



## Тесты ситуационных суждений: развитие подходов к пониманию оценки в анализе конкретной ситуации

БОРОВИКОВА Наталия Васильевна

БЕЛОУСКО Екатерина Валерьевна

Консалтинговая группа ВІТОВЕ, Санкт-Петербург, Россия

**Аннотация.** Цель статьи заключается в анализе и обобщении основных направлений в исследовании тестов ситуационных суждений. Актуальность обусловлена в первую очередь тем, что в российском дискурсе теоретическая основа тестов ситуационных суждений раскрыта неполно. Наряду с этим в последние несколько лет в зарубежной литературе произошло переосмысление тестов ситуационных суждений в том, как они разрабатываются и какие области знаний позволяют диагностировать. **Результаты.** Представлены взгляды на тесты ситуационных суждений с точки зрения их зависимости или независимости от контекста. Это позволило лучше осознать теоретическую основу, на которой разрабатываются тесты ситуационных суждений, и выделить два основных подхода — традиционный и конструктивный. Тесты ситуационных суждений, основанные на конструктах, по сравнению с традиционными имеют прочную теоретическую основу и являются более универсальными и адаптивными под различные профессии. Описаны ключевые характеристики тестов ситуационных суждений: формат заданий, тип инструкции и формат ответа. Определены области применения тестов и проанализированы основные особенности параметров качества тестов ситуационных суждений: внутренняя структура, надёжность, валидность, подверженность фальсификации и эффектам научения. **Ценность статьи** заключается в аналитическом обзоре существующих подходов к осмыслению тестов ситуационных суждений, которые позволяют обеспечить более глубокое понимание инструмента диагностики и могут стать основой для будущих исследований в отечественной науке. Тесты ситуационных суждений — один из самых недооценённых методов диагностики персонала в России. Широкие возможности их применения для различных управленческих задач и разнообразие форматов тестовых заданий открывают перспективы изменения практики оценки персонала в отечественных организациях. Тесты ситуационных суждений могут стать инструментом, который, с одной стороны, даёт обоснованные данные для принятия решений руководителями, а с другой — пользуется доверием сотрудников и вызывает у них интерес.

**Ключевые слова:** тесты ситуационных суждений, кейс-тесты, контекст-зависимые знания, контекст-независимые знания.

## Введение

Тесты ситуационных суждений впервые привлекли внимание научного сообщества благодаря своей способности предсказывать производительность труда. На протяжении многих лет они использовались в качестве инструмента отбора, который позволяет имитировать рабочие условия с низкой точностью и делать выводы о будущей успешности кандидатов. В этой парадигме тесты ситуационных суждений больше рассматривались как метод проверки процедурных знаний<sup>1</sup>. Впоследствии сфера их применения значительно расширилась. Сначала одни исследователи пытались выявить единый конструкт, который измеряют тесты ситуационных суждений, например практический интеллект и так называемые «неявные знания»<sup>2</sup>. Другие исследователи придерживались мнения, что тесты ситуационных суждений могут измерять несколько конструктов. Всё это позволило изменить взгляд на природу теста ситуационных суждений и посмотреть на него как на метод. Исследователи постепенно пришли к согласию, что тесты ситуационных суждений могут оценивать когнитивные способности, личностные особенности и межличностные феномены. На данный момент исследователи шагнули ещё дальше и смогли выделить тесты ситуационных суждений, которые не зависят от контекста ситуации.

Все эти открытия сопровождались изучением свойств тестов ситуационных суждений с учётом широких возможностей применения их как метода. Накал научных дискуссий вокруг тестов ситуационных суждений связан с тем, что несмотря на удобство метода, его очевидную валидность и вовлекающий формат, есть несколько факторов, влияющих на их достоверность. Существует множество исследований, которые рассматривают, каким образом подход к разработке содержания, тип инструкции, формат ответов и формат заданий влияют на достоверность диагностических выводов. Так, например, использование видеоформата, как правило, даёт более валидные оценки, а применение «знаниевых инструкций» помогает снизить эффекты, связанные с искажением результатов. В этом обзоре мы покажем различия в подходах, которые во многом определяют направленность в применении работодателями тестов ситуационных суждений.

## Определение теста ситуационных суждений

**Тест ситуационных суждений** (*Situational judgment tests, SJT*) — метод диагностики, основанный на моделировании реальной рабочей ситуации, содержащей некоторую проблему, которую участник должен проанализировать и выбрать из предложенного списка действий соответствующий подход к её решению (Lievens, Peeters, Schollaert, 2008; Lievens, de Soete, 2015). Встречаются различные вариации перевода *Situational judgment tests* на русский язык: «тест ситуативных суждений» (Личутина, 2017), «кейс-тест» (Шатров, Лурье, 2015; Шмелев, 2006), «тест оценки ситуации» (Лурье, 2011), «ситуационный тест» (Пескин, 2015). Метод основан на предположении, что на основе моделирования рабочих ситуаций можно предсказать, насколько хорошо человек будет выполнять работу (McDaniel, Nguyen, 2001). Пример ситуационного суждения приведён в таблице 1.

1 **Процедурные знания** (также иногда называемые практическими знаниями, императивными знаниями или перформативными знаниями) — это знания, полученные при выполнении некоторой задачи. В отличие от описательных знаний (также известных как «декларативные знания», «пропозициональные знания»), которые включают в себя знание конкретных фактов или утверждений, процедурное знание предполагает способность человека делать что-то. *Прим. ред.*

2 **Неявные знания** (англ. *tacit knowledge*) — подразумеваемые, имплицитные, не выраженные формально знания. Термин введён М. Полани в конце 1950-х гг. При этом он писал о процессе, а не о форме знания, однако его термин был использован для обозначения вида знаний, которое полностью или частично не эксплицированы (не формализованы). *Прим. ред.*

Таблица 1. Пример теста ситуационного суждения

<p>Вы являетесь руководителем финансового отдела крупного производственного предприятия, специализирующегося на производстве пластиковой тары. Перед вами стоит стратегическая задача по оптимизации структуры капитала предприятия, поскольку преобладание заёмных средств отрицательно сказывается на финансовой устойчивости компании. Проведённый анализ свидетельствует о наличии убыточного актива - цеха по производству ПЭТ преформ. Средства от продажи цеха могут погасить значительную часть кредитов, что способствует нормализации финансового состояния предприятия. В то же время, по некоторым прогнозам, эта производственная линия может быть прибыльна в будущем, однако вы в этом не уверены.</p> <p><b>Выберите наиболее эффективные действия, которые можно предпринять в этой ситуации.</b></p>	<p>1) Поднять цены на производимый в цехе товар для того, чтобы увеличить его прибыльность, не допустить закрытия цеха и увольнения людей.</p>
	<p>2) Постараться как можно дольше сохранять цех в числе активов компании, отложить момент принятия решения и внимательно следить за развитием событий.</p>
	<p>3) Продать цех без промедления для решения стратегической задачи и сохранения финансовой устойчивости компании.</p>
	<p>4) Заложить время на дополнительный анализ и принять решение продать цех в случае, если по его финансовым показателям не будет обнаружено прироста.</p>

### Ситуационное в тестах ситуационных суждений

Тесты ситуационных суждений (ТСС) принято рассматривать в качестве инструмента, который измеряет контекст-зависимые знания. Само понятие «ситуационное суждение» означает, что на ответ участника влияют условия ситуации. Однако недавние исследования показали, что ТСС могут измерять и контекстно-независимые знания (Lievens, Motowidlo, 2016; Krumm et al., 2015; Motowidlo, Crook, Kell, Naemi, 2009; Motowidlo, Hooper, Jackson, 2006), а значит, суждения в ТСС не всегда являются ситуационными (Krumm et al., 2014; Lievens, Motowidlo, 2016).

Контекстно-зависимые ТСС разрабатываются в русле традиционного подхода, где описание ситуации имеет решающее значение для определения правильности или неправильности вариантов ответов. Контекстуализация достигается за счёт подробного изложения аспектов работы, максимально близких к реальным. Кандидаты должны представить себе ситуацию и решить, как они будут вести себя в соответствии с представленными ситуационными обстоятельствами. Ситуация может варьироваться по степени своей реалистичности, детальности, сложности и понятности. Дж. Ратман, Р. Шерман и Д. Фандер считают, что есть три вида информации о ситуациях: сигналы, характеристики и типы (Rauthmann, Sherman, Funder, 2015):

- *сигналы* относятся к физическим стимулам в окружающей среде, они объективны и поддаются количественной оценке, но не являются значимыми для восприятия ситуации (например, количество присутствующих людей или имеющейся в помещении мебели);
- *характеристики* используются для описания ситуаций и являются менее объективными, чем сигналы, поскольку требуют интерпретации, часто определяются правилами и нормами поведения;
- *типы ситуаций* представляют собой совокупность сигналов и / или характеристик, встречающихся одновременно, они полезны для передачи широкой информации о ситуациях (например, ситуация «встреча с руководителем» передаёт информацию о том, как необходимо вести себя в этих обстоятельствах).

Сама реакция тестируемого на ситуацию зависит от того, как он воспринимает и интерпретирует совокупную информацию о ней. Модель решения тестов ситуационных суждений, предложенная М. Мартин-Раф и Х. Келлом, подтверждает это. По мнению авторов модели, решение тест ситуативного суждения представляет собой последовательный процесс. Отправной точкой является знакомство с ситуацией и формирование впечатления о ней. Затем тестируемый определяет для себя желаемый исход этой ситуации и выбирает действие, которое поможет достигнуть эту цель (Martin-Raugh, Kell, 2019). Интерпретации тестируемого во многом зависят от предыдущего опыта в сходных ситуациях, а также от индивидуальных особенностей восприятия. Считается, что без контекста ситуации тестируемый имеет недостаточное количество информации для решения задачи и не может определить правильный ответ (Rauthmann, Gallardo-Pujol et al., 2014).

Контекстно-независимые ТСС охватывают общие знания в предметной области, то есть некоторые общие правила эффективного поведения, применяемые в широком диапазоне ситуаций. Общие знания предметной области не получены из конкретного опыта работы, а отражают фундаментальные процессы социализации, такие как воспитание и образование (Corstjens, Lievens, Krumm, 2017). Исследования ТСС без описания ситуации показывают, что выбор способа действия определяется особенностями самой личности (Brown, Jones, Serfass, Sherman, 2016).

