

## 《心理学报》审稿意见与作者回应

题目：认知行为实验研究中最佳素材容量的选择与确定：多元概化理论应用

作者：罗照盛，郭小军

### 第一轮：

#### 审稿人 1 意见：

本文从多元概化理论的视角，对心理学中三个经典认知试验的素材用量问题进行了研究，发现在 IAT 测验上，相容与不相容任务的素材容量为 50 时最佳；在图-词干扰范式中，素材容量以 48 时为最佳；对线索提示范式，50%有效提示中，有效与无效提示最佳素材容量为 35 最好。这个研究的积极意义在于：

- 1、为认知试验中材料数量的确定提供了一种新思路和方法，为确保试验信度、减少材料损耗、控制试验时间、降低疲劳效应等方面具有积极价值。
- 2、多元概化理论原本用于测验数据的分析，作者把该理论引入心理学中的认知试验研究，这是对概化理论应用范围的一种创新和拓广。
- 3、文章关于多元概化理论的应用是恰当的，计算方法是正确的，所得结果基本可靠。

#### 意见 1：

“但是，正如作者在文章中所指出的，心理学的试验数据应当先进行必要的转换，如对数变换，否则，各种主效应和交互效应之间的差别有可能被夸大。建议作者使用  $\ln$  转换后的数据对所做分析进行验证。若所得结果基本一致，则同意发表该文。”

#### 回应：

根据专家意见，作者进行了相应的数据处理和分析，具体结果见提交的附件，由于内容较多，没有附到正文。根据分析结果，对数据进行对数转换只是改变了数据的表现形式，并没有改变研究的结论。

#### 意见 2：

“最后，作者的摘要太罗嗦，不必要的句子可以删除。”

#### 回应：

作者根据审稿专家的意见，已经删去了部分不必要的内容。

#### 审稿人 2 意见：

##### 意见 1：

“论文主要研究“运用多元概化理论研究认知行为实验研究中最佳容量的选择与确定”问题。选题具有一定的意义。由于论文没有专门的文献综述，仅在前言中对部分前人的研究作了简要介绍，没有提及以往的心理实验研究中是否有对实验素材容量的研究，或者是否有运用多元概化理论研究心理学实验素材容量的成果。然而，这一研究与教育考试领域运用多元概化理论研究考试的试题量的原理完全一样。因此，认为这一研究选题有一定意义，但并不一定有重大意义。”

##### 回应：

根据专家意见，我们再次对国内外文献进行了更深入的检索，均未发现对认知行为实验中的

最佳素材容量问题进行专门的研究，或运用多元概化理论对实验素材容量进行研究的文献。

2. 论文根据 IAT、图-词干扰、线索提示三种范式的实验研究数据，运用多元概化理论对每种范式下实验素材容量与实验结果的可靠性关系进行了研究。每个范式研究中，G 研究效应在不同任务上的方差与协方差分量估计结果的解释正确，不同任务上方差分量的解释正确，D 研究中组合全域的各方差分量估计结果的解释正确，不同任务对总方差贡献比例分析正确，D 研究中不同任务样本量与可靠性指数关系分析基本正确。

意见 2:

“论文对三个范式研究过程的描述，尤其是实验程序的描述过于简单，缺少实验数据收集整理过程的描述。因此，难以判断作者研究依据的实验数据的可靠性。如果实验数据本身有较大的误差，论文得出的实验素材最佳容量的结论就不一定可靠了。”

回应:

这确实是我们的疏忽，我们已按照审稿专家意见，补充了关于实验过程的描述内容。

意见 3:

““讨论与结论”部分得出的最后结论是：“由此可见，多元概化理论在探讨实验素材容量选择与确定问题上具有重要的指导意义”，似乎论文主要是在论证多元概化理论对探讨实验素材容量选择与确定是否有指导意义，其实论文只是运用多元概化理论研究实验素材的最佳容量问题，不是研究多元概化理论对探讨实验素材容量选择与确定是否有指导意义。因此这一结论与论文的研究内容是不一致的，需要修改。类似地，“摘要”的最后一句也同样陈述了这一结论，显然是不必要的。”

