

《心理学报》审稿意见与作者回应

题目：通过对比进行学习：多重样例变异性与先前知识的作用

作者：郭建鹏 杨凌燕

第一轮

审稿人1意见：

意见1：该文章主要从样例变异的类型和学生先前知识的角度出发，考察两者对一元一次方程解答的影响。与以往的实验室人工问题研究相比，该研究以真实的教育教学环境作为基础，其研究结果具有更强的生态效度与推广性，得出的一些有意义结论与教学建议也能作为课程教学改革的依据。但仍有一些问题有待作者进一步的考虑与澄清：在前言部分，建议作者把要解决的问题在最后一段用简要的文字进行总结提炼，明确地表述出来。在最后的结论部分，把展望之类的内容放在讨论部分。

回应：感谢审稿专家对论文的意见和建议，对我们进一步改善论文质量很有帮助。我们根据您的意见对论文进行了细致的修改。我们已经把“引言”改为“问题提出”，并围绕研究问题对文献综述进行了提炼和精简，力求突出与研究问题相关的文献。并根据您的建议，在最后一段总结概括了要解决的问题，以及相应的研究假设。另外在结论部分，已经概括出具体研究结论，并把展望与不足的内容放在讨论部分。具体请见“1 问题提出”（p.8-10）、“4 讨论”、“5 结论”（p.23-27）部分。

意见2：作者将“对比类型或解法组”定义为部分与对比类型组相似，部分与对比解法组相似，但两种关键特征又是单独变化的。既然是单独变化的，其配对方式也与对比类型组或对比解法组相似，那么是否可以认为该组的实验材料与对比类型组或对比解法组相同？如果相同，为何要设置这一条件？另外，作者对前三种变异类型进行了举例，但没有对第四种变异类型举例，建议补充。

回应：感谢审稿专家的建议。“对比类型或解法组”的样例配对中有一半（3对样例、共6个样例）与对比类型组相似，对比相同解法不同类型的方程组；另一半（3对样例、共6个样例）与对比解法组相似，对比不同解法相同类型的方程组。设计这一条件的目的是为了在“类型_解法_类型/解法组”这一实验条件中，学生只是对比类型和解法的单独变异，并没有对比两者的同时变异。这样既能保证该组被试各对比了一半的类型单独变异（18个样例）和一半的解法单独变异（18个样例），也能保证他们所对比的样例数与其他组一样，从而可以考察类型和解法是否有同时变异的效果（对比类型_解法_类型+解法组）。由于这第四种变异类型与对比类型、对比解法这两种类型的题目类似，因此我们就没有进行举例。我们对论文的表达进行了相应的修改，以求更为准备地表达意思。具体请见p.12。

意见3：文中提到“前后测的题目相同”是否指代题目内容相同？如果是，这是否会导致熟悉性、练习效应的出现，从而影响后测的结果？

回应：本研究的前后测题目是相同的。在很多类似的学习实验中（如Rittle-Johnson等人Journal of Educational Psychology, 2007, 2009; British Journal of Educational Psychology, 2012; Star等人Learning and Instruction, 2008），都采用相同的题目进行前后测，这样能够更好地了解学生学习之后的进步程度。为了避免您提到的熟悉、练习效应，一般都是在预测一周之后进行后测。在本研究中，学生完成预测后一周才开始进行3天的学习阶段，预测完成后的第11天

才进行后测。在统计分析中，前测分数也是作为协变量进行控制的。根据您的建议，我们修改了“2.2实验设计”和“2.4实验程序”部分，以更清楚说明，具体请见p.11, 15-16。

意见4：在1.2 中，作者提到“可以根据学生审辨出来的关键特征来划分学生的先前知识”，但后面的分析中又将能否使用简便方法作为不同先前知识程度的划分依据，前后对于先前知识的划分并不统一，究竟是如何区分的？简便方法组与常规方法组的划分是根据前测中是否至少使用过一次简便方法解题，这种划分方式与标准是否合适？

回应：本研究根据学生在前测中是否使用过简便方法解题把学生分为简便方法组与无简便方法组。这种方法与Rittle-Johnson等人（2007, 2009a,b）以及Guo等人（2014）研究中对被试的分类方法保持一致。如果学生在前测中使用过简便方法解题，那么我们认为他们对于解法这个关键特征的审辨比较好，对方程的解法比较熟悉。数据分析结果也显示这类学生的各种前测分数都显著好于前测没有使用简便方法的学生。所以我们认为先前知识高的学生要比先前知识低的学生审辨出更多的关键特征。前测使用简便方法的学生对于解法这个关键特征的审辨要好于前测没有使用简便方法的学生，他们的先前知识也较高。我们在后面的结果讨论部分也是根据学生审辨的关键特征进行讨论的，对于前测没有使用简便方法的学生，由于方程的问题类型和解法都是学习的关键特征，因此他们需要首先对比这两个特征的单独变异再同时对两者的同时变异。

