



开放科学  
(资源服务)  
标识码  
(OSID)

# 结构化摘要对论文传播效益的影响研究

宋欣雨<sup>1</sup> 周金元<sup>1,2</sup> 司宇<sup>3</sup>

1. 江苏大学 科技信息研究所 镇江 212013;
2. 江苏大学 图书馆 镇江 212013;
3. 江苏大学 计算机科学与通信工程学院 镇江 212013

**摘要:** [目的/意义] 基于倾向值匹配法 (PSM) 探讨结构化摘要在论文传播上的意义, 以推动 PSM 的发展, 及管窥结构化摘要的前景, 为促进知识传播提供新的视角。[方法/过程] 归纳出作者数量、作者学术年龄、论文篇幅、参考文献数 4 个影响论文传播效益的特征变量, 及论文下载量、被引量 2 个衡量论文传播效益的结果变量, 选取 6 种图书情报期刊开始采用结构化摘要前后共 1 年时间连续发表的 2313 篇论文作为样本, 使用 PSM 进行量化分析, 使论文传播效益波动的“果”客观科学地归因于采用结构化摘要的“因”。[结果/结论] 采用结构化摘要的图书情报期刊论文在下载量和被引量上有优于未采用结构化摘要论文的倾向, 结构化摘要在一定程度上提高了论文传播效益。数据上, 构建用于论文传播效益影响分析的数据集, 为后续研究提供参考; 方法上, 是 PSM 在文献计量领域的有益尝试, 呼吁学界加大对因果推断的重视; 结论上, 首次关注到结构化摘要对论文传播的价值。

**关键词:** 倾向值匹配; PSM; 结构化摘要; 下载量; 被引量; 因果推断

**中图分类号:** G203 G35

## Research on the Influence of Structured Abstracts on Dissemination Efficiency of Papers

SONG Xinyu<sup>1</sup> ZHOU Jinyuan<sup>1,2</sup> SI Yu<sup>3</sup>

1. Institute of science and technology information, Jiangsu University, Zhenjiang 212013, China;
2. Library of Jiangsu University, Zhenjiang 212013, China;
3. School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang 212013, China

**Abstract:** [Objective/Significance] Based on propensity score matching (PSM), this paper discusses the significance of structured abstract in paper communication, in order to promote the development of PSM, and peek into the prospect of structured abstract, so as to provide a new perspective for promoting knowledge dissemination. [Methods/Processes] The number of authors, the academic age of the authors, the length of the paper and the number of references are summarized as four characteristic

**作者简介** 宋欣雨 (1997-), 硕士研究生, 研究方向为科学计量与学术评价, E-mail: 985706928@qq.com; 周金元 (1965-), 硕士, 研究馆员, 研究方向为情报检索、文献计量与信息分析; 司宇 (1998-), 硕士研究生, 研究方向为深度学习。

**引用格式** 宋欣雨, 周金元, 司宇. 结构化摘要对论文传播效益的影响研究 [J]. 情报工程, 2023, 9(3): 92-102.

variables that affect the communication efficiency of the paper, the number of downloads and citations are two result variables that measure the communication efficiency of the paper. The 2313 papers published continuously in six Library and Information journals for one year before and after the introduction of structured abstract are selected as samples. Using PSM for quantitative analysis, the “effect” of the fluctuation of the paper’s communication efficiency can be objectively and scientifically attributed to the “cause” of the structured abstract. [Results/Conclusions] The downloads and citations of library and information journal papers using structured abstract tend to be better than those without structured abstract papers, and structured abstract improves the communication efficiency of papers to a certain extent. In terms of data, a data set was constructed for the analysis of the impact of the communication benefit of the paper, so as to provide reference for the subsequent research. In terms of methods, it is a useful attempt of PSM in the field of bibliometrics, and appeals to the academic community to pay more attention to causal inference. In conclusion, it is the first time to pay attention to the value of structured abstract for paper dissemination.

**Keywords:** Propensity score matching; PSM; structured abstract; downloads; citations; causal inference

## 引言

高水平期刊是科技强国建设中不可或缺的支撑。随着中国科技创新步伐日益加速，打造一批具有广泛影响力的期刊成为中国科技界的重要任务之一。摘要（abstract）作为论文内容的高度凝练，是信息组织的标引对象，是辅助科研人员检索和阅读的重要依据，因此成为论文的重要组成部分。现行国标 GB/T 6447-1986 将摘要定义为“以提供文摘内容梗概为目的，不加评论和补充解释，简明、确切地记述文献重要内容的短文”，并规定了摘要的五要素：目的、方法、结果、结论、其他重要信息<sup>[1]</sup>。清晰、准确的摘要协助读者快速了解文章内容，提高了学术交流的效率。在《关于推动学术期刊繁荣发展的意见》<sup>[2]</sup>等政策的引领下，如何利用摘要提升期刊影响力，促进学术传播，是学界共同关注的课题。

