

[首页 >> 学术争鸣](#)

谈文学排行榜的评价指标与权重设置

王兆鹏

一、排行榜是一种特殊的研究方法

我们相继推出《唐诗排行榜》和《宋词排行榜》之后，引起了比较广泛的反响和争议。无论是肯定还是否定，是鼓励还是指责，是学理性的批评还是情绪性的意见，都表明读者对唐诗宋词的热烈关注，也表明文学排行榜这种形式，自有其学术意义和方法论价值。我们在这里要讨论的，不是排行榜有没有必要做的问题，而是如何改进，如何做得更科学、更完善的问题。通过讨论，我们希望把文学的定量研究提升到规范化、标准化、科学化的学术层级。

文学排行榜，是以计量分析方法为手段来观察和分析作家作品影响力的一种特殊研究方法。对于学者来说，做文学排行榜，不是一种娱乐方式，而是探求一种科学的研究方法。因为研究文学的影响力，是文学研究的一个重要组成部分，文学研究必须对不同作家作品的影响力作出定位和评价。过去我们习惯用定性分析的方法，凭主观的印象去描述和区分哪些作家影响力“很大”，哪些作家影响力“较大”，哪些作家“有一定的”影响。当文学研究发展到成熟阶段时，我们已不能满足于那种标签式的定性描述。面对文学史著作告诉我们的哪些作家作品影响力“很大”，哪些作家作品影响力“较大”，哪些作家作品的影响力“一般”或“较小”时，我们不禁要追问：这些比较和定位，究竟有多少数据作依据？没有数据的支撑，凭什么对作家作品影响力的大小高低进行定位和比较？怎样衡量作家作品的影响力是“很大”还是“较大”还是“较小”？“很大”与“较大”“较小”之间究竟是一种什么样的数量关系？一般读者，了解哪些作家作品影响较大，哪些作家作品影响较小，就可以了，但文学研究的学者和专家，必须在学理上弄清楚：究竟多大的影响力才算“很大”，多大的影响力才算“较大”或“较小”，作家作品的影响力怎样来计量衡量。

单项量化数据的影响力排行，比较容易操作，如一本书或一个唱片的发行销售量排行等，不需要数理统计的转换分析，直观而简单，我们所要讨论的是综合影响力的排行榜。

在电视市场研究中，目前全国各主要电视台对自己的各频道栏目的考核，采取的就是这种综合排行榜的方式，实行综合积分排队，排在最后的实行末位淘汰。各栏目由于节目类型不同、播出时段不同、所在频道不同、频道的覆盖率不同、观众的构成不同，相互之间不能直接比较。比如老年人喜欢的京剧节目在白天播出，不可能与晚上黄金时间播出的大众综艺节目的收视率相比。所以要进行权重修正，然后进行排队。观众看不看电视，看哪个频道的什么节目，是一种主观的选择和感受。读者读不读文学作品，读什么样的文学作品，读哪位作家的哪篇作品，也是主观的选择和感受。不同的电视节目之间可以进行量化的收视率综合排行，不同的作家作品之间当然也可以进行量化的传播影响综合排行。

在社会学领域，综合排行榜方法也很常见。国家(城市)综合实力排行榜、大学排行榜就是人所熟知的例子。无论一个国家，还是一座城市，其综合实力都是由一大堆指数构成，如国内生产总值、人均GDP、非农业人口比重、人均工资、教育支出、每百万人医院和图书馆数量，等等，最后转换成统一可比的得分，变成排行榜数据。目前流行的城市幸福指数排行也是如此。

我们认为，排行榜是一种优化处理，是将复杂的现象简单化的认识方式。传统学术是以主观感觉判断为基础，强调学者的心性顿悟、个人积累，常常缺乏规范统一的评价尺度。现代学术依托科学的发展，尤其是计算机和统计学的介入，不断走向理性、规范、科学、严谨。就像传统的中医正在不断走出主观感觉，不断走向科学严谨一样。如何将这种优化处理的研究做到科学化、标准化，是当代学者的责任。

二、排行榜的指标选择

收藏文章



周访问排行	月访问排行	总访问排行
杜甫《望岳》赏析	杜甫《春夜喜雨》赏析	白居易《长恨歌》赏析
阳光下的罪恶 (3)	柳永《雨霖铃·寒蝉凄切》赏析	杜甫《春望》赏析
民国时期自然灾害下的人与社会	杜甫《客至》赏析	宋代家族与文学【第二章】
辛弃疾《青玉案·元夕》赏析		

