## МЕТОДИКА

УДК 159.9.07, 004.5, 612.821, 159.928.235

# КОМПЬЮТЕРНОЕ ТЕСТИРОВАНИЕ И ГЕЙМИФИКАЦИЯ: ПЕРСПЕКТИВЫ МОНИТОРИНГА ФУНКЦИОНАЛЬНОГО СОСТОЯНИЯ РАБОТНИКОВ В ЭПОХУ КОМПЬЮТЕРИЗАЦИИ ПСИХОДИАГНОСТИКИ

### А. Г. Шмелев, И. А. Лисица

В статье рассматриваются инновации в области обучения и оценки персонала, связанные с внедрением процессов геймификации компьютерных средств обучения и оценивания (психодиагностики). В теоретикометодологической части статьи дается операциональное определение термина «геймификация», а также обсуждаются факторы, препятствующие успешной геймификации, — снижающие валидность компьютерных психодиагностических методов, построенных на использовании игровых элементов. В экспериментально-методической части статьи приводится описание новой компьютерной методики «Ложки», предназначенной для оценки функционального состояния. Эта краткая (трехминутная) методика представляет собой игровую модификацию таких классических тестов внимания, как «корректурная проба Бурдона» и «устный счет по Крепелину». Авторы публикуют данные о достаточно высокой надежности и валидности новой методики, подтверждающие возможность ее применения в качестве инструмента для мониторинга (оперативного отслеживания) функционального состояния.

*Ключевые слова*: оценка персонала, компьютерная психодиагностика, геймификация, умственная работоспособность, функциональное состояние, корректурная проба, подсчет визуальных предметов.

**Шмелев Александр Георгиевич** — доктор психологических наук, профессор кафедры психологии труда и инженерной психологии МГУ имени М.В. Ломоносова. E-mail: alshmelyov@yandex.ru

Лисица Ирина Анатольевна — кандидат психологических наук, ведущий специалист Инновационного центра «Гуманитарные Технологии» (Москва). *E-mail*: lia@ht ru

## Об игровом подходе в психодиагностике

Термин «геймификация» чаще всего встречается в современных прикладных разработках в связи с компьютерным обучением (один из лучших обзоров в этой области — Hamari et al., 2014; см. также статью "Gamification of learning" в англоязычной Википедии). Под этим процессом имеется в виду не полное превращение процесса обучения в игру, а использование в процессе обучения определенных игровых элементов: игровых целей («обыграть компьютер» или «поставить рекорд», например), игровых объектов и персонажей, практически непрерывной обратной связи об успешности действий учащегося (игровая обратная связь), зрительных и звуковых игровых эффектов и других элементов. В последние годы сфера приложения геймификации расширилась. Особенно значима геймификация в сфере рекрутмента и ассессмента — привлечения и оценки персонала (Snider, 2015). Главная цель всякой геймификации в ходе любого компьютеризированного процесса — повышение внутренней мотивации пользователя на выполнение деятельности (см. сайт германской консалтинговой компании www.gamificationsoftware.com).

Первые работы по развитию игрового подхода были выполнены в российской экспериментальной психологии и компьютерной психодиагностике еще в конце XX в. (Джерелиевская, Шмелев, 1993; Шапкин, 1999; Шапкин, Варашкевич, 1999; Шмелев, 1984; Шмелев, Гребенюк, 1994). К сожалению, экономический кризис в России в 1990-е гг. привел к упадку этого нового направления из-за отсутствия платежеспособного высокоинтеллектуального спроса в области оценки персонала. Впрочем, и на Западе о геймификации в прикладной психологии широко заговорили только в XXI в., когда появилось поколение психологов, для которых компьютерная игра с детских лет стала понятным, привлекательным и привычным видом деятельности. Основная цель игрового подхода в психодиагностике оказалась фактически аналогичной той цели, которая ставится при геймификации обучения и рекрутмента. Ее наиболее кратко и ясно сформулировал В.А. Дюк: «"Включение" игровой мотивации повышает привлекательность процесса тестирования и увеличивает достоверность результатов» (Дюк, 1994, с. 25).