С. Крамм и его коллеги подтверждают, что ТСС можно разделить на контекстно-зависимые и контекстно-независимые. Они провели исследование с помощью командного теста *KSA Teamwork Test*<sup>3</sup>, разработанного М. Стивенсом и М. Кэмпбином (Stevens, Campion, 1996), поделив выборку на две группы. В одну вошли те, кто получил тест с описанием ситуации, в другую — те, кто получил тест без описания ситуации (см. Табл. 2). Авторы статьи пришли к выводу, что контекст ситуации оказывает меньшее влияние на решения тестируемых, чем ожидалось. Тестируемые смогли правильно решить 50–70% заданий без описания ситуации (Krumm et al., 2015).

Таблица 2. Пример заданий из исследования Стефана Крамма

Пример задания из теста Teamwork KSA Test с ситуацией	Пример задания из теста Teamwork KSA Test без описания ситуации
Предположим, вы спорили с несколькими коллегами о том, кто должен выполнить очень неприятную рутинную задачу. Что из перечисленного является наиболее эффективным способом разрешения этой ситуации? а) Поручить своему руководителю принять решение, потому что это позволит избежать любых личных предубеждений б) Составить сменный график, чтобы все выполняли рутинную работу ( <i>Правильный ответ</i> ) в) Позволить работникам, пришедшим раньше всех, выбирать в порядке очереди г) Назначить человека для выполнения задания случайным образом и не менять его	Решите для каждого вопроса, какой из вариантов ответа будет наиболее эффективным или отражать идеальное поведение. а) Поручить своему руководителю принять решение, потому что это позволит избежать любых личных предубеждений б) Составить сменный график, чтобы все выполняли рутинную работу ( <i>Правильный ответ</i> ) в) Позволить работникам, пришедшим раньше всех, выбирать в порядке очереди г) Назначить человека для выполнения задания случайным образом и не менять его

Результаты исследования под руководством С. Крамма вызвали бурную дискуссию и новую волну исследований (Kaminski, Felfe, Schäpers, Krumm, 2019; Schäpers et al., 2020; Lievens, Motowidlo, 2016). Некоторые специалисты предположили, что тестируемые всё равно стараются понять контекст ситуации на основании представленных вариантов действия

<sup>3</sup> Тест командной работы (*Teamwork Test*) предназначен для оценки знаний, навыков и способностей (knowledge, skills, ability, KSA) людей, желающих присоединиться к команде. *Прим. ред.*

(Melchers, Kleinmann, 2016; Harris et al., 2016). По их мнению, сами ответы могут нести в себе потенциальную информацию о классах ситуаций, для которых они предназначены (Krumm et al., 2015; Melchers, Kleinmann, 2016; Harris et al., 2016). Дискуссия возникла в первую очередь потому, что на данный момент отсутствует понимание того, как решаются ТСС вне контекста. Этот процесс ещё не описан. Построение модели мыслительных процессов позволит лучше понять, на что опираются тестируемые при выборе правильного ответа, когда описание ситуации отсутствует. В то же время это поможет открыть новые подходы к защите тестов ситуационных суждений от угадывания и искажения.

Несомненным преимуществом разделения тестов ситуационных суждений на контекстно-зависимые и контекстно-независимые является их возможность применения для разных задач. Формирование контекстно-независимых заданий требует сравнительно меньше ресурсов на описание ситуаций и разработку множества вариантов ответов. Они хорошо подходят для сотрудников, имеющих начальный опыт работы, и могут применяться для отбора кандидатов. Контекстно-зависимые, наоборот, будут полезны в ситуациях, когда сотрудники уже приобрели специфичные для работы знания (Corstjens, Lievens, Krumm, 2017), и основной задачей является оценка текущего уровня развития их компетенций.

## **Характеристики тестов ситуационных суждений**

М. Кэмпбелл, Р. Плойхарт и В. МакКинзи мл. провели анализ эмпирических исследований, опубликованных с 1990 года ( $n = 59$ ), и сформировали структуру характеристик тестов ситуационных суждений по двум направлениям: (1) разработка суждений и способов их оценивания; (2) применение ТСС (Campion, Ployhart, MacKenzie Jr., 2014). Авторы идентифицировали 15 атрибутов тестов ситуационных суждений (см. Рис. 1). Несмотря на высокую популярность тестов ситуационных суждений, наименее проработанными направлениями исследований являются следующие:

- разработка тестов на основе конструкторов,
- применение тестов в обучении и развитии,
- использование в квалификационном тестировании,
- влияние кросс-культурных факторов на результаты тестирования.

Рассмотрим ключевые атрибуты тестов ситуационных суждений.

### **Способ разработки ситуаций**

Известны два подхода к разработке тестов ситуационных суждений: традиционный, сформированный на основе критических инцидентов; конструктивный, основанный на теории. Если в рамках традиционного подхода исследователи фокусировались на интегральном показателе теста, то в последнее время всё больше научных работ направлено на изучение конкретных конструкторов. Исследования тестов ситуационных суждений в русле конструктивного подхода отмечаются как перспективные (Christian et al., 2010). Понимание конструкторов играет особую роль для интерпретации психометрических свойств теста, таких как надёжность, валидность.

#### ***Традиционный подход***

Традиционный подход к разработке ТСС основывается на предсказании будущей успешности человека в работе. В связи с этим описание ситуации в традиционном подходе имеет принципиальное значение. Тест разрабатывается на основе метода критических инцидентов. Эксперты в предметной области определяют набор ситуаций, с которыми сталкивается

человек в должности. Описания ситуаций достаточно контекстуализированы, описанные условия в задании влияют на конечный результат. Успешное решение рабочих дилемм, обычно, требует целый комплекс знаний, поэтому выделение конкретных конструктов затруднено. Задания и варианты ответов получаются достаточно разнородными. Это связано с тем, что целью конкретного задания является не измерение определённого конструкта, а фиксация реакции тестируемого на то или иное событие из рабочей жизни. Баллы, полученные за такой тест, коррелируют с будущей успешностью на работе, но имеют низкие показатели внутренней согласованности. Примерно 81,8% ( $n = 41$ ) тестов ситуационных суждений, описанных в научной литературе, разработаны с помощью метода критических инцидентов (Campion, Ployhart, MacKenzie Jr., 2014).



Рисунок 1. Характеристики тестов ситуационных суждений

Традиционный процесс разработки тестов ситуационных суждений представляет собой последовательность шагов:

- выделение критических инцидентов и описание ситуаций экспертами в предметной области;
- разработка вариантов действий на основе ответов эффективных и неэффективных сотрудников или опытных и новичков;
- разработка ключей для обработки на основе экспертного подхода или эмпирических данных (сравнение ответов эффективных и неэффективных сотрудников) (Tiffin et al., 2019);
- выбор типа инструкции и формата ответа.

**Конструктивный подход**

Этот подход предусматривает, что описание свойств теста осуществляется с точки зрения измеряемых конструктов, а не с точки зрения самого метода измерения (Christian, Edwards, Bradley, 2010). Выделение определённых конструктов помогает уменьшить неоднородность задания (Christian et al., 2010). Сам тест, благодаря наличию нескольких независимых факторов, будет иметь более высокие коэффициенты надёжности (Ployhart, Campion, MacKenzie Jr., 2012).

Если в традиционном подходе разработчиками ТСС должны выступать эксперты в предметной области, то в случае конструктивного подхода ими могут быть специалисты с психологическим образованием. Кейсы разрабатываются исходя из представления, в каких классах ситуаций проявляются те или иные конструкты. Целью конкретного задания является выявление одного конструкта. Описание ситуаций являются более универсальными и менее детализированными. Варианты ответа представляют собой разные степени проявленности определённого конструкта, а не качественно разные категории. Вариантов ответов в таком задании может быть несколько меньше в сравнении с традиционным подходом (Guenole, Chernyshenko, Weekly, 2017; Tiffin et al., 2019). Соответственно, процедура разработки тестов ситуационных суждений, основанных на конструктах, представляет собой последовательность шагов, где:

- разработка ситуаций и распределение их по шкалам в соответствии с признаком, который они измеряют;
- формулирование вариантов ответа;
- определение типа инструкции;
- разработка ключей для обработки;
- выбор формата ответа;
- формирование текстов-интерпретаций баллов.

П. Тиффин, Л. Петон, Д. О`Мара с коллегами провели сравнительную характеристику традиционного и конструктивного подходов к разработке тестов ситуационных суждений (Tiffin et al., 2019), результаты которой представлены в таблице 3. Можно соотнести каждый подход к разработке с контекстуализацией заданий. В традиционном подходе задания максимально зависимы от условий ситуации, а в конструктивном, наоборот, должны быть более универсальными.

Взгляд на конструкты эволюционировал за последние 20 лет. В 2000-х гг. некоторые исследователи высказывали уверенность в том, что существует единственный конструкт ситуативного суждения, который измеряют ТСС. Например, неявные знания (*tacit knowledge*), знания о работе (*job knowledge*), адаптивность (*adaptability*) и контекстуальное знание (*contextual knowledge*). Вместе с тем эти гипотезы не были подтверждены убедительными эмпирическими данными (Christian, Edwards, Bradley, 2010). Раннюю попытку по анализу конструктов в исследованиях предпринял М. МакДэниел и его коллеги. Они выявили, что ТСС измеряют такие конструкты, как когнитивные способности ( $Mr = .33 — .46$ ), дружелюбие ( $Mr = .27 — .31$ ), добросовестность ( $Mr = .25 — .31$ ), эмоциональную устойчивость ( $Mr = .26 — .30$ ), экстраверсию ( $Mr = .30$ ) и открытость ( $Mr = .13$ ) (McDaniel, Nguyen, 2001).

М. С. Кристиан, Б. Д. Эдвардс и Дж. Бредли-Гейст проанализировали 84 исследования по ТСС и пришли к выводу, что большинство исследований не содержат информации об измеренных конструктах или сообщают только интегральный балл по тесту. Исследования, содержащие такую информацию, описывают следующие конструкты: лидерские качества (38%), навыки межличностного общения (13%), навыки командной работы (4%), личностные особенности (10%) и профессиональные знания (3%). Авторы метаанализа предположили, что ТСС хорошо подходят для диагностики поведения в сложных межличностных и управленческих ситуациях и оценки лидерских навыков и навыков межличностного общения (Christian, Edwards, Bradley, 2010). За последние годы были также разработаны ТСС для измерения лидерских качеств, навыков работы в команде (Prewett et al., 2013), эмоционального интеллекта (Libbrecht, Lievens, 2012; Sharma, Gangopadhyay, Austin, Mandal, 2013), межличностных навыков студентов-медиков (Lievens, Sackett, 2012).