排行榜是以数据为基础，数据指标的完备与否就成为排行榜的结果是否科学有公信力的前提。指标设置越多，越趋向科学、合理。问题是，文学作品的许多指标无法量化统计，有些能够量化统计的，又无法取得完整的数据。能够取得完整数据的指标，其价值和置信度也不同。比如我们要谈论一个诗人的影响力，我们就渴望知道：他有多少作品传世，作品有多少版本，历代刻印发行量多大，码洋多少，多少人学习、模仿、谈论过、选过他的作品，海外知名度如何，石刻作品多少次，多少地方有他的遗迹和纪念地，美誉度或恶评如何，画像有多少，有关他的文物拍卖价多高……总之，这些指标越多越详细越好。可惜，它们大部分无法取得准确的统计数据。

我们在做《唐诗排行榜》时，考虑过要统计每一首诗对后世诗人创作的影响，但数据采集相当困难。比如，我们可以知道后世哪些人学过杜甫，但我们无法判断后世诗人哪首诗是学杜甫或学杜甫的哪首诗。举几个例子可以做到，但要作为数据源来统计，就很难进行。做《宋词排行榜》时，为了统计每首宋词对后世词人的影响，我们采集了后世词人追和宋词的情况，但这也只是后世词人学习仿效宋词的一个方面，创作方法和创作精神上的仿效暂时还没有好的办法采集数据来统计。作家作品对他人和后世的影响，有些是表层字句或用韵方式的模仿或化用，有些则是精神层面的借鉴与仿效^①。前者的影响是显性的，后者的影响是隐性的。有时隐性的影响比显性的影响更重要。然而，显性的数据尚难采集，隐性的数据就更难集中采样了。随着数字化技术的发展，计算机的智能识别可在一定程度上解决这个问题。现在的数字化技术，可以进行语义分析和诗句的相似度分析。比如，只要数据库中有《全宋词》和《全唐诗》，计算机就可以自动分析任意一句宋词的遣词用字与唐诗的哪一句相似或相同。随着这种语义分析技术的进一步发展，文学创作中“互文性”（一部作品与其他作品之间的显性或隐性的关系）的分析与数据统计有望成为现实。但在现阶段，还没有多少现成的成果可以利用。早些年，我们对宋代词人的影响力进行排行，采用了存词名次、版本名次、品评名次、研究名次、历代词选名次、当代词选名次等六种指标；而《宋词排行榜》对词作的排行，采用了词选、互联网页、评点、研究论文和唱和等五种指标。这虽然是一种谨慎的做法，但绝对不是一种完善的做法。这些指标，都是显性的，而且是不完全的（有些可用的显性数据尚未采集），隐性的数据根本就没有顾及到。

对比“国家综合实力排行榜”、“城市综合竞争实力排行榜”的数十个指标，我们的唐诗宋词排行榜只用三、五个指标，当然太少，难以让人信服。努力寻找新的排行指标，增加指标的丰富度，借以提高可信度，将是改进文学影响力评价指标体系的主要途径。理想的情况应该找到十种以上的排行数据源，然后进行数据处理。我们目前能考虑到的作品影响力的评价指标，大约可分三类：作品的典范性、美誉度和传阅度。

一、典范性。典范性是就作品对他人创作的影响而言。一部作品的艺术性再高、开创性再大，如果没有人仿效，其典范性就无从确立。因此，典范性必须依存于他人的作品中。作品典范性的测度，有两个层面可以考虑：

