

Workplace Safety Behavioral Tendencies in Highway Maintenance: Development of a Content-Valid Item Pool Based on a Dual-Process Framework

Paul W. Ford, Ph.D.

Walden University & Capitol Technology University

Highway maintenance work exposes employees to persistent safety demands involving live traffic, temporary traffic control, heavy equipment, variable weather, fatigue, and production pressure. The present study reports Phase 1 of a systematic post-exploratory refinement program for a context-specific item pool assessing safety-related behavioral tendencies. A provisional 36-item pool representing seven hypothesized dimensions was evaluated by 10 subject matter experts (SMEs) for relevance and clarity and then examined through cognitive interviews with 6 highway maintenance workers. Item-level content validity indices (I-CVI) ranged from .80 to 1.00 (S-CVI/Ave = .98), and mean item clarity was 3.29 on a 4-point scale. The SME decision audit classified 22 items for retention, 11 for revision, and 3 for removal; subsequent cognitive-interview and readability review produced additional wording refinements, yielding a 33-item revised pool. The near-ceiling relevance profile is interpreted cautiously because 5 of 10 SMEs were drawn from highway maintenance organizations and many panelists shared overlapping safety credentials. Cognitive interviews identified wording clarity and socially desirable phrasing as the most common interpretive problems. The resulting pool remains developmental and is prepared for subsequent psychometric testing; no new claims of reliability, factor structure, predictive validity, or operational use are made for the revised 33-item version.

INTRODUCTION

Highway maintenance work is performed in operational environments where exposure to risk is routine, dynamic, and often time-sensitive. Employees may work adjacent to live traffic, establish or modify temporary traffic control, operate or work around heavy equipment, respond to weather and lighting changes, and complete tasks under production demands. Recent work-zone research continues to identify driver behavior, changing traffic conditions, human performance variability, and work-zone complexity as contributors to worker exposure

Author info: Correspondence should be sent to: Paul W. Ford, Ph.D.

Email: paul.ford1@waldenu.edu

North American Journal of Psychology, 2026, Vol. 28, No. 4, 2303-2330

© NAJP <https://doi.org/10.65696/001c.172492>

and incident vulnerability (Chammout et al., 2024; Rathnasiri et al., 2024). These characteristics make highway maintenance an appropriate context in which to examine safety-related behavioral tendencies.

For this study, safety-related behavioral tendencies are defined as recurring but context-sensitive propensities in how workers notice and interpret hazards, initiate action, respond to social and production pressure, adapt to changing or ambiguous conditions, and engage in behaviors that can increase or suppress risk. The construct is not intended to represent an immutable personality trait, a medical or psychological diagnosis, or a direct measure of safety performance, injury risk, or employability. Instead, it targets person-situation response patterns that may help explain why workers facing similar operational demands respond differently to time pressure, uncertainty, fatigue, social expectations, and changing field conditions.

A substantial body of occupational safety research indicates that unsafe outcomes cannot be explained by environmental hazards alone. Human-error theory distinguishes unintended errors from deliberate violations (Reason, 1990), while meta-analytic and longitudinal research shows that person-related and situational variables both contribute to safety behavior and accident outcomes (Christian et al., 2009; Nahrgang et al., 2011; Neal & Griffin, 2006). Research on broad personality traits has also identified associations between individual differences and accident involvement, although effect patterns depend on context and broad traits are not equivalent to field-specific safety behavior (Clarke & Robertson, 2005). These findings support examining worker tendencies without treating those tendencies as deterministic causes of incidents.

Relationship to Existing Measures

The most direct precursor to the present work is the Workplace Safety Personality Indicator (WSPI-36), an exploratory 36-item instrument previously administered to 187 employees in a U.S. highway-maintenance organization (Ford, 2026). That version used six provisional dimensions and showed uneven internal consistency and only partial support for the proposed multidimensional structure, with particular weakness or overlap in Flexibility Under Pressure and Pattern Processing. Ford (2026) therefore concluded that additional refinement and external validation were needed. The present study is not presented as development of an unrelated new instrument; rather, it is the first systematic content-refinement phase following that exploratory administration. The original item pool is retained for traceability, but the construct architecture, wording, and scoring recommendations are reconsidered here before further large-sample testing.

Other established measures address important portions of the broader safety domain without duplicating the specific target of the present

refinement. The Work Safety Scale assesses perceptions of job safety, coworker safety, supervisor safety, management safety practices, and the safety program (Hayes et al., 1998), while Griffin and Neal's (2000) framework distinguishes safety climate from safety performance and separates safety compliance from safety participation. General personality measures capture broad dispositions associated with accident involvement (Clarke & Robertson, 2005), the Job Adaptability Inventory assesses adaptive performance across work contexts (Pulakos et al., 2000), and measures such as the MSTAT-II assess orientations toward complex, unfamiliar, or insoluble stimuli (McLain, 2009). These measures are appropriate for their intended purposes but operate at different levels of abstraction or criterion domains. The present item pool therefore has a narrower rationale: to integrate several safety-relevant response tendencies into behaviorally anchored items written for highway-maintenance conditions and to evaluate their content and interpretability before psychometric retesting.

THEORETICAL FRAMEWORK

The proposed framework organizes the item domain around two broad process classes: risk activation and risk suppression. Risk activation refers to tendencies that may increase vulnerability to unsafe action when employees operate under time pressure, ambiguity, fatigue, frustration, or social influence. Risk suppression refers to tendencies that may increase the likelihood of protective action, including checking, reporting, stopping work, and maintaining procedural controls. The term dual-process is used here to describe these competing classes of safety-relevant response tendencies. The framework is conceptually compatible with dual-process accounts in which rapid default responses may be modified by more deliberative control processes (Evans & Stanovich, 2013), but it does not assert a one-to-one correspondence between the seven proposed dimensions and canonical Type 1 or Type 2 cognition. The current seven-domain structure is a refinement of the six-domain WSPI-36 architecture reported by Ford (2026): Action Orientation, Pattern Processing, and Protective Orientation were retained conceptually; Human-Driven Motivation was reframed as Normative and Production Pressure Sensitivity; Risk Amplifiers was reframed as Susceptibility to Performance-Shaping Factors; and the heterogeneous Flexibility Under Pressure domain was separated into Adaptive Flexibility and Tolerance for Ambiguity and Degraded Control. These seven dimensions remain provisional organizing constructs rather than established psychometric factors.

Action Orientation

Action Orientation represents a tendency to initiate action quickly, rely on immediate problem solving, or continue task activity before all information has been reviewed. Reason's (1990) distinction between errors, violations, and the conditions that shape human action provides an occupational-safety basis for this domain, while dual-process theory provides a broader account of rapid default responding under conditions in which deliberation is constrained (Evans & Stanovich, 2013). Person-situation research in occupational safety further supports examining individual response tendencies without assuming that they operate independently of context (Christian et al., 2009). The proposed items therefore emphasize field behavior under urgency or incomplete information rather than a general personality label such as impulsivity.

