

文章编号: 2095-1663(2015)01-0063-04

学位论文“查重热”的冷思考

卢 威

(北京师范大学 教育学部, 北京 100875)

摘 要:学位论文“查重”作为遏制学术不端行为的重要手段被广泛使用,但其实际效果是有限的,而且可能会导致一些新的问题。这表现在“查重”不仅对遏制抄袭现象作用有限,而且可能会妨碍文献的正当引用,在一定程度上影响了学术创新,并忽略了不同学科和研究方法的差异。因此必须认识到“查重”制度的局限性,在反对学术不端过程中要尊重科学规律,从主要依靠技术鉴别转向以人为主、以技术为辅的鉴别,转向建章立制,转向尊崇学术规范的文化建设。技术手段的使用须合理有度,要避免其负效应。

关键词:“查重”;学术不端;学术规范;反思

中图分类号: G643.7

文献标识码: A

所谓“查重”,也就是通过学术不端文献检测系统进行的论文相似性检测,目前已经作为预防和制裁学术不端行为的重要做法,被广泛地运用于学位授予工作中。学位授予单位根据检测报告的结果,对学位论文送审、答辩的资格条件作了严格的规定。例如,某高校规定:总相似比 $\leq 30\%$ 的学位论文,视为通过检测;总相似比 $\geq 30\%$ 的学位论文视为未通过检测,需对学位论文进行修改,修改后的学位论文可申请进行一次复检。还有的高校作出了更为严苛的规定,如重合率 $\geq 15\%$ 的,要取消答辩资格。笔者并不否认这些做法对研究生学术不端行为可以起到一定的规范作用,从而扭转不良学术风气;但也认为,此类做法也有一些将现实问题简化的成分。利用网络数据库进行“查重”的正效应不仅未必如预期设想那样明显,反而有可能会带来一些新的问题。

一、“查重”对遏制抄袭现象的作用有限

学位论文“查重”成为答辩的前置程序,是近几年来才出现的。开展这项工作的首要目的是“反抄

袭”。从近年媒体报道的情况来看,学位论文、学术论文抄袭的确是一个比较普遍的现象。据《中国青年报》报道,根据有关专家的调查,在其2007年的样本数据中,高达72%的文章是全文抄袭。^[1]在被曝光的庞大论文抄袭队伍中,不仅有本科生、硕士生和博士生,甚至还有一些高校教师、知名学者和大学校长。可以说,论文抄袭在今天的学术界已经不是什么新奇事件,而是一个人尽皆知的公开秘密。正是在这一背景下,人们不得不求助于科技的力量,于是,学位论文“查重”制度应运而生。

但是,从“查重”制度实施的效果来看,很难说这一做法能够有效遏制抄袭现象的蔓延。正可谓“道高一尺,魔高一丈”,既然有了“查重”政策,自然就会有人想出应付“查重”的对策。一方面,通过“查重”来判断一篇文章是否存在学术不端行为,其主要依据是文字的相似程度。文字复制比达到学位授予单位规定的限制,就可能被认定存在抄袭,进而做出相应的处理。但“查重”是由计算机程序来进行的,而计算机程序并不能像人脑一样,它只能机械地计算有多少文字与数据库收录的已有文献重复,却很

收稿日期:2014-07-07

作者简介:卢威(1985—),男,江苏沛县人,北京师范大学教育学部博士研究生。

难应付同义置换。既然大家对“查重”的这一缺陷心知肚明,那么,有多少抄袭者还会选择一字不改地原文复制呢?他们完全可以把想要抄袭的内容作同义替换,而换一种表达方式后,其通过的可能性显然会大大提高。再者,网络数据库收录的多为中文论文,对于那些抄书、转译外文文献之类,则未必查得出来。

另一方面,有需求就有市场,围绕论文“查重”已经兴起了一门产业,互联网上关于论文“预查重”、传授“反查重”技巧的生意早已不是个别现象。为避免在学校组织的正式“查重”中被认定为抄袭进而影响毕业和获得学位,很多学生纷纷采取到网上进行预先“查重”的做法,并根据检测结果对论文作进一步修正,从而达到学校规定的要求。据媒体报道,在毕业季,有的淘宝网店仅为毕业生提供论文“预查重”有偿服务一项,其月收入就超过100万元。^[2]当然,很多“反查重”技巧的传授并不都是收费的,一些人在互联网上还津津乐道地分享着自己的经验心得。更有甚者,还有人干脆通过代写、买卖论文来骗取学位、职称和荣誉。从这些情况看,作为反抄袭的一种普遍做法,“查重”的检测效度和震慑作用都是有限的。正是因为如此,才有学者提出“学位论文复制比检测作为一种辅助手段是必要的,但对学术不端行为的判定必须依靠专家的鉴别”^[3]。而且,在“查重产业”兴起的背景下,学术不端的花样还在持续翻新,学术规范可谓威信扫地、尊严尽失。

