

АВ-тестирование

A/B-тестирование — это неотъемлемая часть процесса работы над продуктом. Это эксперимент, который позволяет сравнить две версии чего-либо, чтобы проверить гипотезы и определить, какая версия лучше. Должны ли кнопки быть черными или белыми, какая навигация лучше, какой порядок прохождения регистрации меньше всего отпугивает пользователей



A/B-тестирование — это метод сравнения двух версий страницы или приложения, который помогает понять, какая из них работает лучше в контексте какой-то конкретной задачи. **A/B-тест — один из самых популярных методов повышения результативности цифровых продуктов: сайтов, мобильных приложений, SaaS-продуктов, рассылок и других т.д.**

Участники тестирования

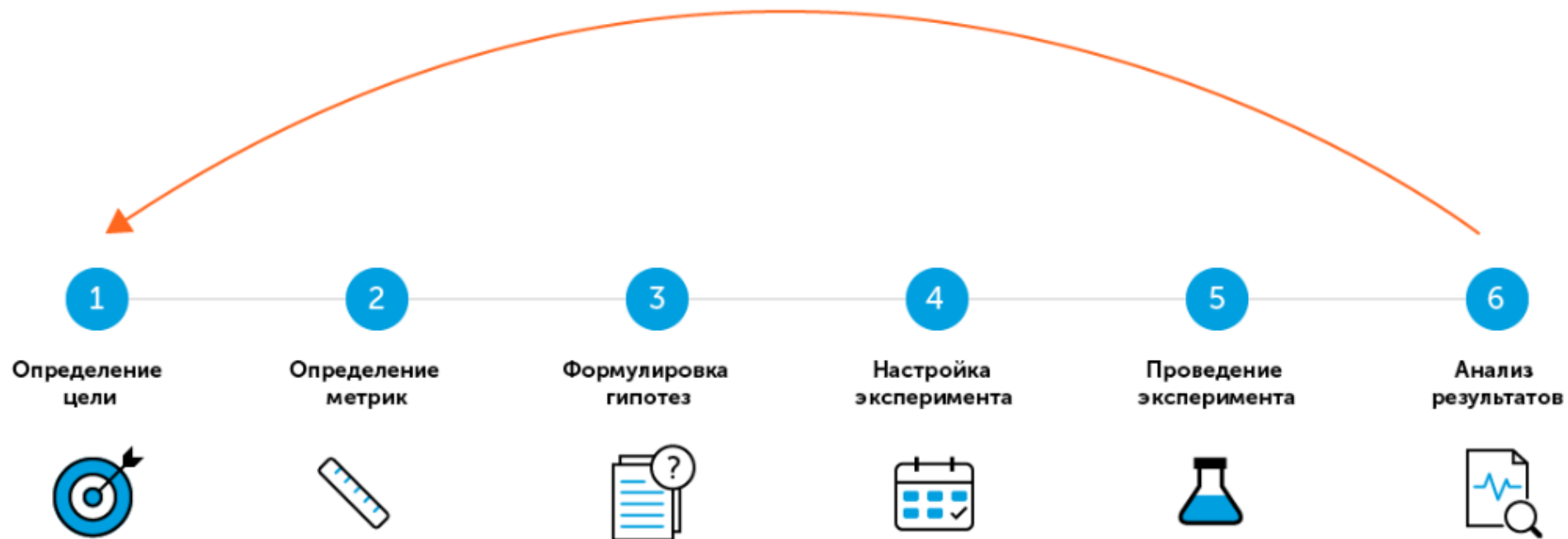
Аналитик участвует в настройке теста:
формирует требования для запуска,
определяет метрики,
проверяет корректность работы,
после окончания анализирует результаты.

Продакты и маркетологи должны понимать, что они могут протестировать, как провести эксперимент и проанализировать его результаты.

Маркетологи используют A/B-тестирование, чтобы привлечь больше людей и лучше конвертировать их в покупку. Чаще всего, они проводят эксперименты при работе с email-рассылками, push-уведомлениями, лендингами и рекламными креативами.

Продакт-менеджеры, в основном, используют A/B-тестирование для улучшения метрик продукта. Например, они могут сравнивать две версии онбординга, разные варианты фичей продукта и их удобство для пользователей.

Этапы АВ тестирования



Марк Цукерберг в одном из интервью признался, что успех Fb кроется в их уникальной системе тестирования: «В любую секунду в мире работает не одна версия Fb, а порядка 10,000».

Тестирование темы рассылки

Гипотеза эксперимента: если в теме рассылки будет личное обращение, то больше пользователей откроют письмо и дочитают до конца, потому что письмо не выглядит как рекламная рассылка.

