

补充材料 1: PRIOR-MA 与核心方法部分条目阐释

开放与可重复荟萃分析报告清单 (PRIOR-MA)¹

章节/主题	#	清单条目
标题		
	1	陈述研究问题，并表明研究类型是系统综述、荟萃分析或是两者兼具。
摘要		
		提供结构化的小结，包括：
		2.1 目标（研究的问题或假设）；
		2.2 文献纳入标准（描述纳入研究的自变量、因变量与研究设计等特征）；
结构化小结	2	2.3 研究综合的方法（用于综合和比较不同研究使用的统计或统计以外的方法，例如效应量的测量，平均法或进行异质性分析时使用的模型）；
		2.4 结果（纳入研究与被试的数量及其特征，相关效应量与置信区间等）；
		2.5 结论（描述文章中存在的优势与局限）。
引言		
		陈述该研究领域相关的问题及其关系，包括：
问题提出	3	3.1 历史背景，已有的系统综述与荟萃分析持有的观点；
		3.2 该研究可能涉及的理论、政策以及实际应用中的议题；
		3.3 该研究可能涉及的相关人群和背景；

¹该标准基于 Prisma Statement 的改编，符合 PRISMA 的版权许可 (the Creative Commons Attribution License)。遵循 PRISMA 的开放精神，PRIOR-MA 采用类似的版权许可 (CC-BY Attribution 4.0 International)，见：<http://www.dx.doi.org/10.11922/sciencedb.00339>。推荐读者进一步阅读 PRISMA 与 MARS 报告标准以获取更多信息。

- 3.4 说明理由：(a) 研究设计选择的理由；(b) 对文献结果进行选择与编码的理由；(c) 对潜在调节或中介的结果进行选择与编码的理由；
- 3.5 对测量方法和其它变量在心理测量特征上的描述。

简要介绍如何对提出的假设进行检验，包括：

目的 24

- 4.1 被试特征（包括动物）的问题，自变量（实验操作，处理或干预）相关的问题，排除可能产生混淆的变量，因变量（相关的结果或标准）及其研究设计的其它特征；
- 4.2 在荟萃分析中使用的整合方法（例如，效应量测量，平均数分析的方法，同质性分析中使用的模型）。

方法

方案预注册 5

陈述是否提前对荟萃分析研究方案进行预注册。如果预注册过，则需要提供相关信息（例如，公开预注册的网址）。

文献纳入标准 6

陈述文献纳入的标准：

- 6.1 变量：自变量（例如，实验的操作，处理、干预或预测变量的类型），因变量（例如，对临床研究中存在的潜在积极与消极效应进行综合后的结果）或相关研究中涉及的变量；
- 6.2 研究设计（例如，样本抽样的方法或实验处理的方式）；
- 6.3 针对同一研究或同一样本存在多次报告或使用同一样本进行多次测量的情况，不仅需要陈述纳入标准的研究特征（例如，自变量、因变量和研究设计等），同时需要陈述报告特征（检索时间、语种、发表日期和发表状态等），以确定哪个研究需要进行报告；
- 6.4 研究纳入中存在的限制（例如，研究的年代、语言、地区或进行报

² 引言部分条目 4 与方法部分中条目的区别在于，前者主要对被试与变量进行简要介绍；后者则需要给出操作化陈述。

告的方式上存在可能因此产生偏倚的困难)；

6.5 文献纳入标准的变化。预注册时提交的文献纳入标准是否在研究开始后发生改变，或者研究过程中文献纳入标准是否发生改变，例如增加或减少纳入研究的变量，作者需要报告发生了什么改变，以及这些改变是什么时间发生，原因是什么；

6.6 如何处理文献中缺失信息，包括与纳入标准的信息相关信息的缺乏（例如，缺乏关于研究设计的信息）及与数据分析相关信息的缺失（例如，没有报告关于结果的量化数据）。