Существенным содержательным тормозом для геймификации психодиагностики стало осознанное вскоре после ее появления препятствие, связанное с низкой экологической валидностью многих популярных стратегических и ролевых игр. Возможность переноса

результатов компьютерной игры из виртуального игрового мира в реальный мир оказалась затруднена тем обстоятельством, что от имени виртуального персонажа многие пользователи-игроки ведут себя вовсе не так, как они ведут себя от собственного имени в реальном мире (понятие «экологическая валидность» обосновано в работах В.Н. Дружинина (2002)). В виртуальной игровой среде пользователи демонстрируют значительно большую свободу и раскрепощенность в своих действиях, ибо последствия ошибок в этом случае не являются тяжелыми и сводятся к банальным затратам лишнего времени на их исправление. Неслучайно первый автор данного сообщения ввел в свое время понятие «мир поправимых ошибок» (Шмелев и др., 1988). В игре пользователь часто демонстрирует совершенно иной профиль мотивации, чем это свойственно ему же в ситуации реального поведения: в игре он чаще проявляет конкурентные (соревновательные) мотивы и конкурентную стратегию поведения, тогда как в реальной жизни ему в большей степени свойственны конформность и кооперативность (Фомичева и др., 1991). Назовем этот фактор, снижающий валидность игрового подхода, «эффектом атипичной мотивации».

В прошлом веке недостаточная экологическая валидность игрового подхода проявилась не только в стратегических и ролевых играх. Уровень достижений игрока в играх на точность-скорость сенсомоторных действий (локомоторные игры, «стрельба» и подобные им) в большей степени отражал не общий уровень сенсомоторных способностей, а лишь различия в степени тренированности игроков в данной компьютерной игре — в степени развития клавиатурных навыков, навыков манипулирования мышкой или джойстиком и т.п. А в 1990-е гг. степень тренированности и объем игрового опыта у разных пользователей еще в значительной степени различались. В настоящее время эти различия нивелируются в силу огромной повседневной практики у большинства пользователей в компьютерной коммуникации и работе с упрощенными интерфейсами, такими, как «тачскрин» — экран, чувствительный к прикосновениям. Этот фактор, снижающий валидность компьютерных игр, следует, очевидно, назвать фактором «специфических игровых навыков». Действие данного фактора заставляет разработчиков применять к игровым методикам не межиндивидуальные, а внутрииндивидуальные (ипсативные) тестовые нормы: сравнивать новые достижения пользователей с их собственными прежними достижениями с целью выявления положительной (рост показателей) или отрицательной (снижение) динамики результатов.

Третьим фактором, снижающим экологическую валидность игрового подхода, является чрезмерное многообразие игровых ситуаций и игровых сценариев, возникающих в современной развлекательной компьютерной игре. Назовем его «фактором игрового разнообразия». Компьютерная игра будет слишком скучной и коммерчески неэффективной, если в ней раз за разом будут повторяться однообразные игровые сцены. В силу этого обстоятельства развлекательная игра всегда отличается от игроподобного теста недостаточной гомогенностью (однородностью) заданий (*items*), т.е. игра лишена того свойства, которое как раз является главным условием измерительной точности-надежности для любого теста.

Тем не менее, несмотря на три указанных выше негативных фактора, снижающих диагностическую ценность компьютерных игр, все-таки разработку игроподобных диагностических тестов (геймифицированных тестов) следует считать достаточно перспективным направлением. И прежде всего мы можем ожидать в самое ближайшее время расширенного использования таких тестов в целях психодиагностического мониторинга функционального состояния работников — уровня мобилизации внимания и актуальной производительности умственной деятельности. Именно в диагностике функциональных состояний имеют значительную ценность такие тесты, которые требуется выполнять многократно и регулярно (в начале и в конце рабочего дня для контроля уровня утомления). В этом случае традиционные бланковые тесты или тесты без игровой обратной связи оказываются слишком скучными и менее валидными, так как быстрее вызывают пресыщение, чем игроподобные методики.

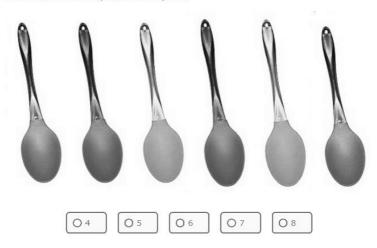
# Компьютеризированный тест «Ложки»

С целью исследования психодиагностического потенциала игроподобных тестов нами был разработан в 2016 г. новый компьютеризированный тест на скорость-точность простых умственных действий под игровым названием «Ложки». Этот тест реализован на базе тестовой онлайн-системы HT-LINE M-TEST силами программистов и методистов Инновационного центра «Гуманитарные Технологии» (см. сайт www.ht-line.ru)<sup>1</sup>. Трехминутная краткая версия этого теста предполагает предъявление

<sup>&</sup>lt;sup>1</sup> Авторы статьи выражают благодарность Евгении Орловой за помощь в подготовке графического материала для второй, улучшенной версии теста «Ложки».

