

白皮书

facebook

利用增效衡量 促进业务发展

2018 年 1 月



目录

- 3 简介
- 4 增效衡量的重要意义
- 4 论文：遵守四步流程
- 5 利用更出色的分析应对业务挑战
- 5 成效衡量委员会促进深度交流
- 6 领先营销者的观点
- 10 Facebook 研究证实增效衡量的作用
- 11 新手入门四步指南
- 11 给所有营销者的实用建议
- 12 利用现代化衡量工具促进业务增长

编著者

Stephan McBride, Netflix
广告科学部总监

Matthew Gerrie, Booking.com
营销科学部总监

Tony Flanery-Rye, eBay
全球增长分析部高级总监

Alok Gupta, Airbnb
数据科学部总监

Maggie Burke, Facebook
成效衡量委员会主管

Jonathan Lewis, Facebook
成效衡量部产品主管

Jesse Goranson, Facebook
成效衡量部总监

Abbey Koslow, Facebook
营销科学部主管

Alex Esber, Facebook
成效衡量部产品营销主管

Brad Smallwood, Facebook
成效衡量和成效分析部副总裁

Chao Chen, Facebook
分析部数据科学家

简介

广告是帮助营销者取得成功的关键因素之一。全球消费者的生活都离不开广告。广告在为人们的生活增添色彩、活力和探索乐趣，有时也会令人感到恼怒和厌恶。随着消费者把关注的目光投向新平台，以及新数据和新技术的不断扩展，一种被视为科学的金科玉律的古老方法逐渐走到台前。这种方法就是随机对照测试，它是一种基础的科学方法，如今被一些广告从业者用做主要工具来研究如何获得竞争优势。新的研究结果也表明，随机对照测试可在 25%-50% 左右的情况下帮助营销者作出更好的决策。

如今的数字领域在不断细分，营销的难度也在日益增加。营销者不仅要更多的营销渠道中做出投资选择，而且还要确保针对各个渠道优化创意、受众，以及日益复杂的购买方式。本文将介绍我们为什么认为营销者应考虑转而采用随机对照测试或其他类似的增效衡量方案，同时以四个品牌为例，介绍这一策略调整带来的显著成效。

增效衡量的重要意义

在过去数十年里，诸如媒体组合模式和末次点击归因等成效衡量工具通常是众多营销者的决策参考。这些工具通常需要使用者具备分析技巧并建立相应的团队。然而，日新月异的广告技术和爆发式增长的数据量不仅改变了现有广告分析方法的适用性和解读方式，而且孕育了新的方法。优秀的营销者认识到，这不仅需要他们改变制定决策的计算方法，而且还需要营销机构改变所采用的方法和技术。顶尖的网络技术公司从最初便树立了“从测试中学习”的意识，所有关于平台功能的决策几乎都要经过 A/B 测试。实践证明这种方法对产品决策具有重要作用，它可以延伸为随机对照测试或其他增效衡量工具，帮助确定最合理的营销预算计划。

增效衡量是营销者判断广告投资是否合理的关键，它可以帮助对比了解花钱投放广告与不投放广告对业务量的影响。例如，节日季营销活动带来的销售额增长量能够为零售商提供充分的依据，去合理安排年末广告预算（只要销售额增长量超过不投放广告所能获得的销量）。

根据所用技术和平台而异，增效衡量方案可采取多种不同的形式，但无论形式如何，其目标始终不变，即通过对比看过广告的人群与没看过广告的人群之间的广告成效差异，计算出广告表现的提升量（也称为反事实分析法，两者的唯一差异是广告展示）。随机对照测试是衡量增量的黄金标准，在系统和数据允许的情况下，从小型营销测试到整个媒体渠道皆可进行评估。

可以说，增量是衡量投资回报的最佳指标。大多数高管都很重视将数据作为决策依据，并根据实际数据分析和衡量营销效果。“高管希望在各个方面都有数据作为依据，如果增效衡量发现某种渠道或组合更有效，则可以重新合理分配预算。” Netflix 广告科学部总监 Stephan McBride 说道，“这种动态调整为公司创造了良机，让他们能根据媒体渠道的实际效果合理分配广告投资。”

论文：遵守四步流程

科学和试验是获得市场优势和可靠效果的基础。选择合适的服务商和工具固然很重要，但赢得竞争优势和业务增长通常要靠公司快速通过测试进行学习和作出判断的能力。

增效衡量为 Airbnb、Booking.com、eBay 和 Netflix 等公司开启了新的商机，各种规模和各个领域的企业都可以发掘这些商机。采用严格的数据驱动型决策的企业将在多个层面上提升业务水平，以实现核心业务目标：增长。

