SUPPLEMENTARY MATERIAL 2: PRACTICAL IMPLEMENTATION GUIDANCE FOR CONVERGE-DIVERGE-RECONVERGE SURVEY ARCHITECTURE

Purpose

This supplementary material provides practical guidance for researchers implementing the converge-diverge-reconverge survey framework in their own multi-role organisational contexts. The guidance synthesises lessons from developing and deploying the engineering workplace survey documented in the main manuscript.

A note on how to read this document. Guidance below is written as the standard a prospective implementation should aim at. It is not a description of the reference deployment, which departs from it in several respects. Where the reference deployment did not meet its own guidance, this is stated at the point the guidance appears, marked **Reference deployment.** *Those divergences are among the reasons the manuscript reports role-specific findings descriptively. Nothing here should be read as a claim about what the deployed survey achieved.*

AT A GLANCE

One page for planning. The sections that follow give the detail behind each item.

The architecture. Every respondent answers a common demographic block, branches to a role-specific module on a single routing variable recorded inside the instrument, then reconverges to a universal comparative block answered by all. Cross-role comparison uses the common and universal blocks only; role-specific items are not comparable across roles by construction.

Five decisions, taken before you build. These complement the five design decisions in Section 3.1 of the manuscript, adding platform choice because it governs what can later be verified.

1. Which constructs are universal and which are role-specific. Test each one: would this item be interpretable to someone in a different role?

2. Module count and option list, set separately. Four to eight modules is manageable; the option list should be as generous as your respondents need, and options may share a module.

3. Module depth, allocated by the analytical weight each branch must carry rather than by your familiarity with the domain.

4. Sequencing, from the three options below.

5. Platform, chosen against your analysis plan rather than your budget. Google Forms, as used in the reference deployment, recorded neither completion time nor randomised section order; other platforms offer these features to differing degrees, often depending on subscription level. Whatever you cannot record, you cannot later claim.

Sequencing: three options, not two.

A. Role modules first, universal block last. Universal items are answered with role experience fresh in mind, but respondents arrive at them having completed unequal amounts of preceding content.

B. Universal block first, role modules second. Removes the exposure differential, but asks for generalised judgements before anything concrete has been brought to mind.

C. Randomised block order within role. The only option that separates a sequencing effect from a role characteristic directly. Requires platform support.

Use C where your platform allows it. Otherwise administer any universal construct that overlaps topically with your branch content before branching, and state that overlap prediction before you analyse.

Five checks on the deployed data, before writing results.

Routing integrity, by cross-tabulating each role-specific block against the routing variable. Structured missingness, confirmed as fully explained by routing. Non-differentiation across exposure groups. Alpha by exposure group. Construct means across exposure groups, with roles compared at equal exposure.

Four traps.

Required fields manufacture completeness. Zero missingness on compulsory items is a property of the form, not a finding about your design.

Platforms that store only submitted responses make non-completion invisible. No attrition claim is available to you, and none should be made.

A gradient across exposure groups is not a sequencing effect until roles at equal exposure have been compared. One construct in the reference deployment showed the largest gradient of all and turned out to vary by role.

Describe the instrument you deployed, not the one you drafted. Getting the block order wrong inverts the question the design turns on and propagates through every section.

1. WHEN TO USE THIS APPROACH

Appropriate contexts

Use converge-diverge-reconverge architecture when:

- Your research question requires BOTH cross-role comparison AND role-specific depth
- Roles in your context experience shared phenomena differently
- You want to capture functional diversity without fragmenting your dataset
- You need to avoid presenting respondents with irrelevant questions
- Sample size permits meaningful subgroup analysis by role ($n \geq 30$ per role as general guidance)

- Your survey platform supports branching logic (most modern platforms do)

Inappropriate contexts

Do NOT use this approach when:

- Roles are too heterogeneous within categories to form meaningful shared modules
- Sample size is insufficient for role-stratified analysis
- Your research focus is purely on universal constructs with role as a simple control
- Resources do not permit complex survey design, testing and validation
- Your platform lacks branching capability
- Time constraints prevent proper testing of all pathways

Reference deployment. Branch sizes ranged from n = 14 to n = 142. Two branches (commercial, n = 14; planning, n = 22) fell below thirty. The $n \geq 30$ figure is offered as planning guidance, not as a threshold the reference deployment met, and small branches there are reported descriptively rather than inferentially.

