial. *The American Journal of Geriatric Pharmacotherapy*, *9*(6), 483-494.
- Kuchibhatla, M. N., Fillenbaum, G. G., Hybels, C. F., & Blazer, D. G. (2012). Trajectory classes of depressive symptoms in a community sample of older adults. *Acta Psychiatrica Scandinavica*, *125*(6), 492-501.
- Kuo, S.-Y., Chen, S.-R., & Tzeng, Y.-L. (2014). Depression and anxiety trajectories among women who undergo an elective cesarean section. *PloS one*, *9*(1), e86653. doi: 10.1371/journal.pone.0086653
- Li, X., McGue, M., & Gottesman, I. I. (2012). Two sources of genetic liability to depression: Interpreting the relationship between stress sensitivity and depression under a multifactorial polygenic model. *Behavior Genetics*, *42*(2), 268-277. doi: 10.1007/s10519-011-9506-x
- M'Bailara, K., Cosnefroy, O., Vieta, E., Scott, J., & Henry, C. (2013). Group-based trajectory modeling: A novel approach to examining symptom trajectories in acute bipolar episodes. *Journal of Affective Disorders*, *145*(1), 36-41.
- McManus, B. M., & Poehlmann, J. (2012). Maternal depression and perceived social support as predictors of cognitive function trajectories during the first 3 years of life for preterm infants in Wisconsin. *Child: Care, Health and Development*, *38*(3), 425-434. doi: 10.1111/j.1365-2214.2011.01253.x
- Nesselroade, J. R., & Baltes, P. B. (1979). *Longitudinal research in the study of behavior and development*: Academic Pr.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535-569. doi: 10.1080/10705510701575396
- Sabiston, C. M., O'Loughlin, E., Brunet, J., Chaiton, M., Low, N. C., Barnett, T., & O'Loughlin, J. (2013). Linking depression symptom trajectories in adolescence to physical activity and team sports participation in young adults. *Preventive Medicine*, *56*(2), 95-98. doi: http://dx.doi.org/10.1016/j.ypmed.2012.11.013
- Sun, Y., Tao, F., Hao, J., & Wan, Y. (2010). The mediating effects of stress and coping on depression among adolescents in China. *Journal of Child and Adolescent Psychiatric Nursing*, *23*(3),

第二轮

审稿人意见 1: 被试流失的问题是两位审稿人都关注的问题，特别是审稿人 2 认为本文能否发表的重要问题。这一稿虽然增加了流失率分析及其讨论，但对于这个局限性及其对推论统计的影响的讨论仍显不足。

回应: 谢谢审稿人的建议。本稿在对这个局限性的讨论从样本流失可能发生的原因、样本的代表性以及推论统计的生态效度三个方面进行探讨。详见正文第 12 页。

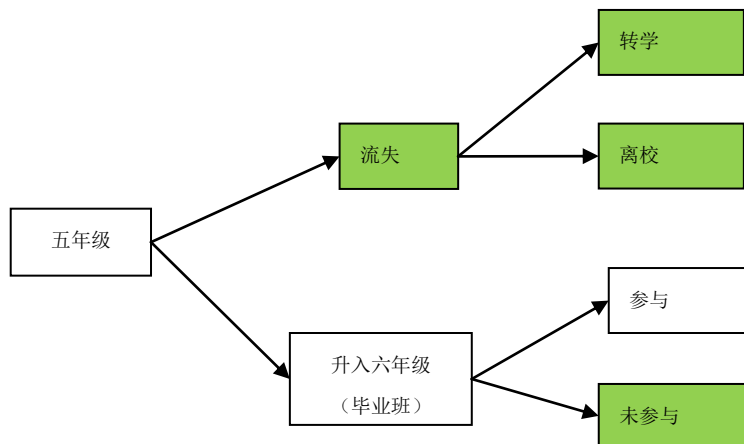
具体分析如下：

W1-W2 的追踪样本是第一轮中的小学四年级、五年级、初一、初二、高一和高二。如下表所示，纵列表示追踪效应，即一轮中的四年级在二轮时升入五年级、一轮中的五年级在二轮时升入六年级，以此类推。

统计结果发现一轮中的五年级、初二年级和高二年级的学生在升入毕业班后，流失样本的抑郁情绪显著高于参与的样本。即，有显著性差异的样本都集中在毕业班（下表中标黄的年级）。

| | | 横断面 | | | | | |
|------|-----|-----|--------------|----|-------------|----|-------------|
| | | 四年级 | 五年级 | 初一 | 初二 | 高一 | 高二 |
| 纵向追踪 | 第一轮 | 四年级 | 五年级 | 初一 | 初二 | 高一 | 高二 |
| | 第二轮 | 五年级 | 六年级 (毕业班) | 初二 | 初三 (毕业班) | 高二 | 高三 (毕业班) |

