

《心理学报》审稿意见与作者回应

题目：接纳承诺疗法的作用机制——基于元分析结构方程模型

作者：任志洪；赵春晓；卞 诚；朱文臻；江光荣；祝卓宏

第一轮

审稿人 1 意见：心理治疗的机制研究是非常重要的选题。作者针对当前比较流行的 ACT 治疗的作用机制应用元分析结构方程模型的方法，针对业已提出的六中心心理机制进行了元分析检验。文献检索规范，文献质量较高，质量控制方法较严格，所得结果可信，结论和展望对未来该领域的研究具有指导意义。全文结构合理，写作文笔流畅，逻辑性较强，建议发表。

意见 1：该文重点论及 ACT 作用机制，在引言中讨论了 ACT 的归属没有太大必要，反而有文不对题之感。2)针对 CBT 与 ACT 的关系，值得认真思考。目前 CBT 更多的认识是整合心理治疗的一个流派，而不是传统意义上 CBT。文中有所界定会给读者一个更清晰的理解。

回应：感谢专家的建议。我们对该问题的理解：原论述中，我们对 ACT 的归属作了梳理，与专家的第 2)问题关联密切。正如专家第 2)问题所提及，学术界有一种观点认为“目前 CBT 更多的认识是整合心理治疗的一个流派，而不是传统意义上 CBT。”，而“整合心理治疗的一个流派”，正是目前兴起的以 ACT、辩证行为疗法(DAT)和正念减压疗法等为代表的行为治疗“第三浪潮”(Dimidjian et al., 2016)。

然而，对于 ACT(也包括辩证行为疗法(DAT)和正念减压疗法等)是不是属于新 CBT，刚开始是有些争议。比如，ACT 的创始人 Hayes 更是将该治疗看作是继行为治疗、认知治疗之后的“第三代行为治疗”(Hayes et al., 2004)；而原有 CBT 阵营研究者则把 ACT 看作是在 CBT 基础上的一种延展(Hofmann & Asmundson, 2008)。

梳理行为治疗的发展脉络后不难发现：行为治疗被喻为“第一浪潮”，它基于条件反射和新行为原则，直接关注有问题的行为和情绪。而“第二浪潮”则强调对非理性思维、病理性认知图式或错误的信息处理方式进行矫正，进而减轻或消除症状，这种以认知改变为重心的治疗被称为“认知行为疗法”(CBT)，当然，CBT 并非放弃行为治疗，而是在行为治疗的基础上强调认知改变的重要性。

尽管大量的研究表明 CBT 治疗多种心理疾病的有效性，仍有学者对 CBT 的真实有效性提出了一些置疑。比如，有研究者指出了三个在 CBT 效果研究文献中“实验研究的异常现象”(Hayes, 2004; Longmore & Worrell, 2007)：其一，成分分析显示，认知干预成分并没有提升治疗效果；其二，CBT 治疗通常伴随着早期症状快速缓解，而这发生在实施认知干预技术之前；第三，以认知为中介变量的改变，并不能预测之后症状的改变。Longmore 和 Worrell(2007)等人通过文献分析，也支持了上述观点。也就是说，CBT 的有效机制正受到挑

战。

幸运的是, 认知行为治疗也在不断发展, 特别是逐渐融合了正念和接纳技术(Hayes et al., 2010)。这些治疗包括接纳与承诺治疗、辩证行为治疗和基于正念的认知治疗, 正逐渐发展行为治疗的“第三浪潮”(Hayes et al., 2010)。

已有研究者对 ACT 与 CBT 的区别与联系作了系统梳理(任志洪, 2012; Hayes, 2016), 简述如下:

(一)ACT 与传统 CBT 的哲学基础不同

传统 CBT 通常基于机械主义哲学角度, 把健康心理功能看作是没有病理性思维。认知通常被比喻为一台机械, 机械的哪一部分出问题, 就对该部分进行修正, 以使整个(人)的机能有效发挥作用。比如说, “我必须赢!”是个错误的信念, 自动化思维, 或有缺陷的图式。治疗师试图去教会当事人以新的、更为合理的认知, 即“我想赢”, 替换这些错误或不合理的认知(Bach & Moran, 2008)。此外, 思维可被看作是情绪或行为的前因, 即不合理思维导致负面情绪评价和功能失调行为。它假定, 当思维更为理性时, 情绪和行为也将获得改善。

