

18 年了，你到底要让我等多久？——论《心理学报》的发表周期（2003~2020）

吕小康

南开大学社会心理学系
南开大学计算社会科学实验室

摘要：采用 R 语言的 **rvest**、**tidyverse** 及 **ggplot2** 系列包，利用《心理学报》官网爬取并统计 2013 至 2020 年间正刊上发表的 2281 篇实质性论文的信息，基于其收稿日期与出版日期计算所有论文的发表周期（发表日期-收稿日期）并按年度汇总统计。结果显示，此 18 年间《心理学报》的年均发表周期在 290~432 天之间，最长的 3 年分别为 2007（432 天）、2015（417 天）和 2014（414 天）年，最短的 3 年分别为 2005（336 天）、2003（294 天）和 2004（290 天），分布较为均匀；单篇文章的发表周期在 45~1561 天（约 4.28 年）之间，方差较大；不论短周期还是长周期文章多集中于 2014 年及之前，2015 年及之后的发表周期波动相对较小。这说明最好（坏）的日子可能已经过去，《心理学报》的投审周期已进入常态化运行的稳定阶段。

关键词：我本将心；奈何明月

1 引言

有首非著名情歌是这么唱的：

你到底还要让我等多久/你可知我有多难受/我不会随随便便爱了就走/我的爱已覆水难收

你到底还要让我等多久/把心掏给你够不够/别再为你的放纵找借口/把话说清楚你再走

眼看着爱情流走/谁愿意善罢甘休/是让爱继续还是要分手/已到了决定的时候

—— 龙晓飞：《到底让我等多久》

声明一下，这歌不是唱给情人的，是唱给审稿人和编委会的。18 年了，号称“中国最具国际影响力学术期刊”的国内心理学刊物——《心理学报》（以下简称《学报》）——究竟是

如何对待我们投稿人（打工人）的？

评价指标当然不一。拒稿率也许是令人伤心的指标，而发表周期（包括投稿检查、外审、修改、接受到出版的全过程）是另一个令人关心的指标。基于数据可获得性的局限与个人精力的有限，这里暂只探讨《学报》的发表周期。其中，发表周期的操作定义为《学报》官方网站给出的收稿日期与出版日期（精确到日），用出版日期减去收稿日期，即得到发表周期的天数（可转为周数、月数或年数，如下图）。

The screenshot shows the journal's header with logos, ISSN (0439-755X), CN (11-1911/B), and publisher (Science Press). The navigation bar includes links like '首页', '期刊介绍', and '投稿指南'. The article information section displays the title '音乐训练与抑制控制的关系：来自ERPs的证据' by Chen Jie, Zhou Yan, and Chen Jie. It lists the submission date (2020-06-14), publication date (2020-12-25), and release date (2020-10-26). The funding sources are also listed at the bottom.

2 方法

2.1 对象

由于 2002 年及之前的许多文章并不在网页上直接公布收稿日期，故只选取 2003~2020 年发表的所有实质性论文，剔除增刊、年度审稿专家名录、会议通讯与纪要、年度总目录（另 2018、2019、2020 未见总目录）、重要心理学家的悼文与诞辰纪念文、特邀编辑发言或专栏介绍等链接；剔除《中国心理学会相关伦理守则与管理条例》等工作条例链接。对所有论文进行重复性检验，确保唯一性。对部分缺失的收稿信息通过下载原文确定收稿日期并进行人工补充。最终得到 2281 篇论文的相关信息（见 GitHub 链接文件：xb_03to20.xlsx）。

不确定 2281 篇文章是否完成对应，但即便有一定误差，应属“爬虫误用”，不影响主要结论。



迷之自信

2.2 工具

采用 R 4.0.1 及 **rvest** 包、**tidyverse** 及 **ggplot2** 包作为爬虫、清洗、分析和绘图工具。

2.3 代码

为避免恶意爬虫造成《学报》网站服务器的不必要负担，以及由于链接更新或页面布局变化造成的链接失效，仅提供基于整理后清洁数据（xb_03to20.xlsx）的分析及绘图代码（见附录）。