## 1 相关研究

20 世纪 50 年代，学界开始研究论文摘要<sup>[3]</sup>。

*Science* 于 1959 年推广使用传统型摘要（traditional abstract），即非结构式摘要（unstructured abstract）<sup>[4]</sup>。作为文章的一部分，传统型摘要通常采取一段式写法，强调语言的逻辑性和突出重点，不需要按照特定的框架结构来组织内容<sup>[5]</sup>。

1987 年，加拿大学者 R.Brian Haynes 倡导在生物医学期刊应用结构式摘要（structured abstract），相比于传统型摘要，其主要特点为信息量更大、层次更分明，具体为在标准格式中（standardized format）展示出文章的研究目的（objective）、实验设计（design）、研究单位（setting）、研究对象（patients and participants）、处理方案（interventions）、评估方法与结果（measurements and results）、结论（conclusions）<sup>[6]</sup>。1995 年，中华医学会系列杂志率先使用结构式摘要，极大推动了其在国内期刊的应用发展<sup>[7]</sup>。

三十多年来，结构式摘要由于其便于检索<sup>[8,9]</sup>、阅读<sup>[10,11]</sup>及记忆<sup>[12]</sup>等优势，已从医学领域逐渐演化推广至其他领域，结构式摘要的内容也随着需求变化而发展。如有包含研究目

的、方法、结果、结论的四要素结构式摘要<sup>[13]</sup>，包含背景、目的、方法、结果、结论的五要素结构式摘要<sup>[14]</sup>，包含背景、目的、方法、结果、结论、评述的六要素结构式摘要<sup>[15]</sup>等。

为了深入探索结构式摘要的实际应用效果，国内外学者对应用结构式摘要的可能性<sup>[16]</sup>、必要性<sup>[17]</sup>、应用现状<sup>[18,19]</sup>等展开了分析。Frederick Mosteller 等<sup>[17]</sup>论述了结构式摘要在教育类期刊中的重要作用，认为这是一种实用的创新，以系统的方式将教育研究成果转移到数以百万计的教育从业者和政策制定者手中。刘雪立<sup>[19]</sup>认为结构式摘要比传统型摘要更易读、易找与易记忆，可以在社会科学期刊中推广。Sandeep Sharma 等<sup>[20]</sup>构建了包含 29 个指标的摘要信息质量评价体系，以 6 本期刊为研究样本，发现结构式摘要比传统型摘要提供了更高质量的信息。James Hartley 等<sup>[21]</sup>设计了针对论文内容的问题，测试读者是否能在摘要中获取问题答案，结果发现相比于传统型摘要，读者在面对结构式摘要时答题的速度和准确率有所提高。可见，结构式摘要的应用优势已得到部分学者认可。但在以往衡量结构式摘要应用效果的相关研究中，研究方法颇具主观性，而单篇论文或多或少地发挥着传播知识的媒介作用，论文的关键价值之一体现在其对未来后续研究的帮助，目前鲜见有研究讨论应用结构式摘要对论文传播效益的影响。因此本文与以往研究有所不同，尝试采用因果推断法分析该问题，兼具了重要的理论和实践意义。在数据层面，构建用于论文传播效益影响分析的数据集，为相关领域研究人员提供数据基础与参考；在方法层面，创新性地使用合适的因果推断法，使论文传播效

益波动的“果”客观科学地归因于采用结构式摘要的“因”；在结论层面，管窥结构式摘要的应用前景、以期为促进知识传播、期刊发展提供值得参考的方法。

## 2 研究设计

### 2.1 构建数据集

回归本文的研究问题，即，使用结构式摘要后，学术论文的传播效益更好吗？传播效益可由认可度高的量化指标衡量。自 Garfield<sup>[22]</sup>提出引文概念后，论文间的关系开始明晰，单篇论文在学界对其他研究产生影响，被引量成为衡量文献影响力的关键指标。随着网络技术发展，下载论文行为能够被及时捕捉，下载量因有效弥补被引量的时滞性问题被广泛关注，成为评估文献影响力的补充指标<sup>[23]</sup>。被下载及被引用表示论文在一定程度上被其他学者认可或使用，从下载及引用行为出发的指标是文献计量学指标研究的重点，下载量与被引量具备作为论文传播效益评价指标的基本条件，即它们是学术传播扩散过程中的有效记录。根据以往研究经验，考虑到数据的代表性及可得性，在这里，传播效益具体表现为论文的下载量与被引量。其中，下载量、被引量峰值约在论文发表后的第 2 年及第 5 年<sup>[24,25]</sup>。因此，本研究讨论的传播效益指的是，已经累积 5 年及以上的较为稳定的文献下载量及被引量。

