# What Likert Scales Cannot See: Informal Governance Mechanisms and the Case for Open-Text Items in Workplace Survey Research

**Jennifer R. Ayres[1a], Ian May[b], Rosmina Bustami [2a], Jethro H Adam[a], Sithara H.P.W. Gamage [3c]**

[a]UNIMAS Water Centre (UWC) Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan, 94300, Malaysia; [b]Cranfield University, College Rd, Cranfield, Wharley End, Bedford, United Kingdom, MK43 0AL; [c]Adelaide University, Mawson Lakes Blvd, SA 5095, Australia.

**ORCID IDs:** [1] 0000-0002-4538-6512 | [2] 0000-0002-8438-8932 | [3] 0000-0001-9209-9111

## Abstract

Structured survey instruments measure what researchers already know to ask about. Likert scales and forced-choice items capture anticipated phenomena but cannot directly name those whose sensitivity makes structured measurement ethically compromised—constraints confirmed in this study's ethics design. Open-text items preserve respondent control over whether and how to name their experience.

Drawing on a cross-sectional survey of 335 engineering professionals across 22 countries, this paper uses thematic analysis of 180 open-text responses to identify four categories that structured items are poorly positioned to recover: informal governance mechanisms (unrecorded hours, unofficial channels, hierarchy override); causal language (how pressure operates, not merely that it exists); emotional texture (coexistence of engagement and harm); and respondent-defined significance (what participants judged the study needed to know). Of 187 respondents providing open-text responses (55.8%), lengths ranged from 4 to 229 words (median 14, mean 18.6, SD 20.4).

Structured and open-text items measure genuinely different things: the former establish prevalence; the latter name mechanisms that scales cannot capture. We propose a mixed-methods survey design as a methodological necessity rather than a design preference wherever informal processes operate alongside power asymmetries.

**Keywords:** open-text survey items; informal governance; workplace pressure; mixed methods; survey design; engineering workforce

## 1. Introduction

Structured survey instruments measure what researchers already know to ask about — and sometimes, what they unconsciously expect to find. The dominance of Likert-scale and forced-choice formats in workplace research reflects an assumption so embedded it rarely surfaces for examination: that the phenomena worth measuring are knowable in advance, nameable by the researcher, and confirmable by the respondent. The instrument, within this frame, is complete. What it does not ask about does not enter the data.

But the choice of what to ask about is never neutral. Structured items are written by researchers who bring existing frameworks, disciplinary assumptions, and prior knowledge to instrument design — often without recognising how those assumptions shape which phenomena are made visible and which remain outside the response space entirely (Schwarz, 1999). In sensitive workplace research, this limitation extends beyond omission. There are phenomena that cannot be directly named in a closed-ended item without the act of naming them compromising the measurement: introducing prompting, creating distress, or producing researcher-defined categories where respondent-defined accounts are what the research requires.

This argument does not preclude alternative structured approaches such as vignettes or indirect questioning but suggests that direct naming of certain phenomena becomes methodologically compromised under conditions of sensitivity and power asymmetry — where the act of naming carries professional or personal risk for the respondent.

This is not a theoretical claim. In designing the instrument used in this study — a global survey of pressures in the workplace across engineering roles — structured items were deliberately avoided for known informal pressure mechanisms, as direct naming would have prompted, introduced measurement bias, and risked participant distress. That constraint was confirmed in the study's ethics design. The open-text item that followed the structured modules was the methodologically preferable and ethically defensible route to that category of data (Dunwoodie et al., 2023).

The case for open text as a legitimate qualitative research tool is well established (Braun et al., 2021). What remains underexamined is the extent to which structured items are ill-placed to recover certain categories of experience in sensitive workplace research — and what it means for research design when those categories include phenomena that are consequential, prevalent, and systematically underrepresented. The normalisation of deviance literature has long observed that the most operationally significant conditions in high-pressure environments are precisely those that resist formal articulation (Vaughan, 2016). Survey design in workplace research has not adequately reckoned with that finding.

This paper makes three contributions. First, it identifies four categories of workplace experience for which open-text items are better positioned to elicit than structured items, without introducing respondent-defined distortion, as demonstrated through a thematic analysis of 180 responses from 335 engineering professionals across 22 countries (Ayres et al., 2026a). Second, it proposes the practitioner-researcher position as a methodological resource in abductive design rather than a source of bias. Third, it argues that structured and open-text items capture genuinely different things — and should therefore be treated as jointly necessary rather than combined optionally in research where informal workplace processes are substantively relevant.

## 2. The Structural Limits of Closed-Ended Survey Instruments

### 2.1 What structured items measure, and what they are less well placed to capture

Structured survey items are instruments of confirmation. They test whether anticipated phenomena exist, how intensely they are experienced, and how frequently they occur — all within a response space pre-specified by the researcher. For hypothesised constructs with established theoretical grounding, this is by design: reliable, comparable, and psychometrically sound (Jebb et al., 2021; Tourangeau et al., 2000). The limitation arises at a different point: not from what structured items do badly, but from what they cannot do without distorting the measurement — specifically, surface phenomena the researcher did not anticipate, or could not name without substituting researcher-defined categories for the respondent-generated accounts the research requires.