企业需要遵守四步流程才能将试验和学习高效应用于营销中。



利用更出色的分析应对业务挑战

了解营销与业务成效之间的紧密联系是当今许多营销者面临的核心问题。得益于衡量方法和技术的进步，这些挑战显著减少了，但要想从中受益，企业必须致力于全面采用更精准的系统。营销目标就是业务目标，企业需要通过测试解决的问题可借助准确的衡量来解答。

衡量增效的能力是获得营销信心及众多决策参考指标的关键。许多企业都认识到，全面的营销成效衡量方法能够带来更好的决策和更高的价值。诸如旅游等行业最适合使用这种方法，其他行业的企业也都可以利用这种方法找到清晰的业务增长之路。

成效衡量委员会促进深度交流

成效衡量委员会于 2016 年 11 月由 Facebook 与合作伙伴共同创立，旨在深度交流数字广告领域的挑战和机遇。委员会汇聚了全球行业领袖、顶级品牌和成效衡量专家，目前正在努力研究如何帮助行业确定短期和长期的业务价值，并鼓励企业积极采用能够提升营销成效的衡量方法。

“互联网成效衡量系统是针对台式电脑主导的时代而设计的。” Facebook 成效衡量部总监 Jesse Goranson 说道，“许多公司都在努力将最初针对 Cookie 时代开发的系统升级为适合移动时代的系统。”

“互联网成效衡量系统是针对台式电脑主导的时代而设计的。许多公司都在努力将最初针对 Cookie 时代开发的系统升级为适合移动时代的系统。”

– JESSE GORANSON, FACEBOOK 成效衡量部总监

Goranson 补充道，“其主要转变之处是从只关注末次点击到全面关注客户购买流程。营销者想要了解和评估可能影响购买决策的所有触点，以及消费流程中不能带来任何增量价值的触点。”

经过全面研究和评估，成效衡量委员会断定，增效衡量不仅对各类企业适用且具有重要意义，而且还可供寻求这类系统的营销者广泛使用。据委员会介绍，实践已经证明常见的营销指标框架和标准效果并不如试验或“从测试中学习”的方式。

根本原因在于数据。然而在营销成效衡量方面，数据透明度及相关信息的精准度仍有尚待解决的问题。大型机构和包装消费品企业关注的是通过可扩展性来树立竞争优势，而成效衡量委员会中的部分企业最感兴趣的是直接评估投资回报。

标准化确实帮助一些企业扩大了规模，但行业标准却是这些委员会成员的次选。深入分析各个营销渠道，并透视大量行为数据或许对企业而言更有帮助。

领先营销者的观点

Facebook 合作伙伴认识到优秀的成效衡量系统已经问世且可供使用。许多行业参与者也纷纷致力于改善自己的成效衡量方法。人们通过长期投资来增强和完善系统，以捕捉精准的决策参考数据。

诸如 Airbnb、Booking.com、eBay 和 Netflix 等领先的营销者都明确了公司要采用严格的成效衡量方式，而这种思维影响着他们投放广告的方式。这四家公司都决定将成效衡量和基础设施放在首位，聚焦能够让客户通过严格的数据衡量方式得出准确结论的发行商。可靠的成效衡量是这些企业的核心，并且能通过各个业务层级进行强化。

虽然行业参与者必须继续寻找更好的方式来获取和验证数据，但增效衡量的核心基础架构已经建立并为许多企业带来效益。虽然 Airbnb、Booking.com、eBay 和 Netflix 在使用增效衡量的方式和认识其重要性上都有不同的经历，但最终都殊途同归。增效衡量为这些企业的内部策略带来了重大（通常是剧烈的）转变。

Airbnb 利用增效衡量来了解投资回报

Airbnb 是全球最大的社区房屋租赁公司，他们依靠效果类广告来实现业务增长目标。与大多数企业一样，他们用来实现业务目标的预算也是有限的，因此最大限度提升营销投资回报是重中之重，而成效衡量在这其中扮演着关键角色。

“成效衡量对两个方面至关重要。” Airbnb 数据科学部总监 Alok Gupta 说道，“第一个方面是了解获得的投资回报和正确计算增长来源。第二个方面是支持不同付费渠道的优化，以便决定如何使用未来的营销预算。”

