Journal of Management
Vol. XX No. X, Month XXXX 1–27
DOI: 10.1177/0149206317745596

The Development and Validation of a Cross-Industry Safety Climate Measure: Resolving Conceptual and Operational Issues

Jeremy M. Beus
Louisiana State University
Stephanie C. Payne
Texas A&M University
Winfred Arthur Jr.
Texas A&M University
Gonzalo J. Muñoz
Universidad Adolfo Ibáñez

Although safety climate research has increased in recent years, persisting conceptual ambiguity not only raises questions about what safety climate really is—as operationalized in the literature—but also inhibits increased scientific understanding of the construct. Consequently, using climate theory and research as a conceptual basis, we inductively articulated safety climate's general content domain by identifying seven core indicators of safety's perceived workplace priority: leader safety commitment, safety communication, safety training, coworker safety practices, safety equipment and housekeeping, safety involvement, and safety rewards. These indicators formed the basis for a generalized safety climate measure that we designed for use across organizations, industries, and construct levels. We then conducted a multilevel construct

Acknowledgments: We thank the review team for its guidance and patience during the review process, and we gratefully acknowledge the financial support received from the Mary Kay O'Connor Process Safety Center at Texas A&M University, National Institute of Occupational Safety and Health Region VI Pilot Projects Research Training Program (RF 1001569), Texas A&M University College of Liberal Arts Cornerstone Faculty Fellowship, FONDECYT Chile (grant #11140488), as well as the Rucks Department of Management Development Fund at Louisiana State University. Earlier versions of this manuscript were presented at the 2011 Society for Industrial and Organizational Psychology annual conference and the 2013 Proceedings of the Academy of Management *(CD), ISSN 1543-8643.*

Supplemental material for this article is available with the manuscript on the JOM *website.*

Corresponding author: Jeremy M. Beus, Rucks Department of Management, E. J. Ourso College of Business, Louisiana State University, 2713 Business Education Complex, Baton Rouge, LA 70803, USA.

E-mail: jbeus@lsu.edu

validation of safety climate using the newly created measure in two separate studies. Results from five samples spanning multiple organizations, industries, and cultural settings revealed that the identified safety climate indicators were parsimoniously explained by an overarching safety climate factor at the individual and workgroup levels. In addition, multilevel homology tests indicated that safety climate's associations with past safety incidents were nearly two times stronger at the workgroup level relative to the individual level, although this difference was not statistically significant. Finally, workgroup-level validity evidence demonstrated expected associations between safety climate and organization-reported pre- and postsurvey safety incidents. On the basis of this supportive evidence, we recommend that this conceptualization and measure of safety climate be adopted in research and practice to facilitate future scientific progress.

Keywords: *safety climate; safety culture; organizational climate; multilevel construct validation; measure development*

> I often say that when you can measure what you are speaking about . . . you know something about it; but when you cannot measure it . . . your knowledge is of a meager and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely . . . advanced to the stage of science, whatever the matter may be.
>
> —Lord Kelvin (19th-century physicist)

Safety climate—employees' shared perceptions of the importance and prioritization of workplace safety (Zohar, 2011)—has received considerable empirical attention across multiple disciplines. This is not surprising, however, when one considers the demonstrated practical importance of safety climate. For instance, high-profile safety incidents resulting in numerous fatalities and substantial ecological damage have been attributed to poor organizational safety climates (e.g., BP Texas City explosion, Deepwater Horizon oil spill; Kunzelman, 2011; Rodriguez, Payne, Bergman, & Beus, 2011). Furthermore, meta-analytic evidence has consistently shown that safety climate is positively associated with workplace safety behaviors and negatively associated with safety incidents across levels of analysis (Beus, Payne, Bergman, & Arthur, 2010; Christian, Bradley, Wallace, & Burke, 2009; Clarke, 2006).

However, despite safety climate's practical relevance and increasing prominence in the eyes of researchers and practitioners, its conceptualization and measurement have suffered from a lack of clarity, as well as a widespread failure to consider the multilevel nature of the construct (Beus et al., 2010; Shannon & Norman, 2009). This stunts scientific progress because it is unclear whether different safety climate researchers are actually studying the same phenomenon and thus whether common inferences can be drawn from their findings.

Although numerous safety climate measures have been developed (e.g., Dedobbeleer & Beland, 1991; Williamson, Feyer, Cairns, & Biancotti, 1997; Zohar, 1980), many of these measures suffer from either limited applicability beyond the examined samples or a combination of content deficiency and contamination (Beus et al., 2010). That is, many safety climate measures contain elements that, although related to workplace safety, are not consistent with the definition of safety climate. For example, perceptions of job risk are included in some safety climate measures (e.g., Flin, Mearns, O'Connor, & Bryden, 2000; Guldenmund, 2000), with higher job risk indicating less favorable safety climates. However, a job's riskiness says little about whether a prevailing safety climate will be favorable or unfavorable.

The inclusion of irrelevant content such as job risk in safety climate measures biases effect size estimates and leads to inaccurate conclusions regarding the magnitude and meaning of safety climate's associations with other relevant constructs (Beus et al., 2010).

In addition, although safety climate is widely recognized as a multilevel construct (Zohar, 2011; Zohar & Luria, 2005), a limitation of past safety climate construct validations is that they have not examined the group-level properties of data in addition—and in comparison—to the individual level (Shannon & Norman, 2009). Thus, it is unknown if aggregated responses to a given safety climate measure maintain the same properties as responses at the individual level or how empirical associations with relevant constructs compare across levels of analysis. Consequently, using extant climate theory and research, we sought to (a) clarify safety climate's content domain, (b) develop a conceptually representative cross-industry safety climate measure, and (c) conduct a multilevel construct validation of safety climate with the new measure. The purpose of these steps is to provide a better understanding of what safety climate is and how it operates across levels of analysis to facilitate future research advancements.

To accomplish these objectives, we followed a sequence of interdependent steps outlined by Chen, Mathieu, and Bliese (2004) to conduct a multilevel construct validation of safety climate. These steps, based on the unitarian view of validity (e.g., Binning & Barrett, 1989; Messick, 1995), expand on accepted construct validation principles and practices by applying them to multilevel constructs. On the basis of this framework, we first describe the nature of safety climate across construct levels and the subsequent development of a generalizable, cross-industry measure consisting of the following core indicators of safety's workplace priority: leader safety commitment, safety communication, safety training, coworker safety practices, safety equipment and housekeeping, safety involvement, and safety rewards. Then, utilizing data from five samples, we report construct-related validity evidence for the measure in two studies. The first study reports individual-level validity evidence in two samples to reduce the measure to a generalizable and practical set of items and to establish the psychometric properties of measure responses. In the second study, we use the measure to provide multilevel validity evidence for safety climate in three divergent organizational contexts by confirming the construct's group-level existence, establishing its group-level factor structure, and estimating its relationships with relevant constructs across levels of analysis.

A Multilevel Construct Validation of Safety Climate

Definition and Conceptualization of Safety Climate Across Construct Levels

The first step for multilevel construct validation involves defining the construct (Chen et al., 2004). For a multilevel construct, this requires defining the construct at each applicable level of analysis and delineating the aggregate nature of the construct and the means through which aggregate assessments are made.

Fundamentally, climate represents socially constructed perceptions of the meaning of work environments (Ashforth, 1985; Schneider, 1975; Schneider & Reichers, 1983). Individuals are driven to form these perceptions out of a desire to establish order and subsequently adapt to their social environments (Schneider, 1975). Because meaning can be attached to numerous facets of organizational life, Schneider and Reichers (1983) contended that climates must be examined as they pertain to something to be practically meaningful (e.g., safety, customer service). Although numerous definitions of safety climate abound (Flin et al., 2000), we adopted Zohar's (2011) definition because of its direct correspondence with the broader climate literature. Zohar's definition builds off Schneider and Reichers's climate definition by describing climates for safety as shared employee perceptions of the priority or importance placed specifically on workplace safety.

As a domain-specific manifestation of climate, safety climate is inherently a multilevel construct (Zohar, 2011). It is based in individual perceptions and emerges at aggregate levels when social interaction and contextual clarity create sharedness in group members' perceptions of workplace priorities (Ashforth, 1985; Schneider & Reichers, 1983). At the individual level, safety climate represents an individual's interpretation of safety's workplace importance, whereas at aggregate levels, safety climate reflects shared perceptions of safety's importance in a group. Theory and empirical evidence suggest that safety climates can vary across organization levels given, for instance, divergent safety priorities enacted by direct supervisors relative to upper management (Zohar, 2000; Zohar & Luria, 2005). In addition, meta-analytic evidence has demonstrated differences in the associations between safety climate and safety-related outcomes at individual and group levels (Beus et al., 2010; Christian et al., 2009), suggesting that different processes may be at work across organization levels.

These cross-level differences may be the result of group-level contextual influences that are not reliably captured by the construct's individual-level manifestation (Bliese, 2000; Bliese, Chan, & Ployhart, 2007; Morgeson & Hofmann, 1999). Whereas an individual's perception of the group's prioritization of safety will be subject to his or her personal idiosyncrasies, the average of group members' perceptions of safety's priority should represent a more accurate reflection of the group's safety climate that better captures the summation of contextual influences (Bliese, 2000). However, despite evidence of such cross-level differences, we are unaware of past studies that tested the psychometric properties of a content valid safety climate measure across levels of analysis (i.e., at the individual and aggregate levels). The failure of past studies to consider multiple construct levels implies the untested assumption of isomorphism (i.e., identical construct meaning and manifestation across levels), which is an assumption that scholars argue few constructs meet (Bliese, 2000; Kozlowski & Klein, 2000).