Pattern Processing

Pattern Processing concerns the balance between reliance on familiar task patterns or experience and attention to smaller environmental changes. Risk-perception research demonstrates that judgments of hazard are shaped by how people interpret cues rather than by objective hazard characteristics alone (Slovic, 1987). Situation-awareness theory likewise emphasizes the acquisition, comprehension, and projection of environmental information and identifies workload, stress, attention, working-memory limits, and system complexity as influences on maintaining awareness in dynamic environments (Endsley, 1995). In highway maintenance, this domain is intended to capture whether task familiarity or concentrated task focus may reduce attention to changing traffic, equipment, or work-area cues.

Normative and Production Pressure Sensitivity

Normative and Production Pressure Sensitivity reflects responsiveness to coworker expectations, supervisory authority, schedule demands, and perceived pressure to maintain production. Safety-climate research established that workers interpret organizational priorities concerning safety and production (Zohar, 1980), and later work linked safety climate, motivation, behavior, and accidents over time (Neal & Griffin, 2006). Job demands and resources also relate to burnout, engagement, and safety outcomes (Nahrgang et al., 2011). Psychological-safety research adds a complementary perspective on whether employees feel able to speak up, question, or interrupt ongoing work (Edmondson & Bransby, 2023). The dimension therefore targets perceived pressure and social response tendencies rather than inferring that workers intentionally disregard safety.

Adaptive Flexibility

Adaptive Flexibility represents adjustment to changing plans or operating conditions while attempting to preserve necessary controls.

Pulakos et al. (2000) developed an eight-dimension taxonomy of adaptive performance, demonstrating that adaptation is a meaningful and multidimensional work-performance domain. Self-regulation theory similarly emphasizes feedback-guided adjustment in pursuit of goals (Carver & Scheier, 1982). The earlier WSPI-36 combined adaptive response, ambiguity tolerance, and related content within Flexibility Under Pressure, which was the least internally consistent subscale and showed evidence of heterogeneous item functioning (Ford, 2026). The present framework therefore treats adaptability as a narrower provisional domain in which the safety-relevant question is whether workers can adjust to changing field conditions while maintaining required controls. Because only three items currently represent this dimension, its content breadth remains provisional despite strong SME ratings.

Tolerance for Ambiguity and Degraded Control

Tolerance for Ambiguity and Degraded Control was designed to capture response tendencies when routines, instructions, or control conditions become less stable. Ambiguity tolerance has a long measurement history (Budner, 1962), and the MSTAT-II conceptualizes it as an orientation toward complex, unfamiliar, or insoluble stimuli (McLain, 2009). Highway maintenance adds a distinct operational concern: uncertainty may coexist with changing traffic control, unfamiliar task conditions, incomplete guidance, or competing demands. The separation of this domain from Adaptive Flexibility also responds to the weak and heterogeneous performance of the earlier Flexibility Under Pressure domain (Ford, 2026). Situation-awareness and human-error perspectives further suggest that uncertainty, workload, and degraded control can alter information processing and action (Endsley, 1995; Reason, 1990). The domain is deliberately treated as exploratory because only two items survived the present review, which is unlikely to represent its full conceptual range.

Protective Orientation

Protective Orientation represents active safety-regulating behavior such as checking equipment and controls, reporting hazards, and stopping work when conditions appear unsafe. Griffin and Neal's (2000) distinction between safety compliance and safety participation demonstrates that safety performance includes both required behaviors and broader participation in maintaining a safe environment. Self-regulation models likewise emphasize feedback, monitoring, and corrective action (Carver & Scheier, 1982). The present dimension is not intended as a general measure of conscientiousness or safety climate; it samples concrete actions that can suppress risk once a hazard or control requirement is recognized.

Susceptibility to Performance-Shaping Factors

Susceptibility to Performance-Shaping Factors addresses conditions such as rushing, fatigue, frustration, and discomfort that may degrade safe behavior. Meta-analytic evidence links job demands and strain-related processes with safety outcomes (Nahrgang et al., 2011), and fatigue has a documented relationship with safety across occupational and transportation contexts (Williamson et al., 2011). Human-error theory similarly emphasizes that performance must be interpreted in relation to the conditions in which behavior occurs (Reason, 1990). Items in this domain therefore focus on workers' perceived vulnerability to common performance-shaping conditions rather than on moralized descriptions of rule breaking.

PURPOSE OF THE CURRENT STUDY

The purpose of the present study was to conduct Phase 1 of a systematic post-exploratory refinement program for a theory-driven, context-specific item pool assessing safety-related behavioral tendencies in highway maintenance work. The earlier exploratory WSPI-36 administration provided preliminary psychometric information but also identified structural and item-level weaknesses requiring additional construct clarification and item refinement (Ford, 2026). Consistent with scale-development guidance that separates content development and validation from subsequent psychometric evaluation (Boateng et al., 2018), Phase 1 was organized into four steps: (1) construct definition and item generation/re-specification, (2) SME content review, (3) cognitive interviewing, and (4) readability review and revision auditing. The objective was to establish a more defensible content foundation and interpretable item wording for the revised pool. No new claims of internal consistency, factor structure, criterion-related validity, predictive utility, measurement invariance, or operational classification are made for the 33-item version.

METHOD

Participants

Two participant groups contributed to Phase 1. The content-review panel comprised 10 SMEs selected to provide expertise in highway maintenance operations, work-zone safety, transportation safety, safety management, safety education, human factors, or industrial-organizational psychology. Selection emphasized direct operational or professional safety experience and relevant credentials. Five experts worked in highway maintenance organizations; the remaining experts represented transportation, construction, independent consulting, and higher education. Table 1 summarizes panel characteristics. The cognitive-

interview sample comprised 6 highway maintenance workers: three roadway technicians, one foreman, and two supervisors, with 1 to 23 years of experience (Table 2). Participants were recruited voluntarily from highway-maintenance operations.

Table 1
Subject Matter Expert Characteristics

ID	Role	Setting	Yrs	Credentials	Primary focus
SME-01	Project Manager	Highway maint.	14	CPM	ATSSA; MUTCD
SME-02	Regional Safety Mgr.	Highway maint.	8	GSP	ATSSA; MUTCD
SME-03	Director of Safety	Transportation	21	CSP; CDS	FMCSA; OSHA
SME-04	Safety Consultant	Independent	14	CSP; CHST	Safety education
SME-05	Division Manager	Highway maint.	12	TCT; TCS	ATSSA; MUTCD
SME-06	Area Manager	Highway maint.	9	TCT; TCS	ATSSA; MUTCD
SME-07	Assistant Professor	Higher education	10	CSP; CIH	Safety education
SME-08	Dissertation Chair	Higher education	8	CSP; CHST	Safety education
SME-09	Project Manager	Highway maint.	26	TCT; TCS	ATSSA; MUTCD
SME-10	Regional Safety Mgr.	Construction	17	CSP	OSHA; NIOSH

Note. CPM = Certified Project Manager; GSP = Graduate Safety Practitioner; CSP = Certified Safety Professional; CDS = Certified Director of Safety; CHST = Construction Health and Safety Technician; TCT = Traffic Control Technician; TCS = Traffic Control Supervisor; CIH = Certified Industrial Hygienist; ATSSA = American Traffic Safety Services Association; MUTCD = Manual on Uniform Traffic Control Devices; FMCSA = Federal Motor Carrier Safety Administration; OSHA = Occupational Safety and Health Administration; NIOSH = National Institute for Occupational Safety and Health.

