**Лабораторная работа № 12 (2020)  
Юзабилити-тестирование путем** **имитации поведения пользователей**

**1. Цель лабораторной работы:**

1.1. Приобретение умений по проведению юзабилити-тестирования.

1.2. Выбор методов тестирования.

1.3. Приобретение практических навыков по построению тестовых сценариев, составления анкет и анализу полученных результатов.

**2. Краткие теоретические сведения**

**Юзабилити-тестированием** является любой эксперимент, направленный на измерение качества интерфейса или же поиск конкретных проблем в нем.

**Методики тестирования**

Для выявления проблем удобства использования приложения, в том числе на ранних этапах планирования и разработки ПО, используется методику двойной проверки:

• Изучение опыта взаимодействия пользователя с приложением **через имитацию поведения пользователей**;

• Проверка соответствия принципам обеспечения удобства пользования и корректного визуального представления в контексте функциональных требований посредством **экспертной оценки**.

При проверке поведения приложения путем **имитации поведения пользователей** мы получаем полную информацию, необходимую для быстрого устранения выявленных дефектов, которые могут негативно сказаться на юзабилити приложения или пользовательском впечатлении в целом. В то же время юзабилити-тестирование не может сделать из плохого продукта продукт хороший; оно всего лишь делает продукт лучше.

Юзабилити-тестирование может быть направлено на:

* **получение количественных данных** (нужных для измерения эргономичности интерфейса);
* **получение качественных** **данных** (нужных для того, чтобы понять, что именно плохо и как это исправить).

**Что нужно для тестирования:**

* респонденты;
* тестовые сценарии;
* метод тестирования;
* рабочее место для проведения теста и отлаженный метод фиксации материала;
* результаты тестирования.

**Респонденты**

Сначала необходимо определить общие требования к респондентам.

**Общие требования к респондентам**

* **Опыт работы с системой.** Общее правило: если оптимизируется интерфейс существующей системы, половина респондентов должна иметь опыт работы (на них можно определить проблемы переучивания при внедрении), а половина нет (на них определяется скорость обучения). Если существуют конкурирующие системы, лучше другая пропорция: треть с опытом работы с предыдущей версией, другая треть – с опытом использования конкурирующих систем, оставшаяся треть – без опыта работы с системой.
* **Уровень компьютерной грамотности.** При прочих равных предпочтительным выбором является реальный, т.е. совпадающий с опытом целевой аудитории, уровень у трех четвертей респондентов и низкий уровень – у оставшейся четверти (на ней можно определить большее количество проблем).

Уровень компьютерной грамотности удобно определять по следующей шкале:

1. **Высокий**. Респондент имеет компьютер на работе и дома, большая часть трудовой деятельности выполняется на компьютере, респондент самостоятельно использует компьютер как средство саморазвития, активно пользуется сервисами в интернете (например, регулярно покупает товары и услуги в онлайновых магазинах).
2. **Выше среднего**. Респондент имеет компьютер на работе и дома, большая часть трудовой деятельности выполняется на компьютере, но респондент не использует компьютер для решения задач, выходящих за пределы его основной деятельности (работает на компьютере «от звонка до звонка» и не больше).
3. **Средний.** Работа с компьютером является частью обычной (трудовой или личной) деятельности в течение двух лет или больше.
4. **Низкий**. Либо на работе, либо дома есть компьютер, но опыт работы с компьютером не превышает двух лет и компьютер не является значимым инструментом в работе.
5. **Очень низкий**. Опыт использования компьютера спорадический, по длительности не превышает трех лет. Компьютер не используется ни на работе, ни дома.