信息来源	7	描述搜索得到所有信息来源（例如，包括数据库中获取的文献信息，联系作者后获取的额外信息）和最后一次搜索日期。
搜索过程	8	至少完整呈现一个主要使用的数据库在线数据搜索策略，包括限定的范围，搜索时使用的关键词、搜索代码等，以确保搜索的可重复性。 陈述每一步研究筛选过程，包括： 9.1 每一步筛选文献的决策方式（如，通过阅读题目、摘要或全文）； 9.2 筛选过程中执行者的胜任力与人数。如果是一位执行者，如何评估提取数据过程的信度问题；如果有多位执行者，如何处理多人进行筛选可能存在的 inconsistency 问题。
数据收集过程	10	描述从报告中提取数据的方法，包括： 10.1 作为摘录数据依据的变量及其种类； 10.2 数据提取过程中执行者的胜任力与人数。如果是一位执行者，如何评估提取数据过程的信度问题；如果有多位执行者，如何处理多人提取可能存在的 inconsistency 问题。
数据条目	11	列出并定义纳入文献中所有与荟萃分析相关信息（例如，文献的作者、发表年份、被试信息、研究设计或控制组信息等）。

单个研究偏倚*	12	描述如何评估纳入的研究本身可能存在的偏倚（说明这种评估是在研究方法过程层面还是结果层面，还是两者都有），以及这种偏倚如何在结果综合中进行处理。
描述性统计指标	13	描述使用的效应量指标（比如，相关系数、Cohen's <i>d</i> ），包括用来计算单个研究效应量公式。如果同一研究存在多个效应量可能引发的非独立性问题（例如，同一研究中存在多个实验处理、多个控制组、多个相关系数或多个指标量等产生的非独立效应量），陈述处理的方法。
结果综合	14	<p>描述进行跨研究效应量综合和对不同研究的效应进行异质性评估的方法（如固定效应模型或随机效应模型的稳健方差估计（robust variance estimation）），具体包括：</p> <p>14.1 选取某一综合方法的原因；</p> <p>14.2 对研究结果进行加权的方法；</p> <p>14.3 估计单个研究与总体结果变异的方法（如，置信区间）；</p> <p>14.4 描述所有的转换/校正（例如，小样本或非等组样本数据）、调整（例如，聚合，数据缺失，或建构不同水平间的关系）以及对这些进行的数据验证）；</p> <p>14.5 如果进行了贝叶斯分析，报告先验分布的选择与模型拟合的评估。</p> <p>注：（如采用更加复杂或者新颖的方法，可以进行适当地调整）</p>
跨研究偏倚	15	评估所有可能导致跨研究证据累积中产生偏倚的可能性（例如，发表偏倚，选择性报告）。
额外分析	16	如果进行了额外分析，需要描述使用的方法（例如，亚群体分析，敏感性分析，元回归分析），指出哪些是预先计划的。
数据可获得性*	17	荟萃分析中使用的数据（即从文献中提取的数据条目）是否公开。如将荟萃分析的原始数据存放在一个公开并可长期存放的在线数据库

