Talegent Approach to

Solution Verification

TALEGENT.

CONSULTING SERVICES

What is
Solution Verification?

Where Talegent Confirms that your Campaign is Effective

- Solution Verification examines an assessment already in use, and establishes whether it measures the right things and how strongly those things relate to performance in the role.
- Talegent correlates assessment scores against manager ratings of job performance. This is the evidence that separates an assessment that looks sensible from one demonstrated to work.
- Every study also examines fairness, testing for differences in how groups score and for differences in how accurately the assessment predicts performance across them.
- Findings translate directly into action: competencies to add or remove, weightings to adjust, and a defensible record of why your selection process measures what it measures.
- Solution Verification services are divided into three categories to suit the data, budget and time available.

Where Data Confirms the Decision

Choosing an assessment is a judgement. Verifying it is a measurement. A Solution Verification study tests the assessment you already run against how people actually perform, and reports the relationship as a number with a confidence interval, not an opinion about whether the competencies look right. Which of the three studies fits depends on the data you can reach.

Talegent Solution Verification
Consulting Services:

Expert Review

Talegent experts conduct a thorough qualitative review of your campaign to evaluate its effectiveness and relevance. Our verification report includes key competencies and actionable recommendations for improvement.

Concurrent Solution Verification

Talegent experts quantitatively review your campaign's solution to assess its effectiveness and alignment with the role. Our verification report includes key competencies, statistical analyses, and actionable suggestions to optimize your recruitment process.

Predictive Solution Verification

Talegent experts conduct a thorough quantitative study of your campaign to evaluate its validity, effectiveness and relevance. Our verification report includes key competencies, deep statistical analyses, and actionable recommendations for improvement.

Choosing the Right Study

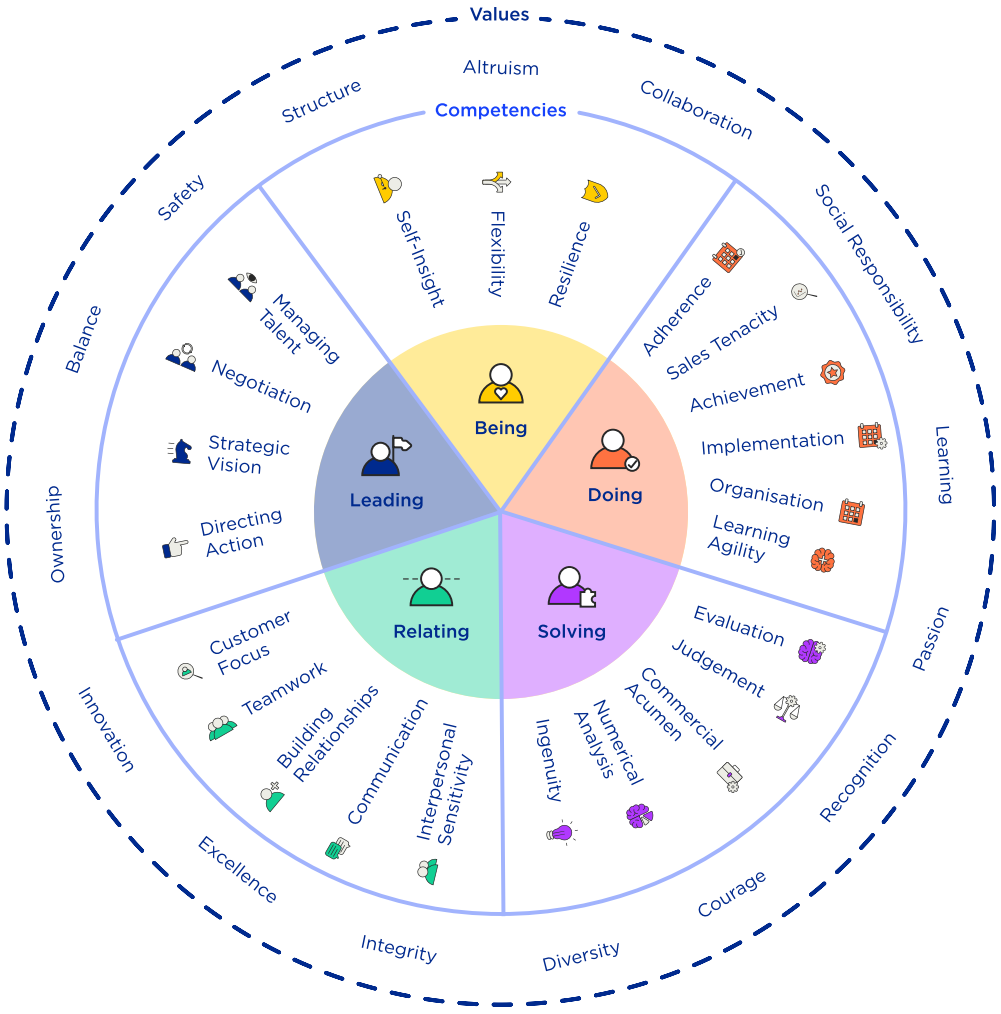
	Expert Review	Concurrent	Predictive
What you need	A live campaign and access to incumbents and managers for interview.	Current employees who can complete the assessment, and their managers to rate them.	Assessment data already held on people hired into the role, and their managers to rate them.
When you can run it	Any time. No performance data required.	Any time you have a settled population in the role.	Once enough hires assessed at application have been in the role three months or more.
What it establishes	**Relevance.** Whether the assessment measures what the role demands.	**Concurrent validity.** Whether scores relate to performance among people doing the job now.	**Predictive validity.** Whether scores at application predict later performance.
Indicative timeframe	3–4 weeks	6–8 weeks	6–8 weeks once the data is in place