ACT 则是将功能情境主义哲学作为其发展的重要基础。功能情境主义与机械论相对立, 其基本前提是实用主义, 也就是说没有“正确”的解决方式, 只有“有用”的方式。其主题是“基于情境”, 强调理解事物与分析问题必须动态地考虑整个事件及其发生的背景, 把行为、刺激和结果作为一个整体单元来考察(Hayes et al., 2010)。ACT 想去理解有机体与历史和情境之间的交互作用, 因此无法单独去解释相互作用中有机体的某一特定部分, 所以为了促使整体(人)健康, 没有必要去处理部分图式或错误信念(Bach & Moran, 2008), 并不把思维看作是行为和情绪的前因。事实上, 思维被认为是在个人独特历史背景下的隐蔽行为, 这与由环境事件引起的公开行为类似, 而分析单元就成为了“环境——环境”的关系。简单来说, 行为(包括公开与隐蔽)都必须在情境中去理解它的发生。基于情境主义观点, 痛苦可以是正常的也可能是有问题的, 而健康并不是没有痛苦。

(二)二者在认知和情绪的作用上有区别也有联系

(1)认知在改变机制中的作用

认知过程构建了 CBT 的核心。认知是思维过程的代名词, 包括自动思维和图式, 它是对世界、自我和未来的一般信念。这些图式决定了个体一般采用的“规则”, 它们往往通过“应该、当然、必须”这种较绝对的形式来表达。当患者把这种非理性的期望强加给自己、他人或这个世界时, 往往会带来不必要的情感困扰。CBT 帮助患者去识别、改变和重新评价这些僵化的规则, 采用一种更为宽松和令人满意的价值系统, 以提升整体的生活满意度。因此, CBT 与 ACT 的工作模式在价值问题这一点上, 具有相似性。

然而, ACT 与 CBT 在治疗中对认知的处理具有重要区别。相较之 CBT, ACT 并没有采用认知三成分模式来处理外显行为(行动)、情绪(主观体验)和认知(思维过程)的区别。相反, ACT 更倾向于把认知看作行为分析中所表示的一种普遍“行为”。从本质上说, “认知”一词在 ACT 与 CBT 中具有不同的意义, 即在 CBT 中认知是思维过程; 而在 ACT 中是个体的私人行为。

(2)情绪在改变机制中的作用

CBT 与 ACT 在处理降低情绪困扰上, 具有许多相似的技术。二者在方法上都是问题聚焦和基于行为的干预, 包括建立良好的咨询同盟, 确定清晰的、可解决的目标。然而, 二者之间有一个重要的不同之处, 即治疗中具体的情绪管理策略。ACT 通过抑制(Suppression)和其他功能失调情绪管理策略处理体验回避和管理不愉快情绪。相反, CBT 主要是关注诱发情绪的刺激本身, 即引起情绪体验的刺激或事件。换言之, CBT 技术主要是先行关注(antecedent-focused)情绪调节, 主要(但不完全是)通过对情绪刺激的重新评价, 来促使对先行关注情绪调节策略的自适应; 而 ACT 主要是反应关注(response-focused)情绪调节, 通过关注不适应的反应关注策略, 来阻止情绪抑制。

综上所述, ACT 疗法与传统 CBT 在哲学基础和对认知和情绪的作用机制上, 有较为明显的不同。

回到专家的问题。简言之, 为什么我们要讨论 ACT 的归属(专家问题 1)? 因为新兴的系列“第三浪潮”疗法, 最开始有些并不承认自己是 CBT 的阵营, 比如 ACT, 上述也详细说明了二者确实有所区别的。几年的磨合之后, 也就有了专家的第二个问题: “目前的观点认为 CBT 是整合心理治疗的一个流派?”, 这句话应该是基于 Hayes 提出的“基于语境的认知行为疗法”, 以说明目前对“第三浪潮”的共识和与传统 CBT 的区别。

因而, 为了更好说明上述问题, 我们对正文的修改, 主要体现两点:

1)补充说明了第一、二、三浪潮的基本特征, 并定义 CBT(传统), 以区分行为治疗的第三浪潮; 2)删除 ACT 归属争议, 强调了 Hayes 提出的“基于语境的认知行为疗法”, 以说明目前对“第三浪潮”的共识和与传统 CBT 的区别。

详细修改如下:

近年来, 新兴的行为治疗“第三浪潮”备受关注。行为治疗被喻为“第一浪潮”, 它基于条件反射和新行为原则, 直接关注有问题的行为和情绪; “第二浪潮”则强调对非理性思维、病理性认知图式或错误的信息处理方式进行矫正, 进而减轻或消除症状, 这种以认知改变为重心的治疗被称为“认知行为疗法”(cognitive-behavior therapy, CBT)(Beck, 1993)。而“第三浪潮”则对心理现象的语境和功能更为敏感, 而不仅仅关注其形式(Hayes, 2016), 典型的疗法包括接纳与承诺疗法、辩证行为疗法(dialectical behavioral therapy), 正念认知疗法(mindfulness based cognitive therapy), 慈悲聚焦疗法(compassion focused therapy)(Hacker et al., 2016); 还包括功能分析心理治疗(functional analytic psychotherapy)和整合行为夫妻治疗(integrative behavioral couples therapy)、图式治疗(schema therapy)和元认知治疗(metacognitive therapy)等疗法(Dimidjian et al., 2016)。