另外，我们根据评审意见对文献综述进行了提炼和精简，突出与研究问题相关的文献。第一稿的1.2内容已经删除。

意见 5：概念性知识主要是一种陈述性、内容性的事实知识，但从表 4 对概念性知识的测量题目上看，其内容与概念性知识的定义相关不大，更像是对知识的操作与问题解决，建议作者再作考虑。

回应：感谢审稿专家的建议。为了把研究结果与已有研究进行对比，本研究的测试题目与Rittle-Johnson等人（2007, 2009a,b）在Journal of Educational Psychology, Star等人（2009）在Journal of Experimental Child Psychology、2008在Learning and Instruction，以及Guo等人（2014）在European Journal of Psychology of Education上的系列研究的测试题目一致。

Kilpatrick等人（2001）认为，概念性知识是“an integrated and functional grasp of mathematical ideas”(p. 118)。The National Research Council对概念性知识的定义是：“comprehension of mathematical concepts, operations, and relations”(p. 5)。Star (2005)的定义是“conceptual knowledge has come to encompass not only what is known (knowledge of concepts) but also one way that concepts can be known (e.g. deeply and with rich connections)’ (p. 408)。

Rittle-Johnson(1999)指出，“conceptual knowledge can be considered a subset of declarative knowledge (which also includes knowledge of unrelated facts).”(p. 3)。Rittle-Johnson(2013)指出，对概念性知识的测量方法有很多种，有些要求对概念的隐性（implicit）知识，如进行类别的选择、判断步骤或答案的正确与否、对比数量、转换数量不同的表征体系等；有些要求对概念的显性（explicit）知识，如解释做出的判断、定义概念、画概念图等。他们认为如果使用多个任务来测量学生的概念性知识则会比较科学。他们对概念性知识的测量是“We measured conceptual knowledge via the ability to recognize and to explain key concepts in the domain, in line with past measures of conceptual knowledge (e.g., Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Hiebert & Wearne, 1996; Rittle-Johnson & Alibali, 1999; Rittle-Johnson & Star, 2007).”（Rittle-Johnson & Star, 2009, p. 532）。

因此，根据上述的定义和做法，本研究使用上述研究中的测试题目。概念性知识被认为是数学知识的整体性和功能性的掌握，体现为识别和解释相关概念的能力。测量学生关于等式、同类项和复合变量的言语及非言语性知识。

意见6: 对于知识的划分，Alexander曾在上世纪90年代提出较为经典的领域知识模型并得到了较为广泛的研究（其中更包含了基于具体教学学科的领域知识研究）。领域知识包含了内容知识、程序过程知识和条件知识，这一知识划分与作者提出的多重样例对比什么（what）、怎样对比（how）和何时对比（when），以及作者用来测试学生解方程能力的三种知识形态是相似的，为何作者在文献综述和理论建构中并未提及该理论及其相关的经典与最新研究？建议对这方面的内容进行补充。作者把知识分为“程序性知识、变通性知识和概念性知识”，这种划分的依据是什么？

回应: 感谢评审专家提出Alexander的领域知识模型。根据专家的建议，我们对这方面文献进行了补充，并对本研究中程序性知识、变通性知识和概念性知识与Alexander对知识的分类：declarative knowledge, conditional knowledge, conceptual knowledge进行了对比。我们把知识分为“程序性知识、变通性知识和概念性知识”的依据，正如对您第5个问题的回答一样。首先是为了把研究结果与Rittle-Johnson等人（2007, 2009a,b）在Journal of Educational Psychology, Star等人（2009）在Journal of Experimental Child Psychology、2008在Learning and Instruction，以及Guo等人（2014）在European Journal of Psychology of Education上的系列研究进行对比。本研究不仅使用了上述研究对知识的这三种分类，而且使用了类似的题目进行测试。这样能保证研究结果的对比性和可解释性。