The response space of a closed-ended item is bounded by the researcher's prior knowledge — what they knew, what their disciplinary framework recognised, what their theoretical commitments made visible. The mechanism through which this operates is rarely deliberate. Schwarz (1999) demonstrated that the framing of survey questions — the concepts named, the response options offered, the order in which items appear — systematically shapes the answers those questions receive. Researchers designing workplace pressure instruments draw on existing theoretical constructs: job demands, role conflict, and organisational support. Items are written to confirm, extend, or challenge those constructs. What they are not written to do is surface governance mechanisms the researcher has professional reasons not to name, or emotional conditions that resist the categories available in the disciplinary literature. The instrument's conceptual boundaries are the researcher's conceptual boundaries — and in sensitive workplace research, those boundaries carry ethical as well as epistemological consequences. A better-designed Likert scale, a longer instrument, a broader item bank can extend the range of anticipated constructs, but none resolve this core constraint: phenomena outside the researcher's prior frame tend not to be measured at all, not measured imprecisely. This is more structural than merely technical.

This has force in workplace research. The normalisation of deviance literature consistently finds that the most consequential risks in high-pressure environments are transmitted through informal channels, absorbed into role expectations, and rendered invisible by the organisational structures surrounding them (Vaughan, 2016). Standard instruments measure pressure: its presence, intensity, and frequency. They are less well placed to recover the mechanisms through which pressure is delivered, the informal governance that sustains it, or the organisational silence that prevents its naming (Detert and Edmondson, 2011). These are not gaps that additional items can close — they reflect what closed-ended formats, in this type of research, are not designed to do.

Star and Strauss (1999) made this argument for formal documentation: informal practices shape outcomes but remain systematically unrecorded, not because they are deliberately hidden, but because no categories exist for recording them. Survey instruments face the same problem. Pressure operating through un-minuted meetings, verbal instructions, and unofficial channels tends not to appear in structured data — not because respondents conceal it, but because no item exists to capture it. The instrument and the informal practice are incommensurable.

There is a further dimension of survey methodology that is underappreciated: the respondent's own relationship to the phenomenon. Structured items assume the respondent can recognise, name, and take a position on the construct being measured. But informal workplace harm frequently operates below the threshold of formal recognition —

experienced as pressure or constraint without belonging to any category the respondent has yet named. Hudson and Weinberg (2025) note that psychometric stress risk instruments tend to underrepresent conditions workers cannot or will not name within researcher-defined categories. What the instrument cannot anticipate, the respondent may still choose to name — but only if the instrument creates space for that naming.

### 2.2 The ethics constraint as evidence of structural limitation

Survey research ethics require that instruments not cause harm to participants. In designing the survey used in this study, this requirement led to a specific methodological decision: known informal pressure mechanisms could not be included as structured items. The ethics design confirmed this. The argument of this section is that this constraint is not merely procedural — it is substantive evidence that these phenomena sit in a category that structured formats, under these conditions, are ill-placed to address.

The argument moves through three steps. First, certain informal workplace mechanisms are ethically inadvisable to name directly as structured items: doing so introduces prompting, identification risk, and the potential for distress. Second, this inadvisability reflects a methodological problem — direct structured measurement cannot, without introducing these risks, produce valid respondent-defined accounts of the phenomena. Third, these constraints together make open-text items the methodologically preferable route. This is not a claim about epistemological impossibility in principle — it is a claim about what direct structured measurement becomes when naming the phenomenon is itself the problem, and when alternative approaches, such as vignettes or indirect items, would face the same ethical constraint in this research context.

Consider what a Likert item about informal hierarchy override does. It introduces a concept to every respondent. It frames experience in the researcher's language. It demands a position. For a respondent who has lived with this mechanism but never named it, the item may not capture their experience — it may construct one. For a respondent currently managing it through professional silence, the item demands acknowledgement that carries real risk, even in an anonymous survey. Research on sensitive survey reporting confirms that respondents routinely edit their disclosures to avoid professional or personal costs, and that this editing reflects the extent of their potential losses (Tourangeau and Yan, 2007).

In each case, naming does something to the respondent before they have answered. That constitutes prompting — and in research involving vulnerable populations, precarious employment, or hierarchical professional settings, it is precisely what ethics review exists to prevent (Dekker, 2011; Israel, 2015). The open-text item is the instrument best placed to preserve the respondent's epistemic authority — their right to decide whether to name their experience, in what language, and to what degree (Braun et al., 2021). Bryman (2006) argues that the complementarity of quantitative and qualitative methods is well established. What is less often argued is the case where that complementarity is not a design preference but a methodological requirement — where the nature of the phenomena makes both methods jointly necessary rather than individually sufficient.

## 3. Study Context and Instrument Design

The data derive from a cross-sectional survey of engineering professionals conducted in January–February 2026 (Ayres et al., 2026b). The survey examined pressures in the workplace across engineering roles and disciplines, combining structured measurement of

anticipated constructs with open-text responses to unanticipated phenomena. A total of 335 participants completed the survey across 22 countries and six continents, spanning commissioning, operations, planning, safety, project management, and related disciplines.

