

FAIL FAST
LEARN SMART

大胆假设，小心求证，试验创新

A/B测试

创新始于试验

王晔 等著

李丰 | 宋星 | 范冰 | 曲卉 | 蒋涛

熊长青 | 欧阳辰 | 孙天澍 | 周喆吾 | 曹成明

联袂推荐



机械工业出版社
CHINA MACHINE PRESS

FAIL FAST
LEARN SMART

大胆假设，小心求证，试验创新

A/B测试

创新始于试验

王晔 等著



机械工业出版社
CHINA MACHINE PRESS

我们如今看到的各种互联网产品的改版以及创新，其背后都离不开试验工作，其中被互联网、移动互联网行业广泛应用的试验方法是A/B测试。A/B测试的本质是试验，作为一种新兴的网页优化方法，通过对比试验增加转化率和注册率，带来产品和运营的创新，以实现各项指标的增长，如今它更是增长黑客所必备的核心思维方式和工作方法。在流量日渐昂贵、精细化运营的大环境下，互联网产品、运营和市场营销人员需要具有A/B测试思维，以测试结果为依据做决策，优化页面，提高转化率，迭代产品。

本书作者因在谷歌总部工作时发现了A/B测试的巨大作用而回国创业，并将自己对试验的理解和多年的经验进行梳理，在书中介绍了试验的本质、A/B测试的概念和价值、A/B测试的方法论和实战步骤，以及电商、金融、教育、旅游、快消品、UGC、PGC、媒体网站、SaaS等行业的应用案例，着陆页、App、网站、服务器端和推荐算法等不同场景的优化案例，还有针对决策者、产品经理、互联网运营、市场营销和工程师的案例。

图书在版编目（CIP）数据

A/B 测试：创新始于试验 / 王晔等著. —北京：机械工业出版社，
2019.2
ISBN 978-7-111-61776-1

I. ①A… II. ①王… III. ①软件—测试 IV. ①TP311.5

中国版本图书馆 CIP 数据核字（2018）第 298834 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

策划编辑：刘 洁 责任编辑：刘 洁

责任校对：李 伟 责任印制：张 博

北京东方宝隆印刷有限公司印刷

2019 年 3 月第 1 版第 1 次印刷

145mm×210mm·7 印张·3 插图·182 千字

标准书号：ISBN 978-7-111-61776-1

定价：69.90 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务

网络服务

服务咨询热线：010-88361066

机工官网：www.cmpbook.com

读者购书热线：010-68326294

机工官博：weibo.com/cmp1952

金 书 网：www.golden-book.com

封面无防伪标均为盗版

教育服务网：www.cmpedu.com

作者简介



王晔
(Sando Wang)

吆喝科技创始人

清华大学电子工程系硕士、耶鲁大学计算机科学博士，中国互联网试验创新领域布道师，“中国创新创业大赛”与“千人计划创业大赛”双料冠军获得者。先后就职于NEC、微软、谷歌等国际知名公司，任职期间负责P2P流媒体优化、P4P/ALTO、Google Adwords质量与移动广告新产品等前沿项目的设计、开发及优化工作，积累了大量试验创新的理论和实战经验。历任IEEE移动计算、GLOBECOM互联网协议与车辆技术评审专家，ACM外部审核专家，2009年对等网络国际研讨会专家组成员，超现代化的通信和控制系统2010、2011年国际会议TPC专家组成员。

王晔于2014年创立吆喝科技公司，率先将硅谷先进的 A/B 测试理论、方法与技术引入中国，与国内外优秀企业一同探索中国特色的试验创新之路。目前，吆喝科技公司已经帮助数千家企业用户成功运行数万组试验，并荣获融资中国2017-2018年度人工智能及大数据产业“最具投资价值公司”称号、2018年盛景全球创新大奖TOP20等奖项。



扫码联系作者获取模板



扫码关注吆喝科技获取免费试用

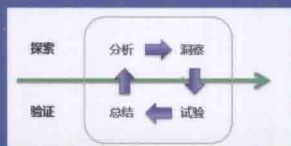
A/B测试是试验文化的核心理论，试验文化是增长黑客理论的本质特征。中国互联网已经进入为增长焦虑的下半场，只有试验驱动才能引领我们迎来下一次创新与繁荣。

