

学习任务单对理科学生学习成效的影响：一项元分析

马欣然，唐欣怡*

西北大学经济管理学院，西安 710127，陕西，中国

摘要：提升学生的学习成效始终是理科教育的核心目标——学生被期望能够进行批判性参与、开展科学推理，并把知识迁移到多样的情境之中。为达成这一目标，学习任务单长期被用于理科课堂。然而，尽管其使用极为普遍，已有的元分析在范围上仍相当有限，往往只聚焦于单一学科，或未考察调节变量的作用。本研究采用元分析方法，分析了取自 33 项研究、涉及 2 465 名学生的 38 个效应量。结果显示，总体效应量大且具有统计显著性（Hedges $g = 1.600$ ），证实学习任务单能显著提升学生的理科学习成效。调节效应分析表明，国家、教育阶段、学科、学习领域与任务单变式均产生显著影响。效应最为突出的情形出现在中学阶段、物理学科，以及民族科学型与多元表征型任务单之中。本研究为学习任务单的有效性提供了循证依据，并强调应审慎设计任务单，使之超越常规练习，成为支持理科教育中有意义学习的工具。

关键词：学习任务单；学习成效；理科教育；元分析

Effectiveness of Worksheets in Science Students' Learning Outcomes: A Meta-Analysis

XinRan Ma, XinYi Tang*

School of Economics and Management, Northwest University, Xi'an 710127, Shaanxi, China

Abstract: Improving students' learning outcomes remains a central goal of science education, as students are expected to engage critically, reason scientifically, and apply knowledge in diverse contexts. Worksheets have long been used in science classrooms to achieve this goal. But despite their widespread use, existing meta-analyses have been limited in scope, often focusing on single disciplines or lacking the influence of moderating variables. This study conducted a meta-analysis, analyzing 38 effect sizes drawn from 33 studies involving 2,465 students. Results revealed a large and statistically significant overall effect (Hedges $g = 1.600$), confirming that worksheets substantially improve students' learning outcomes in science. Moderator analysis indicated a significant effect based on country, education level, scientific discipline, learning domain, and worksheet variation. The most notable effects were observed in the secondary level, in the field of physics, and in ethnoscience and representation-rich worksheets. This study provides evidence-based insights on the effectiveness of worksheets and the importance of thoughtfully designing them as tools that extend beyond routine practice to support meaningful learning in science education.

Keywords: Worksheets; Learning outcome; Science education; Meta-analysis

一、引言

面对全球性挑战与议题,理科教育承诺让今天的儿童具备必要的知识、技能与品格,以在他们即将继承的、日益易变(volatile)、不确定(uncertain)、复杂(complex)且模糊(ambiguous)的 VUCA 世界中立足并有所贡献[1]。然而,这一愿景能否实现,还取决于理科教师采纳以研究为依据的教学策略与干预手段的能力。在当今课堂可用的诸多工具中,学习任务单因其灵活性与低成本而被广泛且持久地使用[2]。学习任务单是由教师设计的学习纸页,其中载有一组任务、练习与视觉辅助材料,用以支持学生的自主练习,或在完成学习活动时起到引导作用[3]。设计良好的任务单为学生提供了进行提取式练习的宝贵机会[4],并使其能够通过结构化任务与引导性提问来运用已有知识[5]。

尽管使用广泛,支持其效果的证据却并不一致。已有若干综述考察了理科任务单如何被开发并用于课堂[6-7],但这些研究只聚焦于单一学科,且缺乏得出确定结论所需的方法学严谨性。Chutami 与 Suhartini[2]曾尝试用元分析综合任务单的效应,但并未遵循 PRISMA 指南来完成严格的定量综合。上述研究也未探讨调节变量如何影响学生的学习成效。这些空白清楚地表明,有必要开展更进一步、更全面的分析,为理科教师提供切实可行且以研究为依据的洞见,以指导他们编制能够提升学生认知、情感与动作技能三类理科学习成效的任务单。

本研究即着眼于此:遵循《系统评价与元分析优先报告条目》(PRISMA)指南[8],对理科教育中学习任务单的既有研究进行定量综合,从而对其有效性获得更严谨、更全面的理解。本研究综合任务单对学生学习成效的影响,考察国家、教育阶段、学科与学习领域等调节变量的作用,并检验哪些变式的任务单最为有效。鉴于以往文献未能识别出究竟是什么让任务单在提升理科学生学习成效上真正奏效,本研究还通过澄清任务单在何种变式下效果最佳,为这一讨论作出贡献。

二、文献综述

理科教育中的学习任务单早已超越了重复练习工具的传统角色。今天,它们被视为能够提升批判性思维能力、培养科学态度的教学材料[4]。以建构主义学习原理为基础,任务单被设计用来让学生通过探索、反思与探究性任务,主动参与知识的建构[9]。这类任务单往往伴随若干活动,要求学生调查真实世界的问题、分析数据,并通过动手参与得出基于证据的结论。若使用得当,引导探究式任务单可通过促使学生进行意义建构与深层思考,在多种科学情境中发展其概念理解、定量推理与元认知能力。支架式任务单则在学习之初提供结构化支持,进而逐步促成自主学习并改善总体学习成效。

