

Глава 2

О нашем исследовании методом отслеживания движений глаз

Мы проводили исследования эргономичности, используя метод отслеживания движений глаз, как для наших клиентов, так для собственных научных целей. Но эта книга написана на основании целого ряда крупномасштабных исследований, прецедовавших четыре главные цели.

- Применить отслеживание движений глаз для проверки теоретических положений, касающихся эргономичного или неэргономичного веб-дизайна.
- Получить данные отслеживания движений глаз для новых изысканий в области эргономики.
- Определить информацию, которую можно собрать только методом отслеживания движений глаз, если это вообще возможно.
- Собрать сведения об апробированных методах и методиках отслеживания движений глаз для исследования эргономичности.

На наш взгляд, это были самые крупные исследования эргономичности веб-сайтов из тех, которые когда-либо проводились методом отслеживания движений глаз.

Для исследования самых разных веб-сайтов и многих задач, которые можно решать в веб, нам пришлось привлечь к эксперименту многих людей. Возможно, для изучения рынка несколько сотен людей — это не так уж и много, но для изучения их поведения такая выборка представляется довольно крупной. (Ведь самое элементарное качественное исследование удобства применения системы требует всего лишь пяти человек, чтобы выяснить проблемы ее эргономичности.) Мы наблюдали за тем, как каждый участник исследования пользовался веб-сайтом, чтобы выяснить, куда он смотрит и куда не смотрит, что было удобно в дизайне сайта и почему.

Кроме того, мы хотели выяснить, сколько посещений пользователей требуется для составления достоверной тепловой карты веб-страницы. И этого мы могли бы добиться только в том случае, если бы одними и теми же сайтами пользовались многие люди, решая на них различные задачи.

Собранные данные

В данном исследовании приняли участие более 300 человек. Мы поставили им 85 задач: одни — весьма конкретные, а другие — не связанные особыми условиями. В результате было зарегистрировано 1,5 млн. фиксаций взглядов. Зарегистрированные и обработанные данные в итоге составили массив объемом более 300 Гбайт.

Результаты исследований, приведенные в этой книге, основываются на 1,5 млн. фиксаций взглядов.

Мы планировали определять разные виды поведения участников исследования количественно, и поэтому нам было очень важно, чтобы все сеансы проводились одинаково. Ради согласованности данного исследования все эксперименты проводились двумя координаторами из компании Nielsen Norman Group по одной и той же методике наблюдения за участниками исследования в соответствии с заранее определенным конкретным протоколом. Лаборатория для этих экспериментов была устроена в центральной части Нью-Йорка — Манхэттене.

Участники исследования

Мы привлекли к исследованиям, главным образом, работающих людей со следующими демографическими характеристиками.

Возраст	18–29 лет: 20%
	30–49 лет: 64%
	50–64 лет: 16%
Опыт использования веб	Небольшой: 39%
	Средний: 29%
	Большой: 32%
Пол	Женский: 58%
	Мужской: 42%
Занятость	Полная: 81%
	Частичная: 15%
	Безработные: 1%
	Домохозяйки: 3%
	Пенсионеры: 0%
Семейный доход	Менее 20 тыс. долл. США: 3%
	20–50 тыс. долл. США: 32%
	Более 50 тыс. долл. США: 65%
Образование	Неполное среднее: 2%
	Среднее: 4%
	Неполное высшее: 25%
	Высшее и другое: 69%
Расовая принадлежность	Белые: 58%
	Черные: 21%
	Латиноамериканцы: 11%
	Прочие: 10%
Постоянное место жительства	Город: 83%
	Пригород и село: 17%

Приведенные выше демографические характеристики были выбраны нами потому, что они представляют пользователей, составляющих большую часть контингента наших клиентов. Ведь если бы мы привлекли людей с меньшим опытом использования веб, число нераскрытых проблем эргономичности могло бы возрасти. К данной категории можно было бы отнести людей с неполным образованием, полученным в учебном заведении, безработных, пенсионеров, нечасто пользовавшихся Интернетом в своей профессиональной деятельности, а возможно, и низкооплачиваемых работников, не применявших вычислительную технику на работе.

Для того чтобы данное исследование оказалось осуществимым, мы проводили все сеансы в одном месте. И благодаря тому что это было сделано в Нью-Йорке, нам удалось привлечь к исследованию людей самого разного уровня подготовки и происхождения. Но если бы данное исследование проводилось в других странах, особенно там, где принято писать и читать справа налево или по-другому реагировать на цвет и анимацию, некоторые результаты оказались бы, скорее всего, иными.

Сеансы качественных и количественных исследований

По традиции методы исследования удобства применения делятся на количественные и качественные. В ходе *количественных* исследований данные эргономичности определяются количественно, тогда как *качественные* исследования направлены на накопление более полных представлений и наблюдений. В наших изысканиях методом отслеживания движений глаз мы пользовались обоими видами исследований.

Например, количественное исследование может показать, что пользователям требуется в среднем 5,2 минуты, чтобы заказать 12 белых роз на сайте цветочного магазина, тогда как качественное исследование позволяет выяснить, что пользователи затрудняются указать дату доставки цветов из-за определенных просчетов в оформлении сайта. Оба вида исследований дают интересные результаты, но для организации успешной торговли цветами более ценным оказывается качественное исследование. Конечно, количественное исследование может дать некоторое расширенное представление об изучаемом явлении, но это представление вторично по отношению к полученным цифрам да и обнаруживается не так часто в силу особенностей проведения количественного исследования.

Отслеживание движений глаз вносит дополнительные трудности, поскольку данные могут быть интерпретированы двумя основными способами: получением совокупного поведения при просмотре среди большого количества пользователей для составления тепловой карты или анализом поведения отдельных пользователей при просмотре с помощью диаграмм взглядов и путем воспроизведения видеозаписей зарегистрированных взглядов.