虽然试验不是什么新鲜事物，A/B测试从字面上也容易理解，但说起来容易做起来难。如果理解有误或是操作不当，就会得到错误的结论，或者起不到期望的效果。那么如何正确地理解和运用下图所示的A/B测试方法论呢？本书将为您一一解答。



A/B测试知识点：

- ✓ 试验的基础知识
- ✓ 试验的战略制定
- ✓ 试验的战术执行
- ✓ 探索、验证闭环
- ✓ 试验假设的产生
- ✓ 试验优先级排序
- ✓ 试验的运行管理
- ✓ 试验的解读归档



为中华崛起传播智慧

封面无防伪标均为盗版

策划编辑◎刘洁

出版咨询◎florences30@163.com

封面设计◎MX DESIGN STUDIO

Q:1765629429



(扫码联系编辑)

赞 誉

广告行业有句老话：“总有一半的广告费是浪费了的，但是不知道是哪一半。”这句话在广告行业一直正确，直到互联网和移动互联网让广告和行为在线化和数据化，这句经典语录终于开始有了不同的可能性。同时，经过接近八年的快速发展，移动互联网用户也从以增量用户驱动进入对存量用户运营。以上这两个趋势和变化，都指向了一个方向：我们可以更加了解用户到底是如何选择、如何决策的，而这个问题的背后就是本书中所深入浅出解释的“A/B 测试”。

——李丰，峰瑞资本创始人

随着数字营销的投入越来越大，我们对数字营销的转化效果也越来越重视。本书介绍的 A/B 测试方法，还有 A/B 测试在营销创意优化的实践经验，给我们很多启发。推荐本书给关心市场营销的朋友阅读，可以帮助大家在日后的工作中持续地提升营销效果。

——宋星，纷析咨询创始人，互联网营销技术与数据专家

过去几年，我们总谈人口和流量红利，今后几年，我觉得应该是数据和方法论红利。而增长黑客和 A/B 测试就是企业急需掌握的两个数据驱动增长的方法，能够先人一步体系化运用的公司，会在下一轮竞争中拥有明显的优势。王晔老师的这本书理论扎实、案例丰富，是不可多得的学习 A/B 测试的教科书。

——曲卉，前增长黑客网（growthhackers.com）负责人，

《硅谷增长黑客实战笔记》作者

伴随人口红利的衰减和移动浪潮的消退，高歌猛进的欢喜时代悄然落幕，人们对挖掘增量和对盘整存量的做法莫衷一是，有人认为要追逐价值，有人索性附庸于流量。“物竞天择、胜者为王”，硅谷的成熟科技公司非常善于利用 A/B 测试等试验手段来精确演算自身商业发展的最佳路径。王晔博士的这本书将为你带来超前沿的一手洞见。

——范冰，《增长黑客》作者，增长官研究院创始人

互联网公司最大的特色是快速迭代和数据驱动，而这背后的关键技术就是 A/B 测试。小到颜色调整、按钮摆放，大到文案设计、产品逻辑，都可以用 A/B 测试来驱动产品的创新优化。可以说哪个团队真正用好、用足了 A/B 测试，哪个团队才成为真正数据驱动的互联网团队。可惜国内用好的团队很少，大公司有自己的产品，王晔从 Google（谷歌）回国创建吆喝科技公司，瞄准这一方向，三年来发展了不少用户，可是发现用好的团队却不多。这也是他撰写这本书的动力，让更多的开发团队更深入地理解 A/B Testing 的重要性、方法以及成果。

“微小的差异决定了谁将生存，谁将毁灭”，细节的累积最终决定成败，决定细节需要数据驱动，数据驱动决策要靠 A/B 测试，希望开发者、产品经理充分使用 A/B 测试，我们的互联网产品就会越来越好。

——蒋涛，CSDN 创始人

数字化时代生存必备的技能之一，就是拥有试验思维并且精于开展针对性的测试。大到商业模式的验证，小到一个广告创意的优化，概莫能外。王晔在 A/B 测试 SaaS 领域创业的同时，还能将自己的丰富经验和实战案例编撰成上乘之作，让人敬佩！

