

РАЗРАБОТКА ТЕСТА ОЦЕНКИ СИТУАЦИЙ НА ОСНОВЕ МОДЕЛИ СИТУАЦИОННОГО РУКОВОДСТВА

Е.В. Лурье¹

Для эффективной оценки персонала в ситуации лимита времени и средств необходимы стандартизированные инструменты. Предлагается использовать для этих целей тесты оценки ситуаций. Представлены результаты разработки теста такого типа, направленного на оценку навыков ситуационного руководства (по модели Херси и Бланшара). Обоснована надежность и достоверность методики, приведены данные исследования критериальной и конструктивной валидности, обсуждаются результаты этой работы и определяются границы использования инструмента.

Ключевые слова: тест оценки ситуаций, Центр Оценки, компетенции, ситуационное руководство, внутренняя надежность-согласованность, критериальная валидность, конструктивная валидность.

Введение

Компетенционный подход, активно применяемый в современном бизнесе и образовании, меняет требования к оценочному инструментарию организационной психологии. Ядром методологии компетенций является ситуационный подход [24]. Можно утверждать, что именно системное обращение к ситуационному и предметному контексту поведения работников дало компетенциям такое важное преимущество и успех² по сравнению с традиционными подходами, основанными на оценке способностей и черт. Формирование нового подхода на уровне критериев оценки влечет за собой коренные изменения в прикладном оценочном инструментарию. Однако в России ситуация с развитием методик оценки продолжает оставаться неопределенной [10].

С одной стороны, нельзя не отметить бурного развития технологий экспертной оценки, реализованных, в первую очередь, в методе Assessment Center – Центров Оценки [8, 13]. Они уже приобрели заслуженную популярность в нашей стране, хотя качественных исследований Центров Оценки в России по-прежнему не хватает. Заметим при этом, что задачи диагностики персонала явно шире возможностей этого метода. Назовем лишь

главные из ограничений Центра Оценки, особенно явно проявившихся в период последнего экономического кризиса:

- высокая стоимость качественно разработанных Центров Оценки;
- жесткие требования к уровню квалификации наблюдателей-экспертов;
- логистические и организационные издержки и ограничения.

Эти особенности Центров Оценки ограничивают возможность их использования, например, в ситуации массового подбора персонала или при первичном отборе в кадровый резерв крупных компаний. А в посткризисный период всё чаще встречаются именно такие запросы на дешевые и короткие технологии оценки.

Спрос на них выявил дефицит в России стандартизированных психодиагностических инструментов, предназначенных для прогноза компетенций [9]. Существующие личностные и мотивационные опросники предоставляют традиционные данные о диспозициях [14]. Однако кросс-ситуативная оценка черт обладает, как известно, «пределом ... надситуативной устойчивости» величиной 0,30 [19]. Этого недостаточно для надежного прогноза в целях отбора персонала. Наиболее распространенные тесты способностей предназна-

¹ Автор выражает благодарность Едренкину И.В., Куликовой В.Г., Ободянской А.Н. и канд. психол. наук Потапкину А.А. за помощь в разработке представленного в статье теста «2Ф».

² По состоянию на 2010 г. в России моделями компетенций располагают по разным данным от 67 до 87 % компаний [10, с. 243].

ны для оценки академического интеллекта [4, 12]. Поэтому и их очевидная и содержательная валидность вызывает серьезные нарекания из-за слабого соответствия специфики заданий особенностям конкретной профессиональной деятельности [12].

В связи с этим представляется особо значимым поиск качественных альтернатив, позволяющих, с одной стороны, приблизиться к валидной оценке компетенций, а с другой – обеспечить оперативность и дешевизну оценки. Одним из перспективных решений этих проблем является выбор тестов достижений, точнее их разновидности – тестов оценки ситуаций³ [23, 27].

Тесты оценки ситуаций

В последние 20 лет исследователи обратили особое внимание на важные преимущества тестов оценки ситуаций (ТОС). По структуре получаемых данных и по содержанию ТОС ближе всего подходят к оценке компетенций и представляют собой стандартизированный метод оценки, в котором респонденту предлагаются ситуации, моделирующие значимые компоненты работы, и варианты возможных действий в них [9; 23, с. 427]. Ситуации могут быть представлены в различном виде: видеосюжеты, изображения и анимационные стимулы, текстовые описания и т. п. В основе моделирования ситуаций лежит метод критических инцидентов Фланагана, который широко применяется и при создании моделей компетенций, и при конструировании Центров Оценки [15]. Современный подход к ТОС был оформлен в работах американских ученых Мотовидло и Стернберга.

Специалисты отмечают высокую значимость данного метода для современной психологической практики [12, 20, 25]. Отметим основные преимущества ТОС:

- высокая очевидная и содержательная валидность метода при удовлетворительных показателях его прогностической критериальной валидности⁴;
- большие возможности их стандартизации по отношению к другим методам, на-

правленным на моделирование ситуаций (например, по сравнению со стандартизацией групповой дискуссии);

- удобство и дешевизна администрирования и обработки при проведении массовых оценочных проектов;
- большая устойчивость по отношению к эффектам дискриминации [23].

Несмотря на отмеченные преимущества, ТОС распространены главным образом, в США [27]. В России существуют интересные разработки и отдельные публикации по данной проблеме, однако пока они отражают скорее эпизодический характер исследований в этой области [11, 20]. В настоящей статье представлены результаты разработки теста оценки ситуаций, направленного на оценку навыков ситуационного руководства.