Для составления надежной тепловой карты приходится усреднять поведение многих пользователей из-за больших различий в движении глаз отдельных пользователей. Как показывают наши исследования, для составления подходящей тепловой карты просмотра отдельной веб-страницы требуется отследить и зарегистрировать движения глаз не менее 30 пользователей. А это означает, что для проверки веб-сайтов на удобство применения необходимо привлечь около 39 участников, поскольку надежно зарегистрировать поведение всех участников эксперимента вряд ли удастся.

Как показывают наши исследования, для составления эффективной тепловой карты просмотра отдельной веб-страницы необходимо отследить и зарегистрировать движения глаз не менее 30 пользователей.

В нашем отчете, свободно доступном по адресу www.useit.com/eyetracking/methodology, приводятся сведения и рекомендации по методике исследования удобства применения с помощью отслеживания движений глаз.

Подробнее о том, сколько пользователей привлекается к исследованию методом отслеживания движений глаз, можно узнать из составленного нами отдельного отчета по методике отслеживания движений глаз. Его можно загрузить по адресу www.useit.com/eyetracking/methodology.

При сборе числовых данных во время сеансов, направленных на формирование тепловых карт, мы не просили участников эксперимента размышлять вслух, поскольку это могло повлиять на продолжительность просмотра ими отдельных элементов страницы. В ходе этих сеансов мы пользовались согласованными методами хронометрирования выполняемых заданий и предварительно составленными методиками оценки успешности их выполнения.

Для сбора данных о более полной качественной реакции мы иногда просили участников эксперимента размышлять вслух, чтобы пролить дополнительный свет на их поведение. Но дополнительные фиксации взгляда во

время размышлений вслух отрицательно сказываются на качестве данных отслеживания движений глаз, и поэтому мы не принимали эти случаи во внимание при подсчете фиксаций взгляда или использовании тепловых карт.

Испытательные сеансы

Каждый испытательный сеанс длился от полутора до двух часов. Пользователей просили сообщать о завершении задания или желании остановиться. Пользователи выполняли задания разными способами (в зависимости от конкретного задания): словесно, заполняя анкету или составляя реферат. Если же задание не ограничивалось особыми условиями, они просто говорили, что завершили его выполнение.

Все задания фиксировались прибором для отслеживания движений глаз, хронометрировались и оценивались координаторами. Иногда на видеокамеру снималось также лицо пользователя во время выполнения задания. Но система периодически выходила из строя, и поэтому мы не всегда прибегали к видеосъемке. Тем не менее в качестве резервного средства использовалась видеокамера, установленная на штативе.

Организация сеансов

Каждый пользовательский сеанс состоял из следующих действий, которые всегда выполнялись в одном и том же порядке.

- Приветствие и установка
- Заполнение формы на согласие участвовать в исследовании
- Заполнение анкеты личных интересов
- Калибровка прибора для отслеживания движений глаз

- Постановка и выполнение заданий
- Оценка наблюдаемого опыта использования веб
- Заполнение анкет после выполнения заданий, в которых пользователи оценивали свою удовлетворенность, неудовлетворенность и уверенность после каждого задания
- Ретроспективное исследование (проводилось редко), в ходе которого участникам исследования предлагалось просмотреть и прокомментировать отдельные части их пользовательских сеансов.

Веб-сайты и испытательные задания

Поставленные задачи представляли собой определенное сочетание заданий, связанных и не связанных особыми условиями. Это означает, что в одних случаях пользователи могли сами выбирать любые сайты безо всяких ограничений, а в других пользователей просили посетить определенный сайт и сделать на нем что-нибудь конкретное.

Система отслеживания движений глаз открывала проверяемый веб-сайт и еще один сайт или пустую страницу для пользователей. Когда пользователям предлагалось посетить какой угодно веб-сайт, мы иногда открывали пустой документ в виде начальной страницы, веб-сайт новостей или какой-нибудь другой сайт, не связанный в заданием, чтобы не оказывать влияния на их первоначальный выбор места посещения в веб.

Количественные задания

Это подборка заданий, проверившихся на большом числе пользователей по методике количественных испытаний.

Поиск и исследование в веб

- Поисковый механизм Google. Каким профессиональным видом спорта занимался Джордж Бретт?
- Поисковый механизм Google. Когда состоялся День Сурка в 2006 году?
- Поисковый механизм Yahoo. Какова максимальная скорость, с которой может плыть акула мако?
- Поисковый механизм Yahoo. Какие документы должен иметь при себе гражданин США, чтобы пересечь канадскую границу на автомашине?
- Поисковый механизм About.com. Если ваш приятель (или приятельница) страдает от изжоги, найдите для него информацию о том, как от нее избавиться.

Чтение

- Журнал *New York Magazine*. Найдите ресторан, в который вам хотелось бы пойти больше всего.
- Газета *New York Times*. Два австралийца получили Нобелевскую премию по медицине. Прочитайте о них статью, чтобы узнать подробнее о самых интересных аспектах их научного исследования.

- Журнал *New York Magazine*. Прочитайте статью “Will Wait for Food” (В ожидании заказанного обеда). Сообщите, какие именно предприятия общественного питания автор статьи считает достойными того, чтобы тратить время на ожидание заказанного в них обеда.
- Сайт Danceworksonline.co.uk. Ваш приятель (или приятельница) — большой поклонник Михаила Барышникова. Узнайте об этом танцовщике побольше, чтобы вам было о чем поговорить со своим приятелем.

Частный сектор

- Сайт [Agere](http://Agere.com). Найдите штаб-квартиру этой компании.
- Сайт [Adelphia](http://Adelphia.com). Ознакомьтесь с особенностями цифровой видеозаписи (DVR).
- Сайт [BNSF](http://BNSF.com). Прочитайте последние новости о железной дороге BNSF Railway.
- Сайт [BNSF](http://BNSF.com). Каковы корпоративные ценности железнодорожной компании BNSF?
- Сайт [JetBlue](http://JetBlue.com). На какой фондовой бирже котируются акции авиакомпании JetBlue?