* **Возраст.** Оптимальная пропорция: три четверти респондентов имеет возраст целевой аудитории системы, оставшаяся четверть старше (на ней можно определить большее количество проблем).
* **Пол** респондентов**(**меньшее влияние на результат) – стоит увеличить количество женщин среди респондентов по сравнению с пропорцией в целевой аудитории, поскольку на женщинах легче определять проблемы при внедрении (женщины, в целом, медленнее обучаются, но, обучившись, лучше работают).
* **Уровень эмоциональной открытости респондентов.** Чем более скован респондент, тем меньше он способен сказать вам ценного. Даже определив наличие проблемы, вы не сможете добиться от него никакой информации о том, что эту проблему вызвало. Существует прекрасный способ решить проблему недостаточной эмоциональной открытости – стоит иметь базу респондентов и использовать их повторно. Респондент, уже знающий на опыте, что в юзабилити-тестировании нет ничего страшного, значительно охотнее идет на контакт и вообще более разговорчив.

Наконец, когда желаемые для теста свойства пользователей уже определены, пришло время подобрать таких респондентов, которые не только удовлетворяют вышеперечисленным требованиям, но и входят в целевую аудиторию системы.

**Сколько нужно респондентов**

В 1992 году Роберт Вирзи (Robert Virzi) предположил, что для теста достаточно пяти респондентов. Через год эстафету приняли Якоб Нильсен и Томас Ландауэр (Jakob Nielsen и Thomas K. Landauer) которые утверждали, что пяти респондентов достаточно для того, чтобы выловить 70% проблем и еще три респондента нужны для того, чтобы довести результативность до 85%. Все три автора писали только о тестировании маленьких систем.

Для более сложных систем придется выполнять несколько разных тестов на разных респондентах; без этого охватить весь интерфейс системы окажется попросту невозможно. В данной лабораторной работе должно быть **не менее 3-х респондентов**.

**Например:**

# Описание респондентов

|  |  |  |  |
| --- | --- | --- | --- |
| **Пользователь** | **Уровень к. гр.** | **Возраст** | **Ур-нь эмоц. откр.** |
| Павел Ж. | 5 | 26 | 3 |
| Дмитрий З. | 3 | 22 | 4 |
| Ольга Ж. | 5 | 26 | 3 |
| Синяк К. | 5 | 19 | 5 |
| Рома Ч. | 5 | 19 | 4 |
| ***Итог*** | *4,6* | *22,4* | *3,8* |

Респондент-1

Иван Михайлов – 32 года, владение смартфоном выше среднего, имеется стаж. Уровень эмоциональной открытости – средний.

Респондент-2

Ирина Красовская – 46 лет, владение гаджетом среднее. При выполнении задания возникли некоторые трудности, связанные с навигацией по приложению, было допущено несколько ошибок, однако впоследствии метода «проб и ошибок» пользователь быстро адаптировался и уже не допускал прежних ошибок.

**Тестовые сценарии**

Тестовый сценарий – это тестируемый аспект системы. Адекватно подобранные тестовые сценарии являются важнейшей предпосылкой качества тестирования.

Сценарии состоят из:

* ***пользовательской задачи;***
* ***тестовых заданий респондентам (заданий может быть несколько);***
* ***значимых эргономических метрик;***

Разберем их подробно.

**Пользовательская задача**

Первым шагом определения сценариев является **определение значимых пользовательских задач**. Эти задачи – исходный материал для составления сценариев.

Что такое пользовательская задача? Это 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). Например, для программы-почтового клиента задачами являются:

* написание и отправка письма
* прием сообщений с сервера
* настройка программы под свои нужды (например, настройка автоматического приема почты через заданные интервалы).

А вот выбор адресата из адресной книжки при написании нового письма уже не является задачей, потому что это действие не самоценно. Это операция, состоящая из многих действий (нажать на кнопку Кому… > выбрать контакт > подтвердить выбор).

При выборе задач для тестирования следует руководствоваться двумя соображениями:

* **Все задачи должны быть реальными**, т.е. выявленными из актуальной деятельности пользователей: желание сделать задания посложнее, чтобы найти разом много проблем, несостоятельно – задания должны быть обычными, поскольку нет смысла искать проблемы, с которыми никто не сталкивается.
* Поскольку протестировать весь интерфейс на всех пользовательских задачах возможно только в идеале, приходится ограничивать себя и **выбирать только важные задачи**.