Because climate is a perceptual construct, it is assessed at the individual level—the locus of perceptions—and then aggregated to form higher-level construct manifestations (Schneider & Reichers, 1983). This requires specifying a composition model to articulate the functional relationships of individual and group-level safety climate (Chan, 1998; van Mierlo, Vermunt, & Rutte, 2009). Based on Chan's (1998) typology of composition models, safety climate, like any manifestation of climate, is best characterized by a referent-shift composition model (Schneider, Ehrhart, & Macey, 2011). A referent-shift model presupposes high levels of agreement, or consensus, among individuals to justify aggregation to a higher construct level (Chan, 1998; Chen et al., 2004). Importantly, the lower-level manifestation of the construct for a referent-shift model—in this case, individual safety climate perceptions—maintains a group-level referent. Thus, although it is an individual's perception, it is the individual's perception of the value that is placed on safety at a collective level. Keeping the group as the referent for climate items is critical given that the true theoretical referent for any type of climate is the aggregate (Glick, 1985; Schneider et al., 2011).

We focused on the workgroup as the aggregate level of particular interest for this study. We did so because individuals interact most often with their immediate supervisor and fellow workgroup members, making the workgroup the most proximal and salient social unit in the organization (Ashforth, 1985) and a setting that is especially likely to foster the creation of climates (Powell & Butterfield, 1978). Research has demonstrated meaningful variability in workgroup-level safety climates within organizations (Zohar, 2000; Zohar & Luria, 2004, 2005), further supporting the value of considering safety climate at this particular level.

Safety Climate's Content Domain

Although safety climate's content domain has yet to be clearly specified, climate theory indicates that a variety of workplace factors can contribute to employees' perceptions of workplace priorities (Ashforth, 1985; Schneider & Reichers, 1983). In fact, any workplace policies, procedures, or practices with implications for workplace safety can conceivably inform individual climate perceptions and subsequent group safety climates. However, for safety climate research and theory to advance, we assert that a single generalizable measure that can be used to assess safety climate across contexts is necessary to move safety climate research from a repeated focus on measurement to an increased emphasis on more substantive issues. The key, then, is to determine which indicators that can be identified across contexts best reflect the value of safety and can representatively establish safety climate's content domain in a single measure.

To help specify safety climate's content domain, Zohar (2011) emphasized that measure content should maintain an overarching focus on safety's perceived priority. This is consistent with the climate literature in general, given that individuals are expected to use contextual cues to inform their perceptions of workplace priorities and the values on which those priorities are based (Schneider, 1975; Zohar & Hofmann, 2012). In the context of workplace safety, a particularly meaningful indicator of safety's priority is leadership's perceived commitment to safety (Hofmann & Stetzer, 1996; Zohar, 2002; Zohar & Polachek, 2014). Leader actions that reflect a commitment to safety (or lack thereof) are informative to individual safety climate perceptions and subsequent group safety climates because safety is most often a leader-driven objective (Barling, Loughlin, & Kelloway, 2002). Absent situations where one's safety or the safety of coworkers will clearly be jeopardized by failing to work safely, workers often choose to discount time- or effort-intensive safety procedures in favor of getting the job done, which is typically that to which the most meaningful rewards are connected (e.g., pay). Thus, it generally falls on organizational leaders to emphasize and consistently reinforce safety as a group priority. Climate research in general underscores the importance of leaders as purveyors of climate, given that (a) climates are a function of the behaviors that employees perceive are rewarded and reinforced (Schneider, 1975; Schneider & Reichers, 1983) and (b) leaders most often administer those rewards (Bass, 1985; Podsakoff, Bommer, Podsakoff, & MacKenzie, 2006). Taken together, one of safety climate's chief indicators should be perceptions of leader commitment to safety. With higher perceived leader commitment, employees are more likely to perceive a positive safety climate in which safety holds a high priority.

A focus on perceived leader safety commitment as a chief safety climate indicator is consistent with a wide and growing body of safety climate research (Flin et al., 2000; Hofmann & Stetzer, 1996; Zohar & Luria, 2004). Zohar (2011) encouraged incorporating content into safety climate measures that is directly and indirectly indicative of leader safety commitment—directly in terms of items that explicitly assess perceptions of leader safety commitment and indirectly through generalized indicators such as the availability of safety equipment and the communication of safety information. Consequently, the majority of existing safety climate measures place either direct or indirect emphasis on leader safety commitment.

In light of these conceptual considerations, we sought to articulate safety climate's core, foundational content domain to address the construct's extant conceptual ambiguity. However, rather than specify which factors best reflect leader safety commitment and safety's corresponding priority based on our knowledge of the construct, we took an inductive approach that incorporated the broader safety climate literature. Specifically, we amassed >1,500 existing safety climate items from 30 years of published safety climate research in varied literatures and across numerous industry settings to guide the determination of safety climate's core indicators and ultimately facilitate the creation of a generalized safety climate measure. This approach was necessary given our purpose of creating a measure that is representative of the safety climate content domain and applicable across industries and settings.

Development of a Cross-Industry Safety Climate Measure

To develop a cross-industry measure of safety climate that articulates a generalizable set of indicators, we first attempted to gather every unique nonproprietary instrument that has been used in the published literature to assess safety climate. To locate these measures, we searched PsycINFO using the keywords "safety climate." In cases where items were not reported in publications, we contacted authors directly for them. Ultimately, this search resulted in 165 studies, from which we were able to obtain 62 of 67 (93%) distinct measures of safety climate. Pooling these measures resulted in a total of 1,504 items.

Second, three safety climate subject matter experts (SMEs; this study's first three authors) evaluated whether each item corresponded to the described theoretical conceptualization of safety climate. For example, we removed items referencing the individual as opposed to the group (e.g., "I regularly contribute to our safety efforts") or personal beliefs instead of perceptions of existing conditions (e.g., "Safety is not the responsibility of only one individual worker"). Other removed items were those that were not directly related to safety ("My supervisor checks on my work very carefully") or that dealt with the nature of the job and not safety's perceived prioritization (e.g., "My job is hazardous"). Through this process we eliminated 570 items. In addition, on the basis of best practices in item writing and measure construction (e.g., Crocker & Algina, 2006; Haladyna, 2004), we removed items that were exact or approximate duplicates of others (428 items), as well as poorly written items (342 items; e.g., double-barreled or ambiguous: "Our management is well-informed about safety problems and it acts quickly to correct them"). Furthermore, we either revised or removed items that were too industry specific for broad applicability (32 items removed; e.g., "My practice setting has clear procedures on what to do when patients arrive with symptoms of respiratory infection"). The SMEs reached consensus on each item's designation, with the process resulting in a pool of 133 unique and conceptually appropriate safety climate items.[1]

Once SMEs agreed on the 133 safety climate items, they familiarized themselves with the items and independently identified safety climate indicator categories that they believed encapsulated the entire item pool. After doing so, the three SMEs met in a series of meetings to reach consensus on (a) what the emergent safety climate indicators should be and (b) where each item should be categorized, with each item being assigned to only one indicator category. The SMEs identified seven indicators that they judged to be sufficiently representative of the pool of safety climate items. These indicators are as follows: leader safety commitment, safety communication, safety training, coworker safety practices, safety equipment and housekeeping, safety involvement, and safety rewards. Descriptions for each of these seven safety climate indicators are provided in the appendix. These categories reflect the direct and indirect indicators of leader commitment to safety and safety's workplace priority that have been assessed most frequently across varying industries and settings and are consistent with the conceptualization of safety climate adopted here.

Due to the practical limitations of administering a 133-item measure in organizational settings, we next gathered data and used factor-analytic techniques to reduce the number of items to a manageable set and determine the most appropriate factor structure to represent safety climate. Because our purpose was to develop a measure that can be used across industry settings, we first gathered responses from employees of multiple industries (Sample 1) to determine safety climate's factor structure; we then confirmed those findings by administering an empirically reduced measure in a single safety-salient context (Sample 2). Analyses based on these two samples were restricted solely to the individual level for initial construct validation. However, we addressed this issue in Study 2 with three organizational samples with workgroup-level data that allowed us to fulfill the remaining steps for multilevel construct validation.

Study 1: Method

Safety Climate Measure

In Sample 1, we assessed safety climate online with 133 items measuring the seven noted safety climate indicators. All items were rated on a 5-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*) and were altered where necessary to have a workgroup-level referent. Respondents were instructed to think of their current supervisor and workgroup when responding to the items. This was primed at the start of the survey by asking respondents to report the number of people in their workgroup (i.e., all those who report to the same supervisor) before responding to the safety climate items.

In Sample 2, safety climate was assessed via paper and pencil with a reduced set of 30 items that were empirically selected per the findings from Sample 1. We describe this empirical selection process in the Results and Discussion section. These 30 items are reported in the appendix and maintained representation of the seven identified safety climate indicators.

Participants and Procedure

The safety climate measure was first administered online to Sample 1, an international sample of employed respondents selected from 10 occupational categories.[2] Of the 1,505 individuals who completed the measure, 60% were male with the majority (83%) being U.S. residents. Most respondents worked in the manufacturing industry (37%), with an average job tenure of 7.27 years (*SD* = 4.88) and an average age of 36.76 years (*SD* = 7.81). All participants were recruited through StudyResponse (Stanton & Weiss, 2002), a nonprofit academic service pairing researchers with individuals willing to serve as research participants. Participants were required to be employed at least part-time and could not be self-employed. Individuals who satisfactorily completed the measures received a $10 gift card to Amazon.com as compensation for their participation.