Airbnb 为优化营销预算分配了大量人力财力。“公司的营销投资回报取决于优化，而优化则依据成效衡量。” Gupta 认为，它们是相辅相成，缺一不可的。“衡量出各渠道的边际效率后，我们就可以了解各渠道的历史效率，从而确定整体投资组合和优化策略。”

“成效衡量对两个方面至关重要。第一个方面是了解获得的投资回报和正确计算增长来源。第二个方面是支持不同付费渠道的优化，以便决定如何使用未来的营销预算。”

– ALOK GUPTA, AIRBNB 数据科学部总监

在预算讨论期间，成效衡量方法并非争论的主题，公司相信 Gupta 的团队会采用现有的最佳方式并非争论的主题，并在对某个渠道的成效衡量信心不足时提出报告。“营销团队拥有跨渠道或营销活动决定营销预算分配的自主权。” Gupta 说道，

“各业务部门的团队成员都知道我们在努力优化固定预算的投资回报，我们的方式是通过长期价值计算和增量来衡量。” Gupta 说道，“人人都知道增量这个概念及其重要性。”事实上，为了强化 Airbnb 对成效衡量和优化的承诺，其营销数据科学团队在过去 18 个月内扩大了 10 倍。

Airbnb 今年的大多数富有影响力的营销工作都源自于衡量增效的试验，以及评估营销渠道的边际效率。这些试验结果促使公司的投资发生了意义重大的转变。各个渠道的预算分配也发生了根本性的变化。Airbnb 的试验和转化提升测试在付费搜索方面带来了一些重要的突破，而付费搜索正是旅游行业激烈竞争的领域。通过利用增效衡量了解的数据，公司在付费展示和付费社交渠道方面获得了更高的回报。

“在理想状态下，A 和 B 两组花费不同但其他条件相同的广告之间的差异即为增量，也就是通过试验获得的额外收益。” Gupta 说道，“在现实中不可能拥有完全相同的条件，所以我们尽量在相同的系统中衡量差异。”

Booking.com 将增效视为数据驱动型试验

“Booking.com 从成立之初便是一家数据驱动型企业。” Booking.com 营销科学部总监 Matthew Gerrie 说道，“许多公司都称自己是数据驱动型企业，但只是口头上说说罢了。我从未见过像 Booking.com 这样的公司，几乎所有的决策或调整都必须经过 A/B 测试，或者有相应的数据作为支撑。”

从公司创立之初至今，每次调整网站前，他们都会首先测试不同的选项或设计方案的转化率。据 Gerrie 所言，当 Booking.com 开始营销投资时，也采用了相同的以数据为支撑的试验方式。

Gerrie 说道：“在 Booking.com，营销者有责任通过 A/B 测试（或其他增效测试）来证明某个方案对业务有增量贡献。”据 Gerrie 表示，这种思维已经在他们公司深入人心，在没有经过增效测试或其他公认的衡量方法的情况下，没有人会接受或确证某个结果。

“我们最重要的工作是确定营销漏斗各个层级应扮演的角色和应使用的渠道，以便尽可能投放最高效的广告。” Gerrie 说道，“我们不会在无法通过增效测试证明回报（包括即时效果或投放广告后不久产生的效果）的渠道中投入资金。”

与其他诞生于网络然后转向电视广告的公司相似，Booking.com 也会针对电视广告采用“从测试中学习”的方法。Gerrie 称，Booking.com 并没有建立一个单独的团队来专门负责电视广告开发，与各种创意和媒体机构保持传统合作关系，而是将营销科学与传统电视广告合为一体。如今的媒体决策要依托数据，因此没有理由割裂营销漏斗的上下两端，因为新媒体和传统媒体的分析员都可以利用自己的知识进行营销分析。

增效衡量如今已在 Booking.com 独具特色的广告效率衡量中扮演主要角色，但以前并非是这样。Booking.com 在前几年非常重视能够持续带来即时预订的广告渠道。“是否拥有持续的投资回报是衡量成功的唯一标准。” Gerrie 解释说。这种理念最终引起了争论，一些测试很快证明网络广告的效果很少能够通过这种老套的方式来衡量。

“现在人人都知道，广告展示或网络视频展示的一些增量价值并不会带来即时预订。这个道理并不复杂，但向企业展示增量价值是关键所在，它能引发思维的转变。” Gerrie 说道，

