Proceedings of ASME 2018 International Design Engineering Technical Conferences & Design Conference
August 26-29, 2018, Quebec City, Canada

DRAFT - DETC2018 - 86061

IS THERE TOO MUCH VARIETY IN THE EXECUTION OF THE VARIETY METRIC?

Sharath Kumar Ramachandran & Scarlett Miller
School of Engineering Design, Technology and Professional Programs
The Pennsylvania State University,
State College, Pennsylvania, 16802
Email: sharath@psu.edu; shm13@psu.edu

Samuel T. Hunter
Industrial and Organizational Psychology
The Pennsylvania State University,
State College, Pennsylvania, 16802
Email: hunter@psu.edu

Mark Fuge
Dept. of Mechanical Engineering
University of Maryland
College Park, Maryland, 20742
Email: fuge@umd.edu

INTRODUCTION
Creativity is a crucial part of the engineering design process as it enables innovation [1] and is the quintessence of new businesses [2]. As such, researchers have devoted substantial effort to developing and testing methods for supporting creative idea development through methods such as TRIZ [3], SCAMPER [4], Brainstorming [5], Brainsketching [6], C-sketch [7], Product Dissection [8], Design Heuristic cards [9, 10] and Gallery [6] techniques. As the methods to develop these solutions increase, the necessity to develop rigorous methods for evaluating the success of these methods also increases [11], because the success of research in this field hinges on the quality of the measurements used to interpret the research findings [11]. However, the robustness of current engineering creativity metrics, or their ability to produce consistent results and avoid misconceptions in the use and deployment of the metric, has been brought to question. The goal of the current study was to determine the rigor and repeatability of the variety metric proposed by Shah, Vargas-Hernandez and Smith (SVS) [12], and its adaptation by Nelson et al. [13], and identify how susceptible the metric is to misinterpretations of its calculation.

MEASURING CREATIVITY IN ENGINEERING DESIGN
Expanding and exploring the design space in the early stages of the design process is necessary to develop successful designs [14]. Because of this, engineering researchers have sought to capture the *variety* of design ideas generated throughout engineering design activities, or how "explored the solution space" is during the idea generation process (pg. 117, [12]). The variety metric is used to counterbalance the quantity measure in design studies [15], because increases in the fluency of ideas must also proportionally increase the *spread* of the ideas [16]. This is important because there is a higher possibility of solving a design problem when a more discrete set of ideas is produced in the initial stages of the design process [14]. This type of divergent thinking has been shown to directly translate to the production of successful design solutions [16, 17].

The engineering community has largely adopted a genealogy tree approach to measuring design variety in an effort to reduce the subjectivity of creativity assessment [15]. The variety metrics commonly adopted in the engineering literature [12] break design variety into four hierarchical branches: the physical principle, followed by the working principle, embodiment and detail. SVS [12] suggested that values of 10, 6, 3 and 1 are assigned to each of these levels, respectively, to ensure that "separation at higher levels will always score a greater total" (pg. 126). In other words, the metric seeks to reward ideas that have more variety at higher abstraction levels. The justification for choosing these values is that ideas that differ in physical principle diverge by a greater extent in the design space [12, 13].

One of the largest challenges with the deployment of metrics that utilize this hierarchical structure is that the metrics commonly deployed throughout the engineering literature [12, 13] fail to define exactly how the principles are defined (e.g., working, physical and/or embodiment) or provide a framework (e.g., [18]) for how to define these principles which may lead to subjectivity in the deployment of these metrics [19]. This is further confused because SVS also mentions that in cases where distinguishing between physical and working principles becomes arduous, only working principles and embodiment can be considered [20]. Because of this, it is up to authors to carefully describe and define these principles in their own work to ensure repeatability of their findings. However, in a review of the literature from technical papers published over the past two decades in IDETC, JMD, IJEE and so forth [11, 14, 15, 21-42], 20 of 25 papers that administered these metrics, failed to provide examples or definitions of the physical and working principles in their study. This is a large problem in design research because "Important qualities characterized by rigorous research are validity and reliability" [43].