испытуемому на экране компьютера 40 тестовых экранов с фотографическими изображениями столовых и чайных ложек разной формы, цвета, ориентации. Тестовые задания предъявляются с применением случайного перемешивания и постепенным повышением сложности: в первом блоке в случайном порядке предъявляются 20 заданий с числом объектов от 5 до 8, во втором — с числом объектом от 9 до 25. На каждом шаге задача испытуемого состоит в подсчитывании числа объектов на экране и выборе одного (правильного) ответа из пяти предложенных. Пример тестового экрана приведен на рисунке.

Сколько ложек изображено на экране:



Тестовый экран в игроподобном тесте «Ложки». Цвета ложек слева направо: голубая, оранжевая, зеленая, оранжевая, зеленая, голубая

Сам по себе прием подсчета числа объектов не является новым и его можно даже назвать «классическим». И академическим, и практическим психологам широко известен один из самых популярных в мире бланковых тестов на продуктивность внимания — «корректурная проба Бурдона», в ходе которой задача испытуемого состоит в том, чтобы обводить и подчеркивать определенные буквы на листе бумаги, подсчитывая при этом их количество. В нашем тесте предъявляется несколько менее однообразный стимульный материал, что придает деятельности испытуемого большую зрительно-эстетическую привлекательность, снижает монотонию и эффект

пресыщения (об этом собраны подтверждающие положительные отзывы от испытуемых, прошедших тест «Ложки»). Определенное напряжение в выполнении данного теста создается благодаря динамическому графическому образу «убывающего времени», отведенного на трехминутный сеанс<sup>2</sup>.

Проверка надежности и валидности теста «Ложки» производилась на испытуемых-добровольцах, постоянных участниках виртуального проекта «Клуб испытателей тестовых технологий» (КИТТ). Клубная форма с постоянной регистрацией и накоплением «индивидуальной истории» для каждого участника является удачным организационным решением для измерения конвергентной (конструктной) валидности между различными тестами. Надежность измерялась как корреляция двух параллельных форм трехминутного теста — вариантов, созданных на базе различных тестовых заданий (по 40 разных заданий в каждой параллельной форме). Значение коэффициента линейной корреляции Пирсона оказалось равным 0.71 (на выборке n=133). С помощью поправки Спирмена—Брауна для теста удвоенной длины надежность теста из 80 заданий (двух трехминутных серий) можно оценить величиной 0.83, что можно считать вполне удовлетворительной величиной для экспресс-тестов, построенных не на самоотчете, а на объективном показателе — количестве правильных ответов.

На 76 участниках клуба, выполнивших и новый тест «Ложки», и старый тест «Краткий тест отбора» (КТО), выявлена высокая значимая линейная корреляция между количеством правильных ответов в тесте «Ложки» и количеством правильных ответов в субтесте «Внимание» из КТО (0.49, значимость на уровне p<0.001). Менее высокими, но тоже значимыми оказались корреляции теста «Ложки» с субтестами числового и вербального интеллекта (0.39 и 0.28 соответственно). А вот с субтестом «Эрудиция» корреляция оказалась, как и ожидалось, незначимой и даже слабоотрицательной (-0.18). Высокая положительная корреляция с субтестом «Внимание» означает, что именно испытуемые из высокой группы по тесту КТО (по шкале «Внимание») считают визуальные объекты быстрее и точнее. А значимую связь с другими субшкалами КТО можно интерпретировать двумя способами: а) с одной стороны, как свидетельство того, что при решении формально-логических и лексических

 $<sup>^2\,</sup>$  В тестах с ограничением времени на экране испытуемого имеется динамический прогресс-бар, наглядно демонстрирующий, что время на выполнение теста убывает.