Тест «2Ф»

Описание теории и конструкта

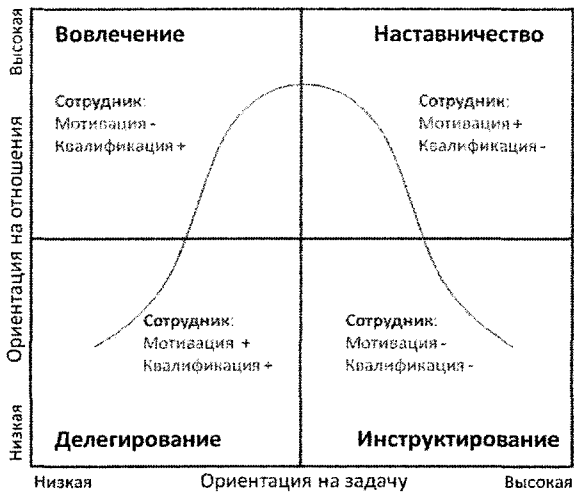
В течение 2008–2009 гг. нами был разработан тест оценки ситуаций «2Ф», предназначенный для оценки навыков ситуационного руководства у менеджеров начального и среднего звена управления. Теоретической основой методики является широко известная в теории менеджмента модель ситуационного руководства Херси и Бланшара [5, 21, 22]. В её основе лежит идея управления подчиненными на основании использования менеджером двух ключевых факторов – мотивации и квалификации сотрудника по отношению к определенной задаче. Комбинация двух этих факторов формирует матрицу размером 2×2, в которой представлены 4 базовые ситуации: сотрудник «хочет», но «не может»; «может», но «не хочет»; «не может и не хочет»; «может и хочет». Каждую из этих ситуаций авторы модели предлагают определить как отдельные стили руководства, имеющие соответствующие названия («Наставничество», «Вовлечение», «Инструктирование» и «Делегирование»). В поведении менеджера эти стили различаются по степени выраженности двух факторов – ориентации на задачу и ориентации на отношения (см. рисунок).

В соответствии с этой моделью от эффективного менеджера требуется точно выявить готовность сотрудника к решению задачи и, в зависимости от результата диагностики, применить один из четырех стилей руководства. Модель носит «нормативный» характер, предписывая руководителям выбирать определенное поведение в зависимости от ситуации. Такая особенность модели делает её

³ ТОС – от англ. Situational Judgment Tests (SJT). В России также используется название «кейс-тест» или кейсовая методика.

⁴ Мета-анализ 95 исследований ТОС, проведенный McDaniel с коллегами в 2001 г., показал среднее значение критериальной валидности $r = 0,34$ и $r = 0,38$ – для инструментов, в которых проводился предварительный анализ должности [26].

удобной основой для проектирования оценочного инструментария для измерения ситуационно заданных переменных.



Модель ситуационного руководства Херси и Бланшара

Выбор модели Херси и Бланшара в качестве «точки отсчета» для разработки теста «2Ф» обусловлен несколькими важными обстоятельствами:

- модель опубликована в научной литературе, теоретически обоснована и валидизирована [5, 22];
- модель достаточно популярна в современном российском бизнес-образовании и в среде HR-менеджеров, формирующих заказ на него [17];
- модель нормативная и при этом ориентирована на изучение трудно нормируемой предметной области управления людьми;
- модель наглядна и убедительна даже для людей, не имеющих управленческого опыта, хорошо согласуется со здравым смыслом, проста в обучении и практическом применении.

Однако для разработки ТОС автору потребовалось в определенной мере уточнить каноническую модель, предложенную американскими исследователями. Одним из спорных мест в теории Херси и Бланшара является её применимость к ситуации управления группой сотрудников, поскольку по отношению к определенной задаче у каждого из подчиненных может быть свой уровень готовности, что, в свою очередь, порождает необходимость применения менеджером разных стилей одновременно⁵. При разработке ТОС

⁵ Херси и Бланшар, разрабатывая собственные тестовые инструменты на основании модели, не

применялось моделирование только ситуации «руководитель – подчиненный». По нашему мнению, такое сужение предметной области методики оправдано и даже необходимо. Во-первых, это согласуется с наблюдением за ежедневной деятельностью руководителей, которые сталкиваются со взаимодействием на индивидуальном уровне чаще, чем с групповой работой [25, с. 97–98]. Во-вторых, чем более четко задана моделируемая предметная область, тем больше шансов представить её ключевые характеристики в коротких кейсовых заданиях.

Другим важным уточнением канонической модели стала операционализация двух ключевых факторов готовности сотрудника к решению определенной задачи – факторов мотивации и квалификации⁶. Для уточнения понятия мотивации предлагается различать три её вида:

- 1) мотивация сотрудника к выполнению конкретной рабочей задачи;
- 2) мотивация сотрудника по отношению к рабочему месту, к выполняемой профессиональной деятельности, лояльность к компании в целом;
- 3) личностная мотивация сотрудника (например, мотивация достижений).

Как можно заметить, виды мотивации располагаются в условном континууме «Ситуативное – Индивидуальное». Такое представление видов мотивации связано со стремлением ограничить моделируемую предметную область (как и в случае отказа от ситуаций «руководитель – группа») и получить более «прозрачную» модель принятия решения при обработке данных о правильности/неправильности ответов тестируемых. Результаты апробации теста «2Ф» показали, что включение в модель этих видов мотивации позволяет уменьшить возникающие у руководителей проблемы при обучении и использовании модели.

Итак, на готовность работника решать определенную задачу решающее влияние оказывает мотивация первого вида. Такая ситуа-

различали готовность группы и готовность сотрудника, предлагая в тесте часть ситуаций для первого и второго случая [17, 21].

⁶ Рабочее название теста «2Ф» (TwoF), или два фактора, было выбрано именно в связи с решающей ролью этих факторов в создании типовых ситуаций, с которыми должен уметь работать руководитель.

ционная мотивация определяется как желание сотрудника выполнять конкретную рабочую задачу, складывающееся из интереса к данной задаче; уверенности, что он эту задачу выполнит; готовности взять на себя обязательства за её выполнение. Именно мотивация этого вида маркирует ключевую информацию в заданиях теста, а остальные уровни маркируют дистракторы. Например, сотрудник может иметь высокий балл по тесту мотивации достижений (дистрактор), но из этого напрямую не следует его желание выполнять конкретную рабочую задачу по составлению объемного и неинтересного для него отчета (ключ).