Таблица 3. Сравнение традиционного и конструктивного подходов к пониманию тестов ситуационных суждений

Параметры	Традиционный подход	Конструктивный подход
Разработка	<p>Разработка ситуаций под руководством экспертов в предметной области</p> <p>Ситуации и варианты ответов контекстуализированы и подробно описаны</p> <p>Разработка заданий основана на ситуациях, которые связаны с производительностью труда</p> <p>В вариантах ответов могут быть задействованы различные черты или способности</p> <p>Используются поведенческая и знаниевая инструкции: «Что вы будете делать?» и «Что необходимо сделать?»</p> <p>Качество определяется критериальной валидностью (корреляция с успешностью в работе)</p>	<p>Разработка ситуаций под руководством специалистов с психологическим образованием</p> <p>Ситуации и варианты ответов носят общий характер и менее контекстуализированы</p> <p>Разработка сценария основана на теории активации черт (<i>trait-activation theory</i>)</p> <p>Варианты ответа описывают различные градации одного конструкта</p> <p>Используются поведенческая и знаниевая инструкции: «Что вы будете делать?» и «Что необходимо сделать?»</p> <p>Качество определяется конструктивной валидностью (корреляция с другими инструментами оценки, измеряющими аналогичные конструкты)</p>
Свойства	<p>Ответы имеют тенденцию быть многомерными, задания — разнородными</p> <p>Низкая надёжность</p> <p>Факторная структура плохо поддаётся интерпретации</p> <p>Более низкая корреляция между баллами и личностными особенностями (особенно при знаниевой инструкции)</p> <p>Имеет более низкий риск фальсификации со стороны тестируемых (особенно при знаниевой инструкции)</p> <p>Сложнее провести параллель между разными формами тестов</p>	<p>Ответы имеют тенденцию быть одномерными задания — однородные</p> <p>Относительно высокая надёжность (например, альфа Кронбаха <math>\geq 0,70</math>)</p> <p>Тест имеет интерпретируемую факторную структуру</p> <p>Более высокие показатели корреляции между баллами и личностными особенностями</p> <p>Имеет более высокий риск фальсификации со стороны тестируемых</p> <p>Проще провести параллель между разными формами тестов</p>

С. Мотовидло для объяснения причины взаимосвязи тестов ситуационных суждений и личностных черт предположил, что ТСС могут измерять так называемые установки на черты (*implicit trait policies, ITP*). Установки на черты можно определить как набор неявных убеждений, установок на психологический конструкт. Это понятие сформировано на основе диспозиционной концепции, согласно которой люди склонны полагать, что присущее им поведение является более эффективным в межличностных ситуациях, и поэтому они выбирают соответствующий своим личностным чертам способ действия. Теория установок на черты объединяет в себе теорию черт, ситуационный подход и теорию соответствия личности и работы:

- *теория черт* предполагает, что человек обладает устойчивыми чертами личности, которые и определяют последовательные и закономерные способы реагирования на происходящие события (то, как человек чаще всего себя проявляет);
- *ситуационный подход*, наоборот, рассматривает реакцию на события как в большей степени определяемую условиями ситуации, нежели устойчивыми качествами личности (то, в каких классах ситуаций человек проявляет те или иные черты);
- *теория соответствия личности и работы* отстаивает позицию, согласно которой комбинация определённых личностных черт может определять успешность человека в определённых рабочих ролях.



Таким образом ТСС, основанные на этом подходе, направлены на измерение установок на черты, которые с наибольшей вероятностью предсказывают эффективность человека на рабочем месте (Motowidlo, Hooper, Jackson, 2006; Motowidlo, Hooper, Jackson, 2006b; Motowidlo, Lievens, Ghosh, 2018). В целом же, данный подход обладает следующими преимуществами:

- разработка ключей к тесту может осуществляться без привлечения экспертов в конкретной предметной области;
- применимость тестов ситуационных суждений может быть более универсальной для разных отраслей и компаний, чем ТСС, основанных на традиционном подходе, поскольку измеряют общие знания предметной области.

С. Мотовидло и его коллеги также считают, что ТСС, основанные на измерении установок на черты, могут быть хорошим инструментом при отборе сотрудников, не имеющих соответствующего профессионального опыта и конкретных знаний о работе (Oostrom, De Soete, Lievens; 2015; Motowidlo, Beier, 2010; Lievens, Motowidlo, 2015).

Таким образом, выбор подхода к разработке тестов ситуационных суждений во многом зависит от задач оценки и требуемого уровня детализации результатов. В русле традиционного подхода ТСС больше опираются на критические инциденты, поэтому разработчиками должны выступать в первую очередь эксперты в предметной области. Традиционные ТСС имеет ряд особенностей: высокая контекстуализация, неоднородная структура теста, низкие показатели внутренней согласованности, одношкальность теста. Основным критерием качества является связь общего балла по тесту и успешности тестируемого в работе. В русле конструктивного подхода ТСС опираются на теоретическую модель конструкторов, поэтому разработчиками могут являться специалисты с психологическим образованием. Отличительными особенностями ТСС на основе конструкторов являются: низкая контекстуализация, однородность структуры теста, удовлетворительные показатели надёжности, более высокая связь с личностными особенностями, возможность создавать тест с несколькими субтестами.

### **Формат задания**

Ситуация может быть представлена в нескольких форматах:

- текстовый,
- анимированный,
- видеоформат,
- иллюстрация.

Наиболее позитивно респонденты воспринимают видеоформат (Chan, Schmitt, 1997; Olson-Buchanan, Drasgow, 2006). Кроме того, ТСС с использованием видеоформата показывают более достоверные результаты. При таком формате варианты ответов могут быть интерактивно встроены в ролик в виде текста или представлены в письменной форме. Гибридные вариации заданий являются наиболее привлекательными, так как позволяют сократить расходы на разработку. Вместе с тем, согласно теории богатства медиавозможностей (*media richness theory*), достоверность тестов ситуационных суждений будет более высокой, если ситуация и варианты ответов будут представлены в едином формате (Schäpers et al., 2019).

Анимированные ролики являются альтернативой видеоформату и тексту, поскольку они дешевле в разработке, чем видео, но создают более реалистичную картину ситуации, чем текст. Тем не менее текстовый формат распространён больше, поскольку он самый дешёвый в разработке, и при необходимости тест легко актуализировать добавлением новых данных (Pollard, Cooper-Thomas, 2015; Lievens, Sackett, 2006; Weekley, Jones, 1997). Кроме этого, в текстовых заданиях намного проще убрать «шумовые» факторы, влияющие на «чистоту» измерения. Самый ненадёжный формат — иллюстрация, поскольку требует от иллюстра-

тора высокого мастерства в передаче ситуации через рисунок. Чаще иллюстрации используются в качестве дополнения к текстовому формату.

### Тип инструкции

Тип инструкции — ещё один атрибут тестов ситуационных суждений. М. МакДэниел и Н. Т. Нгуэн выделили две категории инструкций — знаниевую и поведенческую:

- инструкции, основанные на знаниях, предполагают, что респондент должен выбрать правильный / наилучший ответ или оценить эффективность ответов (McDaniel, Nguyen, 2001);
- инструкции, основанные на поведении, предполагают, что респондент должен выбрать ответ, который соответствует его вероятному поведению в описанной ситуации.

При этом знаниевые инструкции измеряют максимальную производительность, а поведенческие инструкции — типичную эффективность (Lievens, Sackett, Buysse, 2009). В таблице 4 показаны разные формулировки инструкций, представленные в анализе М. МакДэниела и его коллег (McDaniel, Hartman, Whetzel, Grubb III, 2007).

Таблица 4. Вариации формулировки разных типов инструкций

Тип инструкции	Пример
Поведенческая инструкция	Выбрать вероятный способ своего поведения ( <i>Would do</i> )
	Выбрать наиболее и наименее вероятный способ поведения ( <i>Most least likely</i> )
	Проранжировать варианты ответа по степени вероятности поведения ( <i>Rate and rank what you would most likely do</i> )
	Оценить вероятное поведение по шкале Р. Ликерта ( <i>Rate the tendency to perform on a Likert scale</i> )
Знаниевая инструкция	Выбрать, что необходимо делать ( <i>Should do</i> )
	Выбрать лучший ответ ( <i>Best response</i> )
	Выбрать лучший и худший ответы ( <i>Best worst clevenger</i> )
	Выбрать первые два лучших ответа ( <i>Best second best</i> )
	Выбрать первые три лучших ответа ( <i>Best, second, third best</i> )
	Проранжировать ответы по эффективности ( <i>Rate effectiveness</i> )
	Оценить по уровню важности ( <i>Level of importance corts</i> )

### Формат ответа

Важной характеристикой является и формат ответов на ТСС. Наиболее популярные из них:

- единственный ответ (*single-answer*);
- выбор «наилучшего» и «наихудшего» ответов (*dual-answer*);
- оценка всех вариантов ответов (*rate-all*);
- ранжирование всех вариантов ответов (*rank-all*);
- открытый ответ (*free response*).

Существует ряд исследований влияния формата ответа на достоверность и различие в группах (St-Sauveur, Girouard, Goyette, 2014; Arthur et al., 2014; Martin-Raugh et al., 2018; Motowidlo, Crook, Kell, Naemi, 2009; MacKenzie Jr., Ployhart, Weekley, Ehlers, 2010). Так, форматы ответов «ранжирование» и «оценка всех вариантов» считаются наиболее подходящими для использования: они показывают более высокие показатели согласованности и дифференцируемости тестируемых. Вместе с тем эти форматы предполагают удлинение процедуры тестирования за счёт того, что тестируемому необходимо больше времени на то, чтобы оценить варианты ответов и сравнить их.

### Способ разработки ключей

Существует несколько подходов к формированию ключей теста:

- *Теоретическая модель.* Правильные ответы определены в соответствии с теорией, или теория может использоваться для определения наилучших и худших вариантов действия. Методы оценки, основанные на теории, более распространены.
- *Экспертные оценки.* Экспертов просят оценить варианты ответов. Ответы, определённые как «лучшие», оцениваются как правильные, а ответы, определённые как «худшие», оцениваются как неправильные. При данном подходе обязательно должна быть использована проверка согласованности экспертных оценок.
- *Эмпирический подход.* Исследуется связь между ответами на тесты и внешним критерием, например, успешность. Ответы сотрудников, успешных по определённому показателю помечают как правильные, а ответы менее успешных сотрудников — неправильные (Campion, Ployhart, MacKenzie Jr., 2014). Возможна гибридная схема разработки ключей, объединяющая несколько подходов, например, экспертные оценки и эмпирический подход (Oostrom, De Soete, Lievens, 2015).