тестовых заданий требуется понятная и объяснимая мобилизация общекогнитивной функции внимания; б) с другой стороны, как свидетельство того, что при пересчитывании визуальных объектов на экране более эффективный результат достигается с помощью операций «визуального мышления» — определенного навыка группировки и симультанной оценки численности подгрупп визуальных объектов на экране. Таким образом, выявленный профиль связей игроподобного теста «Ложки» с различными субтестами КТО указывает на то, что новый тест является, несмотря на свою краткость (время выполнения всего лишь 2—3 минуты), вполне валидным инструментом для измерения межиндивидуальных различий в производительности (скорости и точности) умственного труда.

Дополнительная исследовательская задача заключалась в **проверке чувствительности теста** «Ложки» к внутрииндивидуальным различиям в функциональном состоянии. На экспериментальном уровне этот вопрос был переформулирован так: может ли динамика результатов этого теста при его последовательном многократном выполнении отразить уровень психической выносливости наших испытуемых? Для проверки этой гипотезы была создана более полная версия данного теста, включающая 4 серии по 40 тестовых экранов с раздельным контролем времени внутри серии и краткими перерывами между сериями. При этом лимит времени на каждую следующую серию постоянно сокращался, о чем испытуемого предупреждали в инструкции перед каждой серией: 180 секунд (3 минуты), 150 секунд, 120 секунд, 90 секунд. Создавался эффект моделирования нарастающего стресса, вызванного нарастающим дефицитом времени на решение задачи.

Среди 260 испытуемых, прошедших полную четырехэтапную версию теста «Ложки», только несколько человек смогли выполнить за 180 и 150 секунд правильно все 40 заданий. За 120 и 90 секунд полностью и безошибочно выполнить задание не смог никто (те, кто предпочитал работать несколько медленнее и без ошибок, как правило, просто не успевали пересчитать все ложки за отведенный срок). Среднее число правильных ответов по сериям выглядит так: 36, 35, 33 и 29.

Естественно при этом, что от первой к последней серии росли показатели стандартного отклонения: 3.2, 3.6, 4.6 и 13.05. Особенно высокий показатель в заключительной серии отражает фактический отказ некоторых испытуемых от попытки выполнить задание за 90 секунд (полторы минуты) и случайный принцип выбора числовых ответов с их стороны. Таким образом, более высокие межин-

дивидуальные различия данный тест выявляет не по абсолютному уровню результатов в первой серии, а именно по степени падения результатов к последней серии, т.е. тест «Ложки» демонстрирует свою диагностическую чувствительность к отрицательной динамике психической выносливости.

По четырехэтапному тесту «Ложки» был построен производный показатель в виде разности между точностью в двух последних сериях и в двух первых сериях. Чем выше эта разность, тем меньше убывание точности, т.е. выше операциональный показатель «выносливость». Для измерения валидности этого показателя по тесту «Ложки» мы рассчитали его корреляцию с показателями участников клуба КИТТ по пяти факторным шкалам тест-опросника «Большая пятерка» и близким к ним факторным шкалам полупроективного теста «Выбор картинок». На 126 испытуемых, выполнивших обе методики, была обнаружена значимая положительная корреляция нового показателя выносливости именно с родственным фактором из «Большой пятерки» — «Эмоциональной стабильностью» (r=0.24, p<0.01). Причем эта связь оказалась более выраженной для полупроективной методики «Выбор картинок», чем для вербального тест-опросника «Большая пятерка». Следует подчеркнуть, что именно для четвертой (заключительной) серии теста «Ложки», в которой моделировался самый жесткий дефицит времени, все корреляции достигали максимума, а для первой серии значимых связей с индивидуально-типологическими особенностями испытуемых выявлено не было. Высокая положительная корреляция со шкалой «Эмоциональная стабильность» означает, что именно испытуемые из высокой группы по этой шкале демонстрировали более высокую выносливость в тесте «Ложки».

В настоящее время производится накопление данных с целью проверки внешней валидности теста «Ложки». Этот компьютерный тест выполняют рабочие одного из предприятий г. Кирово-Чепецка. На небольшой в данный момент выборке (48 человек) удалось обнаружить близкую к значимой корреляцию (r=0.21) между баллом по тесту в первой краткой версии (из одной трехминутной серии) и отсутствием аварийных ошибок в работе. Но эта слабая статистическая связь нуждается в перепроверке на более представительной выборке.

