

Usability-тестирование

Думин Павел Николаевич
dumin.pn@gmail.com



Содержание:

1. Что такое юзабилити-тестирование?
2. Почему именно оно?
3. Что именно измерять?
4. Насколько глубоко удовлетворение?
5. Что нужно для тестирования?
6. Проведение тестирования
7. Анализ результатов
8. Подготовка отчета

Что такое юзабилити-тестирование?

Эксперимент, направленный на измерение качества интерфейса или же поиск конкретных проблем в нем.

Тестирование позволяет:

- Понять, насколько плохо или хорошо работает интерфейс, что может либо побудить улучшить его, либо, если он уже достаточно хорош, успокоиться
- Сравнить качество старого и нового интерфейсов, как следствие – обоснование изменения или внедрения чего-л.
- Найти и опознать проблематичные фрагменты интерфейса

Почему именно оно?

Относительно низкая трудоемкость и стоимость (есть пути понижения стоимости):

- Упрощение понятия юзабилити
- Отказ от сбора количественных данных
- Снижение стоимости оборудования и уменьшение оплаты респондентов (использование только мобильных лабораторий)

Effectiveness и efficiency

ث Usability - the extent to which a product can be used by specified users to achieve specified goals with *effectiveness, efficiency* and satisfaction in a specified context of use.

Эффективность

- Скорость работы пользователя
- Скорость обучения
- Количество человеческих ошибок



Трудоемкость

- Успешность – соотношение выполненных тестовых заданий к невыполненным или выполненным полностью неправильно
- Мощность – соотношение задач из деятельности пользователя к задачам, для решения которых данный продукт предназначен
- Нагрузка на пользователя – количество затраченного «мозгового усилия» + количество действий пользователя на задание

Трудоемкость

- Успешность – соотношение выполненных тестовых заданий к невыполненным или выполненным полностью неправильно
- Мощность – неизмеряемый показатель
- Нагрузка на пользователя – практически полностью проявляется в эффективности

Отказ от количественных данных

- Составляем тест на измерение качественных данных
- Сложность интерпретации обработанных количественных данных

*«There are three kinds of lies: lies, damned lies, and statistics» (c)
Benjamin Disraeli*



Ненадежные цифры

- Реальные пользователи могут отличаться от выбранных респондентов → недостоверные результаты
- Тестовые задания могут неадекватно отражать реальную деятельность пользователя
- Сложность анализа

Парадокс

- 9 групп юзабилити-специалистов тестировали HotMail
- Было выявлено 310 проблем
- ≈ 230 проблем выявила одна группа — остальные группы этих проблем не заметили



Rolf Molich

<http://www.dialogdesign.dk/cue.html>

Что измерять?

Скорость работы пользователя

Метрики:

- длительность выполнения операции
- время, затраченное на обнаружение ошибок
- время, затраченное на исправление ошибок
- количество команд, исполняемых при выполнении операции
- длительность поиска информации в документации
- количество команд, более эффективных, чем использованные пользователем
- снижение производительности при длительной работе

Что измерять?

Ошибки

Метрики:

- процент операций, вызвавших ошибку
- среднее число ошибок на операцию у ОПЫТНЫХ пользователей
- количество ошибок, не обнаруженных и не исправленных пользователями

Что измерять?

Обучаемость навыкам работы с системой

Метрики:

- количество и частота обращений к справочной системе
- длительность периода между началом использования системы и точкой, в которой скорость работы и/или количество ошибок перестает расти/падать
- разница в количестве ошибок/скорости работы у пользователей с опытом пользования системой из без опыта

Что измерять?

Сохранение навыков работы с системой

Метрики:

- разница в скорости работы/количестве ошибок у пользователя после часа работы с системой и у того же пользователя в начале использования системы после долгого перерыва

Удовлетворенность

- Опрос: насколько интерфейс кажется Вам удовлетворительным?
- Наблюдение в контрольные моменты времени – относительная оценка, используется в качестве сравнения

Анкетирование

- Создание формальных анкет – попытка объективизации процесса получения информации (одинаковые вопросы → «одинаковые ответы»)
- Надежных анкет не бывает (см. литературу)

Анкета по словам

Устаревший – Эффективный – Нечеткий – Неудобный –
Замусоренный – ... – Медленный – Круглый –
Печальный – Недружественный – Предсказуемый –
Непонятный – Быстрый – Головоломный – Грустный –
Приятный

- Не требуется адаптация
- Ненадежно

Формальная анкета

В любой момент времени я понимал, что должен сделать дальше

- Никогда
- Редко
- Регулярно
- Часто
- Всегда

- Формальная процедура оценки
- Требуется адаптации

Что нужно для тестирования:

- Респонденты
- Метод тестирования
- Тестовые сценарии
- Рабочее место для теста и отлаженный метод фиксации материала
- Протестированный тест

Общие требования к респондентам

- Опыт работы с системой
- Уровень компьютерной грамотности
- Возраст
- Уровень эмоциональной открытости



Организационные вопросы

- Возложить поиск респондентов на заказчика
- Фокус группы, предоставленные агентствами по подбору персонала
- «Свои люди»
- Выбирать то время/место, которое удобно респонденту. Тестирование – это одолжение
- Вознаграждение

Методы тестирования

- Пассивное наблюдение за выполнением задания
- Поток сознания («думать вслух»)
- Активное вмешательство

Тестовые сценарии

- Тестовая задача
- Значимые метрики
- Задания
- Признаки успешности выполнения задачи

Тестовая задача –

задача, которую ставит перед пользователями их деятельность, и которая имеет самостоятельную ценность для пользователей. Пользовательская задача выполняется как одна или несколько операций (пользовательская операция не имеет самостоятельной ценности).

Метрики задачи:

- Успешность – выполнение респондентами более 90% заданий верно
- Эффективность – скорость: регистрация на сайте выполняется меньше N минут
- Эффективность – ошибки: при вводе 10 форм, количество ошибок не превышает 2
- Эффективность – обучаемость навыкам работы с системой: при выполнении задания 9, отличающегося от задания 2 только вводимыми данными, респонденты не совершают ошибок
- Удовлетворенность – по результатам анкетирования число баллов выросло на 20% по сравнению с предыдущими результатами

Метрики тестовых задач:

- Однозначность
- Полнота
- Краткость
- Отсутствие подсказок
- Точка начала выполнения задания

- «Двойное задание»
- Способы решения

Рабочее место и способы фиксации данных

- Стационарные рабочие места
- Мобильные рабочие места
 - Ноутбук
 - Веб-камера
 - Микрофон
 - Необходимое ПО
 - Средства фиксации времени
 - Тестовые задания (+ вводная форма, анкеты при необходимости)

Проведение тестирования

1. Ввод респондента в задачу
2. Выяснение его ожиданий от системы
3. Тестирование интерфейса
4. Опрос
5. Завершение тестирования

Анализ действий респондентов: ошибки

- Критические ошибки – респондент не разобрался в структуре интерфейса
- Некритические ошибка – оценка частотности, анализ происхождения
- Ошибка задания – переформулировка/исправление задания. Можно попросить респондента

Анализ действий респондентов: замедление работы

- Сложность наблюдения
- Фиксация положения курсора
- Мимика

Отчет по тестированию: общий вид

1. Резюме
2. Основные проблемы (интерфейсные проблемы, проявляющиеся по всему интерфейсу)
3. Частные проблемы (проблемы, проявляющиеся на отдельных экранах)
4. Количественные данные (если они собирались)
5. Приложение 1. Методика эксперимента и условия теста
6. Приложение 2. Описание тестовых сценариев
7. Приложение 3. Описание респондентов.