尽管研究者们对哪些疗法属于“第三浪潮”仍存在争议, 但为了促进疗法的共同发展, 共识多于分歧。Hayes 等人 2011 年对“第三浪潮”进行了修正, 用“基于语境的认知行为疗法”(contextual cognitive behavioral therapy)作统称, 将关注点从原来的哪些治疗方法应该被纳入“第三浪潮”, 转为强调新兴疗法在理论、治疗过程和程序上“开放, 主动和觉察”的特点

(Hayes, Villatte, Levin, & Hildebrandt, 2011)。在这些基于语境的认知行为治疗中, ACT 最受关注, 近年来被引用的频次最多(Dimidjian et al., 2016)。本研究的目的在于系统考察 ACT 的作用机制, 相较之传统 CBT 的特异性, 以及在网络环境中的可迁移性。

意见 2: 在文中谈及 ACT 在机制上较 CBT 具有优势, 具体是哪些优势? 这里的 CBT 是指什么? 有所讨论是否更好?

回应: 感谢专家的建议。

“ACT 在机制上较 CBT 具有优势”该句是指“4.讨论”第二段的首句吗? 我们原句是“ACT 在其所假设的改变机制上较 CBT 具有优势”。

因我们在“1.3 ACT 的作用机制”论述了 ACT 区别于 CBT 的核心机制, 所以在“4.讨论”部分和论点句未作详细铺陈。ACT 所假设的作用机制为: ACT 通过接纳(愿意接触内心体验)、解离(将认知体验为持续的过程, 而非认知过度调节行为)、以己为景(把内在体验作为自身体验的背景, 而不把它看作是体验本身)、此时此刻(能够灵活地接触发生的内部与外部事件, 不作判断)、价值(选择持续行为模式所需的结果, 以建立强化物)、承诺行动(灵活地朝有价值的方向作出行动)等六大成分(Hayes et al., 2006; 曾祥龙, 刘翔平, 于是, 2011; 张婧, 王淑娟, 祝卓宏, 2012), 提升心理灵活性。

当然, 此处的原文我们可能没有论述清晰, 因此, 在正文中作如下两点修改:

ACT 在其所假设的改变机制上, 即心理灵活性及其所包含的 6 大成分, 较传统 CBT 具有优势。另, 如前所述, CBT 是指传统 CBT, 已在全文进行校正。

意见 3: 在所选研究中, 纳入了不同的研究对象和不同的精神障碍或心理问题, 在这些问题中具体的发生和维持机制上可能会有不同, 放在一起来分析, 对研究结果是否会有影响? 值得思考。

回应: 感谢专家的建议。

在方法上, 元分析方法本身一直存在的“苹果和橙子问题”(apple and orange problem) (Wolf, 1986)。所纳入的实证研究, 其研究对象、精神障碍或心理问题类型不尽相同, 各项实证研究的具体操作和数据收集方法具有差异。但纳入元分析的实证研究也具有共性, 都关注 ACT 对精神障碍或心理问题的治疗效果, 这些共性是元分析得以进行的基础。

在元分析中, 专家所提及的此类问题较难避免, 同行大部分也较认可对不同心理问题干预效果量的合并, 比如, 类似地, 对 CBT 的干预效果元分析(Butler, Chapman, Forman, & Beck, 2006)(被引超过 3000 次)。当然, 虽然干预问题不同, 但我们对其结果变量的选择都限定在心理问题的测量上。正文方法部分 2.3.1 所述: “我们借鉴前人的系统选择方法(Gu et al., 2015): 优先选择心理病理的整体测量, 其次选择抑郁和焦虑测量结果……”

当然, 为了让读者也意识到专家所述问题, 我们在研究局限上作如下补充:

(3)所纳入的实证研究, 其研究对象、精神障碍或心理问题类型不尽相同, 各项实证研究

的具体操作和数据收集方法具有差异。这也正是元分析一直存在的“苹果和橙子问题”(apple and orange problem); 但纳入元分析的实证研究也具有共性, 都关注 ACT 对精神障碍或心理问题的治疗效果, 这些共性是元分析得以进行的基础。

意见 4: 注意文中个别错字和笔误。

回应: 文中批注“疼痛障碍”漏掉一个“疼”字, 已补充。其他地方全文校正以蓝色字体体现修改。

.....