——熊长青，亦合资本创始合伙人，九枝兰创始人

有王晔“老司机”带路，A/B 测试之道将充满旖旎风光。

本书将 A/B 测试的来龙去脉、理论实践和经验教训等都介绍得非常清楚。其中丰富的案例令人印象深刻，很多案例都有很好的参考意义，另外，书中介绍的 A/B 测试的基础理论也是非常有价值的，它帮助人们设计更科学的试验方案。

“Fail Early, Fail Often, Fail Forward”（早试错，常试错，敢试错）是硅谷很多科技公司的创新理念。在很多创新的领域中，低成本和低风险的试错是持续获得更大成功的重要基础。试错就需要一个可靠的和科学的试验平台来支持 A/B 测试。我们从 A/B 测试获得的不仅仅是更优的手段，也有数据驱动的精益运营的理念。

数据时代，企业都在不遗余力地收集数据，“小步快跑”地激活数据，探问不止地驱动数据，A/B 测试作为一种重要的试验方法，对于企业提升产品或服务的用户体验有重大的意义。

如果我们来做一次测试，为了优选出市场上最好的 A/B 测试参考书，我相信，无论此书在对照组还是在试验组，都能够获得胜利。

——欧阳辰，品友互动 CTO（首席技术官）

互联网时代，环境快速变化，企业必须采用全新的数据驱动的决策体系和创新模式。试验迭代（A/B 测试）是其中最为核心的一步，也是 Facebook（脸书）、Google、今日头条等知名企业不断创新和增长的秘密和源泉。王晔博士结合自己在硅谷知名企业的见闻和体会，以及与中国数百家企业合作的实践经验，总结出一套企业试验的思维、方法、工具和实战案例，值得所有企业管理者和一线员工好好借鉴和应用。

——孙天澍，USC（美国南加利福尼亚大学）马歇尔商学院教授

王晔深耕增长行业多年，从耶鲁到谷歌，把增长的方法论和工具带到中国互联网，是 A/B 测试在中国很早、很卓有成效的布道者。《A/B 测试》这本书是互联网江湖的“九阴真经”，王晔将各个细分领域的优秀案例深刻剖析，读来宛如按图索骥，不仅为读者描绘了增长天堂的模样，还搭好了云梯，字字珠玑！

——周喆吾，MetaApp 联合创始人

产品经理的日常之一就是“吵架”，每次跟老大过方案或者跟需求方过方案都会吵架，谁都说服不了谁，因为所争执的点通常都是“假设”或者主观的东西。每次吵完都是内伤，虽说“对事不对人”，大家要理解吵架的行为，都是为了把产品做好，对用户负责，对公司负责……如果想解决吵架问题或者降低吵架的概率，大家能开开心心地把事做好，我觉得很好的办法就是做 A/B 测试，通过 A/B 测试，记录下用户的使用情况再进行分析，就能知道哪个方案更靠谱。

——曹成明，起点学院、人人都是产品经理网站创始人兼 CEO

这是一本重量级的 A/B 测试参考书，可以整体提高中国互联网从业者数据驱动的认知和能力。企业具备数据驱动的文化、流程和工具是非常关键的，A/B 测试就是重要的工具之一。希望本书能够帮助推动各行业通过数据挖掘和数据产品的广泛应用，找到健康增长的飞跃曲线，让每个互联网团队都体会到数据科学和试验迭代的巨大价值。

——顾青，DTALK 创办人，E-Bizcamp CEO

“优化”是运营一款产品的上好策略。接手一个业务问题，运营人的

第一思维是“分析数据、找出原因、提出解决方案、上线执行”四步完成，而往往忽略了关键的一环：解决方案的质量。解决方案一旦失败，就意味着浪费运营资源。而试验的价值，就是把这种浪费打消于无形，掌控成功。我曾在博文中说过一句话：“神一样的运营经理，首先是一个把脉高手，其次是个试验狂人，二者缺一不可。”这本书能帮你变身运营高手，是一本实战指南式的好书。

