Reading the Map: A Dual-Lens Protocol for Assessing Systems Thinking Artifacts from Brief Workshops in Engineering Entrepreneurial Settings*

Amin Azad[1,*] and Emily Moore[1]

[1]Institute for Studies in Transdisciplinary Engineering Education and Practice (ISTEP), University of Toronto, Toronto, ON M5S 1A1, Canada
*Correspondence: amin.azadarmaki@mail.utoronto.ca

Abstract

Introduction. Assessing what learners take away from a brief systems thinking workshop is hard. Most evaluations rely on post-session self-report, which is vulnerable to response shift and social desirability, and cannot observe the reasoning the workshop targets. The Iceberg models and causal loop diagrams (CLDs) that learners produce carry visible traces of that reasoning but are rarely used as evidence on their own terms, a gap that is acute in entrepreneurial settings where teams reason about live ventures. **Approach.** We present a dual-lens protocol for these artifacts. Lens1 operationalizes Richmond's systems thinking skills (dynamic, system as cause, closed loop, forest, operational) as features recognizable on the artifact without venture domain knowledge. Lens2 operationalizes opportunity identification signals (event-to-structure conversion, boundary expansion, feedback-driven leverage, surfaced constraints). Both are scored absent/partial/present with a single-page rubric. We demonstrate the protocol on twelve artifacts from a four-hour workshop with seven engineering venture teams. **Results.** The dual lens separates artifacts that demonstrate systems thinking from those that stop at description, identifies which teams reached a leverage hypothesis, and exposes indicator-level variation that aggregate self-report compresses into one number. It can be applied by a single rater in under twenty minutes per artifact. **Discussion.** The protocol is a behavioral, replicable instrument that complements survey measures and is most useful where authentic content is non-negotiable and self-report is insufficient. Priorities for broader adoption are independent rater reliability, an active control comparison, and longitudinal application. KEYWORDS: *systems thinking assessment; artifact based assessment; iceberg model; causal loop diagrams; engineering entrepreneurship education*

1 Introduction

Brief systems thinking workshops are widely used in engineering and entrepreneurship education to introduce interdependence, feedback, and structural reasoning in compact formats (Senge, 1990; Sterman,

*This conference paper presents a condensed account of the protocol; an extended version with the full validity, reliability, and transferability treatment is under review at *System Dynamics Review*.

2000). A recurring challenge is how to assess what learners actually take away. The dominant strategy has been self-report on perceived understanding, collected immediately post-workshop (Dugan et al., 2022); such measures are quick to deploy but systematically vulnerable to response shift bias and social desirability, and they cannot directly observe the reasoning moves the workshop is meant to cultivate (Howard & Dailey, 1979).

The artifacts these workshops generate — Iceberg models and CLDs — are rarely treated as evidence in their own right, yet they carry observable traces of the targeted reasoning: how learners separate events from patterns, push explanation into structure, specify feedback, and connect it to intervention pathways (Richmond, 1993; Sterman, 2000). A recent review reports that artifact-based instruments remain underdeveloped relative to survey and scenario approaches, and calls for instruments that examine behavioral products (Dugan et al., 2022; Grohs et al., 2018). The gap is acute in entrepreneurial settings, where teams reason about heterogeneous live ventures rather than instructor-controlled cases, so any assessment must be interpretable across content without domain expertise, and where self-report is at its most fragile because participants have strong motivational reasons to overstate value (Ardichvili et al., 2003; Clausen, 2020; Howard & Dailey, 1979).

This paper contributes a dual-lens assessment protocol for such artifacts and demonstrates it on one workshop. Lens1 operationalizes Richmond's systems thinking skills at the artifact level; Lens2 operationalizes opportunity identification signals, tying artifact features to the opportunity development literature (Ardichvili et al., 2003). The systems-to-opportunity mapping in Lens2 is our own synthesis, building on our framework linking Richmond's skills to entrepreneurial opportunity identification (Azad et al., 2024). The contribution is methodological: a behavioral, replicable instrument that complements survey measures and can be applied without domain expertise once the indicators are defined.

2 The Dual-Lens Protocol

Vocabulary. A *skill* (Lens1) or *signal* (Lens2) is the construct being assessed; an *indicator* is the observable artifact feature that evidences it; a *descriptor* is the rule that defines the indicator and states what counts as absent, partial, or present. Each indicator is scored on a three-level ordinal scale — absent, partial, present — not as a binary. The partial level is deliberate: in brief workshops the modal outcome on several indicators is an incomplete reasoning move (a structure named but not closed, a pattern implied but not labeled), and a binary scheme would discard that resolution.

