

Как измерить точность оценки персонала?

ЭКСПСИ



Ведущие вебинара



**ЮРИЙ
ШАТРОВ**

Руководитель практики
Digital Assessment



shatrov@ecopsy.ru



**ПАВЕЛ
ДЕГТЯРЁВ**

Директор по развитию продуктов
Автор и ведущий курса «Анализ
данных в HR» НИУ ВШЭ



degtyariov@ecopsy.ru

Российские компании **редко исследуют**, насколько точны их практики и инструменты оценки персонала

Компании проверяют точность ЦО в субъективных метриках (НК РЧК, 2014)



Компании хотят влиять на производительность, но не измеряют размер эффекта (НК РЧК, 2016)



55% компаний считает, что их практика оценки недостаточно обоснована с точки зрения науки и влияния на бизнес (ЭКОПСИ, 2016)

Только 7% выступлений на оценочных конференциях затрагивают тему точности оценки

Выборка — 172 презентации (2015–2019 гг). Исключены презентации провайдеров.

Что такое проверка точности инструмента оценки (валидизация)

Это измерение связи результатов оценки со значимыми внешними критериями*



Инструмент оценки

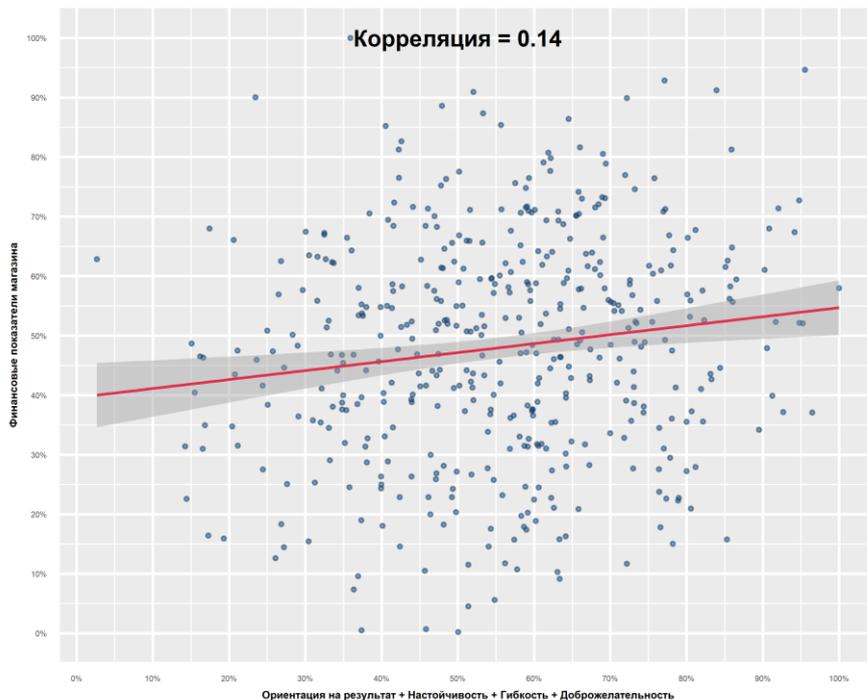


Внешний критерий

*Есть более 80 видов более валидности: понятие «замылилось» (Cambridge, 2014). В оценке самая важная валидная — критериальная. Подробнее: Principles for the Validation and Use of Personnel Selection Procedures, 2018.

Почему важно проверять точность оценки: эффективность

Кейс крупной компании в сфере розничной торговли: сколько денег дополнительно заработает компания, если начнёт отбирать директоров магазина по оценке с валидностью 0,14? Инструмент — вариация теста PiF.



Средняя выручка всех магазинов в год – **120 млрд. руб.**

Средняя рентабельность магазина – **25%**

Худшие директора по тесту дают в среднем на **4%** меньше выручки, чем обычные директора

Правило подбора будет работать только на новых сотрудников, таких – **15%** в год

Как изменится прибыль, если отфильтровать **10%** худших кандидатов на входе при помощи тестирования, отбирая участников с более высокими баллами:

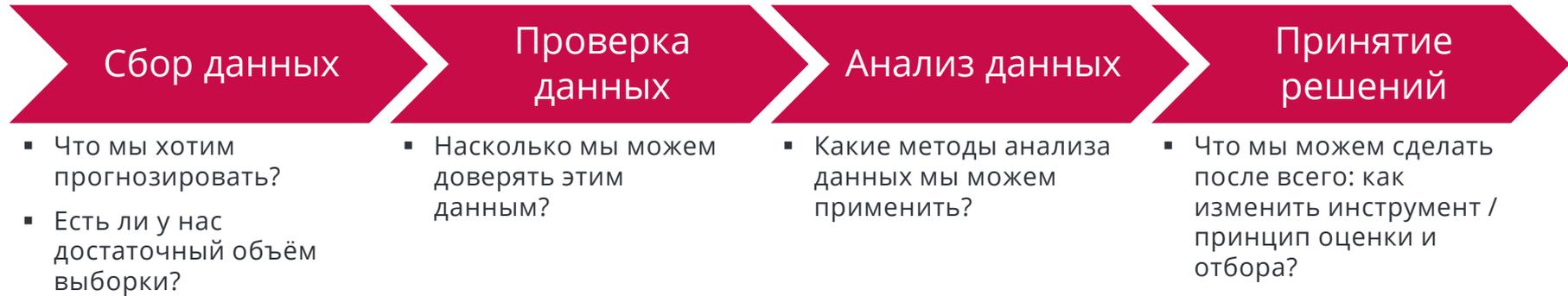
Увеличение прибыли = $120\,000\,000\,000 * 0,25 * 0,04 * 0,15 * 0,1 =$
18 000 000 рублей в год

Почему важно проверять точность оценки: срок работы

Кейс крупного банка: какие операционисты остаются в компании как минимум 3 месяца? Инструмент — опросник мотивации.

Какие сотрудник остаются?	Численное выражение зависимости	Примеры вопросов (все — с обратным ключом)
Неуверенные в себе	Уверенные в себе сотрудники имеют вероятность проработать долго на 27% меньшую , чем не уверенные	<ul style="list-style-type: none"> ▪ Я уверен, что смогу достичь всего, чего захочу ▪ Я знаю, что смогу быть успешен во всём, за что возьмусь
НЕ нацеленные на получение удовольствия от работы	Нацеленные на получение удовольствия от работы сотрудники имеют вероятность проработать долго на 19,6% меньшую , чем не нацеленные	<ul style="list-style-type: none"> ▪ Для меня важно иметь возможность непринуждённо общаться с коллегами по работе ▪ Мне важно, чтобы на работе порой можно было повеселиться
НЕ нацеленные на профессиональное развитие	Нацеленные на профессиональное развитие сотрудники имеют вероятность проработать долго на 18,2% меньшую , чем не нацеленные	<ul style="list-style-type: none"> ▪ Мне хотелось бы заниматься делами, требующими самоотдачи ▪ Моя работа должна быть достаточно сложной
НЕ нацеленные на гармоничные отношения в коллективе	Нацеленные на гармоничные отношения в коллективе сотрудники имеют вероятность проработать долго на 14,6% меньшую , чем не нацеленные	<ul style="list-style-type: none"> ▪ Для меня важно избегать разногласий на работе ▪ Для меня важно относиться к другим людям дружелюбно ▪ Мне кажется важным избегать конфликтов

Этапы проверки точности: вопросы, на которые предстоит ответить



Что мы хотим прогнозировать: внешние критерии для проверки точности оценки и сценарии их применения

Критерий	Сценарий: зачем прогнозировать этот критерий
 <p>Эффективность работы — найти качества (компетенции / способности / ценности / черты) эффективного сотрудника и оценивать всех сотрудников по ним</p>	<p>Во всех HR-процессах:</p> <ol style="list-style-type: none"> 1. Подбор — на входе выбирать лучших 2. Регулярная оценка — давать обратную связь по критериям, которые действительно значимы 3. Развитие — определять зоны развития обучать сотрудников навыкам и качествам, которые будут полезны им и Компании 4. Управление талантами — продвигать только лучших
 <p>Срок работы — понять, какие люди остаются с нами надолго и начать отбирать их</p>	<p>Подбор — чтобы обеспечить кадровую защищённость компании</p>
 <p>Уровень должности / скорость карьерного продвижения — выявить компоненты роста, чтобы видеть самых потенциальных</p>	<p>В двух процессах:</p> <ol style="list-style-type: none"> 1. Подбор — отбирать потенциальных кандидатов на управленческие должности 2. Управление талантами — продвигать на управленческие должности сотрудников, которые будут в них наиболее успешными

Есть ли у нас достаточный объём выборки?

