Journal of Safety Research 98 (2026) 376–389

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

What makes a safety control “enough” for high-energy hazards? Defining and testing acceptable alternative controls when a Direct Control is not feasible

Yaqoob Raheemy [*], Fred Sherratt [1], Matthew R. Hallowell

Construction Safety Research Alliance, University of Colorado Boulder, CO 80309, USA

A R T I C L E I N F O

Keywords:
Serious injuries and fatalities
Construction safety
Safety controls
Alternative controls
Direct control

A B S T R A C T

Introduction: Serious injuries and fatalities (SIFs) in construction have remained stubbornly high despite substantial safety efforts. Prior research suggests that these events frequently arise in contexts where controls are absent, ineffective, poorly defined, or inconsistently implemented. Prevention appears to depend less on adding more controls and more on designing controls that are explicitly intended for SIF prevention, usable in the field, and precise enough to support consistent decisions at the point of risk. Within the energy-based safety framework, the concept of Direct Control has been introduced as a scientifically defined type of control for high-energy hazards with SIF potential. However, Direct Controls are not always feasible, leaving a recurring gap in practice: what constitutes “enough” control when high-energy hazards must be managed without a Direct Control? *Method:* This study addresses that gap through a two-phase research process. In Phase 1, an expert panel of 25 construction safety professionals participated in structured brainstorming, focus groups, and a modified Delphi method to develop a definition and criteria for Alternative Controls. The resulting framework specifies three categories and rules to ensure that Alternative Controls are timely, tangible, targeted, and reliably implemented at the point of risk. In Phase 2, a noise experiment with a separate group of 20 practitioners tested the framework’s usability. Variability in adequacy decisions decreased by 47%, inter-rater agreement improved from fair to substantial (Fleiss’ κ from 0.21–0.32 to 0.61–0.62), and practitioner confidence increased significantly (paired *t*-test, $p < 0.001$). *Results:* Findings indicate that structured criteria for Alternative Controls can reduce decision noise and support more consistent, SIF-focused control decisions when Direct Controls are not feasible.

1. Introduction

Despite significant efforts to reduce serious injuries and fatalities (SIFs) in the North American construction industry, the rate of SIFs remains both alarmingly high and stubbornly constant (Bayona et al., 2024; Hallowell, 2023). Unsurprisingly, numerous programs have been introduced to try to overcome this ‘SIF plateau’ within the occupational safety data, and a growing body of research is specifically dedicated to preventing SIF (Selleck et al., 2023). A notable aspect of this research highlights one key issue: SIF events frequently arise from work contexts where controls are absent, ineffective, poorly specified, or implemented in a fragmented way, such that they fail to interrupt the pathways that lead to severe outcomes (e.g., Ahmadi et al., 2021; Hoffman & Wilkinson, 2011; Saleh et al., 2010; Squillante et al., 2021). In particular, several recent studies have suggested that fatalities in the construction industry often stem from high-risk activities where controls are either undefined, ineffective, or inadequate (Selleck et al., 2022; 2023). For example, and more specifically, an investigation by Winge and Albrechtsen (2018) into 176 accidents revealed that the majority of these incidents could have been prevented if adequate controls (particularly at least two physical controls) had been implemented.

This issue has attracted significant attention among researchers, and there is growing agreement that SIFs must be managed differently from less severe outcomes (e.g., Bayona et al., 2024; Busch et al., 2021). In particular, controls intended to eliminate or mitigate SIFs should be specifically designed for that purpose and clearly defined in operational

* Corresponding author.
E-mail addresses: mohammadyaqoob.raheemy@colorado.edu (Y. Raheemy), fred@nemesisv.co.uk (F. Sherratt), Matthew.Hallowell@colorado.edu (M.R. Hallowell).

[1] This author was formerly affiliated with the Construction Safety Research Alliance at the University of Colorado Boulder.

https://doi.org/10.1016/j.jsr.2026.07.013
Received 11 December 2025; Received in revised form 24 April 2026; Accepted 27 July 2026

terms (Busch et al., 2021; Selleck et al., 2023). High-risk industries such as mining have already moved in this direction by distinguishing controls that address high-consequence events from those aimed at lower-risk events. For example, the International Council on Mining and Metals (ICMM) has criticized the broad classification of almost any safety measure as a "control," arguing that this dilutes attention from measures that are truly critical (ICMM, 2015). In their guidance, training and general safety processes are important but do not qualify as controls for mitigating high-risk hazards that can lead to SIFs. Instead, ICMM emphasizes clearly specified critical controls that are directly linked to serious incident scenarios.

Analogously, within construction, Selleck et al. (2022; 2023) advocate for a more refined approach to safety controls. Their work emphasizes that preventing persistent SIF events depends less on simply adding more controls and more on designing controls that are explicitly SIF-focused and reliable at the point where workers interact with high-risk activities. These studies emphasize the importance of designing controls that are clearly defined, targeted, and structured to ensure they are adequate in mitigating risks associated with activities that have the potential to result in SIFs. Specifically, controls should be (Selleck, 2023; Selleck & Cattani, 2019): (a) defined so they are practical and usable in the field; (b) accompanied by clear, usable criteria that directly guide decisions about high-energy, SIF-relevant hazards; (c) applicable and adaptable across a range of high-risk activities; and (d) specified with sufficient precision to enable consistent and informed decision-making during implementation at the point of risk (Hopkins, 2011).

These emerging insights emphasize the need to define controls that effectively mitigate SIF hazards in construction and raise a central practical question: how should such controls be defined, tested, and implemented? This study addresses that question by focusing on a specific and recurrent problem: determining what constitutes "enough" control when work involves high-energy hazards (i.e., hazards with SIF potential) and a Direct Control is not feasible. The objectives were to: (a) develop a structured definition and explicit criteria for Alternative Controls to be used for high-energy hazards when a Direct Control is not feasible, and (b) evaluate whether this definition and criteria improve decision-making by reducing variability and increasing agreement and confidence among practitioners when judging the adequacy of controls in high-energy scenarios.

2. Context

The concept of control has historically played a fundamental role in safety management (Li & Guldenmund, 2018). The earliest recorded mention of controls in safety legislation can be found in the UK's Factory and Workshop Act of 1878, which mandated machine guarding as the first legally required safety control in mills and factories during the industrial revolution (Sherratt, 2016). Subsequently, controls have been prominent on both sides of an incident – both preventatively in risk assessment and management and retrospectively through the theorizing, modeling, and investigation of accidents (e.g., Bellamy, 2015; Zhang et al., 2022).

Indeed, examples of controls can be noticed prominently in a range of accident models, from early conceptualizations like Heinrich's Domino Theory (Heinrich, 1931) to contemporary frameworks that incorporate systemic and multi-layered control strategies (e.g., Ge et al., 2022; Hoffman & Wilkinson, 2011; Squillante Jr. et al., 2021). These models demonstrate the pervasive role of controls in managing safety risks, yet they also reveal substantial variation in how controls are understood and applied.

Although the vital role of safety controls is widely acknowledged across various studies, the field lacks a consistent and accepted definition of the term 'control' itself (Roelen et al., 2018). Additionally, terms such as 'safeguard,' 'defense,' and particularly 'barrier' have been used with similar meanings or even interchangeably with controls, all of which remain equally vague (Sklet, 2006; De Ruijter & Guldenmund, 2016). There is no common definition for these terms within safety literature or standards (Shahrokhi & Bernard, 2010; Sklet, 2006). This inconsistency in terminology extends to the application of controls, resulting in diverse and often ambiguous classifications. As a result, consistent definitions and classifications of controls remain elusive.

2.1. Controls in safety risk management

Risk management is a domain within safety science where controls are extensively discussed and hold a central role. Safety risk is defined as the likelihood of injury or illness of a given severity (Hallowell & Gambatese, 2009). When modeling risk, it is typically expressed as the product of the frequency or probability of an incident and the severity of its potential outcome (Hallowell et al., 2017). In this process, controls are identified and implemented as part of risk reduction strategies intended to reduce risk to an acceptable level (Liu et al., 2023). This "acceptable level" is often articulated through the As Low As Reasonably Practicable (ALARP) principle or its U.S. synonym, As Low As Reasonably Achievable (ALARA). Thus, the adequacy of controls is analyzed and judged by whether the controls successfully lower the risk to an acceptable threshold. Numerous quantitative, semi-quantitative, and qualitative risk assessment methods exist for this purpose, each selected based on the specific context and complexity (Khan et al., 2015). Common examples include bowtie analysis, fault and event tree analysis, barrier analysis, and layers of protection analysis, all of which are reflected in current risk assessment guidance and standards (Lyon & Popov, 2019; 2016). More recent literature also shows that occupational health and safety risk assessment continues to involve a wide range of methods and substantial variation in how risks are evaluated and prioritized in practice (Liu et al., 2023). This suggests that, although risk assessment is widely used, there is still no single, consistent basis for deciding what level of control is sufficient across different work settings.

According to ALARP, risk reduction should be pursued until the effort, time, or cost involved becomes grossly disproportionate to the benefit gained from any additional reduction (Hurst et al., 2019; Baybutt, 2014; Bowles, 2003). These principles are embedded in the regulatory requirements of numerous countries, most prominently in the UK through legislation such as the Health and Safety at Work Act (1974), with varying legal status (Jones-Lee & Aven, 2011). At the same time, more recent literature continues to show that applying risk acceptance principles in practice is not straightforward. Reviews of occupational risk-acceptance criteria show continued diversity in how acceptable risk is defined and used (Marhavilas & Koulouriotis, 2021), while more recent analysis suggests that commonly used tools do not, by themselves, fully demonstrate that risks have been reduced so far as is reasonably practicable (Ho, 2023). Recent applied work also continues to frame ALARP as a balance among risk, cost, and benefit, while showing that its practical application still depends on threshold setting, context, and judgment (Maselli et al., 2021).