中，方便其他研究者进行重复与在此研究基础上的累积。

分析过程可获得性
*

18

荟萃分析中使用的软件名称及版本号。荟萃分析的代码是否公开，如果公开，在哪可以获得（例如：与数据存在同一在线数据库），方便其他研究者对分析过程的重复。

结果

研究筛选 19

给出筛选后得到的研究数量，给出每一步排除的原因，建议用流程图呈现。

研究特征 20

陈述纳入研究的特征，对每一个纳入的研究变量原则上包括：

20.1 自变量、因变量与主要调节变量的特征；

20.2 重要的被试特征（年龄，性别，受教育程度等）；

20.3 重要的情境变量（例如，环境，日期）；

20.4 研究设计（例如，样本抽样的方法或实验处理的方式）

20.5 在什么地方可以获得该研究的全部数据特征（例如，与作者联系，在补充材料中或预注册网站中）。

单个研究偏倚* 21

尽可能提供每个研究存在偏倚风险的数据，以及质量评估的报告。（见#12）。

单个研究的结果 22

呈现每一个研究综合后的结果（包括阳性与阴性结果），在合适的情况下，推荐使用图表信息进行呈现（例如，森林图）。

结果综合 23

报告每个研究整合的结果（例如，效应量大小，置信区间与异质性程度）。

跨研究偏倚风险 24

描述跨研究偏倚风险，包括：

24.1 陈述是否是（a）未出版的研究与未报告的数据；（b）仅将报告的数据进行整合，如果是，陈述其合理性；

		24.2 是否存在选择性报告；
		24.3 评估发表偏倚的影响大小（例如，漏斗图（funnel plot）、剪补法（trim-and-fill）等）。
额外分析	25	如果做了额外分析，给出结果（例如，亚群体分析，敏感性分析，元回归分析[见#16]）。
讨论		
		陈述主要发现，包括：
		26.1 研究的主要结果，包括荟萃分析的结果；
证据总结	26	26.2 研究证据的总体质量；
		26.3 结果存在的其它解释（例如，混淆变量（confounding），统计检验力）；
		26.4 与先前研究进行整合的结果类似和不同之处。
局限性	27	研究中存在的局限（例如，不一致、不准确与偏倚风险）。
研究推广性与意义	28	对结论的推广性（外部效度）进行描述，并指出对未来研究的意义。
作者贡献与致谢		
作者贡献	29	详细说明各个作者在荟萃分析之中的贡献。
致谢	30	关于对该提供帮助的个人、机构以及资助来源，并列出可能存在的利益冲突。

核心方法部分条目阐释

条目 5. 方案预注册。研究者需要对整个研究进行详细计划，并确保整个计划在荟萃分析实施之前就已经记录在案。报告标准要求作者报告研究是否存在方案，如果存在，需要提

供相应的预注册信息(Moher et al.,2009)。

方案预注册的主要目的在于避免研究者根据自己的需要选择性报告结果,避免摘取与使用数据过程中的随意决策对结果带来的偏倚。在医学荟萃分析方面使用较为广泛的是 YORK 大学建立的 PROSPERO (International Prospective Register of Ongoing Systematic Reviews.www.crd.york.ac.uk/PROSPERO) 系统综述国际预注册平台, 它为研究者们发表方案和信息共享提供了极大的便利 (Booth et al.,2012)。随着方案预注册的便利性的提升, 研究方案的透明性、准确性与完整性也变得极为重要, 也出现了专门为预注册而定制的报告标准 PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocols) (Moher et al., 2015)。PRISMA-P 由 17 个条目组成, 用于指导作者如何对荟萃分析/系统综述方案进行预注册。截止 2018 年 8 月, 已有 42 本杂志可提供荟萃分析预注册的报告, 大部分杂志来自生物行为科学领域 (Quintana, 2018)。在心理学研究中, 目前研究者可以在 Open Science Frameworks (OSF , osf.io)上进行预注册。遗憾的是目前 OSF 尚未对荟萃分析提供相应的模板。

条目 6. 文献纳入标准。荟萃分析预先设定的纳入与排除标准是区别于叙述性综述的一个重要特点。纳入标准的清晰程度影响荟萃分析的效度、可重复性以及最终的结论。荟萃分析报告中对纳入标准的描述不充足会严重影响对稿件质量的评分(Fleming, Koletsis, Seehra, & Pandis, 2014)。PRISMA 基于临床干预研究的荟萃分析, 推荐作者按照 PICOS (participants, interventions, comparisons, outcomes, and study design chosen) (O'Connor, Green, & Higgins, 2008)的格式进行报告。在心理学研究的荟萃分析中, 作者需要在纳入标准中定义并至少报告如下关键内容: 被试群体 (年龄范围、社会经济地位的特征、精神健康状态等)、自变量 (研究者操纵的变量)、因变量 (即主要的测量指标, 比如正确率、反应时间、生理指标、激活脑区等) 与研究设计 (如抽样方式、实验设计、横向对照研究和干预训练研究等)。作者需要清楚地陈述纳入标准的以上特征与原因, 以便读者能够对该报告的偏倚进行判断 (Dwan, Gamble, Williamson, & Kirkham, 2013)。由于研究可能存在多次报告, 一次报告中也可能存在多个研究, 因此文献纳入标准不仅需要陈述上述研究特征 (例如, 被试群体、自变量、因变量和研究设计等), 还需要陈述报告特征 (例如, 时间、语种、发表日期和发表状态等)。

