

TECHNICAL DOSSIER · VERSION 2.0

The science behind the **PMaps** Leadership Assessment

Framework, item design, scoring model, norming logic and the reliability and validity evidence behind every section — for organisations using leadership measurement to inform succession planning and promotion decisions.

Developed with TISS faculty



Quick reference

The questions assessment, HR and I/O psychology teams ask most often about a leadership instrument — and where each is answered in full.

QUESTION	SHORT ANSWER	SEC.
Underlying methodology	A proprietary, behaviourally anchored competency framework built with TISS organisational psychology faculty and industry SMEs.	02
Research behind it	PMaps' own programme: expert workshops, content validation, a 1,274–manager pilot, confirmatory factor analysis and a criterion study against manager ratings.	02, 09
What is measured	Five leadership pillars with eleven sub-competencies, a configurable values module, and configurable cognitive modules.	03–05
How questions were designed	Written against named behaviours, workplace-anchored, balance-keyed, panel-reviewed at CVI > 0.83, then screened on pilot statistics.	06
Scoring and analysis	Item → trait → sub-competency → pillar composites, CFA-weighted, normed against the relevant group, then banded.	07
Where benchmarks come from	Industry benchmarks for the target level, reviewed with the client. Readiness is conjunctive — every section must clear its cut-off.	08
Reliability	Cronbach's α 0.74 – 0.82 on trait scales; four-week test–retest 0.78.	09
Validity	CVI > 0.83 content; CFA RMSEA 0.051 structural; criterion $r = 0.34 - 0.41$ against manager ratings.	09
Fairness and bias	Controlled at authoring, panel review, item screening and norming; group pass rates monitored after launch.	10
What the report contains	Summary, scorecard against benchmarks, competency heat-map, detailed profiler, readiness band and a development plan.	11
Customisation	Values set, cognitive modules, item counts, time limits, benchmarks and report language are configured per role and organisation.	05, 12
Data handling	DPDP, GDPR, ISO 27001 and SOC 2 aligned practices, candidate consent, encryption in transit and at rest, VAPT testing.	12

MEMBERSHIPS, AFFILIATIONS AND RECOGNITION



01 Purpose and scope

The PMaps Leadership Assessment identifies and develops high-potential managers and leaders by measuring observable leadership behaviour, the psychological constructs that drive it, the values a leader personally stands for, and the cognitive ability a senior decision-maker relies on. It is built for three decisions.

Succession planning

Who sits on the bench for each critical role, and how deep that bench actually is.

Promotion readiness

Whether a leader clears the standard for the next level today, and what holds them if not.

Targeted development

Which two or three behaviours, if strengthened, would move the individual and the tier the most.

Design principles

- **Behaviour first.** Every score traces back to items written for a named, observable leadership behaviour — nothing is inferred from a general trait profile.
- **Level-specific.** The standard is the target level, not the current one; the instrument is calibrated for mid-to-senior managerial transitions.
- **Decision-ready output.** Every report ends in a readiness call with the reason attached, so a review panel can act on it in one sitting.
- **Configurable, not bespoke.** The validated core stays fixed; values, cognitive modules and benchmarks flex to the organisation.

What the assessment covers

BEHAVIOUR

Five leadership pillars, eleven sub-competencies.

VALUES

Situational items on the principles a leader stands for.

COGNITIVE

Timed, objectively scored reasoning and accuracy modules.

Scope note. The assessment is designed as a decision-support input, not a decision. It is intended to be read alongside performance history, role context and manager judgement — never as the sole basis for a promotion or an exit.

02 How the assessment was built

The instrument was developed in collaboration with organisational psychology faculty at the **Tata Institute of Social Sciences (TISS)**, together with PMaps content experts and industry subject-matter experts. It rests on PMaps' proprietary competency framework and behavioural science research. All items and constructs are internally developed and empirically validated — the assessment does not use, license or display any external personality trait model.

STAGE	WHAT HAPPENED	EVIDENCE PRODUCED
1	Competency definition — expert workshops with TISS faculty and industry SMEs to name and define the leadership behaviours that separate levels of managerial work.	Validated clusters and behavioural definitions
2	Item authoring — co-authored item pools written against each behavioural definition, at roughly three times the final item count.	Draft item pool with construct mapping
3	Content validation — every item rated for relevance by an independent expert panel and screened on a content validity index.	CVI > 0.83 per retained item
4	Pilot administration — 1,274 managers across Banking, FMCG and Technology in EMEA and APAC.	Item statistics, scale reliabilities
5	Structural validation — confirmatory factor analysis to test that items load on the competencies they were written for.	CFA model fit, RMSEA 0.051
6	Norming — percentile tables built separately for People Leaders and Individual Contributors, with industry norm groups.	Norm bands and industry benchmarks
7	Criterion study — scores correlated against manager performance ratings to test whether the assessment predicts what it claims to.	Criterion validity $r = 0.34 - 0.41$

Kept current

Norms, item statistics and criterion evidence are refreshed on a rolling basis as administration data accumulates.

