

c.case

candidate select



DEEP DIVE: CASE SCORE

case started as a research project in 2015. We have conducted more than 30 validation studies with our customers. This is a summary of the results, including some lessons learned and best practices.



EXECUTIVE SUMMARY

Educational outcomes are noisy and difficult to compare. GPAs don't predict much. We collect vast amounts of data to reveal the signal behind the noise. Our algorithms contextualize and score higher education globally using our case score. [Slide 03-04]



PART 1: EFFICIENCY

The case score enhances automation and reduces screening costs. Candidates with strong case scores are significantly more likely to be hired, even if the score is unknown. With case, applicants can be sorted and fast-tracked efficiently, resulting in higher interview success rates and reduced time to hire. [Slide 05-08]

PART 2: QUALITY OF HIRE

Success in education often requires qualities similar to those needed in the workplace, such as cognitive abilities and personality traits like conscientiousness. Contextualizing grades with the case score significantly improves the prediction of future job outcomes, such as performance reviews and promotions. [Slide 9-11]

PART 3: FAIRNESS & DIVERSITY

Assessments must meet fairness criteria. While algorithms are not per se fair, data shows that bias tends to be more problematic with human assessments. We demonstrate that the case score does not discriminate based on gender, ethnicity, legal status, or sexual orientation. [Slide 12-14]

PART 4: COVERAGE & SET-UP

After a decade of data collection, we achieved global coverage in 2023. Our algorithms now reliably identify and score degrees from over 30,000 universities worldwide. We integrate seamlessly with any system to enable automated and predictive assessments of applicant potential. [Slide 15-18]

case contextualizes performance in education.

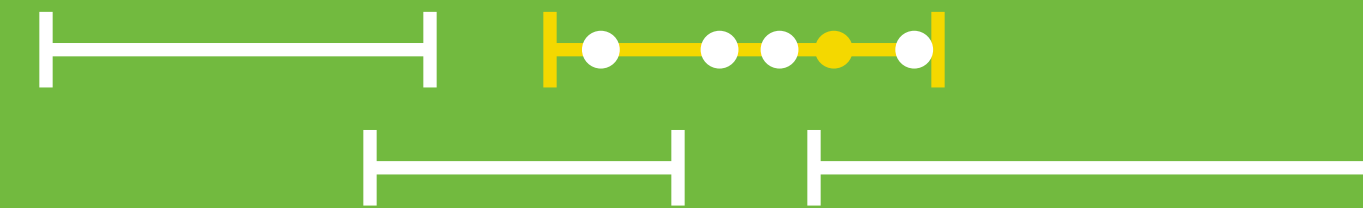


(1)
**WITHIN THE
PEER GROUP**



Performance measures, like GPAs, only convey meaning relative to the respective peers.

(2)
**BETWEEN
CONTEXTS**



Knowing that someone did well in a certain context only means something if that context can be compared to other situations.

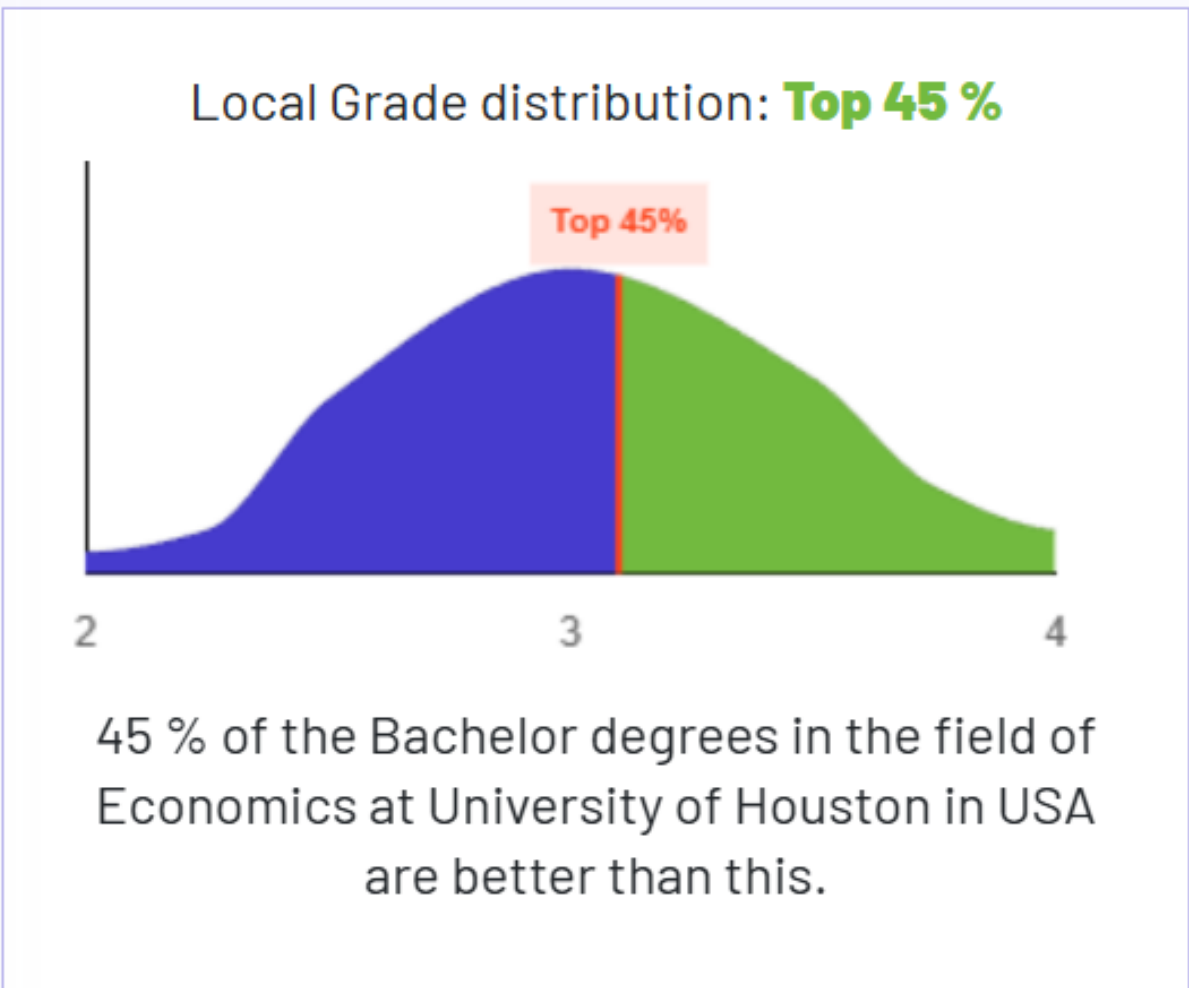
What is *case*?