### **Область применения тестов ситуационных суждений**

Традиционно ТСС применяются в рамках процедуры отбора кандидатов. Аналогично другим методам, таким как интервью или Центры оценки, ТСС представляют собой метод измерения различных конструктов, связанных с работой. В статье С. Мотовидло с коллегами ТСС позиционировались в качестве имитации с низкой точностью. Под имитацией здесь подразумевалось моделирование психологических и физических аспектов работы. Низкая точность означает, что реакции респондента выражаются не в реальных поведенческих актах, а в мыслительных процессах. Центры оценки, напротив, требуют от респондентов демонстрации действий и поэтому считаются имитацией высокой точности. Основное сходство тестов ситуационных суждений и Центра оценки состоит в том, что оба метода основываются на принципе поведенческой согласованности и могут измерять множество конструктов. Тем не менее, Центры оценки являются дорогим инструментом отбора, поэтому применяются на более поздних этапах отбора или для высоких управленческих позиций (Motowidlo, Dunnett, Carter, 1990). Благодаря этому ТСС стали привлекательным инструментом отбора на ранних этапах. Чаще всего они используются для того, чтобы отсеять неподходящих кандидатов.

Отбор — не единственная область применения тестов ситуационных суждений. ТСС могут использоваться в качестве инструмента для определения потребности в обучении и развитии действующих сотрудников, в качестве инструмента для оценки результатов обучения или в качестве учебных материалов для тренировки навыков, связанных с работой, в безопасной среде (Hauenstein, Findlay, McDonald, 2010; Cox, Barron, Davis, de la Garza, 2017). Исследование эффективности тестов ситуационных суждений в обучающей среде показало, что ТСС способствуют лучшему запоминанию знаний и большей готовности применять полученные знания в новых ситуациях (Cox, Barron, Davis, de la Garza, 2017). Исследования — ещё одно направление для применения тестов ситуационных суждений. Если в основу вариантов ответов заложить ценностные ориентации тестируемых, то можно поделить выборку на разные категории для выработки стратегии взаимодействия с ними. Особенно полезным это может быть при формировании коммуникационных кампаний в организации.

Разработка тестов ситуационных суждений, как и других измерительных инструментов, представляет собой многогранный процесс с развилками. Первоначальным этапом разработки является определение области применения теста: отбор, обучение и развитие, исследования. Затем в соответствии с целями выбирается подход к разработке: традиционный или конструктивный. Подход обуславливает порядок формирования ситуаций и вариантов

ответов к ним, а также влияет на выбор подхода к формированию ключей (теоретический, экспертный и эмпирический). Разнообразие заданий обеспечивают типы инструкции (поведенческая или знаниевая), форматы ответов (единственный ответ, выбор «наилучшего» и «наихудшего» ответов, оценка всего вариантов, ранжирование, открытый ответ) и форматы заданий (текстовый формат, анимированный формат, видеоформат).

## Параметры качества тестов ситуационных суждений

Оценка параметров качества теста занимает значимую роль в изучении тестов ситуационных суждений. Эффективность ТСС оценивается на основе хорошо известных психометрических критериев: внутренняя структура, надёжность, валидность, искажение результатов тестирования и эффект научения.

### Внутренняя структура

За годы исследования произошла трансформация представления о тестах ситуационных суждений как о методе, который измеряет единственный конструкт, до метода, который оценивает множество конструктов, таких как когнитивные способности и личностные особенности. По мнению Д. Джексона, А. Лопилато и их коллег (Jackson et al., 2016), в научной литературе сложились три различных взгляда на внутреннюю структуру тестов ситуационных суждений.

Первая точка зрения рассматривает такие ТСС, как метод (*SJT-as-methods*), который не поддается изучению внутренней структуры теста. Оценка измеряемых конструктов должна осуществляться через конструктивную валидность, то есть через корреляцию результатов теста с другим тестом, который измеряет исследуемый конструкт (McDaniel, Whetzel, 2005; Whetzel, McDaniel, 2009). Согласно второй точке зрения, ТСС измеряют те конструкты, которые были заложены в них при проектировании теста (Christian et al., 2010). Таким образом, благодаря целенаправленной разработке можно заложить внутреннюю структуру теста. Согласно третьей точке зрения значительная часть дисперсии объясняется специфическими для ситуации факторами, при этом подразумевается, что ситуации представляют собой основной компонент интересующих конструктов, измеряемых ТСС.

Для определения внутренней структуры тестов ситуационных суждений чаще всего применялся факторный анализ. Однако в результате исследователи получали множество факторов, которые с трудом поддавались интерпретации (Lievens et al., 2008). Как следствие, они высказали предположение, что ТСС конструктивно разнородны на уровне задания, поскольку задание может быть направлено на измерение нескольких параметров производительности. Д. Ветцель и М. МакДэниел подчёркивают, что необходим поиск альтернативных подходов к изучению внутренней структуры теста, который позволит по-новому посмотреть на конструкты ТСС (Whetzel, McDaniel, 2009; Sorrel et al., 2016).

### Надёжность

Большинство исследований надёжности тестов ситуационных суждений были сосредоточены на изучении показателей внутренней согласованности, которые оказывались достаточно низкими. Например, в метаанализе В. Катано, А. Брочу и Ш. Ламерсон было получено среднее значение внутренней согласованности  $\alpha = 0,46$ , в метаанализе М. МакДэниела значения коэффициента  $\alpha$  варьировались от 0,43 до 0,94 (Catano, Brochu, Lamerson, 2012; McDaniel et al., 2001). Кроме этого, были выделены факторы, влияющие на коэффициент внутренней надёжности. Например, более длинные ТСС показывают более высокую согла-

сованность. Другим фактором стал тип задания. Р. Плойхарт и М. Эрхарт обнаружили, что задания, предполагающие оценку эффективности всех вариантов ответов, дают высокие показатели внутренней согласованности  $\alpha = 0,73$ , задания, требующие выбрать два варианта ответа (лучший и худший), показывают согласованность  $\alpha = 0,60$ , а задания с одним правильным ответом  $\alpha = 0,24$  (Ployhart, Ehrhart, 2003).

Низкая внутренняя согласованность тестов ситуационных суждений обусловлена их многомерностью, особенно если тесты разрабатывались на основе метода критических инцидентов. В свою очередь,  $\alpha$  Кронбаха требует, чтобы область построения была однородной, поэтому исследователи пришли к выводу, что показатели  $\alpha$  Кронбаха не подходят для оценки надёжности ТСС. По мнению Р. Плойхарта и его коллег, более высокую внутреннюю согласованность будут демонстрировать ТСС, разработанные на основе теории, поскольку их задания направлены на измерение конкретных факторов (Ployhart et al., 2012).

Лучшими метриками для тестов ситуационных суждений являются надёжность повторного тестирования и надёжность параллельных форм (Lievens et al., 2008; Whetzel, McDaniel, 2009). В целом о надёжности повторных тестирований сообщается редко. Тем не менее можно встретить данные, которые свидетельствуют о том, что повторное тестирование является удовлетворительной мерой надёжности ТСС. В 2012 г. Р. Плойхарт, М. Кэмптон и В. МакКинзи младший провели метаанализ восьми исследований и обнаружили, что средняя надёжность повторного тестирования составляет 0,61 (Campion, Ployhart, MacKenzie Jr., 2014). Р. Плойхарт и М. Эрахарт исследовали зависимость типа инструкции и надёжности повторного тестирования. Поведенческие инструкции показали более высокую надёжность повторного тестирования, чем знаниевые. Ж. Оостром, Б. Де Соете и Ф. Ливенс считают, что эти данные следует интерпретировать с осторожностью, поскольку анализы основывались на небольших выборках от 21 до 30 человек (Oostrom, De Soete, Lievens, 2015).

Исследование надёжности с помощью параллельных форм является более сложным способом проверки качества теста в связи с трудоёмкостью разработки альтернативных форм теста. Д. Ветцель и М. МакДениэл (Whetzel, McDaniel, 2009) приводят в пример два исследования надёжности альтернативных форм: одно принадлежит Ф. Ливенсу и П. Саккетту (Lievens, Sackett, 2007), другое — Ф. Освальду и его коллегам (Oswald et al., 2005). В работе Ф. Ливенса и П. Саккетта описаны три способа разработки альтернативных форм теста.

Первый способ опирается на стратегию случайного распределения: разрабатывается пул заданий ТСС, измеряющих какую-то конкретную область. Случайным образом эти задания распределяются между альтернативными формами теста. Второй способ опирается на стратегию изоморфизма ситуации: формируется несколько заданий из одного и того же критического инцидента. Третий способ опирается на стратегию изоморфизма задания: задания полностью подобны (они измеряют одну и ту же область, описывают один и тот же контекст ситуации и варианты действия), но отличаются друг от друга только формулировкой, описаны другими словами.

Ф. Ливенс и П. Саккетт рассчитали коэффициенты надёжности параллельных форм в ситуации отбора на реальных кандидатах, которые поступали в медицинский колледж. Корреляция между формами для стратегии случайного распределения составила 0.34 ( $N = 703$ ), для стратегии изоморфизма ситуации 0.56 ( $N = 1385$ ), для стратегии изоморфизма задания 0.68 ( $N = 1273$ ) (Whetzel, McDaniel, 2009; Lievens, Sackett, 2012).

Выделяется ещё один вид надёжности — надёжность частей теста. Однако данный тип встречается очень редко. В таблице 5 представлен сравнительный анализ разных типов надёжности. Несмотря на то, что большинство авторов уверены в том, что внутренняя согласованность — неуместный вид надёжности для оценки качества тестов ситуацион-

ных суждений, коэффициент  $\alpha$  Кронбаха остается самой популярной метрикой (Campion, Ployhart, MacKenzie Jr., 2014).

Таблица 5. Сравнение разных видов надёжности (Campion, Ployhart, MacKenzie Jr., 2014)

Тип надёжности	Исследования, n	%	M	Mdn	SD	Range	Min	Max
Коэффициент альфа Кронбаха ( <i>Coefficient alpha</i> )	129	88.4	.57 (.02)	.60	.20	.92	.00	.92
Повторное тестирование ( <i>Test-retest reliability</i> )	8	5.5	.61 (.09)	.66	.26	.72	.20	.92
Надёжность параллельных форм ( <i>Parallel-form Reliability</i> )	5	3.4	.52 (.10)	.57	.21	.54	.22	.76
Надёжность частей теста ( <i>Split-half reliability</i> )	4	2.7	.78 (.01)	.78	.03	.06	.76	.82

### Валидность

Ранее мы упоминали, что в рамках традиционного подхода ТСС разрабатывались на основе метода критических инцидентов, поэтому основным показателем качества теста является связь между результатами тестирования и успешностью сотрудника на работе.

