



БИЗНЕС ДЕЛАЮТ ЛЮДИ

Digital Assessment

Цифровая оценка
персонала

Научные открытия в оценке персонала

Вебинар 20.09.2023

1 приоритет

- Критериальная валидность инструментов оценки
- Инновационные инструменты оценки

2 приоритет

- Оптимальная длительность оценки
- Дискриминационное влияние оценки

3 приоритет

- Обман в интервью
- Точность performance review
- Эффект от повторного тестирования

4 приоритет

- Бизнес-эффект от оценки

Критические вопросы к метаанализу Шмидта (со стороны практиков)



У тестов способностей правда самая высокая валидность?



Почему неструктурированное интервью валиднее структурированного?



Почему у центра оценки такая невысокая валидность?

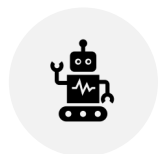
Инструмент	Валидность согласно метаанализу Шмидта и коллег (1998–2016)
Тесты способностей	0,65
Неструктурированные интервью	0,58
Структурированное интервью	0,58
Тесты знаний	0,48
Опросники благонадежности	0,46
Центры оценки и моделирующие упражнения	0,36
Опросники интересов	0,31
Кейс-тесты	0,26
Опыт работы (в годах)	0,16

Валидность инструментов оценки согласно новому метаанализу — Сакетта и коллег (2022–2023)

Инструмент	Данные Шмидта (1998–2016)	Данные Сакетта (2022–2023)	Характеристики инструмента, которые могут повысить критериальную валидность
Центры оценки и моделирующие упражнения	0,36	0,44 для руководителей 0,33 для всех категорий сотрудников	<ul style="list-style-type: none"> • Включение аналитических кейсов для оценки компетенций с интеллектуальной составляющей • Снижение когнитивной нагрузки на наблюдателей — до 6 компетенций в ЦО, до 3 компетенций в одном упражнении • Оценка не только по компетенциям, но и / или по задачам • Оценка по индикаторам, а не просто по компетенциям • Привлечение к оценке и психологов-оценщиков, и представителей бизнеса
Структурированное интервью	0,58	0,42	<ul style="list-style-type: none"> • Структурированность не только вопросов, но и расчета (выставления баллов) • Вопросы про прошлые ситуации, а не про поведение в гипотетической
Тесты знаний	0,48	0,4	<ul style="list-style-type: none"> • Включение в тест конкретных, а не общих тем
Тесты способностей	0,65	0,31	<ul style="list-style-type: none"> • Повышение содержательной валидности — включение к заданиям контента, соответствующего рабочему контексту участников
Опросники благонадежности	0,46	0,31	<ul style="list-style-type: none"> • Включение как прямых (overt), так и личностных вопросов
Личностные опросники	0,31	0,12–0,3 в зависимости от шкалы	<ul style="list-style-type: none"> • Добавление рабочего контекста в вопросы • Применение ситуативных опросников в ситуации отбора, а не нормативных • Использование IRT-расчетов
Кейс-тесты	0,26	0,26	<ul style="list-style-type: none"> • Использование ранжирования или «оценки всех» как формата ответов • Эмпирические (или смешанные) ключи, а не экспертные
Опросники интересов	0,31	0,24	<ul style="list-style-type: none"> • Добавление рабочего контекста в вопросы
Неструктурированные интервью	0,58	0,19	
Опыт работы (в годах)	0,16	0,07	

Источник(и): Schmidt & Hunter, 1998; Schmidt, 2016; Sackett et al, 2022; Sackett et al, 2023; Thorsteinson, 2018; Jackson & Kim, 2013; Wirz, 2013; Sanchez, 2017; Sackett et al, 2017; Dye, 1997; Jackson & Englert, 2011; Saldago, 2015

Какие инновационные инструменты были исследованы в наибольшей степени?