在文献中,任务单有多种称谓,常见的有学生任务单(student worksheets)与科学任务单(scientific worksheets)。学生任务单更强调概念的发现过程,常借助多种媒介提供多样刺激并安排以学生为中心的活动,使学习更有意义、更有效、也更愉快。因此,当这类任务单以动手实验为设计出发点时,便有助于激发学生预测、测量与有效交流的能力。Mahyuna 等[5]强调,以引导探究为基础的科学任务单能显著提升学生的科学过程技能。引导探究式任务单之所以有效,是因为它们易于理解、贴近情境,并能激励学生开展科学探究。当它们被嵌入问题式学习(PBL)或 STEM 整合教学时,在促进学生投入与高阶学习方面会更为有效。Yalyn 等证明,面向 PBL 的任务单显著提升了学生的科学过程技能,在假设提出、变量辨识、数据解释与交流等环节上增益尤为明显。

由于结构化、以学生为中心的取向有利于知识建构与自主学习,任务单长期以来都是理科重要的教学工具。随着课堂使用的持续增长,若干研究评估了其对学习的影响,也促成了对新近发现进行综合的元分析。Chutami 与 Suhartini[2]对相关国内期刊的十篇文章作了元分析,报告了中等至极高的效应量,并得出结论:

在理科学习中使用学生任务单对学生的学习成效有影响。类似地, Putri 等[6] 聚焦任务单对物理学习成效的影响, 分析了 30 篇文章并显示出积极结果; 其分析发现, 在初中阶段采用情境教学 (CTL) 模式时, 任务单对学生学习成效具有高至极高的效应。在另一项综述中, Syahdi 等[7] 综合了 20 项研究, 表明任务单确实对提升自然科学中的批判性与创造性思维能力有正向作用。Anjani 等 则把范围扩展到物理、化学与生物, 基于 25 项研究揭示任务单对学生的高阶思维能力具有很好的影响。

尽管以往的元分析报告了任务单在理科教育中的积极效应, 仍有若干局限未被解决。最突出的是, 早前的综述并未聚焦任务单对学生跨理科学科的总体学习成效的影响[6]。已有证据显示, 融入探究式学习的任务单能促进学生的批判性思维、创造性思维、高阶思维与问题解决能力[7], 但这些综述大多集中于思维技能, 并未纳入基础性的学习成效。此外, 也有一些研究尝试考察调节变量, 涉及学科领域、教学模式与年级水平[6]; 即便如此, 它们仍未考虑任务单变式、地理位置与学习领域等其他关键调节变量, 部分研究还缺乏 PRISMA 指南所强调的方法学严谨性。因此, 本研究旨在检验任务单在提升学生理科学习成效方面的总体有效性, 并通过调节效应分析评估其影响如何随国家、教育阶段、学科与任务单变式而变化。理科教师对强化自身技术、教学法与学科内容知识 (TPACK) 的终身追求, 意味着这一研究领域亟需循证洞见。由此, 本研究为任务单类教学材料的开发提供了具有实证依据的建议, 并推动理科教育中以研究为基础的实践。

三、研究问题

本研究采用元分析方法考察学习任务单对理科学生学习成效的有效性。具体而言, 本研究拟回答以下问题:

1. 学习任务单对理科学生的学习成效有何影响?
2. 当把纳入的研究按下列变量分组时, 学习任务单对学生学习成效的影响是否存在差异:
(a) 学习领域; (b) 国家; (c) 教育阶段; (d) 学科; (e) 任务单变式。
3. 纳入的研究采用了哪些具体类型的学习任务单, 又是如何使用的?

四、研究方法

(一) 研究设计

本研究采用元分析方法, 考察学习任务单在提升理科学生学习成效方面的有效性, 并遵循《系统评价与元分析优先报告条目》(PRISMA) 指南[8-12]。元分析是把各独立研究的数据综合为单一效应估计值, 以回答一个聚焦的研究问题。合并来自不同研究的数据, 使我们得以考察研究设计、实施、分析与结果上的差异如何影响总体处理效应。为更好地理解效应量之间的差异, 本研究开展了调节效应分析, 以探究各研究间变异的潜在来源。纳入的调节变量包括学习领域、国家、教育阶段、学科与任务单变式。

(二) 文献检索程序

本研究开展了广泛的文献检索, 目标为 2015 年 1 月 1 日至 2025 年 6 月 1 日之间发表的实证研究。借助 Publish or Perish (PoP) 工具, 以 “worksheets” “science education” “learning outcomes” 与 “quasi-experimental” 为检索词, 从 Google Scholar、Crossref、OpenAlex 与 Scopus 中检索文献; 另通过 Elicit 与 Scispace 进行人工检索, 作为数据库结果的补充。共识别到 3 069 条记录, 其中 3 001 条来自数据库、68 条来自人工检索。经在线工具与人工双重查重, 排除 276 篇, 余 2 793 篇进入筛选。标题筛选剔除 2 446 条不相关记录, 摘要筛选进一步将范围缩小至 130 篇待作全文评阅。经充分评估后, 最终有 33 项研究符合纳入标准并进入元分析, 从中提取出 38 个效应量。