Существует несколько предположений о том, почему же ТСС предсказывают успешность на работе. Первое объяснение — принцип поведенческой согласованности. Второе объяснение заключается в том, что ТСС измеряют намерения тестируемого. Третье — ТСС измеряют конструкты, которые являются надёжными предикторами выполнения работы, такие как: знания, практический интеллект, когнитивные способности или личностные черты. М. МакДэниел в 2001 г. провёл анализ исследований тестов ситуационных суждений (McDaniel, Nguyen, 2001). Скорректированная корреляция между ТСС и производительностью работы составила 0,34. При этом было обнаружено, что если ТСС разрабатывались на основе анализа работы, то они показывали более высокие показатели валидности (0,38 против 0,29) (McDaniel, Nguyen, 2001).

В метаанализе 2007 г. была показана инкрементная валидность<sup>4</sup> тестов ситуационных суждений (McDaniel et al., 2007). ТСС обеспечивают значимое инкрементное прогнозирование по «Большой пятёрке» (возрастающая достоверность от 0,06 до 0,07), умеренные инкрементные прогнозы когнитивных способностей (возрастающая достоверность от 0,03 до 0,05) и небольшую степень инкрементной достоверности по совокупности когнитивных способностей и «Большая пятёрка» (от 0,01 до 0,02) (McDaniel, Hartman, Whetzel, Grubb III, 2007). Исследования валидности тестов ситуационных суждений проводились в рамках образовательных программ, а именно при вступительных экзаменах. При этом было обнаружено, что ТСС в большей мере предсказывает будущие результаты, чем показывает текущие, в особенности те, которые связаны с межличностными отношениями. Таким образом, ТСС может определять потенциал человека (Tiffin, Paton, O'Mara, MacCann, Lang, Lievens, 2019).

Помимо связи с производительностью труда, ТСС продемонстрировали относительно высокую корреляцию с показателями когнитивных способностей или личностными чертами. М. МакДэниел и его коллеги сравнили баллы, полученные по двум типам инструкций тестов ситуационных суждений, с баллами, полученными по тестам когнитивных способностей и личностных особенностей. Они обнаружили, что ТСС с поведенческими инструкциями, как правило, больше коррелируют с личностью, а ТСС с знаниевыми инструкциями — с когнитивными способностями. Эти результаты позволяют сделать вывод о том, что можно изменить конструктивную валидность тестов ситуационных суждений, изменив тип инструкции (McDaniel, Hartman, Whetzel, Grubb III, 2007).

<sup>4</sup> **Инкрементная валидность** (англ. *incremental* — «приращение, прибыль») — один из компонентов критериальной валидности, прогностической валидности теста, отражающий практическую ценность методики при проведении отбора. Инкрементная валидность может быть выражена количественно при помощи коэффициента валидности. Показатель инкрементной валидности указывает на роль теста в улучшении отбора лиц для реальной деятельности, степень улучшения результативности процедуры отбора по сравнению с традиционной, основанной на анализе объективных сведений, документов, бесед, приёме с испытательным сроком. *Прим. ред.*

В исследовании М. С. Кристиана, Б. Д. Эдвардса и Дж. Бредли скорректированные коэффициенты корреляции с эффективностью работы варьировались от 0,19 до 0,43 в зависимости от измеряемых конструктов. ТСС, измеряющие профессиональные знания и навыки, имели самый низкий критерий валидности ( $r = 0,19$ ), а личностные особенности ( $r = 0,43$ ) и умение работать в команде ( $r = 0,38$ ) показали наиболее высокую связь с показателями работы (Christian, Edwards, Bradley, 2010). Можно отметить, что в целом ТСС имеют удовлетворительные показатели основных психометрических параметров. Накопленные данные свидетельствуют о том, что существует большой потенциал в применении новых психометрических моделей, которые позволят лучше изучить природу конструктов, измеряемых с помощью тестов ситуационных суждений (Sorrel et al., 2016).

### **Социальная желательность и искажение результатов**

Изучение склонности человека к социальной желательности и намеренному завышению тестовых баллов является важным вопросом в процедуре отбора. Такие эффекты могут приводить к тому, что компании будут брать на работу тех людей, которые в действительности не подходят по квалификации. Исследования по преднамеренному искажению результатов тестов ситуационных суждений проводились в большей степени в рамках академической среды, когда респондентов просили улучшать свои результаты немерено. Хорошо изученный фактор, который оказывает влияние на достоверность результатов, — тип инструкции. При поведенческой инструкции человека просят предсказать, как он будет действовать. Однако в большинстве случаев тестируемый выбирает не тот ответ, который соответствует его поведению, а тот ответ, который, по его мнению, более правильный и может принести ему больше баллов (Oostrom, Köbis, Rona, Cremers, 2017; Peeters, Lievens, 2005).

Знаниевые инструкции, наоборот, менее подвержены фальсификации со стороны респондента, поскольку знание подделать сложнее. Если человек не знает, как правильно действовать в этой ситуации, то ему труднее выбрать лучший ответ. В ситуации отбора, когда цена результатов теста высока, рекомендуется использовать именно знаниевую инструкцию, так как в любом случае тестируемые будут стремиться показать максимум того, на что они способны (Nguyen, Viderman, McDaniel, 2005). Недавнее исследование показало, что использование инструкций по типу ложного консенсуса позволяет повысить устойчивость теста к преднамеренному искажению результатов. Эффект ложного согласия — склонность проецировать свой опыт, убеждения и взгляды на других людей (Oostrom et al., 2017).

### **Эффект научения**

Испытание на этапе отбора привело к появлению целой индустрии, предлагающей услуги по обучению успешному прохождению тестов. Поэтому важной характеристикой качества теста является его устойчивость к эффектам научения. Эффект научения заключается в том, что тестируемые стремятся улучшить свои результаты за счёт использования различных стратегий, основанных на знании о тесте.

Следует отметить, что существует не так много публикаций на тему эффекта научения и тестов ситуационных суждений. Первые исследования проводились в лабораторных условиях, когда результаты не влияли на судьбу тестируемых. В сравнении с тестами способностей ТСС демонстрировали более устойчивые показатели к преднамеренному улучшению своего результата. Однако последние эмпирические свидетельства показали, что в ситуации реального отбора эффект научения может достигать высоких значений ( $SD = 0,5$ ). В связи с этим вопрос научения в тестах ситуационных суждений остается открытым и требует дальнейшего исследования (Cullen, Sackett, Lievens, 2006; Lievens, Buuse, Sackett, Connelly, 2012; Lievens, Buuse, Sackett, 2005).

Мы придерживаемся мнения, что необходимо иметь несколько альтернативных форм в ситуации, если результаты тестирования используются для принятия управленческих решений. Даже если ситуация касается конкретных знаний, в которой тестируемому сложнее улучшить свои результаты за счёт научения, он всегда может воспользоваться стратегиями, основанными на психологии. Понимание основных маркеров, по которым тестируемые определяют правильные ответы, помогут разрабатывать более качественные и устойчивые к эффектам научения тесты (Шмелев, 2013).

ТСС проходят обязательную процедуру проверки основных психометрических характеристик теста. Для каждого подхода при разработке ТСС есть наиболее подходящие критерии качества, которые должны учитываться исследователями. Основными показателями валидности ТСС являются критериальная валидность при традиционном подходе и конструктивная валидность при конструктивном подходе. Наиболее подходящими показателями надёжности являются надёжность повторного тестирования и надёжность параллельных форм. Оценка внутренней согласованности и проверка внутренней структуры теста уместны только в русле конструктивного подхода. Защита от искажения в тестах ситуационных суждений достигается за счёт типа инструкции: знаниевые — менее подвержены фальсификации, в то время как поведенческие чаще поддаются искажению.

## Заключение

Обзор современных исследований, посвящённых тестам ситуационных суждений, показывает, что в последние годы существует тенденция к переосмыслению этого инструмента оценки. Выделение контекстно-зависимых и контекстно-независимых заданий поставило под сомнение, насколько ТСС могут считаться методом ситуационной диагностики, если контекст не всегда имеет значение. Вместе с тем это позволило лучше осознать теоретическую основу, на которой разрабатываются ТСС, и выделить два основных подхода — традиционный и конструктивный. ТСС, основанные на конструктах, по сравнению с традиционными имеют прочную теоретическую основу и являются более универсальными и адаптивными под различные профессии. В свою очередь традиционные ТСС опираются на конкретные знания о том, как и что делать в определённых рабочих ситуациях.

Понимание преимуществ и ограничений обоих подходов поможет разработчикам проектировать подходящие решения по задачам оценки. В свою очередь, это требует запуска новой волны исследований, направленных на изучение влияния подхода на психометрические свойства теста. В целом вопросы валидности и надёжности тестов ситуационных суждений на данный момент остаются частично решёнными и требуют дальнейшей проработки. В связи с многомерной природой тестов ситуационных суждений традиционные методы определения валидности и надёжности, основанные на классической теории тестирования, являются неподходящими критериями качества для тестов ситуационных суждений. Например, факторный анализ показывает трудно интерпретируемые данные, коэффициенты внутренней согласованности  $\alpha$  Кронбаха имеют низкие значения. В исследовательском дискурсе назрела потребность в поиске альтернативных подходов к оценке психометрических параметров тестов, которые преодолеют эти накопленные ограничения. Другим важным направлением исследований может стать изучение процесса решения тестов ситуационных суждений при наличии и при отсутствии ситуации. Глубокое понимание модели решения заданий при отсутствии контекста позволит усилить ТСС в отношении искажений и угадывания правильных ответов.



ТСС — один из самых недооценённых методов диагностики персонала в России. Широкие возможности их применения для различных управленческих задач и разнообразие форматов тестовых заданий открывают перспективы изменения практики оценки персонала в отечественных организациях. ТСС могут стать инструментом, который, с одной стороны, даёт обоснованные данные для принятия решений руководителями, а с другой, — пользуется доверием сотрудников и вызывает у них интерес.