Client-specific studies

Local validation against your own performance data can be run where sample sizes allow, and folded into your norms.

03 Competency architecture

Leadership behaviour is measured on **five pillars**, each built from named sub-competencies. Every pillar score is composed from its sub-competencies; every sub-competency is composed from items written specifically for it.

Managerial Acumen

Reading a business situation quickly and accurately, obtaining the information that matters, weighing the alternatives and choosing a course of action that protects the organisation's position.

Problem Solving

Planning & Organization

Strategic Thinking

Self Management

Owning the consequences of one's actions and handling the emotional load of the role maturely — the composure that keeps a unit steady under pressure.

Sense of Ownership

Emotional Maturity

Team Capabilities

Investing time in relationships, handling friction openly, and creating a cooperative environment where people contribute without hesitation.

Team Collaboration

Conflict Management

Team Management

Passing on knowledge and experience so others grow, and working comfortably across cultural and background differences to keep the business running.

Team Mentoring

Understanding Diversity

Drives

The internal energy to set high standards and push personal limits, and the ability to mobilise that energy in others toward shared goals.

Work Motivation

Motivating Others

Where an organisation runs its own leadership framework, PMaps maps it onto these validated constructs — your language in the report, validated measurement underneath.

04 Values module

Leadership behaviour tells you *how* someone leads. The values module tests *what they stand for* — whether a principle is one the leader personally holds or one they can merely recite. Items are situational: each presents a realistic workplace trade-off and asks the leader to choose and prioritise, so the socially expected answer is not always the scoring answer.

Customer ownership

Treating client outcomes as a personal mission, not a departmental target.

Personal integrity

Holding oneself to the standard one sets for others, including when it is costly.

Team first

Choosing the collective result over individual credit when the two compete.

Authored to your values. The three themes above are illustrative. The module is written to the organisation's own stated values or leadership credo — typically three to five values, each with a defined set of observable behaviours — and content-validated before launch. Values scores feed the readiness decision; strengths and development notes are drawn from behaviour and cognitive ability only.

05 Cognitive modules

Cognitive sections measure the raw accuracy and reasoning a senior decision-maker depends on. They are short, timed and objectively scored — there is a right answer.

MODULE	WHAT IT MEASURES	WHY IT MATTERS AT LEADERSHIP LEVEL
Verbal Reasoning	Interpreting written information, following an argument and drawing accurate conclusions from dense text.	Policy, regulation and board papers arrive as prose.
Numerical Ability	Working with numbers and data — computation, interpretation and reading a trend.	Whether a leader owns the numbers or depends on someone else to read them.
Attention to Detail	Accuracy and error-checking under time pressure; spotting variation in information.	Errors that pass a leader's desk are expensive to reverse.
Abstract Reasoning (optional)	Pattern recognition and inductive logic independent of language or numeracy.	Added for roles demanding high strategic analysis.

Cognitive sections are customisable. The modules included, the item counts and the time limits are set to the requirement of the role and the organisation — a commercial population may weight Numerical Ability, a risk or compliance population Attention to Detail, a strategy population may add Abstract Reasoning. Modules can also be dropped entirely where cognitive ability is not part of the promotion standard.

06 Item design and construction

The behavioural core is the **PMaps Trait Inventory** — a proprietary inventory of 144 items assessing the behavioural dimensions that underlie the leadership competencies. A typical full leadership battery runs five sections and around 190–200 items in roughly 90 minutes; the exact construct is configured per client.

How an item is written

1. **Derive items from traits and the work behavior they drive.** Each competency is broken down into the underlying traits and the observable work behaviors those traits produce. The scored items then sum into the sub-competency and competency score basis the weightage defined, so every score traces back to specific traits and behaviors.
2. **Anchor in workplace context.** Items describe recognisable managerial situations rather than abstract self-descriptions, which reduces aspirational answering and improves face validity with senior candidates.
3. **Balance the keying.** Positively and negatively keyed items are mixed within each scale to break response-set patterns such as straight-lining and acquiescence.
4. **Control the language.** Plain, culturally neutral wording; no idiom, no regional reference, no industry jargon. Reading level is held well below the population's working level so the behavioural sections do not become a hidden literacy test.
5. **Panel review before data.** An independent expert panel rates every item for relevance to its intended construct. Items below a content validity index of 0.83 are rewritten or discarded before pilot.
6. **Statistical screening after data.** Pilot data screens each surviving item on distribution, discrimination, item–total correlation and factor loading. Items that fail to behave — poor loading, cross-loading onto a second competency, or near-zero variance — are cut.