Совершение покупок

- Сайт [JCPenney](http://JCPenney.com). Приобретите подарок якобы для новорожденной своего приятеля или приятельницы.
- Сайт [Kiehl's](http://Kiehl's.com). Приобретите что-нибудь для себя.
- Сайт [Headset Zone](http://HeadsetZone.com). Приобретите такие наушники, которые подошли бы вам лучше всего.
- Сайт [Sears](http://Sears.com). Приобретите увлажнитель для помещения площадью 19 м², потратив не более 125 долл. США.
- Закупки на конкретных сайтах в быстром темпе. Щелкните на конкретном товаре, как только его найдете.

Качественные задания

Это подборка качественных заданий, опробовавшихся на меньшем количестве пользователей, которым предлагалось размышлять вслух во время работы в Интернете.

- Кем был Фиорелло Ла Гардия?
- Вы собираетесь с друзьями покататься этой зимой на лыжах в штате Колорадо. Спланируйте свое путешествие.
- Какое знаменательное событие произошло в Гальвестоне, штат Техас, в 1900 году? Каковы его последствия?
- Завяжите беседочный узел. (Для этого мы давали испытуемым кусок веревки.)
- Допустим, что вы собираетесь вложить 10 тыс. долл. США в пенсионный план. Найдите наилучший способ вложения своих средств в такой план.
- Какая из двух пород собак больше всего вам подходит: каменистый терьер или фараонова гончая?

- Ваш приятель или приятельница боится подхватить птичий грипп. Найдите в Интернете информацию, которая успокоила бы его (или ее).
- Предположим, что вы собираетесь приобрести новый дом, чтобы вложить в него свои средства. Найдите ипотечную компанию, которая сможет предоставить вам наилучшие услуги и цены на недвижимость.
- Воспользуйтесь сайтом компании Sony, чтобы приобрести компактную цифровую фотокамеру стоимостью менее 450 долл. США.
- Допустим, что вы собираетесь приобрести цифровой проигрыватель музыкальных записей. Воспользуйтесь сайтом компании Sony, чтобы подобрать наиболее подходящую вам модель проигрывателя.
- Найдите ближайший к вам магазин с продукцией компании Sony.
- Приобретите себе (своему приятелю или приятельнице) что-нибудь подходящее на сайте mrchocolate.com.
- Приобретите новые футболки для себя или кого-нибудь еще на сайте panic.com/goods.
- Подберите себе рубашку на сайте neimanmarcus.com.
- Есть ли рядом с вами магазин Neiman Marcus?
- Допустим, что вы собираетесь приобрести телефон с услугами мобильной связи у оператора T-Mobile. Выберите наиболее подходящие для вас варианты.
- Убедитесь в том, что мобильная связь оператора T-Mobile надежно действует во всех местах, которые вы посещаете.
- Допустим, что вы подали декларацию о налогах пять недель назад. А теперь вы ожидаете частичного возврата суммы уплаченных налогов, но пока еще ничего не получили. Что вы собираетесь делать? Обратитесь, например, за справкой на сайт irs.gov.
- Вы пожертвовали 100 долл. США в Фонд послушных собак (Good Dog Foundation). Подлежит ли эта сумма налогообложению?
- Приобретите себе (своему приятелю или приятельнице) что-нибудь понравившееся на сайте mrcoffee.com.
- Предположим, что ваш приятель (или приятельница) разбил или иным образом испортил кофейник модели ECDT84. Выясните, можно ли приобрести новый кофейник той же модели, обратившись на сайт mrcoffee.com.
- Посетите сайт usatoday.com, чтобы найти для себя что-нибудь интересное.
- Посетите сайт youtube.com, чтобы найти для себя что-нибудь интересное.
- Посетите сайт vogue.com, чтобы найти для себя что-нибудь интересное.
- Что собой представляет Фонд современной моды (Vogue Fashion Fund)?
- Посетите сайт victoriasssecret.com, чтобы найти для себя что-нибудь интересное из предлагаемых товаров.

- Сделайте закупки бакалейных товаров на сайте freshdirect.com.
- Что такое Skype? Интересует ли вас это вообще? И почему интересует либо не интересует?
- Ныряет ли дикая утка за пищей? И если ныряет, то как долго она обычно остается под водой?
- Поищите для себя телевизор с плоским экраном, воспользовавшись сайтом компании Panasonic.
- Закажите в авиакомпании United Airlines для себя и своего приятеля или приятельницы билеты до Сан-Франциско и обратно, чтобы провести в этом городе неделю в июне или июле.
- Какие команды являются лидерами чемпионата национальной ассоциации университетского баскетбола? Обратитесь за справкой на сайт периодического издания *Sports Illustrated* (si.com).
- Каким образом ветроэнергетическая установка генерирует электроэнергию?
- Какие системы для видеоигр рекомендуются журналом *GQ*?

Измерения

Как и в ряде других наших исследований, мы проводили измерения и получали большинство результатов данного исследования качественным методом, наблюдая за поведением пользователей и пытаясь выяснить, что именно они собираются делать. В данном исследовании мы сосредоточили также свое внимание на следующих моментах: на что именно смотрели испытуемые, как долго они смотрели на отдельные элементы оформления и на что они вообще не смотрели. И наконец, мы постарались выяснить, почему они смотрели или же не смотрели на определенные элементы оформления. И в меньшей степени мы прислушивались к комментариям пользователей как во время выполнения заданий, так и в промежутках между ними.

Помимо качественных измерений, мы пользовались четырьмя основными способами количественных измерений, которые нередко применяются в исследованиях удобства применения: хронометрированием времени выполнения задания, оценкой успешности выполнения задания, выявлением ошибок и выяснением субъективной удовлетворенности испытуемых результатами. Кроме того, мы приняли во внимание количество фиксации взглядов испытуемых на различных стадиях выполнения заданий, а также их полезность или бесполезность.

Субъективная удовлетворенность

Собирая результаты объективных измерений, не менее важно узнать мнение самих пользователей по поводу того задания, которое им пришлось выполнять. Это помогает выяснить, насколько оказались удовлетворенными (полностью или частично) пользователи. Ведь иногда пользователь может завершить задание довольно быстро, но остаться неудовлетворенным, или же потратить на него много времени, не особенно беспокоясь о конечном результате.