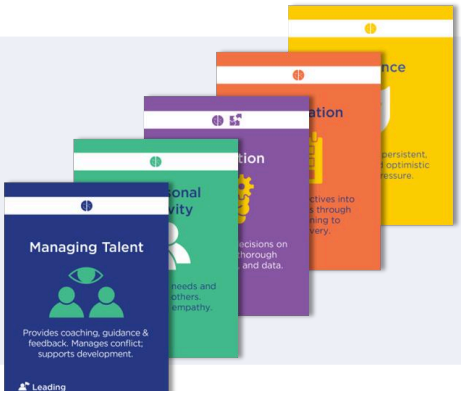
Talegent Competency Model and Cards

The Talegent competency model underlies Talegent's Talent Assessments and comprises of a comprehensive set of 23 competencies, grouped into 5 key clusters. Competencies are selected for role-specific assessments to identify top talent across employment levels and industries. The model provides a logical, practical, and consistent approach to communicating and interpreting candidates' results.

Talegent's Competency Model contains 23 competencies across 5 clusters:

Being | Doing | Leading | Relating | Solving





The **Talegent Competency Cards** are designed to help you identify the key competencies that are important for success in a particular role. Once selected, they can be used to build your psychometric assessment. These cards describe the key behavioral competencies employees need to do their job effectively.

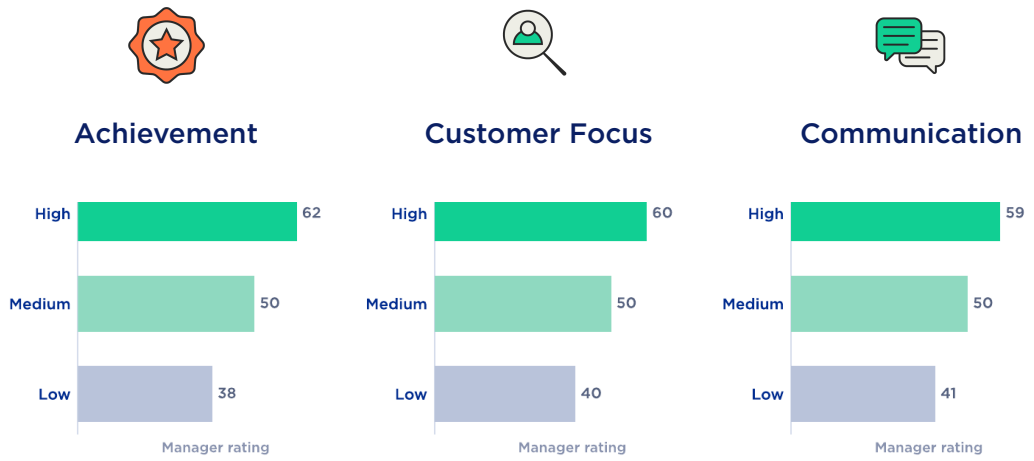
TALEGENT.