审稿人 2 意见:xb18-290 一稿中, 作者使用 Meta-analytic structural equation modeling(MASEM)来综合已发表研究中的数据, 探索 ACT 治疗中, ACT 理论中各个成分/因素是否对治疗效果起到中介作用。工作量较大, 在心理咨询与治疗方面有一定的理论意义。在阅读本稿件中, 本人有如下想法, 与作者商榷:

意见 1:在中介分析中, 作者试图以 ACT 的 6 个成分/因素与心理灵活性作为中介, 以治疗手段作为二分的自变量/预测变量, 以治疗前后心理症状的改变为因变量/被预测变量。同时作者也指出, “心理灵活性作为 6 大 ACT 作用机制的统称”(p12)。这让我有一个疑惑: ACT 疗法、心理灵活性与其他的 ACT 成分、以及治疗效果之间应该是怎么的逻辑关系? 也就是说在中介分析中, 将心理灵活性与 6 个 ACT 成分作为同一水平/层次上的变量是否合适? 是否存在更加复杂的模式: 治疗→某个 ACT 成分 → 心理灵活性 → 治疗效果, 或者是: 治疗 → 心理灵活性 → 某个 ACT 成分变化 → 治疗效果。

回应: 感谢专家很有启发的建议。

理论上确实可能存在专家所假设的作用机制。从理论假设上来说, 6 个成分应是一阶因素, Hayes 用心理灵活性一词作为 6 大机制的统称, 那么心理灵活性应该是二阶因素。所以, 可能存在 ACT 治疗→某个 ACT 成分 → 心理灵活性 → 治疗效果, 或者是: ACT 治疗 → 心理灵活性 → 某个 ACT 成分变化 → 治疗效果。

但在实际研究中却存在困境。从测量上说, 最早仅有测量整体心理灵活性的接纳和行动问卷(Acceptance and Action Questionnaire, AAQ), 在较多研究中仅使用 AAQ 测量 ACT 整体的作用机制。随后几年虽然也针对 6 大成分发展了相应的测量工具, 但正如我们在“5 临床研究启示”中所述: “现有研究主要使用的 AAQ 和 AAQ-II 更趋向于测量全局的心理灵活性, 而对 6 大具体核心机制的检验较少。”然而, 元分析是基于已有实证研究的再分析, 而在纳入的实证研究极少同时测量了心理灵活性及 6 大成分的前提下, 无法计算出变量之间的协方差矩阵, 因此在本研究中, 我们无法去检验更为复杂的 ACT 作用机制模型。

考虑到该问题确实在后续研究中有重要的意义, 我们把这一点, 补充在“本研究局限”中, 以期后续实证研究能对该问题有所觉察。正文补充如下:

(5)从理论假设上来说,6个成分应是一阶因素, Hayes 用心理灵活性一词作为6大机制的统称,那么心理灵活性应该是二阶因素。所以,可能存在更为复杂的机制模型,比如,ACT 治疗→某个 ACT 成分 → 心理灵活性 → 治疗效果,或者是:ACT 治疗 → 心理灵活性 → 某个 ACT 成分变化 → 治疗效果。然而,元分析是基于已有实证研究的再分析,而在纳入的实证研究中极少同时测量了心理灵活性及6大成分,因此在本研究中,我们无法检验更为复杂的 ACT 作用机制模型。

意见 2:作者在纳入标准中,包括了“对中介变量进行前后测量评估”,但是是否包含对这些中介变量、治疗效果测量本身的合理性进行评估?也就是说,在纳入的研究中,对中介变量(如心理灵活性)的测量工具是否有足够好的信效度?治疗结果(心理症状)的测量工具是否有足够好的信效度?对中介变量和治疗结果测量中,是否存在共同测量偏差?

回应:根据我们的理解,专家上述问题涉及到测量工具信效度和共同测量方法偏差两方面问题。我们分开解释。

其一,测量工具信效度问题。诚如专家如言,测量工具本身的信效度对研究结果会有一些影响。但我们面临一些“无力”:关于测量工具本身的信效度问题。因原始研究使用的都是标准化的测量工具,每个都发表在同行评议期刊上,有相应的信效度检验报告,而量表本身的信效度不是本研究最关注的核心问题,同时元分析无力改变已经发生的事实。我们也曾考虑到报告量表的信效度,但是好像在本研究中对后续分析也没有实际的使用价值。鉴于同类元分析结构方程的研究(Gu, Strauss, Bond, & Cavanagh, 2015)对该问题都采取“默认”了原始研究使用的测量工具是有效的(事实上,使用低信效度的测量工具也较难被发表),我们也未进一步深究。如审稿专家对该问题有好的处理办法,也恳请告之,我们再作处理。