Because of the large number of items in the measure, we embedded five check items to use as a method of screening out nonserious responders (cf. Huang, Curran, Keeney, Poposki, & DeShon, 2012). These items were written such that serious respondents who were paying attention to the items should have disagreed—for example, "My supervisor insists that we vandalize company property." As a predetermined decision rule, we excluded respondents who failed to disagree to three or more of these items; 562 individuals (37%) were retained as serious respondents in Sample 1. Because of the large number of nonserious respondents, we report detailed demographic characteristics for this sample in the online supplement for comparative purposes (see Table S1).

To avoid capitalizing on sample-specific idiosyncrasies when conducting initial confirmatory factor analyses (CFAs), we cross-validated our findings (Crocker & Algina, 2006) by randomly splitting Sample 1 into two subsamples[3] (developmental sample, n = 333; cross-validation sample n = 229). Specifically, incorporating all 133 safety climate items, we first used the developmental sample to identify the most representative safety climate factor structure as well as the highest-loading items that represented that factor structure. Then, we used the cross-validation sample to confirm the factor structure identified in the developmental sample with the reduced set of highest-loading items. Table S1 in the online supplement compares these two subsamples and shows that both were similar in demographic composition.

We used Sample 2 to substantiate the factor structure obtained in Sample 1's two subsamples. Specifically, we estimated the model's fit based on data from a single safety-salient organizational setting. Sample 2 consisted of 595 production employees of a large defense company. The majority of respondents were male (69%), with an average age of 41.97 years (SD = 12.86) and an average organizational tenure of 7.55 years (SD = 9.21).

Study 1: Results and Discussion

Descriptive statistics and variable intercorrelations for Samples 1 and 2 are reported in the online supplement (see Table S2). We conducted initial CFAs by using all 133 safety climate items in Sample 1's developmental subsample. We used CFA to evaluate a primary model based on climate theory and two plausible alternative models that could also be reflective of the seven identified safety climate indicators. First, we tested the fit of a second-order safety climate model (the primary model) consisting of the seven identified first-order factors or indicators and an overarching second-order factor. A second-order factor structure is consistent with expectations based on climate theory that the identified indicators of safety's workplace priority will be reflective of a latent overarching safety climate construct. Then, we tested two alternative safety climate models: a single-factor model and a first-order model with seven factors corresponding to the seven identified safety climate dimensions.

The proposed second-order model fit the data well in the developmental subsample ($\chi^2_{[8,638]}$ = 19,887.10, p < .05; comparative fit index [CFI] = .76; standardized root mean square residual [SRMR] = .05; root mean square error of approximation [RMSEA] = .06), as indicated by RMSEA and SRMR values. Although the CFI value is below standards for acceptable fit (Hu & Bentler, 1999), the large number of items in this initial analysis precluded the model from receiving a more acceptable CFI estimate, which penalizes models for complexity. Of the three safety climate models that we tested in the developmental subsample, the single-factor model was the poorest fit to the data ($\chi^2_{[8,645]}$ = 22,471.25, $p < .05$; CFI = .71; SRMR = .05; RMSEA = .07), whereas the seven-factor model had approximately identical fit to the proposed second-order model ($\chi^2_{[8,624]}$ = 19,784.99, $p < .05$; CFI = .76; SRMR = .05; RMSEA = .06). However, because a second-order factor structure is more consistent with climate theory and can explain the high intercorrelations among first-order safety climate factors, we deemed the second-order model to better represent the safety climate construct. Consequently, we adopted the second-order model for the remaining analyses.

Next, to reduce the number of items, we used the CFA results from the second-order model in the developmental subsample to identify the four highest-loading items in each of the seven first-order factors, with the exception of leader safety commitment. Because of the noted centrality of leader safety commitment as a direct indicator of the priority of safety, we retained the six highest-loading items for this factor to give it greater content representation. The selection of at least four items per dimension is consistent with best-practice recommendations in measure development (Harvey, Billings, & Nilan, 1985; Hinkin, 1998). We believe that this number of items per dimension is also sufficient to provide adequate content representation for each indicator. These steps resulted in a reduced 30-item safety climate measure, which we report in the appendix. We then reestimated the second-order model in the developmental subsample by using the 30-item measure and found a noteworthy improvement in fit relative to the full measure, particularly with its CFI value ($\chi^2_{[398]}$ = 840.74, $p < .05$; CFI = .95; SRMR = .04; RMSEA = .06). We then conducted a multigroup CFA that confirmed the factorial invariance of the second-order factor structure across the developmental and cross-validation subsamples of Sample 1 ($\chi^2_{[826]}$ = 1,641.27, $p < .05$; CFI = .95; SRMR = .05; RMSEA = .06).

To establish the relevance of these findings in a single safety-salient organizational setting, we administered the reduced 30-item measure to Sample 2 and tested to see if the specified second-order factor structure was a good fit to the data. Consistent with the results of Sample 1's two subsamples, the estimated factor structure fit Sample 2's data well ($\chi^2_{[398]}$ = 1,206.66, $p < .05$; CFI = .94; SRMR = .05; RMSEA = .06). Taken together, Samples 1 and 2 provide strong support for safety climate's second-order factor structure at the individual level of analysis by demonstrating converging results from a heterogeneous sample of respondents from multiple industries (Sample 1) and a large organizational sample of production workers (Sample 2). The confirmed factor structure is noteworthy because not only does it articulate the lower-order indicators that best represent safety climate's content domain, but it also confirms the existence of a second-order factor that explains common variance across these indicators.

As an additional check of the seven identified safety climate indicators, we solicited the feedback of several well-published international safety scholars (e.g., Julian Barling, Michael Burke, Mark Griffin, Dov Zohar) to help confirm the pairing of our 30 items with the specified factors. Specifically, 11 scholars with an average of 16 years of safety research experience (*SD* = 6.74 years) completed a survey in which they categorized each safety climate item (30 total) with the one dimension that they deemed most representative of that item. Across the 11 scholars, there was 83% agreement with their item categorizations and our own. The chief source of disagreement came from categorizing some items as indicators of "leader safety commitment" instead of categorizing them into the more specific intended categories, such as "safety involvement" or "safety rewards." In open-ended comments, more than one SME indicated that some items could be reflective of, say, safety involvement and leader safety commitment. Given our contention that the factors outside of leader commitment to safety are ultimately indirect reflections of leadership's safety commitment—and safety's overarching priority—in combination with empirical evidence confirming the same with a common second-order factor, we believe that this level of agreement is satisfactory. In fact, if categorizing items as indicators of leader safety commitment is considered correct across the board, the safety scholars' agreement levels with our categorizations increase from 83% to 93%. This provides additional evidence to support the safety climate indicators confirmed in Samples 1 and 2.

Establishing safety climate's core indicators and individual-level factor structure is only a first step in the validation process. To adequately test the validity of a multilevel construct rooted in individual perceptions (Zohar, 2011), data are needed from samples at individual and aggregate levels. Consequently, Study 2's purpose was to address the remaining steps identified by Chen et al. (2004) for conducting multilevel construct validations. Specifically, we (a) tested for sufficient within-group agreement and construct variability among aggregate units (i.e., workgroups) to verify safety climate's workgroup-level existence, (b) assessed safety climate's psychometric properties (i.e., factor structure, reliability) across construct levels, and (c) examined safety climate's associations with theoretically relevant constructs across levels of analysis to extend the results of Study 1. We fulfilled this final step in two parts: first in a homology test that evaluated safety climate's associations with self-reported safety incidents at the individual and workgroup levels; second, in a workgroup-level analysis of safety climate's predictive validity that considered relationships with theoretically relevant constructs, including organization-reported safety incidents and injuries that were reported in the 6-month periods before and after the safety climate assessment.

With regard to our homology tests, we propose that a proportional (rather than an identical or metaphoric) theory of homology most accurately describes safety climate's relationships with safety incidents across construct levels. Proportional theories of homology posit that *X-Y* relationships will be proportionately stronger or weaker across levels of analysis (Chen, Bliese, & Mathieu, 2005). On the basis of previous theory and empirical evidence (Beus et al., 2010; Christian et al., 2009), we posit that the associations between safety climate and safety incidents (e.g., injuries) are proportionately stronger at aggregate levels relative to the individual level; this expectation is rooted in the premise that safety climate's theoretical referent is the aggregate where, correspondingly, its strongest and most meaningful associations should reside. Although individuals' recollections of safety incidents may reveal idiosyncratic effects on their safety climate perceptions that can be colored by individual differences or personal experiences, aggregating across individuals to the workgroup level (given sufficient within-group agreement) should provide a more reliable estimate of the influence of safety incidents on workgroup safety climate that captures meaningful group characteristics as well as the clustering of individual differences within groups (Bliese, 1998; Bliese et al., 2007). We expect the confluence of these factors to result in proportionately stronger workgroup-level associations between safety incidents and safety climate relative to the individual level.

Study 2: Method

Multilevel Organizational Samples

To fulfill the remaining steps for multilevel construct validation, we administered the 30-item safety climate measure to employees in three organizational samples: Samples 3–5. We administered the measures in Samples 3 and 4 via paper and pencil and in Sample 5 via an online questionnaire. To be included in this study's multilevel analyses, respondents from each sample had to identify their workgroups, and at least two employees from each workgroup had to provide usable responses.