对爬虫做简要提示如下：

（1）首先提取并拼接所有期的链接，再次提取所有文章的链接，最后人工剔除所有上述“非实质内容”的文章链接再开始解析网页进行爬虫。

（2）2003 年第 1 期至 2012 年第 8 期的文章没有 doi，而是按卷、期和文章号等呈现链接，同时这些文章在文章页（指这一文章单独的网页链接）还有“审回日期”；而之后的文章根据 doi 呈现链接，且这些文章的文章页没有“审回日期”。为简单起见，可对这两部分进行单独爬虫，最后再行合并。

3 结果

话不多数，直接看表。

表 1 《心理学报》年度发表周期（2003~2020）

年份	均值（天）	SD（天）	N（篇）
2003	294	103	118
2004	290	137	101
2005	336	136	112
2006	397	135	110
2007	432	185	127
2008	378	157	138
2009	364	150	125
2010	417	165	114
2011	357	126	136
2012	367	126	147
2013	354	140	123
2014	414	147	159
2015	376	131	133
2016	408	114	142
2017	387	138	139
2018	397	123	126
2019	396	119	117
2020	358	125	114

从表 1 可见，2003~2020 的 18 年间，《心理学报》的年均发表周期在 290~432 天之间，最长的 3 年分别为 2007（432 天）、2015（417 天）和 2014（414 天）年，最短的 3 年分别为 2005（336 天）、2003（294 天）和 2004（290 天），分布较为均匀。

同时还有一个好消息和坏消息。好消息是，近 5 年来（2016 年以来），年均发表周期总体呈下降趋势；坏消息是，每年发表的文章总数呈肉眼可见的下降趋势。建议编辑部统计一下投稿数量，以计算出投录比。

以下是不怕搞事的两个表。

表 2 10 篇最短等待期文章

标题	作者	天数	发表刊期
我国民众对 SARS 信息的风险认知及心理行为	时勘等	45	2003, 35(4)
工作记忆中汉字与空间的分离及动态优势半球的 ERP 效应	王益文等	50	2004, 36(3)
谈谈心理学研究设计的基本逻辑（副标题略去）	莫雷	62	2004, 36(2)
老化和其他研究领域中心知心理学与认知神经科学的整合	Lars-Göran Nilsson	62	2009, 41(11)
ERP 研究反映感数与计数的不同脑机制	罗跃嘉等	76	2004, 36(4)
关键字母法对中国学习不良儿童英语词汇记忆的作用	林崇德等	76	2004, 36(4)
杏仁核 c-fos 的表达与条件反射性免疫抑制	林文娟等	79	2004, 36(4)
CHRM2 基因 rs1824024 多态性与青少年早期抑郁的关系	王美萍等	79	2010, 42(8)
工作记忆的保持与操控：腹侧和背侧额叶皮层的特异性功能磁共振成像活动	Sara Pudas 等	80	2009, 41(11)
“知道自己知道”与“知道自己不知道”——一项有关知道感（FOK）的脑成像研究	罗劲等	81	2004, 36(4)