The instrument comprised universal modules administered to all respondents and role-specific modules delivered through skip-logic branching (Dillman et al., 2014). Universal structured modules covered anticipated pressure constructs — schedule compression, resource shortage, workload excess, peer pressure dynamics, safety culture, and organisational support — alongside demographic and employment variables. These items established prevalence, tested hypothesised relationships, and enabled cross-role comparison. They were not designed to surface mechanisms, causality, or informal governance — nor were they well placed to do so under the ethical and methodological conditions of this research.

The final universal item was an open-text question: "Please describe any experiences of workplace pressure you feel are important for this study." Several deliberate design decisions shaped this item. Placing it last in the universal module meant respondents had completed all structured items before encountering it — reducing the risk that open-text framing would contaminate scale responses, while ensuring that respondents who found the structured items inadequate had a space to say so. The wording was chosen specifically to hand definitional authority to the respondent: not 'describe any additional pressures' (which implies the structured items were primary) and not 'describe informal pressures' (which would have introduced exactly the category prompting the ethics constraint was designed to avoid). Respondents were explicitly invited to describe positive experiences, neutral observations, or no additional experiences at all. The item architecture was designed to be genuinely open rather than implicitly directive—and the presence of 25 positive and protective responses (PPF, 13.9% of all coded instances) confirms that this openness was genuine rather than nominal.

A five-theme, 23-code framework was developed through abductive thematic analysis of 180 codeable responses, with inter-rater reliability $\kappa = 0.836$ (Landis and Koch, 1977). Full coding details and substantive findings are reported in Ayres et al. (2026a). This paper uses the same dataset to develop a methodological argument: the engineering context is the empirical ground; the instrument design and what it enabled are the contribution. Figure 1 illustrates the sample and coding flow.

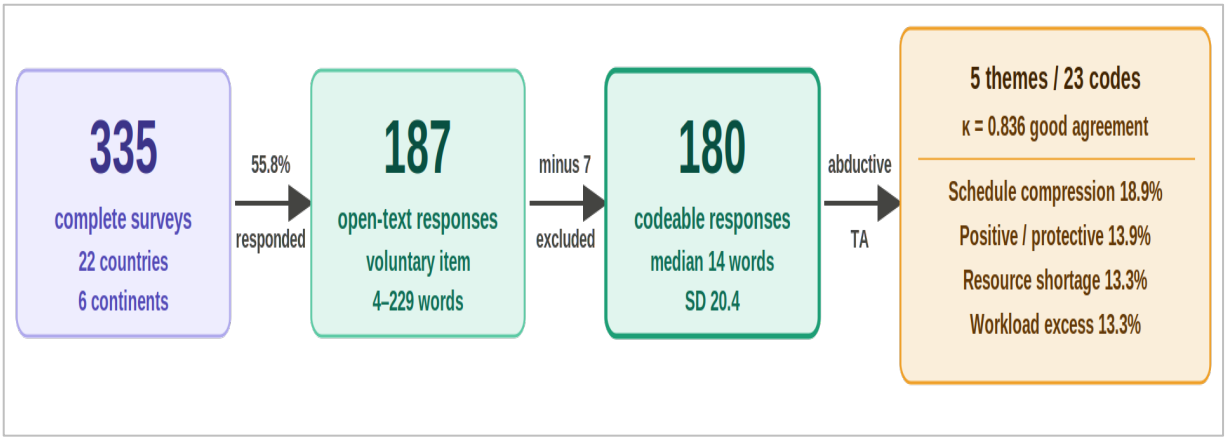


***Figure 1: From 335 complete surveys to five-theme framework: the 55.8% voluntary open-text response rate and wide word count range (4–229 words) suggest engagement extending well beyond the most aggrieved respondents.***

## 4. Four Categories of Data That Open-Text Items Are Better Placed to Recover

What do structured items return? Numbers, distributions, and correlations. What does open-text return? Mechanisms, causal accounts, and respondent-defined descriptions that scale scores cannot represent. Figure 2 maps the conceptual pathway — what each instrument contributes, and why direct naming of certain phenomena creates measurement problems in sensitive research contexts.

