Research Article

A Tribute to Robert J. Mislevy Part 2: Beyond the Test—Designing Job-Aligned Assessments for Agricultural Workers

Maria Elena Oliveri, *Research Associate Professor, College of Engineering, Purdue University*

Aria Immanuel, *University of Massachusetts*

Kevin Boyle, *Director of Organization and Workforce Development, Equitable Food Initiative*

Alice Linsmeier, *Organizational and Workforce Development Manager, Equitable Food Initiative*

Scopus Abstract

This study illustrates how the principles outlined in "Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts" (Oliveri & Immanuel, 2026) may be applied within a specific (agricultural) industry to make assessment more responsive to a multicultural, multilingual workplace context. Like other industries, agriculture in the United States relies on skills rarely taught explicitly or reflected in conventional assessments. Strawberry harvesters, for example, must routinely interpret safety protocols, sanitation standards, production logs, and supervisor notes, yet most literacy assessments emphasize tasks such as reading short passages and answering multiple-choice questions unrelated to job requirements. This disjuncture leaves learners underprepared for their jobs' literacy

demands. To address this disjuncture, we used two frameworks—Evidence-Centered Design (ECD) and Target Language Use (TLU)—to align assessment design with actual workplace literacy practices. We investigated the following question: *What types of texts, literacy practices and job-related documents do strawberry harvesters most frequently encounter at work*? To answer this question, an environmental scan was conducted that included reviews of training materials and training videos, interviews with agricultural employers and employees, and alignment with the Occupational Network database. Results from these analyses informed a test blueprint for developing valid, job-aligned assessments tailored to agricultural work. Future research may extend this investigation into other industries or high-demand occupations to examine how the ECD and TLU frameworks may enhance workforce readiness, improve training relevance, and expand upskilling opportunities for underserved adult learners across industries.

Structured Abstract

- **Background**: This study demonstrates how the principles outlined in "Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts" (Oliveri & Immanuel, 2026) may guide the design of workplace-relevant assessments in the agricultural industry. Like other industries emphasized in that cross-industry analysis, the agriculture industry relies heavily on adult workers who are culturally and linguistically diverse and have had limited access to formal education. Yet the literacy and numeracy practices these workers routinely use on the job are rarely represented in traditional assessments. Strawberry harvesters, for example, regularly interpret safety protocols, sanitation requirements, production notes, and equipment guidelines; however, most literacy assessments rely on generic tasks disconnected from adult learners' workplace realities. This misalignment raises important fairness concerns: workers may appear to lack literacy skills perhaps not because they cannot read, but because assessments fail to represent how reading and meaning making are enacted in context. To address this disjuncture, we draw on Evidence-Centered Design (ECD) and Target Language Use (TLU) frameworks and assessment design principles to guide the identification, representation, and assessment of workplace-relevant literacy practices for adult learners.
- **Literature Review**: Research in workforce development and assessment consistently shows that contextualized, scenario-based assessments better represent real-world competencies than traditional multiple-choice tests built around decontextualized tasks (Jonsson & Blåsjö, 2020; Straesser,

2015; Visén, 2021). Studies of adult literacy further suggest that literacy and numeracy are not generalizable skills, but practices embedded in specific tasks, environments, goals, and constraints (Oliveri & Immanuel, 2026). While these findings have informed assessment development in some high-demand industries (e.g., healthcare and manufacturing), there remain few concrete, industry-specific examples illustrating how to design contextualized workplace assessments using a principled assessment design approach. In particular, prior research offers limited guidance on how to systematically specify workplace contexts, communicative purposes, and text types and how to translate these elements into evidence-based assessment tasks. This study addresses that gap by illustrating how ECD and TLU can be used together to ground assessment design in the actual language practices workers engage in during agricultural work, thereby supporting the development of assessments that are relevant, fair, and instructionally useful.

- **Research Question**: This study addresses the following question: *What types of texts, literacy practices and job-related documents do strawberry harvesters most frequently encounter at work?* This question anchors assessment design in workers' lived experiences and directly supports the principle of contextual fidelity.
- **Method**: Three data sources were examined to address the research question using a descriptive, exploratory approach situated within applied workforce and assessment research: (a) an environmental scan of training materials and instructional videos, (b) interviews with employers and employees in the Southern California agricultural industry, and (c) occupational descriptors from the Occupational Network (O*NET) database. These data sources were analyzed to identify literacy practices in agriculture broadly and in strawberry harvesting specifically to inform the design of principled, workplace-aligned assessment tasks.
- **Results**: Findings indicate that reading is a frequent and essential workplace activity across roles, though purposes and text types vary by job function. General laborers most often read standard operating procedures and safety signage; crew leaders engage with compliance forms and coordinate tasks using written and verbal instructions; and supervisors work with regulatory documents, logs, and training materials. These patterns suggest that workplace-aligned assessments should vary task types, purposes, and complexity to reflect differences across job roles and responsibilities.
- **Discussion**: This study illustrates how ECD and TLU can be used jointly to strengthen the validity and fairness of workplace assessments by grounding tasks in job-relevant contexts. While multiple item formats are possible, scenario-based assessments and situational judgment tasks are particularly well suited to capturing how literacy is enacted under realistic workplace conditions. Such designs not only generate more accurate

evidence of workers' literacy practices but also support learning and instruction through formative, job-aligned feedback.

- **Conclusion**: By integrating input from employers and employees, this study identifies job-relevant literacy tasks central to strawberry harvesting and delineates the TLU domain for this occupational context. The processes for identifying the TLU domain illustrate how a principled approach can be used to develop test blueprints for designing contextualized, workplace-aligned assessments serving culturally and linguistically diverse adult learners. As an applied extension of "Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts" (Oliveri & Immanuel, 2026), the study provides a concrete example of how assessment design principles can support the more accurate, meaningful, and equitable measurement of workers' literacy practices, thereby enhancing workforce readiness and the usefulness of assessments in training programs. Thus, rather than reporting learner outcomes, this study demonstrates how empirical workplace data can be translated into task and evidence models using the ECD framework.
- **Directions for Future Research**: Future studies should use the principles outlined in this study to develop workplace-aligned assessment task prototypes, pilot test them, and refine them with workers in Job Zones 1–3, including non-native English speakers. Additional research can explore how ECD may be used to anticipate future literacy demands as agricultural technologies and regulatory requirements evolve, helping ensure that assessments remain responsive to changing workplace conditions.

Keywords: adult education, agriculture workforce, contextualized assessment, domain analysis, evidence-centered design, literacy practices, reading assessment, strawberry harvesters, Target Language Use domain, workforce development

1.0 Background

The U.S. labor market is evolving rapidly as automation, artificial intelligence, and regulatory pressures reshape how work is organized and performed (Acemoglu & Restrepo, 2022; Ajjawi et al., 2020). These shifts are especially salient in agriculture, where workers increasingly engage with complex tools, safety and sanitation protocols, compliance documentation, and production workflows under physically demanding and time-sensitive conditions. Although foundational literacy skills remain essential for safe and efficient work, the ways these skills are enacted (e.g., interpreting packaging labels, following harvest procedures, or applying safety standards) differ substantially from the decontextualized reading and writing tasks emphasized in conventional instructional and assessment programs (Suárez-Álvarez et al., 2024).

These labor market changes heighten the stakes for adult education and workforce assessment systems. When assessments fail to reflect how skills are used on the job, they risk misrepresenting workers' competencies, weakening instructional relevance, and limiting opportunities for advancement. Such misalignments are particularly consequential for multilingual workers and those with limited formal schooling, for whom decontextualized assessments may introduce construct-irrelevant barriers unrelated to job performance.

Although this article is designed to stand on its own, it is intentionally paired with a companion study "Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts" (Oliveri & Immanuel, 2026). The companion study operates at a cross-industry level, using national labor-market projections and occupational databases to identify broad trends in workplace literacy and numeracy demands. The present study (Part 2) provides an industry-specific instantiation, illustrating how macro-level insights may be translated into assessment-relevant representations of literacy practices when grounded in field observations, worker interviews, and workplace artifacts from agricultural settings. Together, the two studies demonstrate a principled, evidentiary approach to workplace assessment design without duplicating analytic claims or empirical results.

Drawing on previously conducted environmental scans and interviews with agricultural workers and supervisors, this study focuses on strawberry harvesting to illustrate how assessment design principles can be applied in a labor-intensive, multilingual context. Guided by two complementary frameworks, Evidence-Centered Design (ECD) and Target Language Use (TLU), the analysis shows how assessment design can be grounded in the literacy practices workers use in daily work, rather than in proxy measures of general skills. Consistent with this orientation, literacy is conceptualized as situated interaction with workplace texts, including reading, interpreting, documenting, and coordinating action rather than as decontextualized reading or writing ability.

1.1 Why a Principled Assessment Design Approach Matters

Despite rising skill demands, instructional and assessment systems have not kept pace with changes in how work is performed. Many programs continue to emphasize discrete, decontextualized tasks rather than the complex, situated literacy practices that workers

use during daily work (Oliveri & Tannenbaum, 2017). This academic–workplace misalignment is especially consequential in labor-intensive fields such as agriculture and manufacturing, where workers often rely on job-embedded learning, peer support, and multimodal communication (Suárez-Álvarez et al., 2024).

Workforce disruptions further underscore the urgent need for assessment and training systems that support mobility, reskilling, and recognition of workers' existing competencies. The recent closure of the Tyson Foods beef plant in Lexington, Nebraska, resulting in the loss of more than 3,000 jobs in a town of roughly 11,000 illustrates how sudden economic shocks can destabilize entire communities (Bedayn, 2025). For displaced workers, particularly immigrants and multilingual adults, reemployment often depends on their ability to demonstrate existing competencies and identify upskilling pathways across adjacent industries. When assessments rely on decontextualized literacy measures, they can become barriers rather than bridges to reskilling and mobility.

Agriculture faces a parallel but structurally distinct challenge. Persistent labor shortages, driven by demographic shifts, immigration policy, and the physically demanding nature of farm work, strain growers' ability to recruit and retain skilled workers. At the same time, agricultural work is becoming more complex due to increased regulation, digital recordkeeping, technology, and mechanization. Workers must adapt continuously, yet their literacy, numeracy, and decision-making skills often remain invisible because there are few valid, trusted ways to document them.

1.2 Limitations of General Assessments and the Need for Contextualized Measures

Large-scale assessments such as the Program for the International Assessment of Adult Competencies (PIAAC) and WorkKeys provide valuable information about adults' general literacy and numeracy skills. However, these assessments often rely on hypothetical or broadly framed tasks that abstract away from the tools, workflows, and constraints of specific occupations. PIAAC focuses on adults' literacy, numeracy, and problem-solving skills to examine how well they can use these skills in broad life and work contexts. WorkKeys measures applied reading, math, and document skills aligned to the National Career Readiness Certificate to assess generalized indicators of career readiness. Although both assessments contribute important insights, their broad focus limits their application to occupations such as agriculture where workers have limited English proficiency and schooling, and where literacy demands are tightly coupled with tools, workflow norms, and regulatory contexts. Addressing assessment needs within this context may require more targeted assessments of focal skills within a more specific domain.

Research in adult learning and the learning sciences underscores that meaningful learning requires assessments and instruction embedded in real-world contexts (Haigler, 2021; Mislevy, 2018; Perkins & Salomon, 1988). General skills rarely transfer automatically across settings; workers' ability to apply literacy or numeracy depends on their familiarity with job-specific tools, procedures, interpersonal expectations, and routines (Barnett & Ceci, 2002; Perkins & Salomon, 2012). Assessments that ignore these contextual features risk misrepresenting what workers know and can do, limiting their opportunities for upskilling or reskilling and motivating the need for clearer principles on how to design workplace-aligned assessments for adult learners.

1.3 Core Assessment Design Principles for Developing Workplace-Aligned Assessments

To respond to these challenges, this study advances and illustrates a set of workplace assessment design principles. The first four principles also were presented in "Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts" (Oliveri & Immanuel, 2026). In this study, we also outline two additional principles that emerged from direct collaborations with stakeholders in the agriculture industry.

Principle 1 (Contextual fidelity). Assessment tasks should mirror the tools, workflows, representations, and communicative purposes of real workplaces. A high level of contextual fidelity aims to strengthen validity in the form of construct representation and support more accurate interpretations of workers' literacy and numeracy skills through the alignment of assessment demands with actual job practices in real-world settings.

Principle 2 (Task selection grounded in job relevance). Assessments should target the knowledge, skills, and abilities (KSAs) that matter most to evaluate proficiency in job-relevant tasks in specific roles, rather than relying on decontextualized tasks (Haigler, 2021). The identification of which KSAs matter should come from multiple sources including environmental scans, workplace observations, job analysis, secondary databases such as the Occupational Network database, *and* frontline workers. Workers hold essential, situated knowledge about how tasks are performed, how texts and numbers are interpreted under real conditions, and which errors carry meaningful consequences. Workers' understanding of KSAs may come from various sources such as interviews, focus groups, collaborative task walkthroughs, and reviews of artifacts and work tools to surface how workers interact with the tools, coordinate work with each other, and other contextual constraints often invisible in formal job descriptions. Positioning workers as collaborators and co-designers strengthens construct validity by helping to ensure that assessment tasks reflect not only surface features of work but also the deeper strategies and judgments that enable skilled performance (Eraut, 2000; Lave & Wenger, 1991; Penuel et al., 2007, 2011).