Sample 3 consisted of 634 employees from 90 workgroups in a Chilean mining company. Utilizing the noted inclusion criterion, we retained 547 employees (86%) embedded within 75 workgroups (83%) for analyses (*M* = 7.29 employees per workgroup, *SD* = 5.19). Of the retained respondents, 91% were male with an average age of 36.33 years (*SD* = 10.20) and an average organizational tenure of 3.88 years (*SD* = 3.47). For this sample, the measures were first translated from English to Spanish by two psychology doctoral students fluent in both languages, using accepted translation-back-translation techniques (Brislin, 1970).

Sample 4 consisted of 228 English-speaking contractors at a U.S. petrochemical refinery. We retained 195 respondents (86%) from 47 workgroups (*M* = 4.15 employees per workgroup, *SD* = 2.64) for analyses. Of the retained respondents, 95% were male with an average age of 36.86 years (*SD* = 11.94) and an average organizational tenure of 10.62 years (*SD* = 9.13).

Finally, Sample 5 consisted of employees of a large U.S. petrochemical company; 535 of 705 (76%) respondents provided complete, usable responses, and we were able to retain 504 employees from 64 workgroups (*M* = 7.98 employees per workgroup, *SD* = 7.38) for analyses. Respondents reported an average organizational tenure of 16.66 years (*SD* = 11.26). The organization did not permit demographic questions about sex or age in the questionnaire.

To ensure that we had sufficient statistical power for our analyses—particularly at the workgroup level—it was advantageous to combine the three organizational samples into a single sample. To substantiate this decision, we conducted a multigroup CFA to determine whether responses were factorially equivalent across samples. In support of the second-order factor structure established in Study 1, this analysis confirmed equivalent factor structures across the three samples ($\chi^2_{[1,254]} = 4{,}195.02$, $p < .05$; CFI = .92; SRMR = .06; RMSEA = .08); that is, constraining safety climate responses across these samples to be represented by a second-order factor structure with seven first-order factors was a good fit to the data. Consequently, we proceeded to combine Samples 3–5 for the purposes of examining the psychometric properties of measure responses across levels of analysis and for testing a proportional theory of homology. The combined sample consisted of 1,249 employees embedded in 186 workgroups.

Measures

Safety climate. We assessed safety climate in Samples 3–5 via the reduced 30-item measure. We also assessed safety climate by using a reputable 10-item measure from Zohar (2000) in Sample 5 only. We did this for comparative purposes, which we elaborate on later.

Self-reported safety incidents. Respondents in each sample were asked to report the number of injuries and noninjurious incidents (e.g., fires, equipment damage) that occurred in their workgroups over the preceding 6 months. These represent theoretical antecedents of safety climate given that they occurred prior to safety climate assessment (Beus et al., 2010). Based on previous meta-analytic evidence, these variables should demonstrate small to moderate negative associations with safety climate at the individual and group levels of analysis (Beus et al., 2010; Christian et al., 2009).

Self-reported workgroup safety behavior. We assessed self-reported workgroup safety behavior in each sample via Griffin and Neal's (2000) measures of safety behaviors—specifically, required rule-following safety-related behaviors (safety compliance; 4 items) and nonprescribed discretionary safety-related behaviors that contribute to workplace safety (safety participation; 4 items). To facilitate group-level analyses, we altered the referent of these items to reflect the workgroup (e.g., "My workgroup carries out work in a safe manner"). Meta-analytic evidence suggests moderate to strong positive associations between safety climate and safety behavior across levels of analysis (Christian et al., 2009).

Organization-reported safety incidents. In Sample 5 only, we were given access to organization records of workgroup-level safety-related events that occurred in the 6 months before (435 reported events) and after (601 reported events) our safety climate assessment. We excluded events that were attributed primarily to contractors, as contractors were not invited to participate in the safety climate assessment. This reduced the number of events to 328 before the assessment and 432 after the assessment. The types of events reported in the database included worker injuries, fires, chemical/gas leaks and spills, automobile accidents, observed safety rule violations, equipment failures, and general process abnormalities. Because of inconsistent organizational labeling of reported incidents, the first two authors independently coded each event based on the description provided and then arrived at consensus regarding event designations. The events of particular interest for this study—and where there were sufficient incidence rates and between-group variability for workgroup-level analyses—were injuries and generalized safety incidents. We operationalized injuries as reports of physical harm to workers, ranging from minor injuries requiring only basic first-aid treatment (e.g., cuts, heat exhaustion) to more serious injuries that resulted in time off (e.g., severe burns, broken bones).

We included all relevant reported safety-related events in our operationalization of safety incidents (e.g., injuries, leaks, spills, automobile accidents, process abnormalities). Exceptions to this were reported events that were beyond the control of organizational employees (e.g., weather events, security issues involving nonemployee third parties), which are therefore less likely to affect or be affected by workgroup safety climates.

Study 2: Results and Discussion

Safety Climate's Factor Structure Across Levels of Analysis

A critical step in multilevel construct validation is testing the psychometric properties of measure responses at each construct level of interest (Chen et al., 2004). However, before conducting these analyses in our combined sample, we tested for sufficient between-group variability (intraclass correlation coefficients: ICC[1], ICC[2]) and within-group agreement ($r_{wg[j]}$) in safety climate perceptions to determine whether safety climate could meaningfully be considered to represent a workgroup-level construct across these contexts (Bliese, 2000; Chen et al., 2005). We estimated ICC(1) and ICC(2) values to verify that scale responses showed sufficient variability among groups as well as adequate mean stability. These estimates were supportive in both cases, demonstrating that meaningful proportions of safety climate variability were explained by group membership (ICC[1] = .22) and that workgroup means represented fairly stable estimates (ICC[2] = .65) based on the relatively small sizes of the examined workgroups.

Likewise, tests of within-group agreement revealed that workgroup members generally shared corresponding perceptions of workgroup safety climates. We calculated two $r_{wg(j)}$ estimates through moderately skewed (median $r_{wg[j]}$ = .97) and uniform null (median $r_{wg[j]}$ = .99) response distributions since these realistically demonstrate the lower and upper bounds of within-group agreement levels across these samples' workgroups. Uniform response distributions assume that individuals within a workgroup are equally likely to endorse each response option, whereas a moderately skewed response distribution assumes negative skew, or a greater likelihood of responding favorably to items (LeBreton & Senter, 2008). Such a distribution is theoretically plausible in these contexts given the possibility that respondents preferred not to rate their leaders' or coworkers' commitment to safety poorly. Median $r_{wg[j]}$ values across workgroups were statistically significant whether based on critical values (see Smith-Crowe, Burke, Cohen, & Doveh, 2014) for uniform distributions ($r_{wg[j]} \geq .78$) or moderately skewed distributions ($r_{wg[j]} \geq .87$). These indices, in combination with ICC(1) and ICC(2) estimates, provide evidence to suggest that safety climate exists at the workgroup level across these three contexts. We also note that internal consistency estimates of safety climate responses at the individual ($\alpha = .97$) and workgroup ($\alpha = .98$) levels were satisfactory for the combined sample.

Given this supporting evidence, we conducted CFAs at the individual and workgroup levels of analysis to determine whether safety climate's established second-order factor structure was a good fit across construct levels. As already indicated by our multigroup CFA comparing the three combined samples, a second-order factor structure with seven first-order factors was a good fit to individual-level data after accounting for nonindependence due to workgroup nesting ($\chi^2_{[398]}$ = 1,257.96, $p < .05$; CFI = .96; RMSEA = .04; SRMR = .04). We then tested the fit of the second-order model at the workgroup level of analysis. This analysis revealed that mean workgroup-level responses were likewise an acceptable fit to the second-order factor structure established in Study 1 ($\chi^2_{[398]}$ = 1,034.75, $p < .05$; CFI = .90; RMSEA = .09; SRMR = .05).

Taken together, the findings from these three combined organizational samples confirm the workgroup-level emergence of safety climate and establish safety climate's workgroup-level factor structure, providing evidence that safety climate is conceptually similar across individual and workgroup levels. The correspondence of findings from these three samples further supports the generality of this measure in light of differences in language, industry, and national culture.

Testing a Proportional Theory of Homology

To examine the proposed proportional theory of homology for safety climate and relevant safety-related correlates, we tested the relationships between safety climate and two types of self-reported safety incidents at the individual and workgroup levels. Although we assessed safety behavior as a theoretically relevant outcome of safety climate, we did not include it in homology tests, because of the unexpectedly high magnitude of its individual- and group-level correlations with safety climate (r = .70 and .80, respectively). The magnitude of these correlations is likely a combined function of (a) the safety behavior items being adapted to have a group-level referent (much like climate items), (b) the two measures using the same response scale, and (3) the two constructs being proximally located in a cross-sectional survey (Podsakoff, MacKenzie, & Podsakoff, 2012). Descriptive statistics and intercorrelations for the examined constructs at both construct levels in Samples 3 to 5 are reported in the online supplement (see Table S3).

On the basis of our earlier reasoning, we expected the aggregate associations between safety climate and safety incidents to be proportionately stronger than corresponding individual-level associations. To test this proposition, we followed the procedures outlined by Chen et al. (2005) and estimated a series of sequential models to assess scalar and then metric similarity of these relationships across construct levels. In the present context, metric similarity would suggest that safety climate's associations with safety incidents are the same in magnitude and meaning at the individual and workgroup levels, whereas scalar similarity could suggest—as we propose—that the magnitudes of these relationships are proportionately similar but not identical. Before conducting these tests, however, we assessed whether self-reported workgroup safety incidents from Samples 3 to 5 revealed sufficient between-group variability for consideration at the workgroup level. These estimates indicate that a meaningful proportion of the variance for incidents (ICC[1] = .10) and injuries (ICC[1] = .11) can be attributed to workgroup membership. Consequently, we proceeded with our multilevel homology tests.