请自行检索文章的全部作者，可以发现一个铁律：大佬真好，大佬成为大佬真早。是的，我大爷永远是我大爷。

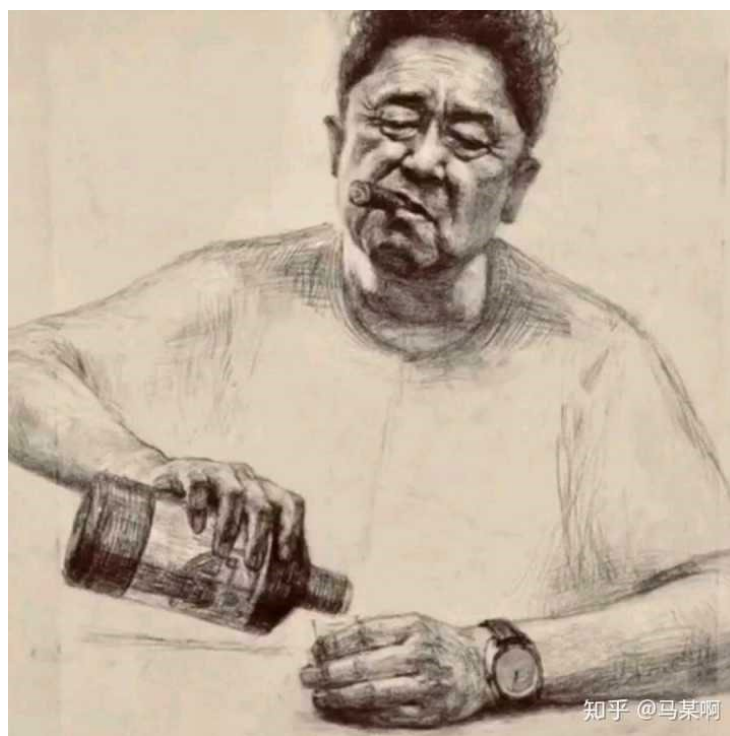
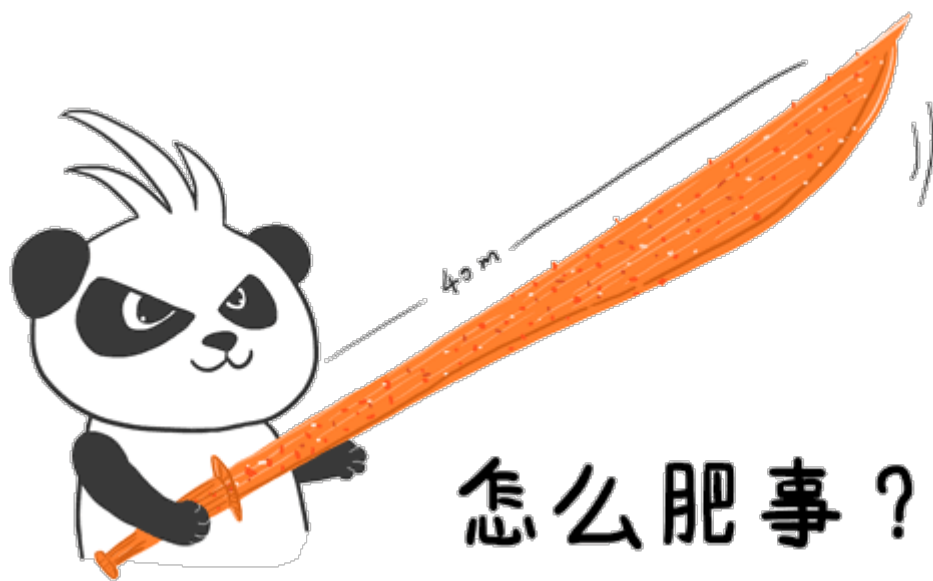


表 3 10 篇最长等待期文章

标题	作者	天(年)数	发表刊期
不熟练中-日双语者同形词和非同形词的隐蔽翻译启动效应	王悦等	1561 (4.28)	2014, 46(6)
我会在谁面前舍弃利益? ——博弈对象的能力与社会距离对名利博弈倾向的影响	谈晨皓等	1278 (3.50)	2017, 49(9)
创造能力不同学生的分类加工过程差异的眼动特点	沃建中等	1197 (3.28)	2010, 42(2)
对中学生的人格建构与学校适应的相关分析	陈会昌等	1131 (3.10)	2007, 39(1)
地震经历对插入空白间隔的汉字知觉干扰效应的影响	杨昕岳等	1121 (3.07)	2013, 45(5)
中国人三重自我建构加工中的相对优先性: 来自 ERP 的证据	王沛等	1111 (3.04)	2017, 49(8)
歧视知觉对城市流动儿童幸福感的影响: 中介机制及归属需要的调节作用	刘霞等	1073 (2.94)	2013, 45(5)
利/害条件下归纳推理的心理效应分离	蒋柯等	1055 (2.89)	2010, 42(11)
多题多做测验模型及其应用	丁树良等	976 (2.67)	2007, 39(4)
项目特征曲线等值的抽样误差	罗照盛等	962 (2.64)	2007, 39(4)