Важными задачами являются:

* частотные задачи, т.е. которые выполняются всеми пользователями и/или выполняются часто;
* все остальные задачи, которые, как вы подозреваете, выполняются в системе плохо;
* задачи, неправильное выполнение которых приводит к крупным проблемам.

**Тестовое задание**

***Тестовое задание*** – это то, что получает от вас респондент, задание, позволяющее провести респондента через фрагмент интерфейса системы и определить характеристики этого фрагмента.

Тестовые задания, помимо того, что должны соответствовать пользовательским задачам, должны обладать еще и следующими свойствами:

* **Однозначность**. Задания должны быть сформулированы так, чтобы исключить их неправильное толкование респондентом. Если респондент поймет задание неправильно, вам почти наверняка не удастся походу теста направить его на правильный путь, не подсказав ему одновременно последовательности выполнения задания.
* **Полнота**. В тексте задания должна присутствовать вся информация, необходимая для выполнения этого задания.
* **Краткость**. Если вы замеряете, скорость выполнения заданий, задания должны быть достаточно краткими, чтобы длительность чтения заданий респондентами не влияла на продолжительность выполнения самих заданий (люди читают с разной скоростью). Если текст задания будет велик по объему, вам придется вручную отсекать длительность чтения для каждого задания, что очень трудоемко.
* **Отсутствие подсказок**. По тексту задания не должно быть понятно, как это задание нужно выполнять. Например, недопустимо использовать терминологию системы – вместо каждого термина нужно расписывать его значение, иначе респонденты просто нажмут кнопки с теми же словами, и вы не выявите никаких проблем.
* В задании должна присутствовать **точка начала выполнения задания**, т.е. должно быть прописано то окно или экран, на котором респондент должен находиться в начале. Если такой информации представлено не будет, респонденты неизбежно будут переходить к другим фрагментам интерфейса, а значит, задание разными респондентами будет выполняться по-разному, что делает бессмысленным все статистические расчеты. Фиксировать начальную точку задания нужно еще в конце предыдущего задания. Если задание начинается с чистого листа, в конце предыдущего задания должно быть написано «вернитесь на главный экран». Если задание должно начинаться с места, на котором закончилось предыдущее задание, предыдущее задание должно заканчиваться словами «закончив, не закрывайте текущее окно/останьтесь на этом экране».

**Например**,

*Пользовательская задача* – Просмотреть прайс-лист и заполнить форму для заказа мероприятия.

*Тестовое задание*:

Вы должны заказать мероприятие, стоимость которого не должна превышать 2000 у.е. и кол-во гостей должно быть в диапазоне от 30 до 50 человек. Кроме того, вы должны указать дату проведения своего мероприятия (день и месяц рождения респондента).

*Тестовый сценарий:*

1. Открыть главную страницу и выбрать вкладку с надписью «Прайс-лист».

2. Просмотреть интересующие мероприятия и их цены, вернуться на главную страницу.

3. Перейти на вкладку «Заказать».

4. Заполнить форму в соответствии с указанными требованиями.

5. Нажать кнопку отправления и получить уведомление об успешной операции.

Тестовые задания  предъявляемые респондентам должны быть распечатаны, каждое задание на отдельном листе, чтобы респондент не мог забежать вперед и прочесть задания, которые он еще не выполнил.