Watch a short introduction and demo video

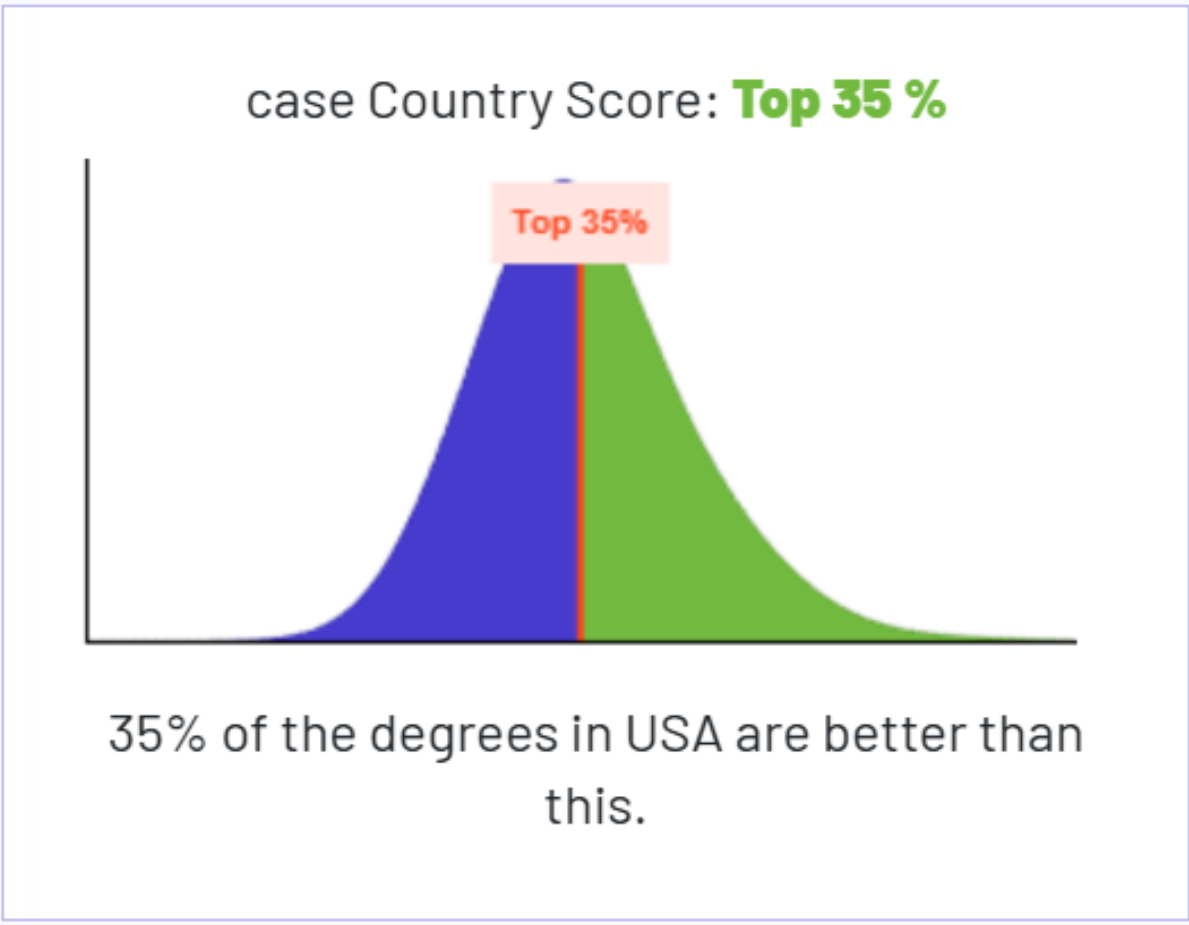


We provide four different metrics to contextualize education.