вариант А
(без изменений)

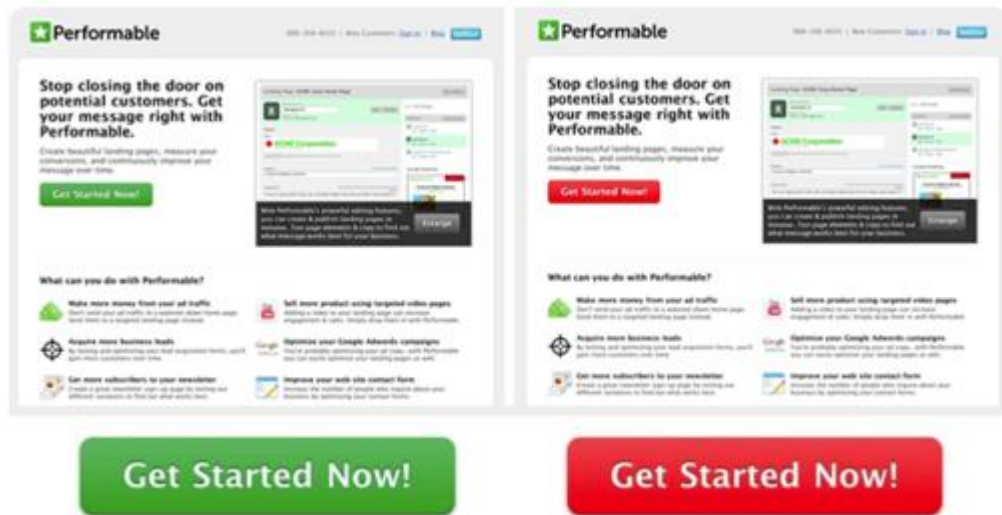
от: skillsetter.io
кому: мне

тема: Новый курс по
продакт-менеджменту в IT

вариант В
(имитация личного обращения)

от: skillsetter.io
кому: мне

тема: Привет, NAME! Хочешь
освоить новую профессию
в IT?



Здесь был протестирован цвет кнопки. Поразительно, но выяснилось, что конверсия красной на 21% выше!

Цели исследования

У исследования есть первоначальная цель — это улучшение показателей, оценка пользовательского опыта. При А/В исследовании посетители могут даже не знать, что они участвуют в тесте. Чаще всего он используется в таких целях:

- **Улучшение юзабилити.** Сюда можно отнести дополнение или упрощение функционала. Может проводиться в дополнение к коридорному тесту, как исследование «в полях».
- **Персонализация коммуникации с пользователями.** С этой целью чаще работают с текстами — на странице, кнопках, уведомлениях, формах обратной связи.
- **Увеличение числа конверсий.** Тут экспериментируют с формами подписки, кнопками, важными элементами страницы.
- **Адаптация действующей системы под актуальные потребности аудитории.**
- **Снижение числа отказов на сайте.** Построение и проверка гипотез, что именно служит причиной: неудобная форма заказа, несоответствие ожиданиям аудитории и т. д. Если в ходе эксперимента показатели страницы улучшаются, можно использовать изменения на все аналогичные страницы.
- **Минимизация рисков при внесении изменений.** Главный плюс метода — продажи не теряются. Из выбранного количества посетителей половина будет видеть старую версию, а потому новые элементы (если они менее эффективны) не приведут к существенному уменьшению числа конверсий.

Распространенные ошибки в проведении исследования

- **Исследование нескольких изменений.** Если нужно повысить конверсию формы обратной связи и при этом вносятся изменения в содержание, цвет кнопок, текст, количество полей — как определить, что из этого сработало? Один элемент вызывает положительную динамику, а другой, наоборот, отрицательную. Результаты смазываются, тестирование получается недостоверным.
- **Использование чужих гипотез.** Учитывайте особенности каждого отдельного бизнеса, функционал сайта и т. д. На ресурсах отличается оптимизация, трафик, ЦА, а потому не все гипотезы будут работать одинаково.
- **Неправильная продолжительность.** Нельзя заканчивать эксперимент раньше времени. Но и слишком надолго растягивать его тоже нельзя. Оптимально — 14 дней.
- **Разовое проведение теста.** Исследование может проводиться несколько раз подряд — тестировать новые гипотезы на основании предыдущих, дополнять изменения на сайте и проверять, какое из них лучше работает.
- **Игнорирование внешних факторов.** Поведение пользователей зависит от многих причин: сезонность, периоды распродаж, праздники и т.д. Из-за этого исследование может давать недостоверные показатели. Проводите тесты нейтральные периоды.
- **Недостаточно полное отслеживание метрик.** Если отслеживать только основной показатель, не учитывая при этом дополнительные, результат будет неполным. Учитывайте связанные показатели — их изменение может дополнить начальную гипотезу и дать материал для построения следующих.