Автоматизированное
видеоинтервью

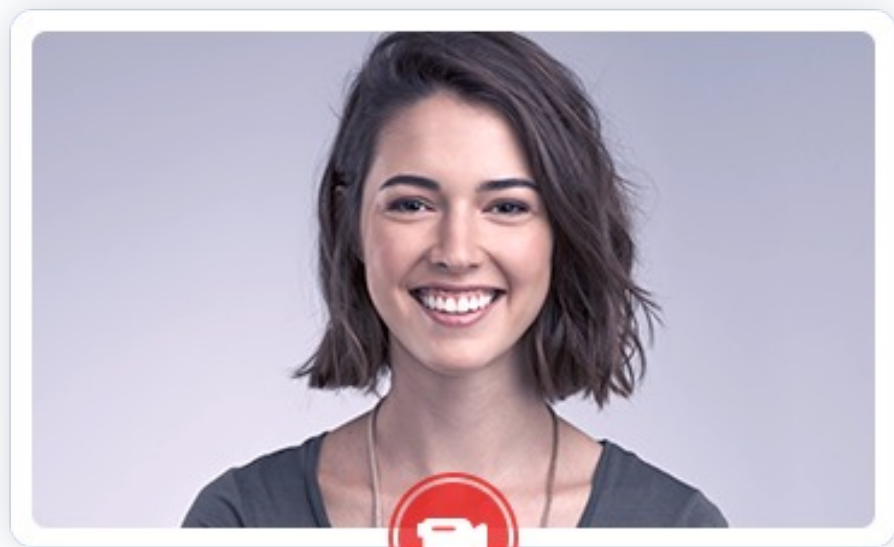


Game-based оценка в
разных проявлениях



Опросники,
использующие
машинное обучение

Что такое автоматизированное видеоинтервью?



01:42

Провайдеры за рубежом



Провайдеры в России



Echo

Sever.ai

Читайте номер HRTimes про искусственный интеллект в HR:

<https://www.ecopsy.ru/insights/iskusstvennyy-intellekt-zion-mozhet-spat-spokoyno/>



ИИ-интервью коррелирует с оценкой человеком на уровне 0,3–0,7 (в измерении личностных особенностей). Это хорошая точность



ИИ-интервью нравится людям меньше, чем интервью с человеком. Люди считают, что ИИ-интервью менее справедливо, менее точно и дает менее ценную обратную связь. Основная причина — люди думают, что меньше контролируют процесс и результат ИИ-интервью



Люди воспринимают ИИ-интервью более позитивно, если:

- У них есть опыт предвзятого отношения со стороны рекрутера / руководителя
- ИИ-интервью используется вверху воронки (для скрининга), но не посередине или на финале
- Они снабжаются *небольшим* количеством информации о том, что и как делает ИИ + аргументами, что ИИ-оценка более формализованная и менее предвзятая

Какая game-based оценка (ГВА) исследована в наибольшей степени?



Тесты способностей в игровой «оболочке»



Кейс-тесты с игровой фабулой

Оценщики раньше сомневались в том, что игры могут использоваться в оценке персонала

2015



Евгений Лурье

Лучше ПОКА не используйте игры для оценки, особенно для отбора, особенно компьютерные симуляции.

[Источник](#)

2016



Юрий Шатров

Геймификация в оценке персонала — это хорошо, но не сейчас.

[Источник](#)

2017



Менее 10% кандидатов хотели бы оцениваться с помощью геймифицированных инструментов.

[Источник](#)


2018





Чарльз Хэндлер

Игры — это весело, а оценка — нет.

[Источник](#)

- 
- GBA хорошо оценивает интеллект (корреляция с тестами способностей 0,4–0,9).
 - Средне — коммуникативные особенности (с личностными опросниками — 0,2–0,5)
 - Слабо — личностные особенности, основанные на установках (0–0,3)

- 
- ### Недостатки GBA
- GBA меньше нравится взрослым участникам или, как минимум, тем, у кого за плечами не много опыта видеоигр
 - Требовательны к устройству и Интернету участника

- 
- ### Преимущества GBA
- GBA больше нравится участникам (молодым — точно)
 - Вызывают у людей меньшее желание обманывать оценку (за счет вовлечения)
 - GBA обычно короче традиционных тестов

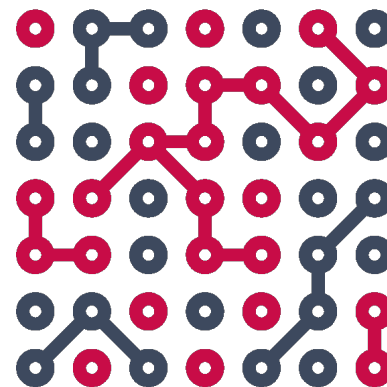
Основной вывод исследователей и практиков: GBA имеет право на жизнь, но не дает таких выгод, которых от неё изначально ожидали

Источник(и): Coovert, 2020; Auer, 2022; Ohlms, 2022; Georgiou & Lievens, 2022; Landers and Collmus, 2022; Ramos-Villagrasa et al, 2022

Как технологии машинного обучения могут повысить точность опросников? На примере delta.ai

УТВЕРЖДЕНИЕ	НАИМЕНЕЕ	НАИБОЛЕЕ
Я достигаю результата		
Я планирую свою работу		
Я выстраиваю отношения		
Я управляю подчинёнными		

Человек отвечает в опроснике, как он себя видит



Ответы человека сравниваются с базой данных сотрудников, кто ранее проходил тест.

Каждый из сотрудников в базе был оценён по компетенциям своим руководителем и коллегами.