Principle 3 (Complexity calibration). Assessment tasks should be intentionally designed so their cognitive, procedural, and interpersonal demands correspond to the expectations of different roles and levels of responsibility. Calibrating complexity allows assessments to distinguish levels of expertise, avoid underestimating job demands, and provide meaningful diagnostic information for instruction and advancement.

Principle 4 (Back-and-forward design linking current and emerging practices). Assessments should anchor tasks in current workplace practices ("back") while anticipating how automation, shifts in regulatory documentation, and technology changes may reshape work ("forward"). This dual time frame may help keep assessments relevant and reduce the risk of rapid obsolescence.

The present study applies these four principles to a specific occupational domain (strawberry harvesting) in one specific industry (agriculture) to illustrate how broad labor-market insights can be operationalized in a specific industry and occupation. By mapping actual agricultural literacy practices onto principled assessment design

frameworks, the study provides actionable guidance for developing assessments that reflect workplace demands and support improved workforce readiness.

In addition to the four core principles, collaboration with the Equitable Food Initiative (EFI), through on-site interviews and participatory working sessions (described later), surfaced two additional principles grounded in the perspectives of frontline workers, crew leaders, supervisors, and employers. This participatory approach provided access to situated knowledge about how literacy, numeracy, and problem solving are enacted under the real conditions of strawberry harvesting, including interpreting safety and sanitation protocols, tracking production, coordinating work with crews, and making time-sensitive decisions under physical, environmental, and linguistic constraints.

This collaboration also revealed that in agricultural workforce contexts (as is the case in other industries), assessments often serve multiple and sometimes competing purposes, including supporting worker learning and advancement, ensuring regulatory compliance, guiding training and instruction, and improving productivity and safety. Correspondingly, stakeholders may hold different views of what constitutes competent performance, which skills matter most, and how evidence should be interpreted. Frontline workers contribute situated knowledge of task execution under real field conditions; supervisors emphasize coordination, efficiency, and error prevention; and employers attend to compliance and operational continuity. These differing perspectives shape every stage of assessment design, from construct definition and task selection to intended uses of results. Attending explicitly to questions of voice, authority, and agency is therefore essential for ethical and valid assessment design, particularly in contexts where workers typically have limited influence over how competence is defined.

This partnership shed light on the following two principles:

Principle 5 (Participatory construct refinement). Engagement with workers revealed that construct definitions derived solely from occupational databases or researcher-led analyses risk omitting critical dimensions of competent performance, such as informal communication, judgment under uncertainty, peer coaching, and adaptive decision making embedded within routine literacy tasks. Involving workers and supervisors in reviewing draft KSA lists, prioritizing tasks based on frequency and consequence, and refining scenarios for realism and feasibility strengthened construct representation and reduced construct-irrelevant variance.

This participatory process aligns with viewing assessment as *practice* and as an *evidentiary argument*, grounding claims about competence in the realities of work as experienced by those who perform it (Bell & Mislevy, 2020; Mislevy, 2018). This process helps ensure that assessment tasks reflect what workers do (not what external observers assume they do), thereby improving validity and fairness, particularly for multilingual and historically underserved populations (Eraut, 2000; Lave & Wenger, 1991; Penuel et al., 2011).

Principle 6 (Legitimacy through stakeholder alignment). Participatory design also highlighted the importance of legitimacy and trust as conditions for valid assessment use. Assessments that align with existing informal evaluative practices such as supervisor observation, peer instruction, and self-monitoring may be more likely to be perceived as fair, credible, and instructionally useful by workers and employers alike.

Incorporating stakeholder perspectives early in task selection and scenario development increases buy-in, clarifies intended uses, and supports an appropriate interpretation of results. This alignment is especially critical in labor-intensive, multilingual settings, where decontextualized assessments may inadvertently disadvantage workers by introducing unnecessary linguistic or procedural barriers unrelated to job performance. This framing of assessment as both a *feedback loop* and a *measurement system* clarifies how results can support learning and improvement while still meeting requirements for consistency and decision making (Bell & Mislevy, 2020; Mislevy, 2018).

Taken together, these principles underscore that assessment quality is not only technical but also social, shaped by whose knowledge is recognized and how competence is defined in practice. By integrating worker voice throughout the design process, this study advances a model of workplace-aligned assessment that complements standardized datasets like Occupational Network (O*NET) with grounded, contextual evidence. This integration strengthens evidentiary reasoning, supports equity, and increases the likelihood that assessments are interpreted and used ethically to recognize skills, guide training, and support advancement in agricultural workforce contexts.

Consistent with this orientation, literacy and numeracy in agriculture are conceptualized not as abstract cognitive abilities but as work-embedded practices enacted through job-relevant activities, including interpreting safety signage, following sanitation and pesticide protocols, documenting production, and coordinating actions within crews. These practices are inseparable from the physical, temporal, and multilingual conditions under which agricultural work is performed.

From an ECD perspective, principled assessment begins with a clear articulation of the domain. This step involves specifying which KSAs matter for competent performance, the situations that elicit those KSAs, and the forms of evidence that support defensible inferences about performance under real working conditions. Applying this logic to agriculture, we integrate national occupational data from O*NET with field-based interviews, observations, and artifact analyses to demonstrate how the six assessment design principles can be operationalized in a labor-intensive, multilingual industry.

As the next section shows, agricultural work employs large numbers of adult learners whose skills are often developed informally on the job and whose access to formal education or English-dominant assessments may be limited. In this context, principled, contextualized assessment design is especially critical for fairness, instructional relevance, safety, and economic mobility.

1.4 Importance of the Agriculture Industry

Of note, although this study does not focus on the high-growth industries highlighted in "Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts" (Oliveri & Immanuel, 2026), we include this study for two reasons. First, the unexpected passing of Dr. Robert J. Mislevy, whose mentorship shaped the conceptual foundations of this work, occurred as this agriculture paper was nearing completion. Given the timeline for this special issue, it was not feasible to generate a new empirical study that simultaneously honored his influence and

met publication deadlines. Second, agriculture remains a critical industry in the U.S. economy: it faces increasing pressures related to automation, efficiency, and compliance; it is essential to national food security; and it employs numerous workers, many adults who are multilingual, non-native speakers of English with limited formal schooling.

Thus, agriculture is a critical industry in the United States, including states such as California, where agriculture is a cornerstone of the economy. The agricultural workforce is predominantly composed of linguistically and culturally diverse populations (National Agricultural Workers Survey, 2021), which underscores the necessity for assessments and training materials that are tailored to the language and other needs of the workforce.

The agricultural workforce also is vulnerable to working conditions that include exposure to extreme heat, long hours, and physically demanding tasks, which can lead to health issues such as heat stress, dehydration, or injuries (National Center for Farmworker Health, 2018). Moreover, agricultural workers encounter barriers to education and professional advancement due to their limited English proficiency, which can impede access to training programs and career development opportunities (Ornelas et al., 2022).

This situation is further compounded by a lack of accessible educational resources and support systems tailored to the unique needs of farm workers. Many existing materials are not designed using universal design principles, which emphasize flexibility, multiple means of engagement, and materials that are usable by individuals with diverse language, literacy, physical, and cognitive profiles (Meyer et al., 2014; Rose & Dalton, 2009). As a result, there is a lack of resources for agricultural training and an associated absence of materials and assessments including safety manuals, pesticide-handling guidelines, or food-safety protocols to inform the meaningful training and assessment of agricultural workers. These barriers reduce workers' opportunities to learn the necessary skills, limit their ability to participate in required training, and can even lead to safety risks when critical information is inaccessible.

Therefore, although the agriculture industry is not considered a *high-growth* industry in the traditional sense, it represents an industry that relies heavily on adult learners (workers with less than a bachelor's degree)—the same occupational groups emphasized in "Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts (Oliveri & Immanuel, 2026). This workforce is largely composed of linguistically and educationally underserved adults, many of whom are multilingual and have limited formal education (Immanuel & Oliveri, 2026). As such, agriculture provides a meaningful context for illustrating how the aforementioned six principles can inform the design of workplace-aligned literacy assessments that reflect job-relevant literacy demands while remaining meaningful and relevant for workforce development and training contexts.

Moreover, this study complements broader cross-industry analyses by providing insights that the O*NET and U.S. Bureau of Labor Statistics (BLS) Employment Projections national datasets alone cannot capture. O*NET is a database sponsored by the U.S. Department of Labor that provides standardized, regularly updated information about the KSAs, tasks, tools, and work contexts associated with hundreds of occupations. It is widely used for job analysis, workforce development, curriculum

alignment, and assessment because it offers a common occupational language across industries and educational systems. The BLS Employment Projections data, by contrast, provides national- and industry-level forecasts of occupational growth, decline, and workforce demand over a ten-year horizon, drawing on economic modeling and labor market trends to identify where employment opportunities are expected to expand or contract.

While these datasets are essential for understanding workforce demand and occupational structure at scale, both aggregate information at a national level. As a result, they cannot fully represent the localized, situational, and adaptive practices through which work (and workplace literacy and numeracy) is enacted in specific settings.

Building on the principled assessment design concepts introduced earlier, this study demonstrates how those principles can be operationalized within the agriculture industry through a focused analysis of strawberry harvesting. The study pursues three interrelated objectives: first, to identify how literacy and numeracy are embedded in actual agricultural work situations, consistent with the principle of contextual fidelity; second, to document the texts, tools, and decision-making demands through which these skills are enacted across roles and levels of responsibility, informing task selection and complexity calibration; and third, to illustrate how participatory, evidence-centered analyses can translate real workplace practices into valid, workforce-aligned assessment tasks. Together, these objectives show how abstract design principles become concrete when grounded in the lived realities of agricultural work, advancing assessments that are meaningful, fair, and instructionally useful for adult learners.

This study is significant for multiple audiences. For the agriculture industry, it provides assessment design prototypes aligned with actual literacy demands, supporting safer, more efficient, and productive work. For measurement researchers, it demonstrates a theoretically grounded approach to translating workplace analyses into scenario-based assessments. For *The Journal of Writing Analytics* audience, it provides empirical evidence on situated workplace literacy practices and shows how contextualized assessments can capture job-relevant literacy practices, provide meaningful feedback, and produce valid inferences about adult learners' performances contextualized in job-relevant settings.

2.0 Literature Review

In this section, we review empirical research on contextualized skill use and assessment design to situate the present study within the broader literature on workplace learning and adult education. Prior research by the National Research Council (2002) and related work in adult learning and assessment underscores that competencies taught or measured in isolation often fail to capture how learners perform in complex, real-world settings. When assessments abstract skills from the contexts in which they are applied, their practical relevance and interpretability for instructional and workforce decision making are diminished (Haigler, 2021).

Across workplace settings, employees are expected to apply foundational literacy, numeracy, and problem-solving skills under conditions that differ substantially from

formal learning environments. These conditions include time pressure, multimodal information sources, collaborative activity, and high-stakes consequences for error: features that shape how skills are enacted and cannot be disentangled from performance itself (Haigler, 2021). Research in workplace learning demonstrates that instruction and assessment are more effective when tasks reflect these actual conditions of use, supporting both skill development and transfer (Barnett & Ceci, 2002; Perkins & Salomon, 2012).

The persistent disjuncture between academic assessments and workplace practices has significant implications for adult learners, particularly those in labor-intensive industries. For multilingual adults with limited formal schooling, decontextualized assessments may obscure existing competencies and introduce construct-irrelevant barriers, while failures in workplace literacy, such as misinterpreting a safety label or procedural instruction, can carry immediate health, legal, and productivity consequences (Devos, 2023; Visén, 2021; Widén et al., 2021). These realities highlight the need for assessment frameworks that explicitly guide the design of job-relevant, context-sensitive tasks.

Responding to these limitations, scholars increasingly advocate for assessments that move beyond isolated test items to evaluate how skills are enacted within realistic workplace scenarios (Hawthorne et al., 2018; Suárez-Álvarez et al., 2024). Contextualized and scenario-based assessments are better positioned to reveal not only whether individuals possess particular skills, but how effectively they deploy those skills under conditions that mirror actual work. Evidence from both large-scale workforce analyses and industry-based case studies suggests that such approaches yield more actionable information for instruction, training, and workforce advancement (Devos, 2023; Haigler, 2021; Visén, 2021).

2.1 A Case for Skill Contextualization

Research on skill transfer provides a strong theoretical foundation for contextualized assessment. Perkins and Salomon (1988), synthesizing over 20 years of research on the generalization of skills, concluded that the assumption that context-independent skills can be readily applied across different situations "has proven to be more a matter of wishful thinking than hard empirical evidence" (p. 19). This finding has direct implications for assessment design: rather than assuming skills learned in isolation will naturally function in real-world contexts, assessments should prioritize tasks embedded in rich, realistic scenarios that reflect the complexity of actual practice. Isolating skills in decontextualized formats risks evaluating knowledge that may have limited relevance to how learners perform or solve problems in meaningful settings (Sabatini et al., 2020). Despite this evidence, a persistent gap remains in the design and development of contextualized, workplace-aligned assessment tasks.

A few illustrative case studies demonstrate how contextualized skills can be assessed in workplace settings. For example, Oliveri et al. (2021) developed simulation-based tasks to assess workplace English communication that demonstrated that scenario-focused assessments go beyond traditional language tests by replicating interactions employees may face on the job, such as responding to customers, collaborating with colleagues, or completing safety briefings. By embedding tasks in workplace situations,

contextualized assessments capture not only language knowledge but also their application in real-world contexts.