Полученные результаты дают возможность утверждать, что разработанный нами игроподобный тест, дающий испытуемому обратную связь об уровне точности-скорости его работы каждые 2—3 минуты, отражает не только устойчивые индивидуальные

особенности, но и определенную динамику функционального состояния и может быть использован для контроля и самоконтроля<sup>3</sup> за уровнем готовности человека к выполнению работы, требующей повышенной скорости и точности-надежности умственных действий, т.е. в ходе управления опасными производственными объектами (транспортными, энергетическими и т.п.).

+ \* \*

Определенная «безыскусность» теста «Ложки» вполне может вызвать такой вопрос: «А нельзя ли получить тот же самый полезный эффект мониторинга функционального состояния с помощью таких же непродолжительных, но более увлекательных компьютерных игр, как, например, «Тетрис-спринт» (модификация известной игры «Тетрис», выполняемая всего лишь за 2 минуты)?». Мы считаем это в принципе возможным. Для более точного ответа требуется провести специальное исследование. Но наш опыт исследования компьютерной игры «Тетрис» (под консультативным надзором автора этой игры А.Л. Пажитнова) показал, что «Тетрис» более чувствителен, чем подсчет объектов на экране, к такому вредному с точки зрения диагностической точности фактору, как «специфический игровой опыт». У опытных игроков в «Тетрис» формируется особый навык управления вращением фигурок на экране и т.п., в то время как в массовом и общедоступном компьютерном тесте функционального состояния от пользователя-испытуемого должна требоваться самая обычная активность — устный счет и реакция выбора одной из точек сензитивности на экране. Кроме того, многие игроки с увлечением играют в тот же «Тетрис», не ограничиваясь одной короткой серией, что приводит уже к эффекту неоправданной потери рабочего времени.

Общий вывод: геймификация в психодиагностике должна решать непростую задачу выбора золотой середины между «пресыщением от монотонии» и «чрезмерной увлекательностью» в компьютерной игровой деятельности.

<sup>&</sup>lt;sup>3</sup> Многократное использование теста «Ложки» одним человеком, как и многократная игра на компьютере, становится возможным благодаря включению механизма «случайного перемешивания заданий», реализованному, в частности в данном тесте. Следует обратить внимание, что этот прием невозможно реализовать в бланковом тестировании, он доступен только в режиме компьютерного тестирования.

#### СПИСОК ЛИТЕРАТУРЫ

Джерелиевская М.А., Шмелев А.Г. Опыт взаимодействия методов в компьютерной диагностике коммуникативных диспозиций // Вестник Московского университета. Сер. 14. Психология. 1993. № 3. С. 66—69.

Дружинин В.Н. Экспериментальная психология. СПб.: Питер, 2002.

Дюк В.А. Компьютерная психодиагностика. СПб.: Братство, 1994.

Фомичева Ю.В., Шмелев А.Г., Бурмистров И.В. Психологические корреляты увлеченности компьютерными играми // Вестник Московского университета. Сер. 14. Психология. 1991. № 3. С. 27—39.

*Шапкин С.А.* Компьютерная игра: новая область психологических исследований // Психологический журнал. 1999. Т. 20. № 1. С. 47—67.

*Шапкин С.А.*, *Варашкевич С.А.* Воздействие компьютерной игры на функциональное состояние пользователя // Вестник Московского университета. Сер. 14. Психология. 1999. № 3. С. 25—33.

*Шмелев А.Г.* На пути к компьютерной психодиагностике // Вестник Московского университета. Сер. 14. Психология. 1984. № 2. С. 13—17.

Шмелев А.Г., Бурмистров И.В., Зеличенко А.И., Пажитнов А.Л. Мир поправимых ошибок (психология компьютерных игр) // Вычислительная техника и ее применение. Компьютерные игры. 1988. № 3. С. 16—84.

*Шмелев А.Г., Гребенюк Г.А.* Диагностика менеджерского стиля с помощью компьютерной игры и тест-опросника // Вестник Московского университета. Сер. 14. Психология. 1994. № 2. С. 59—66.

*Hamari J., Koivisto J., Sarsa H.* Does gamification work? A literature review of empirical studies on gamification // Proceedings of the 47th Hawaii International Conference on System Sciences (Hawaii, USA, January 6—9, 2014).