Lens1 (systems thinking skills). We operationalize five of Richmond's skills (Richmond, 1993, 1998) as artifact features recognizable on either an Iceberg or a CLD. *Dynamic thinking*: the challenge is framed as behavior over time (a pattern/trend layer, or time-varying variables) rather than a one-off event. *System as cause*: outcomes are attributed to internal structure and assumptions rather than external shocks (structure/mental-model layers, or closed causal chains). *Closed loop* (CLD only): at least one closed feedback chain with labeled polarity. *Forest thinking*: boundary breadth, counted as actor categories named beyond the focal user. *Operational thinking*: variables specified as mechanisms rather than placeholder labels.

Lens2 (opportunity identification signals). Lens2 asks whether the Lens1 moves connect to intervention pathways. *Event-to-structure conversion*: a chain from a surface symptom to a structural driver and a candidate intervention placed at the structural layer. *Boundary expansion*: an intermediary, gatekeeper, platform, or regulator whose constraints alter the stated opportunity. *Feedback-driven leverage*: a leverage point named within a feedback structure with an articulated propagation path (Meadows, 1999). *Surfaced constraint*: at least one feasibility or adoption constraint (cost, approval, trust, switching cost, access).

The complete absent/partial/present rules are consolidated in Table 1, which is designed to be used on its own: a rater scores any system map by reading down the nine rows and selecting, per row, the best-fitting column. We do not collapse the scores into a single total; the value lies in the *profile* (which skills are present, which stop at partial, and whether Lens1 evidence is matched by Lens2). A compact report is the count of present and partial indicators within each lens. Annotation takes a competent rater ten to twenty minutes: catalogue surface features, apply Lens1 then Lens2 with a one-to-two-sentence justification anchored to a specific feature, and (with multiple raters) score independently before reconciling at the indicator level.

Table 1: The dual-lens scoring rubric. Score each indicator **absent (0)**, **partial (1)**, or **present (2)** by matching the artifact to the best-fitting column. Closed loop thinking is assessable only on tools that represent feedback explicitly (CLD, stock-and-flow).

Indicator	**Absent (0)**	**Partial (1)**	**Present (2)**
Lens1 — Systems thinking skills (Richmond)			
Dynamic thinking	Static or event-level framing only	Change over time implied but not named as a pattern	Explicit pattern/trend layer (Iceberg) or time-varying variables (CLD)
System as cause thinking	Outcomes blamed on external shocks	Mix of endogenous and exogenous attribution	Outcomes attributed to internal structure, constraints, or assumptions
Closed loop thinking *(CLD only)*	Only linear cause–effect chains	Closed chain without polarity, or an unclosed plausibility chain	At least one closed feedback loop with labeled polarity
Forest thinking	Only focal user, team, and product	One or two actor categories beyond the focal user	Three or more actor categories beyond the focal user
Operational thinking	Placeholder labels only (e.g., “delays”)	Mix of mechanism-level labels and placeholders	Mostly mechanism-level variables/drivers (e.g., “turnaround time”)
Lens2 — Opportunity identification signals			
Event to structure conversion	No descent below the event/pattern layer	Chain reaches structure but names no intervention	Symptom $\rightarrow$ structural driver $\rightarrow$ candidate intervention at structure
Boundary expansion	No actors beyond the focal user	Extra actors named but no opportunity implication	Added actor whose constraints alter the stated opportunity set
Feedback driven leverage	No leverage point identified	Leverage point named without propagation logic	Leverage point named *with* an articulated propagation path
Surfaced constraint	No feasibility/adoption constraint surfaced	A constraint alluded to but not specified	At least one feasibility/adoption constraint named explicitly

3 Demonstration

The demonstration workshop was a four-hour in-person session for early-stage engineering venture teams, structured around the Iceberg Model and CLDs. Of twenty-nine teams that registered interest, seven attended against pre-specified inclusion criteria (active ideation to early-validation stage, a working team of at least two, a live venture challenge, and full availability); 22 participants brought ventures spanning digital platforms, hardware and deep tech, health technology, workforce and marketplace systems, sustainability, and consumer applications. Twelve artifacts from six teams (six Icebergs and six CLDs) form the demonstration set; one team did not produce a CLD and is not included. The first author scored all artifacts using the four-step procedure, independently of the parallel focus-group and survey strands (Azad, 2026). Ethics approval was obtained from the University of Toronto Health Sciences Research Ethics Board (Protocol #47007), with informed consent including use of anonymized artifact images.