此外, 荟萃分析作者需要报告在纳入过程中可能存在的限制 (例如, 语言、区域或者报告类型) (Appelbaum et al.,2018)。

条目 7. 信息来源。 进行荟萃分析时，为保证文献的全面性，通常需要通过搜索电子数据库。但任何一个电子数据库均无法涵盖所有的信息来源，因此作者还需从不同数据库与其它途径收集信息。来源细节的缺失容易导致其他人难以重复检索过程，因此作者需要对信息来源进行规范化报告。PRISMA 建议在数据库中获取信息至少需要报告数据库名称、平台或提供者，以及每次搜索开始与结束的日期(Liberati et al.,2009)。报告搜索中最早与最晚的记录可以避免数据库更新后造成读者的错误解读。除了数据库的搜索之外，作者还应该采用其他方法对相关研究进行收集，比如私人通信等(Liberati et al.,2009)。对信息来源的充分报告为读者提供了衡量文献覆盖程度的重要参考。

条目 8. 搜索过程。除了条目 7 中描述信息来源外，作者还需要报告完整和全面的数据库搜索策略，以保证其他研究者能够重复该搜索过程。PRISMA 建议作者至少明确报告一个主要使用的数据库的搜索策略，包括布尔运算结构 (boolean connectors)、关键词、自动扩展³ (automatic explosion) (Appelbaum et al.,2018)。搜索实施中可能存在的限制同样需要报告，比如时间、资金支持与语言等原因导致某些研究的信息无法获取等。作者应该清楚、直接地描述这些限制，以便读者衡量报告的质量。

条目 9. 研究筛选。文献搜索完成后，作者需要从大量搜索到的信息中，根据研究文献纳入标准进行筛选。针对这一流程，PRISMA 与 MARS 均建议作者报告筛选过程中的每一步纳入与排除的决策依据（例如，通过阅读标题、摘要或全文，查看全文的频率如何，因为什么关键信息排除了某项研究）。PRISMA 标准的优势在于使用流程图 (<http://www.PRISMA-statement.org/>) 对文献筛选过程进行呈现，流程图清晰地将确认，筛选，合格性，纳入四个关键步骤进行呈现(Moher et al.,2009)。通过筛选流程图在结果部分的呈现，读者对每一步纳入的文献数量，以及被排除的文献的原因能够一目了然。

此外，两位及以上研究者同时筛选能有效降低排除相关文献的概率 (Edwards et al., 2002)，也有助于解决一些难以明确判断是否符合纳入标准的文献(Cooper & Ribble,1989)。为提升报告的客观性以及避免在筛选过程中避免错误，作者应该报告每一阶段的执行者人数，以及他们的具体工作。如果是多人共同完成，需要报告在筛选过程中出现的争议内容以及争议解决的方法。

³ 自动拓展指的是在数据库进行搜索时，为了保证文献的全面性，数据库自动对相关关键词进行拓展。例如在 Pubmed 中搜索 MRI 时，其会自动拓展为 Magnetic Resonance Imaging 等多个关键词进行搜索。

条目 10. 数据收集过程。当作者完成对文献的筛选之后，接下来就需要从这些合格的研究中摘取与荟萃分析研究问题相关的信息，以进行下一步的荟萃分析或者讨论。同样，两位作者独立提取数据所犯的差错少于一位作者提取数据后由另外一位作者核实 (Buscemi, Hartling, Vandermeer, Tjosvold, & Klassen, 2006)。为了缩小误差与减少潜在的偏倚，建议由两位及以上的研究者独立进行数据收集并进行报告。

在数据收集的过程中，作者应描述为减少偏倚和错误的措施。例如，作者可能因为纳入研究的信息不清晰或者不完整而联系原始作者，那么需要描述如何联系原始作者、询问了什么问题、获得了哪些必要的信息、哪些数据以及最终哪些信息被包括在荟萃分析之中。