以一轮中参与调研的五年级学生为例，流失的样本有几种途径：（1）升学时流失，转校或者离校；（2）已升入相应的班级，但未参与调研。总体来说，下图中标绿的均为流失样本。从概率上来说，转学和未参与的概率在各个年级是一样的，那么毕业班独有的特点就是部分学生迫于升学压力而选择提前离校。并且，已有研究证实同年龄段的非在校青少年的抑郁情绪显著高于在校青少年（中国青少年心理健康调查课题组，2013）。



因此，本稿对流失样本局限性的分析主要是从在校学生和同龄非在校生的角度进行的。然而，这只是基于现有研究成果的假设，未来的研究应该纳入非在校的青少年，进一步验证该假设是否成立。

意见 2: 审稿人 2 最关注的另一个问题是关于年龄效应（不同 cohort 组）和发展效应（同一 cohort 组随时间的变化）之间的关系如何分离。作者对这个问题给与了充分的答复，但这些信息并没有完全整合到正文中。

回应: 谢谢审稿人的建议。这部分信息已整合至正文 2.3 数据处理部分，详见正文第 4 页。

第三轮

编委意见:

意见 1: 研究部分未有专门对恋爱因素对青少年抑郁的影响的团体，讨论部分对此不宜做过多引申。此部分讨论中的一句话“恋爱对情绪的不良影响被拉长了”，比较口语化，建议进一步推敲。

回应: 谢谢审稿人的意见。本版修改稿中，讨论部分更多地立足于本研究的发现，并在此基础上引申和探讨。删除了对恋爱问题的过度引申。详见正文第 10-12 页。

意见 2: 对于本研究得到的人际关系对中国青少年抑郁的影响，作者在讨论中提及集体主义文化的作用。还有一个角度是否可以探索一下：本文的被试样本是否多数为独生子女？独生子女在处理人际关系方面常常存在问题，这个因素可能是被试对人际关系压力感受的影响因素之一。与此同时，独生子女的生长环境是否也会影响到被试在心理发展方面的成熟的时间，也可以进一步探索。

回应: 谢谢审稿人的意见，这是一个很好的角度。

现有的研究中，有些研究发现独生子女的抑郁情绪显著高于非独生子女（樊敏杰, 2015; Tseng et al., 1988），也有研究发现非独生子女的抑郁情绪高于独生子女（Yang, Ollendick, Dong, Xia, & Lin, 1995），还有研究未发现两者之间的差异（Falbo & Polit, 1986; Falbo & Poston, 1993; 张樱樱和童辉杰, 2007）。研究者更倾向于将研究间的不一致性归结为样本和测量方面的差异。

本研究中 68% 的样本为独生子女，因此将独生子女与否作为控制变量之一进入模型，结果发现独生子女与否本身对青少年抑郁情绪发生发展的预测作用不显著。在控制了独生子女的影响之后，人际关系压力和学业压力的预测作用依然显著（详情请见正文 P9 的表 5，和 P10 的表 6）。

参考文献:

- Falbo, T., & Polit, D. F. (1986). Quantitative review of the only child literature: Research evidence and theory development. *Psychological Bulletin*, 100(2), 176–189.
- Falbo, T., & Poston, D. L. (1993). The academic, personality, and physical outcomes of only children in China. *Child Development*, 64(1), 18–35.
- Tseng, W.-S., Kuotai, T., Hsu, J., Chiu, J., Yu, L., & Kameoka, V. (1988). Family planning and child mental health in China: the Nanjing Survey. *The American Journal of Psychiatry*, 145(11), 1396–1403.
- Yang, B., Ollendick, T. H., Dong, Q., Xia, Y., & Lin, L. (1995). Only children and children with siblings in the People's Republic of China: Levels of fear, anxiety, and depression. *Child*

Development, 66(5), 1301–1311.

樊敏杰. (2015). “一胎化”政策对健康状况影响机制的实证研究——基于独生子女与非独生子女的比较分析 (博士学位论文). 西南财经大学, 成都.

张樱樱, & 童辉杰. (2007). 独生子女与非独生子女心理健康的元分析. *苏州科技学院学报: 社会科学版*, 23(4), 112-115.