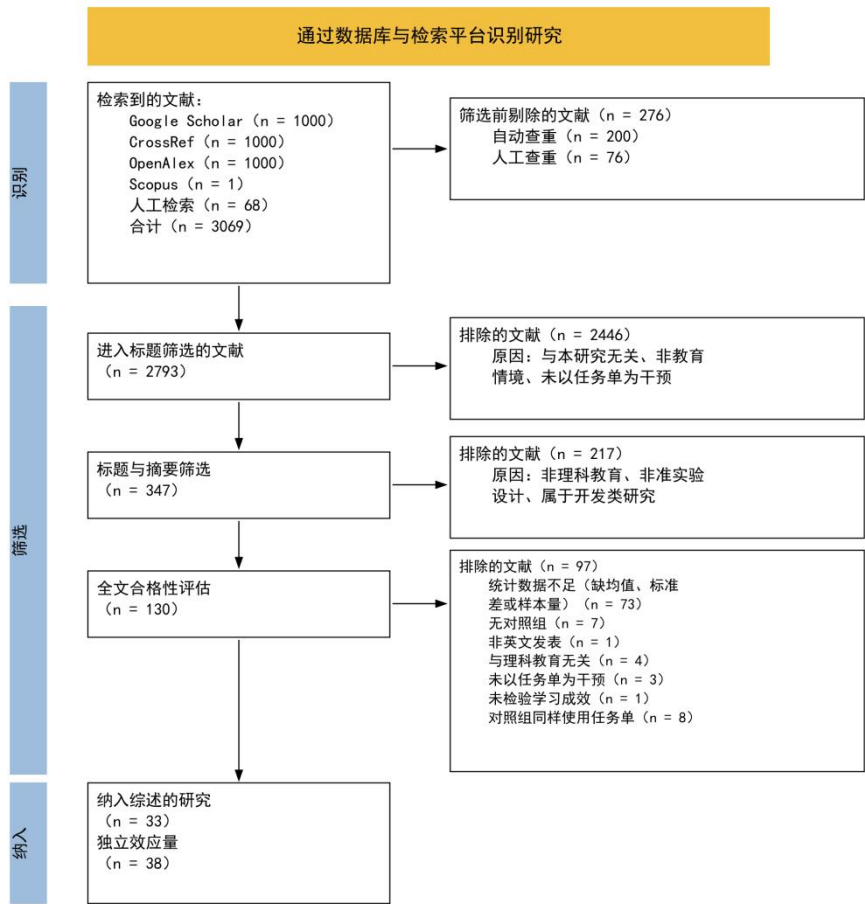


Figure 1 Flow Diagram Using the PRISMA Protocol

图 1 依据 PRISMA 方案的检索流程图

注：依据原文图 1 重绘并译为中文。

（三）纳入标准

为确保只有相关文章进入本综述，遴选过程遵循研究者拟定的一组标准。纳入的文章必须：

（a）可获取在线全文；（b）发表于 2015 年至 2025 年之间；（c）以英文撰写；（d）在小学、中学或高等教育情境中开展；（e）采用准实验设计，且对照组接受常规教学；（f）聚焦理科教育；（g）报告充分的定量数据，包括后测均值、标准差与样本量。

在评阅的 130 篇全文中，有 97 篇因下列原因被排除：未使用任务单（n = 3）、对照组未采用常规教学（n = 8）、无组间比较（n = 7）、缺失定量数据（n = 73）、未聚焦理科教育（n = 4）、未检验学习成效（n = 1）、以及非英文发表（n = 1）。由此剩余 33 项研究符合纳入条件，并从中提取出 38 个效应量——因部分研究报告了多个与本研究范围一致的学习成效。

（四）编码特征与程序

研究者依据既定的纳入标准，从 33 项合格研究中提取并编码了 38 个独立效应量的数据。三位研究者使用共享的 Google 表格独立编码与提取；初次编码完成后，另一位研究者对每项研究交叉核对，分歧或错误经讨论解决。我们起初打算纳入干预时长，但因多数合格研究未报告这一信息而作罢；受时间所限，我们也无法联系原作者。因此，本元分析中的全部数据均取自全文稿件中可获得的信息。

由于许多纳入研究描述了形态各异的任务单设计，我们依据反复出现的教学特征及其与既有教育理论的契

合度对其分组，并按稿件所述的主导性教学设计特征进行归类。类别包括：常规型、技术增强型、建构主义型、探究型、多元表征型、差异化型与民族科学型。常规型任务单未作教学法或技术上的改造，技术增强型任务单则整合了数字工具与模拟。建构主义型任务单以建构主义理论为基础，该理论主张学习者通过与任务和情境的互动来建构知识。

探究型任务单被单列为一类，因为采用该取向的研究数量已足以支撑一次有意义的调节效应分析。多元表征型任务单纳入了以认知负荷理论与双重编码理论为依据的视觉与符号辅助。差异化型任务单运用支架、引导式笔记等策略来支持多样化的学习者。最后，民族科学型任务单融入本土知识，与文化回应式教学相契合。下表列出各类别及其具体任务单变式与对应的研究数量。

Table 1 Worksheet Variation Categories

表 1 学习任务单的变式类别

类别	具体变式	纳入效应量数 (k)
建构主义型	问题式学习 (PBL)	4
	发现学习	2
	基于问题解决	2
	项目论证式学习	1
	项目式学习	1
多元表征型	过程图像式	4
	面向高阶思维 (HOTS)	2
	概念漫画增强式	1
	基于多元智能	1
	多元表征	1
探究型	探究式	3
	引导探究式	2
	过程导向引导探究 (POGIL)	1
技术增强型	电子任务单	4
	增强现实 (AR)	2
差异化型	混合式学习	1
	引导式笔记	1
	学习风格取向	1
	支架式	1
常规型	—	2
民族科学型	—	1
合计		38

注：各类别小计依次为 10、9、6、6、4、2、1，合计 38，与效应量总数一致。

(五) 效应量计算

33 项研究的效应量使用 Comprehensive Meta-Analysis (CMA) 第 4 版计算。我们选用 Hedges g 而非 Cohen's d 作为标准化效应量指标，因其对小样本偏倚作了校正；并采用随机效应模型，因该模型适用于异质性较高的研究。效应量按 Cohen 的标准解释为小 (0.2)、中 (0.5) 或大 (0.8)。正值表示有利于任务单，负值