其二,共同测量方差偏差问题。

(1)就原始研究的共同测量方法偏差而言

因中介变量和治疗结果变量测量方法相同可能存在共同测量方法偏差,研究者建议(周浩,龙立荣,2004)对共同方法偏差的检验和控制。

研究者首先应该考虑采用程序控制,因为这些方法是直接针对共同方法偏差的来源而设计的。对测量进行时间上、空间上、心理上、方法上的分离,保护反应者的匿名性、减小对测量目的的猜度,平衡项目的顺序效应以及改进量表项目等。纳入的原始研究中,极少部分的研究描述了对共同方法偏差在程序上进行的控制,仅有少部分研究在中介变量和结果变量的测量时序上进行控制。这一点,我们在“3.1 纳入文献基本描述”部分作了描述:“较少研究考虑机制变量时序作用,仅有9篇进行了多点测量,但大部分多点测量使用的是前测、后测和追踪测量(Luciano et al., 2014; Stafford-Brown & Pakenham, 2012; Wetherell et al., 2011; Yadavaia et al., 2014),追踪是在治疗结束后,并非在有效治疗阶段……”

但是,在某些研究情境中,受条件限制,上述的程序控制方法无法实施,或者无法完全消除共同方法偏差,这个时候就应该考虑在数据分析时采用统计方法检验和控制共同方法偏差。

可我们检查已有研究文献发现，纳入文献都没有检验共同方法偏差(即使是高影响因子文献)。

(2)就元分析的共同测量方法偏差而言

就我们所知，目前还没有元分析的研究者关注该问题，在本研究中我们也无法解决该方法学上的问题。因此，我们在“本研究局限”中对该问题作了说明。

正文修改：

(8)MASEM 方向本身具有一定的局限性，它仅是用来检验可能的作用机制，无法证明其因果关系；且应该注意到，原始研究中其中介变量和治疗结果变量测量方法可能存在共同测量偏差，本研究结果可能也受其影响。

意见 3：与上一个问题相关，每个研究的治疗结果测量中均选择了 1 个指标，是否能够清楚地标注出来(例如使用加粗的字体)

回应：感谢专家的建议。我们在提供附件中单独作为一列体现。

意见 4：结果部分报告，有 21 个研究中包括了多个中介变量的测量，但结果中对各个 ACT 成分分开进行中介分析，是否有一些文章包含多个共同的中介变量(如都包括某两个或者三个中介 ACT 成分)? 这些文章是否能够一起进行 MASEM 分析分析，以检验不同的 ACT 成分之间的交互作用。

回应：感谢专家的建议。

一方面是，考虑到一些专家的训诫，研究应该以理论先行而非统计先行。2-3 个成分的交互在 ACT 治疗模型中好像没有实际的理论意义。另一方面，在原始研究中，研究者采用多个中介测量也只是单纯考察每个中介变量的显著性水平，未有检验交互作用，因而无法在元分析中获取变量之间的交互作用的数据。

当然，基于理论先行考虑，确实有一 ACT 扩展理论考虑到 6 大成分关联的。正如我们在“5 临床研究启示”第(2)中所描述的：“ACT 的 6 大机制可以分为三大模块(Villatte et al., 2016)：接纳和认知解离属于“开放”模块，目的是降低思维、情感和感觉的有害反应；价值和承诺行动属于“行动”模块，侧重强化动机和增加有意义行为；而接触当下和以己为景主要是为了促进自我觉察，在“开放”和“行动”模块中都包含，但并没有特意强调。就纳入本元分析的文献而言，多数研究聚集的是“开放”模块，而对“行动”模块的关注较少，仅有 6 篇研究测量了价值。”我们原打算基于该理论构建“开放”—“行动”的检验模型，可惜原始研究中鲜有同时检验该理论的数据($K < 3$)，因而并未作进一步分析。

意见 5：结果部分关于 ACT 与 CBT 比较以及基于网络传播机制检验的中介分析示意图是如何的？中介变量是哪些？这些在方法与结果部分均未进行说明，无法判断方法是否合适。假如我理解正确，在 ACT 与 CBT 的比较之中，使用 ACT v. CBT 作为预测变量 X，某些 ACT

成分作为中介变量 M，治疗的结果作为 Y。这里中介的 ACT 成分是否同质？同样的问题也存在于对网络治疗的分析之中。

回应：感谢专家的建议。

确实存在专家所说的问题，第一稿中，我们确实把纳入 ACT 某一成分的都作为中介变量进行分析。虽然该 6 大成分的测量都指向“心理灵活性”，某种程度上来说，也是有“同质”的理论基础，但考虑到成分本身之间有所区别，直接合并效果量恐抵消了可能的异质性。因此，在 ACT 与 CBT、传统干预相对于网络化干预的编码和计算中，我们都以“心理灵活性”作为唯一纳入中介指标，重新分析了相关数据。详见正文“3 结果”部分。