Despite its apparent clarity, ALARP's practical interpretation remains nuanced due to the inherently subjective determination of what is "reasonably practicable" or "achievable" (Redmil, 2010). Although the term "reasonably practicable" is not explicitly defined in legislation, judicial interpretations often balance the degree of risk against the time, effort, cost, and physical difficulty of implementing risk reduction measures (Melchers, 2001). More recent discussion continues to anchor this interpretation in the long-standing legal idea that a decision requires weighing the quantum of risk against the sacrifice involved in averting it, with gross disproportion remaining central to the judgment (Kotre, 2022). Thus, the threshold of acceptance or sufficiency is not always clear-cut and often varies by context and by the perspectives of decision-makers (Melchers, 2001). Moreover, factors such as organizational priorities, financial pressures, and cultural attitudes toward safety heavily influence how ALARP is interpreted. Efforts to ground the "reasonably practicable" threshold in tangible metrics often involve cost-benefit analysis within the ALARP framework (Hurst et al., 2019).

While cost-benefit analysis is widely recognized, its suitability is questioned because it attempts to monetize intangible factors such as societal or environmental impacts (Ale et al., 2015). This remains an active issue in the literature, particularly because cost-benefit approaches may support ALARP decisions but still operate within broader legal, practical, and ethical considerations rather than resolving them fully (Kotre, 2022). As a result, even where a formal risk management process is used, the threshold for what counts as "enough" control can still vary across settings and individuals.

Consequently, the definition of what constitutes "enough" control remains unsettled and tied to the inherently subjective ALARP principle. In addition, an underlying assumption in control selection and layered risk management is that defined controls, including human actions, will be implemented consistently across construction activities. In practice, construction research shows that this assumption does not hold, and the application of controls varies, introducing decision variability (Selleck, 2023). Overall, although risk management processes and the ALARP principle are central to safety practice, they provide limited task-level guidance on how controls should be selected, combined, and judged in the field, particularly for high-risk activities involving high-energy hazards. In construction, where the application of controls often varies across crews, supervisors, and work conditions, this leaves practitioners without a clear basis for consistent and confident decisions about adequacy.

2.2. The hierarchy of controls

Arguably, the most prominent practical application of controls is within the Hierarchy of Controls (HOC), a framework extensively applied across various fields and frequently cited in regulations and standards. In the United States, NIOSH presents the HOC as a framework for hazard prevention and control, and OSHA recommends its use as part of hazard prevention and control strategies. Similarly, the United Kingdom's Health and Safety Executive incorporates the HOC in its risk assessment and management processes (HSE, 2024). The HOC categorizes controls by their effectiveness, ranking elimination as the most effective type and personal protective equipment (PPE) as the least effective in mitigating risks. The common framework of the HOC, from most to least effective, is shown in Fig. 1.

Despite its widespread use, the application and reliability of the HOC have been debated in numerous studies. For example, some studies have noted that the HOC has largely developed through speculation, including the order of its elements, rather than through systematic empirical validation (e.g., Barnett, 2020; Barnett & Brickman, 1986). In certain cases, the practical application of controls in the traditional HOC order has been questioned, with some researchers even proposing an inverse HOC as a potentially more effective alternative (Omari Shekaftik et al., 2023). There is no consistent version of the HOC, and several variations are in use (Lyon & Popov, 2020). This variation is clearly evidenced by the different versions of the HOC presented in the NIOSH framework and in ANSI standards (e.g., CDC, 2024; ANSI Z10-2012 (R2017); ANSI B11.0-2020). For instance, the NIOSH version includes five steps, as discussed above; however, the ANSI version introduces warnings as a control level before administrative controls, making it a six-level hierarchy. Barnett (2020), for example, conducted a comprehensive study on safety hierarchies, analyzing 66 publications since 1952, including 21 different versions of the HOC, and concluded that these hierarchies show significant variation in conceptual foundation and application. These findings have led some scholars to emphasize the need for more rigorous research and clearer guidance on how such hierarchies should be interpreted and applied in practice (Ajslev et al., 2022; Barnett, 2020).

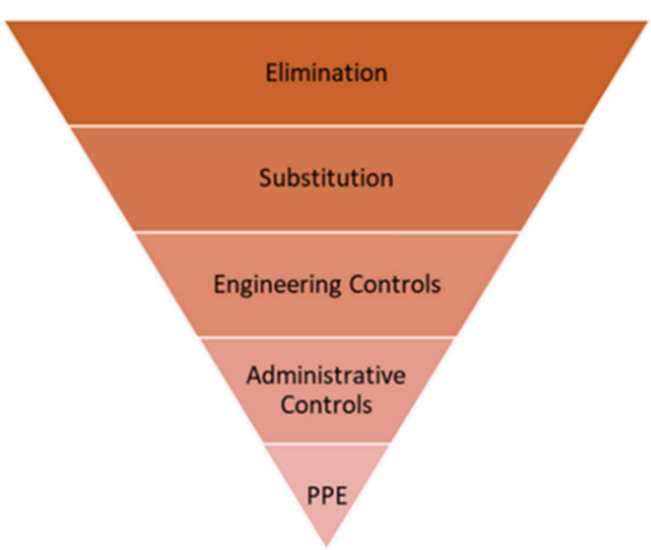

Fig. 1. The hierarchy of controls, adapted from NIOSH (2022).

In practice, the HOC provides little prescriptive guidance on how to layer controls or how to judge when layering is sufficient for a given task. Selection, integration, and verification are typically left to organizations and practitioners (OSHA, 2024), which produces variability across settings. Moreover, there is little evidence that HOC application is calibrated to hazard severity; in many cases the same ordering is applied to both low-risk and high-risk contexts. This results in a classification of control types without operational criteria for adequacy in high-energy hazards.

Evidence from construction further indicates that layered risk management often emphasizes the completion of assessments rather than the effectiveness of the controls themselves, creating a disconnect between hazards and the controls that must prevent energy release (Dekker, 2014; Reason, 2016; Smith, 2018). When controls are selected, line supervision is frequently left to interpret whether they are adequate using general rules rather than explicit, field-verifiable criteria, which introduces decision variability, especially under production pressure and differing risk perceptions.

The HOC is useful for classifying types of controls, but it gives little practical guidance on when a control is "enough" for high-energy hazards. This absence of operational criteria creates uncertainty in decision-making and adequacy judgments. In practice, it shifts effort toward completing paperwork rather than assuring that the controls work as intended (Dekker, 2014; Smith, 2018). It also leaves frontline supervisors to make subjective calls, which introduces variability (Hayes, 2012; Hopkins, 2011). Consistent with this pattern, SIF outcomes have remained stagnant and, in some reports, have worsened (Selleck, 2023).

2.3. Energy-based safety and direct control

Within recent literature, energy-based safety has emerged as a scientifically rigorous framework for mitigating SIFs in construction. A foundational study by Hallowell et al. (2017) empirically demonstrated a direct and measurable relationship between energy magnitude and injury severity, providing an objective basis for differentiating between SIF and non-SIF events. Specifically, the study analyzed a wide range of workplace incidents and quantified injury outcomes based on the energy released during these events. The findings revealed that energy magnitudes below 500 J were typically associated with first-aid injuries, while energy levels between 500 and 1500 J often resulted in medical treatment cases. Critically, energy magnitudes exceeding 1500 J, defined as 'high-energy,' were frequently linked to SIFs. This threshold of 1,500 J not only allows for clear differentiation between SIF and non-SIF events but also establishes an actionable metric for identifying and prioritizing high-energy hazards in the field. To make the framework more accessible and practical, Hallowell (2023) introduced a set of 13 visual icons representing approximately 85% of documented high-energy hazards

(Erkal & Hallowell, 2023), as shown in Fig. 2. While these icons are used as an aid to simplify hazard recognition, the method described by Hallowell et al. (2017) can be applied to other energy forms not represented by the icons.

Building on this foundation, Hallowell (2023) scientifically defined the concept of 'Direct Control' within the energy-based safety framework. Direct Controls are uniquely designed to target high-energy hazards, identified as precursors to SIFs, and are characterized by three essential attributes: (a) they specifically address high-energy sources; (b) they effectively reduce exposure when properly installed and verified; and (c) they remain effective even in the presence of unintentional human error during work (excluding installation-related errors). Examples of Direct Controls include lockout/tagout (LOTO), machine guarding, fall protection systems, and cover-up measures. These controls are distinguished from other interventions, such as training or warning signs, which are susceptible to human error. Table 1 provides a more operational specification of a Direct Control by showing how its definition can be translated into field-level elements for implementation and evaluation.

The critical role of Direct Control in preventing SIFs is further emphasized by Bayona et al. (2024), who found that missing or insufficient Direct Controls almost always precede a SIF, emphasizing the necessity of having Direct Controls in place to mitigate high-energy hazards effectively.

Ideally, a Direct Control should exist for all high-energy hazards that workers face. However, this is not currently the case. There are circumstances where a Direct Control is simply not feasible, due to operational limitations, environmental factors, or the specific characteristics of the hazards involved in the work. Indeed, research by Erkal et al. (2021) estimated that around one-third of typical construction-related work tasks that include high-energy hazards have no corresponding Direct Control available to keep workers from harm.

Future research should aim to address this gap, either by developing new Direct Controls or by redesigning work practices to eliminate high-energy exposures altogether. However, until such advances occur, practitioners frequently encounter situations where high-energy hazards must be controlled using measures other than Direct Controls. These situations require clear, scientifically grounded criteria for identifying acceptable Alternative Controls that can provide the highest possible level of SIF-relevant protection.

3. Point of departure

The preceding sections show a persistent gap in current safety practice. Frameworks such as ALARP and the Hierarchy of Controls provide high-level guidance, but they do not offer clear, field-usable criteria for judging when controls are "enough," particularly for high-energy work with SIF potential. In practice, this leaves frontline personnel and safety practitioners to make inherently subjective adequacy judgments based on broad principles and local interpretation, which introduces variability and contributes to inconsistent protection in high-energy hazard situations.

In contrast, energy-based safety and the concept of Direct Control demonstrate that it is possible to define SIF-focused controls in precise, operational terms. However, Direct Controls are not always feasible in construction, and when they are not feasible, practitioners still need to decide how to manage high-energy hazards in a way that is as protective and consistent as possible.

This study takes that problem as its point of departure. The following sections describe the empirical process used to construct, refine, and test a criterion-based definition of Alternative Controls for managing high-energy hazards when a Direct Control is not feasible. Phase 1 details how expert judgment was organized and structured to produce a definition, categories, and rules for Alternative Controls; Phase 2 presents an experimental evaluation of whether those criteria reduce decision variability and improve agreement and confidence in real-world STKY scenarios. In developing the definition of Alternative Controls, the study also identified a more basic concept, termed Energy Control, which was used to distinguish measures that directly address a high-energy hazard from measures that support work more generally.