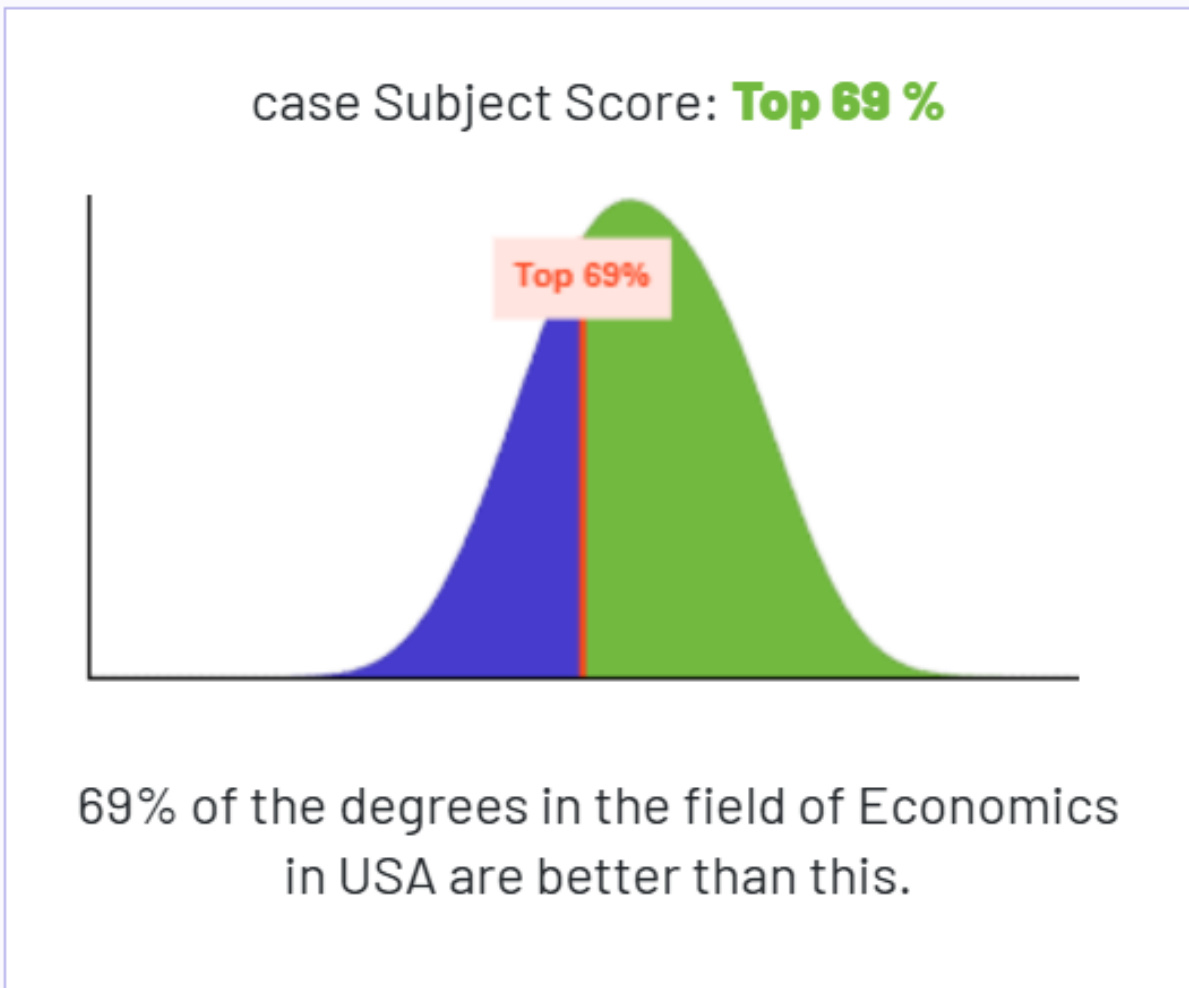
case local score



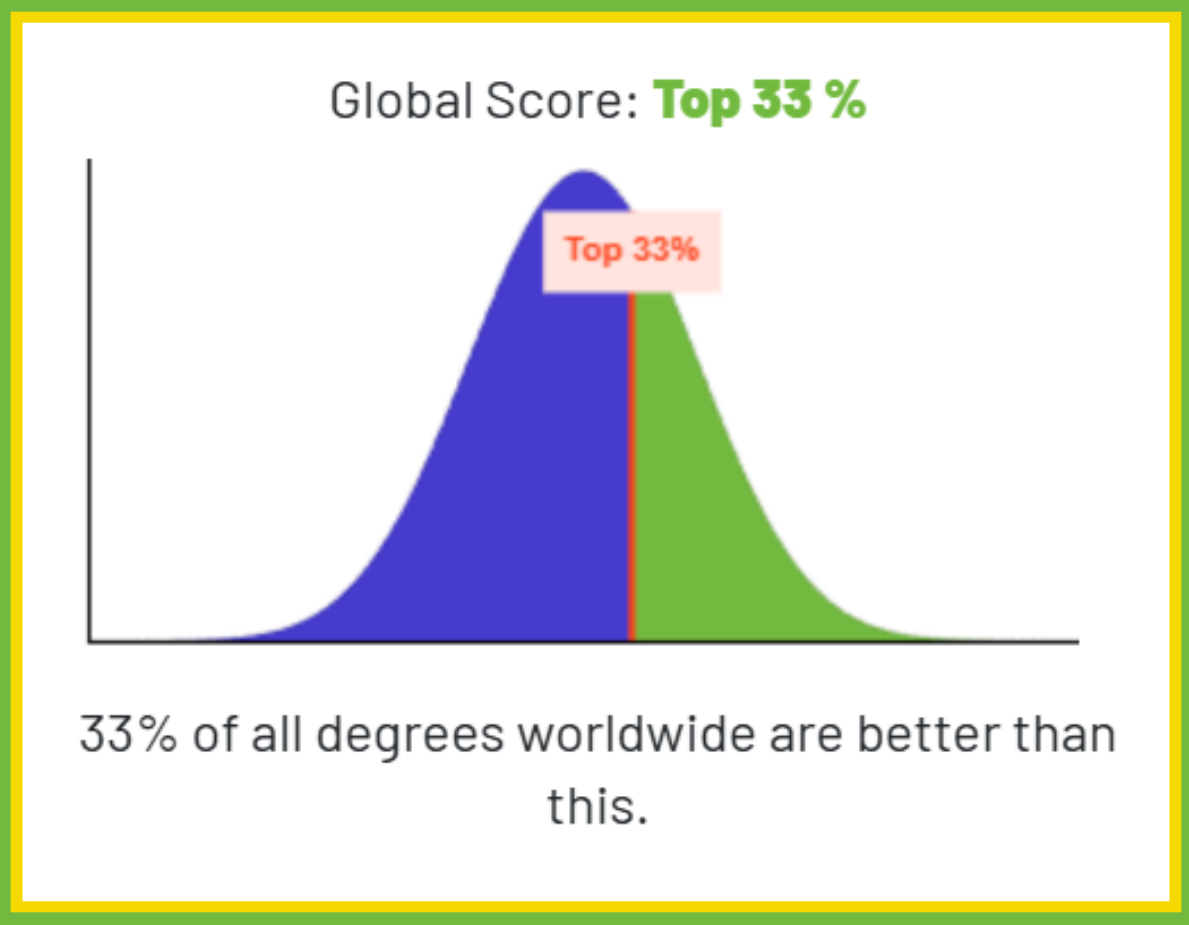
case country score



case subject score



case global score



Uni: University of Houston
Degree: Bachelor
Grade Point Average: 3.1
Subject / Major: Economics
Graduation Year: 2021



Use the **global score** for decision making as it allows to compare different universities, subjects or majors, degree types, years and even countries.

Within country, the country score can be used as well. The other scores are rather to gain a better understanding of a specific degree and have limited comparability.

For simplicity, we will simply refer to “the case score” from here on.



PART 1:

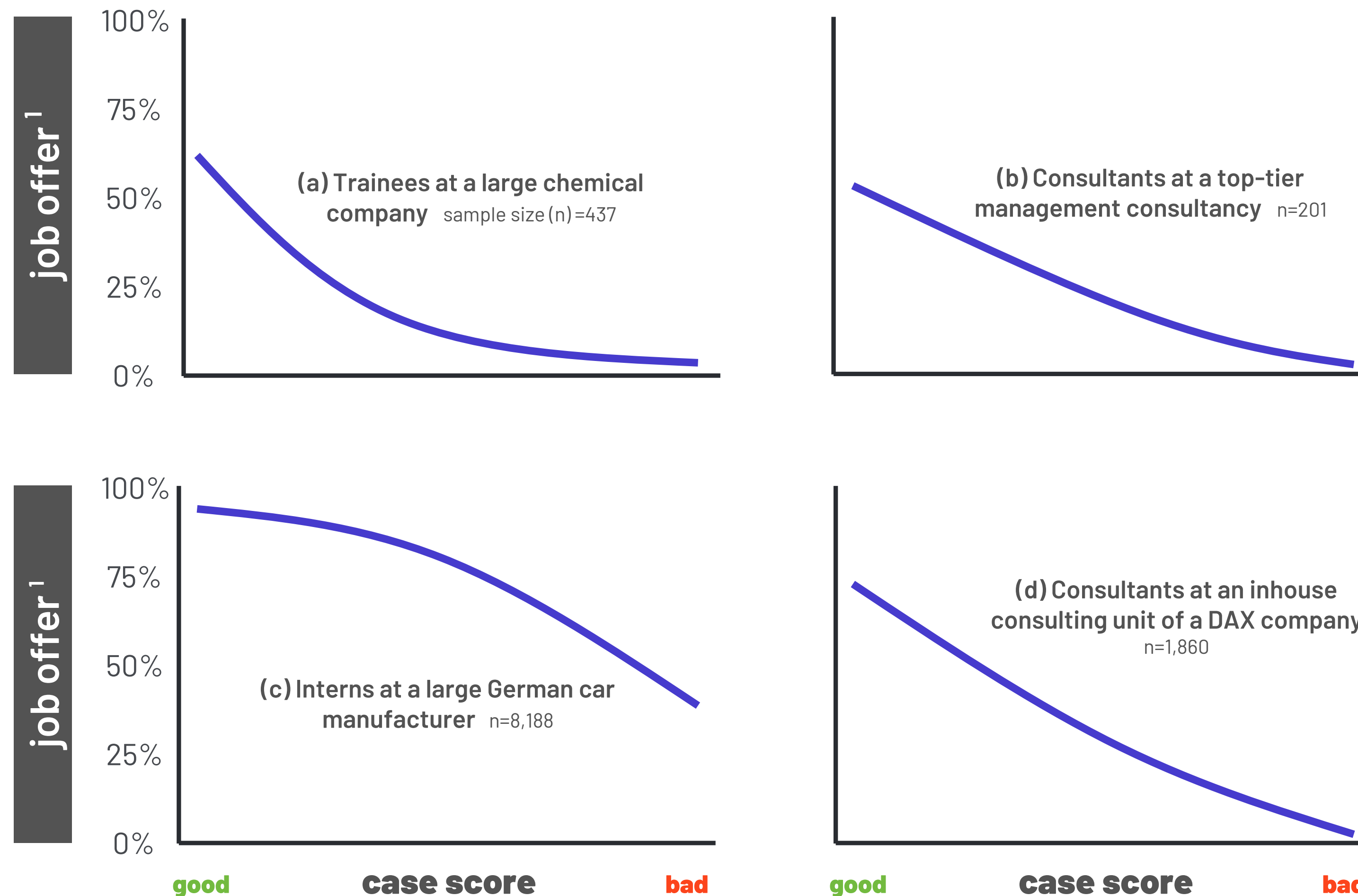
EFFICIENCY

Let's be honest: Applicant diagnostics come at a cost. Screening CVs or conducting interviews requires HR resources, and assessment tests can be time-consuming for applicants – sometimes causing dropouts.

So why are they necessary? Poor diagnostics are even more costly, especially in the long run.

This section looks at hiring processes across various companies and highlights how they optimized efficiency with the case score.

case can predict who you are going to hire before the fact.



Do early assessment results predict job offers at the end of the recruitment process? This is an effective starting point for evaluating efficiency.

The degree of selectivity varies by role, with internship candidates typically having higher chances of receiving offers than full-time applicants.

However, the critical factor is establishing a clear link between initial assessment tools and final job offers. Our case score demonstrates this connection across various positions and sectors.

¹ Probability (logit estimates) to receive a job offer after the final interview at different companies / roles in relationship to the case score.

Further improve your efficiency by combining case with tests.



study with **DHL**
Group

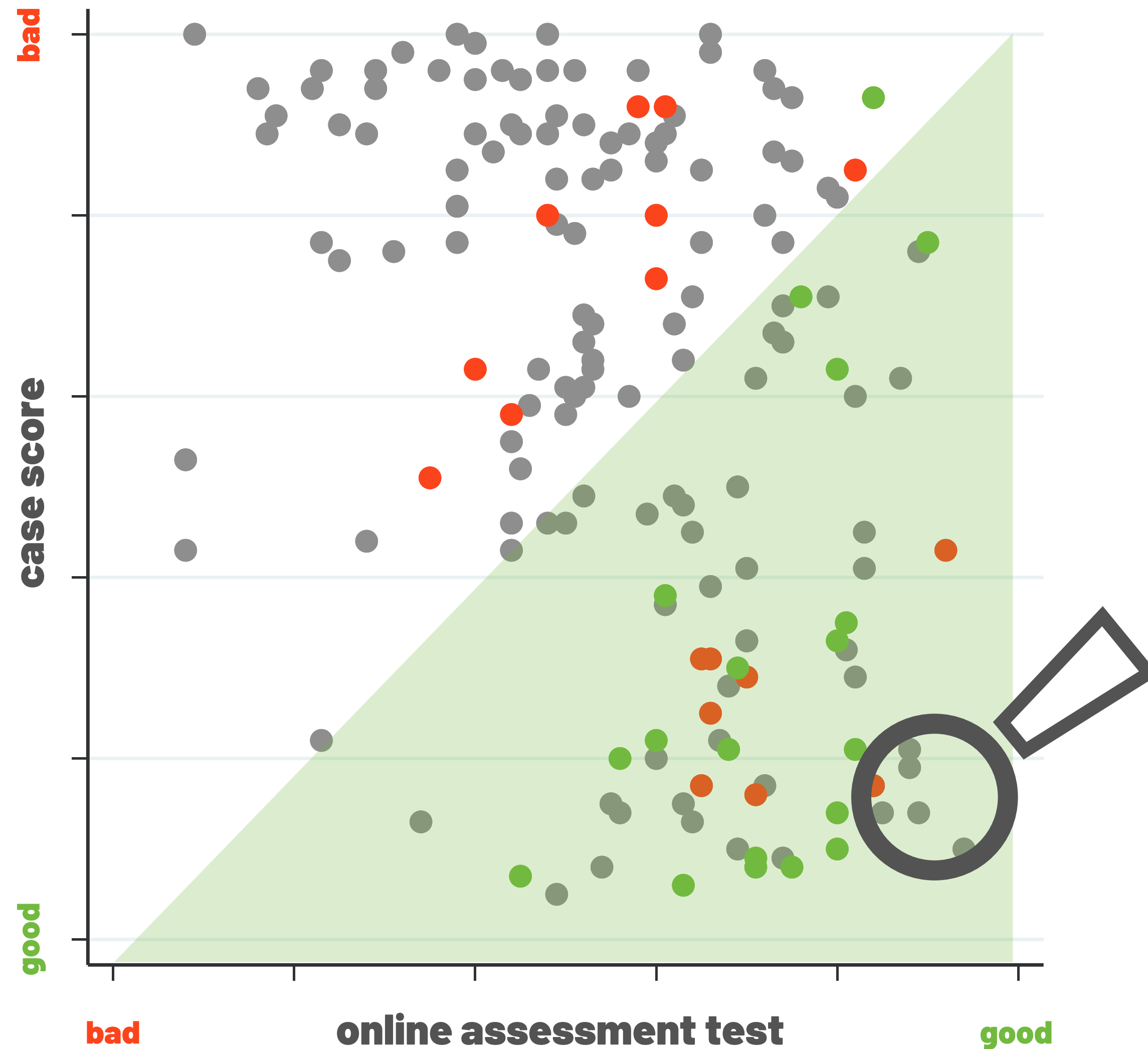
n=287

Combining two criteria boosts efficiency. The graph shows our case score with assessment test results, with the green triangle marking applicants who excelled in both.

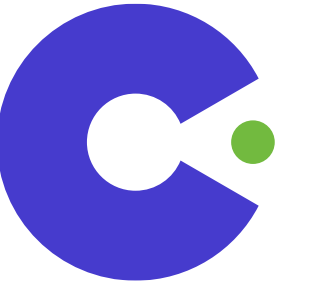
Overall, 54% of candidates received job offers after interviews, but if only those in the green triangle had been invited, the success rate would have been 72%. With a targeted approach, over 80% of recruitment costs could be saved in this example.

One more insight: notice the five grey dots in the bottom right? These top performing candidates withdrew their application before the interview. Next, we'll look at how to fast-track high-potential applicants.