其次，在数学教育研究中，这三类知识一直都被认为是三种关键的数学能力（Hiebert, 1986; Kilpatrick, Swafford, & Findell, 2001; Siegler, 1996; Verschaffel et al., 2009）。概念性知识和程序性知识的划分及其对于数学学习的重要性已经被广泛认可（Carpenter, Franke, Jacobs, Fennema, & Empson, 1998; Hiebert & Wearne, 1996; Hiebert, 1986; Paas & Van Merriënboer, 1994; Silver, 1986）。学生掌握多种解题策略并能够在不同的情境中灵活运用也是一种重要的数学能力（Baroody & Dowker, 2003; Rittle-Johnson & Star, 2007; Star & Seifert, 2006）。变通性知识是领域专家区别于数学新手一个显著的特点（Dowker, 1992; Star & Newton, 2009）。具有较高变通性知识的人在面对不熟悉问题是更能调整已有的方法，并对于领域概念的理解更为深入（Hiebert et al., 1996）。然而很多时候数学教学缺乏灵活性知识，导致学生在解决远近迁移问题时出现很多困难（Hiebert & Carpenter, 1992; National Research Council, 2000）。Star等人（2007）指出，如何发展变通性知识是认知科学和教育一个中心的课题。因此，我们把变通性知识作为概念性知识和程序性知识外的第三种重要的数学知识和能力。

综上，我们在本研究中使用了概念性知识、程序性知识和变通性知识的分类方法。而对于Alexander的领域知识三种分类与我们分类方法之间的对比和联系，我们也在方法部分的2.2.3的测试材料进行了说明。具体请见p.15。

意见7: 整个研究没有从理论上分析为什么出现这样的结果？

回应: 根据专家的建议，我们增加了从样例学习理论的角度分析本研究结果。在讨论部分我们从多重样例对比的对比什么、何时对比、以及怎么对比三个方面分析研究结果。我们指出，样例学习理论采取表面/结构特征的分类方法，探讨多重样例变异性对不同先前知识学生学习的影响。但是已有样例学习理论及相关实证研究并没有回答对比多重样例时应该聚焦于样例的什么特征？应该如何处理多重样例之间的变异性？以及针对不同先前知识的学生应该如何设计多重样例这三个问题。导致已有研究结果出现了很多相互矛盾的地方。本研究采取

关键/非关键特征的分类方法，提出先前知识不同的学生学习时的关键特征也不同，要基于学生学习的关键特征系统设计多重样例的变异性（单独变异和同时变异），让学生在单独审辨每个关键特征的基础上（需要充分的教学支持）经验多个关键特征的同时变异。这些研究发现在一定程度上回答了样例学习理论的三个问题，提供了一个新的研究视角，具有一定的理论价值。另外我们还基于aptitude-treatment interaction (ATI)、expertise reverse effect等理论及研究关于教学方法和学生特质关系的观点，对本研究结果进行了相应的分析。

具体分析请见p.25-26。

意见8：从表6、表7和表8的结果上看，各个实验处理的 η_p^2 普遍较低，因而容易影响结果与结论的可靠性。

回应：本研究中实验处理的主效应都不显著，而且效应值partial eta squared较低,这符合本研究的结果及解释。但是由于本研究的主要结果是发现实验处理和前测是否使用简便方法之间的交互效应,从文中可以看到,当交互效应出现显著的时候,该交互变量的 η_p^2 普遍在0.05-0.09之间,属于中等范围 (η_p^2 : 0.01=small,0.06=medium, 0.14=large, Cohen, 1988)。根据另一位评审专家的意见,我们还删去了原来的表6-8,用文字把统计分析结果报告在正文中,并增加了简单效应分析的 η_p^2 值。我们注意到很多同类研究的 η_p^2 值都在此区间。比如, Fyfe等人2012在Journal of Educational Psychology发表的论文中 η_p^2 介于0.03-0.08之间, McEldoon等人2013在British Journal of Educational Psychology上发表的论文中 η_p^2 基本都在0.04-0.05之间, Richey等人2013在Learning and Instruction上发表的论文中 η_p^2 大部分都在0.04-0.09之间。此外,我们也在文中的不足与展望部分对效应值进行了讨论,本研究作为课堂实证研究可能受到一些无关因素的影响,并且本研究的实验处理主要为学生自行阅读材料,这些因素都会影响到本研究的效应值。未来研究如果进一步加入教师讲授、增加教学支持或同伴讨论等形式则可能会加大样例不同变异类型之间的效果差异。具体请见“3结果”(p. 17-22)和“4.5本研究的不足和未来研究展望”(p.27)。

参考文献:

- Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, 104(4), 1094-1108.
- McEldoon, K. L., Durkin, K. L., & Rittle-Johnson, B. (2013). Is self-explanation worth the time? A comparison to additional practice. *British Journal of Educational Psychology*, 83(4), 615-632.
- Richey, J. E., & Nokes-Malach, T. J. (2013). How much is too much? Learning and motivation effects of adding instructional explanations to worked examples. *Learning & Instruction*, 25, 104-124.