——韩利，《运营实战指南》作者

A/B Testing（A/B 测试）这个词儿很时髦，做互联网产品的人差不多快挂在嘴边了。然而，行业内的人其实多少都知道，国内真正把这个理论落到实处的公司凤毛麟角。这在移动互联网趋势和流量红利爆发的上半场似乎没问题，甚至天经地义，因为那时候跑马圈地是第一要务。然而在移动互联网进入深耕细作的下半场的今天，这个问题就越来越凸显了，如何针对存量用户进行业务的优化？国内互联网巨头和硅谷知名公司在技术驱动的理念和基础设施上的差距已经被越来越多的人认识到了。

这些年看着 Sando 和 吆喝科技的 A/B Testing 产品一路迭代，越来越成熟，也越来越贴合国内用户的使用场景，我深感欣慰。所以每当有人说起 A/B Testing 解决方案时，我也都会毫不犹豫地第一时间就推荐吆喝科技。感谢 Sando 的这本书，帮助我更深刻地理解了 A/B Testing，也让我知道了为什么吆喝科技在这个领域内能独树一帜。

——彭圣才，“菜园子”产品经理社区创始人

百度于 2010 年年初引入 A/B 测试，用于评估网页搜索结果 UI（用户界面交互）的变更，对比不同商业策略对于业务指标的影响；腾讯则

更早在频道首页的改版中采用类似的评价机制。今天，A/B 测试在互联网产品的开发中被越来越多地采用，发挥着无可取代的作用。本书由业内人士撰写，为读者呈现这一领域的全貌。

——王毅轩，百度技术专家

在商业竞争如此激烈的今天，任何成功都需要每个点能够最大化。这种最大化不是想出来的，需要不断假设、验证和总结，这其中 A/B 测试无疑是一种宝贵的思维方式。

——许静芳，搜狗搜索总经理

中国移动 10086 和其微信、微博、App 等互联网服务渠道覆盖 7 亿用户，为实现持续增长和高效转化，我们设立 15 人的数据增长团队，将自动化 A/B 测试应用于图文、海报、菜单、着陆页等关键场景的设计工作中。经过持续运营，商品下单率提升 484.8%，活动转化率提升 112.7%，资讯阅读率提升 185.1%。实践证明 A/B 测试是互联网团队必备的能力，在精细化运营的互联网下半场，本书给了我们更多指导和方向。

——陈庆，中移在线互联网运营中心副总经理

王晔兄将自己多年从事与 A/B 测试相关工作的方法论融合到书中，结合理论、案例实践和工具使用等多个方面介绍了 A/B 测试在互联网应用的内涵与外延，是国内目前介绍 A/B 测试非常难得的一部著作。

——樊聪，前美团点评 A/B 测试系统负责人，现旷视科技技术总监

前言

试验不息 创新不止

1. 从一个故事说起

2013 年的一天早上，我在谷歌（Google）总部广告质量部门的办公室里盯着显示器上的图表，站在我身旁的是穿着正式的广告产品经理。“以现在的试验数据来看，我们申请把流量推上 20% 吧？”“先看看美国地区的细分结果。”类似这样的严肃讨论几乎天天都在发生，在我的办公室内发生，也几乎在每一个其他小团队的办公室里发生。

你可能想问，这样的对话到底是在讨论什么？当时我们团队正在用 A/B 测试的方法来进行一个线上试验。我们需要每天（有时候每小时）观察试验数据，做出项目的下一步决策，如果这个决策的影响可能会比较重大，例如将影响谷歌 20% 的用户，那么我们需要向上级领导汇报以得到其支持。

我们做的试验源于一个大胆的想法，这个想法来自搜索部门的产品经理：搜索产品时试着把谷歌搜索结果（包括广告）里的 URL 换成结构化的域名，例如把“<http://www.appadhoc.com/lpo>”换成“AppAdhoc.com > LPO”，这么做会不会提升用户浏览搜索结果的效率从而提升用户的广告点击率呢？这样的事情从来没有人做过，包括谷歌的竞争对手也没有做过。这样的改动用户会喜爱吗？这样做会不会让用户更容易找到想要的搜索结果？有可能，但也不一定，答案似乎见仁见智。没有人能预测这个项目会不会有收益，更没法预测具体会给广告营收带来 1% 的影响还是 10%、0.1%，或者没有任何影响。