Table 2
Cognitive Interview Participant Characteristics

ID	Role	Supervisory	Years experience
RW-1	Roadway Technician	No	1
RW-2	Roadway Technician	No	3
RW-3	Roadway Technician	No	5
RW-4	Foreman	Yes	16
SUP-5	Supervisor	Yes	12
SUP-6	Supervisor	Yes	23

Note. De-identified region/crew codes were retained in interview notes but are omitted here for brevity. Day-versus-night work status was not retained.

Instrument and Response Format

The provisional instrument contained 36 statements representing the seven hypothesized dimensions. The intended worker-response format was a 5-point Likert-type agreement scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, and 5 = Strongly agree. This format matches the retained supplemental scoring criteria and the five-category response format used in the earlier exploratory WSPI-36 administration (Ford, 2026). An agreement format was retained because the items are self-referent statements about safety-related response tendencies and because five categories provide a neutral midpoint while remaining straightforward for field administration. The present Phase 1 study evaluated item content and interpretation rather than reliability, score distributions, or factor structure. SME relevance and clarity judgments used separate 4-point review scales described below.

Procedure

Phase 1: Step 1 - Construct Definition and Item Generation

The refinement began with the original WSPI-36 item set and the conceptual and psychometric concerns identified in the exploratory administration (Ford, 2026). Items were reconsidered deductively against literature on human error, dual-process cognition, risk perception, social and production pressure, adaptive performance, ambiguity tolerance, fatigue, psychological safety, and self-regulation, and inductively against operational demands of highway maintenance work, including live-lane exposure, temporary traffic control, equipment interaction, changing conditions, and pressure to maintain productivity. The six-domain exploratory architecture was re-specified as seven provisional domains, most notably by separating the earlier Flexibility Under Pressure content into Adaptive Flexibility and Tolerance for Ambiguity and Degraded Control. An internal review then examined redundancy, conceptual overlap, readability, and wording consistency while preserving the 36 original item identifiers for auditability. This approach aligns with the Standards for Educational and Psychological Testing, which emphasize explicit construct definitions, transparent development procedures, and evidence supporting proposed score interpretations (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). Pattern Processing Items 7 and 8 were retained despite close semantic overlap so that their redundancy could be evaluated empirically rather than resolved solely by author judgment.

Phase 1: Step 2 - SME Content Review

SMEs independently rated each item's relevance to its assigned domain on a 4-point scale from 1 (not relevant) to 4 (highly relevant) and clarity

on a 4-point scale from 1 (not clear) to 4 (very clear). I-CVI was calculated as the proportion of experts assigning a relevance rating of 3 or 4, and S-CVI/Ave was calculated as the mean I-CVI across items. Modified kappa coefficients were calculated to describe agreement beyond chance, following the logic described by Polit et al. (2007). Average deviation (AD) indices were used only as descriptive indicators of rating dispersion; no universal AD cutoff was treated as a validity criterion. Decision guidance was to retain items with I-CVI at least .78 and mean clarity at least 3.20, revise items with acceptable relevance but weaker clarity or substantive qualitative concerns, and remove items with weak relevance and limited theoretical or interpretive rescue potential. Numerical rules informed rather than mechanically determined final decisions.

Phase 1: Step 3 - Cognitive Interviewing

Participants completed a think-aloud item review in which they read the items, explained what each statement meant in their own words, and described how they would arrive at a response. Follow-up probes targeted comprehension problems, unclear wording, socially desirable phrasing, double-barreled content, response-mapping difficulty, and mismatch with field language. Notes were coded by item and problem type by a single primary researcher. The same researcher conducted a complete second-pass consistency review to reconcile ambiguous coding decisions and apply final code definitions consistently. This procedure improved internal consistency of the audit trail but did not provide independent coding, inter-rater agreement, or protection against a single researcher's interpretive lens; therefore, the coded issue frequencies are treated as descriptive refinement evidence rather than independently replicated qualitative findings.

Phase 1: Step 4 - Readability and Revision Audit

A Flesch-Kincaid grade-level screen was applied to the provisional and revised item pools. The original pool screened at approximately grade 7.9. After wording simplification and item revision, the retained pool screened at approximately grade 6.2 overall, with item-level estimates from about grade 0.8 to 8.5; no retained item exceeded grade 9.0. Readability estimates were used as screening indicators rather than evidence that all worker groups would interpret every item equivalently. All wording changes and item-removal decisions were retained in an audit trail so the original and final stems could be compared directly.

Ethical Considerations

This investigator-initiated instrument-refinement project was conducted independently and was not submitted to an institutional review board or other formal research ethics committee. Before providing feedback, each SME and cognitive-interview participant was informed of

the project's purpose and scope, what participation would involve, the voluntary nature of participation, and the intended research and instrument-development use of the feedback. Participants had an opportunity to ask questions before proceeding and could decline participation or discontinue their involvement. No compensation was provided. Personally identifying information was not retained in the analytic files beyond de-identified participant labels. The absence of prospective independent ethics review is reported here explicitly for transparency.

RESULTS

Across the 36 provisional items, I-CVI values ranged from .80 to 1.00 and S-CVI/Ave was .98. Thirty-two items (88.9%) received unanimous relevance endorsement (I-CVI = 1.00), one item (2.8%) received .90, and three items (8.3%) received .80. Mean clarity across the full pool was 3.29 on the 4-point scale. The SME-step decision audit recorded 22 items as retain, 11 as revise, and 3 as remove. Subsequent cognitive interviewing and readability editing changed the wording of several items originally marked retain; thus, the SME counts describe the review decision at that step and should not be interpreted as the number of final stems that remained verbatim. The final Phase 1 pool contains 33 items.

Modified kappa values ranged from .79 to 1.00, a pattern best described as strong to near-perfect agreement beyond chance. Mean AD values were .25 for relevance and .35 for clarity, indicating tightly clustered ratings. The convergence of high I-CVI values, high modified kappas, and low dispersion supports strong perceived content relevance but also makes limited rating variability an important interpretive issue. Five of the 10 SMEs came from highway maintenance organizations, two were higher-education safety experts, and several panelists shared work-zone, traffic-control, or professional safety credentials. Such alignment may have increased shared assumptions about what constitutes relevant safety content and may have contributed to the ceiling pattern. A more heterogeneous panel - for example, one including a larger proportion of front-line craft workers as content experts, independent human-factors specialists, occupational-health professionals, and experts from other roadway organizations - could plausibly identify additional disagreements or omitted facets, although greater heterogeneity would not guarantee greater dispersion.

Items 23, 33, and 36 were removed because their combined SME and cognitive-interview evidence indicated low interpretive rescue potential through wording change alone. Dimension-level results are shown in Table 3. Adaptive Flexibility, Protective Orientation, and Normative and Production Pressure Sensitivity showed strong content support. Tolerance

for Ambiguity and Degraded Control had the weakest content-validity profile, and Susceptibility to Performance-Shaping Factors contained two of the three removed items. These patterns suggest that concrete, behaviorally anchored content was easier for experts and workers to interpret consistently than abstract, comparative, or highly socially desirable content.