Utilising your Solution Verification Report

A Solution Verification report tells you which parts of your assessment relate to performance, how strongly, and what to change.

1 What the study found

The value of a verification study is seeing your own assessment data line up against how managers actually rate people in the role. Each competency is tested against the manager's rating of that same competency, so you can see which parts of the assessment are separating strong performers from the rest, and by how much. Competencies that do the most work can then be weighted accordingly; those that do none can be reduced or removed.



2 Recommendations

Findings become specific changes to your campaign, which Talegent can configure in the platform.

Reweight what predicts most

Most solutions weight every competency equally. The study shows they do not contribute equally, and recommends weightings in proportion to what each predicts, raising the value of the score without adding candidate time.

Add or remove competencies

Where a competency shows no relationship with performance it is reduced or removed; where the role demands something not currently measured, it is added.

Act on the fairness findings

Group differences and predictive bias are reported with a recommendation on what to monitor and how often, as volume accumulates.

Know when to revisit

Validity evidence describes the role as it was during the study. The report sets out when to revalidate, and what would trigger doing so sooner.

Figures shown are from a sample Predictive Solution Verification report. An Expert Review reports relevance rather than correlation, and does not produce the chart above.

Talegent Concurrent Solution Verification

A Concurrent Solution Verification establishes whether your assessment relates to how people are performing in the role right now. Current employees complete the assessment, their managers rate their performance against the same competencies, and Talegent correlates the two.

Its advantage is that it can be run immediately. Because the study generates its own data rather than drawing on assessment records accumulated over years, there is nothing to wait for — it suits a solution that has recently changed, a newly configured campaign, or any role where the assessment history does not yet exist.

What it does require is internal buy-in. Employees give up time to complete the assessment and their managers complete a performance survey for each of them, so the study depends on those people being briefed, willing and available. Talegent supports the internal communication needed to secure that.

Process

1

Scoping

Your Client Success rep confirms the role, then helps prepare the internal communications that brief everyone taking part.

2

Data Gathering

Participants complete the assessment, and their managers complete a performance survey rating them against the same competencies.

3

Analysis

Scores are matched to ratings and analysed competency by competency, and the findings written up.

4

Report Delivery

The Talegent Consulting team presents the findings and the recommended weightings.

Pricing

PRICED PER ENGAGEMENT

Talk to us

Timeframe

INDICATIVE TURNAROUND

6–8 weeks

What you receive

- Validity results for every competency, with confidence intervals
- Fairness analysis across the groups your data supports
- Recommended weightings, configured in your platform
- A defensible record against professional standards
- Findings presented by the Talegent Consulting team

TALEGENT CONSULTING SERVICES

Concurrent Solution Verification

Graduate Programme Employee

Prepared for Globex

Globex | TALEGENT.

Report date 4 December 2026
Version 1.0
Prepared by Talegent Consulting

Contents

What's in this report

This report tests whether the Talegent assessment used to select into the Globex Graduate Programme relates to how members of the programme are performing in it.

THE HEADLINE

Six of the seven competencies predicted the manager's later rating of that same competency. The strongest, **Achievement**, correlated **.26** ($p < .001$). Because only hired applicants could be followed into the role, these are conservative estimates.

01 Overview

Purpose and method

Purpose

This report presents the results of a Concurrent Solution Verification for the Globex Graduate Programme. Its purpose is to establish whether the Talegent assessment used at the screening stage predicts subsequent job performance, and to identify where the solution can be strengthened.

Method

Validity is the degree to which an assessment measures what it claims to measure, and — for selection — the degree to which its scores relate to later performance on the job.