二、“查重”可能会妨碍文献的正当引用

“查重”查的是其他文献内容在全文中所占比例的大小,而不是抄袭比例的多少。论文“查重”通过与否,主要取决于被检论文与已发表文献的文字重复比,而且从很多学位授予单位的规定来看,这个复制比要很低才能算作通过。那么接下来的一个问题就是:究竟有多少文字重复才算合适?严格地讲,在全篇论文中只抄袭他人一句话也算抄袭,尽管检测报告显示复制比接近于零;而另一个极端的现象是,文章规范引用了大量的参考文献,用学术评价的术语来表述,即“文献综述充分、参考资料翔实”,但同时核心观点是作者自己的,“查重”则有可能通不过。实际上,我们要求有15%以下的文字重复率也好,30%以下的文字重复率也罢,其背后隐含的意思很明确,文章除了这一小部分与其它文献相同外,其它

大多数都应当是作者自己的分析论述。

不过,这个看似美好的愿望是违背科学规律的。按照托马斯·库恩(Thomas Kuhn)的说法,科学进步一是通过常规科学,二是通过范式革命。提出新范式毕竟仅有极少数人可以为之,大多数学者的研究都属于常规科学。而常规科学则是在前人知识大量累积的基础上,再向前迈出一小步。这就决定了一篇研究成果中的多数内容不可能是我们自己的创见,而更多地是对他人观点的引用、述评与综合。在一篇学位论文中,真正属于作者创见的,只是那一小部分核心观点。文中的大多内容,都是在围绕着这一核心观点服务。而且,即使是范式创新,也需要建立在系统批判旧范式的基础上,这就应当更为充分地研读、分析已有研究文献,并审慎地加以引用和评判。如果大多数学位论文中的大多数内容都是作者自己提出来的,那么很难说其中没有“水分”。

学术研究对资料具有高度依赖性,也即是“有多少资料做多少东西。”^[4]这些资料很大一部分来源于已发表的文献。对文献的引用应当以论述的实际需要为限,而不是应以“查重”规定为限。文章的哪一段话需要引用,以及引用多少,只有作者自己最了解。因此,通过“查重”制度预先地限定引用文字所占全文的比例是不科学的。一味地硬性规定文字重复率,其最后的结果要么是作者尽量减少引文,要么是作者不得不在尽量保持他人观点原意的情况下,将其用自己的“独立表述”改头换面写出来。而从根本上说,充分的引用是知识积累和学术规范的内在要求,优秀作品往往需要“旁征博引”。仅就笔者所在的人文社会科学领域而言,我们不妨翻看一下国内外那些具有广泛影响的论文和著作,很多都是正文篇幅并不长,且包括较多引文,参考文献也占据了不少版面。如《世界贸易宪法》全文11万字,但脚注就占去10万字篇幅,其正文仅有1万字,若考虑正文中还有引用的话,其原创性恐怕不足9%,^[5]但这并没有妨碍它成为法学界的著名论文。不过,如果按照我们今天“查重”的要求,这些文献恐怕连通过的机会都不大,更遑论被认定为名篇佳作!

三、“查重”一定程度上会影响学术创新

创新的重要性在今天这个后工业社会中不言而喻。学术创新的根本在于学术成果的原创性。因此,反对学术不端不仅是学术道德的要求,而且还是

学术创新的保障。而创新程度的高低,在很大程度上取决于对已有知识的重新配置。正如心理学研究所揭示的,“创造力的产生既不是知识本身,更不是知识的简单叠加,而是个体通过对本领域内已有知识的灵活加工。”^[6]而对已有知识的配置和加工则离不开丰富的文献引用。引用他人表述的多少并不是衡量创新程度高低的合理尺度,片面地减少引文只会阻碍创新。