CASE STUDY

The purpose of this case study was to identify what happens as variety is added to the calculation of the variety metric. In other words, what happens if a researcher inappropriately defines the physical and working principle levels and in effects 'flips' or modifies the tree? Or what happens if a researcher does not define these principles and thus leaves it up to interpretation when one deploys these metrics. Will a significantly different variety score be yielded giving the criteria that the variety metric definition and its subsequent weights were provided to provide higher scores for design sets that deploy variety at higher levels of the tree?

In order to determine the impact of these variations, the variety metric was calculated in two different way using a dataset derived from a previous study. Specifically, variety was calculated from an idea set with 934 ideas from 89 first-year students and 52 senior students. The variety metric was calculated in both instances using the method described by Nelson et al. [13]. However, there were two key differences in how variety was calculated, as shown in Figures 1 and 2. Specifically, in Figure 1, the type of motion for the design was used for differentiation at the physical principle level while the type of power was used for differentiation at the working principle level. On the contrary, as Figure 2 demonstrates, the tree physical and working principle definitions were switched. This was done because the authors when analyzing their own work, struggled to define these principles appropriately for the design problem at hand. Thus, in the current study, the metric was calculated both ways and the results were compared.

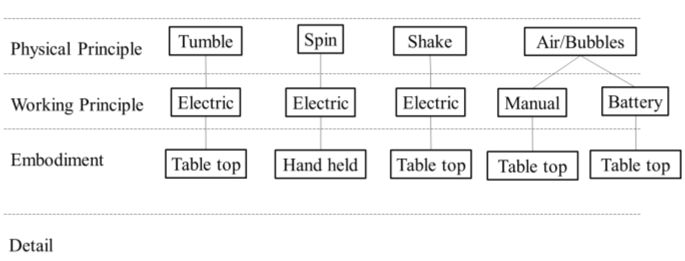

Figure 1. Genealogical tree with assumed physical and working principles resulted in a variety score, V=35.

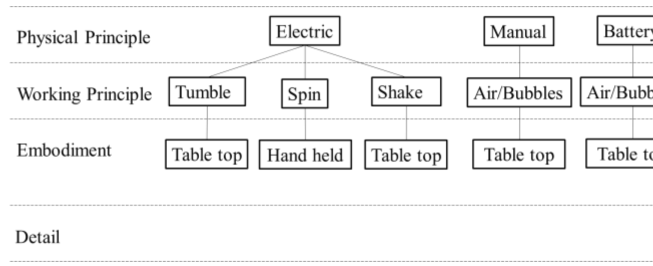

Figure 2. Genealogical tree with physical and working principles switched resulted in a variety score, V=30.

To assess the degree to which both variety scores provided similar values (i.e., interrater reliability) an intraclass correlation coefficient (ICC2) was computed. Results showed that the two methods produced similar ratings to one another as depicted by meeting the threshold of 0.70 (ICC2 = 0.934). While these results are good in that *large* changes to the definitions of the principles do not result in *large variations* in the computation to the variety metric, they bring to question what this metric is actually measuring and if it is really rewarding 'high levels of abstraction.'

CONCLUSION

Exploring the breadth of the design space is essential to produce creative and successful designs [14]. Hence, the need to measure variety is of significant importance. However, the metrics currently used in research have certain implicit assumptions leading to inconsistencies. As indicated in the example above large variations in the deployment of the definitions of the variety metric did not yield significant difference in the variety score. If the variety metric is really "rewarding higher levels of abstraction" then changes to the tree should result in significantly different variety scores. However, results from the study suggest this is not always true. Currently, there are no strict guidelines for the evaluation of ideation methods [29] and the variety scoring relies heavily on the rater's implicit definition of the metric, indicating that the ratings are subjective in nature and, by proxy, potentially biased. Moreover, most studies have failed to provide these definitions and example trees of how the calculation occurred. The consistency of metrics using genealogical trees relies on the fundamental definitions of what constitutes to physical principle, working principle, embodiment and detail. Given the importance of variety in creative idea generation established earlier, this calls for more research that test the rigor and repeatability of existing metrics as well as their predecessors.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1728086.

REFERENCES

[1] Amabile, T. M., 1996, Creativity in context: Update to the social psychology of creativity, Hachette UK.

