

A communication and behavioural assessment framework, openly evaluated.

A research briefing on the Silk Clarity framework: research foundations, construct definitions, assessment methodology, sample outputs, current evidence, explicit non-claims, and the path to independent validation.

Validation status framework used throughout

FOUNDATION

Established research base

PRELIMINARY

Pilot evidence

IN PROGRESS

Current validation work

PLANNED

Independent validation



CONTENTS

Thirty sections, structured for review.

Research foundations, construct definitions, assessment methodology, sample outputs, current evidence, explicit non-claims, and the validation roadmap.

00	Validation Status Framework How every claim is labelled	14	Construct · Adaptability Definition, indicators, non-claims
01	Positioning What Silk is, and is not	14 ^B	Derived · Executive Presence Emergent from Clarity, Credibility, Composure, Influence
02	Research Foundations · Overview Seven bodies of literature	15	Assessment Methodology Five-stage pipeline
03	Trust & Credibility Mayer, Davis, Schoorman; McAllister	16	Sample Output · Evidence Extraction Candidate response, evidence, observations
04	Communication Competence Spitzberg & Cupach	17	Sample Output · Development Guidance Dimension observations, recommendations
04 ^B	Language Proficiency & the CEFR Council of Europe; Canale & Swain; Bachman & Palmer	17 ^B	Communication Development Framework Six dimensions · challenges · development actions
05	Executive Presence Hewlett; Sy & van Knippenberg	18	Reliability Framework Targets, not claims
06	Leadership Communication Mayfield; de Vries	19	Validity Framework Content, construct, criterion
07	Cultural Intelligence & Behavioural Assessment Earley & Ang; Lievens	20	Pilot Programme Preliminary evidence (n ≈ 30)
08	Framework Overview Six dimensions, one emergent analysis	21	Ethics & Data Handling Consent, GDPR, bias monitoring
09	Construct · Clarity Definition, indicators, non-claims	22	Limitations Stated openly
10	Construct · Credibility Definition, indicators, non-claims	23	Explicit Non-Claims What Silk does not measure
11	Construct · Integrity Definition, indicators, non-claims	24	Validation Roadmap Seven-stage scientific maturity
12	Construct · Composure Definition, indicators, non-claims	25	Research Partnerships Four open invitations
13	Construct · Influence Definition, indicators, non-claims	26	Invitation Examine the framework

00 · VALIDATION STATUS FRAMEWORK

What we claim, and on what basis.

This document distinguishes carefully between research foundations, preliminary pilot evidence, validation work currently in progress, and validation that remains planned future work.

FOUNDATION

Established research base

Communication science, organisational psychology, leadership research, and behavioural assessment methodology that informs the framework.

Language used: 'draws upon established research in...'

PRELIMINARY

Pilot evidence

Findings from the initial pilot cohort: scoring consistency, completion behaviour, user feedback, observed behavioural patterns.

Language used: 'preliminary evidence suggests...!' / 'pilot data indicates...!'

IN PROGRESS

Current validation programme

Reliability studies, human-model agreement work, expert review, and construct validation currently being run on the platform.

Language used: 'currently being evaluated' / 'data collection is underway'

PLANNED

Path to independent validation

External researcher partnerships, university collaborations, peer review, and the publication strategy.

Language used: 'planned future work' / 'independent validation will examine...!'

01 · POSITIONING

What Silk is. What Silk is not.

The framework is best understood as a research-informed communication and behavioural assessment system. The distinction from performance prediction is deliberate and material.

SILK IS

A behavioural assessment platform.

- A structured conversational assessment grounded in cross-cultural leadership, trust, communication and organisational psychology research.
- An evidence-based observation system: every score is linked to observable behaviour demonstrated during the assessment.
- A six-dimension behavioural framework: Clarity, Credibility, Integrity, Composure, Influence, Adaptability.
- A development guidance tool for individuals and a structured evidence layer for recruiters and hiring panels.

SILK IS NOT

A performance prediction engine.

- Not a personality test. Silk does not diagnose personality types, traits, or psychological conditions.
- Not a hiring decision tool. Silk does not recommend, accept, or reject candidates.
- Not a performance predictor. Silk does not forecast future job performance or career outcomes.
- Not a substitute for human judgement. Results inform interviews, references, and structured evaluation — never replace them.

WHY THIS DISTINCTION MATTERS

The distinction between behavioural assessment and performance prediction is the centrepiece of Silk's scientific credibility. Behavioural assessment is defensible. Performance prediction requires evidence Silk does not yet hold.

02 · RESEARCH FOUNDATIONS · OVERVIEW

Seven bodies of literature inform the framework.

Each of the following six pages summarises one body of established research, the evidence base it provides, and how it informs the Silk framework. This is briefing-grade, not literature-review depth.