Mislevy et al. (2025) describe a vignette about *Okstaki*, a Blackfoot term referring to the practice of literacy and, more broadly, coming together for communal success. In this Indigenous sociocultural environment, co-author Ms. Crop Eared Wolf drew on the Okstaki worldview to teach Alberta, Canada's Grade 7 science curriculum on thermal energy, heat transfer, insulation, and thermal conductivity. Instead of presenting abstract textbook definitions, she designed activities grounded in land-based knowledge critical for survival, resource use, and storytelling. Students explored questions such as, "How did our people keep homes warm? What materials did we use for insulation before pink fiberglass?" These tasks exemplify Mislevy's notion of contextualization—embedding key concepts, here thermodynamics—within culturally meaningful contexts, allowing learners to leverage their own knowledge and experiences to demonstrate understanding.

Despite these and other illustrative examples (Quellmalz et al., 2010), there remains a need to advance research on principled assessment design, including skill contextualization, beyond current, large-scale assessment approaches used in assessments such as PIAAC and WorkKeys. These assessments often focus narrowly on cognitive or technical KSAs without accounting for the social, cultural, and contextual complexities of real-world applications. For instance, while PIAAC includes some work-related items, tasks are often hypothetical, abstracted from real industries, and presented without the visual, procedural, and social elements that characterize actual jobs (e.g., interpreting safety signage, completing inspection forms, or following written protocols under time or peer pressure) (Organization for Economic Co-operation and Development (OECD), 2021). This gap underscores the need for principled frameworks that systematically guide the design of workplace-aligned assessment tasks.

3.0 Research Question

In this study, we investigated the following question: *What types of texts, literacy practices and job-related documents do strawberry harvesters most frequently encounter at work?* This question responds to the need for contextualized workplace literacy assessments. Rather than relying on generalized skill frameworks, it prioritizes the identification of real-world materials and practices that characterize strawberry harvesting work with the goal of helping to ensure that resulting assessments reflect workplace demands and recognize the proficiencies this population already demonstrates. The question also reflects our participatory approach, which engages agricultural employers, field supervisors, training coordinators, and frontline workers (harvesters) as partners in identifying tasks that are frequent, consequential, and instructionally meaningful.

Addressing this question goes beyond high-level domain modeling by operationalizing TLU information for assessment design. Specifically, it entails specifying the types of evidence that would indicate competent performance, the task formats capable of eliciting that evidence, and the contextual constraints necessary to preserve workplace alignment with real work conditions. In doing so, it offers a concrete

bridge between descriptions of language use in real agricultural settings and the design of assessment tasks that reflect those demands.

4.0 Method

This descriptive, exploratory study situated within applied workforce and assessment research used a data-driven approach to examine its research question. The investigation was guided by two assessment design frameworks (ECD and TLU) and six assessment design principles. The ECD and TLU frameworks provide a structured approach pathway for translating observed workplace practices into assessment specifications, linking real-world task demands to constructs, evidence models, and task designs (Mislevy et al., 2003; Schmidgall et al., 2019).

Three data sources were examined: (a) an environmental scan of training materials and instructional videos, (b) interviews with employers and employees in the Southern California agricultural region, and (c) occupational descriptors from O*NET. These data sources were analyzed to identify literacy practices in agriculture broadly and in strawberry harvesting specifically to inform the design of principled, workplace-aligned assessment tasks.

4.1 Assessment Design Frameworks

Two complementary frameworks (ECD and TLU) were used to guide this study. ECD provided a structured process for translating assessment tasks into assessment designs by articulating the KSAs to be measured, the evidence needed to support inferences about them, and the tasks that would elicit such evidence (Mislevy et al., 2003). The TLU framework defined the real-world texts, communicative practices, and decision-making processes that constitute meaningful performances in specific occupational domains (Bachman & Palmer, 1996, 2010; Schmidgall et al., 2019). By mapping workplace literacy practices into a TLU domain, researchers and practitioners were able to identify the tasks that should serve as the foundation for assessment. Within the ECD framework, the TLU domain specifies what language use looks like in the real world, while ECD specifies how claims about competence can be warranted from assessment performances. In this study, the TLU domain anchored construct definition, while ECD structured the evidentiary argument that linked observed performances to interpretive claims.

4.1.1 The Evidence-Centered Design Framework

ECD offers a structured approach for developing assessments by linking claims about what learners know or can do to the evidence that is needed to support those claims and the tasks that elicit it (Mislevy et al., 2003). The framework is organized into five layers: domain analysis, domain modeling, conceptual assessment framework, assessment implementation, and assessment delivery, each providing a method for systematically documenting and assessing the KSAs required in work contexts (Oliveri et al., 2019).

"Mapping the Skills of Tomorrow—Principled Assessment of Literacy and Numeracy Skills Embedded in U.S. Workplace Contexts" (Oliveri & Immanuel, 2026) illustrates the domain analysis layer, showing how large-scale datasets such as O*NET

can be used to identify literacy and numeracy demands across occupations. O*NET provides a structured national profile of work activities, skills, and knowledge requirements, making it a valuable foundation for identifying broad features of TLU. However, O*NET's strength is also its limitation: its descriptors operate at a national level and cannot capture regional practices, localized workflows, linguistic diversity within worksites, or the specific literacy tasks shaped by crop type, production systems, or employer expectations. As a result, O*NET can indicate *what* literacy demands are likely relevant to agricultural work but not *how* they unfold in the daily routines of workers in a particular region or engaged with a particular crop.

Building on that broader foundation, the present study focuses on the domain modeling and conceptual assessment framework layers, offering a brief example of how the KSAs identified through a national domain analysis can be refined, structured, and linked to task development. Although this paper does not present a full assessment, it demonstrates how on-site qualitative insights can deepen contextual specificity and translate workforce skill analyses into frameworks that more closely reflect actual workplace demands.

Beyond structuring assessments, ECD functions as an anticipatory design framework, one that enables forward-looking assessment development (Oliveri et al., 2021). ECD allows the various stakeholders involved in assessment design, including industry representatives, trainers, and assessment developers, to do the groundwork necessary to ensure that key elements of the test are identified preemptively before major investments are made on test development. This analysis includes determining which KSAs will be assessed, what would constitute evidence of those KSAs, how they will be measured, and which task types will be used. By integrating current occupational data with projections of future workforce trends, ECD helps designers anticipate emerging skill needs and align tasks with the evolving demands of the occupation of interest. This proactive, anticipatory approach allows educators, employee trainers, and assessment designers to prioritize which skills to measure, scaffold instruction effectively, and maintain the relevance of the assessments over time. Our application of ECD as an anticipatory framework offers five key benefits:

1. Turns TLU domain analysis into the evidence-based formal documentation necessary to create assessments.
2. Provides a method to map cross-occupational KSAs to assessment frameworks to support systematic task design.
3. Helps clarify which skills are most critical for workplace success by guiding both instruction and assessment.
4. Enables adaptable assessment design, reducing the risk that instruments become outdated as job requirements evolve.
5. Provides a systematic documentation method that can capture changes while maintaining historical documentation of all changes to an assessment.

ECD has been applied across a range of educational and workforce contexts, including engineering assessments and knowledge-in-use tasks that integrate domain-specific content, cross-cutting concepts, and disciplinary practices (Harris et al., 2019; Pellegrino et al., 2014). Central to these applications is a focus on *skills in use*; that is,

how knowledge and abilities are enacted in actual activity systems rather than demonstrated in isolation.

In agricultural contexts, the notion of skills in use is especially critical. Agricultural work requires workers to integrate literacy, numeracy, perceptual judgment, and procedural knowledge under conditions characterized by time pressure, physical demands, environmental variability, and safety risk. Reading and communication are rarely standalone activities; instead, they are embedded within ongoing tasks such as interpreting pesticide labels while preparing equipment, documenting sanitation compliance while coordinating with a crew, or adjusting harvesting practices in response to changing field conditions. As a result, assessments that treat skills as decontextualized traits risk misrepresenting competence and introducing construct-irrelevant variance, particularly for multilingual workers whose expertise is expressed through action, coordination, and situated decision making rather than through abstract text-based tasks.

ECD provides a principled framework for addressing these challenges by explicitly linking claims about worker competence to observable evidence drawn from realistic task performances. By grounding task design in domain analysis and specifying how particular forms of performance provide evidence for targeted KSAs, ECD supports assessments that capture how skills are mobilized in practice rather than merely whether they can be recalled or recognized. For example, Oliveri et al. (2021) demonstrate how ECD can integrate language use, problem solving, and collaboration within realistic workplace scenarios, resulting in assessment blueprints that are both theoretically defensible and practically relevant.

Applying this approach to agriculture underscores the importance of designing assessments that reflect the *use* of skills *within* work processes, including multimodal communication, informal reasoning, and coordination with others. By foregrounding skills in use, ECD helps ensure that assessments are sensitive to the realities of agricultural labor, support valid inferences about performance, and provide results that are meaningful for training, safety, and advancement decisions.

4.1.2 The Target Language Use Framework

TLU was developed to describe the specific language and literacy skills required by occupational or academic contexts and to guide the principled design of language assessments grounded in these demands (Bachman & Palmer, 1996, 2010). The intent is not to minimize the importance of language proficiency, particularly when assessments inform hiring, upskilling, placement, or safety-critical decisions, but rather to identify the minimum language and literacy competencies necessary to perform tasks accurately, maintain productivity, and avoid preventable hazards that arise from misinterpreting written or spoken information (Holl et al., 2009). From an evidentiary standpoint, these stakes underscore the importance of aligning assessment tasks with workplace demands and of justifying the interpretations and uses of assessment results for specific populations and purposes.

Analyzing a TLU domain involves examining the concrete literacy and communication demands that workers must navigate to perform their jobs effectively. In agricultural contexts, this objective includes identifying the documents workers read or

write (e.g., pesticide labels, food-safety checklists, harvest logs, equipment maintenance forms), the information they must communicate verbally (e.g., coordinating tasks within a crew, reporting safety or quality concerns), the time pressures under which information is processed (e.g., responding to safety warnings or changing field conditions), and the decisions that depend on these literacy activities (e.g., determining field safety following chemical application or documenting quality issues that affect other harvesting processes). Articulating these dimensions clarifies the situations, demands, and decision points that should shape assessment task design and the evidence those tasks are intended to elicit.

Within a TLU-based assessment workflow, considerations of interpretation and use are not treated as a separate or downstream concern but are integrated throughout assessment development. TLU-based assessment design can be described as proceeding through four interrelated stages, each contributing to the overall evidentiary argument:

Task analysis in context – Researchers document the language, literacy, and numeracy practices embedded in actual job activities, capturing how texts are used, what decisions they support, and the consequences of error. Evidentiary reasoning at this stage focuses on construct representation, ensuring that the analysis captures the range and complexity of real workplace demands rather than an idealized or oversimplified version of the job.

Construct modeling – Observed practices are translated into explicit constructs that specify what the assessment claims to measure. Here, the evidentiary focus is on articulating and justifying the link between workplace activities and proposed constructs, making assumptions explicit and open to scrutiny by researchers and other stakeholders.

Assessment task design – Tasks are developed to approximate or simulate job-relevant situations, often using multimodal formats that reflect how information is encountered in practice. At this stage, evidence centers on task fidelity, accessibility, and fairness, particularly for workers with limited formal education or limited English proficiency, to support the interpretation that observed performance reflects targeted competencies rather than construct-irrelevant barriers.

Interpretation and use – A key component of evidentiary reasoning concerns how assessment results are interpreted and used by employers, trainers, and workers themselves. This step includes evaluating whether inferences drawn from scores are appropriate for their intended purposes (e.g., training placement, credentialing, safety assurance) and whether the consequences of use are reasonable, equitable, and aligned with stated goals (Lyons et al., 2025; Messick, 1989; Moss, 1998; Shepard, 1997).

This expanded, use-oriented view of evidentiary reasoning applies to both general and occupation-specific assessments. While general assessments may support comparability across contexts, workplace-specific assessments introduce additional challenges related to contextual dependence, evolving task demands, and diverse stakeholder interpretations. As a result, evidentiary arguments must be revisited over time, rather than treated as a one-time technical exercise (Kane, 2013; Oliveri & von Davier, 2016).

Across industries, TLU analyses consistently show that literacy and communication are embedded in time-sensitive, high-stakes activities (Devos, 2023; Ohlin et al., 2024;

Visén, 2021), whether in agriculture (e.g., interpreting pesticide labels or sanitation protocols), healthcare (e.g., reading patient charts), or manufacturing (e.g., following standard operating procedures). By grounding assessments in these real-world practices and explicitly linking task design to interpretation and use, TLU supports the development of evidence that is meaningful, defensible, and directly actionable for workplace training, safety monitoring, and performance evaluation, enhancing both fairness and practical utility.

4.1.3 Complementarity: ECD and TLU Frameworks

When used together, ECD and TLU complement each other. ECD provides a structured way to translate descriptions of practice (such as activities carried out at work) into formal assessment documentation, including the evidence needed to support claims about KSAs, the task types used elicit that evidence, scoring criteria, and revision history. TLU, meanwhile, guides researchers in describing domain-specific work tasks. These two frameworks have been applied previously in a large-scale assessment where they informed a construct-centered approach to identifying workplace English tasks for the Test of English for International Communication (TOEIC; Schmidgall et al., 2019).

