



Universiteit
Leiden
The Netherlands

Quantifying students' collective creativity in designing soft matter experiments

Bakker, M.A.E.; Logman, P.S.W.M.

Citation

Bakker, M. A. E., & Logman, P. S. W. M. (2026). Quantifying students' collective creativity in designing soft matter experiments. *Journal Of Physics: Conference Series*, 3280.
doi:10.1088/1742-6596/3280/1/012039

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/4309650>

Note: To cite this publication please use the final published version (if applicable).

PAPER • OPEN ACCESS

Quantifying students' collective creativity in designing soft matter experiments

To cite this article: Anna Bakker and Paul Logman 2026 *J. Phys.: Conf. Ser.* **3280** 012039

View the [article online](#) for updates and enhancements.

You may also like

- [Student evaluation of research projects in a first-year physics laboratory](#)
Manjula D Sharma, Alberto Mendez, Ian M Sefton et al.
- [MobLen Model for Enhancing Scientific Creativity of Physics Students: An Alternative in the Covid-19 Pandemic](#)
Dwikoranto, B Jatmiko, E Hariyono et al.
- [Comment](#)

Quantifying students' collective creativity in designing soft matter experiments

Anna Bakker^{1,2,3*} and Paul Logman^{0000-0003-0699-9399, 1,2}

¹ Leiden Institute of Physics, Leiden University, Leiden, The Netherlands

² Leiden University Graduate School of Teaching, Leiden University, Leiden, The Netherlands

³ Christelijke Scholengemeenschap Walcheren, Vlissingen, The Netherlands

*anb@cswalcheren.nl

Abstract. Creativity is an important higher order cognitive skill in solving physics problems. However, it is notoriously difficult and time-consuming to measure. Nowadays, solving complex physics problems is mostly done by multiple research groups in parallel. In such contexts, it is essential that the research community as a whole generates creative solutions to the problem at hand, motivating the need for collective creativity. We apply various methods to quantify collective creativity to undergraduate students' open-ended experimental design proposals in the physics subfield of soft matter. We compare and discuss the three most promising of these methods. After comparison, a probability-based metric of the proposal distribution emerged as the most valid and distinctive quantification of collective creativity. It is capable of resolving year-to-year variations in collective creativity that are not readily apparent from qualitative inspection alone. This new metric can inform evidence-based decisions on how lab courses may be adapted to more effectively stimulate students' collective creativity.

1. Introduction

In physics research many complex problems need to be solved. These complex problems are generally solved by many groups of researchers working in parallel together to come up with creative solutions. This is also the case in the subfield of soft matter physics which tries to create microscopic explanations for macroscopic behaviour of materials, in order to design materials on a microscopic level to deliver specific macroscopically desired properties [1].

Various studies since the 1970's have shown that creativity can be nurtured to quite a large extent with the most recent thorough study showing that only about 23% variance in applied creativity has a genetic origin [2]. At our university we want to nurture applied creativity in the first year bachelor lab in the relatively new field of soft matter. In this lab course we give students an open inquiry assignment in the field of soft matter, thereby handing the students scientific agency to solve soft matter problems. Without this precondition, we would not be able to monitor our students' creativity, since students will be very limited in showing creativity without scientific agency. To be able to find out whether the setup of our course indeed stimulates students to be creative, we need to objectively measure students' creativity while designing their experiments. Once this can be done with sufficient precision, it becomes feasible to optimize the lab course in stimulating students' creativity.



2. Theoretical framework

Creative thinking is a higher order thinking skill, like problem-solving, critical thinking, and more [3]. Stimulating student creativity is essential as it promotes innovation, stimulates other higher order thinking skills, fosters personal development, and it better reflects the nature of real world problem solving in students' future workplaces. Lab courses which provide students with agency and decision-making opportunities are a natural environment for developing creativity and problem-solving skills [4, 5, 6].

Our aim is to stimulate students' collective creativity in the physics subfield of Soft Matter. Soft matter is characterized by non-rigid materials such as foams, (non-)Newtonian fluids, granular materials and colloids. [1]. Research in this field focuses on linking microscopic structure to emergent macroscopic phenomena in soft matter materials. In our first year Bachelor lab, the objective for students is to merely measure and characterize such macroscopic phenomena, such as surface tension, viscosity, impulse transfer, etc. [1]. To encourage students' creativity the students design their own experiment by combining a material with a physical phenomenon.