则有利于常规教学。调节效应分析采用混合效应模型，考察国家、教育阶段、学科、学习领域与任务单类型上的差异。发表偏倚通过经典失效安全数（Classic fail-safe N）、Begg-Mazumdar 检验与漏斗图检视加以评估。所有分析均采用 95% 置信区间与 $p < 0.05$ 的显著性水平。

五、研究结果

（一）总体效应与异质性

本元分析综合了取自 33 项研究、涉及 2 465 名学生的 38 个效应量，以评估学习任务单对理科学习成效的影响。结果发现存在显著异质性（ $Q(37) = 512.047$, $p < 0.001$; $I^2 = 92.77\%$ ），因而采用随机效应模型。总体效应量大且显著（Hedges $g = 1.600$, $SE = 0.175$, 95% CI [1.257, 1.942], $z = 9.151$, $p < 0.001$ ），这意味着学习任务单能实质性地改善理科学习成效。

Table 2 Overall Intervention Effect Size and Heterogeneity Analysis

表 2 总体干预效应量与异质性分析

模型	效应量数	Hedges g	标准误差	方差	下限	上限	z 值	p 值	Q 值	df	p 值	I ² (%)
固定效应	38	1.223	0.046	0.002	1.142	1.324	26.545	0.000	512.047	37	0.000*	92.77
随机效应	38	1.600	0.175	0.031	1.257	1.942	9.151	0.000	—	—	—	—

注：下限与上限为 95% 置信区间；Q 为同质性检验值；df 为自由度；* 表示 $p < 0.05$ ，具有统计显著性。原表 Q 值栏印作“.047”，此处按正文所述的 $Q(37) = 512.047$ 补全。

Table 3 Forest Plot and Distribution of Effect Sizes

表 3 各效应量的分布（森林图数据）

研究	Hedges g	标准误差	方差	下限	上限	z 值	p 值
Afrianti & Lufri (2022)	1.551	0.339	0.115	0.887	2.216	4.575	0.000*
Aiman 等 (2020)	0.965	0.265	0.070	0.445	1.485	3.635	0.000*
Amila 等 (2018)	4.471	0.527	0.277	3.439	5.504	8.489	0.000*
Benedicto & Bulay (2024)	3.330	0.552	0.305	2.248	4.412	6.031	0.000*
Chotimah 等 (2023)	0.597	0.285	0.081	0.039	1.155	2.096	0.036*
Dewi 等 (2018a)	0.802	0.265	0.070	0.283	1.322	3.026	0.002*
Dewi 等 (2018b)	2.085	0.318	0.101	1.462	2.709	6.555	0.000*
Dewi 等 (2018c)	0.765	0.264	0.070	0.247	1.283	2.896	0.004*
Ekantini (2018)	0.013	0.255	0.065	-0.487	0.512	0.050	0.960
Febriana 等 (2019)	1.726	0.257	0.066	1.222	2.230	6.715	0.000*
Fernanda 等 (2024)	2.206	0.315	0.099	1.589	2.822	7.010	0.000*
Fernando 等 (2021)	9.139	0.832	0.692	7.509	10.770	10.987	0.000*
Gani 等 (2017)	1.891	0.330	0.109	1.244	2.539	5.726	0.000*

Hasanah 等 (2021)	0.852	0.176	0.031	0.508	1.196	4.851	0.000*
Hasnunidah (2023)	4.246	0.329	0.108	3.601	4.890	12.918	0.000*
Hatiti 等 (2021)	0.673	0.245	0.060	0.193	1.153	2.749	0.006*
Herlina 等 (2025)	1.917	0.312	0.098	1.304	2.529	6.135	0.000*
Hidayatulloh 等 (2021)	0.576	0.260	0.080	0.066	1.086	2.214	0.027*
Layawon (2024)	0.998	0.296	0.087	0.418	1.577	3.373	0.001*
Maknun & Herman (2024)	1.271	0.365	0.133	0.556	1.987	3.483	0.000*
Maysara 等 (2024)	0.594	0.290	0.084	0.025	1.163	2.046	0.041*
Nadrah (2023)	1.031	0.263	0.069	0.515	1.547	3.916	0.000*
Nurdin 等 (2018)	0.600	0.261	0.068	0.089	1.111	2.301	0.021*
Nurulsari 等 (2023)	1.716	0.299	0.090	1.130	2.303	5.735	0.000*
Okut 等 (2023)	0.495	0.257	0.066	-0.008	0.998	1.928	0.054
Putri & Sani (2023)	0.802	0.246	0.060	0.320	1.284	3.260	0.001*
Rudibyani (2020)	1.656	0.279	0.078	1.110	2.202	5.942	0.000*
Sindani 等 (2021)	0.639	0.297	0.088	0.056	1.222	2.150	0.032*
Suniasih & Sujana (2023)	1.575	0.290	0.084	1.006	2.144	5.426	0.000*
Sutarto 等 (2018a)	2.600	0.323	0.104	1.967	3.232	8.856	0.000*
Sutarto 等 (2018b)	3.157	0.356	0.127	2.458	3.856	8.856	0.000*
Sutarto 等 (2018c)	3.058	0.450	0.123	2.372	3.745	8.731	0.000*
Sutarto 等 (2018d)	3.403	0.372	0.139	2.673	4.133	9.141	0.000*
Wicaktini 等 (2020)	0.576	0.252	0.064	0.081	1.070	2.283	0.022*
Widarti 等 (2019)	1.222	0.288	0.083	0.658	1.786	4.248	0.000*
Yanto (2024)	1.476	0.271	0.074	0.945	2.008	5.445	0.000*
Yurtyapan & Kandemir (2021)	0.012	0.223	0.050	-0.425	0.450	0.055	0.956
Zaini 等 (2022)	-0.045	0.255	0.065	-0.544	0.455	-0.175	0.861
合并效应	1.600	0.175	0.031	1.257	1.942	9.151	0.000*