4. The research process

This research was conducted in two phases. The first phase focused on establishing what would be considered 'enough' in terms of acceptable Alternative Controls able to mitigate SIFs when a Direct Control is not feasible. This phase involved four steps: (a) convening a panel of safety experts; (b) holding a brainstorming session to generate comprehensive ideas regarding Alternative Controls; (c) conducting focus groups to establish the initial definition and criteria for acceptable Alternative Controls; and (d) employing the Delphi method to further refine, build consensus, and finalize the definition and criteria.

The second phase involved conducting an experiment with an external team of safety practitioners to assess the extent to which the definition and criteria improved consistency in selecting adequate Alternative Controls for high-energy hazard cases. This phase evaluated whether the definition reduced decision variability, strengthened inter-rater agreement, and increased practitioner confidence when judging

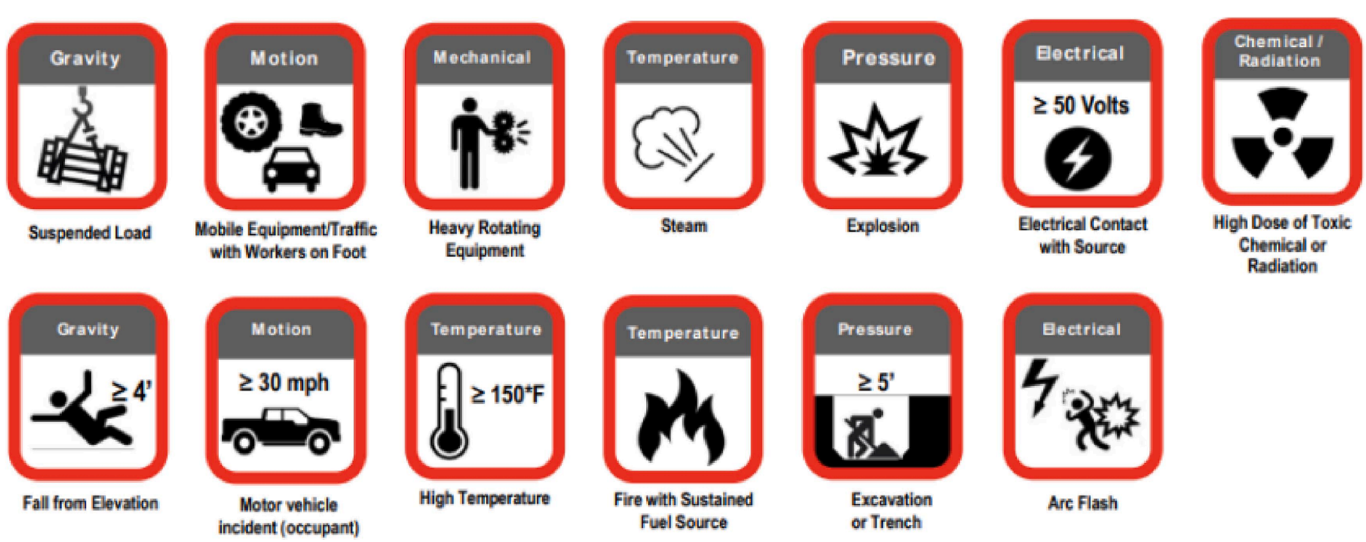

Fig. 2. Common high-energy icons, adopted from Hallowell (2023).

Table 1
Operational requirements for direct control (Hallowell, 2025).

Element	Direct control requirement	Examples/Explanation
Physical or technical object or state	The definition of Direct Control requires a control that specifically addresses the high-energy source and is effective when properly installed and verified. In field terms, this means the control must exist as an actual object, system, or controlled state, not as a general instruction or intention.	Lockout/tagout establishes a controlled de-energized state. Machine guarding is a physical barrier installed at the point of hazard. Cover-up measures are insulating barriers placed over energized parts. Excavation support systems are installed systems that control ground movement.
Interface to exposure	The definition requires that a Direct Control reduce exposure to the high energy hazard. In field terms, this means the control must remove the hazardous state (i.e., high-energy hazard), isolate the worker from the source, or interrupt the transfer of harmful energy to the worker.	Elimination: lockout/tagout removes exposure by de-energizing the source. Isolation: machine guarding separates the worker from moving parts. Isolation: cover-up measures prevent contact with energized parts. Mitigation: a fall protection system interrupts the transfer of harmful energy in a fall event.
Performance requirement	The definition requires that the control be effective when properly installed and verified. In field terms, this means that the control must meet a condition that allows it to function as intended against the high-energy hazard.	For lockout/tagout, effectiveness means the source is de-energized and secured. For machine guarding, effectiveness means the guard remains in place and prevents contact. For cover-up measures, effectiveness means the barrier remains intact and provides the needed insulation or separation. For excavation support, effectiveness means the system maintains ground stability as intended.
Verification method and frequency	The definition explicitly includes the condition that the control is effective when properly installed and verified. In field terms, this means the control must be confirmed as in place and functioning before work begins and, where needed, during the work.	Lockout/tagout requires confirmation that isolation has occurred before exposure begins. Machine guarding requires confirmation that the guard is installed and remains in place. Cover-up measures require confirmation that the protection has been applied correctly and remains in place. Excavation support systems require confirmation that the installed system matches the intended protective arrangement.
Degradation or failure modes	The definition requires that the control remains effective despite unintentional human error during work. In field terms, this requires attention to the ways a control may lose integrity or stop functioning as intended in practice.	Machine guarding may lose effectiveness if removed, displaced, or damaged. Cover-up measures may lose effectiveness if incomplete, shifted, or torn. Excavation support systems may lose effectiveness if installed incorrectly, damaged, or not maintained in the required condition.
Dependence on context	Whether a control functions as a Direct Control depends on the full control arrangement and the way it is applied in the field. The same work activity may or may not satisfy the definition depending on the design and features of the system in use.	A crane operator moving a suspended load may be vulnerable to human error if the control depends only on operator action. However, if the system includes features such as an automatic brake or a load-limiting device, the overall arrangement may satisfy the criteria for a Direct Control, depending on how those features act on the high-energy hazard.
Sensitivity to design, installation, or verification error	A Direct Control is intended to remain effective despite unintentional human error during the work itself. However, errors in design, installation, or verification may still prevent it from functioning as intended.	A hard physical barrier may satisfy the criteria for a Direct Control, but engineering miscalculation, incomplete assembly, or missed inspection may compromise its effectiveness in the field.
System character of many Direct Controls	Many Direct Controls function as systems rather than as single elements. In such cases, all energy-mitigating components must be present and perform as intended for the control to qualify and remain effective.	Fall protection systems, lifting and rigging arrangements, and excavation support systems depend on multiple components working together. Failure in one component may reduce the effectiveness of the entire control system.

the adequacy of controls in high-energy scenarios. The overall research process is illustrated in Fig. 3 and discussed in detail in the following sections.

Participants in both phases were volunteer members of the Construction Safety Research Alliance. Participation was voluntary, and informed consent was obtained before data collection. Participants were informed about the purpose of the study, what participation involved, and how the data would be used. In line with the ethical research practices of the Construction Safety Research Alliance and its host institution, the University of Colorado Boulder, studies in this research program are reviewed by the host institution's IRB Office. All data were kept confidential and protected in accordance with established data security procedures.

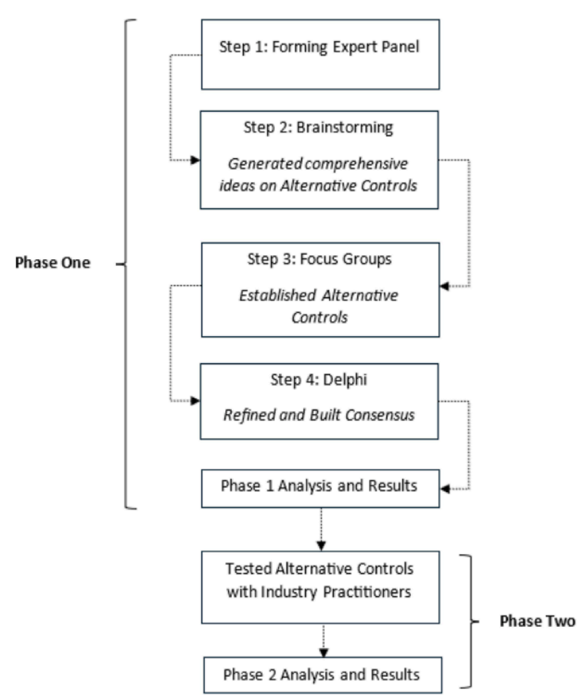

Fig. 3. Overall Research Process.

5. Phase 1 – Method

5.1. Step 1: Expert panel convened

The research began by convening an expert panel of construction safety professionals with extensive experience in the specific topic area (Hallowell & Gambatese, 2010). The panel for this study was assembled from the members of the Construction Safety Research Alliance, following the minimum requirements recommended by Hallowell and Gambatese (2010): (a) a minimum of seven years of professional experience in the construction industry, (b) at least four years of experience in the construction health and safety sector, and (c) a bachelor's degree or higher.

With regard to panel size, 10–50 participants are recommended, but reliable results can be achieved with 10–15 experts (Gambatese et al., 2008). This study exceeded these thresholds by convening twenty-five safety professionals, each meeting all required qualifications and representing a cumulative total of more than 500 years of professional experience. Panel members came from heavy civil, industrial, utilities, energy, and specialty construction sectors across the United States and Canada, ensuring broad exposure to diverse high-energy hazard environments. Table 2 summarizes the panel's background and qualifications.

5.2. Step 2: Brainstorming – Exploring what 'Enough' control looks like

The objective of this step was to initiate a discovery phase with the expert team, to explore a wide range of ideas and establish initial understandings of what makes a control 'enough' for mitigating high-energy hazards when a Direct Control is not feasible. Such situations are often referred to as 'STKY' – situations in which no direct control exists, and the high energy present becomes the 'Stuff That Kills You.' In order to facilitate this process, the panel first collated a library of STKY situations from real-life to form the basis of discussions. The team was provided with a standardized template designed to capture a brief description of the high-energy hazard, an explanation of why a Direct Control was not feasible, and a description of other controls that could be applied.