2.1 Collective creativity

Defining creativity remains an unsettled issue [7]. The American Psychological Association [8] defines creativity as "the ability to produce or develop original work, theories, techniques, or thoughts. A creative individual typically displays originality, imagination, and expressiveness". Contrarily, low creativity arises when students rely heavily on external influences such as peers, instructor examples and readily available online experiments. This leads to an overrepresentation of certain subjects and experimental designs. Accordingly, we aim to quantify originality. An individual can only be truly original if their contribution is unprecedented, which is characteristic of cutting-edge research. However, we do not expect this level of originality from our students, although isolated cases do occur [9]. Instead, our aim is to quantify if students are producing work that is original relative to their peers and the instructional context. Specifically, we examine the distribution of student-produced research along the two Soft Matter dimensions described above: materials and physical phenomena. We therefore adopt the framework of collective creativity, defined as creative innovation emerging from a group [10] and seek quantitative measures to capture it.

2.2 Measuring collective creativity

Measuring students' individual creativity remains an unsettled issue as well and there is a need for steps towards more concrete and robust measurement strategies, despite a large number of existing instruments [7]. In their literature review, Said-Metwaly et al. describe four distinct approaches to measuring creativity in their literature review: *process*, *person*, *product*, and *press* [7].

The first approach looks at cognitive *processes* underlying creativity. Measurements are done via divergent thinking tasks [11]. This approach may contain bias due to scoring and sample size. The second approach looks at *personal* characteristics behind creativity. Measurements are done via personality trait assessments [12]. Due to self-reporting this may result in subjective outcomes. The final approach (*press*) looks at the interaction between the creative individual and the context but this approach appears to suffer from a lack of research-based evidence.

The (third) *product*-based approach is most aligned with our needs as it can focus on the experiments students design resulting from an open-ended task that allows for creative outcomes. The Consensual Assessment Technique [13] relies on expert judges to rate the level of

creativity of these student products. However, judge selection is nontrivial and introduces subjectivity. Our aim is to develop a more objective and transparent method of quantifying the creativity of the students' products.

Individual creativity remains difficult to assess because all four approaches listed by Said-Metwaly et al. suffer from inherent subjectivity. We argue that this arises because creativity can only be meaningfully evaluated relative to peers and the instructional context. While the *product* approach is closest to our aim, we further reduce subjectivity by adopting the framework of collective creativity [10], analogous to modern physics research practices. We aim for an objective and sensitive measure of collective creativity, as it enables quantitative evaluation of how course design influences students' collective creative output.

3 Methods and Findings

In the studied soft matter lab course, first year bachelor students are given an open-inquiry assignment, granting them agency in addressing soft matter problems. Students work in pairs to encourage collaboration and multi-perspective problem solving.

Initially, guidance was provided through example experiments. To encourage the students to take agency in their experimental design, we replaced providing examples with an inspiration matrix, which we will discuss in section 3.1. Each pair of students submits a research proposal, which they execute after approval by the teaching team.

Over the years, the course underwent several structural changes. Some were intentional, like introducing the inspiration matrix in '19-'20 and introducing mandatory Q&A sessions in '24-'25. Others were incidental, including changes in instructors in '19-'20 and '24-'25 and Covid-restrictions in '20-'21. An overview of these changes is provided table 1.

Table 1: Overview of all changes made to the soft matter lab course over the years. Unless otherwise indicated, changes persist in later years. Omitted years saw no major changes.

Academic year	Changes made with respect to previous year
'19-'20	Introduction of the inspiration matrix. Removal of "ready-made" example ideas from syllabus. New lecturer.
'20-'21	Lecturer from '18-'19 returns. Due to Covid-restrictions, students had to perform their experiments alone, at home and with no access to the materials in our bachelor lab. Removal of inspiration matrix from syllabus; Reintroduction of example ideas.
'21-'22	Reintroduction of inspiration matrix, with positive and negative indications in the inspiration matrix of combinations of materials and phenomena of previously (un)successful experiments.
'22-'23	Removal of peer review during proposal writing process.
'24-'25	Mandatory Q&A-session on research question and feasibility before the students hand in their final proposal. New lecturer. Rewrites of the syllabus. Removal of positive and negative indications in the inspiration matrix.

3.1 Development of the inspiration matrix

During the initial design of the inspiration matrix for the '19-'20 academic year (Fig. 1a), we adopted two axes: materials and phenomena. The listed materials and phenomena were initially

selected based on instructor expertise and subsequently refined based on student-produced experiments, by adding categories, or removing or splitting existing categories, with the goal of maximizing the coverage of the feasible soft matter subfield. This subfield is constrained by the students' prior knowledge and experimental skills and the practical limitations of the lab. This process resulted in the matrix that was used in this research (Fig. 1b). For the '25-'26 academic year, the matrix was further updated to include additional phenomena, such as heat conduction and vortices, to encourage experiments targeting these properties (Fig. 1c).

The inspiration matrix has both a descriptive and a predictive function. We use its descriptive function to classify students' research proposals, after which we quantify their distribution across the matrix, thereby quantifying the collective creativity at the cohort level.