我只想问列位，摸着自己的良心问问自己，当时有没有对审稿人产生过如下心理活动：



给小编：插入动图 1 替代静态图

据说，现在不搞点 gay 里 gay 气的图，都赶不上时代了。看来我不给你们露一手，大家还真以为我这个 R 语言教师是吃素的（郑重声明：我在饮食上真的接近吃素，但在学术上完全是杂食动物）：

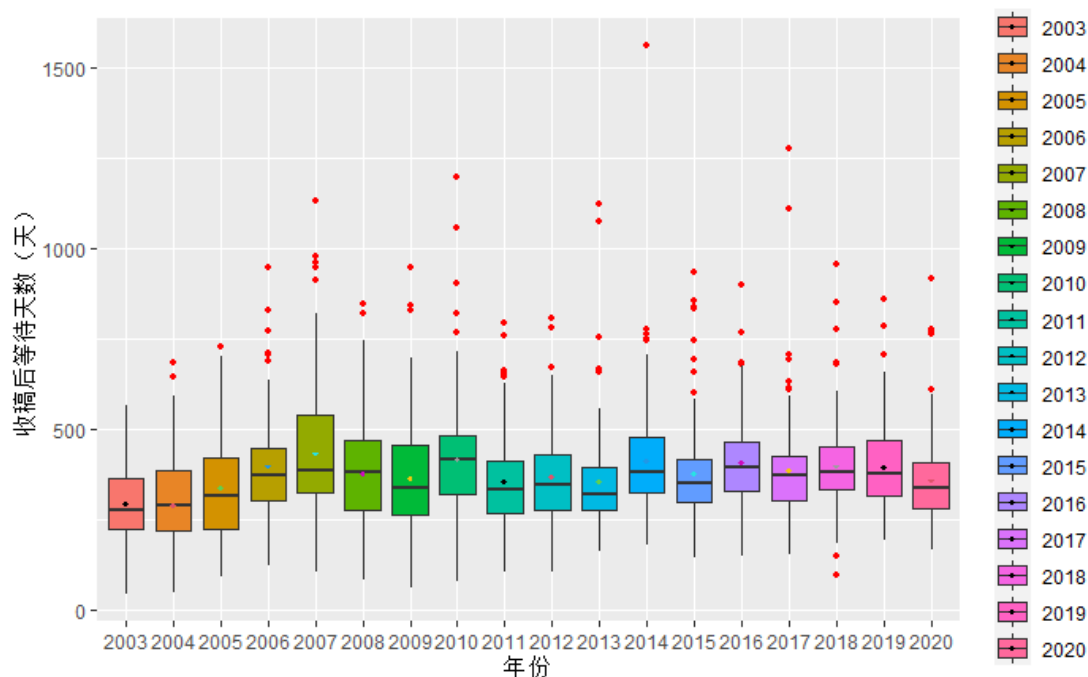
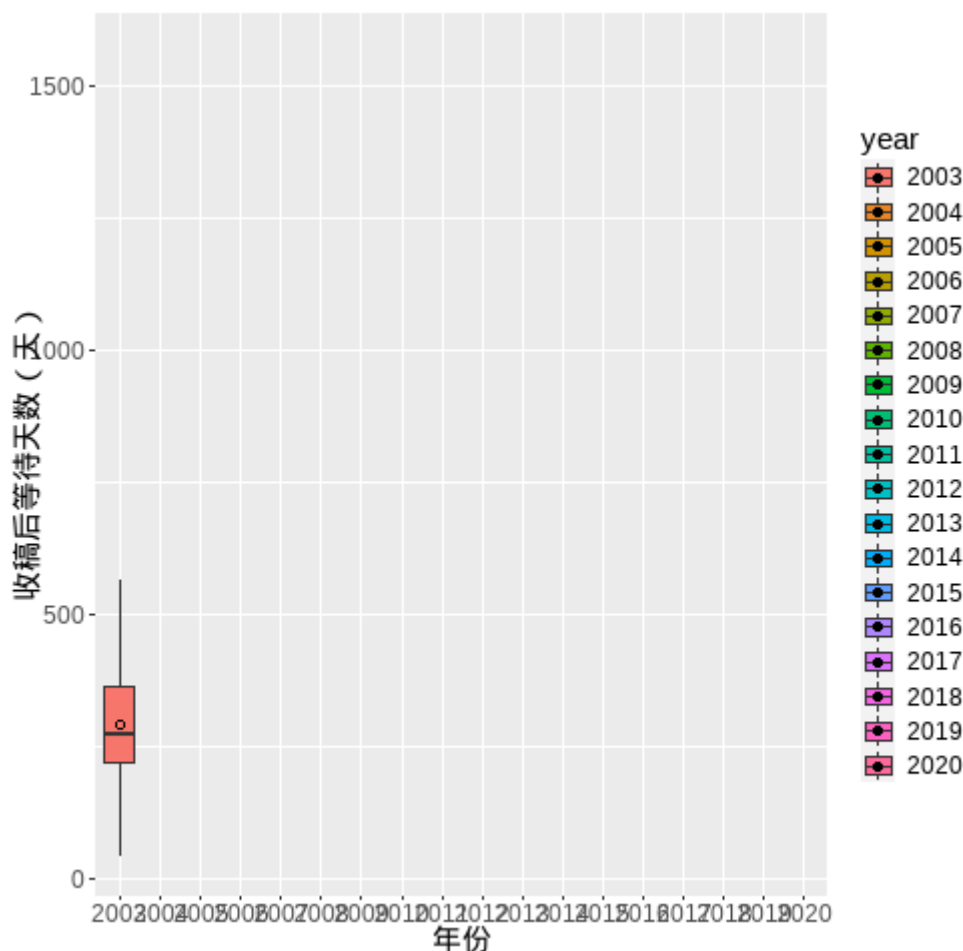