图书情报学是社会科学领域中较为注重稿件格式规范的学科之一，2005 年之后，部分图书情报期刊开始采用结构式摘要<sup>[26]</sup>，为进行对比研究提供了可能，CSSCI 来源期刊在一定程

度上代表了我国该领域期刊的较高水平，现今 18 种图书情报 CSSCI 来源期刊(2021-2022)中，

采用结构式摘要 5 年及以上的有 6 种，其结构式摘要包含的具体要素如表 1 所示。

表 1 图书情报期刊结构式摘要包含要素

期刊	摘要格式
《情报科学》	“目的/意义”、“方法/过程”、“结果/结论”、“创新/局限”
《情报理论与实践》	“目的/意义”、“方法/过程”、“结果/结论”“局限”
《情报杂志》	“研究目的”、“研究方法”、“研究结论”
《数据分析与知识发现》	“目的”、“方法”、“结果”、“局限”、“结论”
《图书情报工作》	“目的/意义”、“方法/过程”、“结果/结论”
《现代情报》	“目的/意义”、“方法/过程”、“结果/结论”

于 2022 年 12 月在 CNKI 数据库中选取以上 6 种期刊采用结构式摘要年份前后连续共 1 年的论文 2313 篇作为数据集，对各期刊内部在相近时间内的论文进行对比研究，基本排除论文质量、学科差异、期刊级别、出版模式及时间等干扰因素。

## 2.2 研究方法

传统的量化分析方法缺乏对内生性 (endogeneity) 和异质性 (heterogeneity) 问题的考虑<sup>[27]</sup>。在准确判断因果关系的问题上，需要一种能够有效提供因果证据的方法，以避免因选择性偏差或互为因果等产生的结论误差。基于自然变化形成的准实验对照数据为因果推断提供了良好的数据基础。1983 年，统计学家 Rosenbaum 等<sup>[28]</sup> 提出倾向值匹配法 (propensity score matching, PSM)，通过均衡实验组与对照组的干扰因素 (confounding variables) 分布，减少了选择性偏差 (selection bias)，加强了对因果关系的解释力度，保障了研究结果的稳健性。倾向值匹配法易于理解、操作简便，相比于回归估计更加直观<sup>[29]</sup>。该方法逐渐应用到企

业绩效、个体发展、医疗效果评价等领域，但在科研管理评价中鲜有涉及。国外有研究人员使用 PSM 探索被评价为高质量的“VIP”论文 (VIP, very important paper) 对被引量的影响<sup>[30]</sup>，国内则有研究人员运用 PSM 验证即时开放获取对论文被引量及下载量的促进效果<sup>[31]</sup>。本研究尝试采用该方法，减少论文样本的选择性偏差，讨论论文是否采用结构式摘要与其下载量、被引量的关系，希望通过实证得到可靠稳定的研究成果，并推动 PSM 等因果推断方法多领域、深层次的发展。

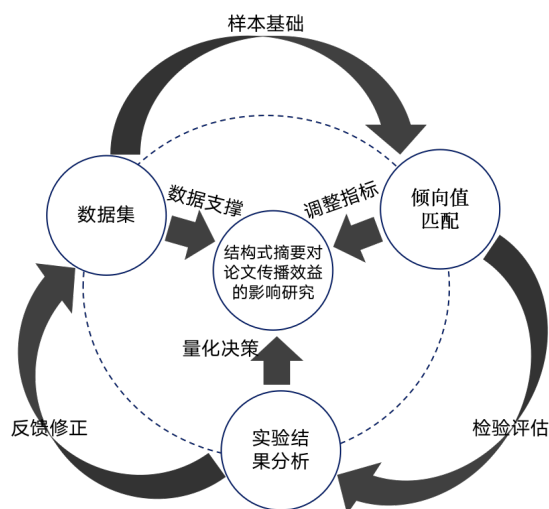


图 1 研究思路

### 2.2.1 变量设置与模型构建

设置变量包括研究变量、特征变量、结果变量。研究变量为“论文是否采用结构式摘要”，为二分变量，即，未采用结构式摘要的论文为对照组，取值0，采用结构式摘要的论文为实验组，取值1。根据以往研究<sup>[31-33]</sup>，考虑作者数量、作者学术年龄、论文篇幅、参考文献数及期刊影响因子作为可观察到的

论文下载量及被引量的影响因素，因本研究是在各期刊内部进行的对比，所以不对期刊影响因子进行控制，此外，借鉴 Larivière 等<sup>[34]</sup>的做法，将第一作者认定为主要负责人，作者学术年龄指标仅考虑第一作者。结果变量为论文 CNKI 下载量与被引量。各变量值获取或计算时间为 2022 年 12 月，变量说明如表 2 所示。