由此看来,虽然抄袭现象越普遍,就越制约学术创新,但是作为旨在反抄袭的“查重”制度,却在一定程度上也影响了知识创新和进步。正是由于它可能会迫使一些作者在需要引用的文献面前谨小慎微,生怕“越雷池半步”,落下“过度引用”乃至“抄袭剽窃”之嫌,进而容易导致“脱离文献”搞研究,由此影响了对已有知识的配置和加工。这就是“查重”制度妨碍正当引用文献所导致负效应的延伸。如果作者一五一十地将所需文献引证过来,则很可能会影响“查重”结果,进而可能会危及其答辩资格。于是在写论文时,一些作者就不得不考虑怎样下笔,从而保证“查重”顺利通过。为此,有些人可能就会尽量减少对文献的引用,让论文通篇都充斥着“自己的话”。而学术创新的一个铁律是“站在巨人的肩膀上”,就人文社会科学领域而言,缺乏对前人文献资料的充分引证和分析,光凭“自己的话”能够使论述达到怎样的高度,恐怕很难说得清楚。

通篇充斥作者“自己的话”的论文,不仅可能达不到一定的理论高度,而且很可能是重复劳动。为了通过“查重”,作者要尽量减少引证他人的成果,这就很容易导致对一些重要著述的忽略。当作者自以为提出一个新观点并为之沾沾自喜的时候,实际上,这个观点可能已经被前人用其它的表述方式说过了。尽管这种观点上的重复算不上抄袭剽窃,但至少属于低水平重复,我们做别人已经做过的工作,对知识发展的贡献根本没有多大意义,充其量是对作者本人的一次锻炼而已。当然,达不到理论高度、低水平重复之类尚且可以算作“无益且无害”的话,那么,最为可怕的是,由于缺乏对已有文献在引用基础上进行的深入分析,这些“自己的话”甚至可能难以自圆其说。论文提出的观点看似言之成理,实则经不住推敲。诚如有学者所言:“抄袭和剽窃虽然有悖于道德伦理,但并不以提供假冒伪劣信息为目的,而那些建立在学术沙滩上的空中楼阁式的创新学术,则无疑将成为一堆需要后人花费大量时间和精力来

加以清理的学术烂尾楼。”^[7]试想,如果这样的问题大量充斥于很多论文之中,那将会误导多少后续研究?其负面影响显而易见。

四、“查重”忽略了学科和研究方法差异

不同学科、专业的学位论文,在选题、方法、表述和结论等方面大不一样,这也是常识。但是,很多学位授予单位的“查重”制度却偏偏背离了这一常识,硬是“一刀切”地给不同学科、专业的学位论文划上一条统一的上限,规定文字重复率超过了一定比例就认定为不通过。这种做法对于不同学科专业的研究,其影响程度是不一样的,因为不同学科对不同类型的文献的依赖程度并不一样。一般而言,人文社会科学对已发表文献的依赖比较明显,特别是一些经典文献,即使发表多年,仍然被学界广泛引用。而在某些自然科学领域,实验则是学位论文观点创新的重要途径,因此,自然科学领域学术不端的主要手段不仅包括抄袭,还包括数据造假。抛去效果不论,“查重”更多地是针对抄袭剽窃,但对实验中的数据造假却很难应付。

“查重”不仅忽略了不同学科、专业之间的差异,即使在同一学科、专业,“查重”制度对不同选题和研究方法的影响也大有不同。实证研究主要是需要获取大量的一手资料,对二手文献依赖性较弱。例如,对于量化研究,即使用同一套问卷,测量对象不同也会得出不同的数据,并进而影响结论;对于质性研究,就同一问题不同人也有不同的回答,个案研究的结论甚至不能被推广。而一篇学位论文若从事理论研究,例如,通过思辨提出新思想、新论断,或者梳理、考察某一思想学说的发展历史,则更多地需要“从文献到文献”的分析,必须大量依赖对已有文献的考察。因此与实证研究相比,理论性研究在“查重”中可能面临更大的风险。趋利避害是人之本能,笔者曾经问过身边一些研究生同学,为什么写学位论文往往喜欢选择量化研究,得到的回答居然十分出乎意料:这就是量化研究更有利于通过相似性检测!一般情况下,选择什么题目进行研究要么是兴趣使然,要么出于所承担课题的需要;而且按理来说,有什么样的研究问题,就应采用与之相适应的研究方法,这是最基本的要求。如果“查重”制度也在某种程度上左右着题目和研究方法的选择,就不能不引起我们的深思。