注：* 表示 $p \leq 0.05$ ，具有统计显著性。原文该表为一幅森林图截图，左侧为上述数值、右侧为对应的效应量点线图（横轴 -5.00 至 5.00 ，左侧“有利于对照组”、右侧“有利于任务单”），本文按可编辑表格重建其数值部分，图形部分的信息见图 3 及正文描述。原图三处 z 值印刷有误：Febriana 等印作“6.71.5”、Maysara 等印作“2.0246”、Sindani 等印作“21.150”，此处按小数位与置信区间推定为 6.715、2.046 与 2.150。Nurdin 等 (2018) 与 Yanto (2024) 在原文参考文献表中无对应条目，故未编号。

表 3 呈现了 38 个效应量各自的效应值、标准误、方差、 z 值与 p 值。效应量的跨度很大，从 -0.045 到 9.139 ，多数落在正值区间。38 个效应量中有 35 个达到统计显著 ($p < 0.05$)，另有 3 个不显著 ($p > 0.05$)。原文的森林图以方块表示每项研究的效应量、以水平线段表示其 95% 置信区间，底部菱形则表示合并后的总体效应量。该图凸显了正向效应的一致性以及各研究间量级的差异，强化了“支持使用学习任务单”这一总体结论的稳健性。

（二）调节效应分析

鉴于异质性显著，本研究就国家、教育阶段、学科、学习领域与任务单变式五个变量开展了调节效应分析，结果显示各组间差异显著。最大的效应出现在印度尼西亚 ($g = 1.657$)、中学阶段 ($g = 1.777$) 与物理学科 ($g =$

2.069)。各学习领域均呈现强效应，其中认知领域最高（ $g = 1.660$ ）。在各类任务单变式中，民族科学型（ $g = 4.471$ ）与多元表征型（ $g = 2.772$ ）的效应量最大，常规型最小（ $g = 0.520$ ）。

Table 4 Moderator Analysis of Included Studies

表 4 纳入研究的调节效应分析

变量	类别	k	效应量	下限	上限	Q 值	p 值
国家	印度尼西亚	34	1.657	1.294	2.019	32.798	0.000*
	菲律宾	3	1.509	0.233	2.786		
	土耳其	1	0.012	-0.425	0.450		
教育阶段	中学	31	1.777	1.373	2.180	6.737	0.034*
	高等教育	3	0.892	0.209	1.538		
	小学	4	0.874	-0.055	1.839		
学科	物理	19	2.069	1.530	2.608	31.816	0.000*
	生物	10	1.426	0.724	2.129		
	化学	3	1.161	0.557	1.766		
	地球科学	2	0.980	0.592	1.367		
	小学科学	3	0.847	-0.089	1.782		
	环境科学	1	0.013	-0.487	0.512		
学习领域	认知	35	1.660	1.281	2.039	13.222	0.001*
	动作技能	1	1.476	0.945	2.008		
	情感	1	0.825	0.539	1.112		
任务单变式	民族科学型	1	4.471	3.439	5.504	51.827	0.000*
	多元表征型	9	2.772	1.701	3.844		
	建构主义型	10	1.632	0.978	2.285		
	技术增强型	6	1.090	0.394	1.786		
	差异化型	4	0.952	0.440	1.464		
	探究型	6	0.950	0.546	1.354		
	常规型	2	0.520	-0.478	1.517		

注：下限与上限为 95% 置信区间；* 表示 $p < 0.05$ ，具有统计显著性。校核发现：国家、教育阶段、学科与任务单变式四组的 k 值合计均为 38，唯“学习领域”一组合计为 37；据附录表，情感领域实为 2 个效应量（Dewi 等 2018c、Hasanah 等 2021），原表该栏应为 2，此处照录原值并加以说明。

（三）学习任务单的变式分布

图 2 显示了各类变式被纳入研究采用的频次。建构主义型（ $k = 10$ ，26.3%）与多元表征型（ $k = 9$ ，23.7%）最为常见，其次是探究型与技术增强型（各 $k = 6$ ，15.8%）。差异化型（ $k = 4$ ）、常规型（ $k = 2$ ）与民族科学型（ $k = 1$ ）出现较少。