Есть много факторов, которые влияют на необходимый размер выборки: уровень значимости, статистическая мощность, репрезентативность (предельная ошибка выборки), стабилизация факторных нагрузок. Мы упростили их все до трёх значений.

Минимум:

100 человек на одну должность

Уже лучше:

500 человек

Больше не требуется:

1 000 человек

Насколько мы можем доверять данным? KPI vs. Оценка руководителя

ПРИ ОЦЕНКЕ РЕЗУЛЬТАТИВНОСТИ ПО ЖЕСТКОЙ ФОРМУЛЕ



КАК ДОЛЖНО БЫТЬ

- План по KPI выполняется стабильно из года в год (стабильность «хороший-плохой» — **60-70%**)
- Результаты по KPI в основном зависят от сотрудника



КАК ЕСТЬ

- На выполнение плана по KPI во многом влияют случайные факторные, а не закономерности
- Совпадение на уровне **20-30%**
- Результаты по KPI зависят от сотрудника в среднем **на 20%**

Как с этим бороться:

- Считать не % KPI, а отклонение от среднего % по аналогичным должностям в тот же период
- Проверять конкретные KPI на нестабильность — корреляция одного KPI по одному сотруднику за несколько периодов (6-12 мес.) должна быть от 0.5

ПРИ МЕНЕДЖЕРСКОЙ (РУЧНОЙ) ОЦЕНКЕ РЕЗУЛЬТАТИВНОСТИ



КАК ДОЛЖНО БЫТЬ

- Руководитель ставит итоговую оценку основываясь на всех результатах, достигнутых в течение периода — **целостный временной горизонт оценки**
- Руководитель пытается объективно оценить эффективность



КАК ЕСТЬ

- На оценку сотрудника гораздо больше влияют результаты последних месяцев — **близорукость планирования**
- Руководитель пытается не только объективно оценить, **но и удержать сотрудника** (завышение оценок, везде «4» и «5»)

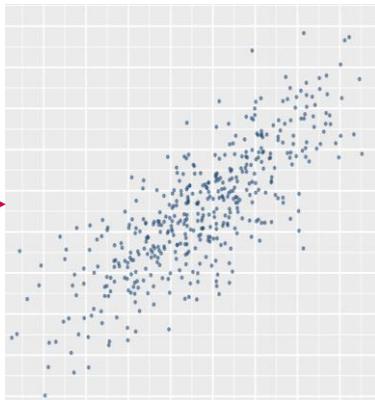
Как с этим бороться:

- Собирать оценки руководителей чаще раза в год и брать среднее от них
- Ранжировать оценки разных руководителей
- Верифицировать оценки (кадровый комитет, оценка через уровень)

Какие методы анализа мы можем применить?

Корреляция

Оценка	Эффективность
80%	100%
60%	80%
100%	120%
70%	90%



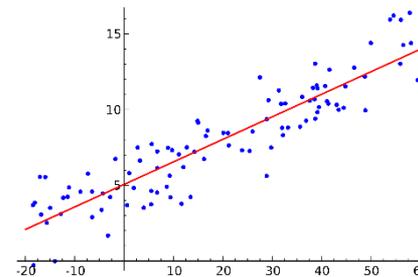
Это мера статистической связи между двумя параметрами. Показывает, насколько изменения в результатах оценки связаны с изменением во внешнем критерии.

Сценарий использования:

1. Нужно проверить один параметр
2. Данные не в числе

Регрессия

Оценка_1	Оценка_2	Эффективность
82%	38%	100%
61%	42%	80%
103%	46%	120%
67%	34%	90%



$$\text{Эффективность} = 0\% + 1 * \text{Оценка}_1 + 0,5 * \text{Оценка}_2$$

Регрессия (линейная) — модель, которая предполагает, что связь между несколькими результатами оценки и одним внешним критерием можно отобразить как прямую. Показывает, как необходимо сложить несколько результатов оценки, чтобы получить максимально близкое к целевой переменной число.

Сценарий использования:

1. Нужно проверить несколько параметров
2. Данные в числе

Какие показатели связи можно считать достаточными?

Не верьте метаанализам

ВАЖНО: в зарубежных исследованиях всегда приводятся скорректированные показатели связи (коррекции — специальные поправки, которые учитывают неточность внешнего критерия). Тесты способностей валидны на уровне 0,65 только после коррекций (Schmidt, Oh, & Shaffer, 2013). Связь тестов способностей с эффективностью БЕЗ коррекции — 0,26 (Schmidt, Shaffer, and Oh, 2008). Коррекции имеют научное значение, наша позиция — не применять их в оценке персонала.

