Design Creativity Workshop 2012
June 6, 2012, Texas A&M University, College Station, Texas, USA

CreativeIT Tools for Assisting and Evaluating Creativity and Problem Framing in Early-Stage Human-Centered Design

Mark Fuge and Alice Agogino
University of California at Berkeley, Berkeley, CA

While there has been an active body of recent work in creativity tools and evaluation metrics for embodiment design, the scope of design extends beyond that of just embodiment, but also to that of service, business model, and experience design. Our hypothesis is that the same aspects of creativity measurement and stimulus that have been applied to embodiment design can be expanded to this larger realm of conceptual design, provided that they are adapted appropriately. We present some of our steps in this direction: 1) Meta4Explorer, a metaphor-based creativity stimulus tool that assists design teams in problem framing and concept generation, and 2) some initial results of a creativity evaluation metric for this domain, demonstrating the importance of user needs and product/service ecosystems in defining conceptual variety. We hope to use the workshop as a forum to discuss how existing creativity tools and evaluation mechanisms can be leveraged outside of typical engineering domains, addressing a broader area of conceptual design.

1. CreativeIT, Creativity, and Human Centered Design

The development of CreativeIT tools – software that improves human creativity – has become increasingly important in industry and academia alike (Benami 2002, Brown 2011, Kurtoglu et. al. 2009). Recent advances in AI, visualization, computational linguistics, and other scientific fields have spurred a growth in CreativeIT tools, particularly in areas such as Design by Analogy or Bio-mimetic Design (Linsey et. al. 2008, Vattam et. at. 2010). Likewise, the need to evaluate and improve CreativeIT tools has encouraged the development of metrics for measuring a tool's effect on creativity. Within the domain of mechanical product design, prior research has defined several potential metrics for measuring the creativity of different sets of mechanical products, including aspects of creativity such as novelty, usefulness, variety, and quality (Shah

2003, Sarkar and Chakrabarti 2008, 2011, Maher 2010). Many empirical studies have used these metrics to quantify the effects of analogy, fixation, timing, solution quality, sketching, and many other factors on human design creativity. These studies of human problem-solving are critical in improving CreativeIT and idea generation methods, and they would not have been possible without a foundation of creativity metrics.

Unfortunately, many of these tools and metrics are created to affect or measure creativity within the stage of design referred to as *embodiment design*, where the user needs and functionality have already been selected. When this is not the case, such as during earlier stages in the design process, many of the current tools and metrics are difficult to apply; in particular, the dominant metric of *concept variety* introduced by Shah (2003) is not particularly appropriate for sets of concepts produced prior to embodiment design.

To see this why this might be the case, consider the following representative mission of a design team we studied: "improve user satisfaction with bus transportation." After conducting several user interviews, the team generated many conceptual designs around a few potential user needs, such as helping short-statured users maintain stability or providing adequate space during crowded bus times. If the team decides to focus on stability during embodiment design, current metrics can reliably judge concept variety by comparing the different physical or working principles used to provide stability (Shah 2003). It is less obvious how these metrics should be applied when comparing concepts that solve different user needs. For example, solutions for providing extra passenger stability and solutions for increasing space during crowded bus times both help fulfill the overall goal of increasing user satisfaction; however, they cannot be accurately compared on variations in their physical or working principles. As a result, the practice of comparing physical and working principles no longer provides a good measure for concept variety.

Likewise, concept generation tools such as Design by Analogy, TRIZ, or Bio-mimetic Design are incredibly useful tools for finding alternative ways of stabilizing a human being in a moving vehicle. However, they are less useful for helping a design team transition from "providing extra passenger stability" to solutions that "increase space during crowded bus times," since these goals are not functionally related. This type of shift is referred to as problem framing, a core part of Human-Centered Design, and has large impacts on the resulting cost of the product (Hey et. al. 2008), yet isn't assisted by current CreativeIT tools.

Our work attempts to bring some of the concepts from design creativity in embodiment design to the earliest stages of design, where the functional specifications are not well defined. To stimulate workshop discussion about this area, we plan on presenting two recent pieces of work: 1) a metaphor-based interactive CreativeIT application, called Meta4Explorer, that leverages the power of conceptual metaphor to assist teams in product re-framing, and 2) experimental results on conceptual variety in early-stage conceptual design which indicate that both the specific product-service ecosystem as well as the need addressed by the design impact the perceived variety of a concept among design experts. We present these results with the goal of spurring interest and conversation around how to extend current tools to earlier stages of design, as well as what kind of metrics are most appropriate for judging creativity in those earlier stages.