Аналогичная схема разделения ситуационного и личностного контекстов была реализована и для второго фактора. Представления о квалификации работника позволили предложить разделение ее на три вида:

- 1) опыт и компетентность в решении подобных задач в данной организации;
- 2) опыт работы в определенной должности;
- 3) профессиональное образование, сертификации, «красивое» резюме.

Готовность работника зависит, в первую очередь, от квалификации первого вида. Ситуационная квалификация – это профессиональный опыт и компетентность сотрудника, позволяющие ему качественно выполнить конкретные виды работ в организации.

Подготовительный этап разработки теста

После проведения содержательной работы по определению конструкторов была использована типовая схема для разработки заданий и апробации инструмента [2, 3]:

- подготовительный этап (авторская разработка, экспертный отбор и редакция пунктов – описаний кейсов, компоновка рабочей версии теста);

- исследовательский этап (первичная апробация, коррекция пунктов, проверка внутренней согласованности и проведение основного пилотажного исследования, уточнение процедуры тестирования и поиск оптимальной схемы скоринга баллов). Отдельным шагом на этом этапе стало изучение и проверка надежности и валидности теста;

- стандартизация и разработка интерпретаций (формирование выборки стандартизации, проведение сбора данных и стандартизация инструмента, утверждение схемы скоринга и описание принципов интерпретации).

Спецификация разработанного теста «2Ф» представлена в табл. 1.

Авторами-разработчиками кейсов и экспертами являлись 7 бизнес-тренеров, активно ведущих обучающие курсы по менеджменту и использующие модель ситуационного руководства в практике преподавания и консультирования управленческого персонала, имеющих собственный опыт работы руководителем.

Итоговый набор пунктов теста «2Ф» представлен 9 описаниями ситуаций, согласованно отобранных экспертами из первоначальной совокупности 14 ситуаций. Для диагностики каждого из четырех стилей руководства предназначены по 2 ситуации. Кроме того, в тест включена 1 ситуация-дистрактор, в которой определение уровня мотивации и квалификации сотрудника не предусматривалось. Окончательная редакция описаний была выполнена после сбора 20 протоколов и содержательной обратной связи от участников тестирования.

Для всех кейсов был разработан одинаковый набор ответов, моделирующий цикл управленческих функций [7, с. 22]:

- функция «планирование» соответствовала задаче определения уровня мотивации и квалификации подчиненного по отношению к описанной в кейсе задаче;

- функция «организация» моделировалась выбором способа постановки задачи сотруднику;

- функция «мотивация» (частный случай функции регулирования) и «контроль» моделировались выбором оптимальных действий по мотивации и контролю подчиненного в конкретной ситуации.

Важным преимуществом использования такого подхода к моделированию функций управления является описание деятельности менеджера в динамике (от анализа ситуации до контроля выполнения задачи). Это повышает содержательную и очевидную валидность инструмента, делает бланк к тесту удобным дидактическим материалом в обучении ситуационному руководству.

Формат работы с тестом – «карандаш и бумага», среднее время работы с тестом – 25 минут (при условии отсутствия в инструкции каких-либо временных ограничений при выполнении теста). Инструкция предлагает поиск лучшего («наиболее правильного», а не «наиболее привычного») варианта ответа, по-

Спецификация теста оценки ситуаций «2Ф»

| Параметры | Характеристики | Основания для выбора характеристик |
|--|--|---|
| Тип теста | Тест оценки ситуаций | <ul style="list-style-type: none"> • Широкие возможности по стандартизации. • Соответствие компетенционному подходу |
| Формат пунктов: описания ситуаций | Ситуации, важные в деятельности менеджеров начального и среднего звена («критические инциденты»). Содержание не привязано к отрасли организации, деятельности в организации. Объем описания – 1/2 листа формата А4 | <ul style="list-style-type: none"> • Требования, предъявляемые к качественным ТОС [23]. • Необходимость создания универсального инструмента для менеджеров с различным опытом руководства сотрудниками. • Дидактическая значимость для целей развития тестируемых респондентов |
| Формат пунктов: описание действий в ситуациях | Одинаковые для всех ситуаций. Основанные на функциях управления. Выбор из заданных альтернатив возможного поведения руководителя | <ul style="list-style-type: none"> • Обоснованность с точки зрения конструктивной валидности. • Соответствие управленческой деятельности. • Расширенные возможности по интерпретации и обратной связи |
| Количество шкал | Одна – основная (ситуационная управленческая эффективность), 4 вспомогательных (эффективность использования 4-х стилей руководства) | <ul style="list-style-type: none"> • Результаты психометрической проверки. • Контроль содержательной валидности |
| Количество пунктов | 9 кейсов (по 2 – для диагностики каждого стиля плюс 1 дистрактор); 5 блоков ответов – для каждого кейса (по 8 пунктов на каждую из 4-х субшкал, 32 пункта в основной шкале) | <ul style="list-style-type: none"> • Баланс времени работы и надежности измерения показателей. • Борьба с фактором знания модели ситуационного руководства (дистрактор). • Моделирование управленческого действия в динамике (от планирования до контроля) |
| Процедура тестирования | Тест типа «карандаш – бумага»: ситуации описаны в виде текста, ответы регистрируются на бланке. Без ограничения по времени (среднее время 25 мин). «Знаниевая» инструкция | <ul style="list-style-type: none"> • Сохранность стимулов и ответов, защита интеллектуальной собственности; • Отсутствие значимых аргументов в пользу лимитирования времени тестирования. • Контроль надежности и достоверности |
| Обработка результатов, принцип построения ключей | Смешанный тип: построение ключей на основе теоретической модели и стат. данных выборки стандартизации | <ul style="list-style-type: none"> • Большая универсальность выводов. Возможность построения обоснованных дидактических рекомендаций по развитию. Соответствие разнородной выборке (управленцы российских и международных компаний) |
| Области применения | Оценка потенциальных и действующих менеджеров начального и среднего звена управления. Развитие управленческих навыков | <ul style="list-style-type: none"> • Результаты психометрической проверки. • Опора на теоретически проработанную модель. • Положительная обратная связь от практиков в оценке и развитии персонала, от руководителей |

сколько тест «2Ф» предназначен для оценки эффективности действий, а не для выявления диспозиционных черт или стилевых предпочтений респондентов⁷.