In the context of adult agricultural workers, integrating ECD and TLU allows researchers to identify specific literacy demands (e.g., reading procedural texts, interpreting labeled agricultural products, interpreting documents to make safety-related decisions) and systematically create scenario-based task prototypes. Building on these foundations, the present study uses TLU-informed domain analysis within the ECD framework to investigate the reading demands of strawberry harvesters and translate them into scenario-based assessment tasks. The processes inherent in these frameworks guide the study by defining the literacy constructs of interest, ensuring that assessment tasks reflect real-world performance, and providing a replicable methodology for the design of future occupationally focused assessments.

4.2 Sampling Plan

We used purposive sampling to identify three representative roles within the strawberry harvesting workforce. Purposive sampling is a deliberate, non-random approach to selecting study participants or cases based on their relevance to the research question. Rather than aiming for statistical representation, researchers identify the individuals, groups, or contexts most likely to provide rich, insightful, and meaningful data. This approach is particularly useful in studies that seek to understand specific experiences, practices, or skillsets, as it allows the researcher to focus on information-rich cases. For example, in workforce studies, purposive sampling might involve selecting workers who occupy particular roles, possess certain skills, or have experience with specific workplace tasks. By selecting participants in this way, researchers can gather detailed and contextually grounded evidence, which is essential for informing assessment design, instructional strategies, or the development of scenario-based tasks. Using purposive sampling, we selected the following three roles given their distinct responsibilities and skill requirements:

1. General laborer (e.g., harvester) – responsible for fieldwork, including picking, sorting, and packaging strawberries.

2. Crew leader – supervises general laborers, ensures adherence to quality and safety standards, and communicates production goals.
3. Supervisor – oversees multiple crews, manages compliance with workplace protocols, and coordinates with management and training staff.

To situate these roles within a national framework, we linked them to O*NET, a U.S. database that provides standardized descriptions of occupations. General laborers align with *Farmworkers and Laborers, Crop, Nursery, and Greenhouse* and *Agricultural Equipment Operators*. Crew leaders and supervisors correspond to *First-Line Supervisors of Farming, Fishing, and Forestry Workers*. O*NET also assigns each job to a Job Zone, which indicates the typical amount of preparation required:

- Job Zone 1 – requires little formal education and only brief on-the-job training.
- Job Zone 2 – requires a high school diploma or equivalent and several months of training.
- Job Zone 3 – requires an associate degree or longer, more intensive training.

These classifications help situate strawberry-field roles within national occupational standards while clarifying the skill expectations and training demands associated with each level.

The study did not collect personally identifiable information, and no formal Institutional Review Board (IRB) approval was obtained, as this was a practice-oriented case study conducted in collaboration with industry partners. Therefore, standard demographic details and attrition metrics are not reported.

4.3 Data Analysis

To investigate our research question—*What types of texts, literacy practices and job-related documents do strawberry harvesters most frequently encounter at work?*—we employed a mixed qualitative analytic approach that integrated (a) field-based observations and document analysis, (b) stakeholder interviews and participatory working sessions, and (c) alignment with O*NET, nationally recognized occupational standards.

This multi-source approach was designed to help ensure that identified literacy practices reflected actual workplace demands and could be systematically translated into assessment-relevant representations of the TLU domain. Each of these data sources contributed unique information to identifying the KSAs and informing potential task development within an evidentiary reasoning framework. Document and artifact analyses enabled the identification of the *official literacy demands* of agricultural work (e.g., texts, procedures, and compliance requirements that define what workers are formally expected to read and follow). Field observations extended this analysis by showing *how literacy is enacted in situ*, revealing how written information is interpreted under physical, temporal, and collaborative constraints. Interviews with workers and supervisors provided a third, distinct layer of evidence by illuminating *how competence is recognized, learned, and evaluated in practice*. Interview data made visible aspects of literacy use not fully captured in documents or observations, such as when written information triggers immediate action, how workers shift attention between physical

labor and text, and how supervisors translate written rules into oral instructions and judgments.

Together, these complementary data sources support an evidentiary argument: documents define the target domain, observations characterize task conditions, and interviews clarify what counts as credible evidence of competent performance; each source informs a different layer of the task and evidence models within the ECD framework.

4.3.1 Environmental Scan and Workplace Text Analysis

An environmental scan was conducted at two strawberry farms located in Southern California over three, two-day site visits. During the site visits, observations of how workers engaged with workplace texts under real working conditions—hazard warnings and safety signage, sanitation requirements and compliance notices, pesticide and herbicide labels, equipment manuals, and inspection and compliance checklists—were conducted. Observations focused on the *purposes* for reading (e.g., safety, compliance, production), the *conditions* under which reading occurred (e.g., time pressure, physical constraints, environmental exposure), and the *consequences* of misinterpreting information read.

The environmental scan also included conducting a review of training materials, workplace documents, signage, forms, and training videos used by the farms for onboarding new or seasonal workers to identify any embedded reading and communication demands and obtain further insights into literacy skills in use. The videos featured agricultural workers performing routine tasks, such as identifying weeds, applying herbicides, following safety protocols, and handling operational documents, all of which provided insights into not only what texts workers encounter but also how they engage with them under real conditions.

A sample protocol for coding workplace artifacts along dimensions that capture how literacy is used in practice is provided in Appendix A1. Example dimensions include text type, functional purpose, frequency of use, cognitive demand, decision-making requirements, numeracy demands, multimodality, and consequences of error. Coding makes explicit which aspects of literacy matter for performance and why, helping distinguish between low-stakes routine reading and high-stakes decision-dependent tasks. This approach aligns with domain and construct modeling practices, which emphasize grounding assessment design in direct evidence of real-world performance rather than hypothetical scenarios. Note that the example coding protocol is not intended as a finalized taxonomy; rather, it serves illustrative purposes and needs further research to validate it.

4.3.2 Early Assessment Design Interviews

Another source of evidence for defining the TLU domain included data from two sets of interviews. The first set drew on interviews previously conducted with representatives from ten agriculture-related organizations across California. These interviews were exploratory and strategic in nature and were designed to inform early assessment design decisions by eliciting high-level perspectives on the demand, feasibility, and intended

uses of workplace-aligned assessments and credentials in agricultural contexts (see Appendix A2 for sample questions from these interviews). Given the vulnerability of many agricultural workers, particular care was taken to ensure voluntary participation, confidentiality, and the non-evaluative use of findings.

Participants in this initial interview set represented a broad range of stakeholders who shape assessment use and consequences, including labor organizations, community-based organizations serving agricultural workers, community colleges and training providers, farm employers, and individuals with direct frontline work experience. Collectively, these stakeholders provided insight into how assessments are perceived, what purposes they are expected to serve (e.g., training, hiring, advancement, compliance), and what constraints or risks might arise if assessments are poorly aligned with workplace realities.

Data from these interviews focused on four areas central to assessment development: (1) *perceived value and legitimacy* of workplace-aligned assessments for workers and employers; (2) *implementation considerations*, including language access, technology availability, time constraints, and integration with existing training systems; (3) *equity and consequences*, such as concerns about assessments being used for gatekeeping rather than advancement, or failing to recognize skills acquired informally on the job; and (4) *alignment with actual work practices*, including skepticism toward assessments that rely on abstract or decontextualized tasks.

Findings from these interviews revealed persistent gaps between existing training materials and the skills and pathways required for hiring, promotion, and wage progression. Stakeholders emphasized the importance of employer buy-in, transparent connections between assessment results and wage or advancement opportunities, multilingual delivery, and technological feasibility in field-based settings. Interviewees also highlighted the risk of using assessments that include tasks or language demands that do not reflect how work is performed, which could undermine both validity (construct representation) and trust.

These early interviews played a critical role in shaping the subsequent phases of the study by clarifying assessment design requirements and conditions for assessment development. Specifically, they underscored the need for assessments that (a) are grounded in actual workplace demands, (b) incorporate frontline worker perspectives, and (c) are designed with attention to use, interpretation, and consequences. As such, this initial interview set informed decisions about which occupations to study, which literacy practices warranted closer examination, and why direct engagement with frontline workers was essential for task selection and construct representation in later stages of the research.

4.3.3 Interviews with Frontline Workers

To capture frontline workers' perspectives and document how literacy and numeracy demands are enacted in daily agricultural work, supplemental interviews with approximately 20 frontline workers and facilitated structured participatory working sessions were conducted during three two-day site visits at two strawberry farms. Strawberry harvesting was selected as the focal commodity because of its labor intensity,

safety-critical tasks, and multilingual workforce, making it a particularly informative context for examining workplace-aligned literacy and numeracy practices.

Participants represented various occupational roles involved in strawberry harvesting, including *piscadores* (harvesters), *mayordomos* (crew leaders), *estibadores* (stevedores), *mantenimiento* (maintenance staff), and *tractor y riego* (tractor and irrigation operators). All participants spoke Spanish and had limited to no English proficiency; accordingly, all worker-facing data collection was conducted in Spanish to help ensure linguistic accessibility and to elicit actual descriptions of work practices. The interview questions (see Appendix A3) were designed to gather the following insights:

- Workers' and supervisors' perspectives on the KSAs required to perform their jobs effectively, including how new workers learn these skills and how competence is recognized informally on the job.
- Interest in potential assessments or credentials and the perceived value of assessments for recognizing skills, supporting training, or enabling advancement.
- Practical and contextual considerations—such as time constraints, language use, technology access, trust, and feasibility—that may directly inform assessment task design and delivery in field-based environments.

Interviews were semi-structured and open-ended, with follow-up prompts used to clarify responses and probe for concrete examples.

*4.3.4 O*NET Data*

To complement field-based observations and interviews, O*NET data was used to link the observed literacy tasks to nationally standardized occupational descriptors. The following reading-relevant KSAs were extracted from the O*NET database:

- Task frequency – how often workers are expected to perform key activities, such as reading instructions or monitoring production information.
- Cognitive complexity – the level of effort required, including the attention, judgment, and problem-solving demands associated with workplace tasks.
- Job-level descriptors – indicators of preparation requirements (e.g., education, training, experience) for roles comparable to general laborers, equipment operators, and supervisors.

These variables allowed us to anchor our field findings in a broader occupational framework while still capturing the specific literacy demands unique to strawberry production.

5.0 Results

We investigate our research question—*What types of texts, literacy practices and job-related documents do strawberry harvesters most frequently encounter at work?*—through the analysis of three complementary data sources to characterize workplace literacy demands in the agricultural industry and map them to O*NET job descriptors to inform assessment design. First, we conducted an environmental scan and analysis of video data of workplace texts and learning artifacts—including signage, manuals,

checklists, and training videos—to document the literacy demands embedded in real working conditions. Second, we conducted a qualitative analysis of interviews and participatory working sessions with frontline agricultural workers, supervisors, and farm managers to capture how skills are learned, used, and evaluated in daily strawberry harvesting work. Third, we examined relevant O*NET occupational descriptors to situate observed practices within nationally recognized definitions of job tasks, KSAs, and work contexts.

Together, these data sources provide converging evidence about the skills required for strawberry harvesting and supervision, as well as the context in which those skills are enacted. The sections that follow present results from each source in turn: (1) results from the environmental scan of workplace texts and practices, (2) findings from frontline worker interviews, and (3) insights from O*NET data.

5.1.1 Results of Environmental Scan and Observational (Video) Data Analysis

The results of the environmental scan and video data analysis are summarized in Table 1.

Table 1

Video Analysis of Workplace Literacy Practices

Video Segment	Literacy Task	Observed Behavior	Implications for Assessment
Segment 1	Interpreting herbicide labels.	Worker reads label, identifies chemical concentration, and determines application method.	Assessments should measure comprehension of technical vocabulary and procedural instructions.
Segment 2	Navigating multilingual materials.	Worker consults translated safety protocol to ensure understanding.	Assessments must be linguistically inclusive and consider varying language proficiencies.
Segment 3	Following written safety guidelines.	Worker checks off completed safety steps on a checklist.	Assessments should evaluate the ability to follow sequential instructions accurately.

The video analysis highlighted several critical aspects of workplace literacy in agriculture:

Technical literacy – Workers must comprehend specialized terminology and instructions to perform tasks safely and effectively (e.g., interpreting pesticide labels, equipment manuals, or production checklists in the strawberry fields).

Multilingual contexts – Many workers speak languages such as Spanish or Mixteco, requiring assessments that are accessible to all (e.g., a task may involve reading a safety sign in Spanish or interpreting instructions verbally in Mixteco).

Procedural understanding – The ability to follow written procedures is essential, requiring assessments that measure this skill accurately (e.g., washing hands before handling fruit, completing harvest logs, or assembling boxes according to weight and quality standards).

Video observations further highlight context-dependent, multilingual literacy practices, emphasizing how real-world reading demands extend beyond what is captured in traditional, decontextualized assessments.

5.1.2 Results of Interviews with Frontline Workers

Interview data from frontline agricultural workers and their supervisors were analyzed to understand how workers learn job-relevant skills, how they prefer to demonstrate knowledge, and what conditions support or constrain assessment in farm settings. In what follows, we present results of each of the five key areas associated with the interview questions in Appendix A3, with particular attention to implications for workplace-aligned assessment design. Rather than treating learning and assessment as abstract activities, workers consistently framed both in relation to doing the work correctly, safely, and efficiently under real field conditions.