“我们会使用测试组和对照组的方式来衡量所有参数，根据研究的漏斗层级来寻找不同关键绩效指标 (KPI) 的增效情况。最终，这种方法会带来预订转化量的提升或增加。”

eBay 构建严格的成效衡量方法

“成效衡量是 eBay 的命脉。” 电商先锋 eBay 全球增长分析部高级总监 Tony Flanery-Rye 说道，“想想能够支持 12,600 名员工办公，并面向全球销售数百万种商品的基础设施，舍入和小数点都需要谨慎对待。”

“成效衡量是 eBay 的命脉”

- TONY FLANERY-RYE, EBAY 全球增长分析部高级总监

这就不难理解这家公司从上至下都采用严格的成效衡量。“我们不遗余力确保 A/B 测试不受其他因素的干扰。” Flanery-Rye 说道，公司会测试投放广告前后的等效状态，衡量用户看到广告后行为或操作的持续时间。“我们非常注重收集确切可靠的数据用于关键判断。”

Flanery-Rye 指出，追求效果最好的测试方式是 eBay 的宗旨，因为归因能最好地诠释公司的增效测试所发现的成果。

他说，尽管增效测试结果和按归因规则取得的成果之间永远存在差异，eBay 通过严格的测试消除了这些归因差异，帮助公司发现了不准确的数据和未经证实的假设，这些都与归因决策息息相关。

“复杂而强大的系统让我们开展一些重要分析，然后将这些分析结论付诸实践。”

Netflix 通过增效衡量支持业务决策

娱乐巨头 Netflix 的广告科学部总监 Stephan McBride 表示，增效衡量让公司得以了解营销工作对提升业务运营效率和实现主要增长目标的效果。“增效衡量这种成效衡量方法让我们能真正了解应该采取哪些措施来调整营销策略，进而改善业务效果。” 他说。

McBride 称，Netflix 一直在根据内部测试和学习来提升产品体验。这家公司极其重视试验，它的企业文化推崇通过随机和反事实对照组测试来确定特定的产品功能、设计方案或广告目标是否更有助于实现公司的核心目标：让现有会员满意，并吸引新会员加入。

McBride 指出，科学是 Netflix 遵守的中心原则，在整个公司上下，他们会例行对产品开发、推荐测试和内容投放算法的试验，以获得科学的结果和信息来帮助决策。Netflix 成功的产品开发周期广受赞誉，他们正在努力将这种成功拓展至广告成效的衡量方法。

“几年前，我们就希望能更深入地了解广告花费所带来的增量效果，并加以改善。” McBride 说道，“在我们的企业文化中，管理者渴望抓住能够显著提升业务成效的机会。”

Netflix 鼓励不同团队的员工定期提供反馈，并查看数据科学团队的研究记录和成果。“我们对衡量方法和数据的兴趣非常浓厚，我们有很多机会吸引和鼓励大家交流对衡量方法的观点。” McBride 说道。

“Netflix 的企业文化注重对假设进行评估，所以这里并没有什么禁忌。我们认同任何人的观点都可能是正确的或错误的，而我们的最终目标是证实或验证这些假设。”他说，“我们认为应该通过评估，确定这些假设是否确实能够帮助我们做得更好。有时候，我们需要数据来做出确认或评估。”

Facebook 研究证实增效衡量的作用

为帮助了解增效衡量相对于其他衡量方法对决策的影响，Facebook 评估了 2015 年 2 月至 2017 年 9 月期间，广告主通过增效衡量工具（转化提升调研）开展的 580 次具有统计学意义的 A/B 测试。¹

这些测试对比了不同的营销要素，例如创意策略或版位，以便了解哪些要素更有助于实现相应的业务目标。

他们对 A/B 测试结果的评估使用了两种不同的衡量系统：增效衡量和单日点击归因（将转化归因至转化事件发生后一天前用户点击的广告）。Facebook 在衡量的测试中发现，在回答哪些营销策略最有效的问题上，这两套成效衡量系统平均有 23% 的情况存在差异。此外，如果结果出现差异时广告主选择了错误的测试单元，则可能错过平均 64% 的单次转化费用 (CPA) 提升。

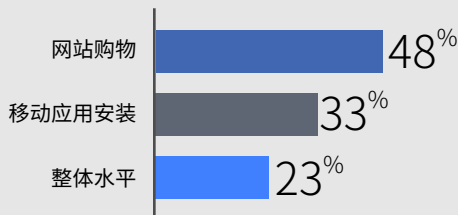
尤其是在网站购物量方面，两套系统的结果存在差异的情况达 48%，此情况下的 CPA 平均相差更达 79%。