КОМПЕТЕНЦИЯ	ОЦЕНКА
Дисциплинированность	67
Стремление к достижениям	21
Управление подразделением	41
Эффективная коммуникация	59

Алгоритмы машинного обучения прогнозируют, какую оценку по компетенциям получил бы сотрудник с таким паттерном ответов, если бы его оценивали руководитель и коллеги.

Как длительность инструментов влияет на валидность и на восприятия кандидатами?



В длительности оценки есть оптимум:

- Очень короткие инструменты — недостаточно точны
- Очень длинные инструменты уже не прибавляют в точности



Есть позиции на рынке кандидатов, на которых даже оптимальные с тз длительности инструменты (опросники больше 20 мин, интервью дольше 45 минут) — не стоит применять

Инструмент	Очень короткие	Очень длинные
Опросники	<10 мин	>45 мин
Тесты способностей	<15 мин	>45 мин
Интервью	<20 мин	> 60 мин



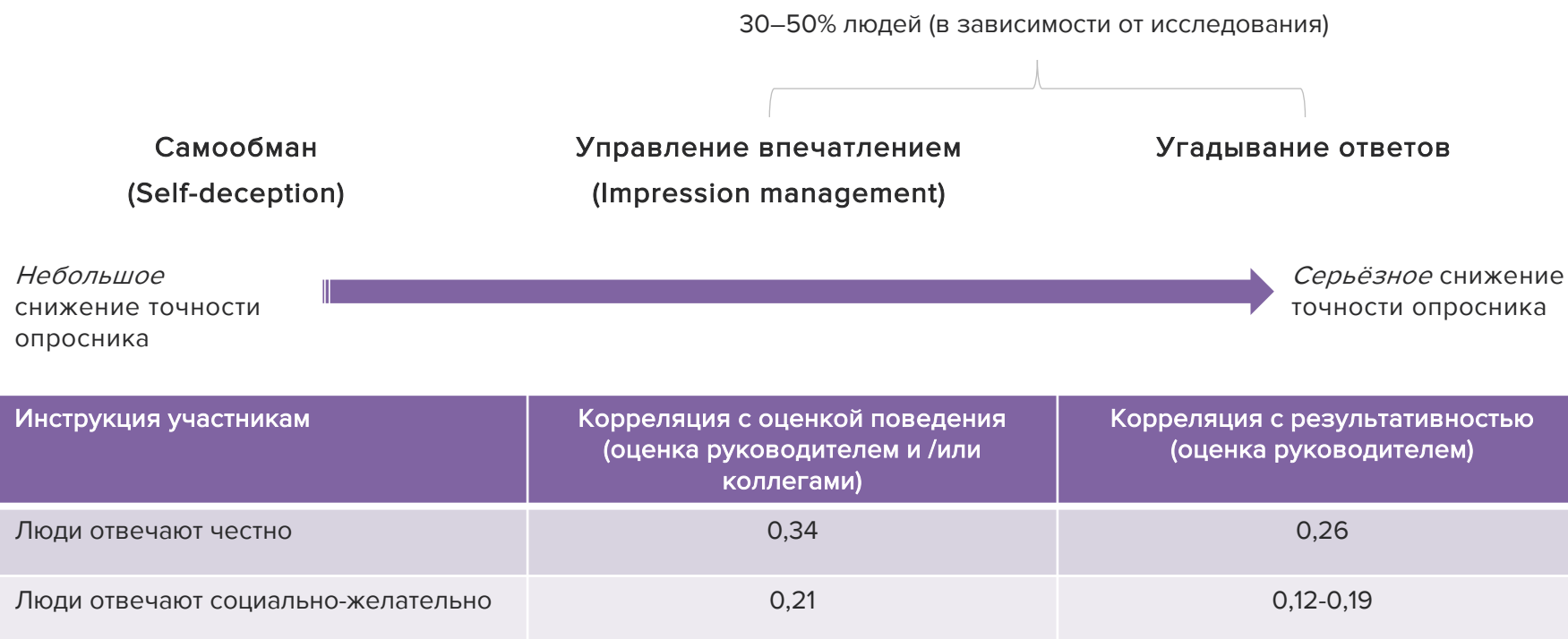
В остальном — если кандидат отказывается проходить долгую оценочную процедуру, это, в среднем, (а) кандидат более низкого качества; (б) может быть показателем того, что он сомневается в своих силах

Не стремитесь к очень короткой оценке


Как инструменты оценивают представителей разных групп?