## Литература

- Личутина, Э. И. (2017). Тест ситуативных суждений как метод диагностики решений профессиональных проблем. *Экономика и менеджмент инновационных технологий*, 1, 158–161.
- Лурье, Е. В. (2011). Разработка теста оценки ситуаций на основе модели ситуационного руководства. *Психология. Психофизиология*, 5(222), 30–41.
- Люсин, Д. В., Овсянникова, В. В. (2013). Измерение способности к распознаванию эмоций с помощью видеотеста. *Психологический журнал*, 34(6), 82–94.
- Пескин, М. (2021). *Ситуационный тест «Сценарии»*. SHL.ru URL: <https://www.shl.ru/o-kompanii/novosti-i-sobitija/2021/situacionnij-test-scenarii>
- Шатров, Ю. И., Лурье, Е. В. (2015). О формате кейс-теста. В сб.: Н. А. Батулин и др. (ред.) *Современная психодиагностика России. Преодоление кризиса: сборник материалов III Всероссийской конференции по психологической диагностике. В 2-х т. Т. 1.* (331–335). Челябинск: Издательский центр ЮУрГУ.
- Шмелев, А. Г. (2006). Компьютеризированные кейс-тесты как инструмент профессиональной психодиагностики. В сб.: *Личностный ресурс субъекта труда в изменяющейся России*. Кисловодск.
- Шмелев, А. Г. (2013). *Практическая тестология: тестирование в образовании, прикладной психологии и управлении персоналом*. М.: Маска.
- Шмелев, А. Г. (2002). *Психодиагностика личностных черт*. СПб.: Речь.
- Arthur, W. Jr., Glaze, R. M., Jarrett, S. M., White, C. D., Schurig, I., Taylor, J. E. (2014). Comparative Evaluation of Three Situational Judgment Test Response Formats in Terms of Construct-Related Validity, Subgroup Differences, and Susceptibility to Response Distortion. *Journal of Applied Psychology*, 99(3), 535–545.
- Brown, N., Jones, A., Serfass, D., Sherman, R. (2016). Reinvigorating the Concept of a Situation in Situational Judgment Tests. *Industrial and Organizational Psychology*, 9(1), 38–42.
- Campion, M. C., Ployhart, R. E., MacKenzie Jr., W. I. (2014). The State of Research on Situational Judgment Tests: A Content Analysis and Directions for Future Research. *Human Performance*, 27(4), 283–310.
- Catano, V. M., Brochu, A., Lamerson, Ch. D. (2012). Assessing the Reliability of Situational Judgment Tests Used in High-Stakes Situations. *International Journal of Selection and Assessment*, 20(3), 333–346.
- Chan, D., Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, 82(1), 143–159.
- Christian, M. S., Edwards, B. D., Bradley, J. C. (2010). Situational Judgment Tests: Constructs Assessed and a Meta-Analysis of their Criterion-Related Validities. *Personnel Psychology*, 63, 83–117.
- Corstjens, J., Lievens, F., Krumm, S. (2017). Situational Judgment Tests for Selection. In: H. W. Goldstein, E. D. Pulakos, J. Passmore, C. Semedo (Eds.). *Handbook of the Psychology of Recruitment, Selection, and Employee Retention* (226–246). John Wiley & Sons Ltd.

- Cox, C. B., Barron, L. G., Davis, W., de la Garza, B. (2017). Using Situational Judgment Tests (SJTs) in Training: Development and Evaluation of a Structured, Low-Fidelity Scenario-Based Training Method. *Personnel Review*, 46(1), 36–45.
- Cullen, M., Sackett, P., Lievens, F. (2006). Threats to the Operational Use of Situational Judgment Tests in the College Admission Process. *International Journal of Selection and Assessment*, 14, 142–155.
- Guenole N., Chernyshenko O. S., Weekly J. (2017). On Designing Construct Driven Situational Judgment Tests: Some Preliminary Recommendations. *International Journal of Testing*, 17(3), 234–252.
- Harris, A. M., Siedor, L. E., Fan, Y., Listyg, B., Carter, N. T. (2016). In Defense of the Situation: An Interactionist Explanation for Performance on Situational Judgment Tests. *Industrial and Organizational Psychology*, 9, 23–28.
- Hauenstein, N., Findlay, R. A., McDonald, D. P. (2010). Using Situational Judgment Tests to Assess Training Effectiveness: Lessons Learned Evaluating Military Equal Opportunity Advisor Trainees. *Military Psychology*, 22, 262–281.
- Jackson, D. J. R., LoPilato, A. C., Hughes, D., Guenole, N., Shalfrooshan, A. (2016). The Internal Structure of Situational Judgement Tests Reflects Candidate Main Effects: Not Dimensions or Situations. *Journal of Occupational and Organizational Psychology*, 90, 1–27.
- Kaminski, K., Felfe, J., Schäpers, P., Krumm, S. (2019). A closer look at response options: Is Judgment in Situational Judgment Tests a Function of the Desirability of Response Options? *International Journal of Selection and Assessment*, 27, 72–82.
- Krumm, S., Lievens, F., Hüffmeier, J., Lipnevich, A. A., Bendels, H., Hertel, G. (2014). How “Situational” is Judgement in Situational Judgement Tests? *Journal of Applied Psychology*, 100, 399–416.
- Libbrecht, N., Lievens, F. (2012). Validity evidence for the situational judgment test paradigm in emotional intelligence measurement. *International Journal of Psychology*, 47(6), 438–447.
- Lievens, F., Buyse, T., Sackett, P. R. (2005). The Operational Validity of a Video-Based Situational Judgment Test for Medical College Admissions: Illustrating the Importance of Matching Predictor and Criterion Construct Domains. *Journal of Applied Psychology*, 90(3), 442–452.
- Lievens, F., De Soete, B. (2015). Situational Judgment Test. In J. D. Wright (ed.). *International Encyclopedia of the Social & Behavioral Sciences*. 2<sup>nd</sup> edition (13–19). Oxford: Elsevier.
- Lievens, F., Motowidlo, S. (2016). Situational Judgment Tests: From Measures of Situational Judgment to Measures of General Domain Knowledge. *Industrial and Organizational Psychology*, 9(1), 3–22.
- Lievens, F., Sackett, P. R. (2007). Situational judgment tests in high-stakes settings: Issues and strategies with generating alternate forms. *Journal of Applied Psychology*, 92(4), 1043–1055.
- Lievens, F., Sackett, P. R. (2006). Video-Based Versus Written Situational Judgment Tests: A Comparison in Terms of Predictive Validity. *Journal of Applied Psychology*, 91(5), 1181–1188.
- Lievens, F., Buyse, T., Sackett, P. R., Connelly, B. S. (2012). The Effects of Coaching on Situational Judgment Tests in High-Stakes Selection. *International Journal of Selection and Assessment*, 20(3), 272–282.
- Lievens, F., Peeters, H., Schollaert, E. (2008). Situational Judgment Tests: A Review of Recent Research. *Personnel Review*, 37(4), 426–441.
- Lievens, F., Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, 97, 460–468.
- Lievens, F., Sackett, P. R., Buyse, T. (2009). The Effects of Response Instructions on Situational Judgment Test Performance and Validity in a High-Stakes Context. *Journal of Applied Psychology*, 94(4), 1095–1101.
- MacKenzie Jr, W. I., Ployhart, R. E., Weekley, J. A., Ehlers, C. (2010). Contextual Effects on SJT Responses: An Examination of Construct Validity and Mean Differences Across Applicant and Incumbent Contexts. *Human Performance*, 23, 1–21.

- Martin-Raugh, M. P., Kell, H. J. (2019). A Process Model of Situational Judgment Test Responding. *Human Resource Management Review*, 100731.
- Martin-Raugh, M. P., Anguiano-Carrasco, C., Jackson, T., Brenneman, M. W., Carney, L., Barnwell, P., Kochert, J. (2018). Effects of Situational Judgment Test Format on Reliability and Validity. *International Journal of Testing*, 18(2), 135–154.
- McDaniel, M. A., Nguyen, N. T. (2001). Situational Judgment Tests: A Review of Practice and Constructs Assessed. *International Journal of Selection and Assessment*, 9, 103–113.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., Grubb III, W. L. (2007). Situational Judgment Tests, Response Instructions, and Validity: A Meta-Analysis. *Personnel Psychology*, 60, 63–91.
- Melchers, K. G., Kleinmann, M. (2016). Why Situational Judgment Is a Missing Component in the Theory of SJTs. *Industrial and Organizational Psychology*, 9(1), 29–34.
- Motowidlo, S. J., Beier, M. E. (2010). Differentiating Specific Job Knowledge from Implicit Trait Policies in Procedural Knowledge Measured by a Situational Judgment Test. *Journal of Applied Psychology*, 95, 321–333.
- Motowidlo, S. J., Crook, A. E., Kell, H. J., Naemi, B. (2009). Measuring Procedural Knowledge More Simply with a Single-Response Situational Judgment Test. *Journal of Business and Psychology*, 24(3), 281–288.
- Motowidlo, S. J., Hooper, A. C., Jackson, H. L. (2006). A Theoretical Basis for Situational Judgment Tests. In: J. A. Weekley, R. E. Ployhart (eds.). *Situational Judgment Tests: Theory, measurement, and application* (57–82). Mahwah, NJ: Erlbaum.
- Motowidlo, S. J., Hooper, A. C., Jackson, H. L. (2006b). Implicit Policies about Relations between Personality Traits and Behavioral Effectiveness in Situational Judgment Items. *Journal of Applied Psychology*, 91(4), 749–761.
- Motowidlo, S. J., Lievens, F., Ghosh, K. (2018). Prosocial Implicit Trait Policies Underlie Performance on Different Situational Judgment Tests with Interpersonal Content. *Human Performance*, 31(4), 238–254.
- Nguyen, N. T., Biderman, M. D., McDaniel, M. A. (2005). Effects of Response Instructions on Faking a Situational Judgment Test. *International Journal of Selection and Assessment*, 13(4), 250–260.
- Olson-Buchanan, J., Drasgow, F. (2006). Multimedia Situational Judgment Tests: The Medium Creates the Message. In J. A. Weekley, R. E. Ployhart (eds.). *SIOP organizational series. Situational judgment tests: Theory, measurement, and application* (253–278). Lawrence Erlbaum Associates Publishers.
- Oostrom, J. K., De Soete, B., Lievens, F. (2015). Situational Judgment Testing: A Review and Some New Developments. In I. Nikolaou, J. K. Oostrom (eds.). *Employee Recruitment, Selection, and Assessment: Contemporary Issues for Theory and Practice* (172–189). Routledge / Taylor & Francis Group.
- Oostrom, J. K., Köbis, N. C., Ronay, R., Cremers, M. (2017). False Consensus in Situational Judgment Tests: What Would Others Do? *Journal of Research in Personality*, 71, 33–45.
- Oswald, F. L., Friede, A. J., Schmitt, N., Kim, B. K., Ramsay, L. J. (2005). Extending a practical method for developing alternate test forms using independent sets of items. *Organizational Research Methods*, 8(2), 149–164.
- Peeters, H., Lievens, F. (2005). Situational Judgment Tests and Their Predictiveness of College Students' Success: The influence of faking. *Educational and Psychological Measurement*, 65(1), 70–89.
- Ployhart, R. E., Campion, M. C., MacKenzie Jr, W. I. (2012). *Reliability and Situational Judgment tests: A Review of the Literature*. Paper Presented at the 27th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

