

## ОСОБЕННОСТИ ПРОВЕДЕНИЯ КОМПЛЕКСНОЙ ОЦЕНКИ ЭРГНОМИКИ ЧЕЛОВЕКО-МАШИННОГО ВЗАИМОДЕЙСТВИЯ

*А. А. Маркина, Д. А. Костюк*

*Брестский государственный технический университет, Брест*

Для измерения эффективности работы пользователя можно выделить два принципиально различных подхода. В первом подходе активно используются экспертные оценки и самосообщаемые параметры (то есть сведения, полученные от респондентов). Этот подход не поддается значительной автоматизации и затратен по времени из-за опросов, хронометража и видеопотоколирования, больше подвержен влиянию человеческого фактора при обработке результатов. Также для ряда параметров, таких как физическая нагрузка, могут быть получены лишь оценки по косвенным признакам и самосообщаемым параметрам. Альтернативный подход предполагает экспресс-оценку состояния пользователя с помощью приборов, позволяющих регистрировать параметры, связанные с физической и когнитивной нагрузкой. Он также имеет ряд ограничений. Биометрические измерения предоставляют большой объем данных, поддающихся автоматической обработке, однако они позволяют оценивать эмоциональное состояние и когнитивные процессы лишь по косвенным признакам, что даёт в ряде случаев лишь приблизительную картину.

В ряде работ нами рассматривается объединение обоих подходов в рамках комплексного usability-тестирования [1], нацеленного на получение картины процессов, имеющих место в ходе человеко-машинного взаимодействия, максимально полной и максимально абстрагированной от влияния человеческого фактора в процессе его оценки. В рамках такого объединённого исследования выполняется совмещение самосообщаемых параметров, оцениваемых по классическим методикам, с анализом данных, получаемых от биометрических устройств бытового сегмента (фитнес-трекеров, энцефалографов, айтрекеров). При этом требуется принятие ряда организационно-методических мер, направленных на обеспечение адекватности исследования – как специфичных для тестирования на основе самосообщаемых параметров или приборной оценки, так и направленных на их эффективное совмещение в рамках серии экспериментов.

Наиболее доступными (а потому и наиболее распространенными) для задач оценки эргономики программных продуктов являются внутригрупповые эксперименты, предполагающие использование небольшой группы испытуемых. Проведение исследования облегчается тем, что в данном случае отсутствует необходимость деления пользователей на группы, введения контрольной группы и использования дисперсионного анализа, как это потребовалось бы в случае с межгрупповым тестированием. Также к достоинствам подхода можно отнести экономию времени, отсутствие сложных статистических расчетов, относительную легкость в интерпретации результатов, а к недостаткам – теоретическую возможность влияния индивидуальных особенностей испытуемого на результат эксперимента [2]. Также в ряде случаев первичный эксперимент может быть проведен как внутригрупповой с последующим дополнением при более высокой выборке до межгруппового с помощью дополнительных межгрупповых связей, таких как гендерные и возрастные различия.

Для проведения тестирования целесообразно предусмотреть следующие этапы:

— Составление сценария использования, включающего типовые задания для использования программного продукта с нарастающей сложностью, что позволяет отследить тенденцию обучаемости. Задания составляются исходя из опыта респондентов.

— Набор участников одной возрастной группы, имеющих сходный опыт работы с продуктами-аналогами. Оговаривается время проведения каждой стадии эксперимента.

— Инструктаж (приветствие, описание мероприятия, целей исследования, метрик и тестов, подписание соглашения на предоставление и обработку персональных данных).

— Вводное интервью (заполнение анкеты участника, проверка уровня владения продуктами-аналогами).

— Заполнение опросников, позволяющих провести предварительную психометрическую оценку респондентов. В этом качестве в различных экспериментах нами опробовались опросники «Актуальное состояние» и «Доминирующее состояние», опросник реактивной и личностной тревожности Спилберга-Ханина, а также проективная методика "восьмицветовой тест Люшера".