Итак, мы попросили испытуемых ответить на три вопроса сразу же после выполнения задания, но прежде, чем предложить им прокомментировать задание или же ответить на конкретные вопросы (по крайней мере, так было в наших количественных исследованиях с участием большого числа пользователей). Эти вопросы обычно были следующими.

- Насколько трудно было выполнить задание?

Очень просто

Очень трудно

1 2 3 4 5 6 7

Ответ на этот вопрос позволяет выяснить, насколько сложным оказалось задание, если оно вообще было таким.

- Какую удовлетворенность или неудовлетворенность доставила работа над этим заданием?

Большая удовлетворенность

Большая неудовлетворенность

1 2 3 4 5 6 7

На первый взгляд, этот вопрос об удовлетворенности заданием позволяет выяснить то же самое, что и первый вопрос. Но мы нередко получали совершенно разные ответы на него, поскольку он на самом деле связан в большей степени с содержимым и набором доступных средств, чем с удобством использования. Например, одному из участников исследования удалось найти довольно интересную информацию о документах, которые нужно иметь с собой при пересечении американо-канадской границы на автомашине. Но это задание могло оказаться для него трудным, и поэтому он обвел кружком категорию трудности 6, отвечая на первый вопрос. А поскольку этот участник исследования выяснил ряд интересных особенностей, то он оказался в достаточной степени удовлетворенным своей работой, обведя кружком категорию удовлетворенности 3.

- Насколько вы уверены или не уверены в том, что выполнили задание полностью (или дали правильный ответ)?

Полная уверенность

Полная неуверенность

1 2 3 4 5 6 7

Этот вопрос об уверенности в результатах позволяет пролить свет на те ситуации, в которых пользователи выполнили задание, но не уверены, что дали правильный ответ. Или же они довели задание до логического завершения, но не осознали или не поверили в то, что сделали все, что от них требовалось.

Например, на одном или нескольких веб-сайтах могли быть предоставлены противоречивые сведения, и поэтому один из участников исследования не знал, какому из этих источников следует доверять. Если во время исследования он проявит слабое доверие к источнику информации, то в реальной ситуации он, скорее всего, поищет столь важную для него информацию в других источниках или же обратится к кому-нибудь за помощью. С другой стороны, вводящая в заблуждение или неверная информация может быть подана на веб-сайте в весьма доверительной форме, чтобы

Вместо устных ответов на подобные вопросы мы даем участникам исследования ручку, чтобы они обвели кружком выбранные ими оценки. Благодаря этому мы можем проверить правильность ответов на любые интересующие нас вопросы, поскольку участники исследования иногда путают положительный конец шкалы оценок с отрицательным. Кроме того, наличие ответов на бумаге дает возможность легко ввести данные в электронные таблицы впоследствии, а во время исследования работать оперативнее, ставя больше задач перед участниками исследования.

Ответы каждого участника исследования мы усредняли, чтобы получить общую оценку субъективной удовлетворенности результатами по каждому заданию. Для этого мы проводили анкетные опросы после каждого количественного задания.

Хронометрирование

Для изменения продолжительности выполнения конкретного задания мы запускали часы или таймер, как только участник исследования прочитал (или услышал) задание, уяснил его и обратился к системе. В лаборатории для исследования веб-сайтов это был момент, когда участник клал кисть руки на мышь и направлял взгляд на экран. Если же выполнение задания прерывалось, например, из-за аварийного сбоя в системе, мы останавливали таймер. То же самое мы делали, когда участник исследования сообщал о том, что задание выполнено. Мы не останавливали таймер, если участник завершал задание, но не сообщал об этом, поскольку иногда он не был уверен в том, что задание выполнено, и поэтому хотел продолжить работу.

Принимая во внимание значение конечного результата выполнения задания, перед началом сеанса очень важно пояснить участникам исследования, что они должны непременно объявлять о завершении задания, иначе эксперимент будет продолжаться. По завершении первого задания некоторым участникам исследования приходилось об этом напоминать. Именно поэтому мы посчитали целесообразным дать участникам исследования сначала возможность поупражняться на одном задании в ходе этого или других количественных исследований. Но мы не предупреждали их о том, что это всего лишь упражнение, а просто считали его таковым.

Для хронометрирования сеансов мы пользовались обычным спортивным секундомером. Нас могут упрекнуть за это, поскольку имеются и более совершенные способы хронометрирования. Но во время сеанса, особенно с отслеживанием движений глаз, координаторы должны следить за столькими аспектами, определяющими правильный ход эксперимента, что им проще иметь под рукой обычный секундомер. Единственный недостаток такого подхода заключается в том, что большинство современных электронных секундомеров издадут звуковой сигнал, когда их останавливают и запускают. Это, конечно, удобно для того, кто проводит хронометрирование, но мешает участникам исследования расслабиться и забыть о том, что они находятся в исследовательской лаборатории. Правда, наша коллега Эми Шейд и ее муж разобрали электронный секундомер и удалили на его задней стенке алюминиевую фольгу, издававшую звуковой сигнал.

Успешность выполнения

Успешность выполнения определялась тем, насколько участник исследования доводил задание до конца. Если он сообщил об окончании задания, то действительно ли он его завершил? А если завершил, то насколько? Мы не оценивали задания как выполненные или не выполненные, поскольку это привело бы к необъективной, более заниженной оценке дизайна веб-сайтов, чем он есть на самом деле.

Выявление ошибок

Мы следили за тем, сколько раз участники исследования выбирали неверный элемент интерфейса. Например, участник исследования мог выбрать кнопку *просмотра* содержимого тележки для покупок, неверно считая, что он выбрал кнопку *ввода* товаров в тележку для покупок. Или же он мог щелкнуть на ссылке *Связи с инвесторами*, чтобы найти пресс-релизы, когда информация для прессы на самом деле располагалась по ссылке *Средства массовой информации*. А если участник исследования работал во внутренней корпоративной сети, то он мог набрать *Ф.И.О.* своего коллеги для поиска его телефона на сайте этой сети, тогда как в действительности ему нужно было бы найти эту информацию в каталоге сотрудников.