- Applicants without an interview (not invited OR application withdrawn)
- Applicants that did not receive a job offer after the interview
- Applicants that received a job offer after the interview



You should fast-track high-potentials with case.



study with **SIMON KUCHER**

n=465

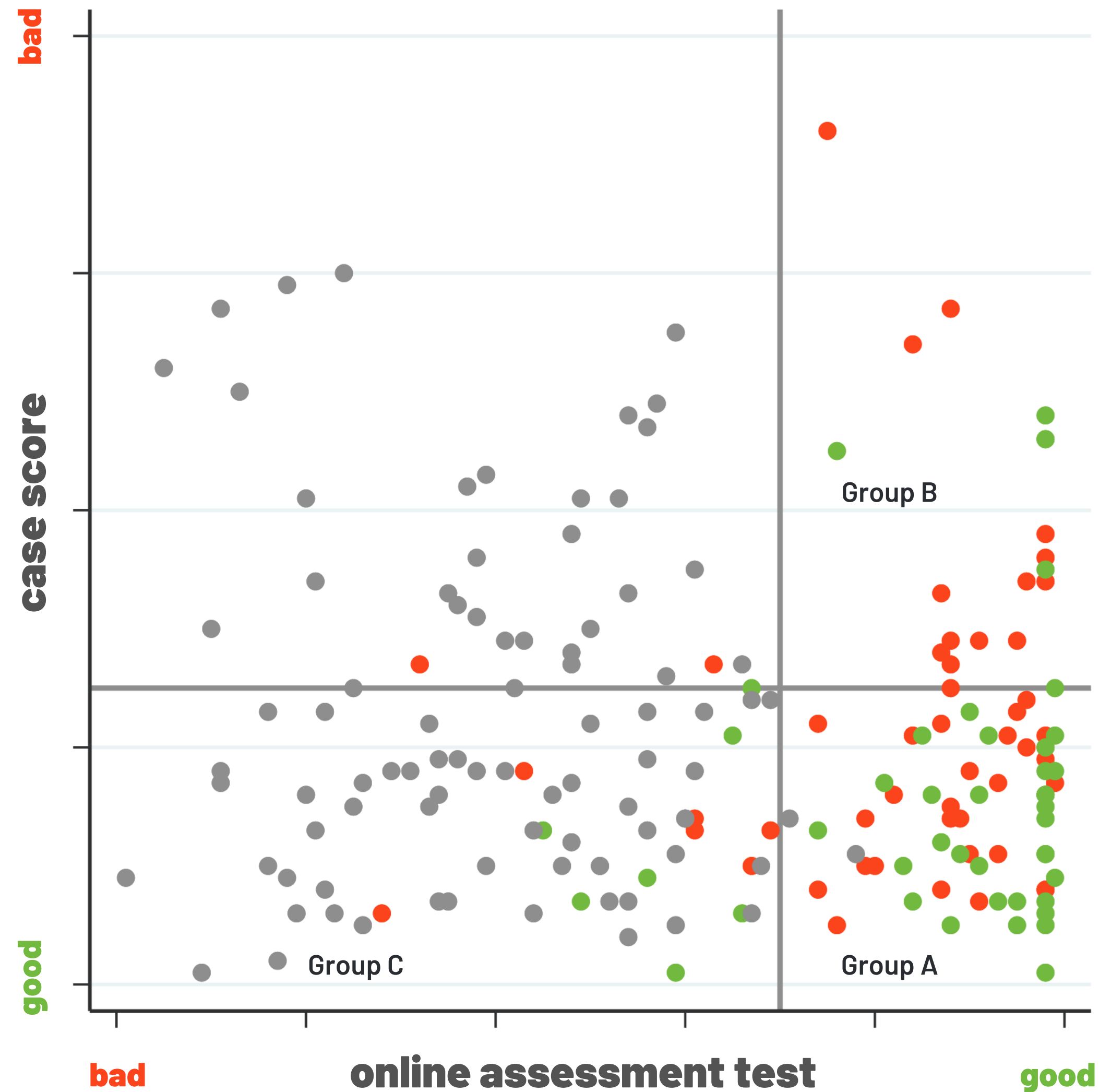
Top candidates often receive competing offers quickly. So, how can we identify and fast-track them?

At Simon Kucher, an assessment test has long been part of the hiring process. With the addition of the case score, they explored whether allowing certain applicants to skip the test would impact interview quality.

- Group A: Strong in both criteria, hired 53% of the time
- Group B: High test score but low case score, hired 28%
- Group C: Low test score but high case score, hired 54%

The data supports fast-tracking: candidates with top case scores skip the test and go directly to interviews, while others can qualify through strong test results.

- Applicants without an interview (not invited OR application withdrawn)
- Applicants that did not receive a job offer after the interview
- Applicants that received a job offer after the interview





PART 2:

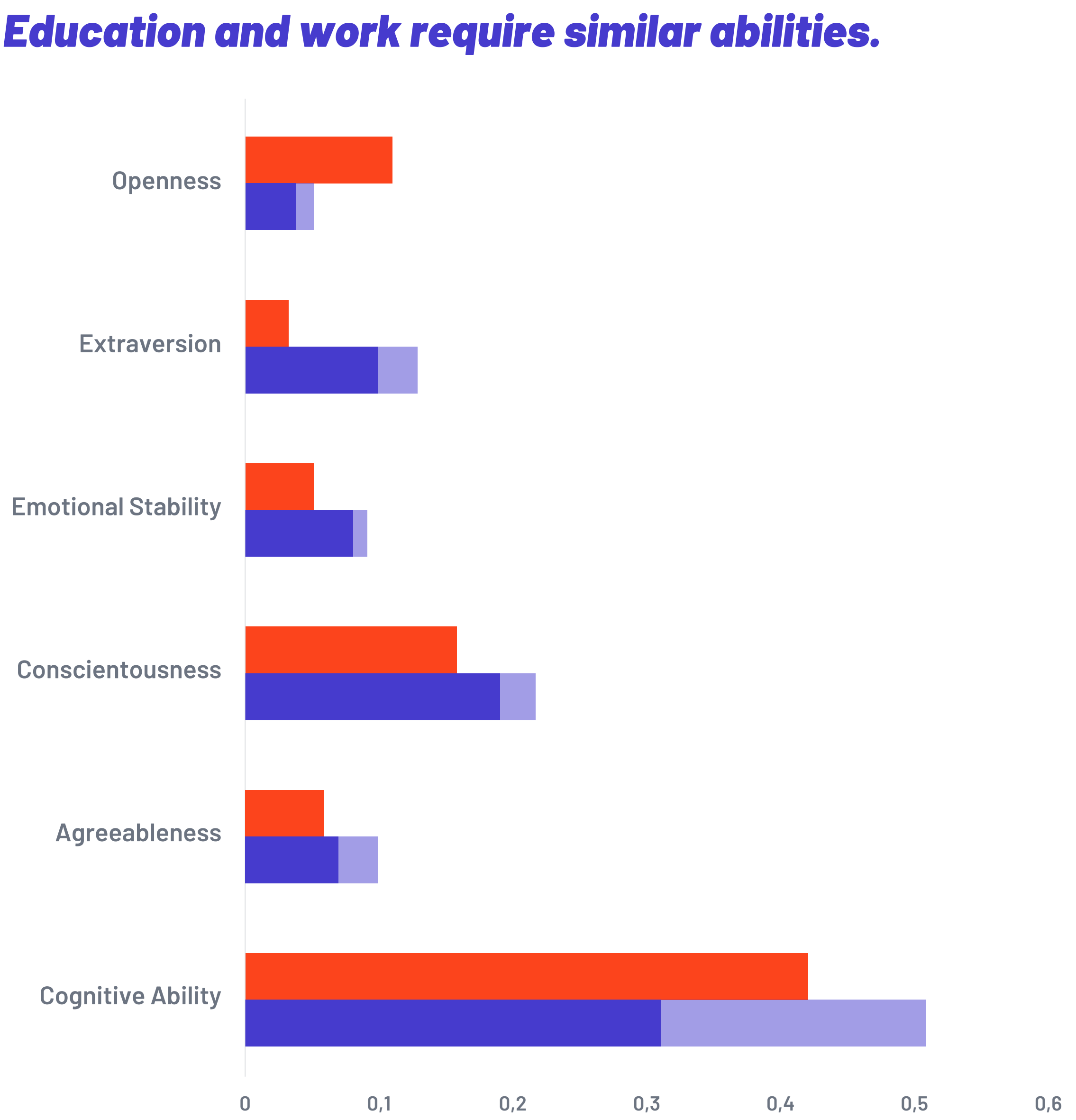
QUALITY OF HIRE

Let's stay humble: Hiring is inherently risky, and we often overestimate our ability to identify the right people. Even the best criteria can only account for about 25% of the variation in later job performance. People are multifaceted, and their future potential is challenging to predict.

This uncertainty should not prevent us from striving to do our best. While no assessment is perfect, well-designed evaluations have a clear, positive impact on organizational performance.

This second section examines how effective assessments — and education in particular — are at predicting job outcomes later in life.

Education and work require similar abilities.



Performance in education (case score)

Job performance



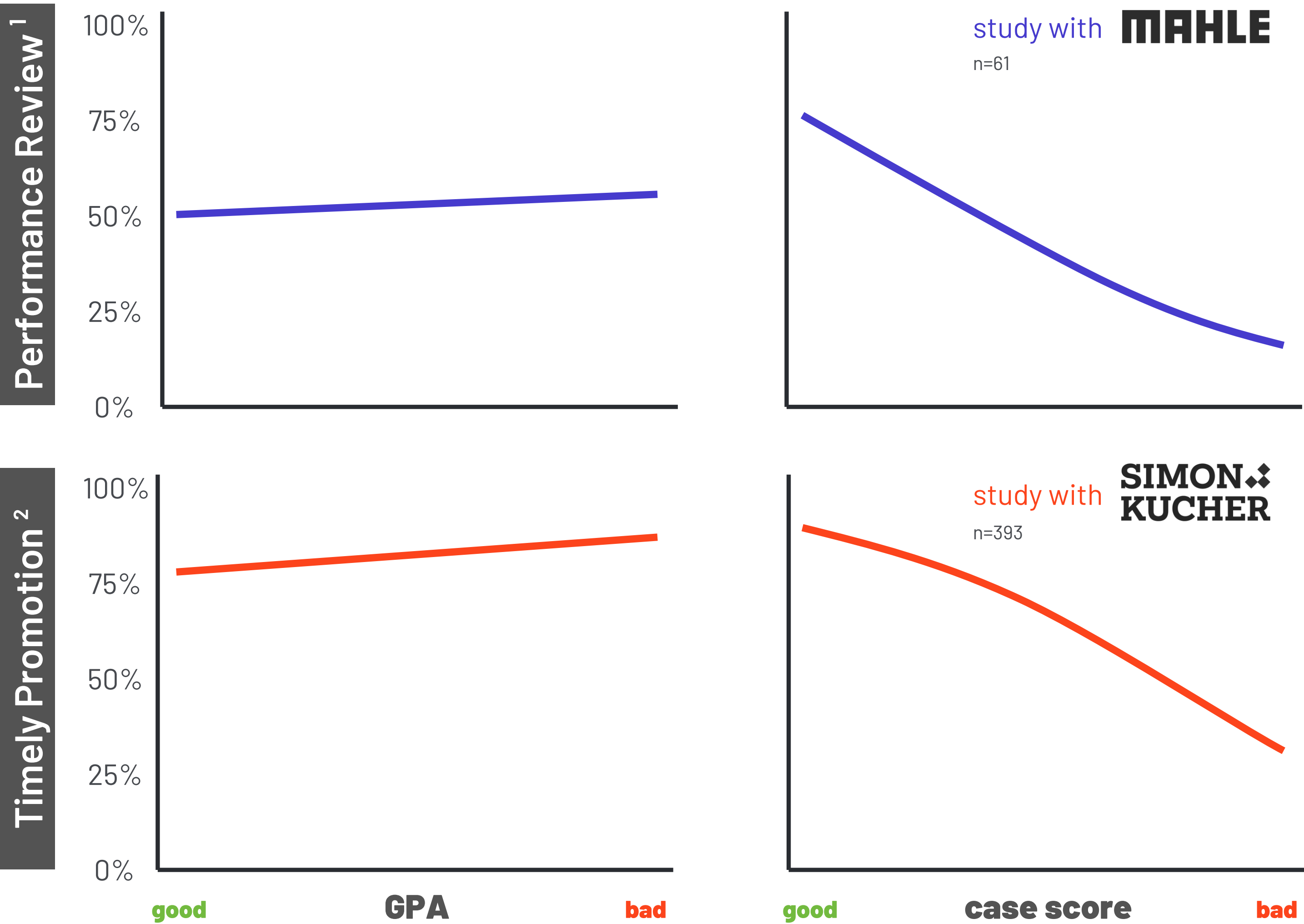
Who succeeds in the labor market? It's a challenging question because human characteristics are hard to quantify, and their impact varies by job type.

Research has identified certain traits that are generally rewarded. Cognitive ability has a positive effect on job outcomes, varying in impact but standing as the strongest single predictor. Personality traits like conscientiousness also correlate positively with job performance.

So, why focus on education? Educational performance is influenced by similar characteristics. Our findings reveal a clear relationship between these traits and the case score, suggesting that it will predict job success. Next, we'll explore whether this holds true in practice.

There is an ongoing academic debate on how these correlations should be statistically corrected. As a result, the estimates can vary quite a bit as indicated in the graph. The correlations with the case score have not been corrected at all and are by design a bit lower. Estimates for case are based on the Fachkraft 2030 study (N=6,807) and a study with the test provider ITB consulting (n=4,844). The other results have been taken from the following meta-analysis, which also provide additional context to the (very) interested reader: Almlund et al. (2011), Handbook of the Economics of Education, Chapter 1 – Personality Psychology and Economics; Schmidt, Hunter (1998), The Validity and Utility of Selection Methods in Personnel Psychology; Sackett et al. (2021), Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range.

case reliably identifies high performing employees.



Probability (logit estimates) to ¹ receive a very positive review 2 years after being hired / ² be promoted within the expected time frame



Any assessment must predict job performance post-employment. This is challenging to prove, as performance is conceptually complex and legally difficult to measure.