意见 6：与元分析报告规范相关的疑问。元分析在实证研究中占据很高的地位，不少研究者甚至将其作为证据力最强的研究。但不可否认的是，元分析也可能受到“可重复危机”的影响。元分析过程中，研究者需要做出许多决定，这些决定会影响到元分析的结果，但是读者/审稿人往往无法得知。正是这些原因，导致了同一研究问题的不同元分析可能得到相反的结论，例如：

<https://www.sciencemag.org/news/2018/09/meta-analyses-were-supposed-end-scientific-debates-often-they-only-cause-more>。根据最近国内多位专家撰写的关于元分析文章报告标准(预印本，即未经同行评议，见：<http://www.chinaxiv.org/abs/201809.00177>)，本稿件的方法与结果部分报告可能还有如下几点可以报告，供作者参考：

回应：感谢专家提供的参考文献，我们尽力借鉴。

意见 7：元分析的方案是否进行预注册？在《学报》的自检报告中，鼓励作者对自己的研究进行预注册，这包括元分析。而且目前国际上已经有了元预注册的模板 <http://www.prisma-statement.org/Extensions/Protocols.aspx>。如果能够说明是否进行预注册，将对读者有所帮助。

回应：感谢专家的提醒。因该元分析工作始于 2016 年，成稿于《学报》新自检报告发布之前，因此未能根据建议预注册。另，通常看到的元分析文献，都是有预注册交待预注册链接或信息，未预注册的并没有特意说明。所以，我们在正文部分找不到合适的地方来说明该问题。

意见 8：文献搜索，是否搜索了中文数据库？

回应：感谢专家的提醒。

我们检索过中文数据库，因仅有一篇我们自己的 ACT 实证研究符合纳入条件，为了避免偏差，本研究仅考虑了英文数据库。ACT 进入中国也不久，干预研究对咨询师素质要求较高，期待不久的将来能得以更好地发展。

意见 9：文献筛选：是否由 2 位或者以上的作者进行筛选？一致性如何？

回应：感谢专家的提醒。文献筛选过程使用 Noteexpress 管理，过程可重复，因此由第二作者先筛选，第三作者核对，有异议文献与第一作者协商一致解决。正文也作如上补充说明。

意见 10：使用 R 语言和 metaSEM，应该进行引用并且说明相应的版本

回应：感谢专家的提醒。已正文补充相应版本号。

使用 R 语言(Ver.3.5.2)中的 metaSEM 包(Ver.1.2.0)进行 TSSEM 分析(Cheung, 2015)。

意见 11：数据摘取过程中，两位作者独立进行摘取的结果的一致如何？

回应：感谢专家的提醒。初稿中，我们采用第二作者摘取，第三作者核对的形式，经专家提醒，我们采用更严格的方式重新进行编码。具体为：对数据摘取分两类：一类是描述研究特征的基础数据，我们采用第二作者摘取，第三作者核对的形式；另一类是真正纳入统计分析的核心数据，我们采用第二作者和第三作者分别编码，求得评分者一致性信度 kappa 系数为 0.89，根据在 0.75 及以上认为一致性非常好的判别标准(Orwin, 1994)，说明本研究编码具有较高的一致性。最后，与第一作者协商一致后确定最终编码。

正文也作如上补充说明。

意见 12：Nfs 的值是多少？能够在文章中提供？另外，是否可以使用 funnel plot 让读者直观地判断出版偏见？

回应：感谢审稿专家的提醒。根据最近一项元分析归纳(Turner, 2013)：在元分析中，出版偏差是难于避免的。出版偏差源于研究者、审稿人和编辑更倾向于接受阳性结果的研究。而“偏差”一词涉及选择性出版(比如 file drawer problem)，或选择性报告研究结果，同样也包含其他现象，包括延迟出版、重复发表、出版地偏差和引用偏差等。

虽然元分析中出版偏差难于避免，但研究者试图用一些统计指标考察可能的出版偏差及影响。失安全系数(fail-safe N , N_{fs})和 funnel plot 都是“评价”出版偏差的常用指标。funnel plot 虽然直观，但也比较主观，在本研究中 funnel plot 图 3 (变量)* 7(机制)=21，学报也不允许同时呈现 21 张的图示；失安全系数(fail-safe N , N_{fs})也存在使用局限：“One needs to remember that a fail-safe calculation is neither a method of identifying publication bias nor a method of accounting for publication bias that does exist. It is simply a procedure by which one can estimate whether publication biases (if they exist) may be safely ignored(Rosenberg, 2005).”