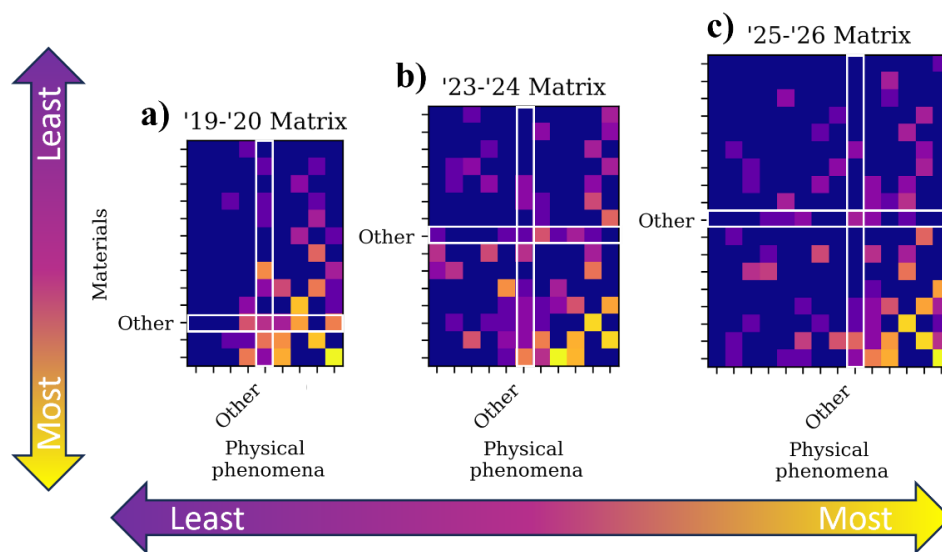


Figure 1. a) All students' proposals since '18-'19 classified using the '19-'20 (initial) inspiration matrix ($n=356$). Materials are sorted from least to most frequently used (top to bottom), and phenomena from least to most frequently used (left to right). The displayed matrices increase in size to reflect the increasing number of categories. For clarity, only the "other" category is indicated. This shows the "other" material category being the third most common category. This high frequency indicates that this early matrix does not yet capture the full range of materials. A full list of the categories can be obtained from the corresponding author. b) All students' proposals ($n=356$) classified using the '23-'24 inspiration matrix, which is used throughout this study. c) All student proposals ($n=356$) classified using the '25-'26 inspiration matrix. The further reduced frequency of the "other" material category reflects the improved descriptive coverage of this revised matrix, with the amount of proposals in the "other" material category decreasing from 46 to 10 between the '19-'20 and '25-'26 matrix, with. Because the difference in the probability metric result (Sec. 3.3) between the '23-'24 and '25-'26 matrix are only minor ($p_{123-124} = 4.6 \cdot 10^{-14}$ and $p_{125-126} = 7.9 \cdot 10^{-15}$ for the '25-'26 data set), and the '25-'26 matrix was not yet developed when this study started, the '25-'26 matrix is not used further within this manuscript.

3.2 Classification in the inspiration matrix.

At the start of the classification, all student research proposals were compiled per academic year. Each proposal was assigned a unique ID-number, but student names were retained to link IDs to proposals. Proposals were then categorized using the '23-'24 inspiration matrix, using the title, introduction, theoretical framework, and measuring method in the proposal. Categorization was performed by assigning the anonymous ID to the appropriate matrix cell. The number of proposals per matrix cell was counted to generate the heatmaps displayed in Fig. 1 and Fig. 2.

When a proposal plausibly belonged to two categories, the best-fitting category was selected. For the '25-'26 academic year this was done by both authors and a teaching assistant. The inter-

rater agreement was assessed using Krippendorff's alpha, yielding a value of 0.46, indicative of moderate agreement [14]. We hypothesize this imperfect agreement arises from several factors. First, fluid-related categories, like mixtures, suspensions and soap solutions, suffer from taxonomic ambiguity. This was independently noted by two raters and is reflected in the classification discrepancies within these categories. Second, in '24-'25 a new lecturer (Tab. 1) introduced vortices as a viable research topic, but this subject was not yet listed as a physical phenomenon in the '23-'24 matrix. Therefore, proposals involving vortices, which first appeared in the '24-'25 cohort, were difficult to consistently classify. Vortices and other topics are now listed in the updated '25-'26 matrix. Improving the matrix to decrease taxonomic ambiguity and increase coverage of the relevant soft matter subfield is an ongoing process. The currently moderate inter-rater agreement was explicitly incorporated in the error analysis (see Section 3.4). Lastly, the probability metric for collective creativity of the '25-'26 cohort (Sec. 3.3) changes only minimally between the '23-'24 and '25-'26 matrices ($p_{r23-r24} = 4.6 \cdot 10^{-14}$ and $p_{r25-r26} = 7.9 \cdot 10^{-15}$), so the '23-'24 matrix was used throughout this study.