2. KEY DESIGN DECISIONS

2.1 Determining universal vs. role-specific content

Universal modules should include:

- Constructs expected to exist across all roles (job satisfaction, wellbeing, organisational commitment)
- Demographic and work-context variables for cross-role comparison
- Organisational-level perceptions (culture, leadership, values)
- Outcome variables you want to compare across roles

Role-specific modules should include:

- Observable practices, behaviours or decisions unique to a function
- Pressure manifestations that differ substantively across roles
- Technical or functional knowledge only specific roles would possess
- Context-specific challenges not meaningful to other roles

Test question for classification: "Would this question be meaningful and interpretable to someone OUTSIDE this role?" If YES, consider for universal modules. If NO, it belongs in a role-specific module.

Additional test. For each universal construct, ask: "Does this share subject matter with what the role-specific modules ask about?" Where the answer is yes, that construct is where sequencing effects will concentrate if you place it after the branch. See §2.4.

2.2 Selecting role categories

1. **Align with organisational structure.** Use functional divisions recognised in your target context.
2. **Ensure within-role homogeneity.** People within a category should experience similar phenomena.
3. **Balance granularity.** Too broad loses functional depth; too narrow leaves insufficient sample sizes.
4. **Consider analytical needs.** Each role needs sufficient n for meaningful analysis (aim for $n \geq 30$).
5. **Distinguish the option list from the module count.** These are not the same number. A long option list helps respondents find themselves; a short module list keeps the design tractable. Several options can route to one module, and several can route to a minimal fallback.

Pilot testing recommendation: test role self-identification with 10–15 people from the target population. Ask which category best describes their primary role. Revise if people struggle to classify themselves.

Reference deployment. Q20 offered sixteen role options plus a free-text field, mapping to five role-specific modules, a single-item unbranched fallback and an unrouted path. The option list was deliberately longer than the module count: sixteen options were chosen to cover the great majority of occupations in the sampling frame, and the free-text field was retained so that a respondent whose role was absent would not conclude the survey was not for them and abandon it. Four respondents used the free-text field and received no role-specific module. The design accepted the loss of their role-specific data in exchange for not excluding them.

2.3 Balancing module length

The challenge: role-specific modules may require different numbers of items to capture their phenomena adequately. How do you balance depth against completion burden?

Guidance: use completion time rather than item count as your equality metric, and verify it against pilot timing data rather than estimates. Accept minor variation rather than forcing artificial equality.

This requires a platform that records completion time. Confirm before you build that yours does — several widely used free platforms, including Google Forms, do not. If timing cannot be captured, completion-time parity is a design intention you cannot verify, and it should be reported as such rather than as an achieved outcome.

Reference deployment. Pathway lengths were set by self-timed completion during instrument construction. The platform captured no timing data, so no completion-time parity was demonstrated and none is claimed. Item load ranged from 55 to 75 items across pathways, a difference of nineteen role-specific items between the longest and shortest routed pathway. That asymmetry followed from the study's development history rather than a considered allocation.

2.4 Sequencing universal and role-specific content

Three options, not two.

Option A: terminal reconvergence — demographics first, role-specific modules second, universal comparative block last. *(The reference deployment's approach.)*

- *Advantages:* universal items are answered with role-specific experience freshly in mind; creates a narrative flow from context to specific to general.
- *Disadvantages:* respondents reach the universal block having completed different quantities of preceding content. This is not merely a theoretical concern — see the finding below.

Option B: initial convergence — universal comparative block first, role-specific modules second, brief verification items at the end.

- *Advantages:* removes the exposure differential entirely; more conventional; easier to defend.
- *Disadvantages:* asks for generalised judgements before any concrete material has been brought to mind.

Option C: randomised block order within role — randomise whether each respondent receives the role-specific or the universal block first, within each role group.