Для точной фиксации ошибок координаторы испытаний должны действовать очень быстро и хорошо знать структуру веб-сайта или в крайнем случае просматривать видеозаписи испытаний после их проведения. Разделение ошибок на отдельные категории само по себе чревато ошибками, главным образом потому что перекрестные ссылки и замысловатая информационная архитектура приводят к появлению слишком многих путей доступа к одним и тем же страницам и соответствующей информации на веб-сайтах. Например, пресс-релиз может быть доступен по ссылке *Средства массовой информации*, а сама эта ссылка — в разделе *Связи с инвесторами* проверяемого сайта. В связи с этим возникает вопрос, совершил ли участник исследования ошибку, выбрав на этом сайте раздел *Связи с инвесторами*? Ответить на него нелегко, как, впрочем, и на вопросы, связанные с выявлением ложных ориентиров.

Ложные ориентиры

Не менее важно упомянуть и о *ложных ориентирах* — элементах, несвоевременно привлекающих внимание пользователя. Но учесть каждый ложный ориентир чрезвычайно сложно, если вообще возможно. Ведь если человек смотрит на что-нибудь, то еще неизвестно, по какой именно причине он делает это: из интереса или неумышленно. Для того чтобы определить по-настоящему ложные ориентиры, встречавшиеся во время наших исследований, нам пришлось просмотреть и проанализировать немало пользовательских сеансов.

Распознавать ложные ориентиры намного труднее, чем ошибки. А поскольку эта процедура отнимает немало времени и чревата ошибками, мы рекомендуем ее только для хорошо подготовленных групп, исследующих удобство использования веб-сайтов. А многим группам разработчиков и дизайнеров пока еще не силам углубленное исследование и анализ реакции пользователей на ложные ориентиры. Большинство таких групп просто не готовы к этому, поскольку они по-прежнему изучают элементарные

выводы в отношении эргономичности и пытаются реагировать на них без учета отслеживания движений глаз. Поэтому выявлением и устранением ложных ориентиров должны заниматься только специально подготовленные группы разработчиков, которые уже обнаружили и устранили более серьезные недостатки, связанные с удобством использования, в том числе неверные щелчки, т.е. неверный выбор ссылок или кнопок с помощью мыши, или безуспешно выполняемые задания.

Почему многие исследования методом отслеживания движений глаз недостоверны

Мы не первые, кто проводил исследования, применяя отслеживание движений глаз для изучения удобства использования веб-сайтов. Но нам впервые удалось организовать широкомасштабное исследование методом отслеживания движений глаз, применяя достоверную методику изучения эргономичности для регистрации взглядов репрезентативной группы пользователей, выполнявших реальные задания на самых разных веб-сайтах. Данное исследование отличается тремя важными критериями.

- Репрезентативная группа пользователей
- Реалистичность выполняемых заданий
- Большое разнообразие исследуемых веб-сайтов

Мы считаем, что все три перечисленных выше критерия необходимы для того, чтобы прийти к правильным выводам, которые можно обобщить и сделать полезными в качестве руководящих принципов для разработки новых веб-сайтов и изменения уже существующих.

Репрезентативная группа пользователей

Научные исследования, проводимые в академической среде с привлечением студентов в качестве испытуемых, непригодны для изучения удобства оформления веб-сайтов частных предприятий, электронной коммерции, государственных ведомств и других видов сайтов, рассчитанных главным образом на взрослых пользователей и профессионалов в деловой сфере. Но даже если сайт и рассчитан на юных и молодых пользователей, полагаться на результаты научных исследований в академической среде все равно нельзя, поскольку они основываются на психологии студентов элитных университетов как участников исследования, разве что именно эта аудитория является целевой для сайта.

Реалистичность выполняемых заданий

Академические исследования также слабо связаны с проектами коммерческих разработок, поскольку в них не предполагается выполнение реальных заданий делового характера. Удобство применения всегда очень сильно зависит от контекста, а удобство использования веб-сайтов в еще большей степени зависит от содержа-

тельности информационного окружения, чем в остальных областях применения эргономики, например в конструкции кабины самолета. Когда люди пользуются веб-сайтом, им обычно требуется перемещаться в гиперпространстве по десяткам тысяч страниц, и они хорошо осознают, что в этом пространстве могут существовать сотни конкурирующих сайтов, которые в равной степени доступны для посещения. И напротив, в академических исследованиях обычно испытывается весьма узко истолковываемое информационное окружение, а управление переменными этого окружения осуществляется довольно строго, чтобы получить теоретические выводы, необходимые для публикации научного труда. Такой подход к эксперименту пригоден для фундаментальных исследований психологии восприятия. Но если требуется исследовать межкорпоративный сайт, то строго контролируемый эксперимент не даст результатов, отражающих реальный объем продаж, даже если привлечь к исследованию заведующих отделами закупок, а не студентов.

Но даже исследования методом отслеживания движений глаз в коммерческой сфере обычно проводятся по слабо разработанной методике, что вызывает недоверие к результатам таких исследований. Характерная ошибка подобных исследований состоит в том, что они проводятся без постановки перед пользователями реальных задач. Ведь люди совершенно по-разному просматривают веб-страницу в зависимости от того, стремятся они достичь конечной цели в каком-то другом месте или же проверить саму страницу. В ходе многих исследований собираются бесполезные данные, когда участников просят просмотреть конкретную веб-страницу и просто высказать исследователям свое мнение о ней. Но ведь это совершенно нереальное задание, поскольку участники исследования обычно не посещают веб-сайт, чтобы проанализировать его оформление. Это интересно лишь самим дизайнерам. Но когда люди просматривают веб-страницу с целью ее анализа, они обращают внимание на элементы оформления более пристально и систематически, чем в обычной ситуации блуждания по веб.

Когда исследователи ставят нереальную задачу дать комментарии к просмотренной странице, они регистрируют в конечном итоге намного больше фиксаций взглядов на таких элементах, как рекламные объявления, чем это характерно для реального поведения пользователей. Они также обнаруживают, что пользователи читают текст намного внимательнее, чем это при реальном применении веб-сайта.