Lastly, proposals classified in the "other" category of the '23-'24 matrix for either physical phenomena or materials typically fell out of the scope of soft matter, lab constraints or students' skill level. In most cases, these students were required to revise their experimental design and research question prior to execution, and such proposals were therefore excluded from the analysis. Nevertheless, the fact that most outliers (excluding the newly introduced vortices) lay beyond the soft matter domain suggests that the '23-'24 matrix adequately maps the feasible experimental space, providing evidence for the content validity of our approach.

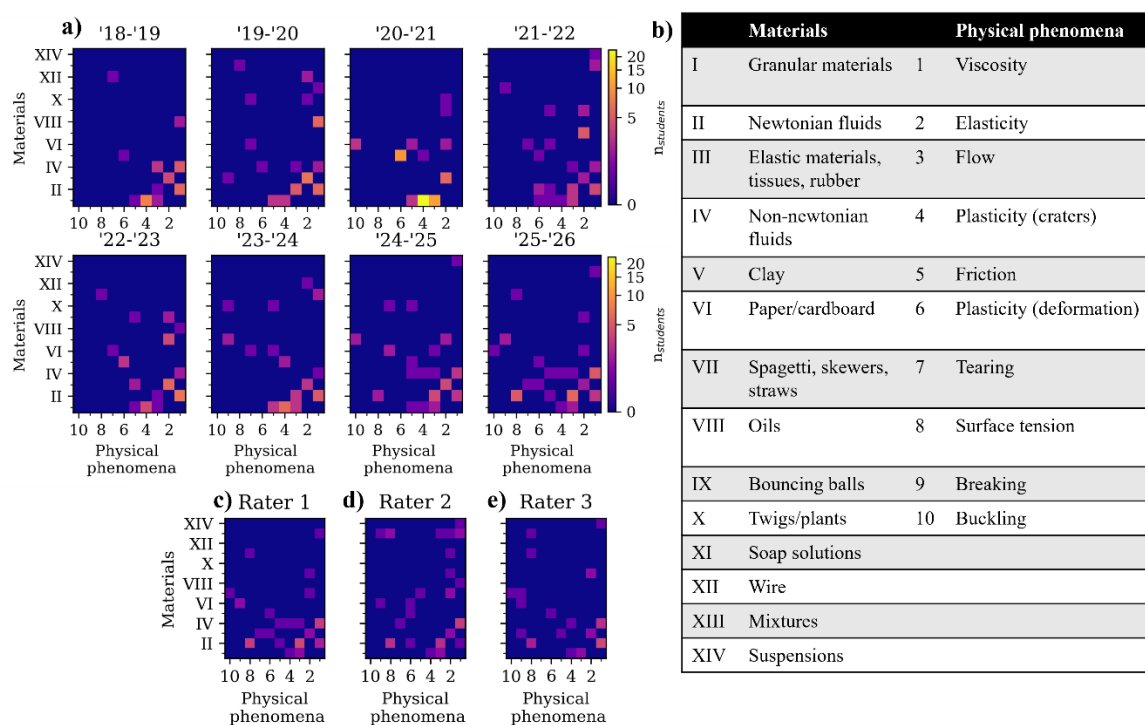


Figure 2. a) Students' experimental subject choices since '18-'19. Sample sizes are respectively $n = 34$, $n = 42$, $n = 63$, $n = 36$, $n = 37$, $n = 35$, $n = 34$, $n = 42$. A clear reduction in creativity is visible in '20-'21, coinciding with an increased sample size due to solo, at-home experiments during Covid restrictions. b) Material and physical phenomenon categories used to classify students' proposals, ordered from most to least frequent across the studied years. cde) '25-'26 students' proposals ($n=41$) classified by three independent raters. Differences are most pronounced for materials II, III, IV and XIII (respectively Newtonian fluids, elastic materials, non-Newtonian fluids and mixtures).

3.3 Quantification methods

To quantify the originality, and thus the collective creativity, of a student cohort we first classified all 356 students' research proposals collected over the eight academic years into the inspiration matrix and visualized the results as yearly heat maps (Fig. 2a).

These heat maps enable qualitative comparison of collective creativity across cohorts. For instance, the marked decline during the Covid lockdowns in the '20-'21 academic year is clearly visible (Tab. 1, Fig. 2a). However, to detect more subtle changes in collective creativity resulting from course modifications, additional quantitative measures are required.