显性层面：如语句的借用（含集句）、化用、追和、模拟、改作（编）等。如五代翁宏《春残》诗的“落花人独立，微雨燕双飞”，被宋人晏几道一字不改地借用到他的《临江仙》词而成为名句。又如五代卢延让《松寺》的“两三条电欲为雨，七八个星犹在天”，被辛弃疾的《西江月·夜行黄沙道中》化用为“七八个星天外，两三点雨山前”。台湾王伟勇教授有《宋词与唐诗之对应研究》^②一书，专门考察宋人词句化用唐人诗句的情况，可以取资作为数据源来统计。只是其书还不全，仅以几位词人为例，而不是全面考察宋词化用唐诗的情况。这种借用、化用、集句的情况，今后可以利用计算机来自动分析和统计。至于宋词化用唐五代词或宋词化用宋词之例，更是举不胜举。如李煜的“一江春水向东流”名句，被欧阳修《踏莎行》化用为“离愁渐远渐无穷，迢迢不断如春水”，被秦观的《满庭芳》化用为“便做春江都是泪，流不尽，许多愁”等。至于风格句式的模拟，如李白的《登金陵凤凰台》“凤凰台上凤凰游，凤去台空江自流。吴宫花草埋幽径，晋代衣冠成古丘。三山半落青天外，二水中分白鹭洲。总为浮云能蔽日，长安不见使人愁”以及《鹦鹉洲》“鹦鹉来过吴江水，江上洲传鹦鹉名。鹦鹉西飞陇山去，芳洲之树何青青。烟开兰叶香风暖，岸夹桃花锦浪生。迁客此时徒极目，长洲孤月向谁明”这两首诗，都有模仿崔颢《黄鹤楼》的痕迹，无论李白是有意比拼，还是一种无意识的效法模仿，后者对前者的影响是明显的，而且有本事记载可以证明。宋词中有不少作品在题序中直接交代是“效”某人或某作之“体”，如苏辙《调啸词·效韦苏州》，郭祥正《醉翁操·效东坡》，李吕《醉落魄·效山谷道人“陶陶兀兀”之句，法其体，作此以遣兴云》，辛弃疾《丑奴儿·博山道中效李易安体》、《暮山溪·赵昌父赋一丘一壑，格律高古，因效其体》、《归朝欢·效介庵体为赋，且以菖蒲绿名之。丙辰岁三月三日也》、《玉楼春·效白乐天体》，吴儆《暮山溪·效樵歌体》，侯真《眼儿媚·效易安体》，刘学箕《贺新郎·用辛稼轩金缕词韵述怀。此词盖鹭鸶林寄陈同甫者，韵险甚。稼轩自和凡三篇，语意俱到。捧心效颦，辄不自揆，同志毋以其迂而废其言》，蒋捷《水龙吟·效稼轩体招落梅之魂》，仇远《合欢带·效柳体》，刘将孙《南乡子·重阳效东坡作》，吕南公《调笑令·效韦苏州作》等。这些都是可以观察、分析得出的

显性数据，应该作为指标进入评价体系。我们可以考虑给每首作品建立借用与被借用、化用与被化用、追和与被追和、模拟与被模拟的档案库，作为数据源备用和统计。

一首诗词作品的影响力或互文性，不仅体现在同类体裁的诗词作品中，也体现在其他体裁的文学作品中。文赋、小说、戏曲等作品，也会化用或引用诗词中的名句或成句。如元曲就经常化用宋词中的名句和语句。人们熟知的范仲淹《苏幕遮》的“碧云天，黄叶地。秋色连波，波上寒烟翠”，被王实甫的《西厢记》[正宫][端正好]点化为：“碧云天，黄花地，西风紧，北雁南飞。晓来谁染霜林醉？总是离人泪。”王维《送元二使安西》中的“劝君更尽一杯酒，西出阳关无故人”，常被明清时期的通俗小说所引用。凡写及友人送别，总是说“正是‘劝君更尽一杯酒，西出阳关无故人’”。“劝君更尽一杯酒，西出阳关无故人”，成为后世小说乃至日常生活中送别的流行话头。这类现象，也应该作为数据源来采集统计。

除了文学作品的效仿和化用之外，还有日常生活语言和书面语言的引用。有些诗词名句，已进入到人们的日常生活语言之中，如“欲穷千里目，更上一层楼”、“每逢佳节倍思亲”、“谁言寸草心，报得三春晖”、“衣带渐宽终不悔”等诗词名句，已经成为熟语或引用语，常常为人所引用。诗词名句被非文学作品引用的情况，也可以作为一个指标纳入统计范围。这种被引用的情况，可以借助语言学的语料库来检索和抽样统计。