The situations represented within the case library were drawn from real-life high-energy hazard scenarios, ensuring practical relevance. As shown in Fig. 2, these cases encompassed a diverse range of high-risk hazard sources, including gravity-related hazards such as falls from elevation and suspended loads with workers operating underneath, mobile equipment hazards where workers were in proximity to moving equipment, electrical hazards involving contact with live sources, and motor vehicle incidents. In addition, other high-energy work contexts commonly encountered across construction operations were also represented. This diversity ensured that the evaluation covered a representative spectrum of high-energy hazards faced in practice.

A brainstorming approach was adopted because it is an efficient way to generate a large volume of ideas, foster openness to criticism, encourage creativity, and refine ideas through collaborative input (Bhandari & Hallowell, 2021). The brainstorming process centered on determining what the expert team perceived as 'enough' control for the STKY situations under normal operations. For each STKY situation, the infeasibility of Direct Controls was first established, before various Alternative Controls were discussed to establish a sense from within the team of what makes a control 'enough.' This discussion included careful consideration of all feasible controls for the situation, as well as consideration of how different controls could be combined and layered. Attributes contributing to adequacy, such as the ability to reduce risks from high-energy hazards or compensation for potential mistakes in other controls, were also discussed and identified.

This step yielded valuable insights into why Direct Controls were not feasible, identified an extensive list of explanations for what other controls (or their combinations) could be deemed adequate for each STKY situation, and thus established the groundwork towards developing a set of rules that define what makes a control an acceptable Alternative, or put simply, 'enough.'

Table 2
Expert panel background.

Criteria	Status
Professional experience in the construction industry	
Minimum	12 years
Maximum	30 years
Average	22 years
Professional experience in the health and safety sector in construction industry	
Minimum	9 years
Maximum	30 years
Average	16.3 years
Educational background	
Bachelor's degree	57 %
Master's degree	43 %
Other	
Professional registration such as Certified Safety Professional or any other	78 %
Member or chair of a nationally recognized committee	52 %

5.3. Step 3: Focus groups – Establishing rules for 'Enough' control

The purpose of this step was to further evaluate the output from the brainstorming session and establish a preliminary set of rules to describe and define what makes an acceptable Alternative Controls. To achieve this, the panel shifted to a focus group model. A focus group involves face-to-face discussions managed by one or more facilitators who encourage conversation through questions and interjections (Bhandari & Hallowell, 2021). Through two focus groups (the second continuing the work of the first), the panel analyzed ten new STKY situations taken from the library, distinct from those explored in the brainstorming phase. These new scenarios were reviewed in-depth, and the ideas generated in Step 2 were carefully considered to develop and establish rules as to the adequacy of different controls in each scenario. For example, discussions covered considerations of when a control can be deemed 'enough,' which controls can or cannot ever be considered adequate, how controls can be layered and in what combinations, and initial categories these controls could be grouped into. During these discussions, the panel also identified the need for a more basic definitional boundary to distinguish measures that directly address the high-energy hazard itself from measures that support work more generally. A key output from this step was therefore the definition of Energy Control itself. The team defined Energy Control as follows:

- Timely: It must be in place during active work when a high-energy hazard is present.
- Tangible: It must be physically present at the worksite during the work.
- Targeted: It must be specifically installed to address the high-energy hazard.

Examples of Energy Controls include physical barriers, fall arrest systems, and de-energizing methods (e.g., lockout/tagout). In contrast, examples of controls that do not qualify as Energy Controls are rules, safe work plans, and training.

In this study, Energy Control is used as a broader definitional concept rather than as a separate category of control. Specifically, it was used to distinguish measures that directly address the high-energy hazard from supporting measures that are important but do not directly control the hazard itself. This distinction was important in the development of the framework because it clarified which measures could count toward adequacy and which could not. Direct controls are a specific type of Energy Control that target high-energy hazards and remain effective despite unintentional human error during the work. Alternative Controls, by contrast, are Energy Controls used only when a Direct Control is not feasible and only when they meet the criteria defined in this study.

At the conclusion of the two focus groups, the expert panel reached an 'agreement-in-principle' on an initial set of rules that constituted acceptable Alternative Controls. The output of the focus groups was the first iteration of the definition and explicit criteria (with categories and rules) for Alternative Controls. This preliminary definition and criteria were then taken forward for further refinement and consensus building in the subsequent step.

5.4. Step 4: Delphi – Testing and securing consensus

The aim of this step was to validate and refine the definition and criteria through further testing, to secure consensus among the panel. To achieve this goal, the Delphi method was used. Delphi is frequently mobilized as a systematic approach to achieve reliable consensus among experts through multiple rounds of rating followed by indirect discussion (Hallowell & Gambatese, 2010; Chan et al., 2001). The standard approach of a Delphi study includes: developing questions for expert evaluation; collecting responses; evaluating consensus; reporting anonymous feedback to the experts; and iterating the process in multiple rounds until the desired consensus is achieved (Hallowell & Gambatese, 2010; Erkal et al., 2021).

In this study, the method was modified to allow for greater flexibility and interaction among the experts. Rather than enforcing rigid communication, structured discussions were facilitated by a moderator to enable more open and collaborative exchanges. This approach was essential for refining the definition and criteria of the Alternative Controls, allowing the collection of feedback during discussions and the iterative reevaluation of questions in subsequent rounds. Although flexibility in communication allows for unencumbered discussions, it can also result in a high susceptibility to cognitive biases (Chan et al., 2001). In this study, potential biases were identified and addressed by following the mitigation guidelines from Bhandari and Hallowell (2021), which are set out in Table 3.

As in traditional Delphi applications, achieving consensus among the expert panelists required multiple rounds of ratings, discussions, and modifications. Following the guidelines provided in Chan et al. (2001), and Hallowell and Gambatese (2010), this study selected three rounds of Delphi. To measure consensus, statistical methods like standard deviation and Kendall's coefficient of concordance can be used (Hallowell & Gambatese, 2010; Chan et al., 2001). This study applied a strict threshold, considering consensus achieved only if no more than one participant disagreed with the proposed definition of Alternative Controls. If more than one disagreed, consensus was not reached. This threshold was pre-established and agreed upon before the Delphi process began.

The number of participants varied across the three rounds of the Delphi process, with 16 in the first round, 14 in the second, and 12 in the third. This variation was due to participant availability for subsequent rounds. Nevertheless, the number of participants never fell below 12, meeting the minimum recommended threshold for Delphi studies (Hallowell & Gambatese, 2010). However, the effect of this was mitigated within this study by not removing questions where consensus was achieved in previous rounds.

Although the Delphi was conducted in a face-to-face workshop, it was managed online via Qualtrics to collect responses from the expert panels while ensuring some anonymity. Participants were presented with nine new real-life STKY scenarios (cases encompassed a diverse range of hazard sources, including gravity-related hazards such as falls from elevation and suspended loads with workers operating underneath, mobile equipment hazards where workers were in proximity to moving equipment, electrical hazards involving contact with live sources, and motor vehicle incidents etc.) and again each was distinct from those used in the brainstorming and focus group phases. For each scenario, participants were given three different controls or combinations of controls and were asked to evaluate whether they were enough for each specific STKY situation when applying the definition and criteria developed in Step 3. An example of one such STKY situation is provided in Fig. 4. The three controls for each scenario were deliberately designed to vary in their alignment with the preliminary definition of Alternative Controls: one did not meet the criteria, one partially met them, and one fully met them. This approach enabled meaningful discussions among the expert team, identifying gaps, inconsistencies, and areas of disagreement within the definition, allowing the expert panel to refine the rules and definitions and ensure that the final version reflected a shared consensus. The analysis and outcomes of this phase are detailed in the next section.

Table 3
Controls for bias.

Bias	Control
Collective unconscious	Voicing all opinions during each round of the feedback discussion
Contrast effect and/or Primary effect	Randomizing the order of questions and cases for each panel member and for each round
Von Restorff effect	Conducting multiple rounds of surveys
Myside bias	Including the expert's panel discussion/reasons in the controlled feedback delivery
Recency effect	Remove individuals who have provided the cases; conducting multiple rounds of surveys
Dominance	Anonymous reporting and feedback of responses

5.5. Phase 1 – Analysis and results

Steps 1 through 3 of Phase 1 facilitated the establishment and development of the definition and criteria for acceptable Alternative Controls for STKY situations. To ensure expert consensus on the definition and criteria, Step 4 mobilized the Delphi study explained above, with consensus calculated and determined as follows: (a) calculating the total number of responses for each control (defined as n); (b) setting the consensus threshold criteria by calculating it as $n-1$, where the threshold was adjusted based on the total number of responses for each case and question due to variations in the number of respondents during each phase; (c) obtaining the total count of 'Yes' (coded as 1) and 'No' (coded as 0) for each control; (d) finding the maximum value (m) between 'Yes' and 'No'; (e) and finally consensus on the adequacy of a control was considered achieved if m (the maximum of the counts of 'Yes' or 'No') was greater than or equal to the threshold $n-1$. An example of the calculation is provided for each of the three controls in case 1 in Table 4.

Since each control was designed to target specific categories or rules of Alternative Controls, for any control where agreement was not achieved, the moderator discussed it with the expert team (considering cognitive bias controls) to gather feedback and refine the specific rules or categories to capture and address the necessary refinements. For example, one defined output from the focus group was that safety signage cannot be considered Alternative Controls unless it informs of something unusual, hidden, or unexpected, and provides critical information about a high-energy hazard that is not visible (e.g., electricity). However, through the modified Delphi process, this rule was refined further: signage can be considered part of an Alternative Controls if: (a) it communicates information about something unusual, hidden, or unexpected, providing critical details regarding a high-energy hazard that is not visibly apparent (e.g., electricity); and (b) it includes writing, symbols, or visuals that ensure complete understanding, taking into account the preferred language and literacy of all exposed employees. Similarly, where the initial focus groups labelled the categories of Alternative Controls as (A), (B), and (C), through the Delphi more specific titles were assigned to better describe these categories: (a) Physical Obstacles, (b) Dedicated Monitoring, and (c) Visual Reminder. This

Description	
Three excavators being used to layout lining for berms for containment cell. All three grab the lining with their bucket and walk backwards with it in the bucket, ground personnel have radios and communicate before each move forward or backward. Personnel must place lining in bucket of excavator and ensure operator claps down on right spot of lining with skids or lining will tear when pulling.	
For the above situation, please answer/select your options for below question:	
Controls	**Is the control enough (by the Framework provided)?**
Control 1: Spotter	Yes / No
Control 2: Painted Lines and Spotter	Yes / No
Control 3: Marker Cones and Proximity Detection Device on the Machines	Yes / No

Fig. 4. Example STKY situation.