- *Advantages:* preserves grounding for half the sample while producing a within-role comparison that separates sequencing effects from role characteristics directly. Neither Option A nor Option B can do this.
- *Disadvantages:* requires platform support for randomised section order, which not every platform or subscription level offers.

What the reference deployment found. Option A does carry a cost, but this design cannot fully measure it. Two of the four universal constructs varied significantly with the quantity of role-specific content preceding them. The first, workplace pressure, shares subject matter with every role-specific module: it was flat across the two lowest exposure levels and higher at the two highest. Whether roles at equal exposure differed depends on the effect-size measure. With η^2H, role differences were small at one and five preceding items but close to the exposure gradient at seven; with ε^2, they were of similar size to the gradient at one and seven preceding items. Sequencing therefore remains a possible explanation for pressure, but it is not established. The second, organisational culture, showed the larger gradient, but roles at equal exposure differed more than the exposure groups did on both measures, which points to role rather than position in the instrument. The remaining two constructs, addressing wellbeing and team dynamics, showed no gradient surviving correction for multiple comparisons. Contextual priming is therefore not an unqualified advantage of Option A. Any effect is likely to concentrate where module content and universal constructs overlap in subject matter, and it cannot be read from the exposure gradients alone: without the equal-exposure comparison, culture would have been misread as a second sequencing effect.

Recommendation. Use Option C where your platform supports it. Otherwise: where a universal construct overlaps substantially in subject matter with your branch content, administer it before branching (Option B for that construct); where contextual grounding is the priority and topical overlap is limited, Option A is justified. Document whichever you choose, and state the overlap prediction before you analyse, not after.

2.5 Handling roles without dedicated modules

Not every role needs a role-specific module. Reasons to skip:

1. **Heterogeneity** — the category is too internally diverse
2. **Research focus** — your primary questions do not centre on this role
3. **Sample size** — expected n is too small for meaningful role-specific analysis
4. **Resource constraints** — limited time for module development

When skipping role-specific modules, you have two workable arrangements:

- **Minimal fallback.** Route these respondents to one or two short items of general relevance (for example, supervisory scope), then to the reconvergence block. This keeps their exposure non-zero and gives you at least one comparable variable.
- **Direct to reconvergence.** Route them straight to the universal block with no role-specific content at all.

In either case: explain the decision in your methods, note in limitations that these roles lack functional depth measurement, and report how many respondents were affected.

Reference deployment. Nine of the sixteen role options routed to a single-item fallback (Q65, supervisory scope) reaching n = 142 — the largest single pathway in the instrument. Four free-text respondents went directly to reconvergence with no role-specific content. Both groups are reported separately throughout, and the fallback group is not described as a role-specific module.

3. IMPLEMENTATION STEPS

Step 1: Define your research questions

- **Which universal constructs overlap topically with your planned role content?** Answer this now; it determines your sequencing decision in Step 3.

Step 2: Select role categories

- Finalise your **module** count — four to eight is manageable; beyond eight becomes unwieldy
- Set your **option** list separately, and make it as generous as your respondents need. Options may share modules. Include a free-text option so that respondents whose role is absent are not driven to abandon the survey, and decide in advance how those respondents will be routed and reported

Step 3: Design branching architecture

- Determine sequencing (see §2.4 — three options)
- Decide the routing destination for free-text and unmatched responses

Step 4: Build survey in platform

Platform considerations: capabilities such as recording completion time, storing partial responses and randomising section order vary between platforms and between subscription levels of the same platform. Confirm that the platform and plan you will use support everything your analysis plan requires before you build.

Choose the platform against your analysis plan, not only your budget. If you intend to report completion-time parity, attrition by pathway, or a randomised sequencing control, confirm that your platform and subscription level support it; otherwise, you will discover the gap after fielding rather than before.

Step 5: Pilot test

Test all pathways with actual users from the target population.

1. Recruit enough testers to cover every pathway — two to three per role category is the working minimum, and more is better where role categories are numerous
2. Time each pathway *(if your platform permits — see Step 4)*

Reference deployment. Fifty-two testers took part in the pilot deployment, well above the working minimum.