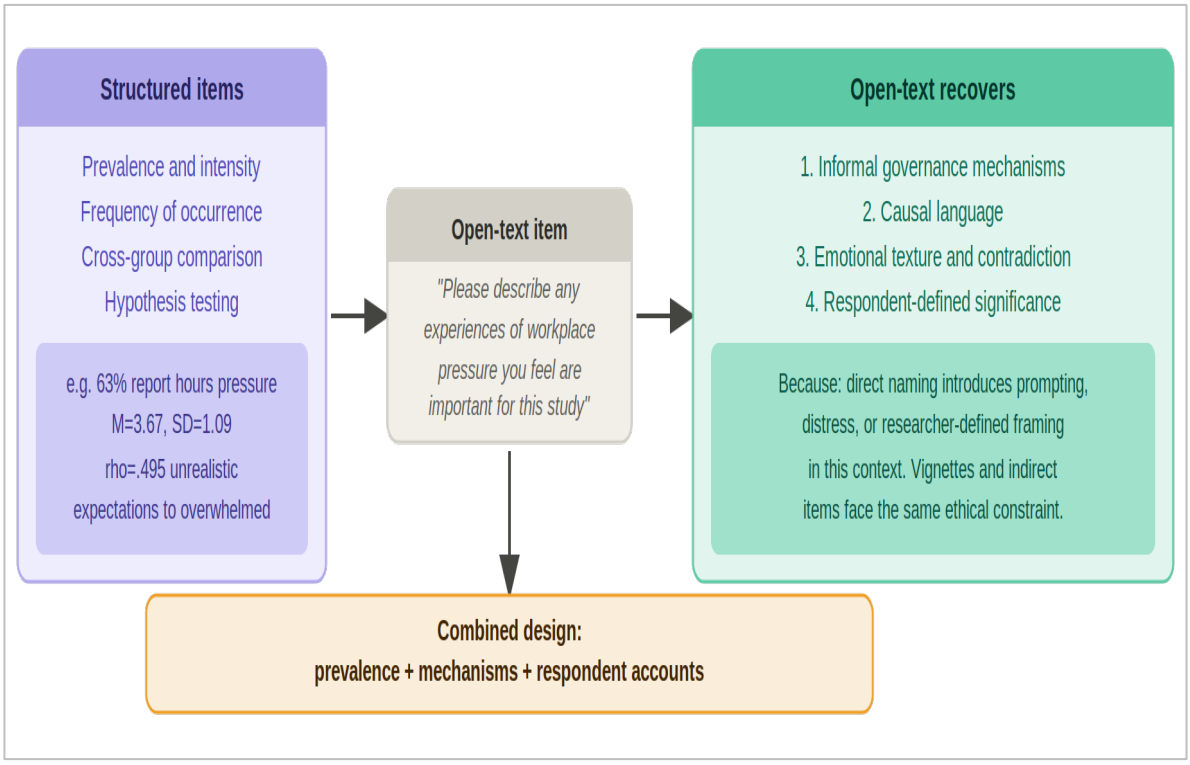


***Figure 2: What each instrument contributes and why.***

Figure 2 illustrates that the structured items recover prevalence, intensity, and hypothesised relationships. Open-text recovers informal governance mechanisms, causal language, emotional contradiction, and respondent-defined significance — categories that, in this research context, structured naming cannot reach without methodological compromise.

The four categories that follow illustrate this distinction empirically. All open-text responses are presented as respondent accounts — interpretations and descriptions of experience, not documentary evidence of objective conditions. Illustrative responses do not appear in the parallel findings paper (Ayres et al., 2026a).

### 4.1 Informal governance mechanisms

The most methodologically significant finding is that informal governance operates through channels that leave little trace in structured data and cannot be directly anticipated without introducing respondent-defined distortion.

One respondent named an unrecorded working practice in ten words:

> *"We get pressured to work longer hours than is recorded."*

The structured data established that 63% report hours pressure (M=3.67, SD=1.09). This response names what the scale cannot: a systematic gap between actual and recorded hours—

a governance concern invisible to an instrument that measures whether pressure is experienced but not whether its consequences are concealed.

A second response describes informal governance in seven words:

> *"To act and then seek approval after."*

Hierarchy override, described as a normalised procedure: instructions acted upon before authorisation; approval retrospectively secured. Structured items measuring formal compliance would not surface this inversion. The respondent has named a parallel governance pattern in which the formal sequence is routinely reversed — leaving no trace in documents subject to formal audit (Feldman and Pentland, 2003).

Both responses describe what Vaughan (2016) identifies as normalisation of deviance at the governance level: practices that violate formal procedure become standard operating procedure through repetition, institutional tolerance, and the absence of correction. Unrecorded hours become expected. Acting before authorisation becomes efficient. Neither appears in the formal record. Neither would appear in any structured survey item, because giving either a formal name in a structured item would constitute precisely the kind of prompting that makes direct structured measurement ethically and methodologically untenable in this context. The open-text item did not ask about unrecorded hours or hierarchy override. It asked respondents to describe what they felt was important. That two respondents independently named these mechanisms — across different roles, countries, and employment contexts — is the methodological finding. The instrument created the conditions; the respondents supplied the evidence.

## 4.2 Causal language

The correlation between unrealistic expectations and being overwhelmed was $\rho=.495$, $p<.01$. A number that confirms a relationship. A number that cannot be named by the mechanism producing it.

One respondent named the mechanism directly:

> *"Deadline never changes no matter how far behind schedule we are — this leads to shortcuts and extremely long hours."*

Not a report of schedule pressure — a causal account. A fixed-deadline regime absorbs slippage through informal means, such as shortcuts and unrecorded hours. The correlation would appear whether pressure was transmitted through legitimate workload management or through informal pressure to conceal schedule failure. The open-text account distinguishes between them in terms that no scale position could represent.

A second response locates pressure at the level of contract structure:

> *"In my previous roles I have often been pressured by senior management due to the nature of DB contracts being run by civil engineers with zero knowledge of complex biological processes. Assumptions are made at the start of a project that a complex biological system will work as soon as the construction is finished. And the programme has been written that way and is totally undeliverable."*

The structured data would record high management pressure. It would not record that the pressure is structurally inevitable given how the contract was framed. This reflects Hollnagel's (2018) gap between work as designed and work as experienced: structured instruments tend to measure the former; open-text accounts tend to reveal the latter.