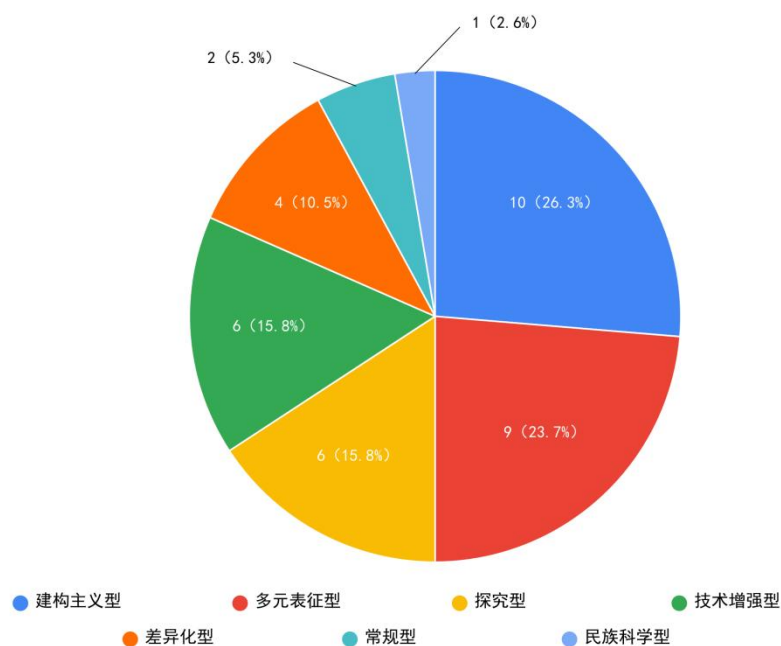


Figure 2 Distribution of Worksheet Variations

图 2 学习任务单变式的分布

注：依据原文图 2 重绘并译为中文；扇区比例与配色照原图。

（四）发表偏倚

本研究通过多项检验评估发表偏倚。漏斗图呈现出轻微的不对称，提示可能存在效应量小或为负的缺失研究；Egger 检验（截距 = 12.98, $p < 0.001$ ）与 Begg-Mazumdar 秩相关检验（ $\tau = 0.66$, $p < 0.001$ ）也支持这一判断。不过，经典失效安全数检验显示，需要额外 8 666 项零效应研究才能把 p 值抬高到 0.05 阈值以上。该检验的具体输出为：观测研究的 Z 值 = 29.662 98，双尾检验，使 p 值大于 α 所需的缺失研究数 = 8 666，观测研究的 p 值 = 0.000 00，对应 α 的 Z 值 = 1.959 96， $\alpha = 0.050$ 00，观测研究数 = 38。结果的总体格局与稳健的失效安全数估计表明：虽然存在一定程度的发表偏倚，但它并未对本元分析结论的效度或稳健性构成严重威胁。

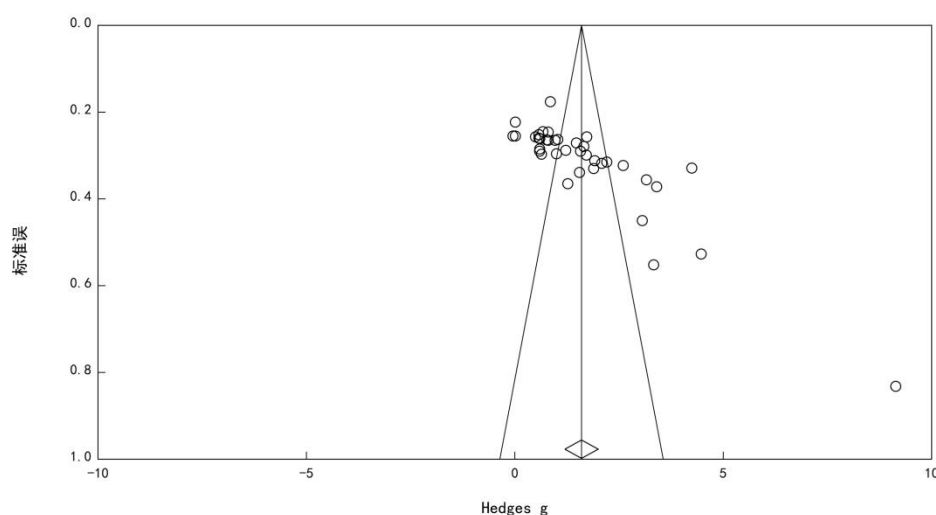


Figure 3 Funnel Plot of Standard Error by Hedges g

图 3 标准误对 Hedges g 的漏斗图

注：依据原文图 3 重绘并译为中文；38 个散点按表 3 的 Hedges g 与标准误逐一定位，漏斗线以合并效应量 $g = 1.600$ 为轴、按 $g \pm 1.96 \times SE$ 作出。

六、讨论

本研究的发现为学习任务单在理科教育中的有效性提供了有力的实证支持。总体效应量 ($g = 1.600$) 属于大效应，证实任务单能显著提升学生的学习成效；森林图以直观方式展示了任务单对学生学习的积极影响，与本元分析的定量结果相互印证。这些发现与以往的元分析一致[2][6]——后者报告，当任务单被嵌入以学生为中心的学习活动之中时，其效应可达中等至极高水平。本研究在此基础上作了拓展：样本更大、更为多样，遵循 PRISMA 框架，并考察了更广泛的调节变量。Q 值 512.047 与 I^2 值 92.774% 表明，所考察的 38 个效应量之间存在实质性异质性。既有方法学文献指出， I^2 超过 75% 即意味着变异远超随机因素所能解释的范围 (Higgins 等, 2003)。这也进一步说明，分析可能影响任务单效果的调节因素十分重要；教学情境、任务单设计、学科内容与学习者特征上的差异，很可能是这些变异的来源 (Higgins & Thompson, 2002)。

学习领域同样呈现出大而显著的效应量，其中以认知领域的结果最为突出。这意味着任务单让学习变得更有意义，进而提升了学生的认知能力。其次是动作技能领域与情感领域，效应量分别为 $g = 1.476$ 与 $g = 0.825$ 。学习过程大多围绕学生的认知发展展开，而对情感与动作技能发展的探索仍显不足。不过已有研究表明，任务单——尤其是在采用引导探究模式时——能有效培养学生的科学态度；而要把任务单真正融入科学态度的培养，就应把发现与批判性思维有意识地嵌入科学过程之中。同样地，Lufri 等发现，以 PBL 为基础的学生任务单能显著增强学生的情感能力，促使他们在提出与回应问题时承担责任并作批判性思考。与此同时，任务单也通过让学生主动参与科学过程来帮助其发展动作技能。未来研究应进一步探讨任务单如何影响情感与动作技能的发展，特别是在实践性与价值取向较强的理科主题中。