Trust & Credibility

Mayer, Davis & Schoorman (1995); McAllister (1995); Colquitt et al. (2007).

Establishes ability, benevolence, integrity as the canonical trust antecedents — and credibility as the foundation of professional influence.

Communication Competence

Spitzberg & Cupach (1984; 2011); Wiemann (1977); Rubin (1990).

Operationalises communication as an observable, behavioural construct — appropriate and effective in context.

Language Proficiency & CEFR

Council of Europe (2001; 2020); Canale & Swain (1980); Bachman & Palmer (1996).

Establishes observable communication performance descriptors across proficiency levels and contexts — language as demonstrable competence, not background trait.

Executive Presence

Hewlett (2014); Sy & van Knippenberg (2021); Goffee & Jones (2009).

Frames executive presence as an emergent perception drawn from gravitas, communication, and appearance — not a personality trait.

Leadership Communication

Mayfield & Mayfield (2018); de Vries et al. (2010); Bass & Avolio (1994).

Connects communication style — direction-giving, empathy, meaning-making — to leadership effectiveness across cultures.

Cultural Intelligence

Earley & Ang (2003); Van Dyne et al. (2012); Ng & Earley (2006).

Defines CQ as a behaviourally observable capability to function effectively across cultural contexts.

Behavioural Assessment Methodology

Lievens & Sackett (2017); Motowidlo et al. (1990); Arthur et al. (2003).

Establishes situational judgement and behavioural observation as evidence-based assessment methodologies with documented validity.

03 · TRUST & CREDIBILITY

The science of why people are believed.

Mayer, Davis and Schoorman's (1995) integrative trust model identifies three antecedents of professional trust: ability (competence and expertise), benevolence (positive intent), and integrity (consistency between word and action). These three constructs have been replicated across decades of empirical work and remain the canonical framework for organisational trust. Colquitt and colleagues' (2007) meta-analysis (k = 132) confirms each antecedent has independent predictive validity for trust judgements and downstream behaviours. McAllister (1995) further distinguishes cognition-based trust (from competence signals) from affect-based trust (from interpersonal warmth) — both demonstrably present in workplace communication. Credibility research (Hovland, Janis & Kelley; O'Keefe) consistently shows that perceived expertise and trustworthiness — communicated through observable verbal and behavioural cues — are the two strongest determinants of message acceptance and influence in professional contexts.

KEY FINDINGS

- Trust is constructed from three observable signals: ability, benevolence, integrity.
- Credibility is conferred, not declared — and is read from behaviour, not credentials.
- Inconsistency between stated value and demonstrated behaviour erodes integrity faster than any other signal.
- Competence signalling and warmth signalling are distinct and both required for sustained professional trust.

HOW THIS INFORMS SILK

Silk's Credibility and Integrity dimensions are operationalised directly against this literature. Observable indicators (evidence-based reasoning, accountability language, ownership of outcomes, consistency under challenge) are scored as behavioural signals associated with each construct — never as personality inferences.

04 · COMMUNICATION COMPETENCE

Communication as an observable behavioural construct.

Spitzberg and Cupach (1984; 2011) define communication competence as the demonstrated ability to communicate in ways that are both appropriate (fitting the context, audience, and norms) and effective (achieving the communicator's legitimate goals). Critically, competence is observable behaviour — not a self-perception, not a personality trait, not an aptitude. Wiemann's (1977) Communicative Competence Scale and Rubin's (1990) Communication Competence Assessment Instrument establish that competence can be reliably evaluated from observed conversational behaviour, with documented inter-rater reliability when criteria and indicators are explicit. More recent reviews (Spitzberg, 2015; Beebe & Mottet, 2016) emphasise that competence is contextual: the same behaviour may be competent in one professional setting and incompetent in another. Assessment must therefore evaluate behaviour against the audience and stakes the communicator is responding to.

KEY FINDINGS

- Communication competence is behavioural — it is what people do, not what they are.
- Competence requires both appropriateness and effectiveness; either alone is insufficient.
- Reliable assessment requires explicit observable indicators, not impressionistic ratings.
- Competence is contextual: assessment must include the communication setting as a variable.

HOW THIS INFORMS SILK

Silk's Clarity, Influence, and Adaptability dimensions extend directly from the competence literature. The assessment provides structured professional scenarios so that observed behaviour can be evaluated against an appropriate and effective standard for the specific context.

04B · LANGUAGE PROFICIENCY & THE CEFR

A descriptor system for observable communication performance.

The Common European Framework of Reference for Languages (Council of Europe, 2001; Companion Volume 2020) operationalises language proficiency as observable communication performance across six reference levels (A1-C2). Each level is defined by 'can-do' descriptors — what a user can be observed to do — across reception, production, interaction and mediation.

