Designing for fusion in statistics education: A variation theory approach

Hanan Innabi[1] and Per Nilsson[2]
[1]University of Gothenburg, Sweden; [2]Linnaeus University, Sweden
hanan.innabi@gu.se

This paper presents a design approach for teaching statistics in STEAM education, grounded in variation theory and enriched by real-world contexts drawn from Education for Sustainable Development (ESD). Variation theory serves as the central theoretical framework, offering a way to support learners in discerning critical aspects of the object of learning through structured variation. A key feature of variation theory is the coordination of part–whole relationships, enabling students to learn statistics not as a set of fragmented procedures, but as an integrated and coherent whole. Within this framework, we suggest that sustainability-related problems provide powerful contexts for supporting the development of statistical reasoning, particularly when aligned with the interdisciplinary goals of STEAM education. The paper outlines how these elements can be brought together in a coherent instructional design to foster purposeful, conceptually grounded learning in statistics.

RETHINKING STATISTICS EDUCATION IN STEAM

In today's data-intensive world, statistical literacy is essential for informed citizenship and professional decision-making. This is especially true in STEAM disciplines, where data driven reasoning underpins evidence-based thinking and innovation. Students in STEAM fields must not only understand statistical methods but be able to apply them in complex, real-world contexts that often involve uncertainty, multiple variables, and evolving datasets (Friedrich et al., 2024). Statistics education, therefore, plays a crucial role in preparing STEAM learners to engage with evidence-based inquiry and socially relevant challenges.

The GAISE College Report ASA Revision Committee (2016) recommends two important emphases for modern statistics education: (a) teaching statistics as an investigative process of problem-solving and decision-making, and (b) giving students experience with multivariable thinking. The first calls for helping students understand statistics as a coherent, inquiry-oriented discipline, rather than an unrelated collection of formulas and procedures. The second reflects the complexity of real-world problems, where valid interpretations often depend on considering multiple interacting variables. These emphases are particularly relevant in STEAM education, where the statistical demands of real-life problems exceed simplified textbook exercises.

THEORETICAL AND EDUCATIONAL FRAMEWORK

This paper is grounded in Variation Theory (VT) and shaped by the educational priorities of STEAM and Education for Sustainable Development (ESD). These priorities guide why and how statistics is taught, focusing on preparing students to use statistical reasoning to make sense of complex, real world problems. To support these educational priorities, we draw on VT as a central pedagogical and theoretical framework. VT offers powerful principles for designing learning opportunities that help students discern critical aspects of complex object of learning. According to VT, learning occurs through structured patterns of variation, including contrast, generalization, and fusion (Marton, 2015). Each pattern contributes to students' ability to see what needs to be seen.

Contrast enables learners to distinguish critical features by encountering variation in one dimension while holding others constant. Generalization supports the abstraction of key aspects across diverse situations. Fusion, the focus of this paper, involves coordinating multiple aspects simultaneously to form a new, integrated whole. According to VT, learning typically follows a trajectory involving the patterns of contrast, generalization, and fusion. Contrast is foundational for meaning making, as it allows learners to discern a critical aspect by experiencing it against what it is not. Without contrast, there is no basis for recognizing a feature as meaningful. Generalization, by contrast, does not introduce new meaning but helps refine it by encountering variation in other aspects while keeping the critical aspect constant, learners deepen their understanding and extend it to new situations. Fusion marks a more advanced stage of learning, in which multiple critical aspects are discerned simultaneously and

In: Birk, L., Loth, G., Jotzo, L., Binder, K., & Frischemeier, D. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025, Münster, Germany*.

coordinated as part of a coherent whole. In this way, the patterns of variation do not just support isolated insights but structure a trajectory toward increasingly integrated and sophisticated understanding.