全部研究均在亚洲开展。印度尼西亚 ($k = 34$) 的总体效应量最高 ($g = 1.657$)，其次是菲律宾 ($k = 3$)，同样报告了大效应 ($g = 1.509$)；土耳其仅有一项研究 ($k = 1$)，效应量可忽略不计 ($g = 0.012$)。印度尼西亚与菲律宾之所以出现较大效应，可能与其教育政策强调以学生为中心的策略、21 世纪技能的整合以及主动学习工具的使用有关。这些发现印证了 Lee 的研究：尽管任务单是全球通用的教学工具，但在亚洲，任务单的使用

更多是课程的正式组成部分,而不仅仅是补充材料。研究地域上的这一局限也表明,有必要在更多元的教育体系中开展全球性研究。

调节效应分析还显示,教育阶段显著影响任务单的有效性($p = 0.034$)。效应最强出现在中学阶段($g = 1.777$, $k = 31$),高等教育($g = 0.892$, $k = 3$)与小学阶段($g = 0.874$, $k = 4$)的效应则较小。学生在认知发展与学习准备度上的差异可能是造成这一结果的原因:认知差异影响学习者把握科学概念的方式,而社会经济与文化背景则可能影响其准备度。因此,同龄儿童在接纳更广泛的外部影响时并不总是处于同一准备水平(Wahyuni 等, 2021)。由是观之,任务单在小学、中学与高等教育三个阶段上有效性的差异,或可归因于学生在认知成熟度以及加工结构化学习材料的准备度上的不同。

学科同样被发现具有显著效应($p = 0.000$),说明任务单的影响随科学分支的不同而变化。物理的效应量最高($g = 2.069$, $k = 19$),这很可能是因为该学科的抽象概念得益于多元表征型任务单中常见的可视化与多重表征。Orulebaja 等强调,有效的物理学习会以方程、图示、图像与文字等多种形式呈现,而在这些形式之间建构意义能够增强概念理解与问题解决能力。物理之后依次是生物($g = 1.426$)、化学($g = 1.161$)、地球科学($g = 0.980$)与小学科学($g = 0.847$),均呈现大效应。这些学科的一个共同点在于,任务单被设计用来让学生主动参与知识建构[9],而这些学科恰恰需要对复杂系统与过程的理解。相形之下,环境科学($g = 0.013$)的效应极小且不显著;但由于仅有一项研究,该结果应审慎解读。

本研究的分析还揭示出任务单在设计与实施上的多样变式,各变式之间的差异均达到显著。其中,Amila 等唯一的一份民族科学型任务单取得了最大的正向效应($g = 4.471$, $k = 1$)。这一大效应很可能应归因于其文化回应式取向,即把新的科学概念与学生既有的文化知识联系起来。类似地,Gay 主张,借助熟悉的叙事进行情境化能同时增强学生的理解与切身相关感。但须注意的是,该变式仅有一项研究,因此结果虽有前景,仍应审慎解读;尽管如此,本研究的发现仍指向把科学与本地文化相联系的教学潜力。多元表征型任务单同样持续取得大效应($g = 2.772$, $k = 9$)。可以看到,这类任务单常被用于抽象概念极为普遍的化学与物理之中。例如 Widarti 等报告,借助多重表征改善了学生对分子间作用力的理解。Gates 认为,不同的视觉材料能为学生提供多种“看见”信息的途径,是帮助他们理解困难概念的有效方式。然而,这类任务单能否成功取决于设计是否周密。正如 Rasch 与 Schnotz 所告诫的,整合不当的视觉材料可能造成冗余或加重认知负荷。因此,多元表征型任务单的有效性,很可能来自它们能够增强而非干扰核心学习——即尽量减少不必要的重复,并确保视觉材料被高效地用于促进学习。

建构主义型任务单是使用最频繁的类型,并持续表现出强效应($g = 1.632$, $k = 10$)。Hasnunidah 以环境污染为题使用了项目论证式学习任务单。既有研究支持这类任务单能够促进更深层的理解与元认知发展[5][9]。但若缺乏清晰的结构,这类任务反而可能让批判性思维尚在形成之中的学生不堪重负。因此,尽管证据显示建构主义型任务单潜力强劲,其有效性仍依赖教师循序渐进地引导学生从依赖走向独立。技术增强型任务单的大效应($g = 1.090$)则反映了数字化干预在理科教育中不断增长的价值与演进。正如文献所强调的,设计良好的数字任务单能改善学生的态度与动机,并促进自主学习。例如 Okut 等实施电子任务单,以支持人体系统单元的自定步调学习。但这一转向也引出了关于学生的设备与网络接入、以及使用这些设备所需技能的重要问题。随着课堂日益数字化,未来研究应探讨学生对技术的感受以及他们需要何种支持。当技术不是作为附加物、而是作为支持探究、可视化与反馈的工具被使用时,技术增强型任务单便能在传统课堂与数字课堂中都高效地深化理科学习。

探究型任务单在理科学生的学习成效上同样呈现出稳定的大效应($g = 0.950$, $k = 6$)。Chotimah 等为气

体动理论开发了引导探究式任务单, 使学生能够主动参与科学意义建构。既有研究表明, 探究型任务单能有效强化数据分析、基于证据建构解释等科学过程技能[5], 本研究所见的强效应支持了这些早前的发现。这意味着, 这类任务单的影响或许在更强调动手学习与问题解决的学习环境中最为明显; 在这种情形下, 探究型任务单既为学生提供了结构, 也提供了更深入探索观念的机会。差异化型任务单产生了与探究型相近的大效应($g = 0.952$, $k = 4$)。Febriana 等 使用了以排泄系统为内容的引导式笔记任务单, 其中关键概念部分给出, 由学生在讲授过程中补全。已有研究显示, 无论通过支架还是混合式学习模式, 差异化型任务单都有助于理解, 在学生进度不一的多样化课堂中尤为有用。