如果按照“传统”的企业决策流程，产品经理将围绕这个想法做一些用户调研，然后将项目汇报给领导，如果领导认为这个想法值得一试（很多情况下领导会劝产品经理放弃），会组织会议进行讨论，因为会议上有人喜欢这个想法，而有人不喜欢，最终的会议决议可能会放弃这个想法。

谷歌的做法不一样，产品经理将想法口述给领导，领导同意运行一个“2% 流量”的试验。技术团队花了几天时间完成了研发和测试，试验很快就上线了。2% 的搜索流量被采样进入试验，其中 1% 的用户作为对照组会看到 URL（如 <http://www.appadhoc.com/lpo>），另外 1% 的用户作为试验组会看到域名（如 AppAdhoc.com > LPO）。这两组用户的广告点击率被准确地采样统计，然后对比分析，得出实时的试验结果。试验结果不是很显著，但是似乎试验版本的样本均值略好一点（如 +1% [-2%, +4%]），也就是说从采样样本来看试验组的广告平均点击率高于对照组，但是从统计意义来看还不能确定两者谁更好。通常情况下，试验结果不显著是因为样本量不足，所以我们希望做更多的分析，然后向领导建议将试验推送给更多的用户（20%），获得更多的试验样本，以期待有可能得到更明确的试验结果。

这样的试验项目在谷歌很常见。具体来说，谷歌每个月都会运行 1 000 个以上的试验项目。每个试验项目的参与者都来自这个大公司的各个部门：产品经理、工程师、销售人员、客服人员、法务人员、质检人员、策略研究者、市场人员等。项目的负责人通常是产品经理，他会管理工作进度及协调公司资源。项目的其他参与者会向各自的部门领导汇报，保证项目决策得到相关部门的支持。

从谷歌的实践来看，这种跨部门组建的试验项目小团队很有战斗力，谷歌内部几乎所有成功的项目都是这么落地的。我觉得这种成功来自于

试验项目的管理得当、目标明确、路线清晰，以及特别重要的——我们可以利用强大的 A/B 测试。

2. A/B 测试带来了很多好处

像“URL 改成域名”这样的项目几乎都是通过 A/B 测试试验系统来实施的。由于这个原因，在谷歌内部，“项目”这个词几乎已经被“试验”所完全替代。广泛使用 A/B 测试为谷歌带来了如下长久的巨大的好处：

- ✓ **确定可预测的业务提升：**每个试验项目的收益在完全上线之前就可以精确衡量，甚至精确到 0.01%（注意，考虑到谷歌的业务体量，营收增加 0.01% 相当于每年多赚或者少赚数百万美元）。如果一个试验项目会带来负增长，这个项目很可能就不会上线；只有带来正增长的试验项目，才会加大投入并最终推广给全量用户。通过 A/B 测试精确预测每个项目的商业回报，然后有选择地上线项目，谷歌可以确保每年广告收入增长约 20%（约 100 亿美元规模），从而保证利润和股价持续十多年的攀升。
- ✓ **低风险、高效率的试错：**试验项目必须经过小流量的灰度发布阶段（如 1% 的流量），只有确定达到了业务预期，并且没有故障，没有过负载，没有用户投诉，没有违背政策监管，没有其他风险，才会推广给更多用户。这样大幅度降低了决策风险，把可能的损失降到最低。同时，A/B 测试排除了试验之间的互相干扰，小流量试验可以大量并行进行，大幅度提高了试错效率，把互联网迭代优化的速度推到了极致。因此，谷歌常常有上千个试验并行运行。
- ✓ **创新的企业文化：**谷歌是一个巨型公司，在全球各地有好几万名优秀的员工，公司有复杂的组织架构。这样的公司很容易滋生“大公司病”，每一个决策都可能因为受影响的部门太多而遭遇重重阻碍。

A/B 测试的低风险、高效率，以及过往的成功实践，持续鼓励公司的新老员工开拓思路并大胆创新，避免了故步自封的“大公司病”问题。创新的企业文化，是企业长久生命力的源泉。小团队创新的工作方法，使阿米巴企业管理方法可以成功落地。

当然，有些项目并不能通过 A/B 测试来做，如谷歌的电视棒项目（Chromecast）是一个典型的从 0 到 1 的创新项目。对于这种没有用户基础的新产品，我们虽然不能用精确、方便的 A/B 测试做试验，但是可以用其他的试验方法，如最小化可行产品（MVP）的市场测试方法。