— Демонстрационный показ работы с продуктом.

— Оценка ожиданий от работы с продуктом. Респондентами заполняется стандартный международный опросник System Usability Scale (SUS), обсуждаются основные вопросы использования, особенности функционала.

— Настройка системы. Выполняется подключение и калибровка имеющегося биометрического оборудования, а также настройка видеопотоколирования и захвата видео с экрана.

— Работа с продуктом. Происходит выполнение тестовых заданий с ведением протокола менеджером, где фиксируются реакции пользователя.

— Сбор итоговых впечатлений. Заполняется стандартный международный опросник и Post-Study System Usability Questionnaire (PSSUQ), проводится ретроспектива.

Важно, что в ходе эксперимента менеджеры, регистрируя время и реакцию пользователей, не вмешиваются в ход выполнения заданий. В наших экспериментах менеджеры подбирались из круга респондентов, были с ними знакомы, что должно положительно сказываться на уровне стресса. В любом случае ответы на вопросы «Как это сделать?» должны быть расплывчаты («А как вы сами думаете?», «А что бы вы сделали в реальной жизни?»), т. к. это мотивирует респондента разбираться с системой и не смещает фокус. Также респонденты знают о видеонаблюдении, но внимание в ходе теста на этом не акцентируется.

Для уменьшения «эффекта последовательности», вызванного индивидуальными особенностями испытуемых, может быть проведен статистический межгрупповой анализ результатов серии тестов с разделением на две экспериментальные группы по виду теста без введения контрольной группы. Также предварительное обучение респондентов позволяет снизить эффект новизны программного обеспечения, а невозможность исключения фоновых естественных эффектов в ходе эксперимента целесообразно учитывать при обработке данных с помощью удаления невалидных результатов.

Список метрик, которые имеет смысл учитывать при проведении исследования, включает показатели результативности (время выполнения задания, ошибки, процент выполненных заданий), нагрузку (частоту сердечных сокращений, концентрацию внимания, фиксации взгляда) и впечатления респондентов (уровень ожиданий и уровень удовлетворения пользователя).

Время и успешность выполнения задания относятся к числу основных показателей и используются только в сравнении. Успешность выполнения выражается в бинарном коде (выполнил/не выполнил).

Проблемы, с которыми столкнулись пользователи, регистрируются как менеджерам эксперимента, так и с помощью видеозахвата экрана. После тестирования проводится ретроспектива: с пользователем обсуждаются задания, вызвавшие проблемы, проигрывается запись, анализируется реакция и поведение в ходе теста. Это позволяет классифицировать проблемы, выявить более значимые для пользователя, рассчитать частотность проблем (сколько пользователей с ними столкнулись).

Уровень ожидания отражает отношение пользователя к продукту и предоставляемую им комфортность работы, а уровень удовлетворения – оценку удобства использования системы после прохождения теста. Эти метрики можно получить с помощью опросников SUS и PSSUQ. Кроме этого, пользователям можно предложить выбрать из списка эпитеты, которые могут описать их впечатления от программного продукта в соответствии с методикой Microsoft Desirability Toolkit [1].

Нагрузка может оцениваться в соответствии с опробованной ранее экспресс-методикой [3], по которой регистрируются быстрота выполнения заданий, физическая и умственная нагрузка, направление взгляда пользователя [4, 5]. Для оценки концентрации внимания могут использоваться метрики доступных энцефалографов потребительского сегмента (например, метрика «Attention» энцефалографа Neurosky Mindwave) либо оценка на основе соотношения альфа- и бета-ритмов [6]. Физическую нагрузку целесообразно оценивать по частоте сердечных сокращений, измеряемой фитнес-трекером, а наиболее доступными средствами регистрации направления взгляда на по нашему опыту являются потребительские айтрекеры фирмы Tobii [7].