На первом листе нужно выводить **вводную форму**:

*Уважаемый [Имя респондента]!*

*Предлагаем Вам выполнить ряд заданий, предназначенных для оценки простоты и удобства использования [Наименование системы]. При выполнении заданий чувствуйте себя свободно. Целью исследования является оценка качеств изучаемого интерфейса, а не Вас лично. Если Вы что-то сделаете неправильно, это будет значить, что интерфейс и только интерфейс нуждается в улучшении.*

*При выполнении заданий Вы должны действовать так, как считаете нужным. Например, если Вы решите воспользоваться Справкой, вы можете это сделать, не спрашивая разрешения экспериментатора.*

*Обратите внимание, что Ваши действия и слова записываются для дальнейшего изучения, но все собранные данные останутся строго конфиденциальными и будут доступны только исследователям.*

*Внимательно прочитайте задание и точно следуйте изложенным в нем инструкциям.*

*Старайтесь довести выполнение каждого задания до конца, но если во время выполнения задания Вы поймете, что не можете или не хотите его заканчивать, сообщите об этом экспериментатору и перейдите к следующему заданию.*

*Пожалуйста, переворачивайте страницу с заданием только тогда, когда выполните задание на открытой странице.*

*Если Вы не понимаете какое-либо задание, не стесняйтесь, переспросите проводящего тестирование специалиста.*

**Методы тестирования**

**Проверка качества восприятия.** Тест позволяет определить, насколько легко интерфейсу обучиться. Поскольку существует разница между понятиями видеть и смотреть, а запоминается только то, что увидено, необходимо обладать уверенностью в том, что пользователь увидит если не всё, то уж хотя бы всё необходимое. А значит – запомнит, благодаря чему в будущем ему не придется обшаривать меню в поисках «чего-то такого, что, я точно знаю, где-то здесь есть».

Сама по себе методика проста. Пользователю даётся задание, связанное с каким-либо отдельным диалоговым окном. Пользователь его выполняет. Через несколько минут пользователя просят нарисовать (пускай даже грубо и некрасиво) только что виденное им окно. После чего рисунок сравнивается с оригиналом.

Разумеется, пользователь запоминает только то, что ему кажется актуальным в процессе работы с окном (плюс еще что-нибудь за того, что ему показалось интересным, да и то не всегда). Это один из тех редких случаев, когда срабатывает ограничение на объем кратковременной памяти, так что количество запомнившихся элементов управления не может быть выше порога.

Пример проверки качества восприятия. Пользователь, которому нужно сменить шрифт абзаца на Arial из всего диалогового окна выбора шрифта в MS Word запоминает только часть элементов управления. Разумеется, он помнит, что помимо них были и другие, но точно вспомнить остальные элементы он, как правило, не может.

**Метода пяти секунд.** Суть метода пяти секунд заложена в его названии: вы предлагаете человеку в течение пяти секунд посмотреть на макет, затем убираете его и спрашиваете, что он увидел или запомнил. Этот тест позволяет узнать, какие элементы на странице больше всего «западают в душу» пользователям и на что они обращают внимание в первую очередь. Это лучше всего работает при тестировании рекламных баннеров и промоматериалов. Например, так можно сравнить эффективность рекламных мест на странице или протестировать содержание самих баннеров, спросив после пяти секунд: «О чем был баннер на макете?» или «Какой баннер вы запомнили?»

**Метод оценки восприятия дизайна.** Метод оценки восприятия дизайна позволяет понять вызывает ли дизайн целевые эмоции. Дизайнеру вместе с макетом интерфейса выдаётся список прилагательных, которые должны характеризовать дизайн сайта, то есть какие эмоции дизайн должен вызывать.

Впервые эту анкету предложили исследователи из Microsoft Usability Laboratory как очень быстрый способ оценки удовлетворенности. Анкета очень проста. Респонденту предъявляется лист бумаги с набором случайно подобранных прилагательных, одна половина которых скорее позитивна, вторая – негативна. Респонденту предлагается подчеркнуть слова, которые, на его взгляд, применимы к продукту. После того, как анкета заполнена, подсчитывается разница между числом негативных и позитивных терминов.

После чего, готовый интерфейс в нескольких вариантах цветового оформления предоставляется респондентам, и также даётся шкала с этими прилагательными, характеризующими общее впечатление, которое они должны оценить. В итоге, выбирается тот дизайн, который вызывает у пользователей нужные эмоции. Если по сумме ожиданий они оказываются примерно равными, выбирается по лучшим ожиданиям: ключевые плюс дополнительные.