To assess scalar similarity, we estimated models testing the proposed relationships (i.e., safety climate predicted by prior accidents and injuries) at the individual (Model A) and workgroup (Model B) levels, with the individual-level model accounting for the nonindependence due to workgroup nesting and the workgroup-level model being unit size weighted. We then estimated a workgroup-level model (Model C) that constrained the estimated parameters to be a multiplicative function of the individual-level estimates from Model A. This test resulted in computing a scaling factor that equated the individual-level estimates to the corresponding group-level estimates; results for these tests are reported in Table 1. To evaluate scalar similarity, we next compared the fit of the unconstrained group-level model (Model B) and the constrained group-level model (Model C) and did not find a significant decrease in fit ($F_{[2, 175]} = 0.68, p > .10$).[4] These results provide initial support for scalar (or proportional) similarity for these individual- and workgroup-level relationships, which means that the scaling factor of 1.85 can be applied to both predictors—self-reported injuries and incidents—to equate the individual- and group-level models.

We next tested for metric similarity to determine if the individual- and workgroup-level relationships were meaningfully different in magnitude (i.e., whether the estimated scaling factor was statistically significantly different from 1; Chen et al., 2005). Failing to support metric similarity suggests that multilevel relationships are proportionately similar in magnitude (scalar similarity), not identical (metric similarity). However, contrary to expectation, results supported metric similarity for the associations between safety climate and safety injuries and incidents ($F_{[1, 177]} = 2.16, p > .10$), indicating that the scaling factor of 1.85 is not

Table 1
Testing Scalar Similarity Between Safety Climate and Safety Incidents

Model: Predictor	*B*	β	*SE*	*t*
Model A				
(Intercept)	0.00		0.04	0.00
Injuries	−0.05	−0.12	0.03	−1.68
Accidents	−0.00	−0.08	0.00	−3.82*
Model B				
(Intercept)	−0.02		0.05	−0.45
Injuries	−0.11	−0.22	0.04	−2.92*
Accidents	−0.00	−0.11	0.00	−1.53
Model C: Scaling factor	1.85	0.26	0.53	3.51*

Note: These analyses were conducting with the combined sample (Samples 3–5). Model A = individual-level model (n = 771). Model B = workgroup-level model (n = 178). Model C = model constraining workgroup-level parameters to be a multiplicative function of individual-level estimates from Model A. *B* = unstandardized regression coefficient. β = standardized regression coefficient.
*$p < .05$.

significantly different from 1. Stated differently, these results specify that workgroup-level associations between safety climate and two safety incident types (i.e., injuries and incidents) are not substantially different in magnitude from corresponding individual-level associations. This suggests that an identical theory of homology is a more accurate representation of safety climate's multilevel associations with self-reported safety injuries and incidents than a proportional theory. Nevertheless, it is noteworthy that, although not significantly different from 1, the scaling factor indicates that workgroup-level safety climate incident effect sizes are 1.85 times the magnitude of corresponding individual-level associations.

Testing Safety Climate's Workgroup-Level Predictive Validity

In addition to establishing a theory of homology to characterize safety climate's multilevel associations, it is imperative to verify safety climate's predictive validity as part of the validation process (Hinkin, 1998; Messick, 1995). Because the focal level of theory for safety climate is the group (Glick, 1985; Schneider et al., 2011), it is most important to assess predictive validity at the group level. Additionally, to mitigate concerns of common method bias and thereby strengthen statistical conclusions, it is advantageous to gather time-lagged and other-source-reported criteria when evaluating predictive validity (Podsakoff et al., 2012). Consequently, as a final step in this multilevel construct validation, we examined safety climate's workgroup-level associations with organization-reported safety incidents and injuries that occurred in the 6-month periods immediately before and after safety climate assessment.

We evaluated predictive validity solely in Sample 5, as this was the only organization to allow access to archival safety records. Descriptive statistics and intercorrelations among the variables of interest for these analyses are reported in Table 2. As can be seen, safety climate's workgroup-level correlations with self- and organization-reported safety events (i.e., small to moderate and negative) are consistent with theoretical expectations and

Table 2

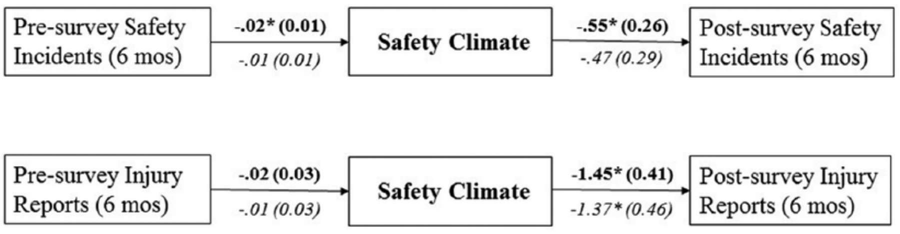
Descriptive Statistics and Intercorrelations for Sample 5

Variable	*M* (*SD*)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Safety climate (overall)	3.99 (0.42)	(.98)													
2. Leader safety commitment	4.24 (0.50)	.93*	(.99)												
3. Safety communication	4.19 (0.47)	.93*	.96*	(.95)											
4. Safety training	3.84 (0.54)	.84*	.65*	.68*	(.96)										
5. Coworker safety practices	3.90 (0.43)	.82*	.68*	.65*	.68*	(.94)									
6. Safety equipment and housekeeping	3.98 (0.42)	.84*	.70*	.70*	.69*	.73*	(.88)								
7. Safety involvement	4.12 (0.47)	.93*	.87*	.90*	.77*	.67*	.73*	(.97)							
8. Safety rewards	3.53 (0.50)	.86*	.74*	.74*	.73*	.66*	.66*	.76*	(.91)						
9. Safety climate (alternate)	3.94 (0.37)	.81*	.73*	.71*	.70*	.58*	.68*	.81*	.76*	(.91)					
10. Incidents (self-report)	1.33 (1.73)	−.28*	−.29*	−.30*	−.21	−.17	−.35*	−.24	−.18	−.21	—				
11. Injuries (self-report)	0.26 (0.48)	−.09	−.01	−.04	−.18	.00	−.18	−.03	−.11	−.08	.50*	—			
12. Incidents (presurvey)	3.81 (5.95)	−.23	−.12	−.12	−.26*	−.19	−.38*	−.17	−.25	−.18	.49*	.27*	—		
13. Injuries (presurvey)	0.58 (1.12)	−.06	.03	.02	−.07	−.07	−.16	.01	−.15	−.04	.27*	.25*	.69*	—	
14. Incidents (postsurvey)	4.11 (6.12)	−.18	−.10	−.11	−.23	−.08	−.32*	−.13	−.18	−.19	.46*	.25*	.70*	.19	—
15. Injuries (postsurvey)	0.47 (0.99)	−.25*	−.16	−.22	−.25*	−.18	−.31*	−.19	−.29*	−.21	.29*	.20	.41*	.11	.63*

Note: Correlations are at the workgroup level of analysis ($n_{workgroups} = 64$, $n_{individuals} = 504$). Safety climate was assessed via the present study's 30-item measure. Safety climate (alternate) was based on Zohar's (2000) 10-item measure of safety climate. Incidents (pre- and postsurvey) and injuries (pre- and postsurvey) were reported by the organization. Numbers in parentheses on the diagonal are coefficient alphas.

*$p < .05$, two-tailed.

Figure 1
Path Models of Safety Climate's Associations With Safety Incidents and Injury Reports



Note: Path models were estimated with data from Sample 5 at the workgroup level ($n = 64$). All path estimates are unstandardized, with standard errors in parentheses; bolded path estimates were obtained via this study's safety climate measure; italicized path estimates were obtained with an alternate safety climate measure from Zohar (2000). $*p < .05$.

existing meta-analytic estimates in terms of the direction and magnitude of associations (Beus, McCord, & Zohar, 2016; Beus et al., 2010; Christian et al., 2009). This provides preliminary evidence of predictive validity for our workgroup-level safety climate responses. Building on these initial confirmatory findings, we estimated path models that simultaneously analyzed the associations between workgroup safety climate and organization-reported safety injuries and incidents (pre- and postassessment). Because "organization-reported safety incidents" is an inclusive variable that subsumes injurious and noninjurious safety-related events (e.g., fires, leaks, spills), we did not include safety incidents and injuries in the same model. Rather, we estimated two path models, with one model including safety incidents and one including only injuries.

The path models, as well as the path estimates for each model, are depicted in Figure 1. The path models were estimated in Mplus version 7 (Muthén & Muthén, 2012) with full information maximum likelihood estimation; because postsurvey incidents and injuries are counts with high frequencies of zero events (i.e., no incidents or injuries over the 6-month period), we modeled these dependent variables through a zero-inflated Poisson distribution. For the model including organization-reported safety incidents, path estimates indicated significant negative associations between pre- and postsurvey safety incidents and workgroup safety climate. In other words, workgroups that reported fewer safety incidents in the 6 months prior to safety climate assessment tended to have significantly more favorable safety climates, as reflected by aggregate scores derived from our safety climate measure. Likewise, workgroups with more favorable safety climates tended to report significantly fewer safety incidents in the 6 months following safety climate assessment.

