



команда роста на аутсорсе
для системного улучшения ключевых метрик

Чек-лист по A/B-тестированию: От гипотезы до результата

Пошаговый гайд для увеличения конверсии и избежания ошибок

1. Подготовка к A/B-тесту (Цели, Гипотезы, Метрики)

□ Раздел 1.1: Определение бизнес-цели и ОЕС (Overall Evaluation Criterion)

Четко сформулируйте, какую бизнес-метрику вы хотите улучшить (например, рост выручки, увеличение количества регистраций, снижение отказов). Это будет ваша ОЕС.

□ Раздел 1.2: Формулирование сильной гипотезы (по шаблону)

Используйте шаблон: 'Если мы изменим [X] на [Y], то [Z] вырастет/уменьшится на [N]%. Пример: 'Если мы уберём поле «Комментарий» из формы заказа, количество завершённых заказов увеличится на 5%'. Гипотеза должна быть основана на данных (аналитика, опросы) (Source: direct.yandex.ru).

□ Раздел 1.3: Выбор Guard-метрик (защитных метрик)

Определите метрики, которые не должны ухудшиться в результате изменений (например, средний чек, время на сайте, количество страниц за сессию). Это поможет избежать негативного влияния на другие важные показатели (Source: career.hh.ru).

□ Раздел 1.4: Приоритизация гипотез (фреймворки ICE/PIE)

Оцените гипотезы по Impact (влияние), Confidence (уверенность) и Ease (легкость реализации) или Potential (потенциал), Importance (важность) и Ease (легкость). Выбирайте те, что имеют наибольший потенциал при разумных затратах.

2. Планирование эксперимента

□ Раздел 2.1: Выбор объекта тестирования (заголовки, кнопки, формы и т.д.)

Начните с элементов, которые сильнее всего влияют на ключевую метрику и находятся 'above the fold' (видимы без прокрутки). Примеры: заголовки, кнопки призыва к действию, формы, структура страницы, предложения, цены, скорость загрузки (Source: direct.yandex.ru).

□ Раздел 2.2: Определение типа теста (A/B, A/B/n, MVT, A/A, MAB)

Для начинающих: A/B-тест (сравнение двух версий). Для нескольких вариантов: A/B/n. Для сложных взаимодействий (требует много трафика): Мультивариантное тестирование (MVT). Для проверки корректности системы: A/A-тест. Для оптимизации распределения трафика: Многорукие бандиты (MAB) (Source: direct.yandex.ru).

□ Раздел 2.3: Расчёт размера выборки и минимально обнаруживаемого эффекта (MDE)

Чем меньше разница в конверсии между вариантами, тем больше трафика нужно. Для большинства сайтов достаточно 5–10 тысяч визитов на вариант для каждого варианта (Source: article_content). Используйте онлайн-калькуляторы для определения необходимого размера выборки.

□ Раздел 2.4: Определение длительности теста (минимум 2 полных недели)

Проводите тест минимум одну полную неделю, чтобы учесть поведение пользователей в разные дни, включая выходные. Оптимально — 2 полные недели, чтобы охватить два цикла рабочей недели и выходных (Source: article_content).

□ Раздел 2.5: Выбор инструмента/плагина для тестирования

Бесплатные: Яндекс Метрика ('Эксперименты'), Microsoft Clarity (для сбора гипотез). Для WordPress: EasyTest, Split Test For Elementor. Платные: Optimizely (от \$99/мес), VWO (от \$99/мес) (Source: article_content).

3. Реализация теста

□ Раздел 3.1: Создание тестовых вариантов (вариант В, С и т.д.)

Убедитесь, что тестовые варианты отличаются только одним параметром, если это А/В-тест. Используйте визуальные редакторы или отдельные URL.

□ Раздел 3.2: Настройка распределения трафика (50/50, 80/20 и т.д.)

Обычно трафик делится 50/50. Если гипотеза рискованная, можно направить на тестовую версию 20% трафика (80/20) (Source: article_content).

□ Раздел 3.3: Проверка корректности работы теста (А/А-тест или пробный запуск)

Перед основным тестом проведите А/А-тест: покажите обеим группам одинаковую версию. Если результаты существенно различаются, есть ошибка в настройке (Source: direct.yandex.ru).