[2] Brands, R. F., and Kleinman, M. J., 2010, Robert's Rules of Innovation: A 10-Step Program for Corporate Survival, John Wiley & Sons.

[3] Altshuller, G. S., 1984, Creativity as an exact science: the theory of the solution of inventive problems, Gordon and Breach.

[4] Eberle, B., 1996, Scamper on: Games for imagination development, Prufrock Press Inc.

[5] Osborn, A., 1957, Applied Imagination, Scribner, New York, NY.

[6] VanGrundy, A., 1988, Techniques of Structured Problem Solving, Van Nostrand Reinhold Company, New York, NY.

[7] Kulkarni, S., Summers, J., Vargas-Hernandez, and Shah, J., 2000, "Evaluation of collaborative (C-Sketch) as an idea generation technique for engineering design," Journal of Creative Behavior, 35(3), pp. 168-198.

[8] Toh, C. A., and Miller, S. R., "Exploring the utility of product dissection for early-phase idea generation," Proc. ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers.

[9] Daly, S. R., Yilmaz, S., Christian, J. L., Seifert, C. M., and Gonzalez, R., 2012, "Design heuristics in engineering concept generation," Journal of Engineering Education, 101(4), pp. 601-629.

[10] Daly, S. R., Christian, J. L., Yilmaz, S., Seifert, C. M., and Gonzalez, R., 2012, "Assessing design heuristics for idea generation in an introductory engineering course," International Journal of Engineering Education, 28(2), p. 463.

[11] Jablokow, K., Teerlink, W., Yilmaz, S., Daly, S., Silk, E., and Wehr, C., 2015, "Ideation Variety in Mechanical Design: Examining the Effects of Cognitive Style and Design Heuristics," (57106).

[12] Shah, J. J., Smith, S. M., and Vargas-Hernandez, N., 2003, "Metrics for measuring ideation effectiveness," Design studies, 24(2), pp. 111-134.

[13] Nelson, B. A., Wilson, J. O., Rosen, D., and Yen, J., 2009, "Refined metrics for measuring ideation effectiveness," Design Studies, 30(6), pp. 737-743.

[14] Henderson, D., Helm, K., Jablokow, K., McKilligan, S., Daly, S., and Silk, E., "A Comparison of Variety Metrics in Engineering Design," Proc. ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers.

[15] Hernandez, N. V., Okudan, G. E., and Schmidt, L. C., "Effectiveness metrics for ideation: Merging genealogy trees and improving novelty metric," Proc. ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 85-93.

[16] Torrance, E., 1972, "Predictive validity of the Torrance tests of creative thinking," The Journal of Creative Behavior, 6(4), pp. 236-262.

[17] Acar, S., and Runco, M. A., 2017, "Latency predicts category switch in divergent thinking," Psychology of Aesthetics, Creativity, and the Arts, 11(1), p. 43.

[18] Hirtz, J., Stone, R. B., McAdams, D. A., Szykman, S., and Wood, K. L., 2002, "A functional basis for engineering design: reconciling and evolving previous efforts," Research in engineering Design, 13(2), pp. 65-82.

[19] Naim, A. M., and Lewis, K. E., "Rethinking Design: The Formal Integration of Engineering Innovation into a Design Process," Proc. ASME 2011 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 465-477.

[20] Shah, J. J., "Identification, measurement and development of design skills in engineering education," Proc. DS 35: Proceedings ICED 05, the 15th International Conference on Engineering Design, Melbourne, Australia, 15.-18.08. 2005.

[21] Shah, J. J., Smith, S. M., Vargas-Hernandez, N., Gerkens, D. R., and Wulan, M., "Empirical studies of design ideation: Alignment of design experiments with lab experiments," Proc. ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 847-856.

[22] Song, S., and Agogino, A. M., "Insights on designers' sketching activities in new product design teams," Proc. ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 351-360.

[23] Chusilp, P., and Jin, Y., 2005, "Impact of Mental Iteration on Concept Generation," Journal of Mechanical Design, 128(1), pp. 14-25.

[24] Okudan, G. l. E., Ogot, M., and Shirwaiker, R., "An investigation on the effectiveness of design ideation using TRIZ," Proc. ASME 2006 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. 953-961.