图 1：《心理学报》的发表周期（2003~2020 年）

注：箱线图（boxplot）的箱（box）中的点表示年发表周期均值；红点表示离群值（outliers）。

再据说，得动图者得天下。那就让它动一动：



给小编：插入动图 2 替代静态图

相关文件及图片见 [GitHub 链接文件](#)。

4 讨论

写论文最怕写讨论，投论文最怕被慢刀子一拖二骂三毙掉。子曾经曰过：稿固有一死，但死要死得痛快，还要死得明白。投稿与审稿皆是一种修行，投稿需耐心，审稿凭良心。然写与发、投与审之间总存在一种近似永恒的短期矛盾，这或可视为一种科研界的特有文化。此之谓“文化”者，超越应然之境而实然之存在也。借用弗洛伊德之语，本我与超我间的矛盾，或只能通过“一仆二主”之自我的勉力协调、做好批评与自我批评，从而假装维持一种充满正能量的学术打工人姿态。

当然，像《学报》（包括《进展》等）公开一些脱敏数据的做法，或有助于投稿者更为客观地认识到投稿周期固有的漫长待审状态，从而培育合理的“投稿后等待”心态。从预期管理入手做好自我心理建设，可能是每一个投稿人可以尝试的自我修养。反过来，对审稿人而言，在力所能及、保证质量的前提下尽早审回稿件，可能属于一种可以追求的学术美德。

此外，鉴于爬虫过程中若干棘手验证问题，建议《学报》（和《进展》）进一步优化网站设计、署名规范等内容。包括：更正存疑链接；修正缺失日期；修正部分标题放在<h3>标签下的<p>段落中造成爬虫无效或者数据结构混乱的情形（见下方示例，爬虫只需要几分钟，

确定这些 bug 用了好几天的带娃的时间);《审稿专家名录》及其他由编辑部统一发表文章的作者署名,可统一为“编辑部”,而不是“bjb”“其他”等混用。相关存疑篇目见 GitHub 附录文件 xb_problematic_links.xlsx,编辑部可对照修正。

示例:

与多数文章页不同的标题放置形式(网页源代码):

```
▼<h3 class="abs-tit">
... <p>进化心理学: 心理科学的未来发展</p> == $0
```