Table 4
Example of Delphi calculation.

Case 1 Participant	Control 1	Control 2	Control 3
1	0	0	1
2	1	0	1
3	1	0	1
4	1	0	1
5	1	0	1
6	0	0	1
7	1	0	1
8	1	0	1
9	1	0	1
10	0	0	1
11	1	0	1
12	1	0	1
13	0	0	1
14	1	0	0
15	0	0	1
16	1	0	1
Count *(n)*	16	16	16
Threshold *(n-1)*	15	15	15
Count of 'Yes' *(coded as 1)*	11	0	15
Count of 'No' *(coded as 0)*	5	16	1
Maximum of Yes/No *(m)*	11	16	15
Consensus *(achieved if m ≥ n-1)*	No	Yes	Yes

process was repeated for all cases and for each round. Upon completion of the third round, the final results were reported to the participants. A summary of the consensus achieved in each round is shown in Table 5. Ultimately, the outcome of this process was a refined and agreed definition and criteria, consisting of a scientific definition of Alternative Controls, their categorization into three types, and explicit criteria for their application in work tasks involving high-energy hazards where no Direct Control is available.

Table 5
Delphi rounds.

Survey Rounds	Total number of questions / controls	Number of control / questions consensus achieved on
Round 1	27	16
Round 2	27	20
Round 3	27	27

5.6. *Definition and criteria for acceptable alternative controls*

The expert panel defined three distinct categories of acceptable Alternative Controls:

Physical Obstacle: an obstruction that blocks the path or hinders progress toward a high-energy hazard. Examples include cones with toppers, danger tape, and fence around equipment.
Dedicated Monitoring: devoted and continuous attention to the high-energy hazard. Examples include monitoring alarms, proximity detection devices with two-way warnings, and a gas monitor.
Visual Reminder: a visible warning of the presence of high-energy hazard. Examples include signage, temporary traffic lights, painted lines, etc.

To be counted as acceptable Alternative Controls, the following rules must be met.

When a Direct Control is not feasible, there must be at least two Alternative Controls, and they must represent different categories (Physical Obstacle, Dedicated Monitoring, or Visual Reminder). This minimum threshold reduces coupling of human-error pathways in high-energy contexts (Winge & Albrechtsen, 2018).
Alternative Controls must specifically target the high-energy hazard (intentionally and not incidentally). That said, when considering Alternative Controls for high-energy hazards, they should be specifically designed for the high-energy hazard so that they are effective.
All workers creating or exposed to the hazard must have the following that are directly related to the high-energy hazard and corresponding categories of the Alternative Controls:

a. Verified specialty training related to the hazard and application of the controls (beyond general training)
b. Safe Work Plan
c. Quality Pre-Job Brief

It is necessary to emphasize that training, safe work plans, and quality pre-job briefs must explicitly address relevant high-energy hazards and Alternative Controls. For example, a safe work plan should clearly outline the high-energy hazards related to the tasks, including the specific controls being implemented. Additionally, workers should receive the required specialty training to safely perform these tasks. Furthermore, during the quality pre-job brief, high-energy hazards present and the designated Alternative Controls must be explicitly discussed to ensure proper understanding and implementation.

The following, while critical to overall safety management, do not by themselves meet the definitional criteria for Energy Control and are therefore not counted as Alternative Controls:

a. Experience
b. Rules
c. Situational awareness of the person at risk (i.e., a spotter is not covered by this rule).

More importantly, to be considered as Alternative Controls, the following specific definitions must also be met:

Spotters: must be individuals whose sole duty during the period of exposure to an existing or potential high-energy hazard is the continuous monitoring of all interactions with the high-energy hazard. Spotters must also have continuous communication with all workers involved in the creation of and/or exposure to the high-energy hazard (i.e., both the machine driver and the adjacent worker on foot). Spotters may not be exposed to energy hazards equal to or greater than 1,500 J without either their own Direct Control or Alternative Control measures in place.

Flaggers: Flaggers must provide dedicated attention to the management of a specific high-energy hazard.

Communication: must be clear, unambiguous (i.e., in a shared preferred language), uninterrupted, specific to the high-energy hazards, and communicated via a clear and dedicated device/channel. This level of communication must include all workers involved in the creation of and/or direct exposure* to the high-energy hazards (e.g., crane operator and all workers receiving the suspended load). **within 6ft.*

Signage: Signs must adhere to the following criteria:

a. Signage must communicate information about something unusual, hidden, or unexpected, providing critical details regarding a high-energy hazard that is not visibly apparent (e.g., electricity).
b. Signage must include writing, symbols, or visuals that ensure complete understanding, taking into account the preferred language and literacy of all exposed employees.

Lastly, the following criteria were deemed essential for Alternative Controls to be effective:

a. They are relevant only when implemented during the execution of the work.

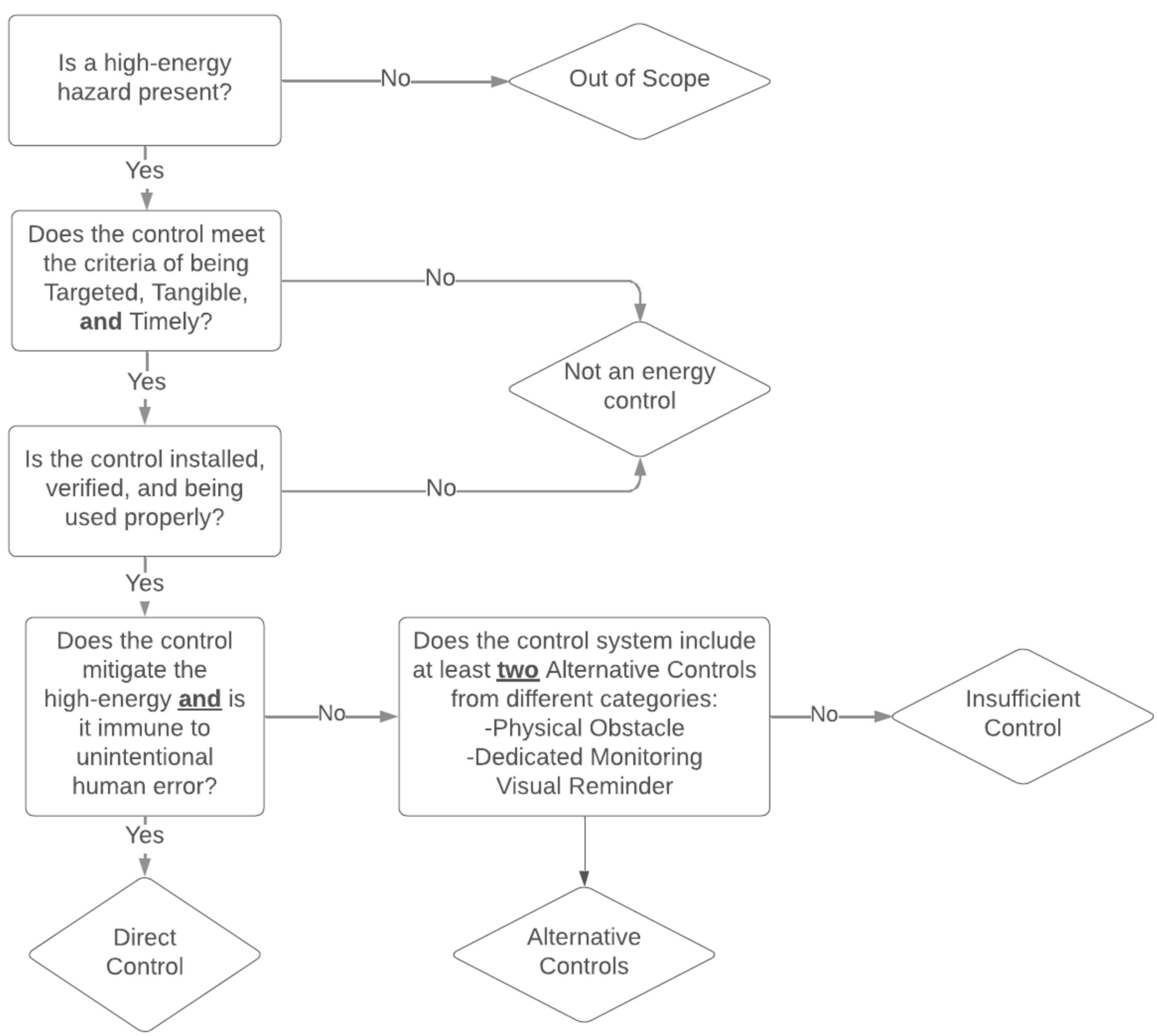

Fig. 5. Model for evaluating controls for high-energy hazards using the definitions of Energy Control, Direct Control, and Alternative Controls. Adapted from Hallowell (2025)

b. They are primarily designed to mitigate human error rather than solely reduce the energy magnitude.

To further clarify the relationship between Energy Controls, Direct Controls, and Alternative Controls, Fig. 5 presents the decision logic for evaluating controls in work involving high-energy hazards. Specifically, the figure shows how the framework begins by determining whether a high-energy hazard is present and then whether the candidate control meets the criteria of being targeted, tangible, and timely. It then distinguishes between controls that meet the previously established definition of Direct Control and those that, when a Direct Control is not feasible, may be accepted as Alternative Controls if they satisfy the criteria defined in this study.