- Ployhart, R. E., Ehrhart, M. G. (2003). Be careful what you ask for: Effects of response instructions on the construct validity and reliability of situational judgment tests. *International Journal of Selection and Assessment*, 11(1), 1–16.
- Pollard, S., Cooper-Thomas, H. (2015). Best Practice Recommendations for Situational Judgment Tests. *Australasian Journal of Organisational Psychology*, 8, DOI: 10.1017/orp.2015.6.
- Prewett, M., Brannick, M., Peckler, B. (2013). Training teamwork in medicine: An active approach using role play and feedback. *Journal of Applied Social Psychology*, 43, 316–328.
- Rauthmann, J., Gallardo-Pujol, D., Hanes, E., Todd, E., Nave, C., Sherman, R., Ziegler, M., Cooper, A. B., Funder, D. (2014). The Situational Eight DIAMONDS: A Taxonomy of Major Dimensions of Situation Characteristics. *Journal of Personality and Social Psychology*, 107, 677–718.
- Rauthmann, J., Sherman, R., Funder, D. (2015). New Horizons in Research on Psychological Situations and Environments. Rejoinder to the Target Article and 29 Commentaries. *European Journal of Personality*, 29(3), 382–432.
- Schäpers, P., Freudenstein, J., Mussel, P., Lievens, F., Krumm, S. (2020). Effects of Situation Descriptions on the Construct-Related Validity of Construct-Driven Situational Judgment Tests. *Journal of Research in Personality*, 87, DOI: 10.1016/j.jrp.2020.103963.
- Schäpers, P., Lievens, F., Freudenstein, J., Hüffmeier, J., König, C. J., Krumm, S. (2019). Removing Situation Descriptions from Situational Judgment Test Items: Does the Impact Differ for Video-Based Versus Text-Based Formats? *Journal of Occupational and Organizational Psychology*, 93(2), 472–494.
- Sharma, S., Gangopadhyay, M., Austin, E., Mandal, M. (2013). Development and Validation of a Situational Judgment Test of Emotional Intelligence. *International Journal of Selection and Assessment*, 21(1), 57–73.
- Sorrel, M., Díaz, J., Abad, F., Torre, J., Aguado, D., Lievens, F. (2016). Validity and Reliability of Situational Judgement Test Scores: A New Approach Based on Cognitive Diagnosis Models. *Organizational Research Methods*, 19(3), 506–532.
- Stevens, M. J., Campion, M. A. (1996). Teamwork–KSA information guide. Arlington, VA: Vangent.
- St-Sauveur, C., Girouard, S., Goyette, V. (2014). Use of Situational Judgment Tests in Personnel Selection: Are the Different Methods for Scoring the Response Options Equivalent? *International Journal of Selection and Assessment*, 22(3), 225–239.
- Tiffin, P. A., Paton, L. W., O'Mara, D., MacCann, C., Lang, J. W. B., Lievens, F. (2019). Situational Judgement Tests for Selection: Traditional vs Construct-Driven Approaches. *Medical Education*, 54(2), 105–115.
- Weekley, J. A., Jones, C. (1997). Video-Based Situational Testing. *Personnel Psychology*, 50(1), 25–49.
- Whetzel, D. L., McDaniel, M. A. (2009). Situational Judgment Tests: An Overview of Current Research. *Human Resource Management Review*, 19(3), 188–202.

Поступила 22.12.2020



## Situational judgment tests: developing approaches to understanding case assessment

**Natalia BOROVIKOVA**

**Ekaterina BELOUSKO**

*Consulting group BITOBE, St. Petersburg, Russian Federation*

**Abstract.** *Purpose.* The purpose of the article is to analyze and summarize the main directions in the study of situational judgment tests. The relevance is primarily due to the fact that the theoretical basis of situational judgment tests is not fully disclosed in Russian discourse. Along with this, in the last few years in the foreign literature there has been a rethinking of situational judgment tests in how they are developed and what areas of knowledge allow diagnosing. *Findings.* The views on tests of situational judgments from the point of view of their dependence or independence from the context are presented. This made it possible to better understand the theoretical basis on which situational judgment tests are developed, and to single out two main approaches — traditional and constructive. Construct-based situational judgment tests, in comparison with traditional ones, have a solid theoretical basis and are more universal and adaptable for different professions. The key characteristics of situational judgment tests are described: the format of the tasks, the type of instruction and the format of the answer. The areas of application of the tests are determined and the main features of the quality parameters of situational judgment tests are analyzed: internal structure, reliability, validity, susceptibility to falsification and learning effects. *The value of the results.* The value of the article lies in the analytical review of existing approaches to the interpretation of situational judgment tests, which provide a deeper understanding of the diagnostic tool and can become the basis for future research in domestic science. Situational judgment tests are one of the most underrated methods of personnel diagnostics in Russia. The wide possibilities of their application for various management tasks and the variety of test task formats open up prospects for changing the practice of personnel assessment in domestic organizations. Situational judgment tests can be a tool that, on the one hand, provides informed data for decision-making by managers, and on the other hand, it enjoys the confidence of employees and arouses their interest.

**Key words:** situational judgment tests, case tests, context-dependent knowledge, context-independent knowledge.

### References

Arthur, W. Jr., Glaze, R. M., Jarrett, S. M., White, C. D., Schurig, I., Taylor, J. E. (2014). Comparative Evaluation of Three Situational Judgment Test Response Formats in Terms of Construct-Related Validity, Subgroup Differences, and Susceptibility to Response Distortion. *Journal of Applied Psychology*, 99(3), 535–545.

- Brown, N., Jones, A., Serfass, D., Sherman, R. (2016). Reinvigorating the Concept of a Situation in Situational Judgment Tests. *Industrial and Organizational Psychology*, 9(1), 38–42.
- Campion, M. C., Ployhart, R. E., MacKenzie Jr., W. I. (2014). The State of Research on Situational Judgment Tests: A Content Analysis and Directions for Future Research. *Human Performance*, 27(4), 283–310.
- Catano, V. M., Brochu, A., Lamerson, Ch. D. (2012). Assessing the Reliability of Situational Judgment Tests Used in High-Stakes Situations. *International Journal of Selection and Assessment*, 20(3), 333–346.
- Chan, D., Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology*, 82(1), 143–159.
- Christian, M. S., Edwards, B. D., Bradley, J. C. (2010). Situational Judgment Tests: Constructs Assessed and a Meta-Analysis of their Criterion-Related Validities. *Personnel Psychology*, 63, 83–117.
- Corstjens, J., Lievens, F., Krumm, S. (2017). Situational Judgement Tests for Selection. In: H. W. Goldstein, E. D. Pulakos, J. Passmore, C. Semedo (Eds.). *Handbook of the Psychology of Recruitment, Selection, and Employee Retention* (226–246). John Wiley & Sons Ltd.
- Cox, C. B., Barron, L. G., Davis, W., de la Garza, B. (2017). Using Situational Judgment Tests (SJTs) in Training: Development and Evaluation of a Structured, Low-Fidelity Scenario-Based Training Method. *Personnel Review*, 46(1), 36–45.
- Cullen, M., Sackett, P., Lievens, F. (2006). Threats to the Operational Use of Situational Judgment Tests in the College Admission Process. *International Journal of Selection and Assessment*, 14, 142–155.
- Guenole N., Chernyshenko O. S., Weekly J. (2017). On Designing Construct Driven Situational Judgment Tests: Some Preliminary Recommendations. *International Journal of Testing*, 17(3), 234–252.
- Harris, A. M., Siedor, L. E., Fan, Y., Listyg, B., Carter, N. T. (2016). In Defense of the Situation: An Interactionist Explanation for Performance on Situational Judgment Tests. *Industrial and Organizational Psychology*, 9, 23–28.
- Hauenstein, N., Findlay, R. A., McDonald, D. P. (2010). Using Situational Judgment Tests to Assess Training Effectiveness: Lessons Learned Evaluating Military Equal Opportunity Advisor Trainees. *Military Psychology*, 22, 262–281.
- Jackson, D. J. R., LoPilato, A. C., Hughes, D., Guenole, N., Shalfrooshan, A. (2016). The Internal Structure of Situational Judgement Tests Reflects Candidate Main Effects: Not Dimensions or Situations. *Journal of Occupational and Organizational Psychology*, 90, 1–27.
- Kaminski, K., Felfe, J., Schäpers, P., Krumm, S. (2019). A closer look at response options: Is Judgment in Situational Judgment Tests a Function of the Desirability of Response Options? *International Journal of Selection and Assessment*, 27, 72–82.
- Krumm, S., Lievens, F., Hüffmeier, J., Lipnevich, A. A., Bendels, H., Hertel, G. (2014). How “Situational” is Judgement in Situational Judgement Tests? *Journal of Applied Psychology*, 100, 399–416.
- Libbrecht, N., Lievens, F. (2012). Validity evidence for the situational judgment test paradigm in emotional intelligence measurement. *International Journal of Psychology*, 47(6), 438–447.
- Lichutina, E. I. (2017). Test situativnykh suzhdeniy kak metod diagnostiki resheniy professional'nykh problem [Situational judgment test as a method for diagnosing solutions to professional problems]. *Ekonomika i menedzhment innovatsionnykh tekhnologiy*, 1, 158–161.
- Lievens, F., Buyse, T., Sackett, P. R. (2005). The Operational Validity of a Video-Based Situational Judgment Test for Medical College Admissions: Illustrating the Importance of Matching Predictor and Criterion Construct Domains. *Journal of Applied Psychology*, 90(3), 442–452.
- Lievens, F., Buyse, T., Sackett, P. R., Connelly, B. S. (2012). The Effects of Coaching on Situational Judgment Tests in High-Stakes Selection. *International Journal of Selection and Assessment*, 20(3), 272–282.