This was a **concurrent** design. Current members of the Graduate Programme completed the Talegent assessment, and at the same time their managers completed a structured questionnaire rating them against the same competencies the assessment measures. Each competency score was then compared with the manager's rating of that same competency.

A concurrent study delivers evidence quickly, because it does not wait for a cohort of applicants to be hired and settle into the role. That speed is its advantage over a predictive design, and the trade-offs below are the price of it.

Three limitations are built in. Programme members are a **survivor sample**: they were selected in and then stayed, so the spread of assessment scores among them is narrower than among applicants, which pushes correlations down. They completed the assessment knowing it was for **research rather than selection**, so their motivation may differ from that of a candidate competing for a place. And because their managers already know them, ratings may reflect accumulated reputation as much as observed behaviour against each specific competency.

None of this invalidates the result. It does mean the figures here should be read as a conservative floor, and that a predictive study remains the stronger evidence where time allows. Appendix B quantifies the range restriction.

At a glance

Role studied	**Graduate Programme Employee**
Organisation	**Globex**
Service	**Concurrent Solution Verification**
Design	**Predictive · assessment at application, criterion at 3+ months**
Analysis sample	**120 incumbents with complete assessment and rating data**
Fieldwork	**6 October – 14 November 2026**
Result	**6 of 7 competencies predicted their matching manager rating**

02 The study

Sample, design and criterion

INVITED
138
Current Graduate Programme members invited to take part

ANALYSED
120
Complete assessment and manager rating data — 87% of those invited

MANAGERS
22
Provided ratings, a median of 5 direct reports each

Sample composition

Of the 120 people in the analysis, 64 were women and 56 men, aged from 20 or younger to 35 (median band 21–25). Tenure in the programme ranged from 4 to 26 months, with a median of 14. Degree disciplines were spread across commerce (38%), engineering and science (31%), arts and social sciences (19%) and other or not stated (12%). English was the primary language for 79%.

Eighteen were excluded: 12 did not complete the assessment within the window and 6 had no manager rating returned. Non-participants did not differ significantly from the analysis sample on tenure ($t = 0.61$, $p = .54$) or on their most recent programme review rating, so participation bias is unlikely to have shaped the result.

The criterion

Managers completed a structured questionnaire rating each person against behavioural statements drawn from Talegent's work performance model, on a five-point scale. Items combine into a rating for each of the competencies the assessment measures.

Internal consistency across the competency scales was **Cronbach's alpha = .89**. Where two managers rated the same person ($n = 19$), inter-rater agreement was **ICC = .58** — typical for supervisory ratings.

HOW THE COMPARISON WORKS

Each competency is tested against itself. The score a programme member achieved on **Learning Agility** is compared with their manager's rating of their **Learning Agility**. The same is done for each of the seven competencies in the solution.

This is a like-for-like test rather than a general one, and it is the more demanding version: it asks whether the assessment predicts the specific thing it claims to measure, not merely whether high scorers are generally well thought of.

03 Findings summary

The assessment predicts performance

Six of the seven competencies predicted the manager's later rating of that same competency. For a predictive study against supervisory ratings, that is a strong result.

STRONGEST PREDICTOR

.28

Learning Agility, against the manager's rating of it, p = .002

TOP VS BOTTOM THIRD

2.2×

High Learning Agility scorers are 2.2 times as likely to be rated highly on it

COMPETENCIES CONFIRMED

6 / 7

Predicted the matching manager rating at p < .05

Effective performance in the Graduate Programme at Globex

Learning Agility | Achievement | Teamwork | Resilience | Communication | Self-Insight

Ingenuity

Shaded competencies predicted their matching manager rating at p < .05. Full results are in Appendix A.

Ingenuity did not reach significance

At r = .12 its confidence interval crosses zero, so on this evidence its relationship with performance is not established. Either the effect is genuinely small, or managers rating graduates in their first two years have limited opportunity to observe original thinking — early programme rotations tend to reward reliable delivery of defined work.

Sample size matters here. At N = 120 this study had 59% power to detect a correlation of .20 and 37% to detect .15, so a real but modest effect could easily have been missed. We recommend retaining Ingenuity at reduced weight rather than removing it, and revisiting once the cohort reaches more senior rotations.