3. 如何复制这个成功

试验，尤其是 A/B 测试的价值很高。在硅谷巨头公司、华尔街和各种创新企业，试验都取得了巨大的成功，但是在目前的中国市场，试验还不太成熟。在很多行业里，美国企业的试验能力和试验效果都比我们领先很多。在巨头公司的对比中，谷歌每年的试验数量是携程的 10 倍；在 A/B 测试服务商的对比中，Optimizely 上的试验数量比吆喝科技（App Adhoc）上的试验数量要多 10 倍。

其实无论是在中国、美国还是其他市场，还有很多企业没能建立起试验创新的文化。要实现试验驱动的增长，需要正确地在企业内建立 A/B 测试的文化，搭建完善的基础设施，采取正确的工作方法。我们通过多年的工作和观察，在尝试走试验驱动路线的企业内，发现了一些常见的问题，将其大致可以分为以下三类：

（1）决策者缺乏试验的思想，项目决策谨慎但是实施坚决。这种“传统”方法经常遇到的问题是产品团队投入 3 个月用于研发产品的大改版，最后没有业务上的回报，甚至用户反馈新版还不如以前。正确的做

法应该是反其道而行之，大胆假设，小心求证。不做试验的企业虽然还为数不少，但是已经在大量转变。在激烈的市场竞争下，企业越来越以结果为导向，对领导者和业务骨干的要求也越来越高，大企业的各个层级的领导们也越来越重视试错和迭代。

（2）决策者有试验的思想，但是缺少 A/B 测试的正确实施方法。虽然领导希望做 A/B 测试来验证决策，但是实施起来需要投入大量人力和时间成本，容易出错，试验设计、试验配置、试验结论也常常饱受争议，并没有提高企业效率，甚至事倍功半，形式大于内容。这样的组织可能处在学习互联网思维的实践阶段，随着互联网产业的蓬勃发展和影响力加强，它们正在快速改进，走上正轨。

（3）已经在正确地运用 A/B 测试，但是效率低、频率低。很多业务线一年尝试的试验不到 10 个，这样的试错速度并没有比传统方法提高多少，只是保证了错误的决策不上线，并没有真正利用好互联网的强大力量。对这些组织来说，目标应该是将试验数量提高 10 倍，方法是人才培养和文化建设，以及采用更好的 A/B 测试基础设施，鼓励高频、高效的创新项目，增加单位员工的试验产出。

4. 为什么写这本书

意识到这些企业普遍面临的问题后，我觉得 A/B 测试的系统性知识和经验会对各行各业的业务负责人很有帮助，特别是对互联网线上业务的从业者来说，A/B 测试是必备能力，这促成了本书的写作。

A/B 测试在各行各业的很多场景中都已经成熟的应用和不可替代的价值，特别是在互联网行业、科学研究、基础农业、医疗、金融、公共政策、市场营销等领域都非常成功。在互联网行业，一个特别热门的 A/B

测试应用场景就是产品运营的增长黑客。增长黑客借助互联网和 A/B 测试的力量让业务增长的速度远超传统行业，这套方法容易落地，投入产出比高，无数成功的独角兽互联网企业就是利用增长黑客创新创业，并创造了财富的。本书的内容将重点放在互联网产品运营特别是增长黑客的 A/B 测试上。

本书的定位是 A/B 测试的工具书。我们从实践角度出发，介绍了 A/B 测试的理论原理，标杆企业的最佳实践，进而详细介绍 A/B 测试的实际应用场景、落地实施流程和业务产出预期。无论你是决策者还是业务骨干，是产品负责人还是软件工程师，是内容运营人员还是市场营销经理，希望本书的思想和内容都可以帮助你。

在本书的写作中，很多行业专家，特别是吆喝科技的专家们为我提供了丰富的案例素材和专业建议。如果没有他们，这本书无法完成。在此，我要特别感谢李森、沈国阳、张毅飞、柏利锋、刘飞、李想、蒋守战、李翔宇、陈聪、刘泽军等。

王晔 (Sando Wang)

2019 年 1 月