2. Meta4Explorer – Leveraging Metaphors for Problem Framing

In exploring how CreativeIT tools can be leveraged to promote creativity in early-stage conceptual design, we leveraged an interesting parallel between analogies and metaphors in design (Hey et. al. 2008). In both cases analogies and metaphors are used to create abstractions about an object, and it is through that process of abstraction that designers are stimulated to apply design principles across domain – studies conducted in the Design by Analogy field have demonstrated this effect on creativity, primarily in functional design synthesis. Similarly, qualitative evidence exists for a similar process of how designers and business people use metaphors to impart attributes and qualities to their designs or customers, helping frame the design problem in a different way.

Prior work has shown that associative exploration tools, such as the WordTree Design by Analogy tool (Linsey 2008), can help designers explore and different, yet analogous solution in different domains, a process which helps lead to creative ideas. In a similar vein, we hypothesized that a similar approach could be applied to the related concept of metaphors, this time focusing on problem re-framing, rather than on functional idea generation.

Towards this end, we adapted a variety of prior computational linguistics research to reconstruct and effectively visualize potential product metaphors. Specifically, we build off the work of Veale and Hao (2008, 2011) by constructing a bi-partite graph of noun-adjective relationships captured from free-text using Hearst patterns and popular search engines, such as Google, Bing, and Yahoo (Veale and Hao 2008). This bi-partite network is then visualized through an interactive web application, Meta4Explorer (www.meta4explorer.com), that allows designers to type in the characteristics that they want their design to embody, and then providing an interactive graph of potential metaphors that users can explore. For an example of the interface, see Fig. 1. For more on the use of metaphors in design, we direct interested readers to Hey et. al. (2008).

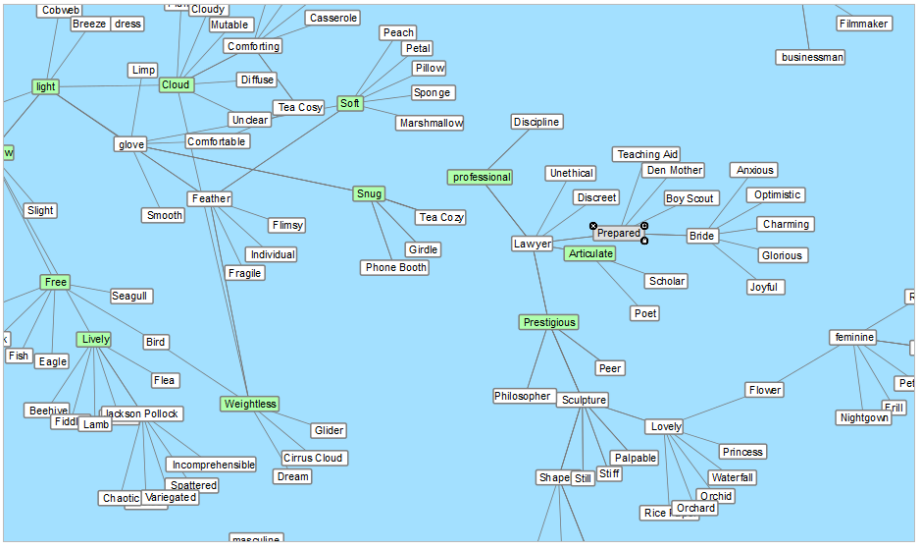

Fig 1 A snapshot of the Meta4Explorer tool. Designers can click on various nodes in the graph to find related attributes or metaphors. Designers are then able to use the metaphor

to re-frame the original design problem, resulting in creative stimulus.

We are currently running experiments which are designed to test various aspects of metaphor generation on concept generation and creativity, and as such do not yet have results to report to that effect. We wish to use the workshop as a venue to discuss how other existing CreativeIT tools might be adapted to earlier stages of the Human-Centered Design process.

3. Evaluating Creativity in the Conceptual Stages of Human-Centered Design

While the above section demonstrated that CreativeIT tools can have a role in the earlier stages of Human-Centered Design, one of major roadblocks to progress in this area is that, unlike embodiment design, the research community has yet to define agreed upon, reliable metrics for judging the creativity of concepts generated during this stage in design. This represents a major intellectual challenge that we believe the community at the Design Computing and Cognition conference, and particularly the Design Creativity workshop participants, is well suited to address.

In evaluating our above Meta4Explorer tool, it became clear that we needed creativity metrics that were broader than the embodiment design metrics currently available (Shah 2003, Sarkar and Chakrabarti 2011). We decided to focus on extending Shah's variety metric, as it is one of the most widely used metrics in the field and also the one that is the most difficult to apply to the conceptual stage of human-centered design.

Towards this end, we collected a corpus of design concepts generated by teams of students in a graduate-level design course at U.C. Berkeley, entitled "Managing the New Product Development Process." We picked a subset of the teams which had produced a large number of concepts, and then selected a representative sample of concepts from these teams to be shown to design experts.