回应:

作者确实没有恰当地陈述本研究的目的和意义，因此已按照审稿专家的意见，对相关陈述进行了修改和删除。

意见 4:

“论文写作存在三个方面的问题，一是很多小节中先呈现数据，再出现文字叙述，这不符合论文的写作要求；二是“讨论与结论”部分只是重复了前言和研究内容部分的部分话语，没有对研究结果进行真正的讨论；三是英文摘要写作质量较差，甚至有句型结构和用词错误，需要修改。”

回应: 关于第一方面问题，作者已按照审稿专家的意见，对相关小节的内容进行了修改和调整。关于第二方面问题，作者充实了部分讨论与结论的相关内容。关于第三方面的英文摘要问题，首先感谢专家提出的具体意见，作者请了两位老师一起对摘要进行了斟酌和修改。

## 第二轮

审稿人 1 意见:

论文经过修改比第一稿好多了，但仍然存在一些问题:

意见 1:

“对三个实验的描述，需要说明有效的反应时取值范围，但实验 2 没有说明。”

回应: 作者已经按修改意见在论文中添加相关描述。

意见 2:

“既然是研究最佳实验素材容量，那么每个实验至少要明确说明实验素材容量，但只有第三个实验说明了实验容量，第 1、2 实验都没有说明实验素材容量，尽管从后面的表格中能够看出素材容量，但最好在实验描述说明一下。”

回应：作者已经按修改意见在论文中添加相关描述。

意见 3:

“三个实验都是通过检查可靠性指数及其增量差的大小来确定最佳容量，但三者均不一致。第一个实验确定的可靠性指数是 0.915，增量差为 0.006；第二个实验确定的可靠性指数是 0.954，增量差为 0.006；第三个实验确定的可靠性指数是 0.973，增量差为 0.003。究竟可靠性指数达到多大为好，从论文看不出一个统一的或者有说服力的说法。其实，就实验而言，关键是要确定有效反应时的允许误差究竟为多大（当然不同的实验，有效反应时不一样，允许误差也可不同），只要绝对误差标准差小于了这一允许误差，相应的素材容量就是合适的容量，而不是仅仅看可靠性指数的大小及其增量差，从这个角度看，每个实验的有效反应时数据应该列出，以便读者可以根据实验数据以及概化理论计算的绝对误差标准差确定研究者给出的最佳容量是否合理。”

回应：关于修改意见的第一方面，概化理论在确定合适容量（比如测验题量、维度等）及测量的可靠性指数时，一般遵循以下几点原则：1、总体可靠性指数（或概化系数）取值合理；2、可靠性指数（或者概化系数）后续取值变化趋势不明显；3、可靠性指数合理的情况下，测量成本尽量减小。所以，本文选用的标准也是可靠性指数取值合理（本文为 0.9 以上），同时可靠性指数增量差异不明显（本文中差异在 0.01 以下），而且实验素材相对不是特别大，最终确定了三个实验的最佳容量。关于修改意见的第二方面，我们已按修改意见在文中对每个实验的有效反应时取值范围进行了描述。

意见 4:

“此外，论文的中英文摘要也有一些问题。例如，中文摘要的第二句是“实验素材是影响实验结果最重要的因素之一”，英文摘要的相应句子是“The size of experiment material is a critical factor that will significantly affect the results of experiment.”综合两句意思和论文内容，似乎应该是“实验素材容量是影响实验结果误差大小的最重要因素之一”，如果是这个意思，中英文摘要的对应句子都不太正确。”

回应：我们对中英文摘要进行了修改，以便使中英文摘要表述的内容能够相互对应。

意见 5：“文中还有少数表述不太合适或者文字错误，详见论文中的修改之处和批注。”

回应：作者仔细通读了论文，并对相关的内容进行了修订。