隐性层面：如创作方法的借鉴、审美范式的遵从、题材表现的启示等。上述“效某人某作之体”，有的就是属于创作方法的借鉴和效法，只是因为作者注明了是效仿某人某作，隐性就变成显性了。有的虽没有标明效某体，实际上就是效某体。只因没有在题序中标明，显性模拟就变成隐性仿效了。如刘过的《沁园春·寄辛承旨》：

斗酒彘肩，风雨渡江，岂不快哉。被香山居士，约林和靖，与东坡老，驾勒吾回。坡谓西湖，正如西子，浓抹淡妆临镜台。二公者，皆掉头不顾，只管衔杯。白云天竺飞来。图画里、峥嵘楼观开。爱东西双涧，纵横水绕，两峰南北，高下云堆。逋日不然，暗香浮动，争似孤山先探梅。须晴去，访稼轩未晚，且此徘徊。在构思和方法上就是效仿辛弃疾的《沁园春·将止酒，戒酒杯使勿近》：

杯汝来前，老子今朝，点检形骸。甚长年抱渴，咽如焦釜，于今喜睡，气似奔雷。汝说刘伶，古今达者，醉后何妨死便埋。浑如此，叹汝于知己，真少恩哉。更凭歌舞为媒。算合作平居鸩毒猜。况怨无大小，生于所爱，物无美恶，过则为灾。与汝成言，勿留亟退，吾力犹能肆汝杯。杯再拜，道麾之即去，招则须来。③这两首词表面上毫无相同之处，但对话体的运用和幽默调侃的风格却相当神似。这类暗中仿效之作，有的并没有研究成果可以参考。又如，苏轼的《念奴娇·赤壁怀古》，不仅成为后人纷纷追和仿效之作(显性)，更开辟出一种风格范型，连其词调名也被后人根据词中的句子改为《大江东》、《酹江月》等，这种风格范型的建立与后人的追步效仿，不容易找到显性的数据，但可以参考古今研究成果，或采用当代诗学专家、词学专家访问表的方式，以采集相关数据。

典范性的作品，被他人文学作品中效仿、化用的情况，可称为“效用率”；被其他非文学类作品(文章和著作)引用，可称“引用率”。效用率与引用率应是衡量一首作品影响力、生命力的重要依据。

作品的典范性，除了采集案例作为数据统计之外，还可以考虑其质量等级之分。开宗立派之作(如苏轼《念奴娇·赤壁怀古》)与一般为人模仿之作应有等级质量的区分。这类级别的区分，或可参考文学史研究的成果。唐宋词排行榜的指标体系，也采用了文学史的叙述数据，但着眼点主要是在量差，不在质差。如何区分和衡量一部作品在文学史进程中的价值和意义，是值得进一步思考的问题。

二、美誉度。美誉度，指读者对作品的评价。我们做唐宋词排行榜时，也采集了历代评论家的评点和今人的研究论著作为指标，但没有区分这些评价是正面的肯定还是负面的否定、是美誉还是恶评。只顾及了评点、研究的量差，没有顾及到质差。一次正面的高度肯定，其影响力也许胜过十次负面的批评。因此，综合统计一篇作品所获的评价，不仅要考虑其所获评价的数量(人次)，更要考量所获评价的质量级别。评价数量的多少(人次)，只能反映一部作品受关注的程度(即关注度)；评价的质量高低，才能反映一部作品的好坏优劣。因此，在采集关注度(被评的次数)数据的同时，还要补充美誉度(被评的级差)的指标。根据历代评论家的不同评价，可以考虑将美誉度分为四个等级：杰作(上等)、佳作(中等)、有瑕疵(下等)、平庸(末等)。然后将四个等级转换成分数，与关注度一起作为指标进入综合评价体系。

不同的作品，在不同的时代、不同的群体、不同的读者那里所获美誉度不同。同一作品，在不同的时代、不同的群体、不同的评论家那里所获美誉度也不一定相同。同一作品，有的可能是正面的评价，有的可能是负面的评价，即使同为正面或负面的评价，但程度有差别。因此，必须先综合每一首作品的美誉度得分，然后才能比较众多作品的美誉度得分。

同一作品，获得的是完全不同的评价，如何得出一个综合的评价结果？是将评价的级别转换成分数进行平均计算？还是参考评论者的身份看谁的评论较准确、较客观、较有影响？这些问题，还需要我们的评价体系更加细化和标准化。细化和标准化，恐怕不能由一个人说了算，需要讨论，达成共识。我们