The CEFR matters here because it converges with the broader communication-competence literature on a single methodological premise: proficiency is established from observable performance against explicit criteria, not from self-perception, aptitude or trait inference. Canale and Swain's (1980) communicative competence model and Bachman and Palmer's (1996) language-ability framework — both foundational to the CEFR — extend this further, distinguishing grammatical, sociolinguistic, discourse and strategic competences as separable observable capabilities. For Silk, the CEFR is not the construct being measured. It is the established empirical reference point for how observable communication performance varies with language proficiency — a variable the validation programme must control for and discriminate against.

KEY FINDINGS

- Language proficiency is operationalised as observable performance, not as trait or aptitude.
- CEFR descriptors define what a user can be observed to do across reception, production, interaction and mediation.
- Communicative competence is multi-componential: grammatical, sociolinguistic, discourse, strategic.
- Convergent and discriminant validation against CEFR is methodologically standard for any communication-performance instrument operating in multilingual professional contexts.

HOW THIS INFORMS SILK

Silk does not measure CEFR level and is not a substitute for a language test. The CEFR informs Silk in three ways: scenarios are calibrated for B2+ professional contexts so that observed behaviour reflects communication craft and not foundational proficiency; the validation programme includes CEFR-aligned convergent and discriminant studies (see §19 and §24); and reported limitations explicitly address language-proficiency effects on Silk dimension scores (see §22).

05 · EXECUTIVE PRESENCE

An emergent perception, not a personality trait.

Hewlett's (2014) widely-cited research on executive presence, drawing on a sample of more than four thousand professionals, identifies presence as an emergent perception formed from three observable categories: gravitas (composure, confidence, decisiveness), communication (clarity, articulate expression), and appearance. Gravitas accounts for the largest share of the perception.

Sy and van Knippenberg (2021) and Goffee and Jones (2009) reinforce that presence is constructed by observers from observable behaviour — it is not a stable internal characteristic. This matters: it means presence can be developed through behavioural change, and it can be assessed without making personality claims.

Crucially, the research suggests that across cultures, individuals perceived as having strong executive presence demonstrate clarity of communication, professional credibility, emotional stability, integrity, and the ability to create confidence in others — a constellation, not a single trait.

KEY FINDINGS

- Executive presence is constructed by observers from behaviour, not transmitted as a trait.
- Gravitas — composure under scrutiny — is the strongest single component of perceived presence.
- Presence is developable: behavioural change drives perceived presence change.
- Cross-culturally, the components recur even where stylistic expression varies.

HOW THIS INFORMS SILK

Silk does not treat executive presence as a standalone scored dimension. Instead, an Executive Presence Analysis is reported as an emergent narrative drawn from observed behaviour across Clarity, Credibility, Composure, and Influence. This is consistent with the construct's research basis.

06 · LEADERSHIP COMMUNICATION

Communication style and observable leadership effectiveness.

Mayfield and Mayfield's (2018) Motivating Language Theory, validated across more than three decades of empirical work, identifies three observable communication categories that distinguish effective leaders: direction-giving (clarity of expectation), empathetic language (acknowledgement of the listener's experience), and meaning-making language (connecting work to purpose).

De Vries, Bakker-Pieper and Oostenveld's (2010)

Communication Styles Inventory identifies six observable communication style dimensions and demonstrates significant relationships between style and leader effectiveness ratings — independent of personality.

Bass and Avolio's transformational leadership work, and subsequent meta-analytic reviews (Wang et al., 2011; Judge & Piccolo, 2004), consistently identify leader communication behaviour — articulating vision, individualised consideration, intellectual stimulation — as central to effectiveness.

KEY FINDINGS

- Leader effectiveness is communicated, not transmitted by hierarchy.
- Direction-giving + empathetic language + meaning-making form a robust observable cluster.
- Communication style explains variance in leader effectiveness above and beyond personality.
- These behaviours can be developed and re-rehearsed; they are not fixed.

HOW THIS INFORMS SILK

Silk's Influence dimension is operationalised from this literature: persuasion, leadership communication, stakeholder awareness, alignment creation, communication impact, and decision framing — each evaluated as observable behaviour rather than as latent disposition.

07 · CULTURAL INTELLIGENCE & BEHAVIOURAL ASSESSMENT

Two converging methodological traditions.