五、结语:规范学术行为应尊重科学规律

笔者无意否定“查重”的积极作用,只不过想要强调,“查重”不是万能的,其效果也并非想象的那样明显。进一步说,它更接近于“防君子不防小人”的制度。“反对学术不端”的话语具有合法化功能,在这一话语下,采取任何措施似乎都是无可置疑的。而且,“合法性机制使得组织不得不接受制度环境里建构出来的具有合法性的形式和做法。”^[8]“查重”制度一经某些学位授予单位率先采用,便迅速地被众多学位授予单位竞相借鉴。但是,善良的目的必须依靠合理的途径来实现,并非任何做法都行之有效。上面列举的种种弊病,说到底都是违背科学规律所导致的结果。为此我们应按科学规律办事。一是破除技术迷信。“查重”制度的弊端之一,在于我们过于迷信技术的力量,用电脑取代人脑,把文章交给网络,机械地作出判断。二是抄袭并不主要取决于文字重复多少,而是取决于是否明确地将引文标记为引用,以及所提出的观点是否出自作者本人;过度引用取决于引用是否超过了必要限度,以及引文是否大量来自同一文献,这也难以通过限制文字复制比加以规定。三是知识发展具有累积性,不是搭建空中楼阁。“查重”限定了文字重复率,可能会影响我们对文献的必要引用。四是对不同学科和不同研究方法的研究成果,不能适用统一的鉴别标准。

总之,“查重”制度对遏制抄袭效果的有限性、对正当学术引用的制约性、对学术创新的潜在不利影响和对不同学科与方法差异的忽略,不利于优秀学术成果脱颖而出;其造成的遵守规范者不敢引用、漠视规范者照抄不误的局面,甚至会导致“劣才驱逐良

才”的后果。为了通过“查重”,各种新的应对手段层出不穷,对本已不够澄明的学术环境造成了二次污染,如是等等,这些都是值得提起注意的。反对学术不端的大方向无疑是正确的;“查重”制度也并非一无是处,更不应该被彻底地摒弃、废除。只不过,我们必须在这一过程中尊重科学规律,规范学术行为终归还是要依靠“人的因素”。因此,我们需要对现行的“查重”制度加以改进,一方面,我们需要从主要依靠技术鉴别转向以人为主、以技术为辅的鉴别,将计算机“查重”结果仅作为进一步判断参考;另一方面,我们最终还是要依靠建章立制,依靠尊崇学术规范的文化建设。总之,技术只是人的工具,我们应合理有度地运用技术手段,避免其产生的负效应。

参考文献:

- [1] 梅刚,李玉波. 论文买卖市场是如何形成的[N]. 中国青年报, 2013-09-23(11).
- [2] 刘霖聪. 淘宝热销论文“查重”服务 网店一月收入过万[EB/OL]. http://finance.cnr.cn/txcj/201405/t20140526_515581978.shtml, 2014-05-26.
- [3] 张军,赵清华. 学位论文复制比检测:问题、对策与思考[J]. 学位与研究生教育, 2013, (11): 38-41.
- [4] 刘献君. 论高等教育研究的规范化[J]. 高等教育研究, 2013, (11): 42-48.
- [5] 邱晨辉. 剔除论文垃圾须迈过评价的坎儿[N]. 中国青年报, 2013-05-25(1).
- [6] 张晓明. 我国博士生创新能力培养误区的解读[J]. 高等教育研究, 2014, (3): 63-67.
- [7] 施爱东. 学术行业生态志:以中国现代民俗学为例[J]. 清华大学学报(哲学社会科学版), 2010, (2): 5-25, 159.
- [8] 周雪光. 组织社会学十讲[M]. 北京: 社会科学文献出版社, 2003: 76.

Reflections on “Similarity Detection” for Degree Theses

LU Wei

(Faculty of Education, Beijing Normal University, Beijing 100875)

Abstract: “Similarity detection” is widely used as an important approach to curbing academic misconduct, but its actual effect is limited and it may lead to some new problems. With only a limited role in curbing plagiarism, it may impede proper citations, affect academic innovation, and ignore the differences between different disciplines and research methods. Therefore, we must recognize the limitations of “similarity detection”, follow scientific laws in fighting academic misconduct, strengthen people-focused examinations instead of relying solely on technology, emphasize rules, and develop a culture to respect academic norms to avoid the negative effects of excessive use of technical methods.

Keywords: similarity detection; academic misconduct; academic norm; reflection