□ Раздел 3.4: Мониторинг технических ошибок и аномалий

Регулярно проверяйте распределение трафика (SRM - Sample Ratio Mismatch). Если вместо 50/50 получается 60/40, тест некорректен (Source: article_content). Отслеживайте технические сбои.

4. Анализ результатов и принятие решений

□ Раздел 4.1: Оценка статистической значимости

Результат считается достоверным, если разница между вариантами не является случайной. Обычно используют порог значимости 95% (вероятность случайности менее 5%) (Source: article_content, career.hh.ru).

□ Раздел 4.2: Интерпретация результатов (победитель, проигравший, отсутствие результата)

Не останавливайте тест преждевременно, даже если один вариант 'побеждает'. Дождитесь достижения статистической значимости. Разница в 2-3% на небольшой выборке может быть шумом (Source: article_content).

□ Раздел 4.3: Анализ влияния на Guard-метрики

Убедитесь, что победивший вариант не ухудшил защитные метрики. Например, рост конверсии не должен сопровождаться падением среднего чека.

□ Раздел 4.4: Выводы и дальнейшие шаги (внедрение, итерация, новый тест)

Если тест успешен и статистически значим – внедряйте изменение. Если нет – анализируйте, почему гипотеза не сработала, формулируйте новую и запускайте следующий тест (итерация).

□ Раздел 4.5: Документирование результатов

Записывайте все детали теста: гипотезу, варианты, метрики, длительность, результаты, выводы. Это поможет избежать повторения ошибок и создаст базу знаний.

Типичные ошибки А/В-тестирования

Избегайте распространённых ошибок, чтобы результаты ваших тестов были достоверными и приносили реальную пользу:

- * Преждевременная остановка теста: Не останавливайте тест, пока не достигнута статистическая значимость, иначе результат может быть случайным (Source: article_content).
- * Тестирование нескольких изменений одновременно: Изменяйте только один элемент за раз, чтобы понять, что именно повлияло на результат. Исключение — MVT, но он требует большого трафика (Source: article_content).
- * Недостаточный объём трафика: При малом трафике тест может длиться месяцы, и за это время могут измениться внешние условия. Для большинства сайтов нужно 5-10 тысяч визитов на вариант (Source: article_content).
- * Игнорирование внешних факторов и сезонности: Не запускайте тесты во время праздников или распродаж, данные будут искажены. Проводите тесты в стабильный период (Source: article_content).
- * Фокус на незначительных элементах: Сначала тестируйте крупные блоки (оффер, цена, структура), которые дают больший потенциальный прирост конверсии (Source: article_content).

- * Ошибки в аналитике и некорректная настройка целей: Убедитесь, что цели настроены правильно и отслеживание работает корректно (Source: article_content).
- * Перекос трафика (Sample Ratio Mismatch, SRM): Проверьте соотношение групп. Если распределение трафика отклоняется от заданного (например, вместо 50/50 получается 60/40), результаты теста нельзя считать достоверными (Source: article_content).
- * Игнорирование сегментов аудитории: Анализируйте результаты по сегментам (устройствам, источникам трафика, географии), так как общий результат может скрывать противоположные эффекты (Source: article_content).
- * Доверие «проверенным» паттернам без собственной проверки: Всегда проверяйте гипотезы на своей аудитории, даже если они основаны на авторитетных источниках. Например, исследование Рона Кохави показало, что скруглённые кнопки не всегда дают заявленный прирост кликов (Source: article_content).

Когда А/В-тестирование не имеет смысла?

Недостаточный трафик	Результаты будут статистически незначимыми (Source: article_content).
B2B-продукты	Если решение о покупке принимают одни люди, а пользуются продуктом другие, тестирование на этой аудитории может быть нерелевантным (Source: article_content).
Затраты превышают выгоду	Если затраты на разработку и запуск теста превышают потенциальную выгоду от него (Source: article_content).
Невозможность корректно разделить аудиторию	Если вы не можете точно и равномерно разделить пользователей на контрольную и тестовую группы (Source: article_content).

Больше интересного в нашем блоге

развивайте бизнес-мышление с нашими исследованиями



Получить пользу