Across the six Icebergs, dynamic thinking and system-as-cause thinking were most frequently present (at least partial in five of six); forest thinking was present in four of six; operational thinking was the most variable indicator, ranging from mechanism-level labels to placeholder language. Across the six CLDs, closed-loop thinking was present with labeled polarity in four, partial in one, and absent in one (a linear plausibility chain); dynamic thinking was present in the two artifacts whose variables were expressed as time-varying quantities; system-as-cause thinking was present in three; and forest thinking was present in two. Reading the Iceberg and CLD scores side by side shows that a team could score differently on the same skill across its two artifacts, information a single-artifact analysis would miss.

Lens2 scores were generally lower, consistent with the more demanding criterion of connecting decomposition to an intervention pathway. Event-to-structure conversion was present in three of six (partial in two); boundary expansion was present in three (partial in three); feedback-driven leverage was present in two (partial in one, absent in three); and surfaced constraint was present in four (partial in two), the most frequent Lens2 signal. Two patterns are worth flagging. First, partial scores are common and informative: they capture artifacts that show the structure of a move without completing it, resolution a binary scheme loses. Second, Lens1 and Lens2 can dissociate — a team can produce a competent structural decomposition without connecting it to a candidate intervention, which is precisely the pattern most relevant to entrepreneurship educators.

4 Discussion and Conclusion

The protocol contributes a behavioral, replicable instrument for brief workshops where self-report is most vulnerable to bias. It operates on artifacts the workshop already produces, requires no additional participant burden, and ties features to recognizable theoretical anchors (Richmond's skills and the opportunity identification tradition) rather than bespoke rubric categories.

Neither lens is intrinsically tied to the Iceberg and CLD. The constructs are learning outcomes that could surface in many representations — stock-and-flow diagrams, behavior-over-time graphs, stakeholder maps, even a written memo or pitch. What the lenses assess is the reasoning move, not the diagram. We focused on the Iceberg and CLD because our prior framework maps Richmond's skills to the representational affordances of exactly these two tools, which is what lets us write confident, tool-specific indicators (Azad et al., 2024). The same measures extend beyond these tools once the indicators are re-specified for the new representation; only closed-loop thinking is intrinsically tool-bound, since it requires a representation that shows feedback explicitly.

The demonstration is illustrative, not a reliability study: it was scored by one non-blind rater on a small single-workshop set, without an active control. The priorities for broader adoption are therefore an independent inter-rater reliability study, an active control comparison against artifacts produced under a matched-dose non-systems-thinking intervention, and a longitudinal extension that scores the same teams over time. The protocol is offered as a complement to, not a replacement for, survey and scenario instruments — most useful where authentic venture content is non-negotiable and self-report alone is insufficient.

Acknowledgments

The authors thank the engineering venture teams who participated in the workshop and shared their mapping artifacts. This research was funded by the NEO Materials Catalyst Fund, with in-kind support from ISTEP, University of Toronto.

References

Ardichvili, A., Cardozo, R., & Ray, S. (2003). A theory of entrepreneurial opportunity identification and development. *Journal of Business Venturing*, *18*(1), 105–123.

Azad, A., Moore, E., & Hounsell, A. (2024). Advancing engineering education: Linking systems thinking skills to the tools through a revised framework. In *2024 ASEE Annual Conference & Exposition Proceedings*. Portland, OR. https://doi.org/10.18260/1-2--46525

Azad, A. (2026). *[Parallel manuscript reporting workshop learning outcomes; Ph.D. Thesis, University of Toronto]*.

Clausen, T. H. (2020). Entrepreneurial thinking and action in opportunity development: A conceptual process model. *International Small Business Journal*, *38*(1), 21–40.

Dugan, K. E., Mosyjowski, E., Daly, S., & Lattuca, L. (2022). Systems thinking assessments in engineering: A systematic literature review. *Systems Research and Behavioral Science*, *39*(4), 840–866.

Grohs, J. R., Kirk, G. R., Soledad, M. M., & Knight, D. B. (2018). Assessing systems thinking: A tool to measure complex reasoning through ill-structured problems. *Thinking Skills and Creativity*, *28*, 110–130.

Howard, G. S., & Dailey, P. R. (1979). Response-shift bias: A source of contamination of self-report measures. *Journal of Applied Psychology*, *64*(2), 144–150.

Meadows, D. H. (1999). *Leverage points: Places to intervene in a system*. Sustainability Institute.

Richmond, B. (1993). Systems thinking: Critical thinking skills for the 1990s and beyond. *System Dynamics Review*, *9*(2), 113–133.

Richmond, B. (1998). The thinking in systems thinking: How can we make it easier to master. *The Systems Thinker*, *9*(2), 1–5.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. Doubleday.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. Irwin McGraw-Hill.