Item formats in use

BEHAVIOURAL SECTIONS

Rating-scale and forced-choice items on workplace behaviour. Untimed but time-tracked; no right answer, so scoring is normative.

VALUES SECTION

Situational judgement items with competing defensible options, scored against an agreed key of preferred and least-preferred responses.

COGNITIVE SECTIONS

Objective items with a single correct answer, strictly timed and delivered from a randomised bank to limit exposure.

INTEGRITY CHECKS

Response-time, consistency and completion checks run in the background; flagged sessions are reviewed before a report is released.

07 Scoring model

Scoring runs bottom-up. Nothing is entered by hand and no scorer judgement is involved; the same response set always produces the same report.

1 Item response → trait score

Keyed responses are summed into the underlying behavioural dimension, with reverse-keyed items inverted.

2 Trait score → sub-competency

Each sub-competency draws on a defined, weighted set of dimensions. Weights come from the CFA loadings, not from opinion.

3 Sub-competency → pillar

Pillar scores are the weighted composite of their sub-competencies; the leadership overall score is the composite of the five pillars.

4 Raw → normed

Raw composites are converted against the relevant norm table — People Leader or Individual Contributor, industry group where available — and reported on a common scale.

5 Normed → band and narrative

The normed score sets the band, and the band selects validated narrative statements written for that construct at that level.

Score bands

BAND	POSITION VS NORM	HOW TO READ IT
High	Clearly above the comparison group	A dependable strength. Deploy it — give the leader work that uses it and let them coach others in it.
Above Average	Above the comparison group	Solid and reliable in familiar conditions; may need support as scope or ambiguity increases.
Average	In line with the comparison group	Functional today. A development area if the next role depends on this behaviour.
Low	Below the comparison group	A focus area. Expect it to show under pressure; pair development with a structural support in the role.

Bands are relative to a norm group, not absolute verdicts. A "Low" on one sub-competency in an otherwise strong profile is a development instruction, not a disqualification.

08 Norming and benchmark logic

A score only means something against a comparison group. Percentile tables are built separately for **People Leaders** and **Individual Contributors**, and by industry group where the sample supports it, so a banking leader is compared with banking leaders rather than a general working population.

How the promotion benchmark is set

Benchmarks are cut-offs on each section that define the standard for the target level. They are set on **industry benchmarks** — the distribution of scores among leaders already performing at that level in comparable organisations — and then reviewed with the client's HR centre of excellence before launch. Each section carries its own cut-off, because the sections are not interchangeable: a numerically strong leader with weak accuracy is a different risk from the reverse.

The gating rule

Readiness is *conjunctive*, not compensatory. A leader is rated **High Potential** only when every section benchmark is met. Missing any one — including a single cognitive section — places them in **Emerging Potential**. Strength in one area is not allowed to paper over a gap in another, because at leadership level the gap is what fails in practice.

HIGH POTENTIAL

Clears every benchmark. Promotion-ready on the assessment evidence; development notes become stretch, not remediation.

EMERGING POTENTIAL

Misses one or more benchmarks. The report names which, so the route to the top band is explicit and time-bound.

Reading a group result

Bench depth

The share of a tier in the top band is the practical read on succession depth for that layer.

Where the tier is held

Pass rates per benchmark show whether the constraint is leadership behaviour or cognitive ability — two different interventions.

Collective strengths

Competencies where the tier clusters high are the capabilities the organisation can lean on now.

Development priorities

Shared low sub-competencies justify a cohort programme rather than individual coaching.

09 Reliability and validity evidence

All figures below are from PMaps' own development and validation programme for this instrument. No external instrument or trait model is used, licensed or reported.

METRIC	TARGET	ACHIEVED	WHAT IT TELLS YOU
Content validity index (per item)	> 0.80	> 0.83	Independent experts agree each item measures the competency it claims to.
Internal consistency (Cronbach's α , trait scales)	≥ 0.70	0.74 – 0.82	Items within a scale hang together — the scale measures one thing.
Test–retest reliability (4-week interval)	≥ 0.70	0.78	Scores are stable over time — you are measuring a trait, not a mood.
Structural validity (CFA model fit, RMSEA)	≤ 0.06	0.051	The competency structure in the report is the structure present in the data.
Criterion validity (r with manager ratings)	≥ 0.30	0.34 – 0.41	Scores relate to how leaders are actually rated on the job.
Pilot and norming sample	—	n = 1,274	Managers across Banking, FMCG and Technology in EMEA and APAC.

Reading the criterion figure

A correlation of 0.34 – 0.41 against manager ratings is a strong result for a behavioural instrument, and is what makes the assessment useful as one input among several. It is not a prediction of any individual's future performance.

Standard error and near-misses

Every score carries measurement error. Where a leader sits within a few points of a benchmark, the recommended practice is a structured interview or a targeted re-test on that section rather than a hard decision on the number alone.