Step 6: Revise and finalise

- **Delete all pilot test data before data collection begins, and do not report analyses of it.** If pilot data has been deleted, no table, figure or thematic analysis drawn from it can appear in the paper. This is easy to breach inadvertently when a pilot-derived figure has been carried forward from an early draft; check for it explicitly before submission

Step 7: Ethics approval

On consent routing. If your consent items instruct respondents to do something — return to an information page, exit if ineligible — confirm whether that instruction is *enforced by platform logic* or merely *presented as text*. These are different claims and must be reported differently. In the reference deployment, skip logic existed on the eligibility item and the role-routing item only; the information-sheet item advised returning to the welcome page but no routing enforced it. Check your own build rather than assuming.

Step 8: Deploy and monitor

- Launch and verify routing on the first live responses
- Check for differential attrition by pathway *(platform permitting — see §5)*

4. COMMON PITFALLS AND HOW TO AVOID THEM

Pitfall 1: role modules vary too much in length

Problem: commissioning gets 20 questions, operations gets 5.

Solution: allocate depth according to the analytical weight each branch will carry, not your familiarity with the domain. Use completion time as the metric where your platform records it. Accept minor variation rather than forcing artificial equality. Where depth must vary, state and justify the variation rather than leaving readers to infer it from module tables.

Pitfall 2: role categories overlap or are unclear

Solution: provide clear definitions; include examples in the role selection question; pilot test self-identification; add a "primary role" qualifier; provide an "Other" free-text option for edge cases, and decide in advance how those respondents are routed.

Pitfall 3: universal constructs scattered through the survey

Problem: mixing universal items into role-specific sections, making cross-role comparison difficult.

Solution: keep the converge–reconverge structure clean. All universal comparative items belong in a single block, whether that block sits before or after the branch. No universal constructs embedded inside role-specific modules.

A caution on describing your own design. Be precise about where the universal block actually sits. Describing a terminal reconvergence design as front-loading its universal constructs, or the reverse, inverts the sequencing question the design turns on and will propagate through every section of the paper. Check the deployed instrument, not the draft.

Pitfall 4: forgetting to test ALL pathways

Solution: test every role pathway before launch; use a checklist; check the data export for each pathway; verify routing in preview AND live modes.

Pitfall 5: using role as both routing variable and measured construct

Solution: keep role selection simple. Do not ask role-specific questions about role identity or role perceptions. Role affects the pathway; it is not an outcome.

Pitfall 6: not documenting sequencing rationale

Solution: document all design decisions explicitly in the methods. Explain the sequencing choice. State your topical-overlap prediction before analysing, so it precedes rather than follows the result.

Pitfall 7: assuming complete data means willing respondents

Problem: the universal block returns no missing values, and this is reported as evidence that the design did not burden respondents.

Solution: check whether your universal items were set as **required fields**, and whether your platform stores partial responses. If items are required and only submitted responses are stored, complete data is guaranteed by configuration before any respondent arrives. It is not evidence about your design. Report it as a configuration outcome. See §5.

5. ANALYSIS CONSIDERATIONS

Data structure

- **Universal variables:** present for all respondents who submitted
- **Role-specific variables:** present only for respondents routed to that module (structured missingness by design)

Analysis implications: cross-role comparison uses universal variables only; role-specific analysis uses a filtered subset; role-specific responses cannot be compared across roles, since respondents answered different questions.

What structured missingness does and does not give you

The genuine advantage is that missingness in the role-specific block is deterministic, fully explained by an observed routing variable, and requires no imputation. This distinguishes the design from split-questionnaire approaches, where randomised allocation is what licenses imputation. That property holds whatever your platform settings.

What it does not give you is an attrition finding. Two cautions:

1. **Required fields manufacture completeness.** If universal items are compulsory, zero missingness on them follows mechanically.
2. **Many platforms do not record non-completion.** Google Forms, among others, stores only submitted responses. A respondent who abandons the instrument generates no record at all. Your achieved sample therefore consists of completers by construction, and no attrition analysis is possible — including the pathway-level dropout comparison recommended in §6 below.

If differential attrition by pathway matters to your research question, select a platform that captures partial responses **before you build**. This cannot be recovered afterwards.