For the model that included the narrower criterion of organization-reported injuries, the path estimate of the association between presurvey injury reports and safety climate was negative but nonsignificant. However, the association between safety climate and postsurvey injury reports was negative and statistically significant. Thus, for the more restrictive criterion of organization-reported injuries, although there was no association between presurvey workgroup injuries and workgroup safety climates, more favorable safety climates were associated with statistically significant decreases in subsequent workgroup injury reports. In sum, results indicate that while safety climate was in general meaningfully associated with broad safety incidents that occurred pre– and post–safety climate assessment, safety climate was related only to future injury reports and not injuries that were reported before safety climate assessment.

To compare the predictive validity of our newly developed safety climate measure with an alternative safety climate measure, we also assessed safety climate in Sample 5 by using a published 10-item measure from a reputable source (Zohar, 2000). As reported in Table 2, aggregate responses to both measures were very highly correlated (r = .81). Likewise, the workgroup-level correlations with self- and organization-reported safety incidents were in the same direction and of similar magnitude. However, as indicated in Figure 1, estimation of the same path models with the alternative measure in place of our measure resulted in somewhat different conclusions. Specifically, only the association between the alternate measure and postassessment organization-reported injuries reached statistical significance. The associations between the alternate safety climate measure and pre- and postassessment safety incidents were negative but not statistically significant, whereas these relationships reached statistical significance with our newly developed measure. Taken together, although workgroup responses to both safety climate measures were highly correlated with each other, there were some noteworthy differences in statistical conclusions that favor the use of the new measure, as it demonstrated better criterion-related validity relative to the alternate measure.

In summary, Study 2 addressed the remaining steps for conducting a multilevel construct validation effort. Specifically, we demonstrated that safety climate data are represented well across construct levels by a second-order factor structure that explains shared variance among the seven indicators identified in Study 1. Study 2 also demonstrated that an identical theory of homology, as opposed to the proposed proportional theory, most accurately describes safety climate's multilevel associations with self-reported safety incidents. Finally, Study 2 revealed that workgroup responses to the newly developed safety climate measure were meaningfully negatively associated with organization-reported presurvey safety incidents as well as with postsurvey safety incidents and injuries. We discuss the implications of these findings in conjunction with Study 1's findings next.

General Discussion

With climate theory and the extant safety climate literature as a conceptual basis, our purpose was to clarify a number of ambiguities regarding the conceptualization and measurement of safety climate across construct levels. To do so, we first created a generalized measure that delineated seven indicators reflecting safety climate's core content domain: leader safety commitment, safety communication, safety training, coworker safety practices, safety equipment and housekeeping, safety involvement, and safety rewards. Then, in two studies incorporating data from five samples of employees, we conducted a multilevel construct validation effort by using the new measure that clarified safety climate's aggregate nature, its psychometric properties across construct levels, and its multilevel relationships with self- and organization-reported safety outcomes.

Theoretical and Practical Implications

As the introductory quote by Lord Kelvin intimates, it is difficult to advance scientific understanding of an organizational phenomenon without sound measurement. Unfortunately, this has been precisely the problem that has plagued safety climate research. We assert that the root of safety climate's measurement problem has been a broad lack of appreciation of the construct's multilevel conceptualization and content domain. As a result, there has been a proliferation of widely divergent measures that have perpetuated safety climate's ongoing conceptual ambiguity. This confusion has persisted despite the existence of a widely cited definition and conceptualization that are consistent with the broader climate literature (i.e., Zohar, 2011).

A plausible explanation for the creation of so many safety climate measures is that there are theoretically an infinite number of organizational policies, procedures, and practices that can reflect safety's workplace importance across industries and settings. Consequently, one of the core contributions of this study is that we compiled 1,500 safety climate items from >60 measures to inductively identify recurrent and theoretically appropriate themes or categories that represent general indicators of safety's workplace importance across industry settings. The purpose of doing this was not to capture all possible lower-order indicators of safety climate but rather to identify a representative set that would be applicable across contexts and would most effectively prime respondents to consider safety's true priority or importance. We are unaware of any other safety climate measures that have been created in a manner that can offer similar assurances of comprehensiveness in measure content.

The seven dimensions that emerged from our literature search represent a conceptual framework that encapsulates safety climate's general content domain. The identification of these dimensions informed the development of our cross-industry measure and our subsequent tests of competing theoretical factor structures. Results revealed that shared variance across these indicators can be explained at the individual and workgroup levels by a higher-order factor representing safety's general workplace value and importance. Consistent with theoretical expectations concerning climate, this suggests that the seven identified indicators represent related manifestations of safety climate that each reflect safety's workplace priority. Although each first-order indicator loaded strongly onto the latent, second-order safety climate factor (standardized loadings at the group level in Samples 3–5 ranged from .68 to .96), it is noteworthy that safety involvement and safety communication were consistently the highest-loading factors across this study's five samples. Thus, it appears that (a) active worker involvement in safety and (b) open communication of safety issues are particularly meaningful indicators of a group's safety climate—likely because effective reciprocal safety communication and active safety involvement by all are clear and compelling evidence to group members that the group values and prioritizes safety.

This study also extended understanding of safety climate by testing its multilevel associations with relevant constructs. Contrary to our expectations, tests of homology revealed that safety climate's multilevel relationships with self-reported incidents and injuries were best reflected by an identical theory of homology, as opposed to a proportional theory. However, although the scaling factor that equated these multilevel relationships was not statistically significantly different from one (1.85), it still indicates that workgroup-level relationships are nearly 2 times stronger in magnitude than corresponding individual-level relationships. Interestingly, the magnitude of this difference is nearly equal to the difference in the same effect sizes reported in Beus and colleagues' (2010) meta-analysis of safety climate–injury relationships. Specifically, the group-level estimate of safety climate's association with prior injuries reported by Beus et al. ($\hat{\rho} = -.29$) is 1.81 times the magnitude of the individual-level estimate of the same relationship ($\hat{\rho} = -.16$). Thus, although group- and individual-level relationships were not proportionately different from a statistical perspective in the present study—perhaps due to statistical power limitations—it is likely that additional studies would further support differences in the magnitude of associations across levels and consequently substantiate the importance of assessing safety climate as a group-level phenomenon.

In addition to clarifying the homologous nature of safety climate's multilevel associations with self-reported safety incidents, we confirmed the workgroup-level predictive validity of safety climate responses. We supported the expectation that the occurrence of negative reported safety events is associated with less favorable safety climates and that such safety climates are connected with a subsequent increase in negative safety events. This underscores the inherently reciprocal nature of climate in that the outcomes that arise from climate can also serve to reinforce the climate (Beus et al., 2010). In other words, a climate-induced safety event (e.g., chemical leak) may not only be an outcome of safety climate but also an antecedent that can adjust group members' perceptions of safety's overarching priority (or the lack thereof).

Importantly, responses to our safety climate measure evidenced improved predictive validity relative to responses from a reputable alternate safety climate measure. We also found comparative advantages to the new measure in its ability to represent safety climate as a group-level phenomenon. Specifically, although within-group agreement was high according to both safety climate measures (median $r_{wg[j]} = .99$ vs. .94), $r_{wg(j)}$ estimates were universally higher for each workgroup with our measure relative to the comparison measure; similarly, greater proportions of between-group variance in safety climate were accounted for when our measure was used (ICC[1] = .20) relative to the comparison (ICC[1] = .13). This pattern of findings suggests an improved ability not only to predict safety incidents but also to more accurately measure safety climate as an emergent group-level phenomenon through our measure relative to the alternative.

Taken together, the results from Studies 1 and 2 provide clarity concerning safety climate's content domain, multilevel dimensionality and factor structure, and cross-level functionality. This is a meaningful contribution to the safety climate literature, as it provides not only a clearer understanding of what safety climate is but also a well-developed means of assessing safety climate and a better understanding of how it operates psychometrically and relates to relevant constructs at different levels of analysis.

A number of practical implications are associated with this study's findings as well. First, the inductive approach taken to develop this generalized safety climate measure should facilitate its use as a benchmarking tool across organizations and industries. In offering this recommendation, we are aware of the practical complexities that are associated with using a 30-item measure for a single construct. For the sake of time, some organizations may be hesitant to permit a 30-item measure for a single construct when there are often several other constructs being measured in the same survey. Furthermore, it can be difficult to obtain a sufficiently large workgroup-level sample size that provides adequate statistical power to analyze the psychometric properties of workgroup-level responses to a 30-item measure.

With these practical considerations in mind, we believe that there is value in offering a short-form version of our safety climate measure when circumstances make it difficult to use the full measure. One reasonable alternative to the full measure is to adapt a shorter version that uses a single representative item to reflect each of the seven core indicators. The second-order factor structure and high correlations among first-order factors demonstrated in Studies 1 and 2 provide some justification for this approach. It effectively constitutes a downward shift in the factor structure such that the second-order factor is replaced by a first-order factor and the first-order factors are replaced by an observed item from each indicator. To accomplish this, we selected one item from each indicator that we believe best exemplifies that particular indicator from a conceptual perspective. The exception to this was leader safety commitment from which we selected the two most representative items given this dimension's focal role in the safety climate construct (Hofmann & Stetzer, 1996; Zohar, 2002; Zohar & Polachek, 2014). This resulted in selecting eight items for the proposed short measure, which are marked with an asterisk in the appendix. Responses to these eight items were represented adequately by a single latent factor at the individual level ($\chi^2_{[20]} = 260.91$, $p < .05$; CFI = .92; RMSEA = .10; SRMR = .04) and the workgroup level ($\chi^2_{[20]} = 126.29$, $p < .05$; CFI = .91; RMSEA = .17; SRMR = .06) via the combined data from Samples 3 to 5, although the RMSEA values were substandard at both levels of analysis.[5] For comparative purposes, we note a near-unity workgroup-level correlation between aggregated responses to the shortened eight-item measure and the 30-item measure ($r = .99$) in Samples 3 to 5. The two safety climate operationalizations shared nearly identical correlations with relevant covariates as well, although it is noteworthy that the full 30-item measure was slightly more strongly correlated in each case; this was likewise true with organization-reported safety incidents and injuries in Sample 5. Thus, although the reduced measure retains some of the core content of the full measure and may be more practically feasible in some cases, we maintain that the 30-item measure is a more complete representation of safety climate's content domain and should be preferred when conditions allow it, particularly for developmental and diagnostic purposes.