См. в интернете: [Product Reaction Cards by Joey Benedek and Trish Miner](http://www.microsoft.com/usability/publications.htm).

Например. следующий набор прилагательных:

Устаревший – Эффективный – Нечеткий – Неудобный – Замусоренный – Тусклый – Яркий – Чистый – Прямой – Ясный – Непоследовательный – Неуправляемый – Привлекательный – Стандартный – Управляемый – Хороший – Интуитивный – Веселый – Любительский – Неэффективный – Опасный – Скучный – Радостный – Безопасный – Жесткий – Раздражающий – Треугольный – Неприятный – Комфортабельный – Холодный – Умный – Бесполезный – Халтурный – Теплый – Светлый – Последовательный – Загадочный – Качественный – Интересный – Ненадежный – Гибкий – Красивый – Некрасивый – Непривлекательный – Полезный – Глупый – Запутанный – Удобный – Понятный – Непредсказуемый – Четкий – Тяжелый – Современный – Легкий – Дружественный – Нестандартный – Плохой – Надежный – Сложный – Простой – Темный – Профессиональный – Медленный – Круглый – Печальный – Недружественный – Предсказуемый – Непонятный – Быстрый – Головоломный – Грустный – Приятный

Обратите внимание, что слова неслучайно даны вперемешку, именно так и нужно предъявлять их респондентам.

**Формальная анкета.** В отличие от анкеты, по словам, эта анкета не может быть использована без адаптации под конкретный проект. Часть ее вопросов порой неактуальна, порой нуждается в смене формулировки. В любом случае, для респондентов-женщин нужно менять половую принадлежность формулировок анкеты.

Анкета представляет собой несколько вопросов (табл. 1), для каждого из которых респондент может выбрать один из пяти вариантов ответа. Обратите внимание, что эта анкета проектировалась только как **пост-тестовая**, ее применение в другом качестве сомнительно.

Таблица 1

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Навигация по приложению удобная? | Нет |  |  |  |  |  | Да |
| Нужно ли внести изменения в функционал приложения | Нет |  |  |  |  |  | Да |
| Система работает достаточно быстро | Нет |  |  |  |  |  | Да |
| Дизайн приложения привлек внимание? | Нет |  |  |  |  |  | Да |
| Нужно ли увеличить функционал поиска информации? | Нет |  |  |  |  |  | Да |
| Легко ли настроить приложение под ваши нужды | Нет |  |  |  |  |  | Да |
| Легко ли было начать работу в приложении; не столкнулись ли вы с существенными трудностями? | Нет |  |  |  |  |  | Да |
| Могли ли вы с легкостью исправить некорректно введенные данные? | Нет |  |  |  |  |  | Да |
| Удовлетворила ли вас скорость работы приложения? | Нет |  |  |  |  |  | Да |
| Во время выполнения заданий чувствовали ли вы себя вполне уверенно | Нет |  |  |  |  |  | Да |
| Легко ли было понять, каким должен быть следующий шаг при выполнении задания? | Нет |  |  |  |  |  | Да |
| Скачали бы вы данное приложение на свой смартфон? | Нет |  |  |  |  |  | Да |
| Смогли бы вы посоветовать приложение своим знакомым? | Нет |  |  |  |  |  | Да |
| Стоит ли взимать плату за скачивание данного приложения? | Нет |  |  |  |  |  | Да |

Результаты нужно подсчитывать по следующему алгоритму:

* центральное значение дает ноль баллов,
* крайние значения дают либо –2 балла (левый вариант ответа),
* либо +2 балла (правый вариант),
* промежуточные значения либо –1 либо,
* +1 балл соответственно.

Сумма баллов является сравниваемым значением.