审稿人 2 意见：

意见 1：该篇文章选题的内容有一定教育意义。但有些问题值得商榷：（1）四组被试是独立进行样例学习还是在老师的指导下进行练习？（2）实验设计是几个因素、几个因变量？要说清楚。（3）做前测的目的是什么？如果是跟后测做对比分析，那在结果分析中应该有前后测的差异分析结果。（4）既然是实验研究，一定要有具体的结果和明确的结论。所以，建议作者修改文章并注意以下几点：文献综述要精简，突出要研究的问题和怎样研究，明确写出实验假设。

回应：感谢审稿专家对论文的意见和建议，对我们进一步改善论文质量很有帮助。我们根据

您的意见对论文进行了细致的修改。我们对您提出的问题结合修改建议回复如下。

根据您的建议，我们精简并提炼了文献综述，把“引言”改为“问题提出”，删去了原来 1.2 的内容，突出研究问题和设计，并在文献综述的最后一段概括了要解决的问题，以及相应的研究假设。具体请见“1 问题提出”（p.8-10）。

意见 2：实验设计要清楚明确，实验程序要单独写，写清楚实验的导语。

回应：感谢审稿专家的建议。我们修改了原来的“实验设计和步骤”部分，改为“实验设计”和“实验程序”两个部分。结合您上述的前两个问题，我们在实验程序里面有介绍被试的学习形式。在三天三节课的干预阶段中，学生主要是进行自学材料，但是在自学之前每节课有15分钟的时候由教师简单介绍实验，并带领全班学生一起解答若干道方程。在实验设计中，我们有介绍本实验采用4（变异类型） \times 2（先前知识）二因素被试间实验设计，以学生在解一元一次方程上的程序性知识、变通性知识、概念性知识和整体表现作为因变量。此外，根据您的建议，我们清楚明确地介绍了实验设计，写清楚了实验导语和具体程序。具体请见“2.2 实验设计”（p.11）和“2.4 实验程序”（p. 15-16）部分。

意见 3：统计分析结果不要列表，要用文字把统计分析结果报告出来。

回应：感谢审稿专家的建议。根据您的建议，我们已经删去表 6、7、8，并把相关的统计分析结果在正文中用文字报告出来。我们还增加了简单效应分析的 η_p^2 值。

另外，回答您前面关于前测的问题。本研究的前测是为了了解学生学习前的先前知识高低（可以和其他相关研究的被试先前知识进行对比），并根据他们的前测表现对学生进行分组。我们还比较了每个实验组在前测分数上差异，结果显示实验组的前测分数不存在显著差别。此外根据您的建议，我们也对比了学生的前后测分数，发现所有学生在全部测试题中的后测分数都显著高于前测分数（ $p < 0.01$ ），表明所有学生通过学习在程序性知识、变通性知识和概念性知识上都有了明显的进步。学生的前测分数也被作为协变量进入到 ANCOVA 分析中，以控制前测分数的影响。

具体请见“3 结果”（p. 17-22）。

意见 4：要根据实验结果讨论问题。要根据实验结果概括出具体的研究结论。

回应：感谢审稿专家的建议。根据您的建议，我们在讨论部分概括出了具体的研究结果，还在最后的结论部分根据研究结果概括出了 3 点明确的研究结论。

我们还把本研究结果与已有相关研究结论联系起来一起讨论，并总结出了本研究结果对教学的实践价值、不足及未来研究展望。此外，根据另一位评审专家的建议，我们增加了从样例学习理论的角度分析本研究结果，从多重样例对比的对比什么、何时对比、以及怎么对比三个方面分析研究结果。具体请见“4 讨论”和“5 结论”（p.23-27）部分。

第二轮

审稿人 1 意见：

意见 1：文章作者就审稿人在一审中提出的疑问进行了回答，审稿对作者的回答大多基本满意，但对意见 3 与意见 7 的回答还不能让审稿人信服。关于意见 3，即前后测题目相同的问题，虽然作者提到其他研究者也是采用了相同的方式进行的，但这并不是一个特别令人信服的证明，因为其他人的研究也同样存者这样的问题。因此，建议作者在讨论中重视这一问题。