Даже постановка реальных задач может привести к недостоверным результатам исследования, если условия испытаний слишком ограничены. Допустим, что требуется выяснить, как покупатели вводят информацию в поле адреса на информационной странице оформления покупок. Для получения точной тепловой карты такой страницы можно поставить перед тридцатью участниками исследования конкретную задачу (ввести свой адрес), а затем попытаться зарегистрировать, каким образом они просматривают заполняемую форму на данной странице как часть достаточно реального задания приобрести что-нибудь для своего дома.

Фактически это приведет к формированию недостоверной тепловой карты, а выводы из ее анализа вряд ли помогут улучшить оформление сайта электронной коммерции, предназначенного помочь посетителям в совершении процесса покупок.

Дело в том, что при такой постановке задачи участникам исследования предлагается перейти на конкретную страницу в отрыве от реального контекста, в котором они обычно доходят до этой страницы в процессе совершения покупок. На отдельной странице они ведут себя совершенно иначе, чем по ходу выполнения последовательности действий, приводящих к этой странице.

В таком случае, по-видимому, имеет смысл предложить участникам исследования совершить весь процесс оформления покупок и зарегистрировать движения их глаз на всех соответствующих страницах, даже если главным образом нас интересует, как они заполняют поле адреса. Конечно, результат такого исследования будет лучше, но он по-прежнему останется не совсем достоверным. Поэтому оптимальным можно считать такое исследование, в ходе которого участникам предлагается совершить покупку на сайте, начиная с поиска подходящего товара и заканчивая его вводом в тележку для покупок. И только в этом случае, попросив их оформить покупку, имеет смысл наблюдать, как они заполняют поле адреса, поскольку их поведение на данной странице будет основываться на их опыте взаимодействия с предыдущими страницами, которые встретились им в течение всего процесса закупки.

Между исследованиями методом отслеживания движений глаз и реальным поведением пользователей все же существует некоторое разногласие. Отслеживание движений глаз регистрирует поведение на отдельной странице, но главным признаком эргономичности веб-сайта служит удобство перемещения по его страницам. Для устранения этого разногласия исследование методом отслеживания движений глаз необходимо проводить таким образом, чтобы его участники перемещались по веб-сайтам как обычно, когда они обращаются к веб для выполнения реального задания. И только в этом случае можно получить достоверные данные о том, как люди смотрят на каждую встречающуюся им страницу.

К сожалению, многие исследования методом отслеживания движений глаз проводятся по следующей методике: участникам исследования сначала показывают одну или две страницы, а затем регистрируют, как они смотрят на них. Но сам показ конкретной страницы уже оказывает влияние на то, как участники исследования на нее смотрят. Ведь когда они заранее знают, что координаторов исследования интересует конкретная страница, они просматривают ее намного тщательнее, чем если бы это была одна из тех страниц, которые встретились им при перемещении по сайту.

Когда участники исследования переходят на искомую страницу, это событие не сопровождается в веб малиновым перезвоном колоколов. В ходе бесчисленных сеансов мы наблюдали, как пользователи достигают нужной страницы только затем, чтобы покинуть ее через несколько секунд, поскольку она выглядела не так, как им хотелось бы. При обычном блуждании по веб людям приходится быстро оценивать страницы и вовремя прекращать свой поиск, если им встретится неподходящая (по крайней мере, они так считают) страница или сайт. И только решив, что страница представляет для них особый интерес, они приступят к ее подробному изучению. Ведь если бы люди подробно изучали все страницы в веб, они никогда бы не достигли желаемого результата, и это им хорошо известно. Поэтому если исследователи сразу же переносят участников своего исследования на конкретную страницу вместо

того, чтобы предоставить им возможность перемещаться по сайту свободно, то они фактически меняют поведение людей (и делают недостоверными результаты своих исследований).

Большое разнообразие исследуемых веб-сайтов и выполняемых заданий

Третий критерий, по которому можно судить о достоверности исследований методом отслеживания движений глаз, заключается в большом разнообразии и широте охвата исследуемых веб-сайтов. В испытании отдельного веб-сайта нет на самом деле ничего дурного, и этим мы часто занимаемся в большинстве своих консультационных проектов. Но результаты исследования одного сайта нельзя обобщить на многие другие сайты.

Результаты исследования удобства использования одного веб-сайта, проведенного с привлечением целевой аудитории, ничего не могут сообщить о том, насколько они характерны для данного конкретного сайта и его аудитории. Возможно, они и носят более общий характер и дают представление об удобстве использования веб-сайтов вообще, но, имея в своем распоряжении только эти результаты, нельзя быть уверенным ни в чем.

Именно поэтому так важно исследовать множество сайтов и их аудиторий. И лишь выявив в поведении пользователей нечто общее, повторяющееся на многих сайтах, можно быть уверенным в том, что обнаружены определенные руководящие принципы эргономичности, которые могут применяться и к тем сайтам, которые не проверялись в ходе исследования.

Оценка фактов

Силами компании Nielsen Norman Group мы испытали более 1250 веб-сайтов, для чего привлекли 3600 пользователей 16 национальностей с четырех континентов. Безусловно, наше исследование методом отслеживания движений глаз оказалось не таким крупным, как проведенное нами обобщенное пользовательское испытание, в частности потому, что мы собирали данные отслеживания движений глаз только от пользователей одной нации. Но интерпретируя результаты нашего исследования методом отслеживания движений глаз, мы могли опираться на опыт тысяч других пользовательских сеансов, которые мы наблюдали, и сравнить их с задокументированными видами поведения по отношению к самому большому разнообразию типов веб-дизайна.

Потребность в обобщении множества веб-сайтов и их пользователей представляет собой еще одну трудность для практического применения большинства академических исследований и даже презентаций на торгово-промышленных выставках и отраслевых конференциях. Академические исследования практически всегда направлены на необычайно узкий ряд типов веб-дизайна — зачастую один или два веб-сайта или же две разновидности одного и того же сайта. А отраслевые презентации и того уже, поскольку на них докладчик обычно представляет собственный проект.