因此，它并非是识别出版偏差的方法，其使用逻辑应当是假设存在偏差的情况下，考察需要多少阴性研究才能推翻已有结论，即，考察的是其结果的稳健性。

此外，因 metaSEM 分析的特殊性，目前很难找到一种合适的统计方法考察其可能的出版偏差。我们在初稿中采用了传统元分析的范式，报告了上述这两项指标。后来，参考同类元分析(Gu et al., 2015)发现，metaSEM 关注多变量的关联(比如本研究中的 X,M, Y)，而非传统元分析的单结果效果量(Y)，鲜有同类研究分别统计和报告每个变量的偏差指标。

鉴于上述考虑，我们拟参考同类研究的报告范式(Gu et al., 2015)，不在正文中报告失安全系数(fail-safe N, N_{fs})和 funnel plot，但在补充材料中额外提供传统元分析三种出版偏差的评估结果，包括失安全系数(fail-safe N, N_{fs})、Begg and Mazumdar rank correlation 和 Egger's regression intercept(见补充分析材料，仅 ACT vs.CBT 的 c 路径可能存在出版偏差，其他参数正常)。虽然我们仍然认为现有这些评估指标并不一定适用于本研究，但此作法应当算是目前满足多方需求较为合适的权宜之计。

正文修改如下：

我们先使用 Begg and Mazumdar rank correlation 和 Egger's regression intercept 评估可能的出版偏差，如果存在偏差，我们进而使用失安全系数(fail-safe Number, N_{fs})检验可能的出版偏差对效果量的影响(Rothstein, Sutton, & Borenstein, 2005)。……本研究中，Begg and Mazumdar rank correlation 和 Egger's regression intercept 分析结果显著的变量，多数的 N_{fs} 值都较大(仅 ACT vs.CBT 的 c 路径可能存在出版偏差)，整体上说，本研究结果具有较强的稳健性(补充分析材料额外提供)。

意见 13：用于元分析的数据与分析代码是否能够存放在公开的数据存在平台(osf.io)或者元分析专用的平台(<http://metalab.stanford.edu/>)

回应：感谢专家的建议。目前我们对论文还未被接受的状态下，提前公开数据，仍心存疑虑。但如有需要，一方面，我们现在可以提供所有数据和分析代码给审稿专家审核(因学报平台限制，只能上传 word 文档，我们把数据另放在“修改说明”中，另上传)；另一方面，如该论文被接收，我们在正文备注感兴趣读者可以联系到作者无偿获取研究中的数据和分析代码。

此外，元分析的特征编码我们已作为附录。当然，如果学报有要求需要把数据分享在公开数据平台，我们也乐于遵从规则。

意见 14：虽然本文主要目的不是讨论 ACT 的治疗效果，但是这个作为最主要的结果，可能也需要在结果中进行报告。

回应：感谢专家的建议。

正如专家所言,本研究的主要目的不是讨论ACT治疗效果,因此未作为主要结果。其一,ACT的治疗效果在众多相关的元分析已作全面的报告,因我们的关注点是作用机制,所以纳入的仅是包含相关机制测量的文献,并不能全面体现ACT的治疗效果,因此我们也觉得失去了报告的意义;其二,我们在表1中报告了r值矩阵,可方便转化为d值。故,为了聚焦研究问题,我们参考了同类元分析结构方程模型的作法(Gu et al., 2015),并未直接在正文报告ACT干预的合并效果量。当然,如前所述,我们可以分享相关数据(包括d值)供感兴趣的研究者参考。

意见 15: 其他小问题。本文在多处引用文献可能不太恰当。例如,在第10页倒数第二段中,“所有我们MASEM分析的模型具有足够的统计力拒绝事实不成立的零假设(Montazemi & Qahri-Saremi, 2015)”。这个引用并不准确, Montazemi and Qahri-Saremi (2015)并非方法类文章,而是使用metaSEM的一篇文章,采用了同样的power analysis。在该文中,引用了A.R. Montazemi, J.J. Pittaway, H. Qahri Saremi, Y. Wei, Factors of stickiness in transfers of know-how between MNC units, *The Journal of Strategic Information Systems*, 21 (2012) 31-57。但这个文章也不是方法类文章,只是做了同样的power analysis,在这个文章的power analysis部分,作者引用了Baroudi and Orlikowski, 1989, *The Problem of Statistical Power in MIS Research*; Cohen, 1988, *In Statistical Power Analysis for the Behavior Sciences*。对于这种情况,要准确的引用,需要引用Cohen或者Baroudi & Orlikowski, 1989。当然,作者需要仔细阅读Cohen或者Baroudi & Orlikowski的文章是否合适。