To probe how the notion of variety extends to earlier stages of design, we asked design experts to rate, between two randomly selected pairs of concepts within the same team, which pair represented greater variety in idea generation. To mitigate any order-dependent anchoring bias, the experts were initially shown all possible concepts together in a large set so that they knew the full range of concepts that had been generated. For a layout of the evaluation environment, see Fig. 2. An example of a high versus low variety comparison can be seen in Fig. 3. These ratings were done separately for each team, by multiple, independent design experts.

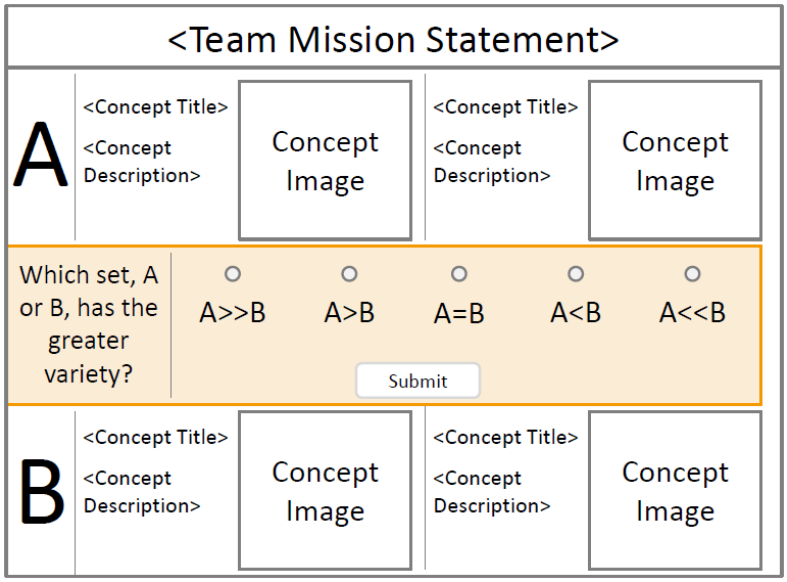

Fig. 2 Design experts were asked to look at two pairs of concepts and to rate which pair had greater variety.

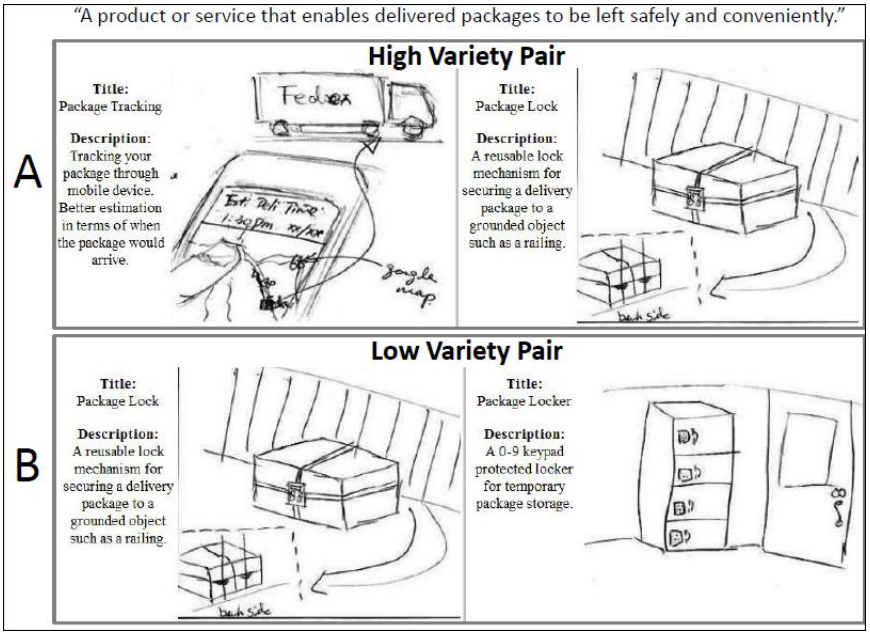

Fig. 3 Given the same mission, to enable "delivered packages to be left safely and conveniently," the concept pair in A was judged to have higher variety than the pair in B.

We hypothesized that if products addressed different user needs or utilized a different product/service ecosystem to provide for a given need, that that pair of concepts would be rated as having higher variety than just those concepts that only varied in their functional attributes, such as the metric originally proposed by Shah (2003). To test this hypothesis,

we extended Shah's metric by including additional latent factors for the product/service ecosystem and for the need addressed by the concept. Each concept was tagged with one of the following categories for product/service ecosystems: Product Only, Service Only, Product combined with Service, Government Program, or Community program. In both the ecosystem and need cases the tagging was performed by the research team, though we are currently developing a rater-independent rubric so that we can remove any researcher bias.