此外，一些数据存在多次出版的情况，比如在样本量、干预的对象或分析结果上存在微小变化而重复出版。这种情况很难被作者发现，且容易引起偏倚。Liberati 等人 (2009) 建议作者将分散在多个报告的同一研究数据进行汇总，并对如何避免数据重复计算实施的措施进行报告（例如，将不同报告作者的姓名进行比较，对比干预措施之间的差异，查看样本量大小以及结果指标是否存在不同）。同时，在可能的情况下，可考虑与原作者联系进行确认。

条目 11. 数据条目。作者需要尽可能地列出纳入研究中收集到的所有变量数据（例如，作者、发表年份、被试信息、研究设计或控制组信息），并对变量进行清晰的定义。尽管有些数据不会在荟萃分析中使用，但这些信息有助于读者了解作者纳入文献的所有信息。如果作者仅对他在荟萃分析中使用的变量信息进行报告，而忽略那些他认为重要但又未能获取的变量信息，可能导致偏倚并误导读者 (Glasziou, Meats, Heneghan, & Shepperd, 2008)。

具体而言，作者在荟萃分析实施之前应该对将要提取的变量信息在研究方案中进行陈述。对这些信息进行清晰地说明，可以有效避免提取的信息与研究的预期不符。此外，在预注册的研究方案中对变量/数据信息进行描述，有助于读者比较预注册方案与最后报告的研究在这些信息上的差异，从而对荟萃分析的质量进行评估。如果作者在研究开始实施后变量发生变化，比如增加或剔除某些变量，应该陈述这些变量变化的原因 (Liberati et al., 2009)。

条目 12. 单个研究偏倚。偏倚是研究结果或统计推断中的一种系统误差，或与真实值的偏倚 (Higgins, 2008)。如果纳入的研究在方法上存在高度偏倚，荟萃分析的结论极有可能具有高度误导性。PRISMA 声明方法部分条目 12 与 15 均涉及偏倚，前者为单个研究的偏倚（包括该方法用于研究层面或结果层面），后者为研究间影响数据合并结果可能存在的偏倚（如发表偏倚）。

单个研究的偏倚评估主要指作者对纳入研究自身效度的衡量，Gates 与 March (2016)认为作者对纳入研究进行评估时，需要关注三方面问题：（1）研究的结果在多大程度上是可靠的，或者这些研究结果值得信赖吗？（2）该研究包含什么类型的偏倚？（3）存在的偏倚如何对荟萃分析的整体质量产生影响？在心理学研究中的偏倚可能出现在样本选取（主要是本科生样本(Henrich, Heine, & Norenzayan, 2010)）、实施过程（是否有注意力检测(Oppenheimer, Meyvis, & Davidenko, 2009)、是否有操纵检验(manipulation check)、统计测量（统计方法是否得当，是否有进行多重比较校正）或写作报告过程之中。

值得注意的是，对单个研究的偏倚的评估，需要研究者对每个研究本身的质量进行评估。对于未经同行评议的文章来说，荟萃分析作者有必要对其进行偏倚评估。但对于经过同行评议并在学术期刊上发表的文章，作者可以根据实际情况选择是否对每个研究进行偏倚评估。有时，荟萃分析包含的文献数量较大时，对单个研究的偏倚评估可能变得过于耗时耗力。因此，我们将其作为拔高标准，即鼓励作者进行报告，但并非基本要求。

条目 13. 描述性统计指标。荟萃分析需要报告使用的效应量指标是什么（例如，相关系数、Cohen's *d*、风险比（risk ratio）或者神经成像中的激活坐标）。在行为研究中，也要说明效应量的转换/计算公式。与此类似，对神经成像的研究进行基于坐标的荟萃分析时，需要报告不同坐标系统转换的标准。

条目 14. 结果综合。为了比较纳入研究的数据，除了报告效应量指标之外，还需要对结果进行综合。综合前需要报告同质性检验的结果，是否存在异质性，选择了何种模型（例如，固定效应模型或随机效应模型）进行综合，以及为什么选择该模型。