На самом деле все инструменты оценки прогнозируют результативность с точностью 0,1–0,4 (без коррекций).

Центр оценки	Тесты способностей	Структурированное интервью	Кейс-тесты (SJT)	Неструктурированное интервью
0,28	0,26	0,24	0,24	0,18

Источники: Arthur, Day, McNelly, & Edens, 2003; McDaniel, Hartman, Whetzel, & Grubb III, 2007; Schmidt, Shaffer, & Oh, 2008; McDaniel, Whetzel, Schmidt, & Maurer, 1994.

Какие показатели связи можно считать достаточными?

Выборка 100 человек	500 человек	1 000 человек
0,2	0,12	0,1

Что такое точный инструмент оценки: полная картина



Инструмент оценки

- Точность критериев
- Точность инструмента
- Точность эксперта (для интервью и центра оценки)



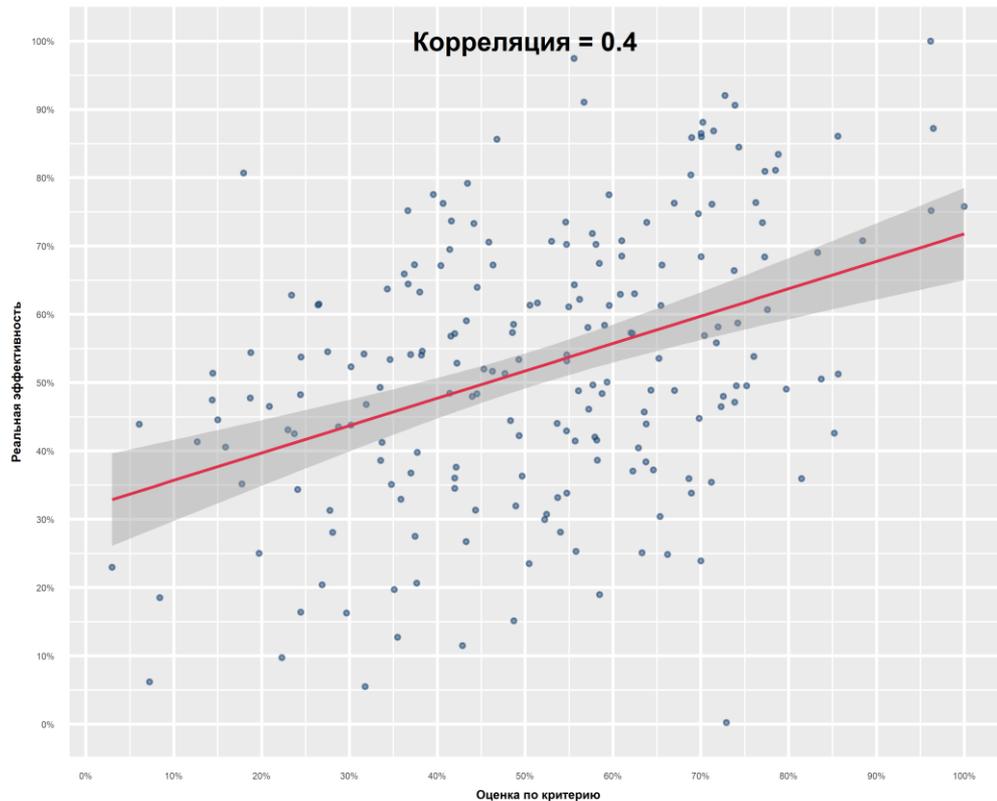
- Достаточная выборка
- Адекватная статистическая процедура