In this context, *statistical tools* refer not just to procedures or formulas, but to conceptual and analytical methods such as measures of variability, confidence intervals, and hypothesis testing understood as reasoning tools that support data-informed decision-making. It is not enough for learners to identify and understand individual statistical tools—they must also grasp how these tools relate to each other and function together in addressing real-world problems (Burrill & Pfannkuch, 2024). As Marton (2015) emphasizes, learning involves moving from an undifferentiated whole to a differentiated set of parts that are then reintegrated into a more powerful understanding.

This view of learning calls for instructional designs that support part whole coordination, allowing students to encounter statistical concepts not in isolation but as interconnected tools serving a common purpose. Research supports this approach: Ingerman (2026) found that students' meaning-making depended on coordinated attention to multiple dimensions of a task. Beyond VT, related work in learning sciences has shown that comparison, invention, and connection-making are critical to building coherence in students' thinking, as emphasized by Chin, Chi, and Schwartz (2016), who highlight the importance of guiding students to recognize structure across cases, an approach consistent with the goals of fusion. Fries et al. (2020), focusing specifically on statistics education, examined how undergraduate students can develop coherence and conceptual integration when instructional design emphasizes comparison, coordination of ideas, and sustained engagement with real-world data.

In statistics education, fusion is particularly important when teaching ideas like variability, statistical inference, and the interpretation of uncertainty. These ideas are often introduced separately and procedurally, which can result in compartmentalized knowledge. However, when students engage with problems that require them to apply and coordinate these concepts returning to the same context at different points in their learning, they are better positioned to develop integrated reasoning. The fusion of statistical tools across time and contexts enhances both retention and transfer of understanding.

As part of an ongoing project, a recent study by Innabi et al. (2026) applied VT in a master's level quantitative research methods course to examine how students' statistical reasoning developed through patterns of variation deliberately designed to support the discernment of critical aspects and their integration over time. A similar approach to what is proposed in this paper was followed: the instructional design employed a multi-level structure of objects of learning and incorporated critical aspects and patterns of variation across descriptive and inferential tasks. Within each level of this structure, the three patterns of variation; contrast, generalization, and fusion, were used to support the development of increasingly integrated understanding. This allowed students to coordinate multiple statistical ideas as elements of an evolving whole rather than as isolated tools.

This paper advocates for course designs that make fusion possible by supporting students in returning to meaningful problems with new statistical tools, encouraging the coordination of parts into a coherent whole (Innabi et al., 2026). The following example illustrates how this design principle can be applied in practice.

EXAMPLE: REVISITING A REAL-WORLD SUSTAINABILITY PROBLEM

In an introductory statistics course for STEAM students, a key challenge is to design learning experiences that help students make sense of statistical ideas in a connected and meaningful way, while progressively learning and applying the full range of statistical tools and related concepts. The overall object of learning in such a course can be framed as: *the capability to transform real world problems into statistical ones using appropriate statistical tools and reasoning*.

To support this, instructors can design a recurring, evolving problem that students encounter at the very beginning of the course and return to at its conclusion. This problem is used initially to introduce basic statistical tools and reasoning and is revisited at the end of the course, where students apply the full range of concepts and methods they have learned. In doing so, they are supported in coordinating and integrating these tools to reason statistically about a complex, real-world issue. One such problem involves exploring the relationship between renewable energy use and CO_2 emissions, drawing on publicly available data from the UN SDG Data Portal (SDG 7: Affordable and Clean Energy, and SDG 13: Climate Action).