通常在荟萃分析开始前需要报告预先设定好使用的描述性统计指标，并在研究方案中对结果综合部分进行报告（详见 PRISMA 声明中条目 5），目的在于让读者可以理解结果的含义，并了解荟萃分析实施中与最初计划发生了多大程度的偏离。

面对同一研究由于各种原因出现多个效应量的情况，经验不足的研究者常忽视荟萃分析进行效应量综合时必须相互独立的原则，使用传统方法或将多个效应量进行简单平均后再处理均违背荟萃分析效应量独立性的原则。针对该状况，作者需要报告处理非独立性问题采取的方法，例如多元荟萃分析法（multivariate meta-analysis）、多元线性模型法和多元整合法等(叶子菁和石伟,2010)。

条目 15. 跨研究偏倚。跨研究偏倚主要来自纳入研究的信息不完整，比如发表偏倚（Publication bias）。发表偏倚是指某些特定研究结果（例如阳性结果）相比其它结果（例如阴性结果）更容易获得接受与发表的现象而导致的偏倚（Higgins, 2008）。通常， $p < 0.05$ 的结果比 $p > 0.05$ 阴性结果的研究更容易或更快地获得发表，这种“选择性”导致某些研究的发表受到压制，使得研究的发表过程不再是一个随机事件。除了未发表的研究导致信息不完整外，荟萃分析作者纳入研究过程的信息缺失也可能导致结果偏倚，这些问题均属于发表偏倚。

Song, Eastwood 和 Gilbody Duley (1999)对医学领域荟萃分析中存在的发表偏倚进行评估，认为约 50%的荟萃分析中存在不同程度的发表偏倚（20%非常明显），10%的荟萃分析结果受到发表偏倚的影响。发表偏倚是很难避免的，但并不是每一个荟萃分析的发表偏倚都会影响结论，因此需要对荟萃分析发表偏倚的大小进行检验。常见的检验发表偏倚的方法包括漏斗图（funnel plot）（叶子菁和石伟, 2010）、剪补法（trim-and-fill）（Duval, & Tweedie, 2000）和失安全系数（Rosenthal's fail - safe number, N_{fs} ）（Rosenthal, 1979）。近年来 P-curve 的方法也开始为研究所采用（Simons et al., 2014）。其中，漏斗图通过可视化的方法来确定偏倚的存在和来源，为后续荟萃分析在方法上的选择提供一个恰当的理论依据，但其使用也有一些注意事项，例如纳入荟萃分析的数量应该大于 10 个、研究间的效应量的标准差应有差异和组间方差异质性的问题（Sterne et al., 2011）。推荐使用多种方法对发表偏倚进行检测并报告。

如果在荟萃分析实施前的研究方案中明确存在关于偏倚风险的评估，但却未真正地实施，作者需要对这一变化的原因做出说明，便于读者衡量与评估荟萃分析报告质量与效度（Liberati et al., 2009）。值得注意的是，如果文献中确实存在严重发表偏倚，采用统计方法进行校正并不能完全矫正偏倚。

条目 16. 额外分析。如果存在额外分析方法，比如敏感性分析或分组分析（subgroup analyses），在引言部分应该对该方法使用的合理性进行介绍，同时在方法部分对该方法的实施进行描述。

条目 17. 数据可获得性。数据是指作者在荟萃分析中用于描述、阐释与统计分析的研究信息，包括来自条目 11 和 13 中收集到的信息。荟萃分析作者需要报告在荟萃分析中所用到的数据，是否能够获得，以及在什么地方获得。根据近年来国际心理学界对实验研究中透明与开放性的要求，作者需要公开其在研究中所用到的数据，如果不公开数据，需要对理由进行说明。