### 4.3 Emotional texture and contradiction

Likert scales assume one-dimensionality. The response space has no position for contradiction — for the coexistence of states, instruments are typically treated as mutually exclusive.

The commissioning subsample reported the highest job satisfaction of any role group — 93.8% satisfied or very satisfied. The same subsample accounted for the most frequently coded qualitative pressure theme: schedule compression, at 18.9% of all coded instances. High satisfaction. High pressure. Both real. The scale returns one score per construct. The open-text returns the relationship between them (Bolton, 2005).

One respondent captured this coexistence in a single sentence:

> *"Pressure is from management that have no clue what commissioning does — my team are great."*

Two scales would return two numbers. The open text returns the relationship: team as a buffer against management failure, with protective experience and structural difficulty coexisting simultaneously.

A second response names the episodic, conditional nature of pressure:

> *"I love my job but feel under pressure at times especially when big projects are starting/finishing."*

High satisfaction. Moderate pressure. Neither score captures the conditional relationship: satisfaction sustained, pressure episodic, spiking at project transitions. The scale returns a static average where the respondent's experience is dynamic.

A third response describes a career trajectory:

> *"My time in the defence industry was always driven by pressure and unrealistic deadlines even when the overall project had slipped by 12 months. Often this made me ill and hating going to work. These days whether age or wisdom I don't particularly care about other people's deadlines or agendas. I make sure that I am performing, given the tools and information at my disposal."*

No structured instrument readily captures temporal change, adaptive coping, or the relationship between seniority and harm tolerance. The coexistence of structural pressure and professional commitment is explained here—not in a scale score, but in the narrative arc of a working life (Bakker and Demerouti, 2007).

### 4.4 Respondent-defined significance

For the three preceding categories, structured data provided a point of comparison. For this fourth, no comparison is available. No structured item addressed what respondents named. That absence is itself the evidence.

One respondent identified informal social performance as a pressure mechanism:

> *"Pressure to socialize or network to be more visible and be valued over others."*

No Likert item addressed this. The survey could not have asked about networking as competitive pressure without naming that category, which would have constituted prompting on a topic concerning the informal allocation of professional recognition. This respondent named it unprompted, in their own language, because the open-text item preserved their right to define what the study needed to know (Fricker, 2007; Nielsen et al., 2020).

A second response named the structural condition that constrains disclosure:

> *"Am on day hire so easy to get rid of when work dries up — unions don't help much either."*

No Likert item addressed this either. Two structural constraints on voice are named simultaneously: precarious employment, where raising concerns risks non-renewal; union protection offering a limited remedy. The relationship between harm and the conditions that silence it cannot surface in an instrument that does not preserve the respondent's epistemic authority.

The response investment data support this interpretation. Of 335 participants, 187 (55.8%) provided open-text responses voluntarily. Respondents who had completed all structured items chose to write more. That behaviour is evidence: the structured items had not captured what they wanted the study to know. Figure 3 shows the distribution of response lengths. Most responses were brief — median 14 words — but the long tail (8 responses of 51–100 words; one of 229 words) reflects respondents who chose to write at length after completing all structured items. That investment is itself evidence that the structured items had not captured what they wanted the study to know.