**Например:**

Критерий оценки участников:

|  |  |
| --- | --- |
| Кол-во баллов | Оценка |
| < 0 | Неуд. |
| > 0 и <10 | Удовлетворит. |
| >10 и < 15 | Хорошо |
| >15 и больше | Отлично |

# Результаты тестирования

|  |  |
| --- | --- |
| Пользователь | Результат |
| **Павел Ж.** | 11 - Хорошо |
| **Дмитрий З.** | 7- Удовлетворительно |
| **Ольга Ж.** | 12 - Хорошо |
| **Синяк К.** | 16 - Хорошо |
| **Рома Ч.** | 14 - Удовлетворительно |

Средняя степень удовлетворенности 12 – Хорошо. Отметка довольно средняя, значит, есть куда расти и есть что менять.

**Значимые эргономические метрики** (что именно измерять?)

Количество измеряемых показателей в конкретном тесте может быть довольно велико.

**Юзабилити** понимается как ***степень эффективности, успешности*** и ***удовлетворенности****,* с которыми продукт может быть использован определенными пользователями при определенном контексте использования для достижения определенных целей.

Главными показателями ***эффективности*** являются *скорость работы пользователя, скорость обучения и количество человеческих ошибок*.

* *Скорость работы пользователя.* Метрики: длительность выполнения операции; время, затраченное на обнаружение ошибок; время, затраченное на исправление ошибок; количество команд, исполняемых при выполнении операции (подразумевается, что чем больше команд, тем дольше их отдавать); длительность поиска сведений в документации; количество команд, более эффективных, чем использованные пользователем; снижение производительности при длительной работе.
* *Ошибки***.** Метрики: процент операций, вызвавших ошибку; среднее число ошибок на операцию у опытных пользователей (именно у опытных, т.к. у неопытных могут действовать и факторы из группы скорости обучения); количество ошибок, не обнаруженных и не исправленных пользователями.
* *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/количество ошибок пользователей перестает расти; разница в количестве ошибок/скорости работы у пользователей с опытом использования системы и без такого опыта.
* *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 длительного перерыва.

Данные, касающиеся производительности, такие (как частота ошибок и время выполнения заданий), оцениваются с помощью методов статистического анализа. Большая часть такого анализа сводится к нахождению среднего значения и стандартного (среднего квадратичного) отклонения, а также к проверке достоверности полученных различий.

Полученные количественные данные могут анализироваться также с помощью стандартных статистических процедур корреляционного и факторного анализа, реализуемых в пакетах SPSS, STATISTICA PRISM, широко используемых в выборочных исследованиях.

Для оценки эффективности существуют следующие формулы:

**Измерение эффективности на основе времени:**



**Расчет общей относительной эффективности:**

**,**

где:

***N*** – общее количество задач;

***R*** – количество пользователей;

***ni,j*** – результат задачи i-го пользователя. Если пользователь успешно завершил задачу, то ***ni,j=1,*** если нет, то ***ni,j=0.***

***ti,j –*** время, затраченное пользователем ***j*** на выполнение задачи ***i.*** Если задача не была успешно выполнена, то время измеряется до момента, когда пользователь покидает задачу.

***Успешность***, т.е. соотношение выполненных тестовых заданий, к невыполненным или выполненным полностью неправильно.

**Признаки успешности выполнения задачи.** Не всегда одно и то же задание можно выполнить единственным способом. Запускать же тест, не зная всех этих способов, некорректно, поскольку дальнейший анализ получится сомнительными. Предположим, респондент **А** выполнил задание способом **А**, а респондент **Б** способом **Б**. Оба респондента справились с заданием, но один все таки лучше другого. Как-никак разные способы, по-видимому, имеют разную эффективность, например, число действий, входящих в способ **Б** в полтора раза выше числа действий способа **А**. Способ **А** в такой ситуации предпочтителен, в идеальной системе (к которой нужно стремиться) все пользователи должны использовать только его.

Кроме того, порой правильный результат теста с точки зрения экспериментатора на самом деле не является правильным, особенно, если предметная область сложна, а юзабилити-специалист знает ее недостаточно. Чтобы убедиться, что правильный результат именно таков, как считает юзабилист, ему необходимо найти специалиста по системе и предметной области, и спросить. Не зная твердо всех способов выполнения задания, вы просто не сможете опознать ошибки.