回应: 因Cohen(1988)的论著年代较久远,已为“公知”概念,且我们引用的文献有更为直接的关联,也考虑到读者查阅的便利性,故直接引用了同类研究文献。感谢审稿专家的提醒,现同时补上原始引用文献。

意见 16: 第21页,提到的新方法,均无参考文献,如果能够补充,对读者会更有帮助。

回应: 感谢专家的提醒,我们补充了引用。

然而,近年来新发展的纵向中介模型分析技术(Grimm, Ram, & Estabrook, 2017),比如潜变量增长曲线模型(Latent Growth Curve Models)、潜变量变化分数模型(Latent Change Score Models)和多水平结构方程模型(Multilevel Structural Equation Modeling)也应值得尝试。

再次感谢两位审稿专家的辛苦付出和专业建议。

第二轮

审稿专家 2 意见：作者(们)很好地处理了本人在一审中的意见。如本人在一审中所述，本研究工作量较大、有较大的意义，推荐发表。

回应：感谢审稿专家的积极肯定。

编委意见：修改稿很好地回应了审稿专家的意见。还有如下小建议：

意见 1：出版偏差建议在结果正文中报告。鉴于漏斗图有很多张所占看见太多，建议做一张表给出失安全系数。

回应：感谢编委专家的建议。考虑到版面限制，我们没有另再加表，经研究，我们已在表 1 中增加了失安全系数的报告。另，出版偏差的其他考察指标，比如 Begg and Mazumdar rank correlation 和 Egger's regression intercept，我们拟在“审稿意见”中作为补充分析材料额外提供。

意见 2：元分析用到的文献，建议在前加*与其它参考文献加以区别。

回应：感谢编委专家的建议。区别于传统元分析纳入参考文献前加*号的标法，本研究较难实现。一方面是因版面限制，另一方面是本研究关注的核心问题为 ACT 作用机制检验，因此，基于行文简练考虑，我们参考同类研究报告形式(Gu et al., 2015)，并没有在正文中报告纳入文献的基本信息表(该表作为附录 1 提供了)。故，元分析纳入的文献绝大多数并没有包含在正文的参考文献中，所以我们作为附录 2 另行提供了完整的纳入文献(发表后放在“审稿意见”中提供)。编委专家如果觉得在正文中插入所有纳入分析文献较为合理(这可能导致引文破百篇)，我们再作处理。

意见 3：目前对 Sobel 检验的批评，建议作者在局限中予以讨论。

回应：感谢编委专家的建议。

因元分析结构方程模型(MASEM)方法学局限，在该文中我们无力在方法学上有所贡献，因此只能用经典 Sobel 法检验。根据编委专家的建议，我们在局限中给予说明该不足。

正文讨论部分修改如下：

(8)MASEM 方向本身具有一定的局限性，特别是本研究对作用机制的考察报告了 Sobel 经典中介检验结果，该法近年来也受到正态分布假设和统计功效较低的诟病(MacKinnon, Fairchild, & Fritz, 2007)；且应该注意到，原始研究中其中介变量和治疗结果变量测量方法可能存在共同测量偏差，本研究结果可能也受影响。

意见 4：还有个别错字(不常用词汇)，如“打断消极心思维”31 页 1 行。

回应：感谢编委专家的提醒。因多人审校，最后合并修改稿时多出了一个“心”字，现已删除，并对全文再作全面审校。

意见 5：作者的文字仍有压缩精简的空间。建议修后发表。

回应：感谢编委专家的肯定。我们对全文进一步压缩。主要修改如下：

目前该 Wrod 文档有 55 页，正文为第 17-32 页，共 15 页，其他为回应、自检报告、参考文献和附录。作为一篇元分析和系统综述论文，应算是非常简约。

当然，为了节省版面，给读者呈现更纯净的研究报告，我们对全文进一步进行精简。原正文 14710 字，删除 2084 字，精简后正文 12626 字。

具体来说，正文精简说明：

(1)引言：原 4309，删除 586 字，精简后为 3723 字。主要精简了对 ACT 干预效果和机制研究的冗余描述。

(2)方法：原 2856，删除 633 字，精简后为 2223。主要精简了两阶段结构方程元分析方法学上的冗余描述。

(3)结果：原 2915 字，无精简。

(4)讨论：3963，删除 559 字，精简后为 3404(符合学报 3500 字以内的要求)。主要删除了元分析方法学本身存在的诸多共性局限。

(5)其他内容精简描述若干字。

再次感谢编委专家的批评指正。