The effectiveness of controls for high-energy hazards depends not only on their technical characteristics, but also on the stage of work at which they are selected and integrated. Opportunities to eliminate or reduce exposure to high-energy hazards arise earliest in design, planning, work methods, site layout, sequencing, and resource allocation (Hallowell, 2025). As illustrated in Fig. 6, the opportunity to implement higher-integrity controls is generally greatest earlier in the work process and declines as work progresses toward active execution. Accordingly, where feasible, Direct Controls and Alternative Controls should be identified early enough to be incorporated into the work plan before mobilization. During execution, however, the adequacy of the control package must be judged against the specific task, work area, and conditions as the work will actually be performed. Where a Direct Control exists and is feasible, its use is required. Where no Direct Control exists, or where infeasibility has been established for the specific work condition, the proposed Alternative Controls package should be approved before work begins and then verified in the field immediately before exposure. In addition, the approved controls should remain physically present, suitable at the point of risk, and monitored for loss of effectiveness. If conditions change such that exposure to a high-energy hazard is created, or if an approved control is no longer present or effective, the work should be replanned before reauthorization.

The framework is intended primarily as a decision-making aid for judging the adequacy of controls for high-energy hazards when a Direct Control is not feasible. It can be applied at organizational, project, and field levels, although its use differs across those levels. At the organizational level, it can inform standards, planning expectations, and training related to high-energy work. At the project level, it can inform work plans, methods, task preparation, and the selection of candidate controls. These upstream decisions are important, but they do not by themselves establish that control is adequate for the work. The final judgment of adequacy must be made against the actual work condition before exposure begins, because the acceptability of the control package depends on the specific configuration of the task, the work area, the workers involved, and the conditions at the point of risk. Where planned conditions and actual conditions do not align, work should not proceed until the control package is reviewed and confirmed for the task as it will actually be performed.

The effectiveness of Alternative Controls also depends on whether creating and exposed workers understand the specific high-energy hazard, the designated controls, and the conditions under which work must not proceed. Accordingly, the safe work plan and pre-job brief should communicate the project-specific controls in explicit, task-specific terms. At a minimum, this communication should make clear the hazard being addressed, the controls required for the work, the conditions indicating that those controls are in place and functioning as intended, and the circumstances requiring the work to be stopped and replanned. Responsibility for maintaining the controls during the work should also be clearly identified. In addition, before exposure begins, understanding should be confirmed through a task-specific check tied to the designated controls rather than through a general acknowledgment of understanding.

6. Phase 2 – Noise experiment

The final stage of the study sought to validate the applicability and usability of the acceptable Alternative Controls produced in Phase 1. To this end, an experiment was designed to test the extent to which the definition and criteria of the acceptable Alternative Controls reduced variability (noise) and improved consistency in decision-making when applied in STKY situations, taking repeated measures into account to isolate the impact of the definition and criteria (Bayona et al., 2023).

A new set of STKY situations was prepared, and a separate expert panel of safety practitioners was convened. Importantly, these practitioners had not participated in Phase 1 of the project and were therefore entirely unfamiliar with the definition and criteria developed during that stage. This design minimized bias and provided a more robust test of the applicability of the definition and criteria. The noise experiment involved 20 safety professionals from both the United States and Canada, representing oil and gas, electric power generation, and infrastructure and building construction. The purpose of the experiment was to quantify baseline variability and assess whether the definition and criteria reduced inconsistency among practitioners when evaluating

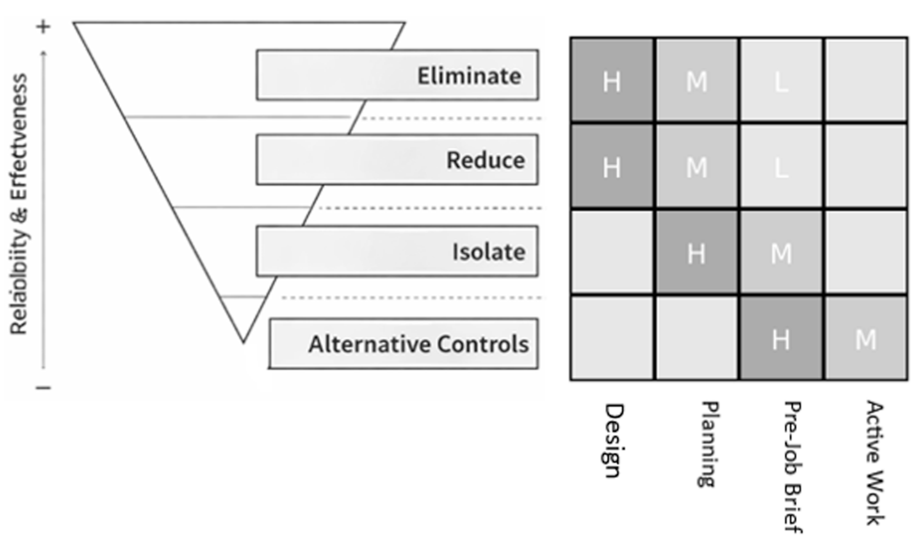

Fig. 6. Relative opportunity to implement Direct and Alternative Controls across stages of work delivery (adapted from Hallowell, 2025). Note: H = high; M = medium; L = low.

high-energy hazard scenarios.

The experiment encompassed three stages: (a) a pre-survey where participants used their existing knowledge to assess whether the controls provided for the STKY cases were adequate. This phase was used to determine a baseline level of noise among the survey participants; (b) introduction of the definition and criteria for acceptable Alternative Controls and training the practitioners; (c) a post-training assessment where participants repeated the pre-survey with different STKY situations using the definition and criteria. In both surveys, participants answered two main questions: "Is this control (or combination of controls) adequate or not?" and "On a scale of 1 to 5, how confident are you in your response?" The training included a detailed explanation of the three categories of Alternative Controls, the adequacy rules, and the specific definitions for spotters, flaggers, communication, and signage. A worked example illustrated how to apply the definition and criteria to a STKY case that was not used in the survey.

A total of 10 STKY situations were used in the experiment. Using a counterbalanced design (as shown in Fig. 7), the 10 cases were split into two surveys (A and B) with five cases each, and the order of cases within each survey was randomized. In the pre-survey, participants were randomly assigned to one of two groups: Group A received survey A, while Group B received survey B. In the post-survey, the cases were crossed over between the groups, with Group A receiving survey B and Group B receiving survey A, as illustrated in Fig. 7. This approach was used to control for ordering effects and to reduce potential cognitive biases among participants (Perreault, 1975).

6.1. Phase 2 – Analysis and results

The aim of Phase 2 of this study was to evaluate whether the definition and criteria and their associated rules could reduce variability and improve consistency when applying acceptable Alternative Controls within STKY situations. To this end, noise was quantified as the variability in responses to the same case, measured using variance — a higher variance indicates greater noise. Thus, the variance for each control within each case was measured for both the pre-survey and the post-survey. The pre-survey established a baseline for empirically measuring noise levels based on participants' existing understanding of applying adequate controls for STKY situations. The findings revealed a high degree of variability in the initial responses. However, after the introduction and training on the definition and criteria, response variability decreased significantly, with the average variance dropping from 0.17 (pre-survey) to 0.09 (post-survey), representing a 47% reduction in noise. Consistency between the experimental groups was statistically compared using the Fleiss' kappa statistic, which assesses the reliability of agreement among raters while accounting for agreement expected by chance (Nichols et al., 2010). Put simply, Fleiss' kappa assesses inter-rater reliability, quantifying the level of agreement between raters and providing a standardized metric for evaluating consistency (Falotico & Quatto, 2015). A higher value reflects stronger agreement and reduced variability, or noise, among raters' assessments (Kahneman et al., 2021). According to Landis and Koch (1977), the interpretation of Fleiss' kappa values ranges as follows: 0.00–0.20 slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; 0.81–1.00 almost perfect agreement.

Findings show a substantial increase in agreement among participants after training on the definition and criteria. Specifically, Fleiss' kappa in Survey A rose from 0.21 (95% CI 0.14–0.29; fair agreement) in the pre-survey to 0.61 (95% CI 0.53–0.68; substantial) in the post-survey, and in Survey B from 0.32 (95% CI 0.24–0.39) to 0.62 (95% CI 0.55–0.70), both reflecting substantial agreement post-training. The non-overlapping 95% CIs indicate a clear, precise reduction in decision noise. These results are shown in Table 6.

Additionally, participants were asked to rate their confidence level when assessing the cases in both the pre-survey and post-survey. The purpose was to examine how the definition and criteria influenced participant confidence—that is, did they make participants feel more confident when implementing Alternative Controls for high-energy hazards? To assess this, a paired *t*-test was conducted to compare the difference in mean confidence levels for the same group of participants at two points (before and after the intervention). The results showed a statistically significant increase in confidence from the pre-survey ($M = 3.993$, $SDdiff = 0.311$) to the post-survey ($M = 4.336$, $SDdiff = 0.311$), with a mean difference of 0.343, $t(19) = 4.937$, $p < 0.001$, 95% CI [0.198, 0.489]. The effect size, as measured by Cohen's d, was 1.10, indicating a large effect. The results of the paired *t*-test are shown in Table 7.

7. Limitations

As with all research, this study has several limitations. Primarily, the data collected from the expert panel were qualitative in nature and reliant on the expertise and experience of the participants. However, the panel members were vetted based on established qualifications for service (Hallowell & Gambatese, 2010), which ensured a high level of professional credibility. Additionally, while established methodologies such as Delphi were incorporated in the process, the reliance on expert opinion carries the risk of cognitive biases. Efforts were made to minimize these biases through structured processes, such as anonymous feedback collection and consensus-based validation (Bhandari & Molenaar, 2020). Nevertheless, the inherent complexity of cognitive biases means that some degree of influence may still have occurred.

Ideally, the Delphi method should be implemented with controlled feedback and discussion provided over multiple iterative rounds to progressively build consensus. However, due to the goals of this study, the method was adapted to allow greater flexibility and interaction among experts. While this modification supported richer discussion, it also introduced the potential for biases. To mitigate this, potential sources of bias were identified in advance, and controls were implemented following validated protocols (Bhandari & Molenaar, 2020; Hallowell & Gambatese, 2010).

An additional limitation is that the study was developed and tested

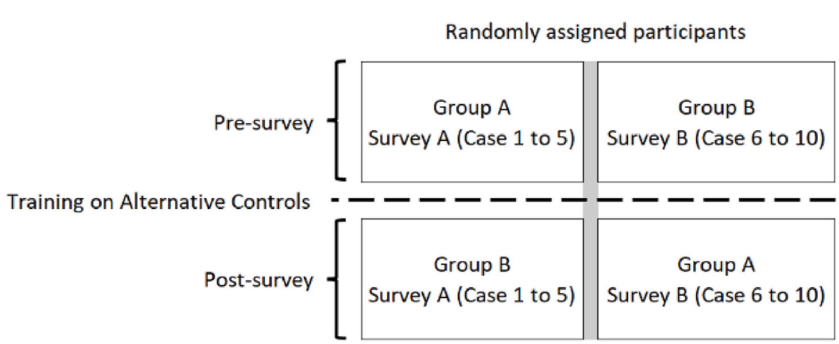

Fig. 7. Noise Experiment Design.