What this means in hiring terms

A programme member scoring in the top third on Learning Agility is **2.2 times** as likely to be rated highly on Learning Agility by their manager as one scoring in the bottom third. The same pattern holds for five of the remaining six competencies.

READING A VALIDITY COEFFICIENT

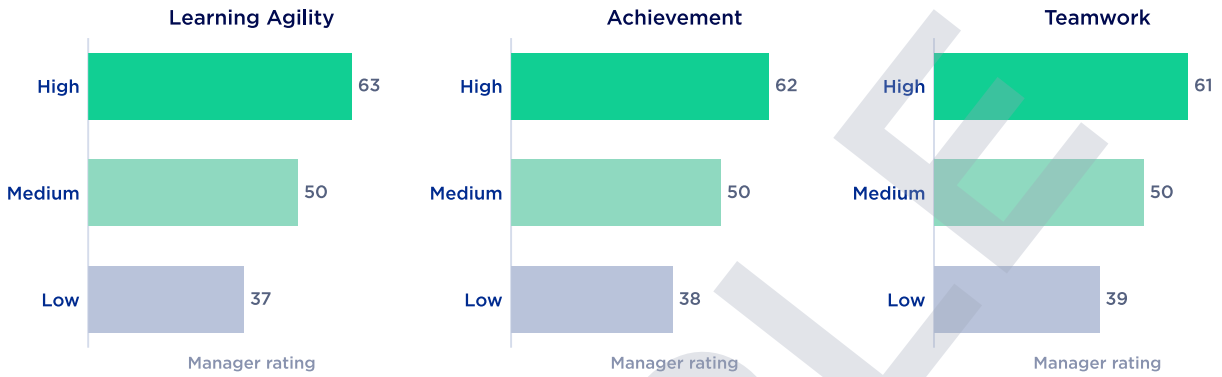
Selection research does not produce correlations near 1. Meta-analytic work puts well-constructed assessments in the .20 to .35 range against supervisory ratings before any correction, so the figures in this report sit where a sound assessment should.

These figures are also conservative, because only hired applicants could be studied. A coefficient reported above about .50 against an independently collected supervisory criterion should be treated with suspicion rather than satisfaction — it usually indicates the predictor and criterion share a source.

04 What the scores predict

Performance by score band

Correlations describe the strength of a relationship but not its shape. The panels below show the mean assessment percentile of people whose managers rated them in bottom, middle and top third on that same competency.



Learning Agility

Grasps and applies new concepts quickly. Enjoys learning, seeks out feedback and adjusts approach in response to it.

The strongest predictor in the solution, which is what a graduate programme should hope to find: members rated in the top third for Learning Agility scored, on average, 26 percentile points higher on the assessment than those rated in the bottom third. In a role defined by rotation through unfamiliar work, getting up to speed is the job.

Achievement

Has high energy combined with a drive to achieve goals. Is motivated by work, tasks and the opportunity to compete with others.

Predicted the manager's rating of Achievement, and more specifically of working consistently hard and following through on commitments — the behaviours managers cite most often when distinguishing between graduates at the same stage.

Teamwork

Works inclusively and prioritises team goals, delivering on commitments to others. Predicted the manager's rating of Teamwork, and of contributing in group settings in particular — unsurprising in a programme structured around project rotations.

05 Fairness

Group differences and predictive bias

A valid assessment can still produce unfair outcomes. Two separate questions have to be answered: do groups score differently, and does the assessment predict performance equally well for each group?

Do groups score differently?

Differences are expressed as Cohen's d, the gap between two group means in standard deviation units. By convention, d below 0.20 is negligible, 0.50 is moderate and 0.80 is large.

Comparison	n	d	Interpretation
Women vs men	64 / 56	**0.11**	**Negligible** NO MATERIAL DIFFERENCE
English vs other first language	95 / 25	**0.24**	**Small** MONITOR

Differences are calculated on assessment scores among the 120 people in the study. No age comparison is reported: the cohort spans 20 to 35 with a median band of 21–25, which is too narrow to support a meaningful split.

Does it predict equally well?