回应：感谢审稿专家的建议。根据您的建议，我们在“本研究的不足与未来研究展望”部分讨论了本研究前后测题目相同可能造成的练习效应，以及我们为了控制这种效应所做的处理，并对未来研究如何更好地消除练习效应提出了建议。

具体分析请见“4.5本研究的不足与未来研究展望”（p.30）。

意见2：关于意见7,即对于结果的理论解释问题，作者仍然未能回答“为什么”的问题，从二修稿件的结论部分来看，基本上就事论事，缺乏理论高度。

回应：感谢审稿专家的建议。根据您的建议，我们在讨论部分增加了一部分内容专门从理论角度解释本研究的结果，回答“为什么”的问题。我们根据变易理论（variation theory）和认知负荷理论（cognitive load theory）的观点，分析了样例对比在发展学生辨别性知识、审辨关键特征中的作用。学习是一种审辨的过程，审辨就需要经验变异。单独变异和同时变异为系统设计多重样例的变异性提供了一种方法，首先单独变异每个关键特征，让学生通过对比分别审辨出每个关键特征。最后再同时变异全部关键特征，让学生通过体验关键特征的同时变化来达到全面的审辨学习。

根据认知负荷理论，不同变异类型的样例对学生会产生不同的认知负荷。同时变异的样例比单独变异的样例具有更高的内在认知负荷（intrinsic cognitive load），学生在对比这些高变异样例的时候很容易出现认知超负荷从而无法有效完成对比。先单独变异再同时变异的方法一方面能够降低学生样例学习时的外在认知负荷（extraneous cognitive load），帮助他们使用较少的认知资源就能够审辨出学习的关键特征；另一方面也能够提高样例学习时的相关认知负荷（germane cognitive load），促进学生积极主动地进行对比和图式构建。本研究还在样例对比中增加了适当的教学解释，帮助学生注意并审辨样例的关键特征，思考、理解样例之间的异同及其体现的原理，在降低学习外在认知负荷的同时，提高了相关的认知负荷。

对于低先前知识学生超负荷的样例对比材料和设计方法对于高先前知识的学生反而可能是合适的，不会超出他们的工作记忆容量；他们能够集中认知资源对比审辨学习的关键特征。而对于低先前知识学生有效的样例对比对于高先前知识的学生则可能是无效的、冗余的，甚至是有害的，迫使他们对比已经审辨出来的关键特征，加重了外在认知负荷并减少了相关认知负荷，从而产生教学效果的逆反，出现专长反转效应（expertise reverse effect）。

具体分析请见“4.3 如何有效进行多重样例对比”（p.28-29）。

审稿人 2 意见：

意见1：论文作者对两位审稿人的意见进行了较为细致的回复，并在论文中进行了相应的修改。但若发表，还需澄清以下问题：作者在文中指出Guo等人（2014）研究中存在的不足，并试图在Guo等人的基础上进行。既然该研究是在先前研究基础上进行的，就应该着重于本研究的新增或改进部分进行深入探讨与论述。但从全文来看，无论是文章的框架结构、研究设计、结果分析及写作风格，都与Guo等人（2014）有很大的相似性。作者在哪些方面有了实质性的改进与贡献？比如，之前Guo等人的研究已将学生先前经验视为重要变量加以操纵，而该研究增加样例对比数量（从原来的30增加到36）及教学指导，这与之前研究中的先前经验的操纵有何实质性区别？该文未针对这种调整（增加样例、教学指导）的意义（或必要性）、理论及文献依据、操纵的依据（如，样题调整到36题的依据与效果？教学指导的独到效果）及其不同效果等展开充分的讨论。建议作者从文献分析、理论构想、研究设计到后续的结果分析等方面进一步围绕该研究所做的改进或调整给予说明，以突出该研究的贡献与创新。

回应：感谢审稿专家的建议。根据您的建议，我们针对本研究与Guo等人（2014）研究在样例对比数量、教学解释和实验被试的先前知识等方面的不同对论文进行了修改，突出了本研

究的不同及贡献。我们在问题提出部分分析了样例对比学习中教学支持的作用，通过文献分析指出，在样例学习中提供教学解释对于学习具有正面的作用，以及指出本研究设计有效教学解释的原则。在研究设计中，我们也比较了本研究与Guo等人（2014）研究在学习材料以及教学解释的不同之处，突出了本研究所做的实质性改进。最后在研究结果分析中，我们专门在“4.2对比相关研究结果”部分比较本研究与Guo等人（2014）研究的结果，讨论教学解释在样例对比学习中的作用，指出教学解释能够降低学生对比样例时的认知负荷，减轻工作记忆的负担的作用，帮助学生把有限的认知资源集中于对比样例的关键特征上。