Table 6
Fleiss' kappa values.

Survey	κ (Pre)	95 % CI (Pre)	Agreement (Pre)	κ (Post)	95 % CI (Post)	Agreement (Post)
A	0.21	0.14–0.29	Fair	0.61	0.53–0.68	Substantial
B	0.32	0.24–0.39	Fair	0.62	0.55–0.70	Substantial

Table 7
t-test results.

Parameter	Statistical Test	Pre-survey mean score [/5] (a)	Post-survey mean score [/5] (b)	Difference (b-a)	p-value	95 % Confidence Interval	
						Lower	Upper
Confidence	Paired *t*-test (N = 20)	3.993	4.336	0.343	<0.001	0.198	0.489

within North American construction settings. As a result, the definition and criteria may reflect assumptions tied to this context, including regulatory expectations, common work organization patterns, language practices, and the types of high-energy hazards commonly encountered in these operations. Although the framework may have relevance beyond North American construction, transfer to other industries, countries, or regulatory systems should not be assumed and requires further validation in those settings.

Finally, while the study's findings were validated through testing with safety professionals and real-world cases, generalizability may be limited by the sample size and sectoral coverage. Validation with a larger sample size across diverse operational environments would be beneficial to strengthen the applicability of the definition of Alternative Controls. Importantly, this study focused on evaluating decision-making consistency rather than directly measuring reductions in SIFs. Future research should extend this work by examining how the use of Alternative Controls definition influences actual safety outcomes in practice, particularly its effectiveness in reducing SIFs. Such studies will be critical to solidify the findings and extend their relevance across the construction industry and beyond.

8. Discussion

This study set out to address a practical problem that remains unresolved in construction safety: how to judge whether a control is "enough" when work involves a high-energy hazard and a Direct Control is not feasible. The results indicate that the framework developed in this study improves the consistency with which practitioners make those judgments. Across the noise experiment, participants showed lower response variability, stronger inter-rater agreement, and higher confidence after training on the definition and criteria for Alternative Controls. These findings suggest that the framework functions effectively as a decision aid for judging adequacy in high-energy scenarios, particularly in situations where practice would otherwise rely on broad rules, local interpretation, or individual experience alone.

The practical implication of this result is not that the framework has been shown to reduce SIFs directly, but that it provides a more structured and consistent basis for decision-making in work involving high-energy hazards. In practice, this may improve the quality of planning and field-level judgment by helping practitioners distinguish between measures that directly address the high-energy hazard and measures that remain important but do not count as acceptable Alternative Controls by themselves. In this respect, the contribution of the study lies in improving the clarity and consistency of adequacy judgments, which is a necessary precondition for adequate control application, but not the same thing as demonstrating improved safety outcomes.

A second contribution of the study is conceptual. In developing the definition and criteria for Alternative Controls, the research also clarified a more basic boundary between measures that directly engage the high-energy hazard and measures that support work more generally. This was captured through the concept of Energy Control and helped define what could and could not count toward an acceptable Alternative Controls arrangement. This distinction is important because one of the persistent problems in safety practice is that nearly any safety-related activity can be labelled a "control," even when it does not directly act on the hazard itself. By clarifying this boundary, the framework helps preserve focus on measures that are intended to control the high-energy hazard at the point of risk.

The framework should also be understood within clear boundary conditions. It is intended primarily as a decision-making aid for judging the adequacy of Alternative Controls for high-energy hazards when a Direct Control is not feasible. Although it can inform decisions at organizational, project, and field levels, it is not a substitute for broader risk management processes, organizational safety systems, formal risk assessment, policy, supervision, or project governance. Rather, it provides a more structured basis for one specific judgment problem within those broader systems: whether a proposed Alternative Controls arrangement is acceptable for the work as it will actually be performed. In this sense, the framework is best understood as complementary to existing safety management approaches rather than as a replacement for them. It addresses a specific judgment problem within high-energy work; it does not, by itself, resolve the broader organizational, contractual, or regulatory conditions that shape how controls are selected and implemented in practice.

For this reason, the findings should also be interpreted with caution. This study did not evaluate changes in exposure, incident rates, or SIF outcomes. Rather, it evaluated whether the framework reduced decision variability and improved agreement and confidence in adequacy judgments. The results therefore support the claim that the framework improves decision consistency under the conditions tested, but they should not be interpreted as direct evidence that the framework reduces SIFs in practice. Whether more consistent adequacy judgments lead to stronger implementation of controls and ultimately to better safety outcomes remains an important question for future research. Further work should therefore examine how the framework performs across a wider range of settings and whether its use is associated with measurable improvements in control quality, exposure reduction, and safety outcomes over time.

9. Conclusion

This study developed and tested a structured definition and explicit criteria for acceptable Alternative Controls when a Direct Control is not feasible for a high-energy hazard. The results show that the framework reduces decision variability, improves agreement among practitioners, and increases confidence in adequacy judgments. These findings support its value as a structured basis for task-level decision-making in high-energy work. At the same time, the study does not establish that the framework reduces SIFs directly; rather, it shows that it improves the consistency with which practitioners judge what constitutes acceptable control in these scenarios. Future research should test the framework across a broader range of settings and examine whether its use is associated with stronger implementation and improved safety outcomes in practice.

CRediT authorship contribution statement

Yaqoob Raheemy: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fred Sherratt:** Writing – review & editing, Supervision, Investigation, Conceptualization. **Matthew R. Hallowell:** Writing – review & editing, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Construction Safety Research Alliance, based at the University of Colorado Boulder, USA. The authors would like to thank everyone who contributed to the research that generated the data for this paper.

References

Ahmadi, O., Mortazavi, S. B., & Asilian Mahabadi, H. (2021). Application and modification of the Tripod Beta method for analyzing the causes of oil and gas industry accidents. *International Journal of Occupational Safety and Ergonomics, 27*(3), 928–937. https://doi.org/10.1080/10803548.2019.1693167

Ajslev, J. Z. N., Møller, J. L., Andersen, M. F., et al. (2022). The Hierarchy of Controls as an Approach to Visualize the Impact of Occupational Safety and Health Coordination. *International Journal of Environmental Research and Public Health, 19*(5), 2731. https://doi.org/10.3390/ijerph19052731

Ale, B. J. M., Hartford, D. N. D., & Slater, D. (2015). ALARP and CBA all in the same game. *Safety Science, 76*, 90–100. https://doi.org/10.1016/j.ssci.2015.02.012

Barnett, R. L. (2020). On the Safety Hierarchy and Hierarchy of Controls. *Am. J. Mech. Eng., 8*(2), 61–68.

Barnett, R. L., & Brickman, D. B. (1986). Safety hierarchy. *Journal of Safety Research, 17*(2), 49–55. https://doi.org/10.1016/0022-4375(86)90093-9

Baybutt, P. (2014). The ALARP principle in process safety. *Process Safety Progress, 33*(1), 36–40. https://doi.org/10.1002/prs.11599

Bayona, A., Bhandari, S., Hallowell, M., et al. (2023). What is a serious injury? A model for defining serious injuries and fatalities. *Professional Safety, 68*(9), 22–30.

Bayona, A., Hallowell, M. R., & Bhandari, S. (2024). The things that hurt people are not the same as the things that kill people: Key differences in the proximal causes of low- and high-severity construction injuries. *Journal of Construction Engineering Management, 150*(8), Article 04024089. https://doi.org/10.1061/JCEMD4.COENG-14545

Bellamy, L. J. (2015). Exploring the relationship between major hazard, fatal and non-fatal accidents through outcomes and causes. *Safety Science, 71*, 93–103. https://doi.org/10.1016/j.ssci.2014.02.009

Bhandari, S., & Hallowell, M. R. (2021). Identifying and controlling biases in expert-opinion research: Guidelines for variations of Delphi, nominal group technique, and focus groups. *Journal of Management in Engineering, 37*(3), Article 04021015. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000909

Bhandari, S., & Molenaar, K. R. (2020). Using debiasing strategies to manage cognitive biases in construction risk management: Recommendations for practice and future research. *Practice Periodical on Structural Design and Construction, 25*(4), Article 04020033. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000521

Bowles, D. S. (2003). ALARP evaluation: using cost effectiveness and disproportionality to justify risk reduction.

Busch, C., Usrey, C., Loud, J., et al. (2021). Serious injuries & fatalities: Why are they constant while injury rates decrease? *Professional Safety, 66*(1), 26–31.

Centers for Disease Control and Prevention (2024) About hierarchy of controls. National Institute for Occupational Safety and Health (NIOSH). Available at: https://www.cdc.gov/niosh/hierarchy-of-controls/about/index.html (Accessed: 10 March 2026).

Chan, A. P. C., Yung, E. H. K., Lam, P. T. I., et al. (2001). Application of Delphi method in selection of procurement systems for construction projects. *Constr. Manage. Econ., 19*(7), 699–718. https://doi.org/10.1080/01446190110066128

Dekker, S. (2014). *Safety Differently: Human Factors for a New Era* (2nd ed.). CRC Press.

De Ruijter, A., and Guldenmund, F., 2016. The bowtie method: A review. Safety Science, 88, pp.211–218. Available at: https://doi.org/10.1016/j.ssci.2016.03.001.

Erkal, E. D. O., & Hallowell, M. R. (2023). Moving beyond TRIR: Measuring and monitoring safety performance with high-energy control assessments. *Professional Safety, 68*(5).