At MAHLE, a German automotive supplier, trainees received performance reviews after two years on the job. The reviews showed no correlation with GPAs but did significantly correlate with case scores.

At Simon Kucher, promotion within the targeted timeframe serves as the primary metric for success. Here, too, no link was found with grades, but a significant correlation appeared with the case score.



PART 3:

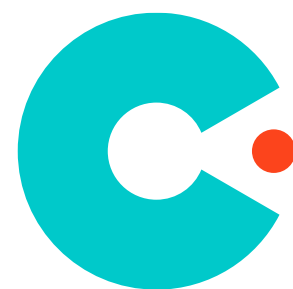
FAIRNESS & DIVERSITY

Let's be fair: Focusing solely on job outcomes is insufficient, as certain labor market practices tend to favor specific groups. For instance, if men are more likely to be promoted, any criterion favoring men will appear to have stronger predictive validity.

Recruitment should work to counteract these biases. We should focus on candidates with future potential, irrespective of gender, race, nationality, or sexual orientation.

This third section explores group differences in hiring practices broadly and in relation to the case score specifically.

Subjective assessments lead to gender differences.



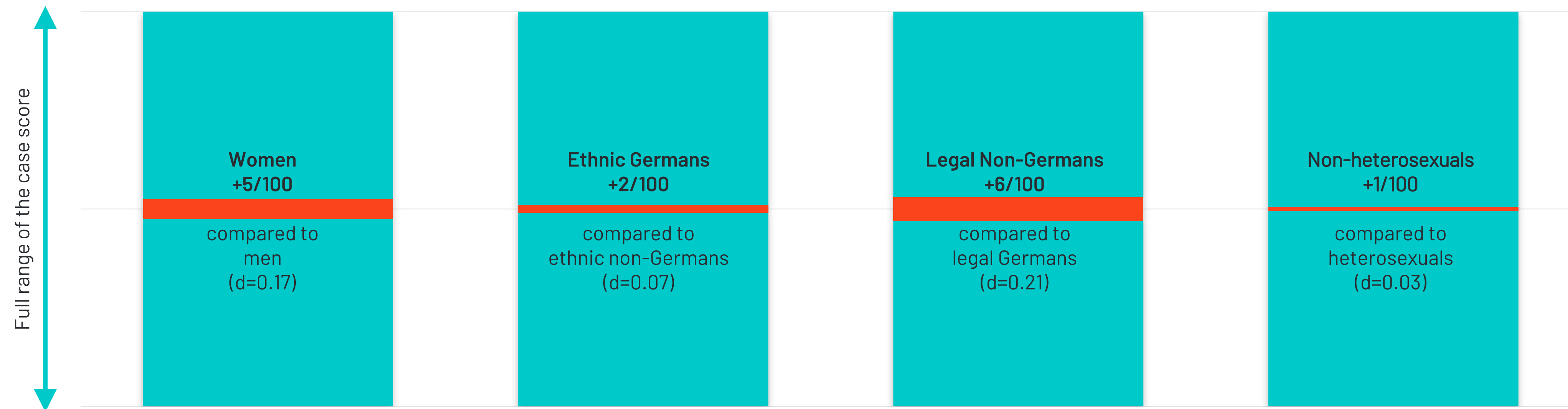
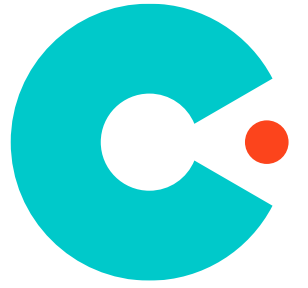
criteria	direction	average (female applicants)	average (male applicants)	effect size (d)
case score	lower scores are better	46.5%	45,5%	0.04
manual screening	lower scores are better	4.1	3.7	0.26
assessment test	higher scores are better	54	57	0.14
interview	higher scores are better	3.1	3.4	0.56

These are assessment results by gender for prospective trainees at a large chemical company. We will look at other criteria like race, legal status or sexual orientation on the next page.

The case score shows the lowest differences by gender, followed by the assessment test, manual screening, and interviews. While manual screening differences are statistically significant, the effect size for interviews is particularly concerning.

Human assessments are prone to bias, and training a large team of recruiters to avoid these biases is challenging compared to using standardized tools. Many biases are unintentional and unconscious, making them difficult to eradicate. Moreover, differences in assessment outcomes can also stem from applicants' behaviors, such as men displaying greater boldness in interviews.

case provides an unbiased and fair performance metric.



This analysis does not rely on recruiting data, as it would be unreasonable to request applicants' ethnicity or sexual orientation, even with the best intentions. Instead, we use a large and representative student sample that includes demographic variables and educational data to compute case scores.

All effect sizes are small, suggesting that the case score is indeed a fair hiring metric, not only for gender but also for ethnicity,

nationality, and sexual orientation. Notably, women slightly outperform men, and foreign students slightly outperform domestic students. Favorable treatment of disadvantaged groups can help to reduce disparities in labor market outcomes more quickly.

Results are based on a multiple regression-analysis to measure group difference in the case score by gender, ethnicity (appearance), legal status (passport) and sexual orientation. The model controls for personality, cognitive ability and language skills. Data is from the Fachkraft 2030 study (n=2.632). Group difference have a small to very small effect size. Given the large sample, differences by gender and legal status are statistically significant at $\alpha = 0.05$. Differences by ethnicity and sexual orientation are insignificant at this level. The small differences reflect differences in input data (e.g., women / mobile students outperforming men / local students) and are not driven by the case algorithm.



PART 4:

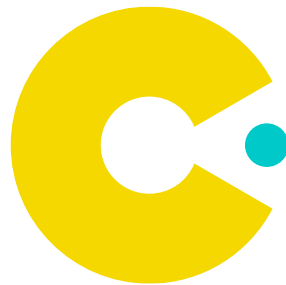
COVERAGE & SET-UP

Let's be clear: We present a strong, data-backed argument that the case score enhances efficiency, hiring quality, and fairness. However, these benefits can only be realized if our scoring achieves high coverage and supports easy integration.

Great coverage requires extensive data, and seamless integration requires experience. We have both.

This final section highlights coverage statistics from current customers and explains how to integrate case into any HR tech stack.

case automatically identifies education and cleans up your data.



University:

Uni Houstn

University of Houston

Graduation Year:

2022

2022

Degree:

B.A.