Результаты исследовательского этапа

Сбор данных и разработка ключей

В выборку исследования вошли результаты тестирования 317 чел., менеджеров различного уровня управления (от начального до высшего). Кроме того, проводилось тестирование сотрудников, не имеющих опыта руководства. Тестирование проводилось как в ситуации «клиента» (на бизнес-тренингах), так и в ситуации «экспертизы» (при оценке персонала). Фактор осведомленности о модели ситуационного руководства был сведен, по возможности, к минимуму: на тренингах испытуемые сначала заполняли тест, а затем изучали лежащую в его основе модель. В случаях наличия достоверной информации о знакомстве респондентов с теоретической моделью Херси и Бланшара результаты таких испытуемых исключались из дальнейшего исследования.

В ходе анализа результатов была уточнена методика построения ключей теста, разработанных на основе модели Херси и Бланшара. В спорных случаях решение о принятии ключа принималось на основе результатов частотного анализа. Таким образом, для диагностики каждого из четырех стилей имеется два кейса, в каждом из них требуется определить мотивацию и квалификацию, а также выбрать три действия (поставить задачу, мотивировать и проконтролировать). За правильное применение действия в соответствии с уровнем готовности подчиненного респондент получает 2 балла, за выбор действия, являющегося условным антагонистом в рамках модели ситуационного руководства – 0 баллов, выбор действий, лежащих между двумя этими оппозиционными действиями, приносит испытуемому 1 балл.

Для расчета сырых баллов по каждой субшкале использовались «штрафы», начисляемые за выбор действий, относящихся к стилю, неуместному для той или иной ситуации.

Проверка надежности и достоверности

Одним из важнейших требований к психодиагностической методике является обос-

⁷ Различия в формулировке инструкции (поиск наиболее привычного или поиск лучшего варианта) значительно влияют на результаты ТОС [26].

нование её надежности. Основной шкалой, прошедшей проверку надежности, стала шкала «ситуационной управленческой эффективности» (СУЭ), объединяющая результативность ответов во всех 8 заданиях. В данном исследовании применялся расчет показателей внутренней надежности-согласованности [16]⁸. Рассчитывались значения коэффициента альфа Кронбаха ($\alpha=0,55$ при количестве измерений $n=32$), коэффициент надежности параллельного расщепления Гутмана при разделении заданий на 2 группы, в каждую из которых вошло по 1 заданию на стиль (Guttman Split-half=0,4 при $n=16$ для каждой из групп). Значения коэффициента альфа Кронбаха составили: для первой группы заданий $\alpha=0,43$, для второй – $\alpha=0,47$.

Проверку надежности прошли также 4 субшкалы, характеризующие эффективность применения каждого из 4-х стилей руководства. Расчет «сырых» баллов для субшкал производился путем определения суммы баллов за решение двух ключевых заданий для каждого стиля за вычетом баллов, начисленных за «неуместное» использование в других шести заданиях действий, относящихся к этому стилю. Рассчитанные значения коэффициента альфа Кронбаха составили: для шкалы «Инструктирование» $\alpha=0,44$ при $n=17$, «Наставничество» $\alpha=0,43$ при $n=14$, «Вовлечение» $\alpha=0,35$ при $n=8$, «Делегирование» $\alpha=0,53$ при $n=23$.

Для проверки устойчивости полученных данных было проведено сравнение распределений случайным образом разделенной выборки (двухвыборочный критерий Колмогорова–Смирнова). Проверка показала достаточный уровень устойчивости данных ($p>0,63$ для шкалы СУЭ, $p>0,79$ для субшкал). Проверка шкал на нормальность распределения показала отрицательный результат (для всех шкал $p<0,01$ по одновыборочному критерию Колмогорова–Смирнова).

⁸ Проверка надежности через разработку альтернативных форм крайне затратна для ТОС и является задачей будущих исследований. Проверка тест-ретестовой надежности, к сожалению, была невозможна в имевшихся условиях сбора данных: респонденты получали обратную связь по результатам тестирования, включавшую рассказ о модели и подробный анализ ответов на тест. Фактор знания модели и ключей обесценивал любое повторное тестирование (хотя этот способ проверки надежности ТОС считается оптимальным [23, с. 430–432]).

Далее был проведен анализ показателей связи результатов тестирования по разработанным шкалам с дополнительными переменными, использованными для установления параметров достоверности и устойчивости теста «2Ф» к различным характеристикам потенциальных выборок и ситуаций тестирования. Значимых взаимосвязей шкал с возрастом ($r \leq 0,03$ при $N=230$) и полом ($p > 0,1$ по U-критерию Манна–Уитни при $N=317$) обнаружено не было. Однако по шкале СУЭ и субшкале «Инструментирование» данные о корреляции результатов в женской части выборки ($N=118$) оказались несколько выше мужской ($N=199$) при уровне значимости $p=0,19$ и $p=0,1$ соответственно.

Проверка устойчивости шкал к различным ситуациям тестирования производилась на выборке $N=442$. Полученные результаты свидетельствуют о достаточной устойчивости результатов тестирования по отношению к этому фактору ($p > 0,15$ для шкалы СУЭ по U-критерию Манна–Уитни). Из пяти шкал только показатели по шкале «Наставничество» значительно отличались в двух выборках: участники оценки персонала получили значительно более высокий средний ранг по данной шкале по сравнению с участниками тренингов ($p < 0,01$).