If we assume that differences in need and ecosystem affect variety, we can then compare this hypothesis to the 170 variety ratings we received from 10 design experts. Figure 4 shows the result, demonstrating that when concepts differed by either the need addressed or the product service eco-system used, design experts rated that pair of concepts as having greater variety when compared with a set of concepts that did not differ in either of those dimensions. We plan on conducting further tests once our evaluation rubric is complete, so as to remove any additional bias, and to explore potential interaction effects with a larger sample size.

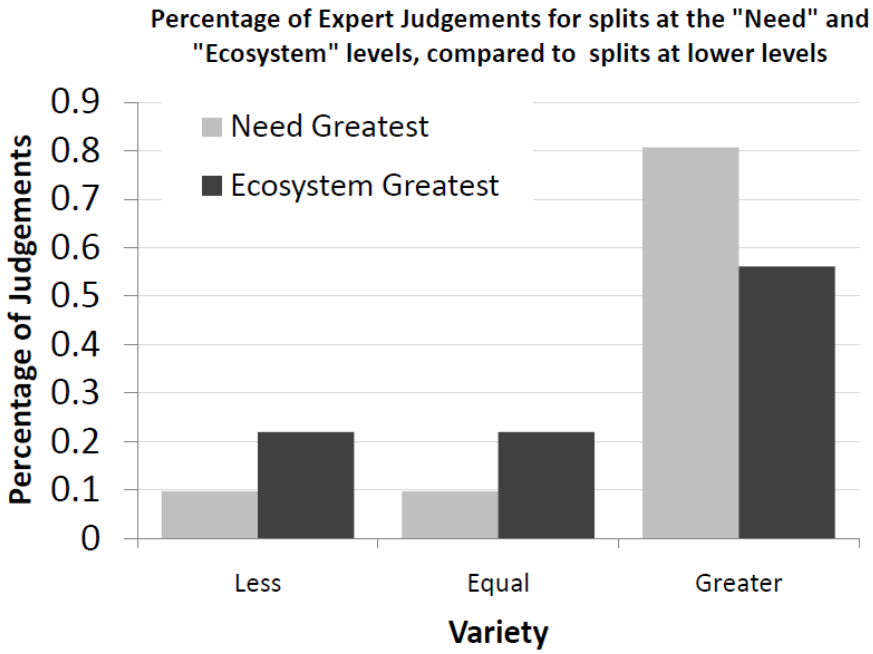

Fig. 4 Concepts that varied in the need addressed or product service ecosystem were judged to have higher variety than concepts that did not.

4. Conclusion

We hope to use this workshop as a venue to discuss how current creativity support tools and evaluation metrics can be used in earlier-stages of the design process. Using our initial work as a case-study, we wish to demonstrate how familiar techniques and methods can be re-applied in this new setting, as well as overview some of the challenges

that we have faced in doing so. We believe that this discussion will provide a rich area for new research, and will broaden the applicability of existing research.

References

Benami, O & Jin, Y (2002). `Creative Stimulation in Conceptual Design'. *ASME Conference Proceedings* **2002**(3624x):251-263.

Brown, D C (2011). `The Curse of Creativity'. In J. S. Gero (ed.), *Design Computing and Cognition '10*, chap. 9, pp. 157-170. Springer Netherlands, Dordrecht.

Hey, J et al. (2008). `Analogies and Metaphors in Creative Design'. *International Journal of Engineering Education* pp. 283-294.

Kurtoglu, T et al. (2009). `An experimental study on the effects of a computational design tool on concept generation'. *Design Studies* **30**(6):676-703.

Linsey, J S et al. (2008). `Increasing Innovation: Presentation and Evaluation of the Wordtree Design-by-Analogy Method'. *ASME Conference Proceedings* **2008**(43284):21-32.

Maher, M L (2010). `Evaluating creativity in humans, computers, and collectively intelligent systems'. In *Proceedings of the 1st DESIRE Network Conference on Creativity and Innovation in Design*, DESIRE '10, pp. 22-28, Lancaster, UK, UK. Desire Network.

Sarkar, P & Chakrabarti, A (2008). `The effect of representation of triggers on design outcomes'. *AI EDAM* **22**(02):101-116.

Sarkar, P & Chakrabarti, A (2011). `Assessing design creativity'. *Design Studies* .

Shah, J (2003). `Metrics for measuring ideation effectiveness'. *Design Studies* **24**(2):111-134.

Vattam, S S, et al. (2010). `A content account of creative analogies in biologically inspired design'. *AI EDAM* **24**(Special Issue 04):467-481.

Veale, T & Hao, Y (2008). `Comprehending and Generating Apt Metaphors: A Web-driven, Case-based Approach to Figurative Language' .

Veale, T & Hao, Y (2010). `Detecting Ironic Intent in Creative Comparisons'. In the *proceedings of ECAI'2010, the 19th European Conference on Artificial Intelligence*. Lisbon.