表 2 变量说明

变量类型	指标名称	描述或计算方式
研究变量	是否采用结构式摘要	论文采用结构式摘要=1，未采用结构式摘要=0
特征变量	作者数量	论文署名作者的数量
	作者学术年龄	第一作者发表最后一篇论文年份-发表第一篇年份+1
	论文篇幅	论文页数
	参考文献数	文末参考文献列表中标示的个数
结果变量	论文下载量	论文发表后在知网的下载量
	论文被引量	论文发表后在知网的被引量

将特征项进行二元 Logit 模型构建，得到预测值即倾向值，其指的是研究对象  $i$  ( $i=1, \dots, N$ ) 在平衡混淆变量后被划分到实验组 ( $Z_i=1$ ) 或对照组 ( $Z_i=0$ ) 的条件概率，是一种“事后随机”，具体公式为：

$$E(x_i) = Pr(Z_i=1|X_i=x_i) \quad (1)$$

当两个样本的倾向值越接近，说明两个样本的特征越接近，干扰因素整体水平情况相当。后续直接针对倾向值进行匹配。

### 2.2.2 样本分组与匹配

根据论文是否采用结构式摘要 (Structured Abstract, SA)，本研究将样本分为实验组 (采用结构式摘要) 和对照组 (不采用结构式摘要)，定义二元研究变量  $SA_i \in \{0, 1\}$ ，其中  $SA_i=0$  代表论文未采用结构式摘要， $SA_i=1$  代表论文采

用了结构式摘要。同时定义  $D_i^0$ 、 $C_i^0$  代表未采用结构式摘要的论文在发表后的下载量及被引量， $D_i^1$ 、 $C_i^1$  代表采用结构式摘要的论文在发表后的下载量及被引量，则是否采用结构式摘要对论文下载量、被引量的实验组平均干预效应  $ATT_D$ 、 $ATT_C$  (Average treatment effect for the treated, ATT) 分别为：

$$ATT_D = E(D_i^1 - D_i^0 | SA_i=1) = E(D_i^1 | SA_i=1) - E(D_i^0 | SA_i=1) \quad (2)$$

$$ATT_C = E(C_i^1 - C_i^0 | SA_i=1) = E(C_i^1 | SA_i=1) - E(C_i^0 | SA_i=1) \quad (3)$$

$E(D_i^1 | SA_i=1)$ 、 $E(C_i^1 | SA_i=1)$  代表实验组采用结构式摘要论文  $i$  的下载量期望及被引量期望， $E(D_i^0 | SA_i=1)$ 、 $E(C_i^0 | SA_i=1)$  代表假如采用结构式摘要论文  $i$  不采用结构式摘要时的下载量期望及

被引量期望，由于反事实情况下  $E(D_i^0|SA_i=1)$ 、 $E(C_i^0|SA_i=1)$  无法被观测，本研究将匹配出与实验组论文特征最接近的未采用结构式摘要的对照论文  $j$ ，用该论文  $j$  的下载量及被引量代替  $E(D_i^0|SA_i=1)$ 、 $E(C_i^0|SA_i=1)$ ，则采用结构式摘要对论文下载量及被引量的平均干预效应为：

$$ATT_D = E(D_i^1|SA_i=1) - E(D_j^0|SA_j=1) \quad (4)$$

$$ATT_C = E(C_i^1|SA_i=1) - E(C_j^0|SA_j=1) \quad (5)$$

常用的匹配方法有最邻近法 (Nearest Neighbor Matching) 及半径匹配法 (Caliper/Radius Matching)，最邻近法将实验组及对照组中倾向值差异最小的个体相比，使得信息被充分利用，半径匹配法则对相邻关系加入公差水平，对匹配程度要求更高<sup>[35]</sup>，本研究将根据 PSM 平行假设检验，评估特征变量均衡性，确定匹配效果较好的方法再计算平均干预

效应。

### 3 实验分析及结果

#### 3.1 论文样本篇均下载量与被引量比较分析

表 3 展示了本研究所用的 CNKI 数据库中图书情报期刊开始采用结构式摘要前后相邻时间窗口共 1 年连续发布的 2313 篇论文下载量与被引量的统计结果，数据获取时间为 2022 年 12 月，每篇期刊论文截止数据收集时均有足够长的时间统计下载量与被引量。值得说明的是，《数据分析与知识发现》在统计时间段内的刊名为《现代图书情报技术》，且因其 2013 年 7 月无发文，故以其曾用刊名检索出发表于 2013 年 6 月—2014 年 6 月的论文，其余 5 种期刊在统计时间段内均未更名与断更。