[25] Nelson, B., Wilson, J., and Yen, J., "A study of biologically-inspired design as a context for enhancing student innovation," Proc. Frontiers in Education Conference, 2009. FIE'09. 39th IEEE, IEEE, pp. 1-5.

[26] Okudan, G. E., Chiu, M.-C., Lin, C.-y., Hernandez, N. V., Schmidt, L. C., and Linsey, J., "A pilot exploration of systematic ideation methods and tools on design learning," Proc. Information Technology Based Higher Education and Training (ITHET), 2010 9th International Conference on, IEEE, pp. 102-107.

[27] Peeters, J., Verhaegen, P.-A., Vandevenne, D., and Duflou, J., 2010, "Refined metrics for measuring novelty in ideation," IDMME Virtual Concept Research in Interaction Design, Oct, pp. 20-22.

[28] Duflou, J. R., and Verhaegen, P. A., 2011, "Systematic innovation through patent based product aspect analysis," CIRP Annals, 60(1), pp. 203-206.

[29] Verhaegen, P.-A., Peeters, J., Vandevenne, D., Dewulf, S., and Duflou, J. R., 2011, "Effectiveness of the PAnDA ideation tool," Procedia Engineering, 9, pp. 63-76.

[30] Hernandez, N. V., Schmidt, L. C., and Okudan, G., "Experimental assessment of TRIZ effectiveness in idea generation," Proc. ASEE Annual Conference, San Antonio, TX.

[31] Schmidt, L. C., Hernandez, N. V., and Ruocco, A. L., 2012, "Research on encouraging sketching in engineering design," AI EDAM, 26(3), pp. 303-315.

[32] Wodehouse, A., and Ion, W., 2012, "Augmenting the 6-3-5 method with design information," Research in Engineering Design, 23(1), pp. 5-15.

[33] Fuge, M., Stroud, J., and Agogino, A., "Automatically inferring metrics for design creativity," Proc. ASME 2013 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers.

[34] Liu, A., and Lu, S. C. Y., 2013, "Impacts of Synthesis Reasoning on Ideation Effectiveness in Conceptual Design," Journal of Mechanical Design, 135(6), pp. 061009-061009-061011.

[35] Verhaegen, P.-A., Vandevenne, D., Peeters, J., and Duflou, J. R., 2013, "Refinements to the variety metric for idea evaluation," Design Studies, 34(2), pp. 243-263.

[36] Doboli, A., and Umbarkar, A., 2014, "The role of precedents in increasing creativity during iterative design of electronic embedded systems," Design Studies, 35(3), pp. 298-326.

[37] Jagtap, S., Larsson, A., Hiort, V., Olander, E., and Warell, A., 2015, "Interdependency between average novelty, individual average novelty, and variety," International Journal of Design Creativity and Innovation, 3(1), pp. 43-60.

[38] Verhaegen, P.-A., Vandevenne, D., Peeters, J., and Duflou, J. R., 2015, "A Variety Metric Accounting for Unbalanced Idea Space Distributions," Procedia Engineering, 131, pp. 175-183.

[39] Grover, R., Emmitt, S., and Copping, A., 2016, "Learning from the past: A conceptual model of independent typology formation in the design studio," Integrated Design at 50: Building our Future, pp. 210-226.

[40] Lim, S. Y. C., Camburn, B. A., Moreno, D., Huang, Z., and Wood, K., "Design Concept Structures in Massive Group Ideation," Proc. ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers.

[41] Coelho, D. A., and Vieira, F. L., 2018, "The effect of previous group interaction on individual ideation novelty and variety," International Journal of Design Creativity and Innovation, 6(1-2), pp. 80-92.

[42] Vargas Hernandez, N., Schmidt, L., Okudan, G., and Linsey, J., "Systematic ideation curriculum effectiveness investigation & deployment to enhance design learning," Proc. Proceedings of IDETC/CIE 2010 ASME Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Montreal, Quebec, Canada, Aug, pp. 15-18.

[43] Fallman, D., and Stolterman, E., 2010, "Establishing criteria of rigour and relevance in interaction design research," Digital Creativity, 21(4), pp. 265-272.