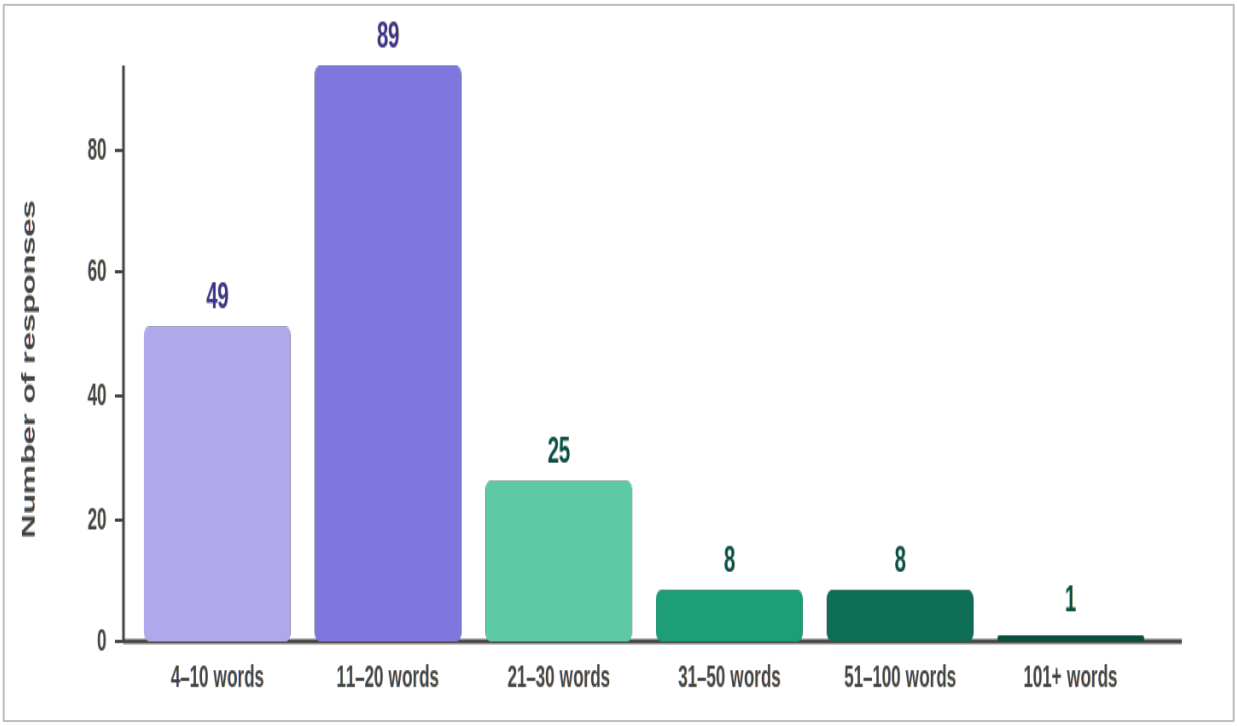


***Figure 3: Distribution of open text answer length***

**Table 1. Four categories of data for which open-text items are better placed to recover.**

| Category | Structured items returned | Open-text revealed | Illustrative response |
|---|---|---|---|
| **1. Informal governance mechanisms** | 63% report hours pressure (M=3.67, SD=1.09). Scale returns a single score — cannot distinguish recorded from unrecorded hours. | Unrecorded hours, unofficial instruction channels, hierarchy override as normalised operating procedure. | *"We get pressured to work longer hours than is recorded"* |
| **2. Causal language** | ρ=.495 (unrealistic expectations → overwhelmed). Confirms association. Cannot explain the mechanism producing it. | Fixed deadline regimes absorbing slippage through shortcuts and unrecorded hours. Contractual knowledge asymmetry as structural cause. | *"Deadline never changes no matter how far behind schedule we are — this leads to shortcuts and extremely long hours"* |
| **3. Emotional texture and contradiction** | CX subsample: 93.8% job satisfaction AND schedule compression most frequent qualitative code (18.9%). Scale cannot hold both simultaneously. | High engagement and structural harm coexisting; team as buffer against management failure; episodic and career-trajectory pressure patterns. | *"Pressure is from management that have no clue what commissioning does — my team are great"* |
| **4. Respondent-defined significance** | No structured equivalent. Absence is the evidence — these phenomena could not be named without prompting. | Informal social performance as competitive pressure; employment precarity as structural constraint on disclosure. | *"Am on day hire so easy to get rid of when work dries up — unions don't help much either"* |

Table 1 summarises the contrast across all four categories, showing what the structured data returned alongside what the open-text accounts revealed. The left column shows what structured items returned — scores and correlations — and the right column shows what open-text revealed: mechanisms, relational dynamics, and respondent-defined conditions that scale positions cannot represent. Quotes are respondents' accounts of experience.

## 5. The Practitioner-Researcher Position as Methodological Resource

The lead researcher entered this study with 25 years of practitioner experience across engineering workplace environments in water, nuclear, and energy infrastructure. All three informal governance mechanisms identified in section 4.1 — unrecorded hours, unofficial instruction channels, hierarchy override — were recognisable from professional practice before data collection began. The network pressure mechanism in section 4.4, and the employment precarity that constrains disclosure, were equally familiar. None could be named in structured items without the attendant risks established in section 2.2.

This might appear to introduce bias: a researcher sensitised to patterns who then finds them. The argument here is more precise. Prior knowledge shaped recognition without predetermining what respondents would say — an abductive logic in which existing knowledge sensitises without foreclosing (Becker, 1998; Timmermans and Tavory, 2012). The open-text responses independently confirmed all identified mechanisms across disciplines, genders, age groups, and employment types. That independent confirmation is what matters. Practitioner knowledge shaped interpretation: respondents generated the evidence (Unluer, 2015).

The ethics constraint and the practitioner's knowledge are connected. It was practitioner knowledge that identified which informal mechanisms carried the ethical risks that precluded direct naming in structured items. A researcher without that knowledge might not have

recognised those risks. The combination — practitioner knowledge informing instrument design, ethics review confirming the constraint, open-text preserving respondent autonomy — created the conditions under which unprompted, independent confirmation became possible (Finlay, 2002). Being unable to name the phenomenon directly is itself informative about its nature.

The self-selection limitation of open-text warrants direct address. Respondents who provide open-text data may differ systematically from those who do not, potentially overrepresenting those with stronger or more salient experiences of workplace pressure. Three features of this dataset partially mitigate the concern, though they do not eliminate it.

First, the 55.8% response rate is substantially higher than typical voluntary open-text completion rates in professional surveys, suggesting engagement extending well beyond the most aggrieved respondents.

Second, the word count distribution reflects diverse rather than uniformly extreme engagement: the median is 14 words, with responses ranging from 4 to 229. A pure venting mechanism would skew more heavily toward longer responses.