表 3 图书情报期刊采用结构式摘要前后相邻时间段的论文下载量及被引量

期刊	论文发表时间	未采用结构式摘要			采用结构式摘要		
		样本量/篇	篇均下载量/次	篇均被引量/次	样本量/篇	篇均下载量/次	篇均被引量/次
《情报科学》	2016.7-2017.6	202	740.02	17.49	190	893.05	23.01
《情报理论与实践》	2016.7-2017.6	218	845.24	17.47	126	945.95	19.10
《情报杂志》	2015.7-2016.6	218	765.27	17.18	208	937.88	20.33
《数据分析与知识发现》	2013.6-2014.6	95	554.03	19.19	80	697.71	17.88
《图书情报工作》	2014.7-2015.6	405	615.12	14.63	230	1108.31	26.10
《现代情报》	2017.1-2017.12	262	697.43	14.90	79	651.72	14.58

可以看出，除《现代情报》外的 5 种期刊，未采用结构式摘要的论文在篇均下载量层面上低于相邻时间发表的采用结构式摘要的论文，除《数据分析与知识发现》和《现代情报》外，其余 4 种期刊采用结构式摘要的论文在篇均被引量上表现更好。即简单均值分析的结论是，采用结构式摘要后，大部分期刊论文传播效益

更好，期刊受关注度更高。但该结论的科学性存疑，一方面，各期刊用于研究的论文样本量不大，易受极端值影响带来较大偏差；另一方面，作者数量、论文篇幅等因素关乎论文下载量与被引量，如只关注是否采用结构式摘要对论文传播效益的影响，应尽可能避免其他因素的干扰。因此，采用倾向值匹配法寻找期刊内作者

数量、论文篇幅等特征相似的论文样本进行匹配对比,减少特征变量之间的差异,使得进行匹配的论文样本之间唯一的区别是“是否采用结构式摘要”。

### 3.2 PSM平行假设检验

尝试采用多种匹配规则以取得较好的匹配效果,具体而言,本研究基于共同取值范围,采用了1:1匹配比例的最邻近法、0.01和0.02卡钳值的半径匹配法。为判断匹配效果,进行PSM平行假设检验,其中匹配成功比例是匹配成功个数与需要匹配个数的比值,在本研究中,各期刊需要匹配个数为“采用结构式摘

要”的样本量。通过平行假设检验检测倾向得分匹配效果,通常可由以下标准判断出匹配效果好:①匹配后“标准化偏差”绝对值小于20%;②“标准化偏差”值在匹配后下降幅度明显;③使用 $t$ 检验法比较实验组与对照组的差异,匹配前有显著性( $p<0.05$ ),匹配后没有显著性( $p>0.05$ );④匹配后 $t$ 检验均没有显著性( $p>0.05$ )。综合各项指标分析,《情报科学》与《数据分析与知识发现》采用1:1最邻近法,《情报杂志》与《现代情报》采用0.01半径匹配法,《情报理论与实践》与《图书情报工作》采用0.02半径匹配法时,匹配效果较好,检验结果表4所示。

表4 PSM平行假设检验

期刊	匹配规则	匹配成功比例	特征变量	状态	实验组	对照组	标准化偏差(%)	$p$ 值
情报科学	1:1最邻近法	100.00%	作者数量	匹配前	2.363	2.515	-15.56%	0.125
				匹配后	2.363	2.368	-0.55%	0.958
			作者学术年龄	匹配前	8.758	7.678	17.84%	0.079
				匹配后	8.758	8.495	4.34%	0.672
			论文篇幅	匹配前	6.842	6.332	13.75%	0.176
				匹配后	6.842	6.763	1.99%	0.847
			参考文献数	匹配前	23.879	21.713	16.00%	0.115
				匹配后	23.879	24.558	-4.72%	0.646
情报理论与实践	0.02半径匹配法	99.21%	作者数量	匹配前	2.571	2.390	18.11%	0.102
				匹配后	2.552	2.624	-7.71%	0.543
			作者学术年龄	匹配前	10.460	9.596	12.14%	0.284
				匹配后	10.264	9.832	6.25%	0.622
			论文篇幅	匹配前	5.968	5.518	35.40%	0.001
				匹配后	5.952	5.984	-2.95%	0.816
			参考文献数	匹配前	15.738	12.454	38.64%	0.001
				匹配后	15.656	15.200	5.44%	0.667
情报杂志	0.01半径匹配法	97.60%	作者数量	匹配前	2.361	2.486	-11.10%	0.252
				匹配后	2.389	2.330	5.36%	0.589
			作者学术年龄	匹配前	8.274	8.216	0.93%	0.924
				匹配后	8.350	7.946	6.51%	0.512
			论文篇幅	匹配前	6.288	6.055	19.96%	0.040
				匹配后	6.246	6.163	7.44%	0.454
			参考文献数	匹配前	15.091	13.156	22.97%	0.019
				匹配后	14.291	14.350	-0.76%	0.939