常规型任务单的效应量相对最小, 但仍属中等($g = 0.520$, $k = 2$)。它们之所以仍在课堂中沿用, 是因为简便易行并能促进主动学习[4]。然而, 其往往偏低的认知要求可能限制了培养深层学习成效的能力——这一顾虑也呼应了 Mayer (2004) 对机械记忆式取向的批评。常规型任务单虽可发挥有用的基础性作用, 但本研究的发现表明, 当它们与其他教学取向或技术手段相结合时, 其教育价值会提升, 效应也随之增大。因此, 教师面临的挑战并非彻底弃用它们, 而是通过探究性提示、技术强化、情境化、以学习者为中心的取向、多重表征或差异化来改造其设计, 从而为学生提供更 meaningful 的学习经验。

七、结论

本元分析考察了取自 33 项研究的 38 个独立效应量, 以判定学习任务单在提升学生理科学习成效方面的有效性。总体效应量大且为正, 表明使用任务单能显著改善理科学生的学习成效。但分析同时揭示出显著的异质性, 意味着这一有效性在不同情境下存在差异。本研究就学习领域、国家、教育阶段、学科与任务单变式五个关键变量开展了调节效应分析, 五个调节变量对效应量均产生了统计上显著的影响。就国家实施情况而言, 任务单在印度尼西亚与菲律宾开展的研究中最为有益; 在教育阶段与学科方面, 最显著、效应最大的分别出现在中学阶段与物理学科。与此同时, 聚焦认知领域、表征丰富或以民族科学为依托的任务单, 对学习成效表现出最大且显著的效应。这些发现凸显出: 在编制任务单以最大化其效果时, 必须充分考虑教育情境、学科内容与学习者需求。

八、建议

本研究为理科教育工作者与课程开发者提供了实践指引。教师宜采用富有参与感的任务单形式, 如建构主义型、多元表征型、探究型、文化贴切型与技术整合型设计。对课程开发者而言, 重要的是编制能够指向认知、情感与动作技能三类学习成效的差异化任务单, 尤其是在视觉支持已被证明有效的物理、化学等学科中。师范院校 (TEI) 应使未来的教师有能力编制以研究为依据、有多媒体支持且具文化回应性的材料。随着数字任务单日益普及, 学校应确保公平的接入并支持学生的数字素养。未来研究应拓宽地域覆盖面, 并探讨干预时长等因素。值得注意的是, 评估情感与动作技能成效的任务单研究明显缺乏, 这是未来研究者可以着力的关键空白。在课堂层面改进理科教学, 是应对全球性挑战的根基所在; 这意味着要把任务单从填空式的纸页, 转变为涵养学生技能、价值观与科学思维方式的工具。

参考文献

- [1] 江绍祥. 智慧校园建设关键技术与应用场景. 中国教育信息化, 2020(12): 1-6.
- [2] 王蕾. 义务教育英语课标实施教学改进. 外语教学与研究, 2021, 53(03): 433-441.
- [3] 王会军. 高校思政课实践教学提质路径. 思想理论教育, 2024(08): 78-83.

- [4] 李艺. 美育融入学科教学实施策略. 中国美术教育, 2021(04): 4-9.
- [5] 褚宏启. 教育公平与教育质量协同发展. 教育理论与实践, 2026, 41(21): 4-8.
- [6] 周光礼. 高校学科建设与科研育人融合. 大学教育科学, 2021(04): 36-44.
- [7] 郑金洲. 校本教研提质增效实施策略. 上海教育科研, 2021(08): 5-10.
- [8] 刘昌亚. 义务教育课后服务长效管理机制. 基础教育, 2025(08): 14-18.
- [9] 丁钢. 传统文化融入现代教育的路径探索. 华东师范大学学报 (教育科学版), 2021, 39(04): 1-10.
- [10] 许晓东. 高校实验室育人功能挖掘. 实验技术与管理, 2020, 38(07): 210-215.
- [11] 沈建华. 劳动教育课程体系开发与实践. 中国教育学刊, 2021(09): 84-88.
- [12] 彭刚. 通识教育课程体系优化建设. 清华大学教育研究, 2021, 42(03): 1-8.
- [13] 孙杰远. 民族地区教育现代化推进策略. 广西师范大学学报 (哲学社会科学版), 2021, 59(03): 108-115.
- [14] 焦建利. 微课赋能课堂教学创新. 中国信息技术教育, 2021(16): 18-22.
- [15] Stein S. Reimagining global citizenship education for a volatile, uncertain, complex, and ambiguous (VUCA) world. *Globalisation, Societies and Education*, 2021, 19(4): 482-495.
- [16] Chutami F, Suhartini S. The effectiveness of using student worksheets in science learning on student learning outcomes. *Jurnal Penelitian Pendidikan IPA*, 2021, 7(4): 587-592.
- [17] Bartoňová M, Kričfaluši D. The methodology for creating worksheets for integrated science//*Proceedings of the 4th International Baltic Symposium on Science and Technology Education (BalticSTE2021)*. 2021: 7-15.
- [18] Burkholder E, Mohamed-Hinds N, Wieman C. Evidence-based principles for worksheet design. *The Physics Teacher*, 2021, 59(6): 402-403.