Внешний критерий

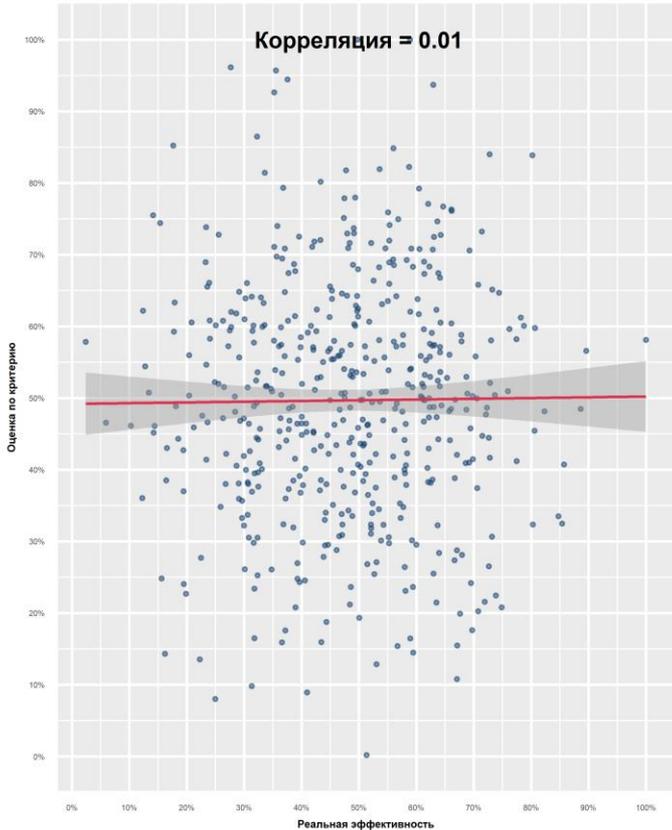
- Точность оценки внешнего критерия

Результаты проверки точности: оптимистичный сценарий

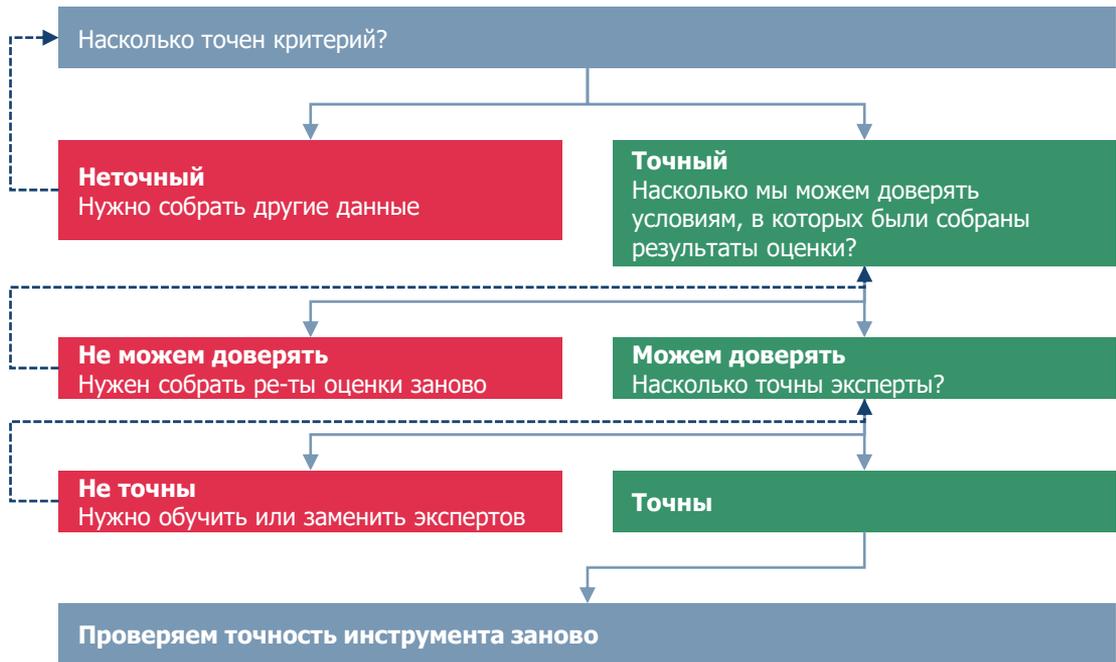


Всё отлично, ваша оценка работает!

Результаты проверки точности: пессимистичный сценарий



Сценарий дальнейшей работы



Результаты проверки точности: пессимистичный сценарий

Точность критерия и точность инструмента

Используем схему параллельного замера — измеряем один критерий двумя инструментами

Точность эксперта

Проверяем несколько экспертов параллельно — разделяем их данные смотрим на них и сравниваем.

Что можно делать после проверки точности инструмента оценки

Отказываться от инструмента, если его точность оказалась низкой (мы уверены в точности исследования)

Вводить пороговые значения для определённых критериев

Отказываться от части критериев, если их точность оказалась низкой и мы уверены в точности исследования



Разрабатывать профиль эффективного сотрудника (метод DEEP от ЭКОПСИ)