Проверка валидности

Критериальная валидность теста «2Ф» проверялась посредством оценки связи результатов по тесту с оценками по компетенциям, полученными в ходе Центров Оценки, а также общим оценочным рейтингом (ООР), построенным на основе этих оценок⁹ [25]. Разработанный инструмент (тест «2Ф») хорошо показал себя в ситуациях оценки персонала без предварительного отсева участников. Значения коэффициентов корреляции Спирмена для показателей по шкале СУЭ с данными ООР в двух проектах Центра Оценки составили: $r=0,64$ при $p \leq 0,05$ (для первого исследования с численностью испытуемых $N=10$) и $r=0,46$ при $p \leq 0,05$ (для второго исследования с $N=25$). Анализ характеристик взаимосвязей СУЭ и 3-х субшкал (кроме субшкалы «Делегирование») с отдельными компетенциями выявил значимые корреляции этих шкал с показателями компетенции

⁹ ООР рассчитывался как среднее арифметическое суммы всех оценок по компетенциям, полученным респондентом по результатам Центра Оценки.

«Управление исполнением» (владение базовыми управленческими навыками) ($r=0,46$ при $p \leq 0,05$ ($N=38$)).

При оценке персонала, проводимой после многоэтапного предварительного отбора участников перед попаданием на Центр Оценки, что проявилось относительно малой численностью выборки очевидно успешных испытуемых ($N=19$), корреляционный анализ не выявил значимых связей между шкалами теста «2Ф» и успешностью при отборе.

Изучение конструктивной валидности осуществлялась посредством поиска наличия / отсутствия связей показателей теста «2Ф» с результатами обследования по профессиональным психологическим тестам, сходным / различным по содержанию измеряемых конструктов. С этой целью применялись профессиональные психологические методики лаборатории «Гуманитарные технологии». В ходе проверки конвергентной валидности была обнаружена значимая связь показателей по субшкалам теста «2Ф» с данными обследования по шкале теста «Индивидуальный деловой стиль-3», наиболее близкого по содержанию и по принципу построения инструмента [6] (табл. 2). Также получены значимые корреляции с показателями по мотивационному опроснику ТИМ ($N=140$), отражающие направленность деятельности успешных (по данным теста «2Ф») руководителей на работу по заданным шаблонам [1].

При этом не обнаружено связей результатов теста «2Ф» с данными по личностному опроснику ТОП-ЮНИТ, основанному на модели Майерс–Бриггс ($N=101$) и с показателями по тест-опроснику эмоционального интеллекта ЭМИQ ($N=43$). Результаты корреляционного анализа связей данных исследования по методике «2Ф» с данными по другим тестам и опросникам представлены в табл. 2.

Обсуждение результатов

Результаты оценки показателей надежности-согласованности (значения коэффициента альфа Кронбаха $\alpha=0,55$ для основной шкалы СУЭ) свидетельствуют, что полученные показатели являются недостаточно высокими по сравнению с типовыми требованиями к тестам [16, с. 140]. Однако для методик типа ТОС этот уровень можно считать приемлемым и в какой-то мере типичным [27, с. 139–140]. В отличие от личностных опросников, в которых пункты, относящиеся к одной шкале, должны быть содержательно очень близки, а ситуации максимально обобщены и усредне-

Таблица 2

Сводная таблица связей шкал «2Ф» с тест-опросниками

| Методика | Шкалы | Ситуац. тестир. | Кол-во респонд. | Инструкт. | Настав. | Вовлеч. | Делегир. | Сит. управл. эффективн. |
|----------|-------------------|-----------------|-----------------|----------------------|---------|---------|----------|-------------------------|
| ИДС | Команда | Общ | 90 | -0,24* | | | | |
| | Задача | Общ | 90 | 0,18 (p=0,09) | | | | |
| | Невмешательство | Оценка | 49 | -0,3* | | | | |
| | Соц. желательн. | Оценка | 49 | | | 0,29* | | 0,29* |
| | Команда | Клиент | 41 | -0,4** | | | | |
| | Соц. желательн. | Клиент | 41 | -0,53** | | | | |
| | Невмешательство | Клиент | 41 | | | 0,32* | | |
| ТОП-ЮНИТ | | Общ | 101 | Связей не обнаружено | | | | |
| ЭмIQ | | Общ | 43 | Связей не обнаружено | | | | |
| ТИМ | Творч. / Независ. | Общ | 140 | | -0,23** | | | |
| | Познан. / Интерес | Общ | 140 | | | 0,18* | | |
| | Творч. / Независ. | Оценка | 70 | | -0,33** | | | -0,27* |
| | Преодолен. / Рост | Клиент | 70 | | | 0,24* | | |
| | Престиж | Клиент | 70 | -0,22 (p=0,06) | | | | |
| | Взаимоотношения | Клиент | 70 | -0,31* | | | | |

Примечания:

1. С подробным описанием методик можно познакомиться на сайте лаборатории «Гуманитарные технологии» по адресу URL: <http://www.ht.ru/cms/component/content/article/101055> (дата обращения 09.01.2011).

2. ** Корреляция значима на уровне $p \leq 0,01$.

* Корреляция значима на уровне $p \leq 0,05$.

ны, в ТОС ситуации намеренно конкретизируются и детализируются. Тест может претендовать на измерение ситуационно заданных конструктов только в том случае, если содержит в себе описания ситуаций, близких к реальным. При этом, конечно, возникает сложность с контролем дополнительных переменных, которые могут влиять на согласованность внутри шкалы. Так приводятся данные различных исследований разработки ТОС, показывающих, что высокие значения коэффициента альфа Кронбаха встречаются только в редких случаях: для методики видео-ТОС Lievens и Coetsier получен результат $\alpha=0,41$ и $\alpha=0,56$, а для ТОС из 20 вопросов получены значения коэффициента альфа 0,29 и 0,32 [27].