续表

期刊	匹配规则	匹配成功比例	特征变量	状态	实验组	对照组	标准化偏差(%)	p 值
数据分析 与知识发 现	1:1最邻 近法	100.00%	作者数量	匹配前	3.150	2.716	36.29%	0.018
				匹配后	3.150	3.250	-8.20%	0.605
			作者学术年龄	匹配前	8.100	7.832	4.39%	0.774
				匹配后	8.100	7.575	8.52%	0.591
			论文篇幅	匹配前	7.412	6.642	55.93%	0.000
				匹配后	7.412	7.425	-0.97%	0.951
			参考文献数	匹配前	9.563	9.916	-6.29%	0.678
				匹配后	9.563	9.950	-7.28%	0.646
图书情报 工作	0.02半径 匹配法	99.57%	作者数量	匹配前	2.357	2.091	23.90%	0.004
				匹配后	2.349	2.249	9.07%	0.332
			作者学术年龄	匹配前	9.826	8.032	25.40%	0.003
				匹配后	9.777	8.712	15.67%	0.094
			论文篇幅	匹配前	7.113	5.632	78.98%	0.000
				匹配后	7.105	7.240	-7.79%	0.405
			参考文献数	匹配前	15.204	9.881	57.84%	0.000
				匹配后	15.035	14.659	3.77%	0.687
现代情报	0.01半径 匹配法	97.47%	作者数量	匹配前	2.418	2.256	15.34%	0.220
				匹配后	2.429	2.403	2.43%	0.880
			作者学术年龄	匹配前	-15.722	8.905	-15.34%	0.338
				匹配后	9.831	9.623	3.16%	0.845
			论文篇幅	匹配前	6.684	6.286	26.99%	0.039
				匹配后	6.649	6.649	0.00%	1.000
			参考文献数	匹配前	16.861	14.916	20.36%	0.122
				匹配后	17.026	16.857	1.73%	0.915

### 3.3 ATT效应分析

对各期刊运用合适的匹配规则，进行 ATT 效应分析，探究匹配成功后实验组与对照组的论文下载量与被引量是否存在显著性差异，匹配后的 ATT 效应分析结果如表 5 所示。

研究显示，《情报科学》与《图书情报工作》的下载量与被引量，《情报杂志》的下载量，t 检验值在 5% 的水平上显著 ( $p < 0.05$ )，表明采用结构式摘要对这三种期刊论文的传播产生了正向影响，ATT 效应值为具体差值，即采用结构式摘要后，《情报科学》论文下载量与被引量平均提高了 189.895 次与 7.516 次，《图

书情报工作》论文下载量与被引量平均提高了 309.929 次与 7.110 次，《情报杂志》论文下载量平均提高了 110.237 次，其余三种期刊未见显著差异。分析其中原因，发现除 CNKI 外，期刊官网、微信公众号也能阅读或下载部分期刊的论文全文，而这部分数据无法精确统计，因此对实验结果存在潜在干扰。不过，6 本图书情报期刊在使用结构式摘要后均未出现下载量或被引量显著下降的现象，就本次调研的情况来看，采用结构式摘要的图书情报期刊论文在下载量和被引量上有优于未采用结构式摘要论文的倾向。



表5 ATT 效应分析

期刊	结果变量	实验组	对照组	ATT效应值	p值
情报科学	下载量	893.047	703.153	189.895	0.019
	被引量	22.763	15.247	7.516	0.004
情报理论与实践	下载量	945.952	854.808	91.144	0.335
	被引量	19.095	17.08	2.015	0.307
情报杂志	下载量	937.885	731.483	206.402	0.029
	被引量	20.327	17.473	2.854	0.259
数据分析与知识发现	下载量	697.712	587.475	110.237	0.299
	被引量	17.438	16.000	1.438	0.662
图书情报工作	下载量	1108.313	798.384	309.929	0.001
	被引量	25.761	18.651	7.110	0.001
现代情报	下载量	651.722	724.013	-72.291	0.460
	被引量	14.582	15.130	-0.548	0.801