Score differences and predictive bias are not the same thing. A group can score lower on average and still have its performance predicted just as accurately — in which case the assessment is doing its job.

Differential prediction was tested by regressing each manager rating on the matching assessment score, group membership, and their interaction. Neither the slope difference nor the intercept difference was significant for either comparison. A given score therefore carries the same performance expectation regardless of group.

WHAT THIS ANALYSIS CAN AND CANNOT TELL YOU

These figures describe people who were **hired**. Applicants who were screened out are not in the study, so this is not an analysis of how the assessment affects who gets through — that requires applicant-level data across the full pool, which Talegent can run separately against your campaign records.

Group sizes here are also modest. Both comparisons should be treated as a baseline to monitor as volume accumulates, rather than a settled finding. Where sample sizes permit, the same analysis can be extended to ethnicity and other characteristics.

06 Recommendations

What to change, and when

The competencies chosen for this solution were the right ones. The main opportunity now is to stop treating them as equally important.

Keep the solution in place

The evidence supports continued use of the assessment for Graduate Programme Employee screening. Six of the seven competencies and the cognitive measure each predicted performance on their own, at levels at the upper end of what selection research achieves against manager ratings. No competency needs removing.

Introduce weighting

The seven competencies currently contribute equally to the overall score. This study now shows they do not contribute equally to performance, so weighting them in proportion to how strongly each relates to its matching manager rating would raise the predictive power of the score without adding a minute to candidate time.

Competency	r	Suggested change
Learning Agility	.28	Increase weight
Achievement	.25	Increase weight
Teamwork	.23	Hold at current weight
Resilience	.21	Hold at current weight
Communication	.20	Hold at current weight
Self-Insight	.19	Reduce slightly
Ingenuity	.12	Reduce, but retain

Talegent can configure these weightings in the platform. We would recommend modest adjustments rather than dramatic ones, because weights estimated from a single sample carry their own error.

Weight the cognitive measure alongside the competencies

Logical Reasoning reached significance on six of the seven rated behaviours it was tested against, most strongly with learning new processes and procedures (r = .29) and understanding technical information (r = .26). For a cohort with little work history, reasoning ability carries more of the predictive load than it would for experienced hires, because most of the work is unfamiliar. It deserves a meaningful share of the overall score rather than being treated as a screening hurdle.

Follow this study with a predictive one

A concurrent design answers the question quickly but at a discount. Its sample is limited to people who were selected in and stayed, and its ratings come from managers who already know them. Running a **Predictive Solution Verification** on the next intake — the assessment at application, ratings after three months in the programme — would test the same question on a wider spread of scores and with the criterion collected independently. The results here are a sound basis for acting now, and a predictive study is how they get confirmed.

Revalidate in 24 months

Validity evidence describes the programme as it was during the study window. A repeat study in two years, or sooner if the programme structure changes, will confirm the weightings and track the group differences reported in section 05.

Appendix A

Validation results

Each competency score at application, correlated with the manager's rating of that same competency, collected at the same time. N = 120 for all analyses.

Table A1. Criterion-related validity of each competency scale against its matching manager rating.

Competency	r	95% CI	p
Learning Agility	**.28**	.11 to .44	.002
Achievement	**.25**	.07 to .41	.006
Teamwork	**.23**	.05 to .39	.011
Resilience	**.21**	.03 to .38	.021
Communication	**.20**	.02 to .37	.029
Self-Insight	**.19**	.01 to .36	.038
Ingenuity NS	.12	−.06 to .29	.192

Correlations are reported exactly as observed, with no statistical corrections applied. Appendix B explains why this understates the relationship, and by roughly how much.

READING THE TABLE

The confidence interval is the range within which the true relationship most likely sits. Where that range excludes zero, as it does for six of the seven competencies, the relationship can be relied on. Ingenuity's interval crosses zero, so on this evidence its relationship with performance is not established.

Appendix A · continued

Cognitive assessment

Cognitive assessment and work performance

The relationship between Logical Reasoning and the work performance criterion was analysed by Pearson correlation. Significant positive relationships were found with understanding technical information, solving complex problems, and a composite critical thinking score.