Job role and skills used at work. Workers described their daily tasks as highly physical, time sensitive, and embedded in routine workflows such as harvesting, handling boxes, following sanitation procedures, and responding to safety conditions. Foundational KSAs were described not as isolated "reading skills," but as skills in use, including recognizing symbols and labels, interpreting short written instructions, coordinating actions with others, and knowing when written information requires immediate action (e.g., stopping work, washing hands, changing procedures).

Harvesters emphasized the importance of these literacy demands, in particular:

- Recognizing safety signage, hygiene reminders, and box labels.
- Understanding that written information communicates rules and expectations.
- Shifting attention between physical labor and short texts without losing accuracy.
- Applying written instructions consistently across repeated tasks.

Supervisors, meanwhile, highlighted the following literacy demands:

- Reading and interpreting longer or more formal documents (e.g., safety manuals, compliance forms, chemical labels).
- Extracting relevant information and translating written rules into oral instructions.
- Monitoring compliance and adapting written guidance to changing field conditions.

Key insights included that literacy demands are role-specific and embedded in action. Workers noted that assessments need to reflect how reading supports decision making, coordination, and safety, not just text comprehension in isolation.

Assessment delivery and platform feasibility. Workers emphasized that any assessment must fit within the realities of farm work. Practical approaches were described as those that:

- Do not require long uninterrupted time blocks.
- Can be completed on-site or integrated into existing routines.
- Rely on minimal technology or use familiar formats (e.g., videos, images, posters).

Workers expressed interest in assessments that could document learning over time, especially when tied to skill recognition, advancement, or acknowledgment of experience. However, feasibility depended on trust, clarity of purpose, and alignment with real job tasks. Key insights included that feasible assessment delivery in agricultural contexts requires short, flexible, low-burden formats that respect time constraints and workplace rhythms and that provide helpful feedback and demonstration of skills needed for farm work.

Assessment versus instructional integration. Interviewees signaled the importance of avoiding the separation of assessment from instruction. Instead, they favored systems that:

- Combine assessment with learning opportunities.
- Provide explanations, demonstrations, or examples alongside evaluation.
- Help workers understand not only whether something is correct but *why.*

Preferred learning supports included:

- Videos and visual demonstrations.
- Clear verbal explanations.
- Opportunities to practice and participate.
- Scenarios that show correct and incorrect approaches.

Workers emphasized patience, repetition, and opportunities to learn together, whether in groups, classrooms, or fields. Insights from this interview section included that assessments are more likely to be accepted and useful when they function as part of a learning and feedback loop, rather than as stand-alone gatekeeping tools.

Language accessibility. All workers emphasized the importance of Spanish (and Mixteco) for accessibility, noting that limited English proficiency creates barriers to understanding written materials and formal assessments. Workers also recognized that crews may include mixed language backgrounds, suggesting value in:

- Bilingual or multilingual presentations.
- Visual supports that reduce reliance on text alone.
- Consistent symbols and formats across languages.

Takeaways included that language accessibility is not optional; it is central to validity and fairness. Multimodal and bilingual design reduces construct-irrelevant variance related to language proficiency rather than job competence.

Worker-generated content and reinforcement. Workers expressed strong interest in contributing to learning by:

- Sharing examples from the field.
- Demonstrating tasks.
- Creating or explaining scenarios based on real experiences.
- Teaching others how they learned to do the job correctly.

Workers described learning as collective and iterative, reinforced through storytelling, peer coaching, and sharing lessons learned across days and crews. Insights

gained from this part of the interview included that worker-generated content can serve as both instructional support and assessment evidence, strengthening relevance, authenticity, and buy-in.

Briefly said, overall takeaways included that workers defined learning and competence in terms of correct performance in context, not abstract knowledge. Preferred assessment approaches emphasized practice, demonstration, explanation, and participation, while challenges included lack of time, physical fatigue, and uncertainty about how to correct errors. These findings reinforce the importance of participatory, workplace-aligned assessment design that captures literacies as they are enacted in real agricultural work, supporting validity, fairness, and trust in assessment use.

*5.1.3 Results of O*NET Data Analysis*

O*NET provides a nationally recognized, standardized taxonomy of Detailed Work Activities (DWAs) and General Work Activities (GWAs) across job roles in various industries including agriculture. DWAs are specific tasks or actions that a worker performs on the job, such as "inspect plants for pests" or "weigh and label harvested produce." GWAs are broader categories of work behaviors that encompass multiple DWAs, such as "monitor processes, materials, or surroundings" or "perform general physical activities." Together, DWAs and GWAs describe the full range of tasks and responsibilities associated with an occupation, providing a framework for linking observed job tasks to nationally recognized occupational standards.

As shown in Table 2, 35 industry-specific job task statements were linked to O*NET DWAs and GWAs, identifying 30 work activities as reading-related literacy requirements.

Table 2

*Results of Matching Job Tasks to O*NET Detailed Work Activities and Associated Skills from O*NET Skills Questionnaire*

EFI Job Task	O*NET Job Task	O*NET Detailed Work Activity	O*NET Skill
Occupation 1: General Laborer			
Adjust or replace equipment components.	Adjust, repair, and service farm machinery and notify supervisors when machinery malfunctions.	Confer with managers to make operational decisions. *)	Listening, Speaking, Writing, Reading
		Maintain forestry, hunting, or agricultural equipment.	Problem-solving
Clean and polish farm equipment.	Adjust, repair, and service farm machinery and notify supervisors when machinery malfunctions.	Confer with managers to make operational decisions. *)	Listening, Speaking, Writing, Reading
		Maintain forestry, hunting, or agricultural equipment.	Problem-solving
Coordinate work between crews.	Direct and monitor the activities of work crews engaged in planting, weeding, or harvesting activities.	Direct activities of agricultural, forestry, or fishery employees. *)	Writing, Speaking, Listening, Reading
Evaluate the condition of the plants (soil, pests, roots, and	Identify plants, pests, and weeds to determine the	Examine characteristics or	Listening, Speaking,

EFI Job Task	O*NET Job Task	O*NET Detailed Work Activity	O*NET Skill
leaves).	selection and application of pesticides and fertilizers.	behavior in living organisms. *)	Reading
Determine the size and quality of the harvested crops.	Weigh crop-filled containers and record weights and other identifying information.	Measure the physical characteristics of forestry or agricultural products.	Mathematics
		Record agricultural or forestry inventory data.	Writing, Reading
Inspect parts of tractors and other farm equipment.	Adjust, repair, and service farm machinery and notify supervisors when machinery malfunctions.	Confer with managers to make operational decisions. *)	Listening, Speaking, Writing, Reading
		Maintain forestry, hunting, or agricultural equipment.	Problem-solving
Handle the tools necessary for the job.	Repair and maintain farm vehicles, implements, and mechanical equipment.	Maintain forestry, hunting, or agricultural equipment.	Problem-solving
Receive instructions from the butler about the day's work or what is going to be cleaned.	Direct and monitor the activities of work crews engaged in planting, weeding, or harvesting activities.	Direct activities of agricultural, forestry, or fishery employees. *)	Writing, Speaking, Listening, Reading
Use a hoe, the appropriate tool depending on the crop for cleaning and weeding.	Plant, spray, weed, fertilize, water, and prune plants, shrubs, and trees, using gardening tools.	Perform manual agricultural, aquacultural, or horticultural tasks.	Problem-solving
Verify contaminants in the field and repair them.	Feel plants' leaves and note their coloring to detect the presence of insects or diseases.	Evaluate the quality of plants or crops. *)	Reading
Occupation 2: Crew Leader			
Check the crew.	Confer with other workers to discuss issues, such as safety, cutting heights, or work needs.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading
Create reports in case there are damaged tools.	Maintain growth, feeding, production, and cost records.	Maintain operational records. *)	Writing, Reading
Create reports of situations found in the field.	Maintain growth, feeding, production, and cost records.	Maintain operational records. *)	Writing, Reading
Day preparation: Ensure crews leave on time.	Confer with other workers to discuss issues, such as safety, cutting heights, or work needs.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading
Evaluate previous work and review results.	Examine and grade trees according to standard charts and staple color-coded grade tags and limbs.	Sort forestry or agricultural materials. *)	Listening, Reading, Mathematics
Give ideas and practices on how to do the job better.	Confer with other workers to discuss issues, such as safety, cutting heights, or work needs.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading
Give instructions to workers about the day's work.	Communicate with coworkers by signals to direct the log movement.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading

EFI Job Task	O*NET Job Task	O*NET Detailed Work Activity	O*NET Skill
Maintain good communication with the supervisor.	Communicate with coworkers by signals to direct the log movement.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading
Mark worker hours on the phone app.	Maintain growth, feeding, production, and cost records.	Maintain operational records. *)	Writing, Reading
Review the production and quality of the day's work.	Mark livestock to identify ownership and grade, using brands, tags, paints, or tattoos.	Mark agricultural or forestry products for identification. *)	Listening, Reading
	Segregate harvested organisms according to weight, age, color, and physical condition.	Classify organisms based on their characteristics or behavior. *)	Speaking, Listening, Reading
Transport crew members and trailers using farm vehicles.	Drive to sawmills, wharfs, or skids to inspect logs or pulpwood.	Drive passenger vehicles.	
Warm up with the crew before working.	Communicate with coworkers by signals to direct the log movement.	Communicate with others to coordinate activities. *)	Writing, Speaking, Listening, Reading
Occupation 3: Supervisor			
Create maps for stewards.	Inspect buildings, fences, fields, or ranges, supplies, and equipment to determine work to be performed.	Inspect equipment or facilities to determine condition or maintenance needs.	
Coordinate closely with the director of human resources, safety, and training on hiring, progressive discipline, performance evaluations, promotions, bonuses, terminations, incident/injury reports, and all other human-resource-related matters.	Communicate with forestry personnel regarding forest harvesting or forest management plan, procedures, and schedules.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading
Guide and support forepersons and crews in basic agronomy and crop activities to ensure work is done to the client's expectations.	Train workers in spawning, rearing, cultivating, and harvesting methods, and in the use of equipment.	Train workers in farming, forestry, or hunting techniques. *)	Writing, Speaking, Listening, Reading
Interact in person and in writing with contractors, vendors, consultants, regulators, or clients.	Confer with managers to determine production requirements, conditions of equipment and supplies, and work schedules.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading
Manage the scheduling and performance of assigned crews and be responsible for the quality of work and efficiency with time and materials.	Communicate with forestry personnel regarding forest harvesting or forest management plan, procedures, and schedules.	Communicate with other workers to coordinate activities. *)	Writing, Speaking, Listening, Reading
Operate independently and professionally in remote settings	Monitor operations to identify and solve problems, improve work methods, and ensure	Monitor operational quality or safety. *)	Listening, Reading

EFI Job Task	O*NET Job Task	O*NET Detailed Work Activity	O*NET Skill
with limited direct supervision.	compliance with safety, company, and government regulations.		
Report dead animals.	Observe animals for signs of illness, injury, or unusual behavior, notifying veterinarians or managers as warranted.	Monitor animal behavior or condition. *)	Listening, Reading
Be responsible for crew safety, time, attendance, job assignments, and compliance with all instructions, rules, standard operating procedure, and state and federal regulations, including OSHA.	Train workers in techniques such as planting, harvesting, weeding, or insect identification, and in the use of safety measures.	Train workers in techniques such as planting, harvesting, weeding, or insect identification, and in the use of safety measures.*)	Writing, Speaking, Listening, Reading
Be responsible for oversight of assigned crews' quality control and safety.	Train workers in techniques such as planting, harvesting, weeding, or insect identification, and in the use of safety measures.	Train workers in techniques such as planting, harvesting, weeding, or insect identification, and in the use of safety measures. *)	Writing, Speaking, Listening, Reading
Supervise and direct the activities of multiple agricultural work crews engaged in various farming tasks on client ranches, including but not limited to preparing fields, planting, irrigating, cultivating, packing, and harvesting.	Perform both supervisory and management functions, such as accounting, marketing, and personnel work.	Direct activities of agricultural forestry or fishery employees. *)	Writing, Listening, Speaking, Reading
Report acres to the client.	Inspect facilities to determine maintenance needs.	Inspect equipment or facilities to determine condition or maintenance needs. *)	
Source and secure basic materials to support assigned crews.	Coordinate dismantling, moving, and setting up equipment at new work sites.	Coordinate forestry or agricultural activities. *)	Writing, Listening, Speaking, Reading
Work closely with management in finance, control, and cost savings.	Calculate monitor budgets for maintenance or development of collections, grounds, or infrastructure.	Monitor financial activities. *)	Listening, Reading

*Note: *) indicates detailed work activities require reading skills. EFI = Equitable Food Initiative. O*NET = Occupational Network Database*

Results shown in Table 2 highlight the central role of literacy across all occupational levels in strawberry harvesting. Of the tasks analyzed, 30 were explicitly reading-related, demonstrating that reading is integral to a wide range of job responsibilities, from interpreting operational instructions and safety protocols to completing compliance forms. Table 2 also shows how each reading-related work activity aligns with industry-specific job task statements, providing insight into the real-world contexts in which literacy is enacted. This mapping not only validates observed workplace practices

against nationally recognized standards but also clarifies which reading tasks are critical for decision making, safety compliance, and task execution, laying the foundation for defining the TLU domain and informing subsequent assessment design.

Tables 3 through 5 break down reading-related work activities by occupational level, illustrating how literacy demands vary across roles. These tables highlight the progression of reading complexity and decision-making responsibility among general laborers, crew leaders, and supervisors. Tables 3 to 5 show which texts and literacy practices occur for each occupational level and how some tasks, such as interpreting regulatory documents or coordinating team activities, become more frequent and cognitively demanding at higher occupational levels.