***Субъективная удовлетворенность пользователя*.** Измерение этой характеристики сопряжено с определенными трудностями. Однако, после оценки восприятия дизайна с помощью набора прилагательных можно рассчитать процент удовлетворенности от продукта по следующей формуле:

,

где:



SА – процент удовлетворенности от продукта;

R – число респондентов (пользователей);

Aj+  – количество положительных ассоциаций (слов) от j-го респондента;

Aj-  – количество отрицательных ассоциаций (слов) от j-го респондента;

Для каждой задачи нужно выбрать свои значимые для нее характеристики интерфейса. С практической точки зрения гораздо удобнее более приземленные характеристики – простые декларативные метрики.

Вот примеры таких метрик:

* **Успешность** – респонденты правильно выполняют 90% заданий.
* **Эффективность** – скорость работы пользователя: регистрация на сайте выполняется меньше чем за 7 минут.
* **Эффективность** – ошибки: при вводе 10 форм количество ошибок ввода не превышает двух.
* **Эффективность** – обучаемость навыкам работы с системой: при выполнении задания **9**, отличающегося от задания **2** только вводимыми данными, респонденты не совершают ни одной ошибки (не считая опечаток).
* **Удовлетворенность** – по результатам анкетирования число баллов выросло на 20% по сравнению с прежними результатами.

**Тестирование**

Если вы ведете заметки, рекомендуется пользоваться следующими тактиками:

* **Скорость работы.** Между заданиями переводите секундомер на новый круг. Если респондент по какой-либо причине отвлекается, ставьте секундомер на паузу.
* **Ошибки.** На листе бумаги ставьте черточку при каждой человеческой ошибке. Удобно ставить мелкие черточки при мелких ошибках и длинные – при ошибках крупных. После теста достаточно посчитать количество черточек. Если вы раздельно считаете ошибки разных типов (например, просто ошибки и отдельно неправильно выбранные элементы меню) лучше пользоваться разными кодами, например, теми же черточками при простых ошибках и буквами М при ошибках, связанных с меню.
* **Проблемы, которые вы видите сразу же.** Кратко записывайте на бумажке сущность проблемы и текущее время (время первым!). Если вы будете точно знать, когда произошла проблема, вам будет легче найти соответствующий фрагмент видеозаписи.
* **Эмоциональные реакции респондента.** Ставьте знак плюса при положительных реакциях и знак минуса – при отрицательных. Реакции, происходящие в момент завершения тестовых заданий, не считаются.

**Анализ результатов**

Анализировать результаты можно как во время, так и после теста.

Оптимальной стратегией является начало анализа во время теста.

Почти любое юзабилити-тестирование направлено на поиск и выявление проблем.

**Анализ количественных данных.** Сами по себе собранные количественные данные практически не способны показать сущность проблемы, они показывают только количество проблем. С другой стороны, они могут быть использованы для сравнения старого и нового интерфейсов. Кроме того, сведенные вместе, они способны показать, где именно выявлены самые значительные проблемы – что полезно, т.к. начинать оптимизацию приятнее всего с областей, обещающих максимальную отдачу.

Четких рекомендаций по представлению количественных данных дать невозможно, поскольку для каждой метрики нужно свое представление. Поэтому необходимо:

* Для каждого типа данных создать отдельную таблицу, при этом в первую таблицу раздела складируйте данные по всем метрикам.
* Если можно нарисовать график, рисуйте график. Для заказчика график бесполезен (заказчики, как показывает опыт, вообще не интересуются не интерпретированными данными – и правильно делают), но зато он полезен вам, так как, рисуя его, вы сможете лучше понять, почему результаты именно таковы.
* Почти в любом тесте среди прочего измеряются успешность и количество ошибок. Для этих метрик используется таблица следующего вида:

Таблица 2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **1** | **2** | **3** | **4** | **5** | **Среднее** |
| **Виктор П.** | 3/4 | 1/13 | 1/4 | 0/4 | 1/2 | 1.2/5.4 |
| **Наталья Б.** | 2/5 | 0/4 | 0 | 2/12 | 1/3 | 1/4.8 |
| **Сергей К.** | 3/4 | 3/9 | 0/4 | 1/8 | 1/2 | 1.6/5.4 |
| **Елена Г.** | 3/5 | 2/9 | 2/1 | 4/19 | 2/4 | 2.6/7.6 |
| **Среднее** | 2.8/4.5 | 1.5/8.8 | 0.8/2.3 | 1.8/10.8 | 1.3/2.8 |  |

где по вертикали отложены тестовые задания, а по горизонтали – респонденты.

В каждой ячейке приведено количество человеческих ошибок (крупные и незначительные отдельно), цветами удобно отмечать проблематичные задания:

* одним цветом задания, от выполнения которых респонденты отказались
* другим – задания, которые респонденты не смогли выполнить полностью
* третьим – выполненные неправильно
* четвертым цветом – задания, выполнить которые не удалось по техническим причинам (читай – «сервер лежал»).

**Представление проблем**

Все выявленные проблемы нужно описывать в порядке убывания их важности. По возможности стоит также дать для каждой проблемы оценку ее деструктивности (в баллах). Без этого будет тяжело оценить, что требует исправления в первую очередь, а что можно оставить на потом.

Если выявленная проблема подтверждается собранными количественными данными, обязательно упоминайте это в описании проблемы. В любом случае, обязательно пишите, у каких респондентов она проявилась.

Вместо того, чтобы писать словами, на каких именно экранах обнаружены проблемы, используйте скриншоты, отмечая проблемные фрагменты именно на них. Если с вашей точки зрения одного скриншота достаточно, чтобы охарактеризовать проблему, ограничьтесь только им.

**3. Задание на лабораторную работу**

3.1. Описать возможные методики Ю-тестирования.

3.2..Произвести Юзабилити-тестирование с помощью имитации  
 поведения пользователей. Для этого:

* Сформировать список требований к респондентам и описать, выбранных вами респондентов.
* Составить тестовые задания, соответствующие пользовательским задачам. При формулировке заданий соблюдать все требованиям к ним.
* Обосновать выбранные методы тестирования; Описать методику эксперимента и условия тестирования;
* Провести тестирование.
* Определить признаки успешности выполнения задач (табл.2). Произвести оценку эффективности.
* Проверить степень удовлетворенности пользователей при работе с вашим интерфейсом с помощью метода оценки восприятия дизайна (анкеты по словам). Сформировать список вопросов для формальной анкеты адаптированный под ваш проект (табл.1) и провести тестирование. Рассчитать процент удовлетворенности.

3. 3. Проанализировать результаты тестирования и на их основе внести исправления в интерфейс.

3. 4. Представить скриншоты модифицированного интерфейса.

3. 5. Составить отчет.

**4. Требования к оформлению отчета**

Отчет должен содержать:

1. Титульный лист;
2. Название и цели работы;
3. Описание этапов выполнения работы в соответствии с заданием;
4. Полученные количественные данные;
5. Выводы по результатам тестирования;
6. Скриншоты интерфейса до и после модификации.

Рекомендуемая литература

* [A Practical Guide to Usability Testing](http://www.amazon.com/exec/obidos/tg/detail/-/1841500208/qid=1118763904/). Joseph S. Dumas, Janice C. Redish. Intellect, Ltd (UK), 1999. ISBN 1841500208.
* [Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests](http://www.amazon.com/exec/obidos/tg/detail/-/0471594032/). Jeffrey Rubin. Wiley, 1994. ISBN 0471594032.
* Описание юзабилити-методов: [Методы веб-дизайна и юзабилити](http://webmascon.com/topics/development/22a.asp)