“效果类营销广告主衡量成效的常用方式是基于点击的归因模型。” Facebook 产品试验部主管 Jonathan Lewis 说道，“我们会研究 A/B 测试中哪些营销策略会被单日点击归因模型选为效果最佳，将它们与增效衡量根据 CPA 选择的策略进行比较，了解测试的关键绩效指标。增效衡量被广泛认为是评估真实效果的最佳方法，但值得注意的是，基于单日点击归因的测试所选的营销要素与严格的对照测试所选的不同。这意味着单日点击归因并非总能提供最佳的预算花费指导。综合考虑错误匹配的规模即可知道增效衡量投资是否值得。”

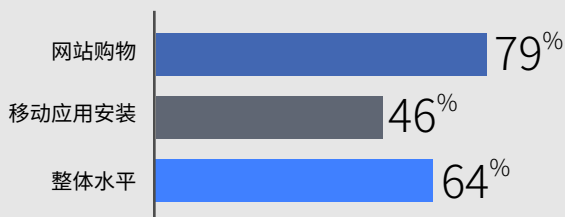
¹Facebook 针对 2015 年 2 月至 2017 年 9 月期间开展的 580 次转化提升 A/B 测试的分析。引用的百分比代表分析的所有测试的整体平均百分比（统计学意义结果 @90% CI）；2017 年 10 月。

研究发现，增效衡量很有价值，能够为营销者作出更高效的营销花费决策提供参考或指引。所作的决策越关键，增效衡量就越重要。

点击归因与成效衡量所选的最有效营销策略存在差异的情况



点击归因与增效衡量所选的最有效营销策略不同时的 CPA 差异



Facebook 针对 580 次转化提升 A/B 测试的分析。引用的百分比代表所有测试的整体平均百分比（统计学意义结果 @ 90% CI）；2017 年 10 月。

新手入门四步指南

采用新的成效衡量方法并非总是易事，但其结果却可能显著提升企业实现业务目标的能力。注重试验并在实践中采用“从测试中学习”方法的企业将获得有利的开局，但所有营销者都应当思考如何通过将增效衡量纳入营销决策进行创新。

按照这四步操作的企业能够深入细致地了解他们的广告投放方法的效果。



给所有营销者的实用建议

来自 Airbnb 的 Gupta 为想要通过增效衡量促进业务发展的营销者提供了一些建议。“首先要确保全面协调统一优化目标，例如最大限度增加固定预算的投资回报。”他说，“其次要充分意识到，每个渠道的回报通常会逐渐减少，也就是投资带来的增效会慢慢降低。因此，这其中的挑战在于寻找能够带来更高投资回报的新渠道。”

来自 eBay 的 Flanery-Rye 表示，企业应该从追踪营销预算开始增效衡量。“首先要做的是找到花费预算最多的地方。找到之后，紧接着就是从小处着手，不断从中学习并做出调整。开展试验并查看结果，总结经验与教训。”他说道，“无论是人工学习还是机器学习，错误都是最好的‘老师’，从错误中可以学到有用的知识和解决问题的方法。”

Lewis 指出，这种严格的衡量方法“与全球一些从未开展测试的顶级广告主形成了鲜明的对比。他们的营销活动创建后就一成不变，并且在结束后 8-12 周才评估营销活动是否有效。”因此，各公司需要招聘和支持懂测试、衡量和学习的人才。

Flanery-Rye 说，公司高管也应该讨论营销投资规模，以及应该采用和遵循的衡量精度。讨论结果可帮助判断公司要取得成功所需的员工数量、合作伙伴和技术投资。“如果要快速作出大量决策，那么你需要的绝不是要三周时间才能运行 A/B 测试的庞大系统。”他说道，“你需要的是能与自己的决策节奏匹配的衡量精度。”

利用现代化衡量工具促进业务增长

虽然可供采用的衡量和归因系统有很多，但我们要面对的事实是消费者行为已经完全改变了数字营销。消费者的注意力越来越分散在不同的设备和渠道。因此，主导当今数字营销的一些方法有时候缺乏透明度和预见性，而透明度和预见性正是企业制定正确的营销预算决策所需要的。

值得庆幸的是，科学技术正在迅猛发展，新的衡量解决方案层出不穷，例如增效衡量，这便于企业将互动与销量更轻松地串联起来。通过开发更成熟和严格的衡量方法，各类企业都可以提升他们的广告投资回报，以更合理和高效的方式使用资源和预算。