Earley and Ang's (2003) Cultural Intelligence (CQ) framework identifies four observable dimensions — cognitive, metacognitive, motivational, and behavioural CQ — and a substantial validation literature (Van Dyne et al., 2012; Ng & Earley, 2006) supports both reliable measurement and predictive utility for cross-cultural effectiveness. Behavioural assessment methodology in organisational psychology — situational judgement tests (Motowidlo, McDaniel, Whetzel & Schmitt), behavioural observation scales, and assessment centre exercises (Arthur, Day, McNelly & Edens, 2003) — has accumulated meta-analytic evidence that observing how individuals behave in structured professional scenarios produces validity coefficients that meaningfully exceed CV review and unstructured interviewing. Lievens and Sackett's (2017) review confirms that behaviourally-anchored assessment, when grounded in explicit observable indicators and structured scoring rubrics, is among the most defensible forms of professional evaluation available to organisations.

KEY FINDINGS

- Cultural intelligence is a behavioural capability, not a personality trait.
- Situational judgement and behavioural observation have documented validity (meta-analytic).
- Structured scenarios with explicit indicators produce reliable, defensible evidence.
- Behaviourally-anchored assessment can be reported without making personality claims.

HOW THIS INFORMS SILK

Silk's Adaptability dimension extends from the CQ literature; the platform's structured conversational assessment extends from the behavioural assessment tradition. Both bodies of research support the position that observable communication and behavioural evidence can be reliably evaluated when criteria are explicit and the scoring rubric is published.

08 · FRAMEWORK OVERVIEW

Six dimensions. One emergent analysis.

The framework evaluates observable communication and behavioural indicators across six research-informed dimensions. Executive presence is reported as an emergent narrative, not a scored dimension.

DIMENSION 01

Clarity

Structure, precision, coherence, ability to explain complexity.

DIMENSION 02

Credibility

Evidence-based reasoning, analytical capability, subject-matter signal.

DIMENSION 03

Integrity

Accountability, ownership, consistency, transparency, ethical reasoning.

DIMENSION 04

Composure

Emotional regulation, stability under challenge, ambiguity tolerance.

DIMENSION 05

Influence

Persuasion, alignment creation, stakeholder awareness, decision framing.

DIMENSION 06

Adaptability

Audience awareness, context sensitivity, cross-cultural intelligence.

EMERGENT ANALYSIS · NOT A SCORED DIMENSION

Executive Presence

Reported as a narrative analysis drawn from observed behaviour across Clarity, Credibility, Composure, and Influence. Consistent with Hewlett (2014) and the research view that presence is an emergent perception, not a personality trait or standalone construct.

09 · CONSTRUCT · CLARITY

Structured, precise, coherent communication.

OPERATIONAL DEFINITION

The ability to communicate ideas with structure, precision and logical coherence — and to explain complexity without losing the listener.

OBSERVABLE INDICATORS

- Communication structure: explicit framing, ordering, signposting.
- Precision: word choice that narrows rather than blurs meaning.
- Coherence: each statement follows logically from the preceding one.
- Logical flow: explicit reasoning chain rather than assertion alone.
- Ability to explain complexity to a non-specialist audience.
- Conciseness: claims made without redundancy or hedging.

WHAT IT IS NOT

- Not articulateness as a personality trait.
- Not eloquence, vocabulary range, or rhetorical style.
- Not a measure of speaking ability in the abstract.
- Not predictive of writing or presentation skill outside structured conversation.

RESEARCH BASIS

Spitzberg & Cupach (1984); Wiemann (1977); Mayfield & Mayfield direction-giving language; Rubin (1990) Communication Competence.

10 · CONSTRUCT · CREDIBILITY

Competence and expertise, observed.

OPERATIONAL DEFINITION

The degree to which communication demonstrates competence, evidence-based reasoning and professional expertise sufficient to confer intellectual trust.

OBSERVABLE INDICATORS

- Competence signals: domain-relevant terminology used accurately.
- Evidence-based reasoning: claims linked to facts, examples, data.
- Strategic thinking: explicit framing of decisions and trade-offs.
- Analytical capability: decomposition of problems into evaluable parts.
- Decision quality: trade-offs surfaced, criteria stated.
- Subject-matter expertise demonstrated through specific examples.

WHAT IT IS NOT

- Not intelligence, IQ, or general cognitive ability.
- Not credentialism — credentials are not assessed.
- Not a measure of subject-matter knowledge against a fixed answer key.
- Not a personality trait or self-confidence measure.

RESEARCH BASIS

Mayer, Davis & Schoorman (1995) ability antecedent; Colquitt et al. (2007); Hovland, Janis & Kelley source credibility; O'Keefe (2002).

11 · CONSTRUCT · INTEGRITY

Accountability and consistency, demonstrated.

OPERATIONAL DEFINITION

The consistency, accountability and ethical reasoning observable in communication — particularly when accounting for difficult outcomes or contested decisions.