General laborers. Table 3 shows the industry-specific job task statements for the corresponding DWA for general laborers. Reading is embedded primarily in routine operational procedures and peer/supervisor communications. Key reading-related tasks include reporting operational adjustments, coordinating activities with supervisors, and evaluating plant conditions or weeds according to standard operational procedures. For example, the DWA "confer with managers to make operational decisions" and the associated job task statements are summarized in the first row, second column. Results shown in Table 3 indicate that general laborers engage with texts embedded in operational and environmental monitoring tasks.

Table 3

*Reading-Related Work Activities from O*NET for the General Laborer Occupation*

Reading-Related Detailed Work Activity	Industry-Specific Job Task
Confer with managers to make operational decisions.	**Report to managers regarding operational tasks**, such as making adjustments or replacements of equipment components; cleaning and polishing farm equipment; and inspecting parts of tractors and other farm equipment.
Direct activities of agricultural, forestry, or fishery employees.	**Communicate with supervisors and coworkers**, such as coordinating work between crews and receiving instructions.
Examine characteristics or behavior in living mechanisms.	**Evaluate plant conditions and weeds**, such as observing soil, pests, roots, and leaves, and identify weeds and apply herbicides, according to a standard operating procedure.
Evaluate the quality of the plants.	**Detect and verify contaminants in the field**, such as the presence of insects or disease.

Crew leaders. As summarized in Table 4, which shows the DWAs associated with job task statements for crew leaders, reading skills are embedded in more complex task statements that involve reviewing and communicating ideas. For example, reading skills are used to provide instructions to general laborers before the working day, to grade agricultural commodities, and to track working progress. As shown, crew leaders were responsible for interpreting compliance documentation, recording field data, and reviewing safety logs. Table 4 shows that crew leaders require higher literacy skills to mediate between operational execution and supervisory responsibilities, engaging with documents that are higher stakes and more complex than those encountered by general laborers.

Table 4

*Reading-Related Work Activities from O*NET for the Crew Leader Occupation*

Reading-Related Detailed Work Activity	Industry-Specific Job Task
Classify organisms based on their characteristics or behavior.	**Review the quality of commodity production** based on the grading system (weight, color, or physical condition).
Communicate with other workers to coordinate activities.	**Communicate with crew members and supervisors**, such as preparing the crew for work, giving ideas, practices, or instructions for completing daily work.
Maintain operational records.	**Create and track equipment and working hours reports**, such as creating a report for damaged tools, problems in the field, or working hours.
Mark agricultural or forestry products for identification.	**Review and mark livestock quality** based on the grading system.

Supervisors. For supervisors, literacy practices are advanced and multifaceted, spanning training, coordination, monitoring, and compliance functions, as seen in Table 5. For example, reading skills are used to provide training materials to workers about safety regulations, suggesting that the supervisor role involves reading safety regulation procedures, inspection forms, labor regulations, safety protocols, and quality assurance protocols. Results shown in Table 5 suggest that reading demands increase in complexity and accountability from general laborers to supervisors. Supervisors require integrated literacy competencies, including the comprehension of complex procedures, regulatory texts, and multi-step operational instructions.

Table 5

*Reading-Related Work Activities from O*NET for the Supervisor Occupation*

Reading-Related Detailed Work Activity	Industry-Specific Job Task
Communicate with other workers to coordinate activities.	**Coordinate workers according to standard work instructions and safety procedures**, such as managing the schedule, assigning crew members, and ensuring work efficiency. **Communicate with external team or stakeholders regarding the work operations**, such as coordinating closely with HR about hiring, training, promotions, bonuses, and communicating with contractors, vendors, consultants, regulators, or clients.
Coordinate forestry or agricultural activities.	**Prepare basic materials to support assigned crews**, such as coordinating dismantling, moving, and settling up equipment at new work sites.
Direct activities of agricultural forestry and fishery employees.	**Supervise multiple work activities on the farm**, such as directing multiple work crews in different farming activities (preparing fields, planting, irrigating, cultivating, packing, and harvesting).
Monitor animal behavior or conditions.	**Observe and report unusual activities,** such as observing dead animals, signs of illness, injury, and notifying veterinarians or managers.
Monitor financial activities.	**Work closely with management in finance, control, and cost savings**.
Monitor operational quality of safety.	**Monitor work progress independently in remote settings**, such as monitoring operations, improving work methods, and ensuring compliance with safety and company regulations.
Train workers in farming, forestry, or hunting techniques.	**Train and support forepersons and crew members**, such as training workers in basic agronomy, standard operating procedures, safety working procedures, and OSHA regulations.

5.1.4 Synthesis of Findings and Evidentiary Implications for Assessment Design

Guided by the ECD framework, this section synthesizes findings across Sections 5.1–5.3 by tracing an evidentiary argument from domain analysis to implications for task design, evidence, and use. Rather than reiterating results by data source, this synthesis integrates interview data, field observations, and occupational analyses to show how empirical evidence about agricultural work can be translated into defensible assessment design decisions.

As outlined in Table 6, in which findings are mapped onto design implications, the analysis begins with the domain analysis that characterizes the TLU domain of literacy in agricultural work. Across data sources, literacy practices are shown to be embedded in safety-critical, compliance-driven, and production-oriented tasks enacted under physical, temporal, and multilingual constraints.

Building on this domain analysis, Table 6 illustrates task model implications by identifying which literacy tasks warrant assessment and under what conditions. High-frequency and high-consequence reading tasks (e.g., interpreting pesticide labels, sanitation notices, and production logs) emerge as construct relevant because errors carry immediate implications for safety, regulatory compliance, and coordination. This synthesis further reveals systematic variation in task complexity across occupational roles, supporting calibration of task difficulty based on role-specific decision-making responsibilities rather than generic or grade-level indicators. This shift from grade-level proxies to role-calibrated task complexity represents an analytic advance for writing assessment, enabling more defensible claims about how literacy functions in real workplace contexts.

The synthesis also informs the evidence model, clarifying what constitutes credible evidence of literacy competence in agricultural settings. Rather than extended text comprehension, evidence is observed in how workers act on written information to select appropriate protective equipment, adjust procedures in response to signage, or accurately document production. These findings support the use of scenario-based and performance-oriented tasks in which evidence is derived from decision making and coordinated action under realistic constraints.

Finally, the synthesis addresses interpretation and use, emphasizing stakeholder alignment as a condition for valid and fair assessment. Because supervisors commonly evaluate competence through observation, demonstration, and peer instruction, assessments that mirror these practices are more likely to be trusted, appropriately interpreted, and used instructionally. Participatory refinement of task scenarios further strengthens the evidentiary argument by ensuring that task features reflect real-world vocabulary, formats, and decision points, thereby reducing construct-irrelevant barriers related to language or formal education.

Together, Table 6 demonstrates how empirical workplace evidence can be systematically translated into principled assessment design decisions, consistent with Mislevy's vision of use-oriented, forward-looking assessment grounded in evidentiary reasoning. It establishes an evidentiary bridge between empirical workplace data and assessment design implications.

Table 6

Crosswalk Linking Findings, Evidentiary Design Principles, and Agriculture-Specific Assessment Implications

Summary of Findings	Elaboration of Assessment Design Principles by ECD Layer	Implications for Assessment Design with Agriculture-Specific Examples
Strawberry harvesters regularly engage with standard operating procedures, safety signage, sanitation protocols, pesticide/herbicide labels, equipment manuals, production logs, and compliance checklists.	**Task Selection (Domain Analysis** → Task Model): Identifies high-frequency, high-consequence literacy tasks that are central to safe and effective job performance.	Design scenario-based reading tasks that require interpreting pesticide labels, identifying required PPE, or responding appropriately to sanitation notices during harvest.
Literacy practices are embedded in physical labor, time pressure, multilingual communication, and collaborative workflows.	**Contextualization (Task Model)**: Literacy is enacted under real-world constraints and in coordination with other actions, not as isolated text comprehension.	Embed short texts, icons, tables, and labels into scenarios that simulate field conditions (e.g., reading signage while coordinating with crew members or responding to changing conditions).
Reading complexity and decision-making demands increase from general laborers to crew leaders to supervisors.	**Complexity Calibration (Task Model → Evidence Model)**: Task difficulty and inferential demands must align with role-based responsibility and decision-making authority.	Use role-differentiated tasks: following posted instructions (general laborer); verifying compliance checklists and coordinating actions (crew leader); reviewing logs and identifying discrepancies (supervisor).
Workers frequently combine written information with oral instructions, judgment, and peer coordination.	**Participatory Construct Refinement (Domain Analysis)**: Worker input reveals construct-relevant skills (judgment, coordination, informal communication) not fully captured in occupational databases.	Include collaborative task elements, such as interpreting a written notice while receiving verbal instructions from a supervisor, reflecting actual literacy-in-use practices.
Current practices reflect evolving regulatory, technological, and documentation demands in agriculture.	**Back/Forward Design (Domain Analysis ↔ Future Use)**: Aligns current TLU practices with anticipated changes in roles, regulations, and technologies.	Design tasks that prepare workers for increased digital documentation, automated reporting, or expanded compliance requirements tied to future upskilling pathways.
Workers and supervisors rely on observation, demonstration, and peer instruction to evaluate competence.	**Legitimacy through Stakeholder Alignment (Evidence Use)**: Alignment with existing evaluative practices supports trust, fairness, and appropriate interpretation of assessment results.	Use performance-based or scenario-based tasks that mirror how supervisors already judge competence, strengthening acceptance and interpretability of results.

Building on that foundation, the upcoming Discussion section interprets these findings in relation to prevailing assessment practices and theoretical assumptions in workforce literacy research. In particular, it examines how commonly used proxies such as grade-level indicators misrepresent workplace literacy demands and how role-calibrated, context-sensitive design offers a more defensible basis for assessment in occupational settings.

6.0 Discussion

A central implication of these findings concerns the limitations of grade-level indicators and other coarse proxies commonly used to characterize workplace literacy demands. Grade-level metrics assume that reading difficulty can be inferred from textual features such as length, vocabulary, or syntactic complexity. The present analysis shows, however, that literacy in agricultural work is defined less by textual complexity than by task consequences, decision-making responsibilities, and contextual constraints. Short texts such as pesticide labels, sanitation notices, or equipment warnings often demand high levels of judgment, risk awareness, and procedural knowledge because errors carry immediate safety or compliance consequences. In contrast, longer narrative passages typically associated with higher grade levels may involve comparatively lower stakes. As a result, grade-level measures systematically underrepresent the cognitive and situational demands of real work and risk mischaracterizing workers' capabilities.

Grade-level indicators are therefore poorly suited to workplace assessment not because they lack technical rigor but because they encode assumptions about learning progression that are misaligned with how literacy and numeracy function in occupational contexts. In workplaces, task difficulty is shaped by consequences of error, time pressure, coordination with others, and accountability for decisions, rather than by textual features alone. Our findings demonstrate that literacy demands increase systematically across occupational roles not in linear grade-level terms but through expanded responsibility for interpreting information, making judgments, and coordinating action. Replacing grade-level proxies with role-calibrated task complexity advances assessment practice by aligning evidence with how competence is enacted and evaluated at work, strengthening the defensibility of assessment-based inferences.

Beyond technical alignment, the study foregrounds the social dimensions of evidentiary reasoning through participatory construct refinement and stakeholder alignment. Interviews revealed that supervisors already evaluate competence through observation, demonstration, and peer instruction, and that workers view learning and assessment as inseparable from performing work correctly and safely. Assessments that mirror these existing evaluative practices are more likely to be perceived as credible, fair, and instructionally useful. Participatory review of task scenarios further strengthens the evidentiary argument by ensuring that vocabulary, formats, and constraints reflect actual workplace conditions, thereby reducing construct-irrelevant variance related to language proficiency or formal education.

Taken together, the findings synthesized in Table 6 demonstrate how empirical evidence from agricultural workplaces can be systematically translated into defensible assessment design decisions. By linking domain analysis, task models, and evidence models within an explicit evidentiary argument, the study operationalizes Mislevy's

vision of principled, forward-looking, and use-oriented assessment. The results (see Appendix A4) illustrate an approach for designing workplace-aligned assessments that are technically sound, socially responsive, and meaningful for both workers and employers in agricultural contexts. Specifically, Appendix A4 provides a practical design guide for translating workplace findings into assessment tasks guided by the ECD and TLU frameworks. The table is intended to be used iteratively rather than sequentially, supporting designers as they move between understanding workplace practices (domain analysis), specifying task features (task model), and defining what constitutes credible evidence of competence (evidence model).

Designers should begin by identifying target roles and conducting job analyses to characterize the TLU domain, with particular attention to high-consequence tasks and decision points. These findings then inform task selection, contextualization, and complexity calibration, ensuring that assessment tasks reflect how literacy and numeracy are enacted in agricultural work. Participatory review with workers and supervisors is recommended at multiple stages to refine task realism, language accessibility, and feasibility. By demonstrating how workplace evidence can be systematically translated into task and evidence models, this study operationalizes Mislevy's vision of assessment as an evidentiary argument grounded in actual language use.