## 4 结论及建议

本研究基于图书情报领域6种期刊采用结构式摘要前后相邻共1年时间内连续发布的2313篇论文数据,运用倾向值匹配法(PSM)检验了采用结构式摘要对论文传播效益的影响。结果表明,有2种期刊采用结构式摘要提高了论文下载量和被引量,有1种期刊采用结构式摘要提高了论文下载量。为得到科学可靠的结论,本研究在共同取值范围内进行了3种规则的匹配,并选择较好的匹配结果用于后续研究。在2.2部分采用了简单均值比较法做了统计分析,受限于极端值及干扰变量,研究结论的科学性存疑。PSM有效缓解了因论文作者数量、作者学术年龄、论文篇幅、参考文献数变量造成的内生性和异质性问题,科学衡量采用结构式摘要对论文下载量与被引量的平均干预效应,强调原因的唯一性,将随机实验思想作为基础,识别出其中的因果关系,并计算出

因果效应,是因果推断法在文献计量学领域内的有益尝试,助力PSM等因果推断法在各领域的深化发展,为判断研究中的因果关系提供可参考的解决方案。

摘要是当前论文的重要组成部分,其极具检索意义,相比于传统型摘要,结构式摘要更强调内容的完整性与叙述的层次性,在实际应用时有诸多优势:①将文章的主要信息以清晰、简洁、有序的方式呈现,使读者更容易理解文章内容和框架,进而提高阅读效率;②层次分明从而更有利于检索,提高文章的可发现性和可读性;③更易于被翻译、记忆和使用,促进跨语言和跨领域的分享、交流与合作。本研究中,采用结构式摘要的图书情报期刊论文比未采用结构式摘要的论文有更容易被下载和被引用的可能性,结构式摘要在一定程度上促进了知识传播。当然,不同学科领域的期刊和读者对摘要的需求和标准存在差异,如文学、艺术等领域的关注点不在于实验、数据等明确的结构化

部分, 结构式摘要或难以适用这些学科。因此, 在适合的学科中可以尝试应用结构式摘要。鼓励科研人员及期刊勇于创新, 不断探索信息化时代中进行学术交流的高效方式, 挖掘提高论文传播效益的方法。

综上, 本研究在数据上, 构建了用于论文传播效益影响分析的数据集, 划分出作者数量、作者学术年龄、论文篇幅、参考文献数 4 个影响论文传播效益的特征变量, 及论文下载量、被引量 2 个衡量论文传播效益的结果变量; 方法上, 用 PSM 探究采用结构式摘要这唯一原因对论文传播效益的影响, 并尝试不同匹配规则以取得较好的匹配效果; 结论上, 客观论证了采用结构式摘要对论文关键价值, 即传播价值的提升存在一定作用, 为论文及期刊增加影响力, 提升传播效益提供新的视角。但本研究仍存不足之处, 考虑到数据的可得性与易得性, 结合前人研究, 仅采用了论文作者数量、作者学术年龄、论文篇幅、参考文献数 4 个特征变量, 研究样本也仅限定在了图书情报领域, 后续可考虑更多混淆因素及学科差异, 这也是未来的研究方向。

## 参 考 文 献

- [1] 中华人民共和国国家标准·文摘编写规则 [EB/OL]. [2022-11-8]. <https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=8713727404B8704955E5CAF11EE3EB59>.
- [2] 中华人民共和国中央人民政府·中宣部、教育部、科技部印发《关于推动学术期刊繁荣发展的意见》 [EB/OL]. [2023-04-23]. [http://www.gov.cn/xinwen/2021-06/25/content\\_5620876.htm](http://www.gov.cn/xinwen/2021-06/25/content_5620876.htm).
- [3] Hartley J, Betts L. Common weaknesses in traditional abstracts in the social sciences[J]. Journal of the American Society for Information Science and Technology, 2009, 60(10): 2010-2018.
- [4] Hartley J, Sydes M. Are structured abstracts easier to read than traditional ones?[J]. Journal of research in reading, 1997, 20(2): 122-136.
- [5] 赵丽莹, 苗秀芝, 国荣. 中文科技期刊采用结构式长摘要的建议 [J]. 编辑学报, 2017, 29(S1): 59-61.
- [6] Ad Hoc Working Group for Critical Appraisal of the Medical Literature. A proposal for more informative abstracts of clinical articles[J]. Annals of Internal Medicine, 1987, 106(4): 598-604.
- [7] Liu X L, Qiao H C, Pan B R, et al. Structured abstracts in Chinese biomedical journals: the current situation and perspective[J]. European Science Editing, 1993, 48: 20.
- [8] 郑梦悦, 秦春秀, 马续补. 面向中文科技文献非结构化摘要的知识元表示与抽取研究 —— 基于知识元本体理论 [J]. 情报理论与实践, 2020, 43(2): 157-163.
- [9] Gerstein M, Seringhaus M, Fields S. Structured digital abstract makes text mining easy[J]. Nature, 2007, 447(7141): 142-142.
- [10] 宋东桓, 李晨英, 刘子瑜, 等. 英文科技论文摘要的语义特征词典构建 [J]. 图书情报工作, 2020, 64(6): 108-119.
- [11] 赵丹宁, 牟冬梅, 白森. 基于深度学习的科技文献摘要结构要素自动抽取方法研究 [J]. 数据分析与知识发现, 2021, 5(7): 70-80.
- [12] Hartley J, Sydes M. Structured abstracts in the social sciences: Presentation, readability, search and recall[J]. European Science Editing, 1995, 56: 6-7.
- [13] Hartley J. Headings in structured abstracts[J]. The British Journal of Psychiatry, 1998, 173(2): 178.
- [14] Hartley J. Clarifying the sub-headings of structured abstracts[J]. European Science Editing, 2007, 33(2): 41-42.
- [15] Hartley J. Is it appropriate to use structured abstracts in social science journals?[J]. Learned Publishing, 1997, 10(4): 313-317.
- [16] 牛燕平, 赵路华, 赵存如. 科技论文结构式摘要的正确运用 [J]. 编辑学报, 1998(3): 134-136.
- [17] Mosteller F, Nave B, Miech E J. Why we need a structured abstract in education research[J].