OBSERVABLE INDICATORS

- Accountability: ownership of decisions and outcomes.
- Consistency: stated values match described behaviour.
- Transparency: limits, uncertainties and unknowns named openly.
- Ethical reasoning: explicit consideration of stakeholders and consequences.
- Reliability signals: commitments framed as falsifiable rather than aspirational.

WHAT IT IS NOT

- Not a moral character judgement.
- Not an honesty test or deception detection.
- Not a personality inventory.
- Not predictive of conduct outside the assessment context.

RESEARCH BASIS

Mayer, Davis & Schoorman (1995) integrity antecedent; McAllister (1995); Colquitt et al. (2007); Treviño et al. (2006) on ethical leadership behaviour.

12 · CONSTRUCT · COMPOSURE

Stability and response quality under challenge.

OPERATIONAL DEFINITION

The capacity to maintain register, reasoning quality and response coherence under unscripted challenge, ambiguity or scrutiny.

OBSERVABLE INDICATORS

- Emotional regulation: tone holds when challenged.
- Stability during challenge: reasoning quality does not degrade.
- Response quality under pressure: structured response despite discomfort.
- Confidence calibration: confidence proportionate to evidence held.
- Tolerance for ambiguity: comfort answering without complete information.

WHAT IT IS NOT

- Not trait anxiety or any clinical construct.
- Not emotional suppression or affect dampening.
- Not introversion / extraversion.
- Not a psychological diagnostic.

RESEARCH BASIS

Hewlett (2014) gravitas; Klein (1998) recognition-primed decision-making; Lazarus & Folkman appraisal theory (behavioural observable correlates only).

13 · CONSTRUCT · INFLUENCE

Persuasion, alignment, and communication impact.

OPERATIONAL DEFINITION

The ability to create understanding, alignment and engagement — to move others through communication rather than authority.

OBSERVABLE INDICATORS

- Persuasive communication: claims framed for the listener's frame.
- Stakeholder awareness: explicit attention to who is in the room.
- Decision framing: making the operative question clear.
- Communication impact: claims that land without dilution or softening.
- Alignment creation: shared understanding constructed in the exchange.

WHAT IT IS NOT

- Not dominance, assertiveness as a trait, or charisma.
- Not a measure of likeability.
- Not manipulation, persuasion-against-interest, or political skill.
- Not predictive of negotiation outcomes in the abstract.

RESEARCH BASIS

Cialdini (2001) influence; Mayfield & Mayfield (2018) Motivating Language Theory; de Vries et al. (2010); Bass transformational leadership.

14 · CONSTRUCT · ADAPTABILITY

Audience awareness, context sensitivity, cultural intelligence.

OPERATIONAL DEFINITION

The ability to adjust communication appropriately to different audiences, contexts and situations — including observable cross-cultural intelligence.

OBSERVABLE INDICATORS

- Audience awareness: register adjusts to who is being addressed.
- Context sensitivity: response calibrated to professional stakes.
- Communication flexibility: shifts in approach as the conversation moves.
- Cultural intelligence indicators: behavioural CQ observable in code-switching.
- Behavioural adaptability: revision of approach when initial framing misfires.

WHAT IT IS NOT

- Not Openness to Experience or any Big Five facet.
- Not 'cultural fit' — Silk does not evaluate cultural fit.
- Not a personality trait.
- Not a predictor of cross-cultural performance in the field.

RESEARCH BASIS

Earley & Ang (2003) Cultural Intelligence; Van Dyne et al. (2012); Giles communication accommodation theory; Spitzberg context-appropriateness.

14B · DERIVED ANALYSIS · EXECUTIVE PRESENCE

An emergent communication construct — not a scored dimension.

Executive Presence is not measured directly. It is reported as a derived analysis that draws on observable behaviour across four of the six measured dimensions.

OPERATIONAL DEFINITION

The degree to which communication creates confidence, trust, authority and professional credibility in others.

DERIVATION

Clarity
Credibility
Composure
Influence



EXECUTIVE PRESENCE

Derived analysis — emergent narrative drawn from behaviour observed across the four feeder dimensions.

ANALYSIS EVALUATES

- Communication clarity
- Intellectual credibility
- Stability under scrutiny
- Influence without authority

WHAT IT IS NOT

- Not charisma
- Not dominance
- Not confidence alone
- Not personality
- Not a leadership prediction
- Not a hiring recommendation

RESEARCH FOUNDATION

Cross-cultural leadership research suggests executive presence is commonly associated with competence, integrity, communication effectiveness, emotional stability and influence, although cultures differ in how these qualities are expressed.

15 · ASSESSMENT METHODOLOGY

A five-stage pipeline. Evidence at every stage.

Each assessment is a structured conversational experience designed to elicit observable communication and behavioural evidence. Every score is traceable to specific evidence in the response record.