Bachelor

Subject:

Major in Economics

Economics

(Current) GPA:

My final GPA is 3.1

3.1



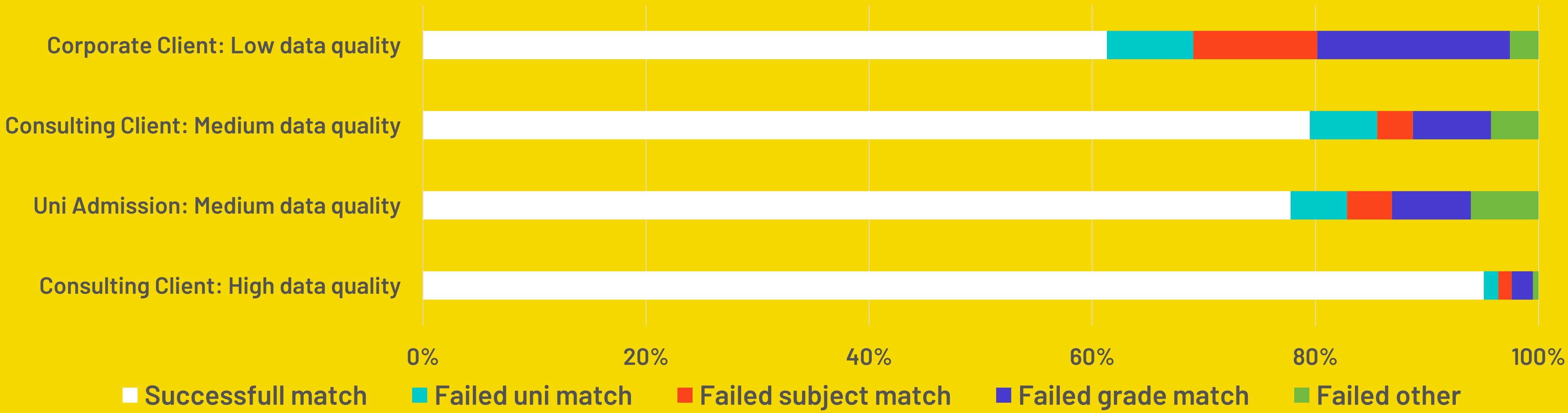
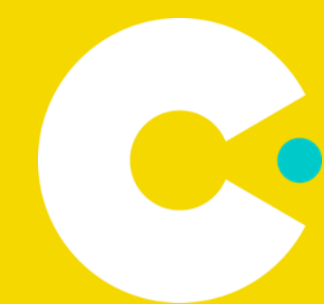
To calculate a score, we require no personalized data – only the five variables shown on the left.

This information is typically provided by applicants through application forms and is sometimes parsed directly from CVs.

Our API identifies and categorizes these entries from free text automatically. Scores are then computed – as demonstrated on slide 4. The categorized data is returned as well to enhance your analytics and reporting capabilities.

The next slide illustrates how accurately these entries are recognized with our current customers.

Our high coverage allows you to reliably identify high-performing applicants globally.



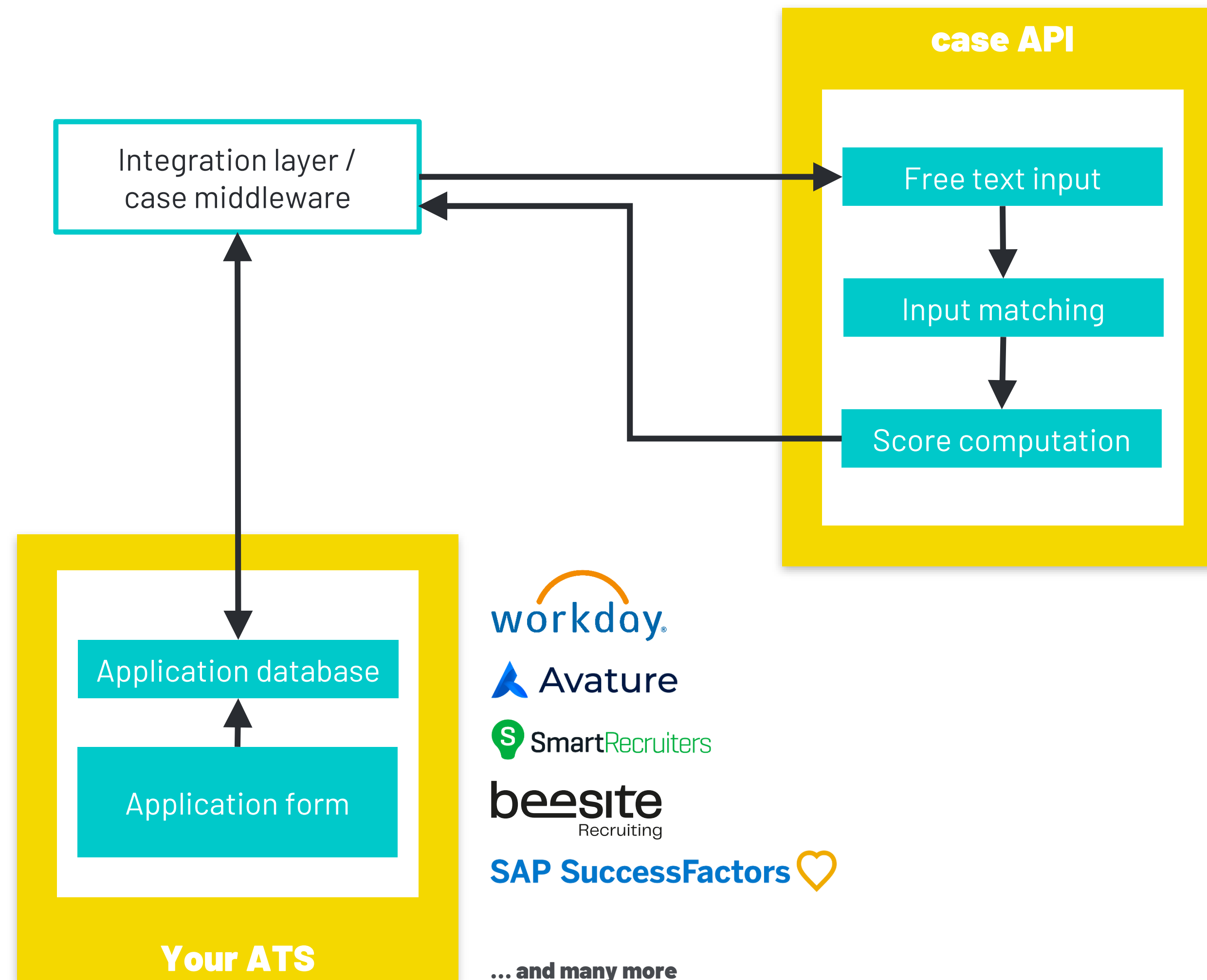
The graph shows the percentage of all data that our API successfully scored automatically. Coverage strongly depends on the quality of data provided. When data is complete and accurate, our API can score over 95% of entries.

The four clients use case scores to evaluate education globally and receive thousands of applications each month. They were selected

to represent the full range of data quality that we observe.

The main reasons for unsuccessful matches are missing or inconsistent data in grades, university names, and subjects. In many cases, producing no score is appropriate – such as when a high school is entered as a university or when a grade has already been converted to a different system.

It is easy to integrate case into any HR tech stack.



There are two options for data exchange:

- Your ATS can push data to our API via an integration layer.
- Our custom middleware can pull data from your ATS and forward it to our API.

Our API then recognizes and scores the degrees, with scores sent back to display instantly in your ATS.

This integration is compatible with any ATS that offers API access. We have ongoing collaborations with many ATS providers and can recommend the best integration approach for your system.