A natural extension of the desire to accurately measure safety climate is the goal to improve it and thereby minimize costly safety incidents. The practical value of the seven identified safety climate indicators (whether assessed with four items or one) is that they offer specific factors that, when improved, can alter a prevailing safety climate and reduce safety incidents. Interestingly, the three specific safety climate dimensions that revealed the strongest relationships with organization-reported incidents and injuries (pre– and post–safety climate assessment) are safety training, safety equipment and housekeeping, and safety rewards. Of the seven safety climate dimensions, these are arguably the most straightforward and tangible dimensions to improve. Offering adequate and timely safety training, providing appropriate equipment and working conditions, and incentivizing safety are all things that are readily within the typical group or organization's capacity to change and may also go the furthest in terms of reducing negative safety incidents. Focusing on these dimensions could also lead to improvements with other dimensions. For example, attaching rewards to safe work behaviors is likely to foster perceptions of greater leader safety commitment, as it would simultaneously indicate that leaders place greater value on workplace safety. Likewise, providing sufficient safety training is likely to enhance worker safety practices, safety communication, and even safety involvement, as employees who are more knowledgeable and capable with regard to safety (i.e., better trained) will likely have greater motivation to be compliant and involve themselves in safety-related issues (Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008).

Limitations and Future Directions

Although this study's findings provide needed conceptual clarity for a construct of demonstrated practical importance, its limitations are worth noting. The rate of nonserious responding in Sample 1 is a concern. While we are confident in the data that we ultimately retained for analysis given our screening criteria, the proportion of respondents who failed these criteria is troubling. However, we note that the second-order factor structure obtained through Sample 1 was substantiated at the individual and workgroup levels in four safety-salient organizational samples. Thus, it is unlikely that the high proportion of nonserious respondents in Sample 1 biased the initial determination of safety climate's factor structure.

An additional limitation is the unexpectedly high correlations between safety climate and safety behavior in Samples 3 to 5. The magnitudes of these correlations are to such a degree as to make the constructs empirically indistinguishable. In addition to the artifactual factors noted earlier, there are substantive reasons to expect safety climate and group behavior to be highly correlated. Theoretical perspectives on climate converge in conceptualizing leader and coworker behaviors as reflecting group priorities and expectations (Ashforth, 1985; Schneider & Reichers, 1983; Zohar & Hofmann, 2012). In other words, although behaviors are a natural by-product of climates, they also inform climate perceptions because they are outward manifestations of what is prioritized and important in the group. Consequently, when climate is assessed—regardless of the domain—it is customary to use behaviorally laden items to capture the phenomenon. An unintended result of doing so, however, is that climate items may correlate particularly strongly with behavior items, especially when those items use the same theoretical referent. In light of this consideration, we recommend that safety climate researchers (and climate researchers in general) take greater care to ensure that hypothesized behavioral outcomes of climate are assessed in a manner that will not unnecessarily inflate statistical relationships. Effective ways of doing so include using different response scales (e.g., frequency instead of agreement based), separating measurement in time, or using other-source reports of behavior (Podsakoff et al., 2012).

We contend that it is difficult to make scientific progress in a field that is plagued by conceptual ambiguity. In light of our findings that clarify the conceptualization and measurement of safety climate across construct levels, there are important issues for future safety climate researchers to consider. First, we know comparatively little about the formation of safety climates (or climates in general) within groups. Although theoretical explanations for climate formation exist (e.g., Ashforth, 1985; Schneider & Reichers, 1983), empirical tests of these explanations, particularly for safety climate, are in short supply. One way of testing this phenomenon would be to survey newly formed groups (e.g., workgroups, work sites, organizations) periodically in their initial stages of development to see how perceptions within groups shift over time and to consider the factors that lead to or detract from within-group agreement (e.g., group composition, opportunities for social interaction, personnel changes). Interconnected with the issue of safety climate formation is the consideration of how safety climates change. Zohar and colleagues (Zohar, 2002; Zohar & Polachek, 2014) have made compelling contributions regarding the role of formal leaders in climate change, but additional factors need to be explored to further understanding of these issues. For example, to what extent do formative events, such as a major injury or workplace fatality, affect safety climates? Beyond the event itself is the consideration of how the group responds to the event (e.g., a learning opportunity vs. an exercise in retribution). In considering safety climate change, we emphasize the need to examine it in terms of the level or favorability of the climate (i.e., mean group safety climate scores) as well as the strength or consensus among group members regarding the climate (i.e., variability around group mean scores). In this way, changes in the quality and pervasiveness of the safety climate can be tracked.

Although we focused our attention on the individual and workgroup levels for this study, it is likewise important to consider other aggregate levels for safety climate research. Climates can have particular importance at the organization level, where the differences among organizations are likely to be starker than those among workgroups nested in the same organization. Unlike workgroups, different organizations generally do not share common leadership or underlying cultural values and assumptions, which likely restricts the observed variance in workgroup-level climates within a single organization. As a result of this workgroup-level range restriction, it is plausible that the emergent properties (e.g., within-group agreement, between-group variance) and predictive validity of organization-level responses to our safety climate measure will be enhanced relative to the workgroup-level results reported here. Future research is needed to determine the extent to which this is the case.

Conclusion

Despite safety climate's demonstrated practical importance, persistent conceptualization and measurement problems have failed to allow science to gain a better understanding of this construct. This study contributes to the extant literature by articulating safety climate's meaning and multilevel nature, inductively defining safety climate's content domain, and creating and validating a cross-industry measure to be used across construct levels. Data from five widely varying samples provided multilevel construct validity evidence for safety climate via the newly developed measure. Given the evidence presented here, we recommend the generalized adoption of our safety climate measure and conceptualization and hope that this will allow safety climate research to advance beyond its past measurement issues to address substantive research questions that will enhance the science and practice of workplace safety.

Appendix

Safety Climate Dimension Descriptions and Items

*Items with an asterisk represent the shortened eight-item measure described in the General Discussion section.

Leader safety commitment. The extent to which employees perceive that their leaders are dedicated to providing a safe workplace.

1. My supervisor strictly enforces the safe working procedures in my workgroup.
2. My supervisor takes a proactive stance when it comes to safety.
3. My supervisor demonstrates leadership by keeping people focused on safety.
4. My supervisor takes the lead on safety issues.
5. My supervisor is committed to improving safety.*
6. My supervisor places a strong emphasis on workplace safety.*

Safety communication. Employees' perception of the effectiveness of communication regarding safety issues.

1. Safety issues are openly discussed between my supervisor and my workgroup.*
2. My workgroup gets timely feedback on safety issues we have raised with our supervisor.
3. My supervisor keeps my workgroup informed of safety rules.
4. My supervisor informs my workgroup when procedure changes affect safety.

Safety training. The extent to which employees perceive that the safety training provided is sufficient to inform all workers on how to work safely.

1. There is adequate safety training in my workgroup.
2. Employees receive safety training when they change work tasks.
3. Enough time is set aside for employee safety training.
4. My supervisor ensures employees have adequate safety training.*

Coworker safety practices. Employees' appraisal of the extent to which their fellow co-workers are committed to workplace safety.

1. My co-workers always follow safety procedures.
2. My co-workers are quick to point out unsafe conditions.
3. My co-workers take safety very seriously.
4. My co-workers are committed to safety improvement.*

Safety equipment and housekeeping. The extent to which employees perceive that they have been provided the proper safety equipment and that working conditions have been maintained sufficiently to ensure worker safety.

1. Employees in my workgroup are given sufficient safety equipment.
2. Efforts are made in my workgroup to provide safe working conditions.
3. Equipment in my work area is checked to make sure it is free of faults.
4. Unsafe conditions are promptly corrected in my work area.*

Safety involvement. The extent to which employees perceive that they are involved in and allowed to contribute to workplace safety decisions.

1. My supervisor consults with employees regularly about workplace safety issues.
2. My supervisor promotes employees' involvement in safety-related matters.
3. My supervisor values employees' ideas about improving safety.
4. My supervisor encourages employees to become involved in safety matters.*

Safety rewards. Employees' perception of the extent to which safety behaviors are reinforced and supported by organizational leaders.

1. The reward system in my workgroup promotes high performance only when work is conducted safely.
2. My supervisor rewards safe behaviors.
3. My supervisor praises safe work behavior.*
4. In my workgroup, employees who work safely get recognition.

Note: All items were answered with a 5-point scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *neither agree nor disagree*, 4 = *agree*, 5 = *strongly agree*).

This measure has been adapted for use at the organization level, and Spanish, French, Italian, and Chinese versions of this measure are available upon request.

As part of ongoing efforts to advance the utility of this measure, we have constructed a normative database that is publically available for benchmarking purposes. To this end, we encourage all users of this measure to contribute their descriptive statistics to the database as a professional courtesy. To contribute information to this database (e.g., descriptive statistics, administration, and sample information), please go to http://safetyclimate.sites.tamu.edu or contact the corresponding author. Thank you.