01

Structured Conversational Assessment

The candidate engages with a sequence of professional scenarios calibrated to elicit observable communication behaviour relevant to the six dimensions. Items are scenario-grade — they have no single correct answer.

02

Behavioural Observation

The response is treated as observable behaviour, not as a self-report. Indicators specified in the construct definitions are looked for, recorded, and time-stamped against the response transcript.

03

Evidence Extraction

For each indicator, supporting quotations and behavioural notes are extracted from the response. This evidence is retained at the case level and is reportable alongside any score.

04

Dimension Scoring

Per-dimension scores aggregate the evidence collected, calibrated against the published rubric. No single response can drive a score; the rubric requires evidence across multiple indicators within a dimension.

05

Human Interpretation

All scoring is reviewed by a human evaluator before any candidate-facing or recruiter-facing report is released. Edge cases, atypical responses, and contested indicators are escalated.

16 · SAMPLE OUTPUT · EVIDENCE EXTRACTION

Candidate response · evidence · observations.

An anonymised, representative example. The scenario asks the candidate to account for a missed product launch to their CEO.

SCENARIO · ACCOUNTING TO THE CEO

Your product launch slipped its committed date. The CEO has asked you to brief the executive team. Please respond in two minutes.

CANDIDATE RESPONSE · VERBATIM EXCERPT

*"We missed by eleven days. Three things drove the slip. One: we under-estimated the security review — that's on me, I owned the timeline. Two: the integration with billing surfaced an edge case we hadn't tested; engineering caught it, which is why we delayed rather than ship. Three: the customer-comms work needed more hand-holding than we'd budgeted."
"Here's where we are now: revised launch is the 14th, security review is signed, billing edge case is fixed and regression-tested. I'd like fifteen minutes after this to walk through the revised customer-comms plan with marketing."*

Integrity · Accountability

"that's on me, I owned the timeline" — explicit ownership before explanation.

Clarity · Structure

Three-cause enumeration with explicit signposting ("One... Two... Three...").

Credibility · Evidence-based reasoning

Specific cause attribution; concrete remediation status named for each cause.

Composure · Response under pressure

Calibrated affect, no defensiveness, ends with a forward-action request.

17 · SAMPLE OUTPUT · DEVELOPMENT GUIDANCE

Dimension observations and development recommendations.

Sample, continued. The dimension-level observations and the developmental recommendations that the platform would surface — for the candidate, and (in anonymised form) for the recruiter.

Clarity	Strong	Structured response under time pressure; explicit enumeration aids the listener.
Credibility	Established	Cause attribution is specific and remediation status named. Could add quantified risk where available.
Integrity	Strong	Accountability stated before explanation. Consistent with stated values.
Composure	Strong	Calibrated affect under accountability pressure. No defensiveness observed.
Influence	Developing	Forward-action request is well-framed; could include explicit stakeholder framing ('for the board / for customers').
Adaptability	Developing	Register appropriate to CEO audience; could vary explanation depth for mixed-seniority audiences.

DEVELOPMENT RECOMMENDATIONS · CANDIDATE-FACING

- Practice extending the forward-action request with explicit stakeholder framing — name who benefits from each next step.
- In mixed-seniority briefings, rehearse a two-pass structure: executive summary first, then operational detail for the operators in the room.
- Continue the three-cause enumeration pattern under pressure — it is a strong, reliable composure-and-clarity signal.
- No development action recommended for Integrity or Composure at this stage; both observed as durable strengths.

17B · COMMUNICATION DEVELOPMENT FRAMEWORK

Structured evidence + development guidance.

Silk is designed not only to identify behavioural evidence but to provide a structured pathway for communication development. Every observation should connect to a practical improvement action.

CLARITY

OBSERVED CHALLENGE

Unstructured responses

DEVELOPMENT

- Response architecture training
- Message sequencing
- Executive briefing practice

CREDIBILITY

OBSERVED CHALLENGE

Unsupported assertions

DEVELOPMENT

- Evidence-based communication
- Strategic reasoning frameworks
- Decision explanation practice

INTEGRITY

OBSERVED CHALLENGE

Limited ownership language

DEVELOPMENT

- Accountability communication
- Stakeholder transparency techniques
- Ethical decision framing

COMPOSURE

OBSERVED CHALLENGE

Reduced clarity under pressure

DEVELOPMENT

- Pressure simulations
- Executive challenge rehearsals
- Ambiguity-response training

INFLUENCE

OBSERVED CHALLENGE

Weak stakeholder alignment

DEVELOPMENT

- Persuasive communication
- Stakeholder mapping
- Alignment-building frameworks

ADAPTABILITY

OBSERVED CHALLENGE

Same communication style for all audiences

DEVELOPMENT

- Audience calibration
- Cultural intelligence
- Context-sensitive communication practice

STRUCTURED EVIDENCE + DEVELOPMENT GUIDANCE

Silk assessments are designed to support communication growth and professional development. Assessment results identify observable strengths and development opportunities rather than fixed traits or permanent capabilities.