Third, and most directly relevant to the bias concern: 25 of 180 codeable responses (13.9% of all coded instances) were classified under PPF — positive and protective factors describing supportive management, good team dynamics, or genuinely satisfying work. Respondents who had completed all structured items chose to write to say their experience was good. If the instrument design were steering respondents toward critical accounts, the PPF category would not exist at the scale it does. Its presence is not only a methodological reassurance — it is evidence that the open-text item was functioning as designed: genuinely open to whatever respondents judged the study needed to know. These mitigations do not eliminate the limitation. They clarify its scope: these are accounts from those who chose to engage, representing what this instrument design, under these conditions, made possible to name.

## 6. Towards Mixed Methods as Methodological Necessity

Structured items returned numbers: 63% reporting hours pressure, 66% role expansion, and correlations confirming hypothesised relationships. Open-text returned accounts: governance concerns, causal chains, emotional trajectories, and structural conditions that constrain voice. Not two versions of the same information at different levels of detail — epistemologically distinct, each recovering what the other, under these research conditions, tends not to.

Bryman (2006) argues the complementarity of quantitative and qualitative methods is well established. What is less often argued is the case where complementarity is not a design preference but a methodological requirement — where the nature of the phenomena makes both instruments jointly necessary. Because open-text responses tend to oversample stronger experiences, prevalence claims require structured data. Because direct structured naming of informal mechanisms risks methodological compromise in sensitive contexts, mechanism claims require open text. Neither method alone produces an adequate account. Together, they produce genuinely different things — which is exactly what is needed.

The methodological logic is likely transferable to other fields where formal structures coexist with power asymmetries. This argument claims transferability of the reasoning, not empirical equivalence across sectors — the data here are from engineering. Four adjacent domains illustrate the principle.

In bullying and harassment research, structured instruments are the norm. Yet the mechanisms through which bullying operates — indirect verbal demeaning, social isolation, informal hierarchy exploitation — are precisely those that ethics constraints make difficult to name directly without introducing prompting or distress. Nielsen et al. (2020) document how emotional and cognitive factors systematically reduce disclosure in structured bullying instruments, producing findings that capture outcomes without adequately accounting for the mechanisms that produce them. An open-text item would not replace established instruments — it would recover what those instruments, by their design parameters, tend not to reach.

In early career vulnerability research, the ethical case for open text is particularly strong. Naming specific forms of precarity in structured items risks both prompting and distress among the most professionally vulnerable respondents. In this study, respondents named financial dependency, scope overload, and fear of non-renewal unprompted.

In occupational health research, structured instruments measure outcomes: anxiety, burnout, and physical health impact. They are less well placed to capture the informal mechanisms that accumulate to produce those outcomes. One respondent in this study provided a detailed account of a near-miss incident involving schedule pressure, authority ambiguity, interpersonal intimidation, and compromised professional judgement. An anxiety scale would return a number; it would return nothing of the causal chain. Understanding why occupational harm occurs requires precisely the kind of contextual account that only an open-text item creates the conditions to produce.

In organisational justice research, structured instruments measure perceived fairness but are less well placed to capture the informal channels through which injustice operates. Neuert et al. (2025) demonstrate that format choice has substantive consequences for what can and cannot be known. In each of these fields, the decision to use structured-only formats shapes which phenomena enter the evidence base — and which do not.

The practical implication is direct: the question at the instrument design stage should not be whether to include an open-text item, but what is likely to be missed without one. In sensitive workplace research, both instruments are needed — not because more data is better, but because different instruments recover genuinely different things.

## 7. Conclusion

Structured survey instruments measure what researchers already know to ask about. The limitation this paper has demonstrated is not one of depth or nuance — it is one of category. Certain phenomena cannot be directly named in closed-ended items without introducing prompting, measurement bias, or participant distress. In sensitive workplace research, that constraint is not a design inconvenience. It is evidence that these phenomena require respondent-defined articulation to be adequately captured.

Four such categories were identified: informal governance mechanisms that formal records do not contain; causal accounts of how pressure operates that correlations cannot convey; emotional textures — contradiction, coexistence, trajectory — that scale positions cannot represent; and respondent-defined significance that no pre-specified item could have anticipated. Structured data established that pressure existed. Open-text accounts established how it operated. Not competing findings — epistemologically distinct contributions that together produce what neither instrument can produce alone.

The ethics constraint documented in this study's design provides evidence beyond theoretical argument: known informal pressure mechanisms could not be directly named as structured items without compromising participant welfare and measurement integrity. Open-text items are not a supplement to structured measurement, but its necessary complement.

The methodological logic is likely transferable to contexts where formal structures coexist with power asymmetries. In bullying research, early career studies, occupational health, and organisational justice, structured instruments are the norm, and the phenomena most resistant to direct structured measurement are often the most consequential. The question at the instrument design stage should not be whether to include an open-text item, but what is likely to be missed without one. Structured items tell us what is present and how often. Open-text items tell us why — and in what form — it operates. In sensitive workplace research, knowing both is not a luxury. It is the point.

**Acknowledgements**

AI tools (Grammarly) were used to support wording refinement and consistency checks during manuscript preparation. All analyses, coding decisions, interpretations, and the final text were developed and verified by the authors.