Когда он рассказывает о том, что именно было обнаружено в результате исследования данной страницы методом отслеживания движений глаз и как затем был усовершенствован сайт его компании, это, как правило, вызывает живой интерес у аудитории. Но компании нередко заявляют о больших успехах в оформлении своих веб-сайтов, не особенно акцентируя внимание на том факте, что их сайты на самом деле призваны приносить прибыль.

Многие отчеты об исследованиях методом отслеживания движений глаз, с которыми нам пришлось ознакомиться, оказались недостаточно убедительными, поскольку в них нарушалось одно (а нередко и все три) требование к составлению обоснованно правильных представлений. Впрочем, из большого числа исследований все же можно составить некоторые представления. Если каждое исследование будет проведено с привлечением репрезентативной группы пользователей, выполняющих реальные задания, то результаты такого исследования, скорее всего, окажутся достоверными, несмотря на то что они применимы только к исследованному веб-сайту. А если объединить результаты многих узконаправленных исследований, то это будет равнозначно проведению одного обширного исследования путем обобщения результатов каждого узкого исследования в отдельности.

По существу, это означает оценку фактов. Так, если проанализировать 20 отчетов об исследовании 20 сайтов и в 19 из них обнаружатся более или менее одинаковые результаты, то можно считать, что выявлен определенный действительно руководящий принцип эргономичности веб-дизайна. Не следует, однако, забывать, что руководящие принципы — это всего лишь принципы или эмпирические правила. Они применимы далеко не к каждому дизайну. А поскольку дизайн взаимодействия сильно зависит от контекста, то особые обстоятельства могут потребовать отклонения дизайнерского проекта от вполне достоверного в остальном руководящего принципа эргономичности.

Если в рассматриваемом здесь гипотетическом примере 19 исследований дали сходные результаты, а одно — противоположные, то надежнее всего ориентироваться на большинство исследований. Ведь вполне возможно, что исследование, давшее результаты, отклоняющиеся от общей тенденции, было проведено по не совсем удачно разработанной методике, и такое случается слишком часто. Но даже если бы эта методика и оказалась надежной, результаты данного исследования могли быть получены по счастливому стечению обстоятельств, т.е. совершенно случайно. Это особенно характерно для количественных исследований, выводы из которых делаются на основании статистического анализа. Как правило, критерий статистической значимости равен $p < 0,05$, а это означает, что вероятность получения неверного результата из-за случайных факторов составляет 5%. Но если проведено 20 исследований, то 5%-ная вероятность получения неверного результата фактически сводится к одному ошибочному исследованию. И если неизвестно, какое именно из этих исследований ошибочное, то доверять нельзя ни одному из них.

В конечном счете, даже если методика и окажется надежной, а результаты — неискаженными случайной ошибкой, результаты исследования, отклоняющегося от общей тенденции, следует в рассматриваемом здесь примере отклонить. Вполне

Выясняя достоверность результатов исследований, проводимых методом отслеживания движений глаз, следует непременно руководствоваться тремя следующими критериями: представительность пользовательской аудитории, реальность выполняемых заданий и разнообразие исследуемых веб-сайтов.

возможно, что данное конкретное исследование было проведено для изучения особых видов дизайна, которые, естественно, отличаются от подавляющего большинства некоторыми особенностями своего контекста.

Существует немало причин для недостоверности исследований, проводимых методом отслеживания движений глаз. Провести такое исследование и проанализировать его результаты правильно намного труднее, чем сделать это традиционными методами исследования удобства использования. И если принять во внимание большое количество проведенных исследований, то нет ничего удивительного в том, что они дали немало сомнительных результатов.

Стоимость исследований, проводимых методом отслеживания движений глаз

В настоящее время общая стоимость исследований, проводимых методом отслеживания движений глаз, остается относительно высокой. В нее входит стоимость аренды или приобретения прибора для отслеживания движений глаз, а также подготовки и проведения качественного исследования. Поэтому, прежде чем приступить к такому исследованию, рекомендуется провести ряд пробных испытаний, чтобы проверить согласованность протокола исследований с применяемой системой и тем самым избежать неожиданностей, когда начнется настоящее исследование, а также проанализировать результаты подобных испытаний.

Ниже приведена стоимость проведенного нами исследования. Следует, однако, иметь в виду, что мы проводили очень крупное и дорогостоящее исследование. А поскольку многие затраты на исследование методом отслеживания движений глаз меняются в зависимости от количества сеансов, участвующих пользователей, объемов данных и средств их хранения, то более мелкое по масштабу исследование обойдется дешевле. Самые крупные фиксированные затраты, которые не изменяются в зависимости от масштаба исследования, связаны с приобретением или арендой оборудования для отслеживания движений глаз, а также с арендой помещения. Но для организаций, в которых имеется достаточно места для проведения подобного исследования и уже приобретен прибор для отслеживания движений глаз, эти конкретные затраты могут оказаться незначительными.

Общие затраты на исследование

В приведенной ниже таблице сведены общие затраты на исследование, исключая временную оплату услуг специалистов по эргономике и аналитиков.

Статья затрат	Стоимость, долларов США
Аренда на пять месяцев помещения в районе Манхэттена, Нью-Йорк	11000
Аренда на пять месяцев оборудования для отслеживания движений глаз	10440
Наем рабочей силы, в том числе 2 тыс. долл. США на рекламу	18000
Денежное вознаграждение участникам исследования	17900
Расходы на выполнение заданий в области электронной коммерции	1790
Расходы на жесткие диски для хранения данных	2000
Итого	61090

Расходы по найму рабочей силы

В приведенной выше таблице расходы по найму рабочей силы выглядят весьма ощутимыми. Сверх этого, нам пришлось отсеять во время найма до 5% людей как непригодных для исследования в связи с недостатками их зрения. Подобные затраты трудно разделить, как, впрочем, и прослушивание записанных на автоответчик звонков.