多数文章页的标题放置形式:

```
<!--中文-->
... <h3 class="abs-tit"> 人类大脑、认知与行为进化的整合模型 </h3> == $0
```

注意多数文章页的<h3 class = "abs-tit">后没再下设<p>标签放入标题文字。宜统一修正。

5 结论

问世间愁为何物,试问投稿人:知否知否,应是拒绝接收。

问世间情为何物,转问审稿人:知否知否,应是屡败屡投。

小编有苦:世间安得双全法,不负如来不负卿。

苍生共祝:努力实现投稿自由与发表自由的有机统一。

以上结论,基于臆测;如有雷同,纯属投射。

致谢

感谢南开大学社会心理学系所有选修本人 R 语言课程的历界本硕同学。

参考文献

仅附“‘审稿等待期’情歌”一曲:

余路还要走多久/你攥着我的手/让我感到为难的/是挣扎的自由
—— 赵雷:《成都》

后记

刚写完此文，便兴冲冲欲将结果与诸研究生分享，勉励其进行心理建设以树立良好的科研观、论文观、发表观。于是就有了如下对白：

我（兴奋）：大家仔细看看我的研究结果，要理解我的良苦用心。投稿后的等待很漫长，被毙掉也很正常，总之要用颗平常心。

学生（摆手）：“老师 duck 不必如此苦心，我们的心理健康得很！这种凡尔赛流派的心理咨询，不适合本土国情。”

我（惊讶）：“敢问有何高招，可破投稿后焦虑与被毙后创伤综合症？”

学生（笑靥如花）：“你看啊，我们就从没给《学报》投过稿；以前没有，现在和可预见的将来应当都不会有……”

果然，聊到这里，师生间的高傲都已经尽数体现了。

作者简介

吕小康，南开大学周恩来政府管理学院社会心理学系教师，南开大学计算社会科学实验室成员，经常自我赞助进行各类无脑研究（请仔细断句，不要与“脑计划”“类脑”等划上等号或不等号）。

附录

相关数据及原始文档见 GitHub 链接：

<https://github.com/xkdog/xb>

部分代码如下：

```
# 载入相关包；如未安装请先安装
```

```
library(tidyverse)
library(lubridate)
library(readxl)
library(openxlsx)
library(ggplot2)
library(gganimate)
```

读入数据并做初步整理

```
xb <- read_xlsx("xb_03to20_all.xlsx") %>%  
  mutate(sub = ymd(sub),  
         pub = ymd(pub),  
         pub_days = pub - sub,  
         pub_years = round(time_length(pub_days, unit = "year"), 2),  
         year = as.factor(year)  
  )
```

最快与最慢发表记录

```
xb %>%  
  arrange(desc(pub_days)) %>%  
  head(10) %>%  
  write_xlsx("toitse10.xlsx")  
xb %>%  
  arrange(pub_days) %>%  
  head(10) %>%  
  write_xlsx("rabit10.xlsx")
```

```
xb %>%  
  arrange(desc(pub_days)) %>%  
  head(10)
```

```
xb %>%  
  arrange(pub_days) %>%  
  head(10)
```

计算年发表均值与标准差

```
xb %>%  
  group_by(year) %>%  
  summarise(mean_pub_days = round(mean(pub_days)),  
            sd_pub_days = round(sd(pub_days)),  
            total_papers = n()) %>%  
  write_xlsx("summary_pub_days_03to20.xlsx")
```

绘图

```
ggplot(data = xb,  
       mapping = aes(x = year, y = pub_days, fill = year)  
       ) +  
  geom_boxplot(outlier.color = "red", outlier.size = 1) +  
  stat_summary(
```

```

    fun = mean,
    geom = "point",
    shape = c(1:18),
    size = 2,
    colour = factor(1:18)
) +
xlab("年份") + ylab("收稿后等待天数 (天)") +
transition_states(
  year,
  transition_length = 2,
  state_length = 2
) +
enter_fade() +
exit_shrink() +
ease_aes('sine-in-out')

```