Table 3
SME Content-Validity Results by Dimension

Dimension	n	I-CVI M	I-CVI SD	Clarity M	Clarity SD	κ^*	Ret. / Rev.-Drop
Action Orientation	6	.98	.04	3.37	.38	.98	4 / 2
Adaptive Flexibility	3	1.00	.00	3.70	.26	1.00	3 / 0
Normative/Production Pressure	6	1.00	.00	3.28	.25	1.00	4 / 2
Pattern Processing	6	1.00	.00	3.17	.27	1.00	2 / 4
Protective Orientation	6	1.00	.00	3.52	.31	1.00	5 / 1
Performance-Shaping Factors	6	.97	.08	3.08	.44	.97	3 / 3
Ambiguity/Degraded Control	3	.87	.12	2.93	.42	.86	1 / 2

Note. κ^* = modified kappa. Ret. / Rev.-Drop reports the SME-step count retained without requested revision versus the combined number marked revise or drop. Overall mean AD relevance = .25 and mean AD clarity = .35; AD was interpreted descriptively rather than against a universal threshold.

Across the six cognitive interviews, clarity problems and socially desirable wording were the most common flagged issues, followed by direct comprehension problems. Of items with coded interview feedback, 6 were understood as intended, 5 were partly understood as intended, and 3 were not understood as intended. Workers often recognized the intended safety concept but reacted to wording that implied an obvious 'good' answer or did not match field language. For example, Item 24 was interpreted as 'I stay calm under pressure,' which did not adequately represent ambiguity tolerance; Item 33 elicited comparative overconfidence rather than a behaviorally anchored field tendency; and Item 32 was more readily understood after fatigue was connected to following detailed safety steps. Table 4 summarizes coded interview issues by dimension, and Appendix C provides the item-level audit.

Table 4
Cognitive Interview Findings by Dimension

Dimension	Flags	Comp.	Clarity	Social	Double	Rev./Rem.
Action Orientation	5	0	3	2	0	3
Adaptive Flexibility	0	0	0	0	0	0
Normative/Production Pressure	4	0	2	2	0	2
Pattern Processing	5	0	3	2	0	2
Protective Orientation	2	0	1	1	0	1
Performance-Shaping Factors	6	1	3	2	0	3
Ambiguity/Degraded Control	4	2	1	1	0	2

Note. Comp. = comprehension; Social = socially desirable wording; Double = double-barreled content; Rev./Rem. = revise or remove. Counts are descriptive coding flags; they are not independent reliability estimates.

DISCUSSION

This Phase 1 study provides content and interpretive evidence for a revised, context-specific item pool intended to capture safety-related behavioral tendencies in highway maintenance. It follows the exploratory WSPI-36 administration reported by Ford (2026), which showed sufficient conceptual signal to justify continued development but also identified weak or overlapping content that required refinement. The principal contribution of the present study is therefore not a new claim of psychometric validity; it is a transparent content-refinement sequence in which the item domain was re-specified, reviewed by SMEs, examined through worker cognitive interviews, screened for readability, and documented through a revision audit. The findings indicate which proposed domains and item forms appear sufficiently relevant and interpretable to justify subsequent psychometric evaluation of the revised pool.

Item Redundancy and Construct Distinctiveness

The similarity between Pattern Processing Items 7 and 8 is substantive. Both items concern missing small environmental changes while attention is concentrated on the task, and their retained stems remain intentionally close. The earlier exploratory WSPI-36 findings already suggested that Pattern Processing contained weaker or overlapping item content (Ford, 2026), so the pair should be treated as an explicit redundancy test rather than assumed to represent distinct content. At Phase 1, both were retained to avoid premature content loss based solely on author judgment. Their retention does not imply that both belong in a final scale. Phase 2 item reduction should examine inter-item correlations, factor loadings, residual or local dependence, response distributions, and whether either item contributes unique content. If the pair functions redundantly, one should

be removed or re-specified. Similar scrutiny should be applied across the entire pool so that content coverage is balanced against item economy.

SME Panel Composition and Ceiling-Level Ratings

The near-ceiling SME profile warrants caution. Thirty-two of 36 items received unanimous relevance endorsement, and the panel shared substantial occupational-safety language, training, and credentialing. Five experts came from highway maintenance organizations and several held overlapping ATSSA, MUTCD, CSP, CHST, TCT, or TCS backgrounds. This composition is defensible for judging highway-maintenance content, but it can also constrain conceptual diversity. Experts who share similar professional socialization may make similar assumptions about what a safety item means or which behaviors matter. Accordingly, the strong CVI results are best understood as evidence of relevance within a highly aligned expert panel rather than as universal agreement across all plausible content experts. A future content-review round should deliberately broaden expert roles and organizational perspectives before treating the domain map as saturated.

Cognitive Interview Coding

Cognitive interviewing added evidence that high relevance ratings did not guarantee clear worker interpretation. However, all notes were coded by one researcher. The second-pass review reduced within-coder inconsistency, but it cannot substitute for independent interpretation or an inter-rater reliability estimate. Issue frequencies and classifications may therefore reflect the researcher's coding decisions as well as participant meaning. This limitation lowers confidence in treating the qualitative category counts as reproducible measurements. The interviews remain useful for identifying concrete wording problems and documenting why specific revisions were made, but Phase 2 refinement would be strengthened by independent coding of at least a subset of interviews or a consensus procedure involving a second trained reviewer.

Construct Underrepresentation

Adaptive Flexibility and Tolerance for Ambiguity and Degraded Control require particular attention to construct representation. Their separation was theoretically motivated and also responsive to the weak, heterogeneous Flexibility Under Pressure domain in the earlier exploratory version (Ford, 2026). In the present review, Adaptive Flexibility retained three items and received strong SME ratings, but three items provide limited coverage of a potentially broad adaptive-performance domain. Tolerance for Ambiguity and Degraded Control is more problematic: after Item 23 was removed, only Items 22 and 24 remained. Two items are insufficient to assume adequate domain representation or a stable standalone latent factor, and the surviving stems may emphasize routine

deviation and control degradation more than the full ambiguity-tolerance construct. The domain is therefore retained only as an exploratory hypothesis for Phase 2. Additional candidate items should be generated to sample uncovered facets of both dimensions and subjected to the same SME and cognitive-review process before they are incorporated into any later pool. If the two-item tolerance domain fails to achieve conceptual and empirical separation, deletion, merger, or re-specification would be preferable to preserving it by assumption.

Implications for the Dual-Process Framework

At the content level, the results are compatible with distinguishing processes that may activate risk from behaviors that may suppress it. Protective Orientation was strongly supported and comparatively concrete, consistent with the view that checking, reporting, and stop-work action are recognizable safety-regulating behaviors. By contrast, items intended to represent ambiguity, fatigue-related vulnerability, or comparative optimism were harder to express without comprehension or social-desirability problems. This asymmetry does not validate a two-process factor structure. It instead suggests that the suppression side of the framework may be easier to operationalize behaviorally than some activation-side constructs. Phase 2 factor analysis must determine whether the seven provisional dimensions, a smaller set of factors, or a different structure best represents observed responses.