常常觉得文学研究找不到问题、找不到题目可做，其实当深度进入定量分析的研究之后，深层次的问题就越来越多，需要研究的问题也越来越多。

三、传阅度。指作品被传播的广度(或说到达率)、被阅读的人次和频次。一部作品传播越广、被阅读的人次越多，表明作品的到达率越高、知名度越高、影响力越大。唐诗宋词排行榜在测度诗词作品的传阅度时，只采用了选本和互联网两个指标。现在看来，确实还远远不够。

就选本而言，不仅采集的选本数量(种数)有待进一步提高和完善，选本种类更需拓展和增补。一首诗词，不仅通过总集性选本(如《唐诗选》、《宋词选》之类)来传播，也通过别集性选本(如《杜甫诗选》、《李白诗选》、《苏轼词选》、《辛弃疾词选》等)来传播，还通过作家别集(如《杜工部诗集》、《李太白诗集》、《东坡乐府》、《稼轩长短句》等)来传播。也就是说，读者阅读一首诗词，既可能用总集性选本来读，也会借助别集性选本来阅读，还会利用作家本人的别集(全集)来阅读。不同类型、不同层次、不同年龄的读者，阅读、了解诗词的读本也会不同。少年儿童，主要通过童蒙读物来阅读，中小學生主要靠教材课本来阅读，汉语以外的海外读者要靠译本阅读。所以，仅就诗词选本(读本)而言，还应把别集性的选本和作家别集、少儿童蒙读物、中小学教材、海外译本等都作为数据源，纳入统计的范围。

读者阅读了解唐诗宋词，不仅是利用书籍形式的选本(读本)来阅读，还会通过歌曲的形式来传唱、聆听和欣赏。在唐代，唐诗是可以入乐歌唱的，入乐歌唱后成为声诗。虽然每一首诗传唱的地域广度和时间长度，我们已经无法知晓，但哪些诗入乐歌唱过，还是有资料可以查考。任半塘先生的《唐声诗》就可以作为数据源来统计，明清时代的琴谱等音乐史料也有记载，可以辑考。宋词在宋代的传唱，更是尽人皆知的事实，只是每首词传唱的具体情形，已不甚了了。但大体而言，在北宋时代，柳永、秦观和周邦彦词传唱较广，是铁定的事实。在统计宋词的传阅度时，如何将这类传唱的隐性资料纳入数据源，也是值得考虑和研究的问题。到了现当代，不少唐宋诗词被谱曲歌唱，有的已成为人们耳熟能详的流行歌曲，如苏轼的《明月几时有》、李清照的《一剪梅》等，有的是依原词谱曲，有的则有改动增补。这类作品也应作为数据源来考虑。还有的词作，被现当代的影视作品谱成歌曲演唱，如近期热播的电视剧《甄嬛传》就将温庭筠的《菩萨蛮》(小山重叠金明灭)谱曲歌唱，颇受观众的欢迎和喜爱，影响很大。温庭筠这首词，原来不为一般读者所熟知，但通过电视剧插曲的反复歌唱而广为人知。这类传播，在当下电子音像发达的时代，已成为比较普遍的现象，也应该作为数据源纳入统计的视野之内。

除了歌曲演唱，绘画和书法、题壁和石刻等传播方式对诗词作品的传播，同样应作为数据源来统计。不少诗词作品，成为绘画题材，画家不仅按诗意作画，还将原诗词题写于画面上，成为一种增殖性传播。如王维的《送元二使安西》，宋代著名画家李公麟曾据其诗意画作《阳关图》，并题王维原诗于画上，成为宋元二代的名画，苏轼、黄庭坚等名流都有题跋^④，书法与绘画相得益彰，不仅实现了传播增殖，更加速扩大了王维原诗的影响力。还有的诗词，常被书写成书法作品，这类书法作品，既是艺术品，也是诗词的特殊传播载体，有着特殊的传播效应。这种诗词入绘画和入书法艺术的情形并非个别现象，也有现成的史料(如上海书画出版社出版的《中国书画文献索引》等)来检索和统计。题壁和石刻，在古代也是一种很普遍而有效的传播方式，题壁和石刻的作品，以及作品的其他遗迹和纪念场所，等等，也应列作数据源进入评价体系。