**References**

Ayres, J.R., May, I., Bustami, R., Adam, J., Gamage, S., 2026a. What Engineers Actually Say: A Qualitative Analysis of Workplace Pressure in a Global Engineering Survey. https://doi.org/10.31224/6614

Ayres, J.R., May, I., Bustami, R., Adam, J.H., Gamage, S.H., 2026b. Global Workplace Pressures in Engineering: A Cross-Sectional Survey of Role-Based Pressures, Peer Influence, and Psychological Outcomes. https://doi.org/10.31224/6598

Bakker, A.B., Demerouti, E., 2007. The Job Demands-Resources model: state of the art. J. Manag. Psychol. 22, 309–328. https://doi.org/10.1108/02683940710733115

Becker, H.S., 1998. Tricks of the trade: how to think about your research while you're doing it, Chicago guides to writing, editing, and publishing. University of Chicago Press, Chicago, Ill.

Bolton, S.C., 2005. Emotion management in the workplace, Management, work and organisations. Palgrave Macmillan, Houndmills, Basingstoke, Hampshire ; New York.

Braun, V., Clarke, V., Boulton, E., Davey, L., McEvoy, C., 2021. The online survey as a *qualitative* research tool. Int. J. Soc. Res. Methodol. 24, 641–654. https://doi.org/10.1080/13645579.2020.1805550

Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done? Qual. Res. 6, 97–113. https://doi.org/10.1177/1468794106058877

Dekker, S., 2011. Drift into Failure: From Hunting Broken Components to Understanding Complex Systems. https://doi.org/10.1201/9781315257396

Detert, J.R., Edmondson, A.C., 2011. Implicit Voice Theories: Taken-for-Granted Rules of Self-Censorship at Work. Acad. Manage. J. 54, 461–488. https://doi.org/10.5465/amj.2011.61967925

Dillman, D.A., Smyth, J.D., Christian, L.M., 2014. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 1st ed. Wiley. https://doi.org/10.1002/9781394260645

Dunwoodie, K., Macaulay, L., Newman, A., 2023. Qualitative interviewing in the field of work and organisational psychology: Benefits, challenges and guidelines for researchers and reviewers. Appl. Psychol. 72, 863–889. https://doi.org/10.1111/apps.12414

Feldman, M.S., Pentland, B.T., 2003. Reconceptualizing Organizational Routines as a Source of Flexibility and Change. Adm. Sci. Q. 48, 94–118. https://doi.org/10.2307/3556620

Finlay, L., 2002. Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. Qual. Res. 2, 209–230. https://doi.org/10.1177/146879410200200205

Fricker, M., 2007. Epistemic injustice: power and the ethics of knowing. Oxford university press, Oxford.

Hollnagel, E., 2018. Safety–I and Safety–II: The Past and Future of Safety Management, 1st ed. CRC Press. https://doi.org/10.1201/9781315607511

Hudson, J., Weinberg, A., 2025. Psychometric stress risk assessments: proceed with caution. Int. J. Workplace Health Manag. 18, 101–117. https://doi.org/10.1108/IJWHM-12-2022-0203

Israel, M., 2015. Research Ethics and Integrity for Social Scientists: Beyond Regulatory Compliance. SAGE Publications Ltd, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom. https://doi.org/10.4135/9781473910096

Jebb, A.T., Ng, V., Tay, L., 2021. A Review of Key Likert Scale Development Advances: 1995–2019. Front. Psychol. 12, 637547. https://doi.org/10.3389/fpsyg.2021.637547

Landis, J.R., Koch, G.G., 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33, 159. https://doi.org/10.2307/2529310

Neuert, C.E., Kunz, T., Gummer, T., 2025. An empirical evaluation of probing questions investigating question comprehensibility in web surveys. Int. J. Soc. Res. Methodol. 28, 367–378. https://doi.org/10.1080/13645579.2024.2391957

Nielsen, M.B., Notelaers, G., Einarsen, S.V., 2020. Methodological Issues in the Measurement of Workplace Bullying, in: Einarsen, S.V., Hoel, H., Zapf, D., Cooper, C.L. (Eds.), Bullying and Harassment in the Workplace. CRC Press, pp. 235–265. https://doi.org/10.1201/9780429462528-8

Schwarz, N., 1999. Self-reports: How the questions shape the answers. Am. Psychol. 54, 93–105. https://doi.org/10.1037/0003-066X.54.2.93

Star, S.L., Strauss, A., 1999. Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. Comput. Support. Coop. Work CSCW 8, 9–30. https://doi.org/10.1023/A:1008651105359

Timmermans, S., Tavory, I., 2012. Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. Sociol. Theory 30, 167–186. https://doi.org/10.1177/0735275112457914

Tourangeau, R., Rips, L.J., Rasinski, K., 2000. The Psychology of Survey Response, 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9780511819322

Tourangeau, R., Yan, T., 2007. Sensitive questions in surveys. Psychol. Bull. 133, 859–883. https://doi.org/10.1037/0033-2909.133.5.859

Unluer, S., 2015. Being an Insider Researcher While Conducting Case Study Research. Qual. Rep. https://doi.org/10.46743/2160-3715/2012.1752

Vaughan, D., 2016. The Challenger launch decision: risky technology, culture, and deviance at NASA ; with a new preface, Enlarged edition. ed. The University of Chicago Press, Chicago London.