There is no established statistical metric for collective creativity. We therefore drew inspiration from biodiversity studies, machine learning, geography and probability theory. In consultation with a statistician, we explored several candidate metrics, of which three passed an initial face validity assessment. A metric must operate on nominal variables to be suitable, as the inspiration matrix categories are nominal. The selected metrics share the property that their values increase as proposals become more evenly distributed across categories, indicating greater originality. The three approaches are the corrected Shannon index (eq. 1) [15], the cosine similarity (eq. 2) [16], and a probability-based measure quantifying the likelihood of a given matrix distribution arising at random (eq. 3). The same symbols are used for the same quantities throughout these equations, even if this means deviating from commonly used symbols. The outcomes of these three approaches are shown in Fig. 3e.

$$\text{corrected Shannon index} = \frac{H'}{H_{\max}} = \frac{-\sum_{i=1}^{N_c} (p_i \ln p_i)}{H_{\max}} = \frac{\sum_{i=1}^{N_c} \left(\frac{n_{p_i}}{N_p} \ln \frac{n_{p_i}}{N_p} \right)}{\ln N_c} \quad (1)$$

The Shannon index is commonly used to quantify biodiversity, where it measures how similar the abundances of different species in a community are [15]. Here, it measures how similar the abundances of each material-phenomenon combination are. In this case N_p denotes the number of student proposals, N_c denotes the number of categories (i.e. the number of phenomena times number of materials). Lastly, n_{p_i} denotes the number of proposals in the i^{th} category and p_i denotes the relative number of proposals in the i^{th} category.

$$\text{cosine similarity} = \cos(\theta) = \frac{(\mathbf{A} \cdot \mathbf{B})}{|\mathbf{A}||\mathbf{B}|} = \frac{\sum_{i=1}^{N_c} n_{p_i}}{\sqrt{\sum_{i=1}^{N_c} n_{p_i}^2} \cdot \sqrt{N_c}} \quad (2)$$

The cosine similarity represents the angle between the observed distribution ($\mathbf{A} = (n_{p_1}, n_{p_2}, \dots, n_{p_{N_c}})$) and the expected, idealized distribution ($\mathbf{B} = (1, 1, \dots, 1)$), if both distributions are represented by an N_c -dimensional vector (Fig. 3a). Additionally to the symbols listed above, n_{c_j} denotes the number of categories containing j proposals. Both the Shannon index and the cosine similarity metrics favour maximally dispersed research proposals, attaining their extrema when each material-phenomenon combination is chosen the same number of times, corresponding with independent and original thought processes taking place in a student cohort.

$$\text{probability} = \frac{\frac{N_c!}{\prod_{i=1}^{N_p} n_{c_i}!} \cdot \frac{N_p!}{\prod_{j=1}^{N_c} n_{p_j}!}}{N_c^{N_p}} \quad (3)$$

For the probability approach, we compute the probability that the observed distribution arises from random assignment of material-phenomenon combinations. A higher probability thus corresponds to a distribution consistent with an independent and original thought process, and therefore greater collective creativity.