Для каких элементов подходит - примеры

- Какие элементы лучше всего подходят для А/В тестирования?
- Некоторые элементы маркетингового актива способствуют конверсии больше, чем другие. Например, изменение одного слова в тексте поста, вероятно, не повлияет на конверсию в значительной степени.
- Чтобы у вас сложилось полное представление хочу показать **7 наиболее эффективных элементов** для проведения А/В тестирования. Вот они:
- Заголовки
- Текст
- Призыв к действию
- Медиаматериалы (Изображения, аудио и видео)
- Объем информации
- Описание продуктов или услуг
- Отзывы

Что тестировать в первую очередь

На сайте можно проверить чуть ли не каждый элемент — заголовки, цены, кнопки, фотографии и прочее. Первый вопрос — с чего начать? Есть три подхода для начала теста:

Способ 1: протестировать страницы с высоким трафиком

Если вы улучшите конверсионность страницы, которая и так получает много трафика, вы сможете получить еще больше конверсий. Вы сможете удостовериться, что изначальный вариант работает по-максимуму, или найдете лучший.

Способ 2: протестировать страницы с самыми плохими показателями

Вы можете начать A/B тестирование на страницах с низкими показателями конверсии. Хуже им не будет, зато вы сможете улучшить их конверсионность в несколько раз.

Способ 3: оценить качественные и количественные данные

Используйте данные юзабилити-тестирований и других исследований. Если вы знаете, что пользователи не могут найти кнопку для бронирования, с помощью тестов вы подберете лучшее место для этой кнопки. Первым тестом может быть гипотеза «Поместить кнопку нужно по центру страницы».

Этапы исследования

Само проведение можно разделить на несколько частей.

1. Построение гипотезы

- На первом этапе проводится **анализ ресурса**. Здесь маркетолог определяет, что именно можно изменить и как это может сказаться на эффективности.
- В гипотезе должны быть сформулированы предположение и результат. Например: «Если мы заменим **конверсионную форму на более простую**, она начнет приносить больше лидов»
- Новая страница с визуальными результатами сделана для запросов, связанных с поиском вдохновения. Например, «Идеи для декора на Хэллоуин» или «Идеи вертикального сада для улицы». Новая функция уже доступна

1. Построение гипотезы

- В идеале гипотезу формулируют на основании исследования пользовательского опыта — **проведения интервью или коридорного тестирования, пятисекундного теста**. Иногда в формировании гипотез помогает изучение отчетов из метрики: **списка конверсий, карты скроллинга, карты кликов**.
- Также в гипотезе важно выделять, какой показатель должен измениться.
- Дополнительно отслеживайте другие метрики — **время пользователей на сайте**, например. Это поможет в дальнейшем собрать более полную картину исследования.

2. Подключение метрик

- Выберите программу, с помощью которой будете проводить тестирование.
- В метриках выберите основные параметры проведения эксперимента:
 - Длительность.
 - Процент разделения аудитории — кто будет видеть версию «А», а кто будет видеть версию «В».
 - Минимальный размер выборки — его можно вычислить с помощью калькулятора. Программа помогает рассчитать, сколько пользователей нужно для участия в эксперименте. Для этого нужно знать базовую конверсию и минимальный видимый эффект — относительную разницу в конверсии в сравнении с базовой. Чем она ниже, тем больше выборка потребуется.

Метрики эксперимента: уровни

- Целевые — показатели, на которые направлено изменение (конверсия, средний чек).
- Опережающие — показатели, хорошо коррелируемые с целевыми, дающие предикт и полезные тогда, когда нет времени ждать основную метрику.
- Guardrail — показатели, на которые направленно влияет изменение, но которые не являются целевыми. Рекомендуется наблюдать за ними и в том числе на их основе принимать решение.

Примеры системы уровней

Пример 1. E-commerce, тест нового UI корзины.

- Целевые: конверсия в покупку, средний чек, ARPU (средний доход на пользователя), ARPPU (средний доход на платящего пользователя).
- Опережающие: добавление товара в корзину на сессию, просмотры товаров на сессию, отток чекаута, ошибки на чекауте.

•Guardrail: время от входа в корзину до её прохождения, доля поисковых запросов из корзины, взаимодействие с рекомендательными блоками в корзине.

Пример 2. Образовательный продукт, тест нового образовательного контента.