INTENDED USE AND MISUSE

The revised item pool is intended for research and developmental use only. Although the earlier exploratory WSPI-36 study used a provisional Safety Risk Index and risk bands (Ford, 2026), those composite scoring and classification procedures are not carried forward as validated features of the present 33-item pool. The current supplemental scoring criteria recommend provisional dimension means only and explicitly advise against a total risk score, weighted composite, cutoff, banding system, or employee classification pending further validation. The revised pool is not validated for hiring, promotion, discipline, fitness-for-duty, return-to-work, or other adverse employment decisions. The Standards for Educational and Psychological Testing emphasize that proposed interpretations and uses of scores require evidence appropriate to the purpose, including construct definition, validity, reliability, and fairness (AERA et al., 2014). Employment applications would additionally require legal review, job-related validation, and adverse-impact analysis, particularly because several items address fatigue, frustration, pressure, and related performance-shaping conditions (Equal Employment Opportunity Commission [EEOC], 2007; Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, &

Department of Justice, 1978). The present study does not supply that evidence.

LIMITATIONS

Several limitations constrain interpretation. First, Phase 1 evidence is limited to content validity, interpretive refinement, and readability screening; internal consistency, factor structure, convergent and discriminant validity, criterion-related validity, predictive validity, and measurement invariance of the revised 33-item pool remain untested. Second, the evidence comes from a small cognitive-interview sample and a 10-member SME panel rather than from a broad multi-site workforce. Third, the SME panel was professionally aligned and weighted toward highway-maintenance management and safety roles, which may have contributed to ceiling-level ratings. Fourth, cognitive-interview coding was performed by one researcher without independent coding or inter-rater reliability. Fifth, the two retained Tolerance for Ambiguity and Degraded Control items and the three Adaptive Flexibility items may underrepresent their intended domains. Sixth, the project did not undergo prospective review by an institutional review board or other independent research ethics committee; participant protections therefore rested on the investigator's informed-participation procedures rather than independent oversight. Seventh, readability estimates do not substitute for empirical evaluation of literacy demands across worker groups. Finally, day-versus-night work status was not retained, a meaningful sampling gap in a highway-maintenance context where night work is common.

RECOMMENDATIONS FOR FUTURE RESEARCH

Phase 2 should begin with a larger, multi-site sample of highway maintenance personnel or comparable field-operations workers. For the current 33-item pool, approximately 330 to 660 participants would correspond to conventional item-to-participant ratios of roughly 10:1 to 20:1, although sample adequacy should ultimately be judged using communalities, factor strength, model complexity, and planned analyses rather than a ratio alone. Exploratory factor analysis with parallel analysis should guide factor retention, followed by confirmatory work in an independent sample when feasible. Dimensions represented by fewer than three well-functioning items should be treated as candidates for deletion, merger, or conceptual re-specification.

Before or alongside large-sample testing, additional candidate items should be developed for Adaptive Flexibility and especially Tolerance for Ambiguity and Degraded Control, and those items should undergo new content and cognitive review rather than being inserted directly into the validated pool. Pattern Processing Items 7 and 8 should be treated as an explicit redundancy test case. Phase 2 should examine item distributions,

corrected item-total relationships, factor loadings, cross-loadings, residual dependence, internal consistency, and convergent and discriminant relationships with established measures of safety climate, safety compliance and participation, conscientiousness, adaptability, ambiguity tolerance, and fatigue-related functioning. Criterion linkage should use time-lagged and exposure-adjusted outcomes where feasible, including near misses, stop-work activity, preventable incidents, and recordable events. Only after those steps should operational scoring or classification be considered.

CONCLUSION

This study presents a theory-grounded, context-specific 33-item pool for further study of safety-related behavioral tendencies in highway maintenance. Building on the earlier exploratory WSPI-36 administration, the present work re-specified the construct architecture and subjected the item pool to structured expert review, worker cognitive interviewing, readability screening, and a documented revision audit. The evidence supports continued development but also identifies clear unresolved issues, including redundancy within Pattern Processing, professional homogeneity of the SME panel, single-researcher cognitive coding, and underrepresentation of Adaptive Flexibility and especially Tolerance for Ambiguity and Degraded Control. The resulting Version 0.2 pool should therefore be understood as an empirically refined developmental instrument and a foundation for subsequent psychometric testing, not as a finished safety assessment or an employment decision tool.

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, Article 149. <https://doi.org/10.3389/fpubh.2018.00149>
- Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of Personality*, 30(1), 29-50. <https://doi.org/10.1111/j.1467-6494.1962.tb02303.x>
- Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin*, 92(1), 111-135.
- Chammout, B., Ahmed, M. O., & El-Adaway, I. H. (2024). Assessing the critical factors influencing worker safety in roadway work zones. *Journal of Construction Engineering and Management*, 150(12), Article 04024180. <https://doi.org/10.1061/JCEMD4.COENG-15205>

- Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology, 94*(5), 1103-1127. <https://doi.org/10.1037/a0016172>
- Clarke, S., & Robertson, I. T. (2005). A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings. *Journal of Occupational and Organizational Psychology, 78*(3), 355-376. <https://doi.org/10.1348/096317905X26183>
- Edmondson, A. C., & Bransby, D. P. (2023). Psychological safety comes of age: Observed themes in an established literature. *Annual Review of Organizational Psychology and Organizational Behavior, 10*, 55-78. <https://doi.org/10.1146/annurev-orgpsych-120920-055217>
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64. <https://doi.org/10.1518/001872095779049543>
- Equal Employment Opportunity Commission. (2007). *Employment tests and selection procedures*. <https://www.eeoc.gov/laws/guidance/employment-tests-and-selection-procedures>
- Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978). *Uniform guidelines on employee selection procedures*. 29 C.F.R. Part 1607.
- Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223-241. <https://doi.org/10.1177/1745691612460685>
- Ford, P. W. (2026). Workplace Safety Personality Indicator (WSPI-36): A behavioral approach to understanding occupational safety risk. *Open Journal of Social Sciences, 14*, 27-37. <https://doi.org/10.4236/jss.2026.147003>
- Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology, 5*(3), 347-358. <https://doi.org/10.1037/1076-8998.5.3.347>
- Hayes, B. E., Perander, J., Smecko, T., & Trask, J. (1998). Measuring perceptions of workplace safety: Development and validation of the Work Safety Scale. *Journal of Safety Research, 29*(3), 145-161. [https://doi.org/10.1016/S0022-4375\(98\)00011-5](https://doi.org/10.1016/S0022-4375(98)00011-5)
- McLain, D. L. (2009). Evidence of the properties of an ambiguity tolerance measure: The Multiple Stimulus Types Ambiguity Tolerance Scale-II (MSTAT-II). *Psychological Reports, 105*(3), 975-988. <https://doi.org/10.2466/PRO.105.3.975-988>
- Nahrgang, J. D., Morgeson, F. P., & Hofmann, D. A. (2011). Safety at work: A meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. *Journal of Applied Psychology, 96*(1), 71-94. <https://doi.org/10.1037/a0021484>
- Neal, A., & Griffin, M. A. (2006). A study of the lagged relationships among safety climate, safety motivation, safety behavior, and accidents at the individual and group levels. *Journal of Applied Psychology, 91*(4), 946-953. <https://doi.org/10.1037/0021-9010.91.4.946>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*(4), 459-467. <https://doi.org/10.1002/nur.20199>

- Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612-624.
<https://doi.org/10.1037/0021-9010.85.4.612>
- Rathnasiri, N., De Silva, N., & Wijesundara, J. (2024). State of the art in work zone safety: A systematic review. *International Journal of Transportation Science and Technology, 13*(1), 14-28.
<https://doi.org/10.1016/j.ijst.2023.11.006>
- Reason, J. (1990). *Human error*. Cambridge University Press.
- Slovic, P. (1987). Perception of risk. *Science, 236*(4799), 280-285.
- Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K., & Connor, J. L. (2011). The link between fatigue and safety. *Accident Analysis & Prevention, 43*(2), 498-515.
- Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology, 65*(1), 96-102.