Full technical output — item-level statistics, factor loadings, scale intercorrelations and norm tables — is available under NDA to your assessment or I/O psychology team on request, with a walkthrough by the team that built the instrument. Write to ssawant@pmaps.in.

10 Fairness and bias control

Fairness is engineered into the instrument at four points, and monitored after launch.

At authoring

Culturally neutral language, no idiom or region-specific reference, and controlled reading level so the behavioural sections do not become a language test.

At panel review

The expert panel screens items for content that could advantage or disadvantage a group irrespective of leadership capability.

At item screening

Pilot data is examined for items that function differently across available demographic groups; items that behave differently for equally capable candidates are removed.

At norming

Separate norm groups for People Leaders and Individual Contributors prevent structural disadvantage to leaders coming from a specialist track.

After launch

Group-level pass rates are monitored so adverse impact is visible early. Where a section shows disparity that is not explained by role requirements, the section is reviewed — the assessment changes, not the candidates.

Accessibility and accommodation. Timed cognitive sections can be delivered with extended time where a documented need exists; the report notes that an accommodation was applied without disclosing the reason. Every candidate receives the same instructions, the same practice items and the same interface.

11 Reading the report

Each leader receives an individual report; the organisation receives group and tier views built from the same scores. The elements are consistent, so a reviewer learns the format once.

Leadership overview and summary

A short narrative snapshot generated from the scored competencies — the strongest signals, the softer ones, and the shape of the leader overall. It restates scored evidence in plain language; it never introduces a judgement that is not in the scores.

Leadership scorecard

Every pillar, value and cognitive module on one page, each plotted against its benchmark range with its band label. This is the page a review panel works from.

Competency heat-map

A colour read across all assessed competencies for fast triage at group level.

Green — strength

Amber — develop

Red — focus

Detailed profiler

For each pillar and sub-competency: the construct definition, the score and band, and a behavioural commentary. Commentary is written in probabilistic language — "may", "tends to" — because the assessment describes likelihood of behaviour, not certainty.

Readiness and development plan

Top strengths, areas to develop, the readiness band with the reason it was assigned, best-fit next step, what to watch under pressure, and development actions linked to micro-learning journeys.

Worked example. A leader scores High on Planning & Organization, Sense of Ownership and Emotional Maturity, Above Average on Team Management, Low on Strategic Thinking, and misses the Attention to Detail benchmark by one band. The report reads: a composed, dependable people-leader whose readiness is held by accuracy rather than by leadership behaviour — so the plan is a second-pass quality routine plus ownership of a twelve-month forward plan with a senior sponsor, not a generic leadership course.

12 Privacy, limitations and support

Data privacy and compliance



- DPDP, GDPR, ISO 27001 and SOC 2 aligned data practices, with regular VAPT testing.
- Explicit candidate consent captured before the assessment begins, with a stated purpose for the data.
- Encryption in transit and at rest; role-based access to reports.
- Documented retention and deletion schedules agreed with the client.
- Candidate feedback rights supported — a development-facing version of the report can be released to the individual.

Limitations and appropriate use

- The assessment measures likelihood of behaviour, not certainty. Commentary is deliberately probabilistic.
- It is a decision-support input, and should not be the sole basis for a promotion, a pay decision or an exit.
- Scores are relative to a norm group. Change the comparison group and the band can move; the underlying behaviour has not.
- A single administration reflects the leader at one point in time. Re-test after a meaningful development period, not within weeks.
- Results should be interpreted by someone trained on the framework, and fed back to the leader. An unexplained report is a wasted one.

Customisation and consulting support

PMaps works with your HR centre of excellence to fit the instrument to your organisation rather than the other way round.

Map your framework. Align your internal leadership model to the validated competency structure, keeping your language in the report.

Build success profiles. Define the standard for each critical role and set the benchmarks that gate readiness for it.

Configure the construct. Choose the values set and the cognitive modules, item counts and time limits that match the role.

Facilitate feedback. Run manager briefings, one-to-one feedback and development-planning workshops after results land.

What a typical engagement looks like

WEEK 1

Align

Map your framework, agree the target level and set the benchmarks.

WEEK 2

Configure

Finalise the values set and cognitive modules; brief candidates and managers.

WEEKS 3–4

Assess

Leaders complete the battery online; flagged sessions reviewed before release.

WEEKS 5–6

Act

Group and individual reports, calibration session, then feedback and plans.

See it on your own roles

Walk through a live report, or run the assessment yourself on a free trial. The full validation pack — item statistics, factor loadings and norm tables — is available on request.

[Book a demo](#)

[Start 7-day free trial](#)

Talk to us

ssawant@pmaps.in

+91 85913 20212 · call or WhatsApp

pmapstest.com