18 · RELIABILITY FRAMEWORK

Targets, not claims.

Reliability work is in progress. The thresholds stated below represent what the validation programme is designed to evaluate against — not results currently held.

INTERNAL CONSISTENCY

PRELIMINARY

 ≥ 0.80 target · Cronbach's α

Pilot α observed in the 0.74–0.83 range across dimensions. Full-sample target ≥ 0.80 .

TEST-RETEST

IN PROGRESS

 ≥ 0.75 target · r (4-week)

Pilot subset $n = 24$ retested at 4 weeks. Data collection ongoing for the full sample.

INTER-RATER (HUMAN)

IN PROGRESS

 ≥ 0.80

target · ICC

Two-rater coding of a stratified subsample currently in field.

HUMAN-MODEL

IN PROGRESS

 ≥ 0.70 target · κ

Agreement between human coders and the scoring engine evaluated on the same subsample.

19 · VALIDITY FRAMEWORK

Content. Construct. Criterion.

Validation evidence is structured against the standard three-class model. Each class is labelled with its current status. Criterion validity is scoped to communication and behavioural outcomes — not general job performance.

Content validity

FOUNDATION

Items reviewed by subject-matter panel; dimension definitions anchored in the published literature cited in §03–§07.

Face validity

PRELIMINARY

Pilot respondents ($n \approx 30$) rated the assessment as 'recognisable' and 'professionally credible' in post-assessment feedback.

Convergent validity

IN PROGRESS

Currently being evaluated against an established communication-competence self-report measure and an observer-rated communication-effectiveness instrument. Additionally evaluated against CEFR-aligned proficiency descriptors at B2/C1/C2 level for language-performance convergence.

Discriminant validity

IN PROGRESS

Currently being evaluated against general personality (Big Five) and general cognitive ability to demonstrate non-redundancy with existing constructs. Additionally evaluated against raw CEFR language proficiency to demonstrate non-redundancy with language ability.

Criterion validity · communication outcomes

PLANNED

Planned: relationship between Silk dimension scores and supervisor-rated communication effectiveness (not promotion, not performance) at 6 and 12 months.

Incremental validity · communication outcomes

PLANNED

Planned: variance in supervisor-rated communication effectiveness explained above and beyond personality, tenure, and interview ratings. Additionally controlled for CEFR proficiency level.

20 · PILOT PROGRAMME

Preliminary evidence.

Initial pilot data is reported below. These are preliminary findings, not validation. The pilot was designed to evaluate scoring consistency, user experience, and observable signal — not to test any predictive claim.

n = 30

SENIOR PROFESSIONALS

22 min

MEDIAN TIME-ON-TASK

93%

COMPLETION RATE

0.74-0.83

OBSERVED A RANGE

Scoring consistency

PRELIMINARY

Internal consistency observed in the target range across five of six dimensions; one dimension (Adaptability) sits below target and the indicator set is being refined.

User experience

PRELIMINARY

Self-reported credibility median 4.6 / 5. Respondents consistently described the assessment as 'respectful' and 'recognisable'.

Behavioural signal

PRELIMINARY

Distinguishable behavioural patterns observed across the cohort, consistent with the framework's six-dimension structure.

Repeat administration

IN PROGRESS

Subset n = 24 retested at 4 weeks. Analysis under way.

21 · ETHICS & DATA HANDLING

Research-grade governance.

All studies operate under explicit informed consent, GDPR-compliant data handling, and a published bias-monitoring regimen.

Informed consent

Every respondent consents at study entry and may withdraw and request deletion at any time.

Data minimisation

Only assessment responses and stated demographic controls are collected. No biometric, video, or audio data.

Anonymisation

Research datasets are de-identified at source; tenant- and respondent-level identifiers stripped prior to analysis.

Bias monitoring

Score distributions monitored across demographic strata. Quarterly bias audit summary published.

Ethics oversight

Studies submitted to an external ethics review prior to data collection. Reviewer panel composition published.

Open methodology

Item construction protocol, scoring rubric, and analysis plans shared with research partners under NDA.

22 · LIMITATIONS

Stated openly.

A research instrument is only as useful as its limitations are honestly stated. Five are material at this stage.

Sample size

Pilot evidence is drawn from a small senior-professional sample ($n \approx 30$). Generalisation beyond this sample requires the full validation programme.