- Целевые: продление обучения, средний доход на платящего пользователя (ARPPU).
- Опережающие: интенсивность обучения, количество ошибок в момент обучения, частота обращений в службу поддержки, технические характеристики качества видео.
- Guardrail: время, проведённое за одним занятием, прерывание занятий, перемотка видео.

3. Проведение теста

Стандартный срок выполнения исследования составляет **10-14 дней**. Первую неделю называют «чистой». В это время статистика только начинает собираться, а пользователи — привыкать к изменениям. **Даже если изначально определяется лидер**, не нужно прерывать эксперимент.

Примеры:

1. Ложноположительное решение

Гипотеза: повышение недельной цены подписки с 1 до 2

Итог: в первые 2 дня эксперимента был зафиксирован статистически значимый результат.

Продакт и аналитик приняли решение принять результат как успешный, а после публикации изменения на всех пользователей ключевой показатель изменился в худшую сторону. Таким образом, приняли за истину то, что ей не является.

2. Ложноотрицательное решение

Проводили эксперимент 2 дня, не видели разницы, остановили

Итог: поскольку эксперимент охватил только 2 дня, мы не учли поведение аудитории в остальные дни недели. Возможно, изменение имеет отложенный эффект: в понедельник пользователь попал в тестовую группу, а в пятницу принял решение.

4. Сбор результатов и формирование выводов

- Для оценки достоверности теста существуют специальные калькуляторы. На основании размера выборки и полученных результатов он помогает определить, можно ли считать результаты теста достоверными.

Компания Server Density протестировала изменения цен и увеличила доход на 114%

Server Density представляет услуги мониторинга серверов. Они решили изменить цены таким образом, чтобы увеличить общую выручку.

Гипотезы было две:

- 1.Повышение цены сократит количество бесплатных регистраций
- 2.Повышение цены увеличит выручку, несмотря на то, что регистраций станет меньше.

На сайте отслеживались и тестировались две цели:

- 1.Бесплатная регистрация
- 2.Покупка плана (upgrade)

Старая версия сайта

Pricing
Awesome clients have awesome discounts

Monthly Yearly £ \$

10% DISCOUNT

Mix and match server monitoring e.g. CPU, disk, MongoDB, etc with website monitoring e.g. up/down status and response time.

\$7 per server per month 40% DISCOUNT

\$8 per server per month 20% DISCOUNT

\$11 per server per month

1 Server

\$11

+

\$2

1 Website

=

\$13 per month

Sign up to the free trial

Цена на ней менялась в зависимости от количества серверов и сайтов, которые нужно было мониторить. Было логично, что чем больше объектов обслуживалось, тем выше была общая стоимость.

Несмотря на то, что большинству клиентов требовалось обслуживание в среднем 7 серверов, им показывалась цена обслуживания одного сервера — \$11+\$2, в надежде, что небольшое число привлечет внимание и это увеличит базу клиентов.

Pricing

Enterprise \$499 / month

- ① Up to **100** servers
- ① Up to **100** websites
- ① **1000** SMS credits
- E-mail & phone** support
- ① **∞** users + permissions

[Sign up](#)

Small Business \$299 / month

- ① Up to **50** servers
- ① Up to **50** websites
- ① **500** SMS credits
- E-mail** support
- ① **∞** users

[Sign up](#)

Starter \$99 / month

- ① Up to **10** servers
- ① Up to **10** websites
- ① **100** SMS credits
- E-mail** support
- 1** user

[Sign up](#)

Тут можно заметить, что 10 серверов и 10 сайтов по старым ценам стоили бы \$130 в месяц, то есть они на самом деле снизили цену, но увеличили среднюю стоимость заказа.

Результат:

	Control (per unit pricing)	Variation (packaged pricing)
Goal 1 (Free Signup)	135 conversions 6.92% conversion rate	100 conversions 5.19% conversion rate
Goal 2 (Paid Upgrade)	20 conversions	15 conversions
Average Order Value (AOV)	US\$ 19.70	US\$ 55.53
Total Revenue	US\$ 394	US\$ 833
Test conducted using Visual Website Optimizer. Chance to beat original 93%		

Бесплатные регистрации:

Исходная страница — 1950 посетителей, 135 регистраций, 6.92% конверсия

Тестируемая страница — 1925 посетителей, 100 регистраций, 5.19% конверсия

Результат: конверсия бесплатных регистраций уменьшились на 24.96%

Покупки плана:

Исходная страница — 1950 посетителей, 20 покупок, средний чек \$19.70, общая выручка US\$394

Тестируемая страница — 1925 посетителей, 15 покупок, средний чек \$55.53, общая выручка \$833

Результат: 114% увеличение общего дохода.