APPENDIX A

Revised Provisional Item Pool (Version 0.2)

Item numbers preserve the original identifiers so that the final wording can be traced to the SME review, cognitive-interview summary, and revision audit. Items 23, 33, and 36 were removed. Items 7 and 8 are intentionally retained as a near-redundant pair for empirical item-reduction testing in Phase 2; their retention does not imply that both will remain in a final scale.

Item	Dimension	Final retained wording
1	Action	I usually act quickly when something needs to be done, even if all the details are not clear.
2	Action	When equipment seems off, I first try to fix it myself before stopping work.
3	Action	I am comfortable making decisions on the fly.
4	Action	When work is moving fast, I am less likely to slow down and check every hazard.
5	Action	I prefer solving problems by taking action rather than by planning.
6	Action	When things get urgent, I trust my own judgment more than the rules.
7	Pattern	When I focus on the task, I can miss small changes around me.
8	Pattern	When I focus on the job, I sometimes miss small changes around me.
9	Pattern	I would rather know the goal than follow each step.
10	Pattern	When a task feels familiar, I rely more on experience than on written checklists.
11	Pattern	Detailed safety steps feel harder to use when the job feels routine.
12	Pattern	I assume things are safe unless something obvious looks wrong.
13	Norm./Prod.	I feel pressure to help the crew finish work, even when safety steps take more time.
14	Norm./Prod.	I feel pressure to keep work moving even if safety rules slow things down.

Item	Dimension	Final retained wording
15	Norm./Prod.	I worry about how coworkers see me if I stop a job for safety reasons.
16	Norm./Prod.	I feel pressure to put production ahead of safety steps.
17	Norm./Prod.	It is hard for me to question a supervisor about safety.
18	Norm./Prod.	I try to avoid causing delays, even if something feels unsafe.
19	Adaptive	I work best when things are changing quickly.
20	Adaptive	I am comfortable adjusting on the fly when plans change.
21	Adaptive	When plans change, I keep required safety controls in place.
22	Ambig./Control	I can work safely even when a task does not follow a familiar routine.
24	Ambig./Control	When work gets hectic, it is harder to keep every safety control in place.
25	Protective	When I notice an unsafe condition, I stop work before continuing.
26	Protective	Even when I am not being observed, I follow the required safety steps.
27	Protective	I check equipment twice before I use it.
28	Protective	I report hazards, including minor ones, when they could affect the work area.
29	Protective	Before work starts, I check that protections and controls are in place.
30	Protective	Before work starts, I think about how my actions affect me and nearby coworkers.
31	Perf.-Shape	I take more risks when I feel rushed.
32	Perf.-Shape	When I am tired, it is harder for me to follow detailed safety steps.
34	Perf.-Shape	I get frustrated when safety slows down production.
35	Perf.-Shape	I feel pressure to push through discomfort to get the job done.

Note. Intended worker response format: 1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree. The pool remains developmental and is not approved for operational employment decisions.

APPENDIX B

Item-Level SME Content-Validity Results

#	Dim.	I-CVI	Cl. M	Cl. SD	κ^*	Decision	Revision note
1	Action	1.00	3.40	0.52	1.00	Retain	
2	Action	0.90	2.80	0.63	0.90	Revise	Social desirability; vague repair wording
3	Action	1.00	3.70	0.48	1.00	Retain	
4	Action	1.00	3.00	0.67	1.00	Revise	Judgmental phrasing; social desirability
5	Action	1.00	3.70	0.48	1.00	Retain	
6	Action	1.00	3.60	0.52	1.00	Retain	Admission-like wording; field applicability

#	Dim.	I-CVI	CL M	CL SD	κ^*	Decision	Revision note
7	Pattern	1.00	3.20	0.42	1.00	Revise	Abstract wording
8	Pattern	1.00	3.30	0.67	1.00	Retain	
9	Pattern	1.00	3.20	0.63	1.00	Revise	Broad wording
10	Pattern	1.00	2.70	0.48	1.00	Revise	Clarity; social desirability
11	Pattern	1.00	3.10	0.57	1.00	Revise	Clarity; social desirability
12	Pattern	1.00	3.50	0.53	1.00	Retain	Softened absolute wording
13	Norm./Prod.	1.00	2.90	0.57	1.00	Revise	Judgmental rule-bending wording
14	Norm./Prod.	1.00	3.40	0.52	1.00	Retain	
15	Norm./Prod.	1.00	3.30	0.48	1.00	Retain	
16	Norm./Prod.	1.00	3.10	0.32	1.00	Revise	Judgmental self-report; pressure conflation
17	Norm./Prod.	1.00	3.60	0.52	1.00	Retain	
18	Norm./Prod.	1.00	3.40	0.52	1.00	Retain	
19	Adaptive	1.00	3.80	0.42	1.00	Retain	
20	Adaptive	1.00	3.40	0.70	1.00	Retain	
21	Adaptive	1.00	3.90	0.32	1.00	Retain	Ability/behavior overlap simplified
22	Ambig./Control	1.00	3.40	0.52	1.00	Retain	Concrete wording; domain exploratory
23	Ambig./Control	0.80	2.60	0.84	0.79	Drop	Low clarity; comprehension failure
24	Ambig./Control	0.80	2.80	0.42	0.79	Revise	Hectic-condition comprehension failure
25	Protective	1.00	3.40	0.52	1.00	Retain	Absolute/ceiling-prone wording
26	Protective	1.00	3.80	0.42	1.00	Retain	Absolute/ceiling-prone wording
27	Protective	1.00	3.30	0.48	1.00	Retain	
28	Protective	1.00	3.10	0.57	1.00	Revise	Social desirability; reporting threshold
29	Protective	1.00	3.90	0.32	1.00	Retain	
30	Protective	1.00	3.60	0.52	1.00	Retain	Behavioral stem simplified
31	Perf.-Shape	1.00	3.30	0.67	1.00	Retain	
32	Perf.-Shape	1.00	3.10	0.57	1.00	Revise	Fatigue wording needed behavioral anchor
33	Perf.-Shape	1.00	2.60	0.52	1.00	Drop	Comparative overconfidence wording
34	Perf.-Shape	1.00	3.60	0.52	1.00	Retain	
35	Perf.-Shape	1.00	3.40	0.52	1.00	Retain	
36	Perf.-Shape	0.80	2.50	0.53	0.79	Drop	Optimism-bias wording; social desirability

Note. Dim. labels are abbreviated for table width. κ^* = modified kappa. Decision reflects the SME review step; later cognitive and readability review produced additional wording changes to some items initially marked retain.