Sample size requirements

Where n is insufficient in some roles: consider combining similar roles on content grounds; report descriptive statistics only for small-n roles; focus comparative analysis on roles with adequate n. Note that subgroup reliability estimates are unstable below roughly n = 30, and multi-group confirmatory factor analysis is not feasible at those sizes — which means measurement invariance across roles will be assumed rather than demonstrated. State this explicitly if it applies.

Choice of tests

Universal-block items are typically ordinal and exposure groups are typically unequal in size, so non-parametric procedures are usually appropriate: Kruskal-Wallis for group comparisons, with eta-squared based on H (η^2H) or epsilon-squared ($\varepsilon^2 = H/(n - 1)$) as effect size, stating which is used, since the two give different values, and Spearman's rho for monotonic association across ordered exposure levels. Where you conduct several construct-level tests, apply and state a correction for multiple comparisons.

A caution on comparing effect sizes. If you compare role variation at constant exposure against a between-exposure gradient, note that these come from tests of different designs with different group counts and sample sizes. Both η^2H and ε^2 are sensitive to these, in different ways. Treat such comparisons as qualitative patterns rather than numerical rankings, and do not present them as a decomposition of variance.

6. REPORTING YOUR SURVEY

Methods section should include:

1. Architecture description: converge-diverge-reconverge with branching logic
2. Routing logic: how respondents are assigned to modules, and how unmatched responses are handled
3. Module structure: items per module, content focus
4. Sequencing rationale, with the topical-overlap prediction stated
5. Completion burden: how pathway lengths were set, and **whether timing data were captured**
6. Pilot testing: how the instrument was verified before deployment
7. Platform capabilities and limitations relevant to your analysis plan

Results section should include:

1. Response distribution by role and by pathway
2. Routing audit against the deployed sample (see below)
3. Completion times by pathway, **if your platform recorded them** — if not, say so rather than omitting the question
4. Attrition analysis, **if your platform records partial responses** — if not, state that non-completion is unobservable and make no completion claim

Limitations section should address:

1. Sequencing effects, where universal items follow role-specific content
2. Role self-selection and potential misclassification
3. Generalisability and sample representativeness per role
4. Module depth trade-offs
5. Measurement invariance, if subgroup sizes preclude testing it
6. Any claim your platform configuration cannot support

Post-deployment verification

Pre-deployment testing establishes that routing logic executes. It cannot establish what differential exposure did to the responses that logic produced. Run these five checks on the deployed data before writing your results:

Check	Procedure	Interpretation
Routing integrity	Cross-tabulate each role-specific block against the routing variable	Any response outside an assigned pathway indicates a logic fault
Structured missingness	Confirm missingness is fully explained by routing	Unexplained missingness suggests burden or comprehension problems
Response behaviour	Compare within-respondent variation and non-differentiation across exposure groups	Narrowing with rising non-differentiation suggests burden-related satisficing
Reliability	Compute alpha by exposure group	Systematic degradation suggests exposure-related measurement instability
Sequencing	Compare construct means across exposure groups; use roles at equal exposure to separate exposure from role	Gradients surviving the equal-exposure control support a sequencing interpretation

7. ADAPTATION TO OTHER CONTEXTS

The framework transfers to multi-role organisational settings where roles are functionally distinct but interdependent, shared organisational conditions manifest differently across roles, and the research aims at both system-level phenomena and role-specific mechanisms. Healthcare, education, emergency services and logistics all have this shape. Role-adapted instruments with common cores have been reported in health settings; transferability to the others is proposed rather than demonstrated.

The routing characteristic must be recordable within the instrument rather than inferred from sampling frame membership, since it is the recorded answer that drives both routing and analysis.

CONCLUSION

The converge-diverge-reconverge framework offers a principled approach to multi-role survey design that balances depth against comparability. Success depends on role categorisation, module design, sequencing chosen with topical overlap in mind, and thorough testing of every pathway.

Two lessons from the reference deployment are worth carrying forward. The first is that comparability is conditional: administering identical items to every respondent does not by itself make their answers comparable, because what precedes those items differs by branch. The second is that a branched instrument needs verification after fielding, not only testing before it — and that some of the checks worth running depend on platform capabilities that must be chosen before the instrument is built.