Другой потенциальный источник снижения надежности тесно связан с «многомерностью» измеряемых в ТОС конструктов [26]. Ситуационная управленческая эффективность (СУЭ), безусловно, относится к таковым. С одной стороны, многомерность порождает трудности с контролем дополнительных переменных, что снижает согласованность пунктов внутри шкалы, что и проявилось в

случае проверки теста «2Ф». С другой стороны, объединенные в одну шкалу разнородные пункты, затрагивающие четыре функции управления, создали достаточно устойчивую структуру – в случае удаления любого из 32 пунктов шкалы не наблюдалось увеличения значения коэффициента альфа Кронбаха.

Содержательно это согласуется с пониманием деятельности руководителя как единого процесса принятия решений, опирающегося на все четыре основные функции [7]. Результаты проверки надежности шкалы СУЭ и субшкал показывают, что о наличии / отсутствии у менеджера навыков ситуационного руководства можно судить только на основании оценки использования им всего арсенала управленческих функций.

Тем не менее требование высокой внутренней надежности-согласованности для теста не отменяется данными о невысоких значениях результатов исследований разработки ТОС в процитированных исследованиях. Получение более высокого значения коэффициентов внутренней надежности-согласованности должно являться одной из основных задач дальнейших этапов исследования.

Важным ограничением для использования инструмента является использование методики «2Ф» в режиме ретеста. Если участники после завершения тестирования получили обратную связь (что представляется необходимым), то повторное использование теста не имеет смысла. Результаты будут отражать лишь уровень научения заполнению данного инструмента с использованием полученных знаний о модели ситуационного руководства. Отсутствие достоверных данных об использовании теста «2Ф» в ситуации компьютерного тестирования приводит к выводу о том, что использование его в виде компьютерной психодиагностической программы при решении практических задач по оценке персонала преждевременно.

Данные исследования критериальной валидности свидетельствуют о том, что разработанный инструмент имеет явные преимущества и ограничения в разных оценочных ситуациях. Тест оказался низковалидным в ситуации обследования выборки испытуемых (N=19), сформированной после строго выполненного предварительного отбора кандидатов на должность регионального менеджера в фармацевтическую компанию. В этом случае не наблюдалось какой-либо связи показателей теста «2Ф» с внешним критерием (оценки по компетенциям), поскольку итоговое решение при выборе из достойных кандидатов принималось руководителем не только и не столько на основании профиля компетенций: кандидаты, проходившие Центр Оценки и заполнявшие тест «2Ф», уже на первых этапах отбора были выделены именно на основании их различий в проявлении компетенций. Кроме того, окончательный выбор кандидатов осуществлялся на основании «побочных» (с точки зрения измерения валидности) факторов – личных симпатий лица, принимающего решение о найме, отсутствия других более достойных кандидатов на рынке труда и т. п.

В ситуации первичной оценки кадрового резерва без предварительного отбора претендентов методика «2Ф» продемонстрировала результаты, хорошо согласующиеся с оценками экспертов. Наличие этих взаимосвязей дает основание считать методику пригодной для дополнительного изучения компетенций в экспертных процедурах оценки (при условии, что участники не были предварительно отобраны).

Особый интерес представляет результат проверки связи оценок по шкалам теста «2Ф» и отдельным компетенциям. Наиболее ценно,

что 4 из 5 шкал показали высокую корреляцию с компетенцией «Управление исполнением». По сути, это есть целевая компетенция, заключающая в себе принципы, лежащие в основе модели ситуационного руководства. Типовое определение данной компетенции, которое использовалось в указанных проектах: *управление исполнением – эффективное применение навыков постановки задачи, контроля и дачи обратной связи непосредственным подчиненным, организаторские способности.*

Именно эта рациональная сторона управленческих качеств, связанная с конкретным оперативным (тактическим) руководством непосредственными подчиненными, оказалась связана с большинством представленных шкал методики «2Ф». Такой результат можно считать также и свидетельством в пользу конструктивной валидности разработанной методики.

Исследование конвергентной конструктивной валидности теста «2Ф» содержит в себе большой материал для анализа и развития инструмента. Статистические данные табл. 2 показывают, что наибольшее количество значимых корреляций с показателями по другим методикам проявилось у субшкалы «Инструктирование». Важно, что выявлена отрицательная корреляция со шкалой «Команда» из тест-опросника ИДС-3, основанного на модели стилей управления Блейка–Муттона и являющегося в силу этого наиболее близким по содержанию и по принципу построения к тесту «2Ф» [6]. По нашему мнению, выявленная связь отражает близость конструктов, положенных в основу двух методик. Авторы ИДС предлагают трактовать высокие результаты по шкале «Команда» как готовность предоставлять подчиненным свободу и самостоятельность в работе, желание определять цели и алгоритм действий в совместной деятельности [6]. Очевидно, что такой подход находится в прямой оппозиции стилю «Инструктирование», предписывающему жесткую централизацию ответственности и полномочий. Учитывая, что стиль «Инструктирование» в модели ситуационного руководства является наиболее деструктивным с точки зрения последствий влияния на подчиненного, а высокие показатели по шкале «Команда» являются неким «идеальным результатом» в варианте Блейка–Муттона, следует признать содержательность и значимость данной связи.

Наличие близкой к значимой корреляции показателей по шкалам «Инструктирование»

и «Задача» (ИДС) ($r=0,18$ при $p=0,09$) предполагает возможность связи конструкторов двух этих наиболее близких по содержанию шкал. Как в модели Блэйка–Муттона [5], так и в модели Херси–Бланшара [22] эти два стиля отражают высокую ориентацию руководителя на задачу при низкой ориентации на отношения. Косвенным подтверждением этому выводу являются данные об отрицательной корреляции показателей по шкале «Инструктирование» со шкалой «Взаимоотношение» мотивационного опросника ТИМ [1]. Значимая отрицательная связь шкал «Инструктирование» и «Невмешательство» (ИДС), полученных при исследовании в ситуации экспертизы (фактически – наличие положительной связи со шкалой «Вмешательство»), отражает стремление руководителя подчеркнуть свою максимальную включенность в дела сотрудников даже при их низком уровне готовности к решению задачи. По нашему мнению, полученные результаты свидетельствуют в пользу конвергентной валидности субшкалы «Инструктирование».