以上诸项，如能逐步建成一个数据采样模型和数据来源库，将会提升文学定量研究的规范化、标准化和科学化水平。

文学作品的影响力，是读者(接受者)的阅读反应。典范性、美誉度和传阅度三类指标分别对应三个层次读者的阅读反应。典范性体现的是创作型读者(作家)的阅读反应和接受情况，美誉度可表明批评型读者(专家、学者)的阅读反应和接受情况，传阅度指标主要反映消费型读者(普通大众)的阅读反应和接受情况。

至于作家影响力的评价指标，可以参考作品的典范性、美誉度和传阅度来设置，但要有所扩展。作家的典范性，是其全部作品典范性的集合。作家的美誉度和知名度，就既要考虑文学领域的评价和影响，还要考虑他在社会、文化、思想、政治、教育等其他领域的影响。像屈原、陶渊明、杜甫和苏轼等人，他们的影响力，远远逸出文学的范围，已深入到国人的道德观、价值观、人生态度和行为方式等方面。这方面的数据，如何获取？是否可以考虑从所有中国文学和人文社会科学的著作文章(包括报纸、杂志、网络)中，搜集整理相关资料数据(包含评价、提及、画像和纪念日、场馆、纪念品等)。处理这些资料数据时，既要统计一个作家被提及的数量频次，也要考虑被提及和评论的内容，根据被提及的方式和评价的高低，分为若干等级，进行“质”的评估。如果有可能，最好能像语言学建语料库那样，也建设一个包含人文社会科学各种著作的文学史料库，以全面检索统计各种人文社会科学论著中引用、提及作家作品的情况。

作家的传阅度，其实就是其作品传阅度的整合，除采集上述作品传阅度的相关数据之外，还应关注

其别集版本的流传情况，包括评注本和抄刻本等。既要考虑版本数量之多少，也要考虑版本种类(评、注、译等)之多少，在现当代，要尽可能反映出每个版本(版次)的印数和发行量，以充分衡量评价的有效度。

在综合考察作品典范性、美誉度和传阅度的基础上，还应考察作家的名篇贡献率，即一个作家，特别是诗人词人，一生中创造和贡献了多少名篇。名篇贡献率，应该是衡量诗人词人影响力和文学史地位的重要指标。我们很难设想，一位在文学史上影响很大、地位很高的诗人，是没有贡献多少名篇的诗人。名篇与名家究竟是什么关系，也可以借名篇的贡献率获得解释和印证。

兹将上述作家作品影响力的评价指标列作一表，以求醒目(见下页表)。

三、排行榜指标的权重设置

权重的设置，主要有客观赋权和主观赋权两种。客观赋权，是以定量分析的方法，根据数据的完备程度、分布密度和活跃程度等计算出各指标的权重；主观赋权，是用定性分析的方法，根据各个指标影响因子的大小来人为设定各指标的权重。主观赋权法，又有专家评判法和层次分析法两种。唐诗宋词排行榜主要用主观赋权法中的专家评判法，虽然有一定的学理依据，但因为是初次尝试，没有前行的成果可资借鉴，也未经众多专家的讨论形成共识，所以对相关权重的设定，只能是“一家之言”，未必能得到公众和专家学者的认可。

文学排行榜评价指标的权重设置，最理想的状态是客观赋权法和主观赋权法相结合。可以考虑：各种类别的指标用客观赋权法，通过软件计算来确定；而各个具体指标，用主观赋权法来设置，定性分析中参用定量数据。比如选本，既要考虑选本这一大类指标在整个评价指标体系中的权重(用客观赋权法计算)，又要考虑每个选本指标的权重差异。试举几个宋词选本为例说明：