Нам удалось сократить расходы по найму рабочей силы до 50 долл. США на каждого нанимаемого человека благодаря тому, что наем проводился людьми, работающими в нашей организации. Конечно, обращение к услугам внешних организаций, специализирующихся на найме рабочей силы, позволяет сэкономить время, но в то же время такие услуги обычно обходятся больше в среднем на 100–200 долл. США на каждого нанимаемого человека.

В связи со столь крупными масштабами исследований система иногда выходила из строя и была не готова к проведению следующего по плану сеанса. В подобных случаях нам приходилось отказываться от услуг участников исследований. Помимо упоминавшихся выше расходов по найму рабочей силы, мы потратили еще несколько тысяч долларов на ее наем в связи с устранением неисправностей в технике отслеживания движений глаз.

Причины, по которым запланированные сеансы не состоялись	Расходы по найму рабочей силы
Неявка участников исследований	49 человек: 2450 долл. США
Отправка участников исследований домой в связи с неисправностями в технике отслеживания движений глаз	18 человек: 900 + 1800 = 2700 долл. США
Итого	5150 долл. США

Помимо дополнительных расходов по найму рабочей силы в связи с устранением неисправностей в технике отслеживания движений глаз, мы дополнительно заплатили тем участникам исследований, особенности строения глаз которых не позволяли откалибровать систему отслеживания движений глаз и нормально работать с ними. Например, большинство очков без специальных линз пригодны для работы с прибором для отслеживания движений глаз. Если оправа очков слишком толстая,

то это иногда нарушает калибровку прибора. А если сетчатка глаз пользователя почти не отличается по цвету от его зрачка, то иногда прибор для отслеживания движений глаз не в состоянии обнаружить зрачок, чтобы следить за ним. Такие люди, прошедшие отбор, получали дополнительное вознаграждение в размере 100 долл. США, если они являлись в лабораторию вовремя, несмотря на то что не могли принимать участие в исследованиях методом отслеживания движений глаз из-за осложнений при калибровке оборудования по их глазам.

Обычно доля людей, не являвшихся на запланированные сеансы, составляла около 20%, когда их в последний момент отпугивали некоторые особенности эксперимента, например необходимость пользоваться своей кредитной картой в ходе исследования. Это типичная норма для наших исследований. А кроме того, нам пришлось отменить или отложить 7% сеансов в связи с устранением неполадок в технике отслеживания движений глаз.

Затраты на потерю данных отслеживания движений глаз и регистрации

В связи с тем, что система иногда выходила из строя, данные некоторых заданий терялись полностью. Тем не менее мы продолжали свои наблюдения и делали заметки в тех сеансах, где это случалось. А кроме того, мы делали резервные видеозаписи, хотя и не могли обратиться к данным отслеживания движений глаз после подобных сеансов. Поэтому нам пришлось привлекать к эксперименту больше пользователей, чтобы их количество было достаточно представительным для нашего исследования.

Кроме того, мы отбрасывали или не собирали данные отслеживания движений глаз, когда оборудование недостаточно хорошо калибровалось по глазам пользователей.

Анализ причин потери данных при выполнении задания	Потери данных при выполнении задания, процентов
Плохая калибровка	15,83
Сбои в работе техники для отслеживания движений глаз	7,19
Итого	23,02

Затраты, связанные с отслеживанием движений глаз

С любым исследованием удобства использования связаны определенные затраты, особенно если в нем участвует более 50 пользователей. Помимо этих затрат, с данным исследованием были связаны особые затраты на отслеживание движений глаз. В представленной ниже таблице приведены только те затраты, которые связаны с частью данного исследования, относящейся к отслеживанию движений глаз. (К ним не относятся расходы на повременную оплату услуг специалистов по эргономике.)

Статья затрат, связанных с отслеживанием движений глаз	Затраты, долларов США
Отправка пользователей домой в связи с неисправностями в технике для отслеживания движений глаз	2700
Привлечение дополнительных пользователей для восполнения потерянных данных (наем, денежные вознаграждения)	9450
Аренда на пять месяцев оборудования для отслеживания движений глаз	10440
Приобретение жестких дисков для хранения данных	2000
Итого	24590 (43% от общих затрат)

Оборудование

В ходе исследования мы пользовались прибором для отслеживания движений глаз типа Tobii 1750 EyeTracker и высокопроизводительным ПК, подключенным к Интернету. Разрешение экрана было установлено равным 1024×768 с 16-разрядной глубиной цвета. (Эти параметры экрана рекомендованы компанией Tobii, и мы сами обнаружили, что более высокие разрешение и глубина цвета требуют слишком больших вычислительных ресурсов.)

На отдельном мониторе мы могли наблюдать (и регистрировать для анализа) точку на экране, которая оказывалась в центре фовиального зрения пользователей. А периферическое зрение система не регистрировала.

Оборудование компании Tobii работало с частотой кадров 50 Гц. В нем применялся метод отслеживания движений глаз с обнаружением положения центра зрачка и роговического блика (PCCR — Pupil Center Corneal Reflection), при котором едва воспринимаемый свет оставляет блик на глазу пользователя, а камера улавливает его и регистрирует.

Система приспособлена к устранению двух недостатков: отклонения и ошибки компенсации движений головы. Каждый из них может стать причиной слабой фиксации взгляда.

- **Отклонение** — это ухудшение калибровки по глазам участника эксперимента. Оно может произойти вследствие изменений в окружающей обстановке испытательной лаборатории, например условий освещения или влажности. Мы следили за этим и повторяли такую калибровку по мере надобности.
- **Ошибка компенсации движений головы** может происходить, если пользователь двигает головой. Поле зрения камеры составляет около 20×15×20 см по ширине, высоте и глубине на расстоянии 60 см от экрана. Этого достаточно для компенсации положений головы, удобных для того, чтобы сидеть в нормальной позе перед экраном компьютерного монитора. Если же пользователь резко двигался, начинал сутулиться или сильно наклоняться вперед либо отклоняться назад, мы исправляли положение и повторяли калибровку по мере надобности. Подробнее о технологии и оборудовании компании Tobii можно узнать по адресу www.tobii.com.