- Lievens, F., De Soete, B. (2015). Situational Judgment Test. In J. D. Wright (ed.). *International Encyclopedia of the Social & Behavioral Sciences. 2nd edition* (13–19). Oxford: Elsevier.
- Lievens, F., Motowidlo, S. (2016). Situational Judgment Tests: From Measures of Situational Judgment to Measures of General Domain Knowledge. *Industrial and Organizational Psychology*, 9(1), 3–22.
- Lievens, F., Peeters, H., Schollaert, E. (2008). Situational Judgment Tests: A Review of Recent Research. *Personnel Review*, 37(4), 426–441.
- Lievens, F., Sackett, P. R. (2006). Video-Based Versus Written Situational Judgment Tests: A Comparison in Terms of Predictive Validity. *Journal of Applied Psychology*, 91(5), 1181–1188.
- Lievens, F., Sackett, P. R. (2007). Situational judgment tests in high-stakes settings: Issues and strategies with generating alternate forms. *Journal of Applied Psychology*, 92(4), 1043–1055.
- Lievens, F., Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, 97, 460–468.
- Lievens, F., Sackett, P. R., Buyse, T. (2009). The Effects of Response Instructions on Situational Judgment Test Performance and Validity in a High-Stakes Context. *Journal of Applied Psychology*, 94(4), 1095–1101.
- Lurie, Ye. V. (2011). Razrabotka testa otsenki situatsiy na osnove modeli situatsionnogo rukovodstva. Psikhologiya [Developing a situation assessment test based on a situational leadership model. Psychology]. *Psikhofiziologiya*, 5(222), 30–41.
- Lyusin, D. V., Ovsyannikova, V. V. (2013). Izmereniye sposobnosti k raspoznavaniyu emotsiy s pomoshch'yu videotesta [Measuring the ability to recognize emotions using a video test]. *Psikhologicheskiy zhurnal*, 34(6), 82–94.
- MacKenzie Jr., W. I., Ployhart, R. E., Weekley, J. A., Ehlers, C. (2010). Contextual Effects on SJT Responses: An Examination of Construct Validity and Mean Differences Across Applicant and Incumbent Contexts. *Human Performance*, 23, 1–21.
- Martin-Raugh, M. P., Anguiano-Carrsaco, C., Jackson, T., Brenneman, M. W., Carney, L., Barnwell, P., Kochert, J. (2018). Effects of Situational Judgment Test Format on Reliability and Validity. *International Journal of Testing*, 18(2), 135–154.
- Martin-Raugh, M. P., Kell, H. J. (2019). A Process Model of Situational Judgment Test Responding. *Human Resource Management Review*, 100731.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., Grubb III, W. L. (2007). Situational Judgment Tests, Response Instructions, and Validity: A Meta-Analysis. *Personnel Psychology*, 60, 63–91.
- McDaniel, M. A., Nguyen, N. T. (2001). Situational Judgment Tests: A Review of Practice and Constructs Assessed. *International Journal of Selection and Assessment*, 9, 103–113.
- Melchers, K. G., Kleinmann, M. (2016). Why Situational Judgment Is a Missing Component in the Theory of SJTs. *Industrial and Organizational Psychology*, 9(1), 29–34.
- Motowidlo, S. J., Beier, M. E. (2010). Differentiating Specific Job Knowledge from Implicit Trait Policies in Procedural Knowledge Measured by a Situational Judgment Test. *Journal of Applied Psychology*, 95, 321–333.
- Motowidlo, S. J., Crook, A. E., Kell, H. J., Naemi, B. (2009). Measuring Procedural Knowledge More Simply with a Single-Response Situational Judgment Test. *Journal of Business and Psychology*, 24(3), 281–288.
- Motowidlo, S. J., Hooper, A. C., Jackson, H. L. (2006). A Theoretical Basis for Situational Judgment Tests. In: J. A. Weekley, R. E. Ployhart (eds.). *Situational Judgment Tests: Theory, measurement, and application* (57–82). Mahwah, NJ: Erlbaum.
- Motowidlo, S. J., Hooper, A. C., Jackson, H. L. (2006b). Implicit Policies about Relations between Personality Traits and Behavioral Effectiveness in Situational Judgment Items. *Journal of Applied Psychology*, 91(4), 749–761.

- Motowidlo, S. J., Lievens, F., Ghosh, K. (2018). Prosocial Implicit Trait Policies Underlie Performance on Different Situational Judgment Tests with Interpersonal Content. *Human Performance*, 31(4), 238–254.
- Nguyen, N. T., Biderman, M. D., McDaniel, M. A. (2005). Effects of Response Instructions on Faking a Situational Judgment Test. *International Journal of Selection and Assessment*, 13(4), 250–260.
- Olson-Buchanan, J., Drasgow, F. (2006). Multimedia Situational Judgment Tests: The Medium Creates the Message. In J. A. Weekley, R. E. Ployhart (eds.). *SIOP organizational series. Situational judgment tests: Theory, measurement, and application* (253–278). Lawrence Erlbaum Associates Publishers.
- Oostrom, J. K., De Soete, B., Lievens, F. (2015). Situational Judgment Testing: A Review and Some New Developments. In I. Nikolaou, J. K. Oostrom (eds.). *Employee Recruitment, Selection, and Assessment: Contemporary Issues for Theory and Practice* (172–189). Routledge / Taylor & Francis Group.
- Oostrom, J. K., Köbis, N. C., Ronay, R., Cremers, M. (2017). False Consensus in Situational Judgment Tests: What Would Others Do? *Journal of Research in Personality*, 71, 33–45.
- Oswald, F. L., Friede, A. J., Schmitt, N., Kim, B. K., Ramsay, L. J. (2005). Extending a practical method for developing alternate test forms using independent sets of items. *Organizational Research Methods*, 8(2), 149–164.
- Peeters, H., Lievens, F. (2005). Situational Judgment Tests and Their Predictiveness of College Students' Success: The influence of faking. *Educational and Psychological Measurement*, 65(1), 70–89.
- Peskin, M. (2021). *Situatsionnyy test «Stsenarii»* [Situational test “Scenarios”]. SHL.ru. URL: <https://www.shl.ru/o-kompanii/novosti-i-sobitija/2021/situacionnij-test-scenarii>
- Ployhart, R. E., Campion, M. C., MacKenzie Jr., W. I. (2012). *Reliability and Situational Judgment tests: A Review of the Literature*. Paper Presented at the 27th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
- Ployhart, R. E., Ehrhart, M. G. (2003). Be careful what you ask for: Effects of response instructions on the construct validity and reliability of situational judgment tests. *International Journal of Selection and Assessment*, 11(1), 1–16.
- Pollard, S., Cooper-Thomas, H. (2015). Best Practice Recommendations for Situational Judgment Tests. *Australasian Journal of Organisational Psychology*, 8, DOI: 10.1017/orp.2015.6.
- Prewett, M., Brannick, M., Peckler, B. (2013). Training teamwork in medicine: An active approach using role play and feedback. *Journal of Applied Social Psychology*, 43, 316–328.
- Rauthmann, J., Gallardo-Pujol, D., Hanes, E., Todd, E., Nave, C., Sherman, R., Ziegler, M., Cooper, A. B., Funder, D. (2014). The Situational Eight DIAMONDS: A Taxonomy of Major Dimensions of Situation Characteristics. *Journal of Personality and Social Psychology*, 107, 677–718.
- Rauthmann, J., Sherman, R., Funder, D. (2015). New Horizons in Research on Psychological Situations and Environments. Rejoinder to the Target Article and 29 Commentaries. *European Journal of Personality*, 29(3), 382–432.
- Schäpers, P., Freudenstein, J., Mussel, P., Lievens, F., Krumm, S. (2020). Effects of Situation Descriptions on the Construct-Related Validity of Construct-Driven Situational Judgment Tests. *Journal of Research in Personality*, 87, DOI: 10.1016/j.jrp.2020.103963.
- Schäpers, P., Lievens, F., Freudenstein, J., Hüffmeier, J., König, C. J., Krumm, S. (2019). Removing Situation Descriptions from Situational Judgment Test Items: Does the Impact Differ for Video-Based Versus Text-Based Formats? *Journal of Occupational and Organizational Psychology*, 93(2), 472–494.
- Sharma, S., Gangopadhyay, M., Austin, E., Mandal, M. (2013). Development and Validation of a Situational Judgment Test of Emotional Intelligence. *International Journal of Selection and Assessment*, 21(1), 57–73.



- Shatrov, Yu. I., Lurie, Ye. V. (2015). O formate keys-testa [About the format of the case test]. In N. A. Baturin i dr. (Eds.) *Sovremennaya psikhodiagnostika Rossii. Preodoleniye krizisa: sbornik materialov III Vserossiyskoy konferentsii po psikhologicheskoy diagnostiki. V 2-kh t. T.1.* (331–335). Chelyabinsk: Izdatel'skiy tsentr YuUrGU.
- Shmelev, A. G. (2002). *Psikhodiagnostika lichnostnykh chert* [Psychodiagnostics of personality traits]. SPb.: Rech'.
- Shmelev, A. G. (2006). Komp'yuterizirovannyye keys-testy kak instrument professional'noy psikhodiagnostiki [Computerized case tests as a tool for professional psychodiagnostics]. In *Lichnostnyy resurs sub'yekta truda v izmenyayushchey Rossii*. Kislovodsk.
- Shmelev, A. G. (2013). *Prakticheskaya testologiya: testirovaniye v obrazovanii, prikladnoy psikhologii i upravlenii personalom* [Practical Testing: Testing in Education, Applied Psychology and Human Resource Management]. M.: Maska.
- Sorrel, M., Díaz, J., Abad, F., Torre, J., Aguado, D., Lievens, F. (2016). Validity and Reliability of Situational Judgement Test Scores: A New Approach Based on Cognitive Diagnosis Models. *Organizational Research Methods*, 19(3), 506–532.
- Stevens, M. J., Champion, M. A. (1996). *Teamwork–KSA information guide*. Arlington, VA: Vangent.
- St-Sauveur, C., Girouard, S., Goyette, V. (2014). Use of Situational Judgment Tests in Personnel Selection: Are the Different Methods for Scoring the Response Options Equivalent? *International Journal of Selection and Assessment*, 22(3), 225–239.
- Tiffin, P. A., Paton, L. W., O'Mara, D., MacCann, C., Lang, J. W. B., Lievens, F. (2019). Situational Judgment Tests for Selection: Traditional vs Construct-Driven Approaches. *Medical Education*, 54(2), 105–115.
- Weekley, J. A., Jones, C. (1997). Video-Based Situational Testing. *Personnel Psychology*, 50(1), 25–49.
- Whetzel, D. L., McDaniel, M. A. (2009). Situational Judgment Tests: An Overview of Current Research. *Human Resource Management Review*, 19(3), 188–202.

Received 22.12.2020