Совместное применение средств определения уровня когнитивной нагрузки на основе относительной эффективности условия и оценки её количественного нейрофизиологического выражения с помощью стандартных измерительных инструментов помогает более качественно выполнить сравнительную оценку эффективности интерфейса, задействовав системный многофакторный анализ, продуктивно выявлять критические участки и ограничивающие факторы человеко-машинного взаимодействия с последующим формированием предложений по их преодолению.

#### Список литературы

1. Костюк, Д. А. Подход к комплексному межгрупповому usability-тестированию для платформы GNU/Linux / Д.А. Костюк, А. А. Маркина // Свободное программное обеспечение в высшей школе: сб-к материалов Тринадцатой конференции. М.: "МАКС Пресс", 2018. – С. 39–44.
2. Tullis T., Albert W. Measuring the User Experience Collection, Analyzing and Presenting Usability Metrics – Morgan Kaufmann, 2013 – 320 p.

3. Маркина, А. А. Система параллельного тестирования эффективности человекомашинного взаимодействия // Тринадцатая конференция разработчиков свободных программ: Тезисы докладов / Калуга, 01-02 октября 2016. – С. 32–37.
4. Применение айтрекеров для юзабилити-исследований ПО в GNU/Linux / А. Дубицкий [и др.] // Четырнадцатая конференция разработчиков свободных программ: тезисы докладов – Калуга, 22–24 сентября 2017 г. – М. : Базальт СПО, 2017. – С. 36–41.
5. Костюк Д. А. Подход к биометрической оценке эргономики графического интерфейса пользователя / Д. А. Костюк, О. О. Латий, А. А. Маркина // Вестник БрГТУ. – № 5 (101): : физика, математика, информатика. – Брест: Издательство БрГТУ, 2016 г. – С. 46–49.
6. Костюк, Д. А. Использование электроэнцефалографии для оценки эмоциональных состояний / Д. А. Костюк [и др.] // Новые горизонты – 2020: сборник материалов VII Белорусско-Китайского молодежного инновационного форума, 17 ноября 2020 года / Белорусский национальный технический университет. – Минск : БНТУ, 2020. – Т. 1. – С. 37-39.
7. Костюк, Д. А. Применение устройств окулографической биометрии потребительского сегмента в исследованиях эргономики человеко-машинных интерфейсов / Д. А. Костюк, А. А. Маркина // Вестник Брестского государственного технического университета. Серия: Физика, математика, информатика. – 2020. – № 5. – С. 78–81.

УДК 004.9

## **ТЕХНИКА ДИНАМИЧЕСКОЙ ВИЗУАЛИЗАЦИИ В ЗАДАЧЕ ПОСТРОЕНИЯ ЭЛЕКТРОННЫХ СРЕДСТВ ОБУЧЕНИЯ**

*А. Ю. Савицкая*

*Брестский государственный технический университет, Брест  
Научный руководитель И. И. Гладкий*

Одним из приоритетных направлений в области повышения качества обучения техническим дисциплинам является разработка и внедрение инновационных образовательных технологий, основанных на применении современных аппаратно-программных средств компьютерной техники. Практика применения компьютерных обучающих систем совместно с традиционными средствами обучения демонстрирует существенное улучшение качества знаний и навыков слушателей. При этом на первый план выходит задача принципиально нового построения содержания учебного материала, деятельности преподавателя и учебной работы студента в компьютерной среде.

В контексте данной задачи наиболее перспективным направлением можно считать использование техники динамической визуализации (ДВ) и элементов виртуальной реальности. Они позволяют студенту самостоятельно исследовать процессы, протекающие в изучаемом объекте, понять основные закономерности, получить всестороннее представление об излагаемом материале. Часто компьютерная программа ДВ позволяет наглядно представить материал, словесное описание которого объемно и достаточно сложно для восприятия, а показ на натурном образце невозможен или затруднен (например, физические, информационные процессы и др.). Важным положительным свойством техники ДВ является возможность визуализации функционирования таких объектов и систем, в которых работа различных взаимоувязанных в систему компонентов полностью или частично совмещена во времени. Большим достоинством ДВ также является возможность имитации и моделирования протекания различных