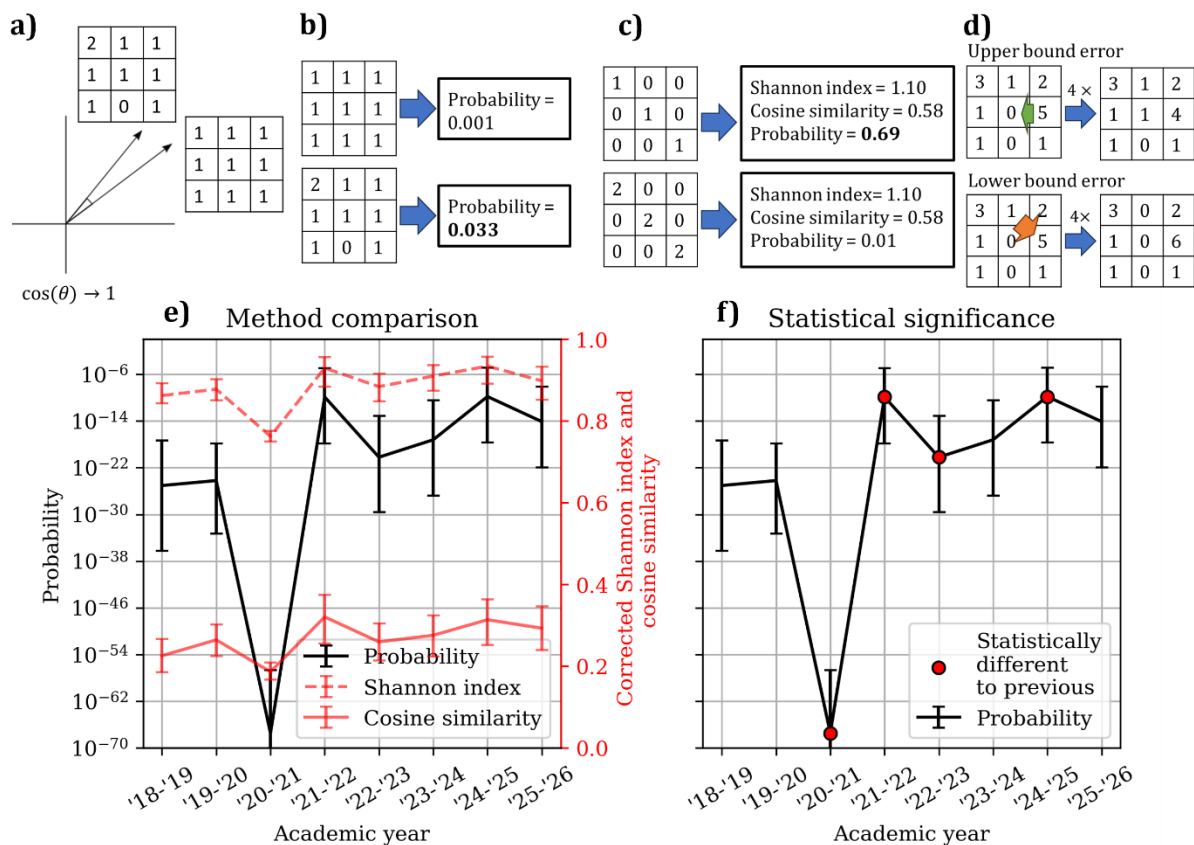


Figure 3 a) Schematic illustration of the cosine similarity calculation. b) Face validity assessment of the probability metric regarding uniform distribution. c) Face validity assessment of the cosine similarity, Shannon index and probability metric d) Schematic illustration of the Monte Carlo error estimation procedure which is repeated four times to estimate the error bars in Fig. 3ef and Fig. 4bc. The upper bound of this error bar is estimated from an adjusted matrix where four proposals are moved from the most-occurring category to four different “0” categories. The lower bound of this error bar is estimated from an adjusted matrix where four proposals are moved from categories with 1 proposals in it to the most-occurring category. e) Comparison of the three creativity metrics across academic years. f) Probability of each distribution shown earlier in Fig. 2a arising by chance. The probability of 1 corresponds to a perfectly random distribution, whereas low probabilities indicate reduced collective creativity. The pronounced dip in ‘20-’21 reflects Covid restrictions, while the statistically significant increase in ‘24-’25, not apparent in Fig. 2a, is clearly resolved here.

To assess validity, we argue that low creativity leads to overrepresentation of specific combinations (Sec. 2.1). In contrast, independent and original thinking does not imply deliberate avoidance of peers’ choices. Consequently, a perfectly uniform distribution is not expected for large cohorts. Unlike the Shannon index and cosine similarity, the probability calculation

explicitly accounts for this effect (Fig. 3bc). As a result, the probability metric demonstrates both face and construct validity.

Moreover, the cosine similarity and Shannon index do not scale appropriately with the number of proposals. Intuitively, the lower distribution in Fig. 3c should reflect less creativity than the upper distribution, as we expect a larger group of students ($n = 6$ versus $n = 3$) to cover a wider range of categories. However, both metrics (unlike the probability-based metric) yield identical values for these cases, indicating limited face validity.

Given the limited validity of the Shannon index and cosine similarity, we therefore proceed exclusively with the probability-based metric.

3.4 Error analysis

To distinguish statistically significant data points over the years despite the moderate inter-rater agreement (Sec. 3.2), we use a Monte Carlo approach to quantify the effect of inter-rater classification uncertainty by simulating minimal reclassifications. We simulate a decrease in the probability metric starting by moving one proposal from a least-occurring category ($n_{c_j} > 0$) to the most-occurring category to calculate the lower bound of the probability metric and we simulate an increase in the probability metric by moving one proposal from the most-occurring category to an empty category ($n_{c_j} = 0$) (Fig. 3d) to calculate the upper bound of the probability metric. Iterated minimal reclassifications increase these upper and lower bounds. This procedure is iterated until the probability metric of the classifications by the three different raters of the '25-'26 dataset fall within these upper and lower bounds (Fig. 4c). This was the case after four iterations (Fig. 3d). Four iterations were then used to calculate the upper and lower bounds of the probability metric over all years (Fig. 3ef, Fig. 4bc). Furthermore, this method is also used to calculate the upper and lower bounds for the Shannon index and cosine similarity metrics.

Finally, statistical significance is assessed by year-to-year comparison. A datapoint is deemed significant if it lies outside the upper and lower bounds of the preceding year (Fig. 3f), yielding four significant data points ('20-'21, '21-'22, '22-'23 and '24-'25).