- Educational researcher, 2004, 33(1): 29-34.
- [18] Nakayama T, Hirai N, Yamazaki S, et al. Adoption of structured abstracts by general medical journals and format for a structured abstract[J]. *Journal of the Medical Library Association*, 2005, 93(2): 237.
- [19] 刘雪立. 结构式摘要应用研究的最新进展 [J]. *中国科技期刊研究*, 1999, 10(3): 185-186.
- [20] Sharma S, Harrison J E. Structured abstracts: do they improve the quality of information in abstracts?[J]. *American journal of orthodontics and dentofacial orthopedics*, 2006, 130(4): 523-530.
- [21] Hartley J, Sydes M, Blurton A. Obtaining information accurately and quickly: are structured abstracts more efficient?[J]. *Journal of information science*, 1996, 22(5): 349-356.
- [22] Garfield E. Citation indexes for science: a new dimension in documentation through association of ideas[J]. *Science*, 1955, 122(3159): 108-111.
- [23] 谢娟, 龚凯乐, 成颖, 等. 论文下载量与被引量相关关系的元分析 [J]. *情报学报*, 2017, 36(12): 1255-1269.
- [24] 方红玲. 我国科技期刊论文被引量和下载量峰值年代——多学科比较研究 [J]. *中国科技期刊研究*, 2011, 22(5): 708-710.
- [25] 胡泽文, 任萍, 沈佳慧. 融合 K 值算法与三指标的神经科学领域“睡美人”论文识别及影响因素探析 [J]. *现代情报*, 2022, 42(3): 147-156.
- [26] 周志超. 中文图情期刊摘要的核心要素与逻辑结构分析 [J]. *情报科学*, 2018, 36(3): 8-12+32.
- [27] 王筱纶, 赵宇翔, 王曰芬. 倾向得分匹配法: 促进数据科学视角下情报学研究的因果推断 [J]. *情报学报*, 2020, 39(11): 1191-1203.
- [28] Rosenbaum P R, Rubin D B. The central role of the propensity score in observational studies for causal effects[J]. *Biometrika*, 1983, 70(1): 41-55.
- [29] Gu X S, Rosenbaum P R. Comparison of Multivariate Matching Methods: Structures, Distances, and Algorithms[J]. *Journal of Computational and Graphical Statistics*, 1993: 405-420.
- [30] Mutz R, Wolbring T, Daniel H D. The effect of the “very important paper”(VIP) designation in *Angewandte Chemie International Edition* on citation impact: A propensity score matching analysis[J]. *Journal of the Association for Information Science and Technology*, 2017, 68(9): 2139-2153.
- [31] 郭进京, 黄奇, 盛姝, 等. 基于倾向值匹配的即时开放获取对论文引用和下载量的影响分析 [J]. *情报杂志*, 2022, 41(7): 166-175.
- [32] 张丽华, 张康宁, 赵迎光, 等. 科研人员职业生涯学术论文相似度及其对被引频次的影响分析 [J]. *情报学报*, 2022, 41(8): 822-831.
- [33] 谢娟, 成颖, 李江, 等. 文化资本与论文影响力的关系 [J]. *情报学报*, 2019, 38(9): 943-953.
- [34] Larivière V, Sugimoto C R, Bergeron P. In their own image? A comparison of doctoral students' and faculty members' referencing behavior[J]. *Journal of the American Society for Information Science and Technology*, 2013, 64(5): 1045-1054.
- [35] 胡永远, 周志凤. 基于倾向得分匹配法的政策参与效应评估 [J]. *中国行政管理*, 2014(1): 98-101.