6.1 Study Limitations

Study limitations include a small sample size and limited demographic information due to proprietary constraints, meaning results are illustrative rather than fully generalizable. Observations and videos capture only a subset of workplaces and may not reflect the full diversity of worker experiences. Additionally, no inferential statistics or effect sizes were calculated, as the study was exploratory; descriptive summaries and qualitative analyses form the primary evidence. Despite these limitations, the findings provide contextually grounded insights that support the development of TLU-informed, workplace-aligned reading assessments and lay a foundation for future research and assessment design.

Although the TLU domain provides robust, actionable insights, limitations include a restricted number of field sites, small sample sizes, and observations that may not capture the full diversity of work practices. Additionally, while O*NET mapping validates alignment with national standards, it cannot fully reflect localized or situational workplace variations. Despite these constraints, the TLU domain may offer a replicable framework for linking job-relevant literacy practices to assessment design, instructional planning, and/or employee training.

7.0 Conclusion

This study demonstrates how a TLU-informed approach, integrated with principles of ECD and back/forward design, can translate environmental scans into contextually grounded, workplace-aligned reading assessments. By engaging frontline workers, supervisors, and industry partners, and triangulating stakeholder input with occupational data, we identified the literacy demands of strawberry harvesting, including interpreting pesticide labels, following sanitation protocols, and documenting production information. Embedding job-relevant texts and tasks into assessment design strengthens

content validity, reduces construct-irrelevant variance, and enhances fairness for linguistically and educationally diverse workers who are often underrepresented in conventional assessments.

Briefly said, across this multi-source analysis, the study illustrates six interrelated principles of workplace-aligned assessment design. The first four (task selection, contextualization, complexity calibration, and back/forward design) help ensure that assessments focus on high-stakes, job-relevant literacy practices; mirror the multilingual and collaborative realities of work; align task difficulty with occupational roles and decision-making responsibilities; and remain relevant as work practices evolve. Extending these principles, the study highlights participatory construct refinement as a fifth principle, demonstrating how engaging workers and supervisors in reviewing KSAs and refining scenarios strengthens construct representation and improves validity by capturing dimensions of performance often absent from formal job descriptions. The sixth principle, legitimacy through stakeholder alignment, emphasizes that assessments are more likely to be trusted, interpreted appropriately, and used instructionally when they align with existing informal evaluative practices and reflect shared understanding among workers, employers, and training providers.

8.0 Directions for Further Research

Although this study offers a model for developing workplace-aligned reading assessments, several limitations shape how the findings should be interpreted and where future research can build. First, the study focused on a single occupational cluster (strawberry harvesters), which enabled us to deeply examine job-specific literacy demands and carefully articulate the TLU domain. However, this focus also limits generalizability. Future research should apply and adapt this ECD- and TLU-based approach to additional agricultural roles (e.g., irrigation technicians, packing plant workers) and to other industries such as manufacturing, logistics, or health care. Examining the approach across multiple industries would help establish its robustness across diverse workplace cultures, languages, and literacy demands.

Second, the study relied primarily on qualitative data sources, including interviews, workplace observations, and environmental scans. These methods provide rich contextual insight but also require interpretive judgment. Future studies should strengthen the empirical base by incorporating direct performance data, such as the results of pilot testing assessment tasks with workers, think-aloud protocols, and task-based trials conducted under realistic workplace conditions. Such methods would help verify that the proposed tasks elicit the intended constructs and reflect how workers engage with texts and information on the job.

Third, while this study advanced an approach to develop workplace-relevant assessments aligned with observed workplace practices, tasks have yet to be developed with workers and administered to workers. Task development and field testing are therefore critical next steps. Future research should prioritize using the ECD and TLU frameworks to inform task development and evaluate the reliability, validity, and fairness of these assessments, including examining score consistency, sensitivity to language background, and relationships between assessment performance and real-world job outcomes. Validation efforts should be aligned with intended uses, recognizing that

both general and occupation-specific assessments require systematic evidence to support their interpretations and consequences.

Future work also should incorporate co-design strategies that engage workers, supervisors, and other stakeholders to ensure assessments are contextually relevant, equitable, and reflective of diverse workplace realities. This participatory approach may not only strengthen technical quality but also address the social and contextual dimensions of validity, such as whose perspective defines competent performance and how tasks support meaningful skill development. Looking ahead, this line of research has the potential to make a broader impact on adult education and workforce development. By validating these assessments with actual test-takers, examining their effects on learning and employment outcomes, and adapting the approach to a wider range of occupations, researchers and practitioners can help create a new generation of workplace literacy assessments. Such tools would not only measure skill but also actively support learning, safety, and mobility for workers in in-demand industries.

Author Note

This research extends the pioneering contributions of Dr. Robert J. Mislevy to assessment theory, particularly his work on principled assessment design and Evidence-Centered Design (ECD). It illustrates Mislevy's assessment design principles and ECD approach by (a) grounding task selection in documented workplace activities, (b) contextualizing tasks within industry practices, (c) calibrating complexity across job roles, and (d) using back-and-forward design to link current job demands with future upskilling. This study contributes to informing principled task selection, contextualization, and role-calibrated complexity (i.e., the alignment of assessment task demands with occupational responsibilities, decision-making authority, and consequences of error), rather than providing evidence of score reliability, predictive validity, or instructional impact (i.e., via grade-level or generalized difficulty metrics).

Moreover, this study aligns with Mislevy's evidentiary reasoning and contributes mainly to the *front end* of the assessment argument by clarifying the domain, articulating task-relevant conditions, and specifying what would count as credible evidence of competence in workplace contexts. Empirical validation remains an important next step for future research. Consistent with Mislevy's sociocognitive perspective, this study treats performance as inseparable from the conditions of work, informing assessment designs that lead to valid, fair, and meaningful inferences for culturally and linguistically diverse populations.

The research reported here was partially supported by the Institute of Education Sciences (IES), U.S. Department of Education, through grant number R305N210031. The opinions expressed are those of the authors and do not represent the views of IES or the U.S. Department of Education.

Acknowledgments

We gratefully acknowledge the Equitable Food Initiative for their invaluable partnership and guidance in the development of this article. We are also grateful to René Lawless for reading earlier drafts and offering insightful feedback that strengthened the manuscript.

Their expertise, insights, and commitment to equitable practices in agricultural labor provided essential context and support, helping ensure that our work reflects the realities and needs of the field.

Author Biographies

Dr. Maria Elena Oliveri is a Research Associate Professor in the College of Engineering at Purdue University specializing in performance-based assessment in educational and workplace contexts. Her work focuses on fairness, equity, and sociocultural validity, advancing culturally and linguistically responsive approaches to assessment design and interpretation. Grounded in sociocognitive and evidence-centered design (ECD) frameworks, her research investigates principled claims about knowledge, skills, and practices as they are enacted in real-world tasks. At Purdue, Dr. Oliveri's work includes leading and collaborating on assessment research related to workforce readiness in semiconductors and cutting-edge technologies, where she designs job-aligned semiconductors and advanced manufacturing skill pathways and associated assessments. Dr. Oliveri's scholarship integrates psychometrics, sociocultural theory, and learning analytics to design assessment systems that treat skills as situated action—embedded in disciplinary, professional, and workplace activity. Recent work examines the use of stealth performance assessment, combining task design, generative artificial intelligence and analytics to support assessment for learning, feedback literacy, and equitable decision-making across engineering, advanced manufacturing, agriculture, and adult workforce education. Across these efforts, Dr. Oliveri emphasizes minimizing construct-irrelevant variance, enhancing fairness, and consequences of testing, and maximizing construct coverage through task-centered assessment systems aligned to actual workplace demands.

Aria Immanuel is a graduate student at the University of Massachusetts Amherst in the department of Educational Policy, Research, and Administration (EPRA). His work focuses on psychometrics, assessment design, workforce-relevant skills, and equity-centered approaches to education and training. His research examines how literacy, numeracy, and problem-solving skills are embedded in real-world occupational contexts, with particular attention to in-demand industries and the agricultural sector. Aria is a co-author of peer-reviewed articles in the Journal of Workplace Assessment that explore innovative methods for identifying and assessing job-aligned skills, including studies on in-demand occupations and agricultural work contexts. Through this work, he contributes to advancing assessment practices that are grounded in authentic workplace tasks and responsive to the needs of diverse learners and workers. At UMass Amherst, Aria engages in interdisciplinary research that bridges education, workforce development, and applied measurement. He is especially interested in translating research findings into practical tools and frameworks that support employers, educators, and policymakers in designing fair, valid, and meaningful assessments. Aria's scholarship reflects a commitment to improving pathways into and through work by ensuring that assessments recognize the full range of skills individuals bring to complex, evolving labor markets.

Kevin Boyle is Director of Organization and Workforce Development of Equitable Food Initiative (EFI). He is responsible for creating an industry-wide framework of frontline agricultural skills, skill-building and credentialing to provide professional career ladders for agricultural workers and supervisors. He is also responsible for working with growers and their workforce in redesigning production systems to meet the needs of future trends in the industry and maintaining high performing work organizations. As the consultant to multiple stakeholders in agriculture and retail, Kevin facilitated the design and collaboration among diverse participants in a systems approach to improving and developing the fresh produce supply chain across the Americas, today known as the Equitable Food Initiative. "Through this work I am committed to developing Better Places to Work that Work Better, for all of the professionals who own, produce and harvest our food." Kevin retired from the telephone industry after 30 years working on issues of workforce education and the design of new workplaces during times of major technological and market driven changes. He also sat on the National Skills Advisory Board as a representative of The Communications Workers of America. Upon retirement he became the principal of Boyle & Associates, Inc. where he consulted with major corporations and Unions in the U.S., Europe, Canada and South America in developing approaches to major changes within their industry supply chain due to market, technical and social drivers. Boyle and his family still own a Century Farm in Dougherty, IA. He resides in Corvallis, OR.

Alice Linsmeier is an Organizational and Workforce Development Manager at Equitable Food Initiative (EFI), a nonprofit certification and skill-building organization that seeks to transform agriculture and improve the lives of farmworkers. As part of the Workforce Development team, Alice has trained labor-management leadership teams in workplace collaboration, problem-solving skills around labor and environmental standards and supports ongoing learning and communication among workers, managers, growers and EFI. She coordinates the international worker-management Leadership Team Committee that directly advises EFI's board of directors. Her current work is in co-creating an industry-wide framework of frontline agricultural skills, skill-building and credentialing to provide professional recognition and career ladders for agricultural workers and supervisors. Alice's greatest teachers were the farmworkers communities she accompanied and worked with for several years in El Salvador. She has continued learning and working alongside farmworker and immigrant communities for the past 30 years in Latin America and the United States. Inspired by popular education teachers in the community, she completed a Master's in International Multicultural Education from the University of San Francisco and earned a Bilingual, Crosscultural Language and Academic Multi-subject (BCLAD) teaching credential and a graduate certificate in Teaching English to Speakers of Other Languages (TESOL) from San Francisco State University.

References

Acemoglu, D., & Restrepo, P. (2022). Tasks, automation, and the rise in U.S. wage inequality. *Econometrica, 90*(5), 1973–2016. https://hdl.handle.net/1721.1/146052

Ajjawi, R., Tai, J., Huu Nghia, T. L., Boud, D., Johnson, L., & Patrick, C. J. (2020). Aligning assessment with the needs of work-integrated learning: The challenges of authentic assessment in a complex context. *Assessment & Evaluation in Higher Education*, *45*(2), 304–316.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests.* Oxford University Press.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford University Press.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? Transfer of learning. *Psychological Bulletin, 128*(4), 612–637.

Bell, C., & Mislevy, R. (2022). Practice, feedback, argument, measurement: A frame for understanding diverse perspectives on teaching assessments. In M. Blikstad-Balas, K. Klette, & M. Tengberg (Eds.), *Ways of analyzing teaching quality: Potentials and pitfalls* (pp. 21–52). https://doi.org/10.18261/9788215045054-2021-01

Devos, N. J. (2023). Reading for the technical workplace: Developing a diagnostic reading assessment for understanding instructional texts. *TESL Canada Journal, 40*(2), 41–61, https://doi.org/10.18806/tesl.v40i2/1393

Bedayn, J. (2025, December 24). An American dream at risk: What happens to a small Nebraska town when 3,200 workers lose their jobs? *Durango Herald.* https://www.durangoherald.com/articles/an-american-dream-at-risk-what-happens-to-a-small-nebraska-town-when-3200-workers-lose-their-jobs

Eraut, M. (2000). Non-formal learning and tacit knowledge in professional work. *British Journal of Educational Psychology*, *70*(1), 113–136.

Haigler, K. (2021). The role of occupational research in complex assessments: Aligning educational practices with workplace requirements. *The Journal of Writing Analytics*, *5*, 259–291. https://doi.org/10.37514/JWA-J.2021.5.1.08

Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & DeBarger, A. H. (2019). Designing knowledge-in-use assessments to promote deeper learning. *Educational Measurement: Issues and Practice, 38*(2), 53–67. https://doi.org/10.1111/emip.12253

Hawthorne, J., Zerr, R., & Kelsch, A. V. (2018). Assessment of general education: An unexpected (but effective) faculty development opportunity. *New Directions for Teaching and Learning, 2018*(155), 65–72. https://doi.org/10.1002/tl.20304

Holl, D., Kolovich, L., Santiago, D., & Andrade, S. (2009). *Evaluation of the limited English proficiency and Hispanic worker initiative*. United States Department of Labor Employment and Training Administration. https://www.dol.gov/agencies/eta/research/publications/evaluation-limited-english-proficiency-and-hispanic-worker

Immanuel, A., & Oliveri, M. E. (2025, April 24). *Co-design(-ing) to align reading assessments with workplace activities in the agricultural industry* [Graduate student eBoard presentation]. National Council on Measurement in Education Annual Meeting, Denver, CO, United States.