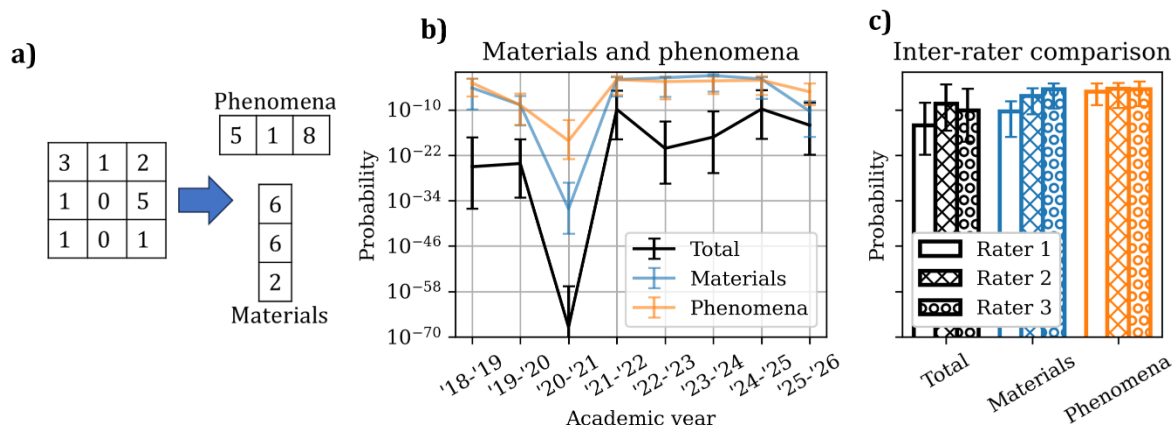


Figure 4 a) Schematic illustrations of the row and column summations underlying the creativity quantification of materials and phenomena shown in Fig. 4bc. b) Probability of each distribution shown earlier in Fig. 2a arising by chance, shown separately for the total, material and phenomenon distributions. c) Calculated probabilities for the different raters (Fig. 2cde) of the '25-'26 data. The close agreement in phenomenon probabilities across raters is evident.

3.5 Phenomena and materials

Lastly, we quantify the creativity of the distribution of students' choices in physical phenomena and materials. To assess creativity in phenomena selection, we count how often each physical phenomenon is chosen, independent of material choice. The probability and associated uncertainty of this distribution are then calculated using the same procedure as for the full proposal distribution. A similar analysis is performed for the material choices (Fig. 4bc).

Analysing the phenomena and materials separately yields additional insights. For example, creativity in materials decreased more strongly than creativity in phenomena during the Covid-restricted year '20-'21. This is consistent with the instructional context, as students were more constrained in available materials (household items for at-home experiments) than in accessible phenomena. Furthermore, we observe that the three raters who classified the '25-'26 proposals show substantially higher agreement for phenomena than for materials, consistent with our qualitative observations (Sec. 3.2).

Lastly, a Cramér's V of 0.53 indicates a relatively strong association [17] between the materials and phenomena, as is physically expected. Nevertheless, students are able to combine some materials with multiple phenomena, and vice versa, to design their experiments (Fig. 1c).

4 Conclusion and outlook

While clear definitions of creativity remain elusive [7], we provide a functional definition of collective creativity: the ability of students to produce work that is original relative to their peers and the instructional context, which shows as a random distribution of research proposals at the cohort level.

Among the three candidate metrics, the probability-based metric proved the most informative for quantifying collective creativity. It exhibits the highest face and construct validity, as it most closely aligns with our definition. The construct validity is further supported by the statistically significant decrease in collective creativity during COVID-19 lockdowns. We applied the same framework to quantify creativity in distributions of physical phenomena and materials. The other methods lacked face validity by failing under varying numbers of proposals and lacked construct validity by rewarding uniform distributions, which are not expected for large cohorts (Fig. 3bc).

The probability based metric constitutes an extension of existing product-based creativity instruments [7]. In its current implementation, it shows moderate inter-rater agreement (Krippendorff's $\alpha = 0.46$), which can be improved to substantial agreement by improving the descriptive resolution of the inspiration matrix by adding new categories (e.g. vortices) and by defining all categories mutually exclusively (e.g. non-Newtonian fluids versus mixtures).

Our results support the choice of materials and phenomena as the two primary axes of the inspiration matrix for soft matter. While any physics subfield is governed by multiple dimensions, we recommend selecting the two most dominant aspects, one of which is typically physical phenomena. At present, we have created and are further developing matrices for the physics fields of Oscillations and Waves (axes are types of oscillations and waves & physical phenomena), Electricity, Magnetism & Electromagnetism (axes are measuring techniques & physical phenomena), Optics & Quantum mechanics (axes are measuring techniques & physical phenomena) and others. In this respect, the inspiration matrix is also proving to be useful in other areas of physics within physics education.