具体分析请见 1.问题提出 (p.12), 2.3.2 学习材料、2.4 实验程序 (p.17-18), 4.2 对比相关研究结果 (p.27-29)。

第三轮

审稿人 1 意见:

意见 1: 审稿人对二修稿件基本满意, 建议采用。

回应: 感谢专家的肯定。

审稿人 2 意见:

意见 1: 作者虽然补充了与之前Guo等人（2014）所做研究的对比等方面的内容, 但并没有把本研究与之前研究最为不同的做法——教学解释、增加样例等作为论文要探讨的核心内容。如果这样的话, 与之前的研究并无实质区别。既然两篇文章在设计上的主要差异表现在教学解释、样例数量方面, 那么本研究要回答的关键问题就应该是为什么要给予更多的教学解释、更多的样例数量? 其必要性是什么? 效果及作用机制是什么? 建议作者思考本文的研究重心究竟应该突出什么, 以便体现其实质性的价值。

回应: 感谢评审专家宝贵的建议。

Guo 等人（2014）的实验得出了与其研究假设不一致的研究结果, 而他们对此不一致结果的解释是: “学生在对比两个关键特征的同时变异之前, 需要足够的时间和机会去单独审辨这两个关键特征。否则当他们面对太多变异的时候就容易混淆, 无法审辨出关键特征, 从而影响学习结果。”为了验证这种说法, 本研究在实验设计上做了两方面的调整: (1) 在样例对比学习中增加了教学支持; (2) 本研究的被试在参加实验之前没有学过相关知识, 学生的先前知识较低。从而得到了与 Guo 等人（2014）不同甚至相反的研究结果, 验证了“学生需要足够的机会和支持单独审辨每个关键特征然后才能对比关键特征的同时变异”这种说法。这两个研究的不同表明在增加教学支持的情况下, 样例变异类型的作用发生了变化。本研究采用与 Guo 等人（2014）相同的实验处理也是为了达到这个对比的目的。因此, 我们认为与 Guo 等人（2014）研究相比, 本研究有实质区别和价值。

类似的例子也可以从本论文的两个主要参考文献中找到。Rittle-Johnson 等人 2009 年在 *Journal of Educational Psychology* 第 3 期和第 4 期上发表的两篇论文“Compared With What? The Effects of Different Comparisons on Conceptual Knowledge and Procedural Flexibility for Equation Solving”和“The Importance of Prior Knowledge When Comparing Examples: Influences on Conceptual and Procedural Knowledge of Equation Solving”。这两篇论文的主要区别在于被试先前知识水平的不同, 一个研究的被试学过解方程, 另一个研究的被试则没有学过解方程, 两个研究的其他方面基本相同。由于被试先前知识水平的不同, 两个研究得出了不同的结论。

我们明白专家希望论文主要讨论教学支持的作用, 以此作为本论文的一个研究重心。

根据专家在上一轮的建议，本论文已经从问题提出、文献综述、实验处理和讨论中对教学支持进行了分析。但是考虑到本研究的实验设计，论文无法对教学支持的作用做过多的铺开论述。这是因为，本研究是二因素实验设计，两个自变量分别为变异类型和先前知识。因此，数据结果只允许讨论样例变异类型和学生先前知识的作用，及二者交互作用对学习的影响效果。虽然教学支持是本文与 Guo 等人（2014）研究的一个主要不同，但是教学支持并不是本研究的一个自变量，实验数据无法说明教学支持如何影响学习，或者教学支持与样例变异类型、先前知识三者之间的交互如何影响学习。这也是我们没有把讨论教学支持的作用作为本论文重心的原因。

根据专家的建议，我们认为未来研究可以在本研究的基础上设计新的实验，考察教学支持、变异类型和先前知识等自变量及其交互作用对学习的影响效果。这样才可以重点讨论教学支持的效果及机制。而这是本研究的设计所无法完成的任务。我们在未来研究展望部分对这一点进行了讨论。

第四轮

审稿人 2 意见：

意见 1：作者对评审意见进行了较为充分的回答与解释，客观地阐述了该研究与之前类似研究的关键区别以及该研究存在的不足，并在正文中做了相应的修改与说明，显示了作者诚恳务实的态度。总体来看，作者对存在的问题给予了可接受的回答，并在正文相关处尽量加以修改，建议发表。

回应：感谢专家的肯定。