Кроме того, данные табл. 2 свидетельствуют, что другие шкалы методики «2Ф» также имеют определенные связи с показателями обследования по профессиональным методиками, прошедшими психометрическую проверку. И, хотя систематический сбор данных по параллельным формам психодиагностических инструментов еще только начат, уже сейчас можно констатировать невысокую, но значимую согласованность данных, полученных при исследовании с помощью разработанных шкал, с надежными и валидными инструментами оценки руководителей.

В представленной матрице связей обращает на себя внимание отрицательная корреляция данных по шкале «Наставничество» (а также данных по шкале СУЭ, полученных в ситуации экспертизы) со шкалой «Творчество / Независимость» в мотивационном опроснике ТИМ [1]. Респонденты, которые значимым мотиватором называли «независимость», стремятся к личной свободе и следуют за предложенными обстоятельствами. В этом плане ситуационный подход, реализованный в тесте «2Ф», предполагает строгую зависимость действий руководителя от обстоятельств (в данном случае, от уровня готовности сотрудника). Сужение выбора действий формальными и достаточно простыми критериями, а также сама ситуация выбора из предложенных альтернатив явно ограничива-

ет деятельность тестируемых и приводит к низким результатам у нацеленных на независимость и творчество людей. Особенно явно эта ситуация проявляется в Центрах Оценки (в ситуации экспертизы), где заполнение испытуемыми Опросника ТИМ приобретает характер декларативно заявляемых мотивов. Характеристики этой связи, проявившейся на достаточно большой выборке ($N=140$), также подтверждают взаимосвязанность измеряемых в тесте параметров с компетенцией «Управление исполнением».

Отсутствие связи результатов теста с возрастом респондентов косвенно подтверждает дискриминативную валидность теста по отношению к классическим тестам интеллекта [4], а отсутствие связей с показателями личностным опросникам ТОП-ЮНИТ и ЭмIQ позволяет сделать вывод о «самостоятельности» инструмента и в отношении подхода, основанного на теории черт [19]. Такой результат показывает необходимость принимать во внимание рекомендации современных американских исследователей о необходимости тщательной проработки вариантов формулировки инструкции в контексте их ориентированности на знания или на предпочтения [26]. В упоминаемом исследовании было показано, что инструкция, ориентирующая на предпочтения, приводит к появлению значимых связей результатов ТОС с показателями по шкалам Большой пятерки. В нашем случае использовалась «знаниевая» инструкция, и связей с похожими по сути конструктами обнаружено не было. Данные о дискриминативной валидности еще раз подчеркивают правомерность выделения ТОС как самостоятельного подхода к оценке, отличного от типовых опросников и тестов.

Выводы

Тесты оценки ситуаций (ТОС) только в последние несколько лет стали приобретать статус самостоятельных инструментов в арсенале российских специалистов в области организационной психологии [9]. Сейчас очевидно, что значимость инструментов такого типа будет только возрастать из-за близости методологических посылок компетенционного подхода и технологии ТОС.

Технология создания ТОС с опорой на зарекомендовавшую себя в практике нормативную модель с узкой областью валидности имеет перспективы как с психометрической, так и с дидактической точки зрения и требует дальнейшей разработки.

Объединив данные исследования критериальной и конструктивной валидности, необходимо провести четкую границу для области валидности разработанной методики «2Ф». Проф. А.Г. Шмелев утверждает, что «тесты правильнее ... использовать не для отбора персонала (во всяком случае, не на последнем этапе отбора), а для отсева кандидатов на ранних стадиях. То есть информативность тестов не симметрична для разных полюсов тестовой шкалы. В случае высокого балла тест НЕ ПОЗВОЛЯЕТ спрогнозировать высокую эффективность деятельности (включается слишком много разных других факторов, способных ее снизить даже в тех случаях, если тест зафиксировал высокий уровень развития искомой способности или компетенции), но зато ПОЗВОЛЯЕТ более надежно спрогнозировать заведомо низкий уровень эффективности деятельности в случае явно низких результатов» [20].

Результаты нашего исследования подтверждают приведенные выше выводы, сделанные применительно к одному из наиболее известных российских ТОС – тесту ИТУПС. Действительно, при нацеленном на отбор персонала исследовании с первичной селекцией кандидатов, тест показал низкие результаты критериальной валидности. Однако при такого же рода исследованиях, ориентированных на диагностику карьерных перспектив «внутренних» сотрудников, характеризующихся не столь жестким регламентированием механизма номинации в кадровый резерв (руководители на местах имеют определенную свободу действий в отборе направляемых в Центры Оценки сотрудников), тест показал себя валидным.

Таким образом, основным назначением теста «2Ф» на основании приведенных данных о его надежности валидности видятся три основных направления его использования:

- первичная оценка и отсев непригодных кандидатов на начальные управленческие позиции;
- оценка кадрового резерва или так называемых HiPo (от англ. high potentials – высокопотенциальные <сотрудники>);
- обучение менеджеров начального уровня навыкам ситуационного руководства.

Литература

1. Алтухов, В.В. Тест индивидуальной мотивации (ТИМ): опыт разработки новой мотивационной методики / В.В. Алтухов //

Психологическая диагностика. – 2006. – № 2. – С. 24–40.

2. Батури, Н.А. *Технология разработки тестов: часть I* / Н.А. Батури, Н.Н. Мельникова // *Вестник ЮУрГУ. Серия «Психология»*. – 2009. – Вып. 6. – № 30(163) – С. 4–14.