Инструмент	Белые / темнокожие — кто получает, в среднем, более высокие баллы	Мужчины / женщины — кто получает, в среднем, более высокие баллы (при отборе на управленческие позиции)
Центры оценки и моделирующие упражнения	Белые ↑↑	Женщины ↑
Структурированное интервью	Белые ↑	Примерно одинаковые
Тесты знаний	Белые ↑↑	Нет данных
Тесты способностей	Белые ↑↑↑	Примерно одинаковые
Опросники благонадежности	Примерно одинаковые	Женщины ↑
Личностные опросники	Примерно одинаковые	Мужчины ↑
Кейс-тесты	Белые ↑↑	Смешанные результаты
Опросники интересов	Белые ↑↑	Нет данных
Неструктурированные интервью	Белые ↑↑	Примерно одинаковые
Опыт работы (в годах)	Белые ↑↑	Мужчины ↑


Источник(и): Sackett, 2022; Herde, 2019; Dean, 2008; Vanhove, 2022; Whetzel, 2008




Источник(и): Konstabel, Aavik, & Allik, 2006; Dilchert et al, 2006; Griffith, & Converse, 2021




В интервью слабо или сильно врет от 60 до 90% людей (в зависимости от исследования)



Незамеченный обман в интервью снижает точность оценки (на 0,03–0,15 в терминах критериальной валидности)



Интервьюерам сложно угадывать, соврал ли человек в конкретном месте — точность угадывания от 53 до 65%



Основные маркеры лжи оцениваемого:

- Меньше количество невербальных проявлений (движений тела и мимики)
- Меньшая демонстрируемая тревожность
- Частое повторение одних и тех же слов
- Дефицит деталей в рассказе
- Логические противоречия в рассказе
- Большое количество отрицаний в рассказе (формулировок от «не», «без»)



Замеры результативности имеют слабое отношение к реальной результативности — корреляция в районе 0,2–0,3



Людам не нравится процесс замера результативности. Еще меньше им нравятся вынужденные рейтинги, когда определенный процент людей точно нужно отнести в категорию неуспевающих

Основной вывод исследователей: performance management сломан как оценочная процедура. Если цель такой оценки — обратная связь и развитие, нужны другие инструменты (борее регулярные и не завязанные на вознаграждение и удержание)



Повторное тестирование (а) в среднем повышает баллы и (б) испытывает влияние регресса к среднему



Повторное тестирование способностей повышает точность оценки. Основная причина — люди уже знакомы с оценочной процедурой и демонстрирует способность, а не умение справляться с тревогой перед незнакомым тестом

Давайте людям ознакомительные версии тестов перед основной



Повторное тестирование опросников немного снижает точность оценки. В ситуации отбора люди дают более социально-желательные ответы (буквально пытаются их угадать)

Бизнес-эффект от оценки: что известно на данный момент

Как понималась зрелость оценочных практик (что связано)	Выборка	Критерий эффективности бизнеса (с чем связано)	Корреляция
Количество оценочных инструментов, имеющих, согласно исследованиям, наибольшую валидность	268 организаций в сфере коммерческих перевозок, Великобритания	Рентабельность	0,14
Количество людей, задействованных в процессе оценки; Количество соискателей на одну вакансию	194 нефтеперерабатывающих производств, США	Рентабельность	0,09
Количество оценочных инструментов, имеющих, согласно исследованиям, наибольшую валидность	890 организаций из разных отраслей и разных стран	Рентабельность	0,12
Соотношение откликов и нанятых кандидатов = избирательность компании в подборе и оценке	359 южнокорейских организаций из разных отраслей	Прибыль	0,11

Источник(и): Delery & Gupta, 2016; Wright et al, 1999; Subramony, 2009; Kim & Playhart, 2014

Как оставаться в курсе новостей в оценке персонала (её научной части)

Подпишитесь на самые важные журналы в мире организационной психологии. Новости о номерах вам будут проходить на почту. Ниже представлены издательства, их сайты и основные журналы

John Wiley & Sons	Taylor & Francis	American Psychological Association	Elsevier	Остальные издательства
<ul style="list-style-type: none">• International Journal of Assessment and Selection• Applied Psychology: An International Review• Personnel Psychology• Journal of Occupational and Organizational Psychology• Journal of Organizational Behavior	<ul style="list-style-type: none">• European Journal of Work and Organizational Psychology• Human Performance• International Journal of Testing	<ul style="list-style-type: none">• Journal of Applied Psychology• Consulting Psychology Journal: Practice and Research• Journal of Personality and Social Psychology	<ul style="list-style-type: none">• The Leadership Quarterly• Research in Organizational Behavior• Personality and Individual Differences	<ul style="list-style-type: none">• Industrial and Organizational Psychology (Cambridge University Press)• Annual Review of Organizational Psychology and Organizational Behavior• Personnel Assessment and Decisions (International Personnel Assessment Council)• Organizational Psychology Review (Sage)

Также есть журнал «Организационная психология», выпускаемый НИУ ВШЭ, на него тоже стоит подписаться