APPENDIX C

Item-Level Cognitive Interview Summary

#	Dim.	Understanding / issue	Interview evidence	Final action
1	Action	No specific issue coded.	-	Retained

#	Dim.	Understanding / issue	Interview evidence	Final action
2	Action	Partly; social desirability, clarity	Worker described keeping equipment running even if long-term damage was possible.	Revised to final Item 2 wording.
3	Action	No specific issue coded.	-	Retained
4	Action	Yes; clarity, social desirability	Original stem sounded like an obvious unsafe answer rather than a pressure-related tendency.	Revised to final Item 4 wording.
5	Action	No specific issue coded.	-	Retained
6	Action	Partly; clarity	Worker emphasized knowing what to do in emergencies without referring to a procedure.	Revised to final Item 6 wording.
7	Pattern	No material problem requiring removal.	Near-redundancy with Item 8 retained for Phase 2 testing.	Retained/revised; final wording standardized.
8	Pattern	No material problem requiring removal.	Near-redundancy with Item 7 retained for Phase 2 testing.	Retained; plain-language edit.
9	Pattern	No material problem requiring removal.	-	Retained/revised; final wording standardized.
10	Pattern	Yes; clarity, social desirability	Experience-versus-checklist contrast needed a familiar-task context.	Revised to final Item 10 wording.
11	Pattern	Yes; social desirability, clarity	"Tedious" was evaluative and vulnerable to socially desirable responding.	Revised to final Item 11 wording.
12	Pattern	Yes; clarity	Workers understood complacency, but original wording was overly absolute.	Revised to final Item 12 wording.
13	Norm./Prod.	Yes; social desirability, clarity	Rule-bending wording sounded judgmental.	Revised to pressure-focused final wording.
14	Norm./Prod.	No specific issue coded.	-	Retained

#	Dim.	Understanding / issue	Interview evidence	Final action
15	Norm./Prod.	No specific issue coded.	-	Retained
16	Norm./Prod.	Partly; clarity, social desirability	Worker framed the issue as completing work without an itemized step-by-step guide.	Revised to pressure-focused final wording.
17	Norm./Prod.	No specific issue coded.	-	Retained with plain-language edit.
18	Norm./Prod.	No specific issue coded.	-	Retained
19	Adaptive	No specific issue coded.	-	Retained
20	Adaptive	No specific issue coded.	-	Retained with plain-language edit.
21	Adaptive	Yes; double-barreled	Original stem blended adapting plans with maintaining controls.	Revised to final Item 21 wording.
22	Ambig./Control	Partly; clarity	Workers heard routine independence more strongly than degraded-control tolerance.	Revised; domain remains exploratory.
23	Ambig./Control	No; comprehension, social desirability	Worker interpreted it as continuing work without being told what to do.	Removed.
24	Ambig./Control	No; comprehension, clarity	Interpreted as "I stay calm under pressure," not the intended construct.	Revised to final Item 24 wording.
25	Protective	No material problem requiring removal.	Original was absolute/ceiling prone.	Revised to final Item 25 wording.
26	Protective	No material problem requiring removal.	Original was absolute/ceiling prone.	Revised to final Item 26 wording.
27	Protective	No specific issue coded.	-	Retained with plain-language edit.
28	Protective	Yes; social desirability, clarity	"Minor" hazard reporting lacked a clear relevance threshold.	Revised to final Item 28 wording.
29	Protective	No specific issue coded.	-	Retained with clearer action anchor.
30	Protective	Yes; clarity	Workers understood the protective concept,	Revised to observable pre-task behavior.

#	Dim.	Understanding / issue	Interview evidence	Final action
31	Perf.-Shape	No specific issue coded.	but the original wording was broad and moralized. -	Retained
32	Perf.-Shape	No; clarity, comprehension	Worker described cutting corners when exhausted; fatigue needed a concrete behavioral anchor.	Revised to final Item 32 wording.
33	Perf.-Shape	Partly; social desirability, clarity	Interpreted as being better than others at avoiding injury.	Removed.
34	Perf.-Shape	No specific issue coded.	-	Retained
35	Perf.-Shape	No specific issue coded.	-	Retained
36	Perf.-Shape	Partly; clarity, social desirability	Interpreted as believing injury or vehicle crashes would not happen.	Removed.

Note. The original source contained a typographical error in a participant revision suggestion for Item 2; that error has been corrected here. Sparse entries indicate that no specific cognitive-interview problem was coded for the item. Coding was conducted by one researcher and is interpreted descriptively.

APPENDIX D

Revision Audit Trail: Original Wording, Identified Issue, and Final Wording

#	Original wording	Issue identified	Final wording	Outcome
1	I usually act quickly when something needs to be done, even if all the details are not clear.	No material wording issue identified.	I usually act quickly when something needs to be done, even if all the details are not clear.	Retained
2	When equipment is not working quite right, I try to make it work rather than stop the job.	Social desirability; vague self-repair wording	When equipment seems off, I first try to fix it myself before stopping work.	Revised
3	I am comfortable making decisions on the fly.	No material wording issue identified.	I am comfortable making decisions on the fly.	Retained
4	I dislike slowing down work to analyze potential hazards.	Judgmental phrasing; social desirability	When work is moving fast, I am less likely to slow down and check every hazard.	Revised

#	Original wording	Issue identified	Final wording	Outcome
5	I prefer solving problems by taking action rather than by planning.	No material wording issue identified.	I prefer solving problems by taking action rather than by planning.	Retained
6	In urgent situations, I rely on instinct more than procedures.	Admission-like rule-violation wording	When things get urgent, I trust my own judgment more than the rules.	Revised
7	I tend to focus more on the overall task than on small environmental details.	Abstract wording; redundancy now flagged	When I focus on the task, I can miss small changes around me.	Revised
8	I sometimes miss minor changes in my surroundings while concentrating on the job.	Plain-language simplification; redundancy now flagged	When I focus on the job, I sometimes miss small changes around me.	Revised
9	I prefer understanding the goal rather than following step-by-step instructions.	Broad wording	I would rather know the goal than follow each step.	Revised
10	I rely on experience rather than checklists.	Clarity; social desirability	When a task feels familiar, I rely more on experience than on written checklists.	Revised
11	I find detailed safety procedures tedious.	Clarity; social desirability	Detailed safety steps feel harder to use when the job feels routine.	Revised
12	I assume things are safe unless something obvious looks wrong.	Absolute/complacency framing	I assume things are safe unless something obvious looks wrong.	Revised
13	I sometimes bend rules to help my team finish work.	Judgmental rule-bending wording	I feel pressure to help the crew finish work, even when safety steps take more time.	Revised
14	I feel pressure to keep work moving even if safety rules slow things down.	No material wording issue identified.	I feel pressure to keep work moving even if safety rules slow things down.	Retained
15	I worry about how coworkers see me if I stop a job for safety reasons.	No material wording issue identified.	I worry about how coworkers see me if I stop a job for safety reasons.	Retained
16	Getting the job done is sometimes more important than following every procedure.	Judgmental self-report; pressure conflation	I feel pressure to put production ahead of safety steps.	Revised