In the early stages of the course, students use descriptive statistics to summarize and visualize patterns in the data. While doing so, they encounter all three patterns of variation: contrast, generalization, and fusion, to support the discernment of critical aspects. For example, by contrasting values above and below the mean, students begin to understand the mean as a balance point. Through generalization, they explore how the mean behaves across different datasets with similar or different structures (e.g., small vs. large samples, or datasets with the same mean but different values). Fusion initially occurs within the concept of the mean itself, as students simultaneously discern multiple aspects, its role as a representative value, its dependence on all data points, and its sensitivity to changes in individual values. As learning progresses, more advanced forms of fusion are supported by integrating the concept of the mean with related concepts such as spread and distribution shape. For instance, students come to see that two datasets may have the same mean but differ in how the values are spread around it or how the shape of the distribution affects interpretation. According to VT, such integration of multiple critical aspects, within and across concepts, reflects more sophisticated learning, as students begin to coordinate different dimensions of variation in a unified way. Later, they apply inferential tools such as correlation, regression, and confidence intervals to draw conclusions about trends and associations. Again, the design uses all three patterns of variation to structure learning experiences. According to VT, the sequence typically begins with contrast, followed by generalization, and finally fusion, but all three may be present at any level of the object of learning to support students in developing more integrated and flexible understanding. For instance, by contrasting strong and weak correlation patterns or wide and narrow confidence intervals, students can discern critical aspects of statistical uncertainty and the strength of evidence.

This design reflects the principle of fusion by allowing students to return to a familiar context with new conceptual tools, progressively building a coherent understanding of the investigative process. At each stage, previously discerned critical aspects are revisited and brought into new simultaneity with emerging aspects, allowing students to experience fusion as they recontextualize what they have already discerned in relation to new learning. For example, initial descriptive analyses (e.g., calculating means or creating plots) are later interpreted alongside confidence intervals and regression results, encouraging students to coordinate their understanding of data distribution, uncertainty, and association. Rather than treating each statistical concept as a separate procedure, the investigative process invites students to integrate them as mutually reinforcing elements of a unified reasoning process. This cumulative use of statistical tools exemplifies fusion as described in VT, where learning involves discerning how differentiated parts come together to serve the overarching purpose of understanding and solving real world problems. It also aligns with the goals of ESD by embedding learning in socially and environmentally relevant contexts, while promoting the type of integrated thinking required in STEAM fields.

CONCLUSION

This paper proposed a design approach that uses VT to foster integrated statistical understanding through fusion. By revisiting meaningful problems in sustainability contexts, students are supported in coordinating statistical tools and concepts over time, promoting deeper reasoning and preparing them for data informed thinking in STEAM fields.

REFERENCES

Burrill, G., & Pfannkuch, M. (2024). Emerging trends in statistics education. *ZDM Mathematics Education*, *56*, 19–29. https://doi.org/10.1007/s11858-023-01501-7

Chin, D. B., Chi, M., & Schwartz, D. L. (2016). A comparison of two methods of active learning in physics: Inventing a general solution versus compare and contrast. Instructional Science, *44*, 177–195. https://doi.org/10.1007/s11251-016-9374-0

Friedrich, A., Schreiter, S., Vogel, M., Becker-Genschow, S., Brünken, R., Kuhn, J., Lehmann, J., & Malone, S. (2024). What shapes statistical and data literacy research in K-12 STEM education? A systematic review of metrics and instructional strategies. *International Journal of STEM Education, 11*, Article 58. https://doi.org/10.1186/s40594-024-00517-z

Fries, L., Son, J. Y., Givvin, K. B., & Stigler, J. W. (2020). Practicing connections: A framework to guide instructional design for developing understanding in complex domains. Educational Psychology Review, *33*, 739–762. https://doi.org/10.1007/s10648-020-09561-x

GAISE College Report ASA Revision Committee. (2016). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016*. American Statistical Association.

Ingerman, Å. (2026). *On phenomenographic analysis of group discussions: Patterns of collaborative meaning-making of physics during problem solving* [Unpublished manuscript].

Innabi, H., Marton, F., & Emanuelsson, J. (2026). Statistical Reasoning and the Variation Theory of Learning [Manuscript submitted for publication].

Marton, F. (2015). *Necessary conditions of learning*. Routledge. https://doi.org/10.4324/9781315816876

United Nations Statistics Division. (n.d.). *SDG Indicators: Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development*. United Nations, Department of Economic and Social Affairs, Statistics Division. https://unstats.un.org/sdgs/indicators/indicators-list/