荟萃分析一般是对已经发表的研究进行的分析，因此其数据一般不涉及伦理或者保密的问题，因此公开数据可能伦理上并无太大的阻碍。将荟萃分析的数据公开，有利于研究者自己及其他研究者重复该荟萃分析，同时有利于一段时间后继续在此基础上进行荟萃分析，因此公开荟萃分析数据对于科学知识的积累有极大的推动作用。但考虑到公开数据近年才逐渐受到研究者的关注，因时，本文暂时将其作为拔高标准，但学术期刊可以根据实际情况，决定是否要求作者将数据存放在可长期保存和容易获取的网络服务器中，并在文章中对数据的可获得性进行说明。

条目 18. 分析过程的可获得性。分析过程指的是对荟萃分析中所有的步骤。包括从文献中提取数据的编码手册、将提取数据进行转化的过程（如 Excel 模板、R 代码等）、荟萃分析的代码等。这些过程与代码对于重复荟萃分析以及对该荟萃分析的质量评估至关重要。同样，考虑到公开分析过程也是近年来出现的新趋势，因此也暂时将其作为拔高标准。学术期刊可以鼓励作者记录荟萃分析的全过程，并将相应的脚本代码与数据一起存放于网络服务器中，并在文章中进行公开。

参考文献

- 叶子菁, 石伟. (2010). 多元整合元分析法评介. *心理科学进展*, 18, 505-510.
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M. & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73, 3-25.
- Booth, A., Clarke, M., Dooley, G., Ghera, D., Moher, D., Petticrew, M. & Stewart, L. (2012). The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Systematic reviews*, 1, 1-9.
- Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L., & Klassen, T. P. (2006). Single data extraction generated more errors than double data extraction in systematic reviews. *Journal of Clinical Epidemiology*, 59(7), 697-703.
- Cooper, H. & Ribble, R. G. (1989). Influences on the outcome of literature searches for integrative research reviews. *Knowledge*, 10, 179-201.
- Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, 95(449), 89-98.
- Dwan, K., Gamble, C., Williamson, P. R., & Kirkham, J. J. (2013). Systematic review of the empirical evidence of study publication bias and outcome reporting bias-an updated review. *PLoS One*, 8(7), 1-37.
- Edwards, P., Clarke, M., DiGuseppi, C., Pratap, S., Roberts, I., & Wentz, R. (2002). Identification of randomized controlled trials in systematic reviews: Accuracy and reliability of screening records. *Statistics in Medicine*, 21(11), 1635-1640.
- Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C. & Antes, G. (1997). Language bias in randomised controlled trials published in English and German. *The Lancet*, 350, 326-329.
- Fleming, P. S., Koletsi, D., Seehra, J., & Pandis, N. (2014). Systematic reviews published in higher impact clinical journals were of higher quality. *Journal of Clinical Epidemiology*, 67(7), 754-759.
- Gates, N. J. & March, E. G. (2016). A neuropsychologist's guide to undertaking a systematic review for publication: Making the most of PRISMA guidelines. *Neuropsychology Review*, 26, 109-120.
- Glasziou, P., Meats, E., Heneghan, C., & Shepperd, S. (2008). What is missing from descriptions of treatment in trials and reviews? *British Medical Journal*, 336(7659), 1472-1474.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83; discussion 83-135.
- Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. New Jersey, USA: Wiley-Blackwell.
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., ... Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical*

Epidemiology, 62, e1-e34.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Plos Medicine*, 6, e1000097.

O'Connor, D., Green, S., & Higgins, J. P. (2008). *Defining the Review Question and Developing Criteria for Including Studies: Cochrane Book Series*. John Wiley & Sons, Ltd.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45(4), 867–872.

Quintana, D. (2018). Remedies for three common issues in meta-analysis. Retrieved August 27, 2018, from <https://doi.org/10.31219/osf.io/5skuv>

Simons, D. J., Holcombe, A. O. & Spellman, B. A. (2014). An introduction to registered replication reports at perspectives on psychological science. *Perspectives On Psychological Science*, 9, 552-555.

Song, F., Eastwood, A., Gilbody, S., & Duley, L. (1999). The role of electronic journals in reducing publication bias. *Medical Informatics*, 24(3), 223–229.

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal (Online)*, 343(5).