Notes

1. This total is not 132, because one item was split into two items, resulting in 133 total items.

2. Respondents in Sample 1 were chosen from nine safety-salient occupational categories (e.g., construction, law enforcement, manufacturing) as well as one nonsafety salient occupational category (i.e., accounting/financial) for comparative purposes. A complete listing of occupations is provided in the online supplement (see Table S1). Multigroup CFA confirmed a common safety climate factor structure across safety-salient and non-safety-salient categories, so respondents from the non-safety-salient occupational category were retained in the final sample.

3. The developmental sample was purposefully made larger than the cross-validation sample to provide greater statistical power to test factor structures with all 133 items. The cross-validation sample was used to test a smaller subset of items and, as such, required a smaller sample to achieve the requisite statistical power.

4. To ensure that we had sufficient power to reject the null hypothesis in these tests (i.e., to fail to support scalar and metric invariance), we followed Chen and colleagues' (2005) recommendation to adopt a more conservative p value, given our comparatively limited sample size at the workgroup level. We thus adjusted the significance level of these F tests to $p < .10$ to ensure sufficient statistical power to detect small model differences.

5. Administrations of the eight-item short measure of safety climate in samples for other projects—where the items were grouped together as a set—revealed better fit for a single-factor solution at the individual and group levels of analysis (i.e., CFI > .90, RMSEA < .08, SRMR < .08). These results are available from the first author upon request.

References

Ashforth, B. E. 1985. Climate formation: Issues and extensions. *Academy of Management Review*, 10: 837-847.

Barling, J., Loughlin, C., & Kelloway, E. K. 2002. Development and test of a model linking safety-specific transformational leadership and occupational safety. *Journal of Applied Psychology*, 87: 488-496.

Bass, B. 1985. *Leadership and performance beyond expectations*. New York: Free Press.

Beus, J. M., McCord, M. A., & Zohar, D. 2016. Workplace safety: A review and research synthesis. *Organizational Psychology Review*, 6: 352-381.

Beus, J. M., Payne, S. C., Bergman, M. E., & Arthur, W., Jr. 2010. Safety climate and injuries: An examination of theoretical and empirical relationships. *Journal of Applied Psychology*, 95: 713-727.

Binning, J. F., & Barrett, G. V. 1989. Validity of personnel decisions: A conceptual analysis of the inferential ad evidential bases. *Journal of Applied Psychology*, 74: 478-494.

Bliese, P. D. 1998. Group size, ICC values, and group-level correlations: A simulation. *Organizational Research Methods*, 1: 355-373.

Bliese, P. D. 2000. Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*: 349-381. San Francisco: Jossey-Bass.

Bliese, P. D., Chan, D., & Ployhart, R. E. 2007. Multilevel methods: Future directions in measurement, longitudinal analyses, and nonnormal outcomes. *Organizational Research Methods*, 10: 551-563.

Brislin, R. W. 1970. Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1: 185-216.

Chan, D. 1998. Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, 83: 234-246.

Chen, G., Bliese, P. D., & Mathieu, J. E. 2005. Conceptual framework and statistical procedures for delineating and testing multilevel theories of homology. *Organizational Research Methods*, 8: 375-409.

Chen, G., Mathieu, J. E., & Bliese, P. D. 2004. A framework for conducting multi-level construct validation. In J. Yammarino & F. Dansereau (Eds.), *Research in multilevel issues: Multilevel issues in organizational behavior and processes*, vol. 3: 273-303. Oxford, UK: Elsevier.

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: 1103-1127.

Clarke, S. 2006. The relationship between safety climate and safety performance: A meta-analytic review. *Journal of Occupational Health Psychology*, 11: 315-327.

Crocker, L., & Algina, J. 2006. *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.

Dedobbeleer, N., & Beland, F. 1991. A safety climate measure for construction sites. *Journal of Safety Research*, 22: 97-103.

Flin, R., Mearns, K., O'Connor, P., & Bryden, R. 2000. Measuring safety climate: Identifying the common features. *Safety Science*, 34: 177-192.

Glick, W. H. 1985. Conceptualizing and measuring organizational and psychological climate: Pitfalls in multilevel research. *Academy of Management Review*, 10: 601-616.

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: 347-358.

Guldenmund, F. W. 2000. The nature of safety culture: A review of theory and research. *Safety Science*, 34: 215-257.

Haladyna, T. M. 2004. *Developing and validating multiple-choice test items* (3rd ed.). New York: Routledge.

Harvey, R. J., Billings, R. S., & Nilan, K. J. 1985. Confirmatory factor analysis of the Job Diagnostic Survey: Good news and bad news. *Journal of Applied Psychology*, 70: 461-468.

Hinkin, T. R. 1998. A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1: 104-121.

Hofmann, D. A., & Stetzer, A. 1996. A cross-level investigation of factors influencing unsafe behaviors and injuries. *Personnel Psychology*, 49: 307-339.

Hu, L., & Bentler, P. M. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6: 1-55.

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. 2012. Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, 27: 99-114.

Kozlowski, S. W. J., & Klein, K. J. 2000. A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations*: 3-90. San Francisco: Jossey-Bass.

Kunzelman, M. 2011. *Deepwater Horizon owner Transocean had poor safety culture, says Coast Guard*. Retrieved from http://blog.al.com/wire/2011/04/deepwater_horizon_owner_transo.html

LeBreton, J. M., & Senter, J. L. 2008. Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11: 815-852.

Messick, S. 1995. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50: 741-749.

Morgeson, F. P., & Hofmann, D. A. 1999. The structure and function of collective constructs: Implications for multilevel research and theory development. *Academy of Management Review*, 24: 249-265.

Muthén, L.K., & Muthén, B. O. 2012. *Mplus user's guide* (7th ed.). Los Angeles: Muthén & Muthén.

Podsakoff, P. M., Bommer, W. H., Podsakoff, N. P., & MacKenzie, S. B. 2006. Relationships between leader reward and punishment behavior and subordinate attitudes, perceptions, and behaviors: A meta-analytic review of existing and new research. *Organizational Behavior and Human Decision Processes*, 99: 113-142.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. 2012. Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63: 539-569.

Powell, G. N., & Butterfield, D. A. 1978. The case for subsystem climates in organizations. *Academy of Management Review*, 3: 151-157.

Rodriguez, J. M., Payne, S. C., Bergman, M. E., & Beus, J. M. 2011. The impact of the BP Baker report. *Journal of Safety Research*, 42: 215-222.

Schneider, B. 1975. Organizational climates: An essay. *Personnel Psychology*, 28: 447-479.

Schneider, B., Ehrhart, M. G., & Macey, W. H. 2011. Perspectives on organizational climate and culture. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology: Building and developing the organization*, vol. 1: 373-414. Washington, DC: American Psychological Association.

Schneider, B., & Reichers, A. E. 1983. On the etiology of climates. *Personnel Psychology*, 36: 19-39.

Shannon, H. S., & Norman, G. R. 2009. Deriving the factor structure of safety climate scales. *Safety Science*, 47: 327-329.

Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. 2008. A review and meta-analysis of the nomological network of trainee reactions. *Journal of Applied Psychology*, 93: 280-295.

Smith-Crowe, K., Burke, M. J., Cohen, A., & Doveh, E. 2014. Statistical significance criteria for the rWG and average deviation interrater agreement indices. *Journal of Applied Psychology*, 99: 239-261.

Stanton, J. M., & Weiss, E. M. 2002. *Online panels for social science research: An introduction to the StudyResponse project* (Technical Report No. 13001). Syracuse, NY: Syracuse University School of Information Studies.

van Mierlo, H., Vermunt, J. K., & Rutte, C. G. 2009. Composing group-level constructs from individual-level survey data. *Organizational Research Methods*, 12: 368-392.

Williamson, A. M., Feyer, A., Cairns, D., & Biancotti, D. 1997. The development of a measure of safety climate: The role of safety perceptions and attitudes. *Safety Science*, 25: 15-27.

Zohar, D. 1980. Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology*, 65: 96-102.

Zohar, D. 2000. A group-level model of safety climate: Testing the effect of group climate on microinjuries in manufacturing jobs. *Journal of Applied Psychology*, 85: 587-596.

Zohar, D. 2002. Modifying supervisory practices to improve subunit safety: A leadership-based intervention model. *Journal of Applied Psychology*, 87: 156-163.

Zohar, D. 2011. Safety climate: Conceptual and measurement issues. In J. C. Quick & L. E Tetrick (Eds.), *Handbook of occupational health psychology* (2nd ed.): 141-164. Washington, DC: American Psychological Association.

Zohar, D., & Hofmann, D. A. 2012. Organizational culture and climate. In S. W. J. Kozlowski (Ed.), *The Oxford handbook of industrial and organizational psychology*: 643-666. Oxford, UK: Oxford University Press.

Zohar, D., & Luria, G. 2004. Climate as a social-cognitive construction of safety practices: Scripts as proxy of behavior patterns. *Journal of Applied Psychology*, 89: 322-333.

Zohar, D., & Luria, G. 2005. A multilevel model of safety climate: Cross-level relationships between organization and group-level climates. *Journal of Applied Psychology*, 90: 616-628.

Zohar, D., & Polachek, T. 2014. Discourse-based intervention for modifying supervisory communication as leverage for safety climate and performance improvement: A randomized field study. *Journal of Applied Psychology*, 99: 113-124.