Erkal, E. D. O., Hallowell, M. R., & Bhandari, S. (2021). Practical assessment of potential predictors of serious injuries and fatalities in construction. *Journal of Construction Engineering Management, 147*(10), Article 04021129. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002146

Falotico, R., & Quatto, P. (2015). Fleiss' kappa statistic without paradoxes. *Quality and Quantity, 49*(2), 463–470. https://doi.org/10.1007/s11135-014-0003-1

Gambatese, J. A., Behm, M., & Rajendran, S. (2008). Design's role in construction accident causality and prevention: Perspectives from an expert panel. *Safety Science, 46*(4), 675–691. https://doi.org/10.1016/j.ssci.2007.06.010

Ge, J., Zhang, Y., Chen, S., et al. (2022). Accident causation models developed in China between 1978 and 2018: Review and comparison. *Safety Science, 148*, Article 105653. https://doi.org/10.1016/j.ssci.2021.105653

Hayes, J. (2012). Use of safety barriers in operational safety decision making. *Safety Science, 50*(3), 424–432. https://doi.org/10.1016/j.ssci.2011.10.002.

Hallowell, M. R. (2025). *Energy-based Safety: A Scientific Approach to Preventing Serious Injuries and fatalities (SIFs)*. CRC Press.

Hallowell, M. (2023). *Safety Classification and Learning (SCL) Model. United States of America*. Edison Electric Institute (EEI).

Hallowell, M. R., Alexander, D., & Gambatese, J. A. (2017). Energy-based safety risk assessment: Does magnitude and intensity of energy predict injury severity? *Constr. Manage. Econ., 35*(1–2), 64–77. https://doi.org/10.1080/01446193.2016.1274418

Hallowell, M. R., & Gambatese, J. A. (2009). Construction Safety Risk Mitigation. *Journal of Construction Engineering Management, 135*(12), 1316–1323. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000107

Hallowell, M. R., & Gambatese, J. A. (2010). Qualitative research: Application of the Delphi method to CEM research. *Journal of Construction Engineering Management, 136*(1), 99–107. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000137

Heinrich, H. W. (1931). *Industrial accident prevention: A scientific approach* (1st ed.). New York, NY; London, UK: McGraw-Hill Book Company.

Hoffman, I., and Wilkinson, P. 2011. "The barrier-based system for major accident prevention: A system dynamics analysis." In Proc., 29th Int. Conf. of the System Dynamics Society, 25–29. Washington, DC, USA.

HSE (Health and Safety Executive). 2024. Managing risks and risk assessment at work. Health and Safety Executive. https://www.hse.gov.uk/ppe/managing-risk-using-ppe.htm.

Ho, P. (2023). Do safety cases demonstrate risks have been reduced so far as is reasonably practicable? An Australian study examining the methods of presenting safety cases. *Safety science, 159*, Article 106042.

Hopkins, A. (2011). Risk-management and rule-compliance: Decision-making in hazardous industries. *Safety Science, 49*(2), 110–120.

Hurst, J., McIntyre, J., Tamauchi, Y., et al. (2019). A summary of the 'ALARP' principle and associated thinking. *Journal of Nuclear Science and Technology, 56*(2), 241–253. https://doi.org/10.1080/00223131.2018.1551814

ICMM (International Council on Mining and Metals), 2015. Health and Safety Critical Control Management Good Practice Guide. London, UK: ICMM. https://www.icmm.com/en-gb/publications/health-and-safety/ccm-good-practice-guide.

Jones-Lee, M., & Aven, T. (2011). ALARP—What does it really mean? *Reliability Engineering and System Safety, 96*(8), 877–882. https://doi.org/10.1016/j.ress.2011.02.006

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. New York, NY: Hachette.

Khan, F., Rathnayaka, S., & Ahmed, S. (2015). Methods and models in process safety and risk management: Past, present and future. *Process Safety and Environmental Protection, 98*, 116–147. https://doi.org/10.1016/j.psep.2015.07.005

Kotre, C. J. (2022). ALARP: When does reasonably practicable become rather pricey? *The British Journal of Radiology, 95*(1138), Article 20220612.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159. https://doi.org/10.2307/2529310

Li, Y., & Guldenmund, F. W. (2018). Safety management systems: A broad overview of the literature. *Safety Science, 103*, 94–123. https://doi.org/10.1016/j.ssci.2017.11.016

Liu, R., Liu, H. C., Shi, H., & Gu, X. (2023). Occupational health and safety risk assessment: A systematic literature review of models, methods, and applications. *Safety science, 160*, Article 106050.

Lyon, B. K., & Popov, G. (2016). The art of assessing risk: Selecting, modifying & combining methods to assess operational risks. *Professional Safety, 61*(3), 40.

Lyon, B. K., & Popov, G. (2019). Risk treatment strategies: Harmonizing the hierarchy of controls and inherently safer design concepts. *Professional Safety, 64*(5), 34–43.

Lyon, B. K., & Popov, G. (2020). Managing risk through layers of control. *Professional Safety, 65*(4), 25–35.

Marhavilas, P. K., & Koulouriotis, D. E. (2021). Risk-acceptance criteria in occupational health and safety risk-assessment—the state-of-the-art through a systematic literature review. *Safety, 7*(4), 77.

Maselli, G., Macchiaroli, M., & Nesticò, A. (2021). ALARP criteria to estimate acceptability and tolerability thresholds of the investment risk. *Applied Sciences, 11*(19), 9086.

Melchers, R. E. (2001). On the ALARP approach to risk management. *Reliability Engineering and System Safety, 71*(2), 201–208. https://doi.org/10.1016/S0951-8320(00)00096-X

Nichols, T. R., Wisner, P. M., Cripe, G., et al. (2010). Putting the Kappa Statistic to use. *The Quality Assurance Journal, 13*(3–4), 57–61. https://doi.org/10.1002/qaj.481

NIOSH (National Institute for Occupational Safety and Health). 2022. Hierarchy of Controls. https://www.cdc.gov/niosh/learning/safetyculturehc/module-3/2.html.

Oguz Erkal, E. D., Hallowell, M. R., & Bhandari, S. (2021). Practical Assessment of potential Predictors of Serious Injuries and fatalities in Construction. *Journal of*

Construction Engineering Management, 147(10), Article 04021129. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002146
Omari Shekaftik, S., Golbabaei, F., & SheikhMozafari, M. J. (2023). An analysis of 'hierarchy of controls' in workplaces and laboratories involving nanomaterials. *Journal of Nanoparticle Research, 25*(12), 245. https://doi.org/10.1007/s11051-023-05891-3
OSHA (Occupational Safety and Health Administration). 2024. Recommended Practice for Safety and Health Programs. https://www.osha.gov/safety-management/hazard-prevention (Accessed January 20, 2024).
Perreault, W. D., Jr. (1975). Controlling order-effect bias. *Public Opinion Quarterly, 39*(4), 544. https://doi.org/10.1086/268251
Reason, J. (2016). *Organizational Accidents Revisited*. Routledge.
Redmil, F. (2010). *ALARP explored*. Newcastle University: School of Computing Science Technical Report Series.
Roelen, A., van Aalst, R., Karanikas, N., Kaspers, S., Piric, S., & de Boer, R. J. (2018). EDefinition, classification, and performance. *Aup advances, 1*(1), 175–189.
Saleh, J. H., Marais, K. B., Bakolas, E., et al. (2010). Highlights from the literature on accident causation and system safety: Review of major ideas, recent contributions, and challenges. *Reliability Engineering and System Safety, 95*(11), 1105–1116. https://doi.org/10.1016/j.ress.2010.07.004
Selleck, R. J., 2023. A critical control approach to preventing fatalities in construction.
Selleck, R., Cattani, M., & Hassall, M. (2023). Proposal for and validation of novel risk-based process to reduce the risk of construction site fatalities (Major Accident Prevention (MAP) program). *Safety Science, 158*, Article 105986. https://doi.org/10.1016/j.ssci.2022.105986
Selleck, R., Hassall, M., & Cattani, M. (2022). Determining the reliability of critical controls in construction projects. *Saf., 8*(3), 64. https://doi.org/10.3390/safety8030064
Selleck, R., & Cattani, M., 2019. Preventing fatalities in the construction industry—a review of critical risk management strategies.
Shahrokhi, M., & Bernard, A. (2010). A development in energy flow/barrier analysis. *Safety Science, 48*, 598–606.
Sherratt, F. (2016). *Unpacking construction site safety*. John Wiley & Sons.
Sklet, S. (2006). Safety barriers: Definition, classification, and performance. *Journal of Loss Prevention in the Process Industries, 19*, 494–506.
Smith, G. (2018). Paper safe: Triumph of bureaucracy in safety management. *Gregory Smith*.
Squillante, R., Jr, Dias, J. C., Moscato, L. A., et al. (2021). A framework for synthesis of safety-related control design to avoid critical faults and pathogenic accidents in the process industries. *Safety Science, 139*, Article 105168. https://doi.org/10.1016/j.ssci.2021.105168
Winge, S., & Albrechtsen, E. (2018). Accident types and barrier failures in the construction industry. *Safety Science, 105*, 158–166. https://doi.org/10.1016/j.ssci.2018.02.006
Zhang, Y., Dong, C., Guo, W., et al. (2022). Systems theoretic accident model and process (STAMP): A literature review. *Safety Science, 152*, Article 105596. https://doi.org/10.1016/j.ssci.2021.105596

Yaqoob Raheemy is a Graduate Research Assistant at the Construction Safety Research Alliance based at the University of Colorado Boulder. He earned a BS in Civil Engineering and MS in Construction Project Management. He is currently completing his PhD in Construction Engineering and Management with a focus on Construction Safety.

Fred Sherratt was Associate Director of Research at the Construction Safety Research Alliance based at the University of Colorado Boulder. She is a Chartered Builder and Chartered Building Engineer, with over ten years of experience working for a large contractor in the United Kingdom. Fred has also worked as a Professor of Construction Sociology in the UK, and her research prioritizes the wellbeing of construction workers and specifically the elimination of serious injuries and fatalities in the construction industry.

Matthew R. Hallowell is a President's Teaching Scholar and Endowed Professor of Construction Engineering at the University of Colorado Boulder. He earned a BS and MS in Civil Engineering and a PhD with a focus on Construction Engineering and Occupational Safety and Health. Before his academic career, he worked in construction as a laborer, project engineer, and quality inspector. Dr. Hallowell specializes in construction safety research and has pioneered leading indicators, hazard recognition programs, safety risk assessment, and precursor analysis. He has published over 100 peer-reviewed papers and has overseen over $3 million in external funding. For his research, he has received the National Science Foundation CAREER Award and the Construction Industry Institute Outstanding Researcher Award.