南宋曾慥的《乐府雅词》编成于南宋高宗绍兴十六年(1146年)，黄升的《花庵词选》成书于理宗淳祐九年(1249年)，明代陈耀文的《花草粹编》初刻于万历十一年(1583年)，清代朱彝尊编选的《词综》始刊于清康熙三十年(1691年)，近人朱彊邨《宋词三百首》初印于1924年，宋词排行榜中采用选本最晚的一种是2003年高等教育出版社出版的《中国古代文学作品选》。从传播、出版的周期(时间长度)来看，如果以2008年统计得出《宋词排行榜》的结果为下限，那么，《乐府雅词》的出版周期已有862年，《花庵词选》的周期有759年，《花草粹编》有425年，《词综》有317年，《宋词三百首》为84年，《中国古代文学作品选》仅5年。传播周期越长，其传阅度应该越高，到达率、知晓率应该越高。如果不加区别地将每个选本入选的数据同等看待，当然是不公平、不合理的。这还仅仅是时间上的传播周期的差别，还没考虑空间上的传播广度。传播广度无法用数据来衡量，但可以依据每个选本的版本种数(印次)、印数(发行量)来评估。发行数百万册的《宋词三百首》肯定比发行几万册的《中国古代文学作品选》的到达率、传播广度要大得多。因此，每一部词选，由于传播周期(长度)和发行量(广度)不同，而具备不同的传播含金量，也就是传播实际效率。因此，在统计不同作品入选频次之前，应对所有词选进行这种含金量的系数评定，使每一部词选都有不同的传播效率系数。一首作品入选不同的词选，即获得不同的系数积分。这与现代电视的收视率统计原理大致相同。

传播周期与传播广度又如何设定分值？唐诗宋词排行榜是按朝代来区分权重，尚嫌粗略。是否可以按年限来计算传播周期的得分？比如，上述《乐府雅词》的周期为862年，如果1年计1分，那么其传播周期得分为862分，而《宋词三百首》得84分，《中国古代文学作品选》得5分(具体统计时当然会对这些数值作标准化处理，否则数值过大，不便于比较分析)。那些出版年代不详的选本，可以据其成书的时代虚拟一个大致周期时限。传播广度的得分，古代词选按版本种类计算(如1种版本得1分)，现代词选按版次计算(重印或再版1次计1分)，现代词选还应考虑其印数的多少。传播周期数与版本(版次)数的乘积，经过标准化处理后转换成选本的效率系数。这些印数和发行量的数据，现在有的是可以获得的。时过境迁之后，就很难说了。所以，我们应该注意搜集和保存这些出版发行量的资料数据。

在考量每个选本的传播周期时，还应考虑每个选本的传播时间差。这二者表面看起来，都涉及选本出版的时间长度，但实际上还是有区别，传播周期是就选本在整个传播历程中的长度而言，传播时间差是指同时代的选本出现的先后早晚。如唐诗选本中，盛唐的选本只可能入选初、盛唐人的诗作，无法入选此后中、晚唐的诗篇，以盛唐选本来统计和比较整个唐代诗人诗作的人选篇次，显然对中、晚唐的诗人不公平，因为他们没有“入选权”。同样，宋人词选中，南宋初期成书的《乐府雅词》，就无法入选南宋中叶以后的词人，如辛弃疾、张孝祥、陆游和姜夔、吴文英等人没有“入选权”；在黄升的《花庵词选》中，宋末的周密、王沂孙、张炎、蒋捷和文天祥等人，也没有机会入选。初唐人与晚唐人相差了二百多年，宋初的词人比宋末的词人早了近三百年，前者有更多的传播机会。把他们进行同等的传播数量统计，必须将这个时间差考虑进去，以权重系数抹平彼此的时间差。评点数据中，对同时代被评的作家作品，也受这种时间差的影响，必须一并考虑其权重以平衡客观数据存在的差异。

古代点评指标和今人研究论著指标的权重，也可以考虑参酌上述选本指标的赋权法，每个指标乘以一定的传播周期和传播广度。

以上只是初步的思考，希望能引发广泛讨论，逐步形成共识，共同推进文学量化的研究。

注释：

①参见尚永亮《中唐元和诗歌传播接受史的文化考察》，武汉大学出版社2010年版，第235～240页。

②台北文史哲出版社2004年修订版。

③二词分别见《全宋词》，中华书局1965年版，第2143、1915页。

④参见王兆鹏《〈阳关图〉与〈送元二使安西〉图画传播》，《中国韵文学刊》2011年第1期。

[关于我们](#) | [联系方式](#) | [意见反馈](#) | [投稿指南](#) | [法律声明](#) | [招聘英才](#) | [欢迎加盟](#) | [软件下载](#)

永久域名: literature.cass.cn E-Mail: wenxue@cass.org.cn

版权所有：中国社会科学院文学研究所 京ICP备05084176号 