Jonsson, C. & Blāsjö, M. (2020). Translanguaging and multimodality in workplace texts and writing. *International Journal of Multilingualism, 17*(3), 361–381. https://doi.org/10.1080/14790718.2020.1766051

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lyons, S., Oliveri, M. E., & Poe, M. (2025). A framework for enacting equity aims in assessment use: A justice-oriented approach. In C. M. Evans & C. S. Taylor (Eds.), *Culturally responsive assessment in classrooms and large-scale contexts* (pp. 88–105). Routledge.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*(2), 5–11. https://doi.org/10.3102/0013189X018002005

Meyer, A., Rose, D. H., & Gordon, D. (2014). *Universal design for learning: Theory and practice.* CAST.

Mislevy, R. J. (2018). *Sociocognitive foundations of educational measurement.* Routledge.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence centered design. *ETS Research Report Series*, *2003*(1), i–29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x

Mislevy, R. J., Oliveri, M. E., Slomp, D., Cropped Eared Wolf, A., & Elliot, N. (2025). An evidentiary-reasoning lens for socioculturally responsive assessment. In R. E. Bennett, L. D. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 199–241). Routledge.

Moss, P. A. (1998). The role of consequences in validity theory. *Educational Measurement: Issues and Practice*, *17*(2), 6–12. https://doi.org/10.1111/j.1745-3992.1998.tb00826.x

National Agricultural Workers Survey. (2021). *Findings from the National Agricultural Workers Survey (NAWS) 2017–2018: A demographic and employment profile of United States farmworkers* (Research Report No. 17). United States Department of Labor. https://www.dol.gov/sites/dolgov/files/ETA/naws/pdfs/NAWS%20Research%20Report%2017.pdf

National Center for Farmworker Health. (2018). *Agricultural worker occupational health & safety.* https://www.ncfh.org/wp-content/uploads/2025/04/fs-occ_health_2018-1.pdf

National Research Council. (2002). *Performance assessments for adult education: Exploring the measurement issues: Report of a workshop.* National Academies Press. https://doi.org/10.17226/10366

Ohlin, S., King, S., Takashima, M., Ossenberg, C., & Henderson, A. (2024). Learning in the workplace: Development of a simple language assessment tool that supports second-level nurse practice. *Nurse Education in Practice, 103983*. https://doi.org/10.1016/j.nepr.2024.103983

Oliveri, M. E., & Immanuel, A. (2026). Mapping the skills of tomorrow—Principled assessment of literacy and numeracy skills embedded in U.S. workplace contexts. Journal of Writing Analytics.

Oliveri, M. E., Lawless, R., & Mislevy, R. J. (2019). Using evidence-centered design to support the development of culturally and linguistically sensitive collaborative problem-solving assessments. *International Journal of Testing*, *19*(3), 270–300. https://doi.org/10.1080/15305058.2018.1543308

Oliveri, M. E., Slomp, D. H., Elliot, N., Rupp, A. A., Mislevy, R. J., Vezzu, M., Tackitt, A., Nastal, J., Phelps, J., & Osborn, M. (2021). Introduction: Meeting the challenges of workplace English communication in the 21st century. *The Journal of Writing Analytics, 5,* 1–33. https://doi.org/10.37514/JWA-J.2021.5.1.01

Oliveri, M. E., & Tannenbaum, R. J. (2017). Insights into using TOEIC scores to inform human resource management decisions. ETS Research Report https://doi.org/10.1002/ets2.12177

Oliveri, M. E., & von Davier, A. A. (2016). Psychometrics in support of a valid assessment of linguistic minorities: Implications for the test and sampling designs. *International Journal of Testing, 16*, 205–219. https://doi.org/10.1080/15305058.2015.1069743

Organization for Economic Co-operation and Development. (2021). *The assessment frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies: OECD skills studies*. OECD Publishing. https://doi.org/10.1787/4bc2342d-en

Pellegrino, J. W., DiBello, L. V., & Brophy, S. P. (2014). The science and design of assessment in engineering education. In A. Johri & B. M. Olds (Eds.), *Cambridge handbook of engineering education research* (pp. 571–598). https://doi.org/10.1017/CBO9781139013451.036

Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331–337.

Penuel, W. R., Roschelle, J., & Shechtman, N. (2007). Designing formative assessment software with teachers: An analysis of the co-design process. *Research and Practice in Technology Enhanced Learning, 2*(1), 51–74.

Perkins, D. N., & Salomon, G. (1988). Teaching for transfer. *Educational Leadership, 46*(1), 22–32.

Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. *Educational Psychologist*, *47*(3), 248–258.

Quellmalz, E. S., Timms, M., & Buckley, B. (2010). *Using science simulations to support powerful formative assessments of complex science learning* [Conference presentation]. Annual Meeting of the American Educational Research Association, San Diego, CA, United States.

Rose, D. H., & Dalton, B. (2009). Learning to read in the digital age. *Mind, Brain, and Education, 3*(2), 74–83.

Sabatini, J., O'Reilly, T., Weeks, J., & Wang, Z. (2020). Engineering a 21st-century reading comprehension assessment system utilizing scenario-based assessment techniques. *International Journal of Testing, 20*(1), 1–23. https://doi.org/10.1080/15305058.2018/1551224

Schmidgall, J., Oliveri, M. E., Duke, T., & Carter Grissom, E. (2019). *Justifying the construct definition for a new language proficiency assessment: The redesigned TOEIC Bridge® tests—Framework paper.* Educational Testing Service. https://doi.org/10.1002/ets2.12267

Shepard, L. A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, *16*(2), 5–24.

Straesser, R. (2015). "Numeracy at work": A discussion of terms and results from empirical studies. *ZDM Mathematics Education 47*, 665–674.

Suárez-Álvarez, J., Oliveri, M. E., Zenisky, A., & Sireci, S. G. (2024). Five key actions for redesigning adult skills assessments from learners, employees, and educators. *Zeitschrift für Weiterbildungsforschung, 47*, 321–343. https://doi.org/10.1007/s40955-024-00288-8

Ornelas, I., Fung, W., Gabbard, S., & Carroll, D. (2022). Findings from the National Agricultural Workers Survey 2017–2018. *WorkforceGPS.* United States Department of Labor, Employment, and Administration Training. https://lmi.workforcegps.org/resources/2023/11/13/21/27/Findings_NAWS_2017-2018

Visén, P. (2021). Tricks of the trade or situated literacy—Disciplinary reading literacy practices in vocational education. *Nordic Journal of Literacy Research, 7*(1), 60–83.

Widén, G., Farhan, A., Nikou, S., Ryan, B., & Cruickshank, P. (2021). Workplace information literacy: Measures and methodological challenges. *Journal of Information Literacy, 15*(2), 26–44.

Appendix A1. Illustrative Coding Framework with Example Code Values

This table provides an initial, illustrative example of how workplace literacy tasks may be coded to support scenario-based assessment design. The examples are intended for demonstration purposes; future research is needed to further refine, validate, and expand this framework across occupations and contexts.

Coding Dimension	Example Code Values	Definition	Low-Demand Example	High-Demand Example
Text Source	Signage; Label; Checklist; Manual; Video; Digital Interface	Origin or format of the text	Heat-stress warning sign	Pesticide label with application table
Text Type	Informational; Procedural; Regulatory	Primary rhetorical function	Informational poster about hydration	Procedural steps for herbicide application
Functional Purpose	Production; Safety; Compliance; Coordination	Why the text is used	Production target notice	Safety and regulatory compliance instructions
Language(s)	Spanish; English; Bilingual; Symbol-based	Language(s) used	Symbol-only hygiene sign	Bilingual chemical label
Frequency of Use	Routine; Periodic; Conditional	How often text is encountered	Daily harvest count log	Conditional use when weeds appear
Cognitive Demand – Comprehension	Low; Moderate; High	Level of interpretation required	Identifying a single warning symbol	Interpreting multiple warnings and exceptions
Cognitive Demand – Decision Making	None; Guided; Independent	Judgment required after reading	Follow posted break time	Select correct herbicide based on conditions
Numeracy Involvement	None; Basic; Applied	Use of quantitative information	Reading a date	Calculating dilution ratios
Consequences of Error	Low; Moderate; High	Impact of misinterpretation	Minor production delay	Worker exposure or crop damage
Multimodality	Text-only; Text + Symbols; Text + Visuals/Video	Representational forms used	Text-only notice	Label with icons and diagrams

Appendix A2. Sample Questions for Initial Set of Interviews

Section 1. Demand and Use Cases for Credentials

Goal: Understand whether credentials are perceived as valuable, for whom, and for which roles and skills.

1. From your perspective, what skills or competencies are most important for success in … (occupations)?
2. For which roles might recognizing or validating skills be especially valuable? Why?
    - In what circumstances do you think a credential would be most useful, such as hiring, promotion, wage increases, or mobility across farms or regions?
3. What potential benefits or drawbacks do you see in introducing credentials?

Section 2. Past and Current Credentialing Efforts

Goal: Learn from existing or prior credentialing models, including barriers and successes.

1. Are you aware of any current or past efforts to develop industry-recognized credentials in … (occupations)?
    - If so, what aspects seemed effective? What challenges or barriers were encountered?
2. Are there credentialing models in adjacent industries (e.g., construction, manufacturing, food processing) that could provide insights?
    - Which features might be applicable or adaptable?
    - Which features might be less suitable or ineffective in this context?

Section 3. Fit with Lived Experience of Frontline Workers

Goal: Help ensure credentials and assessments align with workers' realities, constraints, and strengths.

1. How do frontline workers typically learn new skills in your experience?
    - Through informal peer learning, on-the-job training, formal classes, or other methods?
2. What challenges do workers face in accessing training or credentials?
    - For example, time, language, transportation, or technology?
3. What would make a credential feel fair, respectful, and relevant to workers?
    - For example, language options, hands-on or demonstration-based assessment, or recognition of prior experience?
4. How important is it for workers themselves to help shape which skills are assessed and how assessments are designed?

Appendix A3. Sample Questions from Frontline Worker Interview Protocol

Part 1: Job Role and Skills Used at Work

- What kinds of tasks do you do during a typical workday?
- Which KSAs are foundational (more / less important) to your job?

Part 2: Assessment Delivery and Platform Feasibility

- What types of assessment approaches would you see as practical in farm settings, considering factors such as time, technological access, and work schedules?
- In your view, how could assessments help document learning or track progress over time?

Part 3: Assessment vs. Instructional Integration

- How do you think an assessment system could best support workers; should it focus only on assessment or include some instructional elements?
- What forms of feedback or explanations (for example, videos, written guidance, or other formats) would be helpful in supporting learning?

Part 4: Language Accessibility

- Which languages are most important to ensure assessments are understandable and accessible for all workers?
- How could bilingual or multilingual presentations be designed to support crews with mixed language backgrounds?

Part 5: Worker-Generated Content and Reinforcement

- How might workers create their own images, videos, or examples to contribute to learning or training?
- In what ways could worker-generated materials be used to support understanding or skill development?

Appendix A4. From Workplace Findings to Task Design: An ECD-Oriented Checklist

This design-oriented checklist is intended to support the translation of workplace analyses into assessment prototypes to assist researchers and assessment coordinators in developing principled, workplace-aligned assessment prototypes. It is not a prescriptive recipe nor a substitute for empirical validation. Instead, the checklist makes explicit design decisions that link domain analysis, task features, and intended interpretations, consistent with ECD assessment development.

Step	Design Question	Action	Agriculture-Specific Examples
1	What work roles and decisions matter?	Select the workforce, job roles, focal industry, and progression pathways using data sources.	Identify work roles, such as general laborers, crew leaders, supervisors in strawberry harvesting.
2	What literacy practices occur in real work?	Conduct job analysis through interviews, observations, and document review.	Identify routine and high-stakes literacy and numeracy tasks such as reading pesticide labels, sanitation signage, production logs, safety checklists.
3	What texts and interactions define the TLU domain?	Specify documents, vocabulary, symbols, and oral communication used on the job.	Collect/develop multilingual signage, icons, tables, short procedural texts, verbal instructions.
4	Which tasks are most consequential?	Prioritize high-frequency, high-risk, or high-decision tasks.	Respond to sanitation violations; determine PPE requirements.
5	How complex should tasks be by role?	Calibrate task difficulty and responsibility by occupational level.	Example tasks by roles: follow instructions (laborer); coordinate compliance (crew leader); audit logs (supervisor).
6	What evidence demonstrates competence?	Define observable actions or decisions that count as evidence.	Correctly identify re-entry intervals; flag discrepancies in logs.
7	What task features elicit that evidence?	Design scenario-based or performance-oriented tasks.	Interpret simulated label under time pressure; evaluate compliance scenarios.
8	How should language and modality be handled?	Incorporate multilingual and multimodal supports.	Identify content and language considerations, such as Spanish text, icons, visuals, short audio instructions.
9	How will tasks be reviewed and refined?	Validate tasks with workers and supervisors.	Conduct participatory review of vocabulary, realism, and feasibility.
10	How will tasks remain relevant over time?	Align tasks with future skill demands and upskilling pathways.	Identify digital documentation procedures, evolving regulations, supervisory responsibilities.