#	Original wording	Issue identified	Final wording	Outcome
17	I feel uncomfortable challenging a supervisor about safety.	Plain-language simplification	It is hard for me to question a supervisor about safety.	Revised
18	I try to avoid causing delays, even if something feels unsafe.	No material wording issue identified.	I try to avoid causing delays, even if something feels unsafe.	Retained
19	I work best when things are changing quickly.	No material wording issue identified.	I work best when things are changing quickly.	Retained
20	I am comfortable improvising when plans change.	Plain-language simplification	I am comfortable adjusting on the fly when plans change.	Revised
21	I can adapt plans while still maintaining required safety controls.	Ability and behavior combined	When plans change, I keep required safety controls in place.	Revised
22	I do not need routines to work safely.	Routine independence conflated with ambiguity	I can work safely even when a task does not follow a familiar routine.	Revised
23	I can continue working safely even when procedures are unclear.	Low clarity; comprehension failure	—	Dropped
24	I do not get stressed when things feel chaotic.	Comprehension failure under hectic conditions	When work gets hectic, it is harder to keep every safety control in place.	Revised
25	I would stop work immediately if something seemed unsafe.	Absolute/ceiling-prone wording	When I notice an unsafe condition, I stop work before continuing.	Revised
26	I carefully follow safety procedures even when no one is watching.	Absolute/ceiling-prone wording	Even when I am not being observed, I follow the required safety steps.	Revised
27	I double-check equipment before using it.	Plain-language simplification	I check equipment twice before I use it.	Revised
28	I report hazards even if they seem minor.	Social desirability; vague reporting threshold	I report hazards, including minor ones, when they could affect the work area.	Revised
29	I verify protection or controls before beginning work.	Clearer action anchor	Before work starts, I check that protections and	Revised

#	Original wording	Issue identified	Final wording	Outcome
			controls are in place.	
30	I feel responsible for my own safety and others' safety.	Broad moral wording rather than observable behavior	Before work starts, I think about how my actions affect me and nearby coworkers.	Revised
31	I take more risks when I feel rushed.	No material wording issue identified.	I take more risks when I feel rushed.	Retained
32	I am more likely to ignore safety rules when I am tired.	Fatigue wording needed behavioral anchor	When I am tired, it is harder for me to follow detailed safety steps.	Revised
33	I believe I am better than most people at avoiding accidents.	Comparative overconfidence; social desirability	—	Dropped
34	I get frustrated when safety slows down production.	No material wording issue identified.	I get frustrated when safety slows down production.	Retained
35	I feel pressure to push through discomfort to get the job done.	No material wording issue identified.	I feel pressure to push through discomfort to get the job done.	Retained
36	I rarely think accidents will happen to me.	Optimism-bias wording; social desirability	—	Dropped

Note. Final wording is standardized to Appendix A. Editorial simplifications are included in the audit so Appendix A, Appendix C, and Appendix D use one consistent retained stem for each item.

APPENDIX E

Original 36-Item Pool

Item	Dimension	Original wording
1	Action	I usually act quickly when something needs to be done, even if all the details are not clear.
2	Action	When equipment is not working quite right, I try to make it work rather than stop the job.
3	Action	I am comfortable making decisions on the fly.
4	Action	I dislike slowing down work to analyze potential hazards.
5	Action	I prefer solving problems by taking action rather than by planning.
6	Action	In urgent situations, I rely on instinct more than procedures.
7	Pattern	I tend to focus more on the overall task than on small environmental details.
8	Pattern	I sometimes miss minor changes in my surroundings while concentrating on the job.
9	Pattern	I prefer understanding the goal rather than following step-by-step instructions.
10	Pattern	I rely on experience rather than checklists.

Item	Dimension	Original wording
11	Pattern	I find detailed safety procedures tedious.
12	Pattern	I assume things are safe unless something obvious looks wrong.
13	Norm./Prod.	I sometimes bend rules to help my team finish work.
14	Norm./Prod.	I feel pressure to keep work moving even if safety rules slow things down.
15	Norm./Prod.	I worry about how coworkers see me if I stop a job for safety reasons.
16	Norm./Prod.	Getting the job done is sometimes more important than following every procedure.
17	Norm./Prod.	I feel uncomfortable challenging a supervisor about safety.
18	Norm./Prod.	I try to avoid causing delays, even if something feels unsafe.
19	Adaptive	I work best when things are changing quickly.
20	Adaptive	I am comfortable improvising when plans change.
21	Adaptive	I can adapt plans while still maintaining required safety controls.
22	Ambig./Control	I do not need routines to work safely.
23	Ambig./Control	I can continue working safely even when procedures are unclear.
24	Ambig./Control	I do not get stressed when things feel chaotic.
25	Protective	I would stop work immediately if something seemed unsafe.
26	Protective	I carefully follow safety procedures even when no one is watching.
27	Protective	I double-check equipment before using it.
28	Protective	I report hazards even if they seem minor.
29	Protective	I verify protection or controls before beginning work.
30	Protective	I feel responsible for my own safety and others' safety.
31	Perf.-Shape	I take more risks when I feel rushed.
32	Perf.-Shape	I am more likely to ignore safety rules when I am tired.
33	Perf.-Shape	I believe I am better than most people at avoiding accidents.
34	Perf.-Shape	I get frustrated when safety slows down production.
35	Perf.-Shape	I feel pressure to push through discomfort to get the job done.
36	Perf.-Shape	I rarely think accidents will happen to me.

Note. The original pool is reproduced to preserve traceability from initial generation through revision and item-removal decisions.

APPENDIX F

Provisional Scoring Guidance for the Revised 33-Item Pool

The revised 33-item pool remains developmental. If scores are computed for research during Phase 2, dimension scores may be calculated as mean responses within each retained dimension. No total score, weighted composite, risk band, operational cutoff, or employment classification is recommended. The two-item Tolerance for Ambiguity and Degraded Control score is especially provisional and should not be treated as a stable scale unless subsequent evidence supports it.

Dimension	Items	Provisional score	Interpretation of higher scores
Action Orientation	1-6	Mean of items 1-6	Greater tendency to initiate action before full review or scanning.

Dimension	Items	Provisional score	Interpretation of higher scores
Pattern Processing	7-12	Mean of items 7-12	Greater tendency to rely on broad task patterns rather than smaller hazard cues.
Normative and Production Pressure Sensitivity	13-18	Mean of items 13-18	Greater sensitivity to coworker, supervisory, or production pressure.
Adaptive Flexibility	19-21	Mean of items 19-21	Greater tendency to adjust plans while trying to maintain controls.
Tolerance for Ambiguity and Degraded Control	22, 24	Mean of items 22 and 24	Greater willingness to continue working when routine or control conditions feel less stable.
Protective Orientation	25-30	Mean of items 25-30	Stronger endorsement of active, observable safety-regulating behaviors.
Susceptibility to Performance-Shaping Factors	31, 32, 34, 35	Mean of items 31, 32, 34, and 35	Greater susceptibility to rushing, fatigue, frustration, or discomfort-related degradation in safe behavior.

Note. Higher scores on Protective Orientation reflect stronger endorsement of active safety-regulating behavior. Higher scores on the other dimensions reflect greater endorsement of the targeted response tendency or vulnerability. These directions are conceptual only until Phase 2 supports the proposed dimensional structure.