3. Батури, Н.А. *Технология разработки тестов: часть II* / Н.А. Батури, Н.Н. Мельникова // *Вестник ЮУрГУ. Серия «Психология»*. – 2009. – Вып. 7. – 42(175) – С. 11–25.

4. Бурлачук, Л.Ф. *Психодиагностика: учеб. для вузов* / Л.Ф. Бурлачук. – 2-е изд. – СПб.: Питер, 2010. – 384 с.

5. Герберт, Д. *Организационная психология. Человек и организация: пер. с нем.* – Харьков: Изд-во Гуманитарный Центр / О.А. Шпилова, 2006. – 624 с.

6. Гребенюк, Г.А. *Построение и валидизация тест-опросника для диагностики деловых качеств* / Г.А. Гребенюк // *Психологический журнал*. – 1996. – Т. 17, № 1. – С. 149–156.

7. Кабаченко, Т.С. *Психология управления: учеб. пособие для высших учеб. заведений* / Т.С. Кабаченко. – М.: Педагогическое общество России, 2005. – 384 с.

8. Красностанова, М.В. *Assessment Center для руководителей. Опыт реализации в российской компании, упражнения, кейсы* / М.В. Красностанова, Н.В. Осетрова, Н.В. Самара. – М.; СПб.: Вершина, 2007. – 208 с.

9. Лурье, Е.В. *Психодиагностика на службе современных кадровых технологий* / Е.В. Лурье // *Инициативы XXI века*. – 2010. – № 1. – С. 69–73.

10. Лурье, Е.В. *Две психодиагностики российского бизнеса: обзор тестов и технологий* / Е.В. Лурье // *Ежегодник профессиональных рецензий и обзоров. Методики психологической диагностики и измерения* / под ред. Н. А. Батурина, Е. В. Эйдмана. – Челябинск: Изд-во ЮУрГУ, 2010. – Т. 1. – С. 234–261.

11. Моргунов, Е.Б. *Модели и методы управления персоналом. Российско-британское учеб. пособие* / под ред. Е.Б. Моргунова. – М.: Бизнес-школа «Интел-Синтез», 2001. – 464 с. – (Серия «Библиотека журнала «Управление персоналом»).

12. *Практический интеллект* / Р.Дж. Стернберг, Дж.Б. Форсайт, Дж. Хедланд и др. – 1-е изд. – СПб.: Питер, 2002. – 272 с.

13. Сергиенко, С.К. *Современные отечественные технологии оценки и развития управленческого персонала* / С.К. Сергиенко. – М.: Книжный мир, 2004. – 206 с.

14. Симоненко, С. Сказки и Были о методах оценки персонала / С. Симоненко, Д. Хренов. – М.: Изд-во "De-Tech", 2010. – 184 с.
15. Спенсер-мл., Л.М. Компетенции на работе: пер. с англ. / Л.М. Спенсер-мл., С.М. Спенсер. – М.: НИРО, 2005. – 384 с.
16. Фер, Р.М. Психометрика: Введение / Р. Майкл Фер, Верн Р. Бакарак; пер. с англ. А.С. Науменко, А.Ю. Попова; под ред. Н.А. Батурина, Е.В. Эйдмана. – Челябинск: Издательский центр ЮУрГУ, 2010. – 445 с.
17. Фуколова, Ю.В. Заговор тренеров / Ю.В. Фуколова // Секрет фирмы. – 2007. – № 27 (210).
18. Ван Хувейк, Р. Присущий вам стиль руководства: тест на гибкость и адекватность / Р. ван Хувейк // Директор школы. – 1994. – № 6. – С. 7–13.
19. Шмелев, А.Г. Психодиагностика личностных черт / А.Г. Шмелев. – СПб.: Речь, 2002. – 480 с.
20. Шмелев, А.Г. Анализ результатов прохождения ИТУПС членами Сообщества E-xecutive // HT.ru: сайт Лаборатории «Гуманитарные технологии». 2005. 10 марта. – <http://www.ht.ru/press/articles/print/art76.htm> (дата обращения: 07.06.2010).
21. Hersey, P. *Situational Leadership®. A Summary* / P. Hersey. – Center for Leadership Studies, Inc., 1993.
22. Hersey, P., Blanchard, K. H. (1995). *Behavioral theories of leadership*. In J. T. Wren (Ed.), *The leader's companion* (pp. 144–148). New York: Free Press (Reprinted from *Life cycle theory of leadership*, pp. 94–100, by P. Hersey & K. Blanchard, 1979, *Training and Development Journal*).
23. Lievens, F. *Situational judgment tests: a review of recent research* / F. Lievens, H. Peeters, E. Schollaert // *Personnel Review*. – 2008. – Vol. 37, № 4. – P. 426–441.
24. McClelland, D. *Testing for Competence Rather Than for "Intelligence"* / D. McClelland // *American Psychologist*. – 1973. – Vol. 28, № 1. – P. 1–14.
25. Obermann, C. *Assessment Center: Entwicklung, Durchführung, Trends. Mit originalen AC-übungen*. 4., vollständig überarb. u. erw. Aufl. Gabler, Betriebswirt.-Vlg, Wiesbaden, 2009. – 516 S.
26. *Predicting job performance using situational judgment tests: a clarification of the literature* / M.A. McDaniel, F.P. Morgeson, E.B. Finnegan et al. // *Journal of Applied Psychology*. – 2001. – Vol. 86, № 4. – P. 730–740.
27. Weekley, J.A. *Situational judgment tests: theory, measurement and application*. / J.A. Weekley, R. E. – Ployhart Lawrence Erlbaum Associates, Inc., 2006. – 400 p.

Поступила в редакцию 27 января 2011 г.

Лурье Евгений Валентинович. Руководитель центра разработки оценочных технологий ЗАО «ЭКОПСИ Консалтинг»; lurie.eugene@gmail.com

Eugene V. Lurie. Research & development manager at ECOPSY Consulting; lurie.eugene@gmail.com