Behavioural substrate

Scenario responses are observed behavioural artefacts within the assessment environment. Generalisation to behaviour in the field is the validation programme's central empirical question.

Cohort calibration

Benchmarks are currently calibrated against the cohort the platform has assessed. As the assessed cohort grows, benchmarks will be re-calibrated; current benchmarks should be treated as preliminary.

English-language items

The current item pool is English-language only. Cross-cultural translation and adaptation is part of the planned future programme.

Language proficiency effects

Silk assesses observable communication behaviour and is influenced by — but is not a measure of — language proficiency. The current programme controls for stated CEFR level and monitors score distributions across B2/C1/C2 strata.

23 · EXPLICIT NON-CLAIMS

What Silk does not measure. What Silk does not do.

This page exists so that the scientific, legal and ethical boundaries of the platform are unambiguous. Each statement is absolute. None will be qualified or softened in product copy.

Silk does NOT diagnose personality

Silk does not assess personality types, traits, or psychological constructs. The framework is behavioural, not dispositional. The words 'personality', 'personality test', and 'personality type' do not appear in user-facing copy.

Silk does NOT diagnose psychological conditions

Silk is not a clinical instrument. It does not assess, diagnose, screen for, or rule out any psychological, psychiatric, or developmental condition.

Silk does NOT measure IQ or cognitive ability

Silk does not test intelligence, general cognitive ability, or any cognitive aptitude. It does not produce IQ-equivalent scores or correlates.

Silk does NOT predict future job performance

Silk does not forecast performance, success, tenure, promotion, or any career outcome. The platform is an evidence layer for human decision-making, not a predictive engine.

Silk does NOT make hiring decisions

Silk does not accept, reject, recommend, or rank candidates. No score and no analysis produced by the platform constitutes a hiring recommendation. Human judgement is essential.

Silk does NOT assess cultural fit

Silk does not evaluate 'cultural fit' for an organisation. It evaluates observable communication behaviour against a published rubric — neutral to the hiring organisation's culture.

24 · VALIDATION ROADMAP

Seven stages of scientific maturity.

Validation is treated as a programme, not a milestone. The current position is stated explicitly at each stage. Independent validation is the destination, not the starting point.

01	Foundation Research Seven bodies of established literature reviewed and mapped to the six-dimension framework. Construct definitions published.	COMPLETE
02	Expert Review Subject-matter panel reviewing item content, scoring rubric, and dimension definitions. First-round feedback incorporated.	IN PROGRESS
03	Pilot Testing Initial pilot (n ≈ 30) complete. Scoring consistency, user experience, and observable behavioural signal evaluated.	COMPLETE
04	Inter-rater Reliability Two-rater coding of a stratified subsample currently in field. ICC target ≥ 0.80.	IN PROGRESS
05	Construct Validation Convergent and discriminant validity studies against established communication and personality instruments under way. Includes CEFR-aligned language-proficiency descriptors to demonstrate non-redundancy with raw language ability.	IN PROGRESS
06	External Validation Independent validation by university partners. Pre-registered analysis plans. Conducted by researchers without commercial interest in the outcome. Includes applied-linguistics partners for CEFR-concurrent validation.	PLANNED
07	Longitudinal Validation 12–18-month follow-up studies against supervisor-rated communication effectiveness — not promotion outcomes, not general performance.	PLANNED

25 · RESEARCH PARTNERSHIP OPPORTUNITIES

Four open invitations.

Silk Clarity is actively seeking academic partners. Data and platform access provided under structured research agreements.

OPEN PARTNERSHIP

Construct validation

University partner

Run the independent convergent and discriminant validity studies using an independently-recruited sample. Co-authorship on resulting publication.

OPEN PARTNERSHIP

Communication outcomes study

Research group

Lead a 12-month follow-up cohort study against supervisor-rated communication effectiveness. Outcomes rated blind to baseline. First-author opportunity for lead investigator.

OPEN PARTNERSHIP

Cross-cultural adaptation

International partner

Translate, adapt, and re-validate the item pool in a non-English-language professional cohort.

OPEN PARTNERSHIP

CEFR-concurrent validation

Applied-linguistics partner

Concurrent validation of Silk dimension scores against CEFR-aligned communication-performance descriptors at B2/C1/C2 in a multilingual professional cohort.

26 · INVITATION

*Examine the framework.
Run the studies. Publish the results.*

Silk Clarity is a research-informed communication and behavioural assessment platform that provides structured evidence and development guidance. It is not, and does not claim to be, a system that predicts who will succeed. The first is defensible. The second is not — and is not the position the platform takes.

RESEARCH CORRESPONDENCE

hello@silkclarity.com

We will respond with: the full methodology document, the current pilot evidence pack, the analysis plans for the in-progress studies, and a draft research agreement.