Snider E. Can computer games double as personality assessment tools? // E-journal «Search financial applications» on the website www.techtarget.com. 2015. URL: http://searchfinancialapplications.techtarget.com/feature/Can-computergames-double-as-personality-assessment-tools

Поступила в редакцию 29.11.16 Принята к публикации 15.12.16

# COMPUTER TESTING AND GAMIFICATION: THE PERSPECTIVES OF MONITORING OF THE FUNCTIONAL STATES OF THE PERSONNEL IN THE ERA OF COMPUTERIZATION OF PSYCHODIAGNOSTICS

# Alexandr G. Shmelyov<sup>1</sup>, Irina A. Lisitza<sup>2</sup>

<sup>1</sup> Lomonosov Moscow State University, Faculty of Psychology, Moscow, Russia

Abstract: This short article discusses innovations in training and personnel assessment associated with the implementation of gamification of computer tools for learning and assessing (psycho-diagnostics). In the theoretical part of the article the authors give an operational definition of the term "gamification", and also discussed factors that impede successful gamification — reduces the validity of computer psycho-diagnostic methods based on the use of game elements. In an experimental-methodological part of the article describes a new brief, three-minute computer technique for the estimation of a functional state called "The spoons". This technique is a game modification of the classic "proof-reading test". The authors publish data on a sufficiently high reliability and validity of the new method, confirming the possibility of its use as a tool for monitoring (operational security) functional status.

**Key words:** personnel assessment, computerized psychodiagnostics, gamification, mental capacity, functional states, proof-reading test, account of visual objects.

#### **References:**

Druzhinin, V.N. (2002) Eksperimental'naya psikhologiya [Experimental psychology]. St. Petersburg: Piter.

Dyuk, V.A. (1994) Komp'yuternaya psikhodiagnostika [Computer psychodiagnostics]. St. Petersburg: Bratstvo.

Dzherelievskaya, M.A., Shmelev, A.G. (1993) Opyt vzaimodeystviya metodov v komp'yuternoy diagnostike kommunikativnykh dispozitsiy. *Vestnik Moskovskogo universiteta. Ser. 14. Psikhologiya* [Moscow University Psychology Bulletin], 3, 66—69.

Fomicheva, Yu.V., Shmelev, A.G., Burmistrov, I.V. (1991) Psikhologicheskie korrelyaty uvlechennosti komp'yuternymi igrami. *Vestnik Moskovskogo universiteta. Ser. 14. Psikhologiya* [Moscow University Psychology Bulletin], 3, 27—39.

Hamari, J., Koivisto, J., Sarsa, H. (2014) Does gamification work? A literature review of empirical studies on gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences* (Hawaii, USA, January 6—9, 2014).

<sup>&</sup>lt;sup>2</sup> Innovation Center "Human Technologies", Moscow, Russia

Shapkin, S.A. (1999) Komp'yuternaya igra: novaya oblast' psikhologicheskikh issledovaniy. *Psikhologicheskiy zhurnal* [Psychological journal], 20, 1, 47—67.

Shapkin, S.A., Varashkevich, S.A. (1999) Vozdeystvie komp'yuternoy igry na funktsional'noe sostoyanie pol'zovatelya. *Vestnik Moskovskogo universiteta. Ser 14. Psikhologiya* [Moscow University Psychology Bulletin], 3, 25—33.

Shmelev, A.G. (1984) Na puti k komp'yuternoy psikhodiagnostike. *Vestnik Moskovskogo universiteta. Ser. 14. Psikhologiya* [Moscow University Psychology Bulletin], 2, 13—17.

Shmelev, A.G., Burmistrov, I.V., Zelichenko, A.I., Pazhitnov, A.L. (1988) Mir popravimykh oshibok (psikhologiya komp'yuternykh igr). *Vychislitel'naya tekhnika i ee primenenie. Komp'yuternye igry* [Computer technology and its application. Computer games], 3, 16—84.

Shmelev, A.G., Grebenyuk, G.A. (1994) Diagnostika menedzherskogo stilya s pomoshch'yu komp'yuternoy igry i test-oprosnika. *Vestnik Moskovskogo universiteta. Ser. 14. Psikhologiya* [Moscow University Psychology Bulletin], 2, 59—66.

Snider, E. (2015) *Can computer games double as personality assessment tools*? URL: http://searchfinancialapplications.techtarget.com/feature/Can-computergames-double-as-personality-assessment-tools

Original manuscript received November, 29, 2016 Revised manuscript accepted December, 15, 2016