The observed general upward trend in collective creativity over the years aligns with the intended effects of course redesigns, including the reintroduction of the inspiration matrix in '21-

'22 and enhanced individualized guidance via Q&A sessions in '24-'25, both corresponding to a statistically significant increase. These findings suggest that the inspiration matrix effectively stimulates creative experiment design, and that the Q&A sessions work! We have therefore implemented both elements across all open-ended laboratory assignments at our university.

Acknowledgments

We would like to thank statistician Chip Huisman for his advice on the quantification of collective creativity. Furthermore we would like to thank Parvin Ghalebтарash for performing the independent classification of the '25-'26 student proposals into the inspiration matrix. Analysis scripts can be obtained from a GitHub repository [18].

References

- [1] Barrat, J. L., et al. (2024). Soft matter roadmap. In JPhys Materials (Vol. 7, Issue 1). Institute of Physics. <https://doi.org/10.1088/2515-7639/ad06cc>
- [2] Velázquez, J. A., Segal, N. L., & Horwitz, B. N. (2015). Genetic and environmental influences on applied creativity: A reared-apart twin study. *Personality and individual differences*, 75, 141–146. <https://doi.org/10.1016/j.paid.2014.11.014>
- [3] Marzano, R. J., & Kendall, J. S. (2007). The new taxonomy of educational objectives (2nd ed.). Corwin Press.
- [4] Agustian, H. Y., Finne, L. T., Jørgensen, J. T., Pedersen, M. I., Christiansen, F. V., Gammelgaard, B., & Nielsen, J. A. (2022). Learning outcomes of university chemistry teaching in laboratories: A systematic review of empirical literature. *Review of Education*, 10(2), e3360. <https://doi.org/10.1002/rev3.3360>
- [5] Zoller, U., & Pushkin, D. (2007). Matching higher-order cognitive skills (HOCS) promotion goals with problem-based laboratory practice in a freshman organic chemistry course. *Chemistry Education Research and Practice*, 8(2), 153–171. <https://doi.org/10.1039/B6RP90028C>
- [6] Soomro, S. A., Casakin, H., Nanjappan, V., & Georgiev, G. V. (2023). Makerspaces Fostering Creativity: A Systematic Literature Review. *Journal of Science Education and Technology*, 32(4), 530–548. <https://doi.org/10.1007/s10956-023-10041-4>
- [7] Said-Metwaly, S., Noortgate, W. van den, & Kyndt, E. (2017). Approaches to Measuring Creativity: A Systematic Literature Review. *Creativity. Theories – Research - Applications*, 4(2), 238–275. <https://doi.org/10.1515/ctra-2017-0013>
- [8] American Psychological Association. (2018, April 19). APA Dictionary of Psychology: creativity. <https://www.apa.org/topics/creativity>
- [9] Goesaert W.M. & Logman P.S.W.M. (2024), Easy method to establish the dispersion relation of capillary waves on water jets, *American Journal of Physics* 92(2): 93-99. <https://doi.org/10.1119/5.0144849>
- [10] Acar, O. A., et al. (2024). Collective Creativity and Innovation: An Interdisciplinary Review, Integration, and Research Agenda. In *Journal of Management* (Vol. 50, Issue 6, pp. 2119–2151). SAGE Publications Inc. <https://doi.org/10.1177/01492063231212416>
- [11] Guilford, J. P. (1967). Creativity: Yesterday, Today and Tomorrow. *The Journal of Creative Behavior*, 1(1), 3–14. <https://doi.org/10.1002/j.2162-6057.1967.tb00002.x>
- [12] John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The Big Five Inventory - Versions 4a and 54. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research. <https://doi.org/10.1037/t07550-000>
- [13] Amabile, T. M. (1982). Social psychology of creativity: A Consensual Assessment Technique. *Journal of Personality and Social Psychology*, 43, 997–1013. <https://doi.org/10.1037/0022-3514.43.5.997>
- [14] Krippendorff, K. (1970). Estimating the reliability, systematic error and random error of interval data. *Educational and Psychological Measurement*, 30(1), 61–70. <https://doi.org/10.1177/001316447003000105>
- [15] Krebs, C. J. (1989). *Ecological methodology*. Harper & Row.
- [16] Tan, P.-N., Steinbach, M., & Kumar, V. (2014). *Introduction to Data Mining*. Pearson.
- [17] Rea, L.M., Níames, Parker, R.A., & Allen, R. (2016). *Designing and conducting survey research*. San Francisco: Jossey-Bass.
- [18] Bakker, A. Quantifying creativity analysis [Computer software]. <https://github.com/Unitato99/Quantifying-creativity/>