Table A2. Correlations between the work performance criterion and Logical Reasoning. N = 120.

Work performance criterion — ability to...	r	p
Learn new processes and procedures	**.29**	.001
Understand technical information	**.26**	.004
Solve complex problems	**.24**	.008
Think critically (domain score)	**.22**	.016
Assimilate information from several sources	**.21**	.021
Achieve results (domain score)	**.19**	.038
Acknowledge when they have made a mistake NS	.13	.157

Appendix A · continued

Competency detail

Competency scores and work performance

Correlations between each competency and the specific manager-rated behaviours most closely associated with it. Reported to two decimal places; p values two-tailed.

Table A3. Competency scales against individual rated behaviours. N = 120.

Competency	Work performance criterion — ability to…	r	p
Learning Agility	Pick up unfamiliar work quickly	.30	< .001
Learning Agility	Seek out and act on feedback	.26	.004
Learning Agility	Adjust approach when something is not working	.23	.011
Achievement	Work consistently hard	.27	.003
Achievement	Follow through on commitments	.24	.008
Teamwork	Contribute in group settings	.25	.006
Teamwork	Prioritise team goals over individual credit	.21	.021
Resilience	Stay composed when work is unclear	.23	.011
Resilience	Recover quickly from critical feedback	.20	.029
Communication	Explain their work clearly to non-specialists	.22	.016
Communication	Write concisely and accurately	.19	.038
Self-Insight	Recognise their own limitations	.21	.021
Self-Insight	Ask for help at the right time	.18	.049
Ingenuity	Suggest better ways of doing things NS	.14	.127
Ingenuity	Generate original solutions NS	.10	.277

Appendix B

Method, standards and limitations

Professional standards

This study was designed and reported in accordance with the **SIOP Principles for the Validation and Use of Personnel Selection Procedures** and the **Standards for Educational and Psychological Testing** (AERA, APA and NCME).

Analysis

Relationships were tested by Pearson product-moment correlation, two-tailed, with significance set at .05. At N = 120 the smallest correlation reaching significance is .18. Confidence intervals were derived by Fisher z transformation. All correlations in the body of this report are reported **exactly as observed**, with no statistical corrections applied.

Why the observed figures are conservative

Two features of any predictive study push observed correlations below the true relationship.

The first is **restricted range**, and a concurrent design suffers it twice over. Programme members were selected in, and then stayed, so the spread of assessment scores among them is narrower than among applicants on both counts. Measuring a relationship across part of a range always produces a smaller number than measuring it across all of it. For this sample the standard deviation of assessment scores among programme members was 78% of that among applicants to the programme, against the 82% typical of a predictive study.

The second is **criterion unreliability**. Manager ratings are the practical measure of performance, but two managers rating the same person agree only moderately — here an inter-rater ICC of .58. No predictor can correlate with a criterion more strongly than that criterion agrees with itself.

Applying the standard corrections for both, the strongest competency rises from .26 to **.40** and the weakest significant one from .17 to .26. These corrected figures are the ones most comparable with published validity research. We report the uncorrected values in the body of this report because they describe what was actually measured in this organisation.

Limitations

- **Statistical power.** At N = 120 the study had 79% power to detect r = .25, 59% for .20 and 37% for .15. A concurrent design is limited to current incumbents, so it will usually be less powerful than a predictive study, and modest true effects may not reach significance.
- **Single-source criterion.** Performance data came from one manager per person in most cases. Ratings carry known leniency and halo effects.
- **Subgroup sizes.** The fairness comparisons rest on modest group sizes and should be monitored rather than treated as settled.
- **Restricted range.** Only hired applicants could be studied, so all figures understate the relationship across the full applicant pool.
- **Local evidence.** Findings describe the Graduate Programme at Globex during the study window and should not be transported to other roles without a job comparison.

> **ON THE SIZE OF THE NUMBERS**
>
> Readers comparing this report with others should note that every correlation in the body is **observed and uncorrected**. Many published studies and vendor reports quote corrected coefficients without saying so, and will therefore appear to show stronger results from the same underlying evidence. The corrected figures for this study are given above for that comparison.