Measuring Creativity in Software Development

C. Nelson, B. Brummel, D.F. Grove, N. Jorgenson, S. Sen, R. Gamble
University of Tulsa
800 S. Tucker Drive, Tulsa OK, USA
{courtney-nelson, bradley-brummel, dean-grove,
noah-jorgenson, sandip, gamble@utulsa.edu}

Abstract. Creativity involves choosing to direct resources toward developing novel ideas. Information technology development, including software engineering, requires creative discourse among team members to design and implement a novel, competitive product that meets usability, performance, and functional requirements set by the customer. In this paper, we present results that correlate metrics of creative collaboration with successful software product development in a Senior Software Projects class that is a capstone course in accredited Computer Science programs. An idea management and reward system, called SEREBRO, provides measurement opportunities to develop metrics of fluency, flexibility, originality, elaboration, and overall creativity. These metrics incorporate multiple perspectives and sources of information into the measurement of creativity software design. The idea management portion of SEREBRO is a Web application that allows team members to initiate asynchronous, creative discourse through the use of threads. Participants are rewarded for brainstorming activities that start new threads for creative discourse and spinning new ideas from existing ones.

1 Introduction

Creativity can be understood and measured from a variety of perspectives. Two of the major approaches to measuring creativity are the psychometric approach and the confluence approach [1]. In this paper, we are concerned with creativity as a process not as an individual difference. We follow Sternberg and Lubart's [2] investment model of creativity that conceptualizes it as a decision that anyone can make, but few people do because of potential costs. We also incorporate Amabile's [3] componential model of creativity which requires three critical components for creativity: domain-relevant skills, creativity relevant processes, and intrinsic task motivation. We have implemented a software tool called *SEREBRO* (Software Engineering REwards for BRainstorming Online) to promote creative discourse in a Senior Software Projects class that is a capstone course in the Computer Science Curriculum at the University of Tulsa. SEREBRO organizes students around creativity relevant processes related to software engineering to motivate and assist teams of students with making creative software design and implementation decisions. Our primary interest as presented in this paper is examining whether encouraging the infusion of more creativity in the software development process results in more creative output with respect to the resulting software projects.

Essential cognitive elements for divergent thinking in the creative process include *fluency*, *flexibility*, *originality*, and *elaboration* [4, 5]. Fluency is the frequency of ideas generated. Flexibility involves seeing the problem in new ways or reconceptualizing the problem. Originality is devising novel ideas or synthesis of the problem. Elaboration is extending or fleshing out details. Creativity can be demonstrated in a specific part of a project or evaluated within the full project. Both objective and subjective metrics can be established to measure these elements in the creative process. The overall creativity encompassing the four cognitive elements can also be appraised subjectively. Our project uses both objective and subjective metrics of each element in the creative process to understand and foster creativity in collaborative software engineering projects.

Software engineering is the process of designing and developing a software product. It often involves a team-based approach that is oriented around specific tasks such as communication, planning, modeling, construction, and deployment [6]. The class undergoes training to acquire domain-relevant skills in each task to derive a set of work products, also called artifacts, which include customer requirements, documentation, designs, code, and packaging. Creativity relevant processes manifest themselves in a variety of forms in software development, such as.

- Interface design to achieve the proper functional use and aesthetic appeal
- Combining reusable entities in novel ways to meet innovative product requirements
- Implementation of functionality that is new to the developer's skill set
- Novel strategies for code refinement to increase performance, usability, and security
- Inventive ways to describe how to use a product in the final packaging delivered to the customer

We address the issue of fostering creativity as part of the collaborative software design in the context of a Senior Software Projects course. Since a team of students is involved, discussing what the novelty is and how it is placed in the product is as essential as the tangible design and implementation itself. We are most concerned with the creativity exhibited in the software development process during which a number of intermediate artifacts are produced, as well as the final software project and demonstration provided by the team at the end of the semester.

Fostering creativity in software engineering student deliberations poses unique challenges. First, it is uncommon for seniors in computer science to be forced to think creatively in their area of studies. Though liberal arts and science courses have creative and critical thinking components, in a computing curriculum, there is less attention paid to creativity and "thinking outside the box." Second, sharing and furthering ideas among team members in a collaborative fashion is very dependent on the openness of the team members, the time available to meet, and the different skill sets of the team members with respect to classifying requirements, analyzing a system design, and coding the various components (interface, processing, databases, etc.) that comprise the software. This problem has also been noted by information technology experts as well, leading to new software development processes such as agile development [7]. Third, creativity is often not directly rewarded. Even if ideas are explored, captured and used in idea management tools [8, 9], only the resulting

artifacts are examined by the instructor (i.e. management) and the customer. Therefore, since the creative discourse is not credited to the participants and is not explicit as an artifact itself, there is limited extrinsic motivation to engage in it.

To address these problems and provide a mechanism that supports, recognizes and fosters creativity, student teams use SEREBRO throughout the software development process. SEREBRO combines idea expression and management with a reward system designed to reinforce the usefulness of ideas and their contribution to creative discourse during the software development process [10]. Our objective is to explore the ability of SEREBRO to both capture and enhance creativity. The outcomes of our experiments can in turn be used to improve future information technology development, including introducing appropriate rewards.

2   Team Interaction on SEREBRO

Teams are an essential part of software development. Even when a single member is more creative or has an advanced skill set, the success of the project requires the contribution of all members, especially within a small team. Therefore, our initial experiments focus on evaluating creativity at a team level. While there may be variances in the contributions to the overall project by each member of the team, our analysis ignores that variance in favor of comparing team creativity with overall performance and outcomes.

Team interaction on SEREBRO starts with the designation of a list of topics. These topics are loosely based on the tasks to be performed and the artifacts required for hand-in at each milestone. The software engineering class used for testing SEREBRO had four major milestones over which the artifacts were produced, culminating in a final product presentation and demonstration. Discussions on SEREBRO commence within a specific topic when any team member posts a *brainstorm* node. Upon reviewing the node, another team member can agree or disagree with the post and then input their own idea (in the case of a disagree node) or an enhancement of the idea (in the case of an agree node). We call the agree/disagree process *spinning* an idea. Multiple brainstorms can start the discourse within a single topic. Figure 1 shows an *idea network* with seven brainstorm nodes in blue circles and agree nodes as green triangles. Disagree nodes, though not included in this figure, are upside down orange triangles. The box at the bottom shows the actual post that is seen when a node is 'moused' over.

Figure 1: SEREBRO Idea Network

Discourse on SEREBRO is exhibited by a *thread* of postings, much like an online forum. A short thread is shown in Figure 2 from an actual team when discussing a concept for a video game. Students who are the focus of this experiment are highly computer literate and very familiar with this style of forum posting. Asynchronous postings to SEREBRO, along with email alerts when posts are made, give students the freedom to work on the project from anywhere and at anytime. Though face-to-face meetings are encouraged, they are not essential to the progress of the project and mimic distributed teams. Threads can be *pruned* when they are no longer considered valid. Those threads that embody essential ideas regarding a specific topic, artifact, or direction are *finalized*. Finalization in SEREBRO involves naming the emerging concept as a new topic, generally for the next milestone, and tagging all of the contributing ideas, including brainstorm, agree, and disagree nodes.
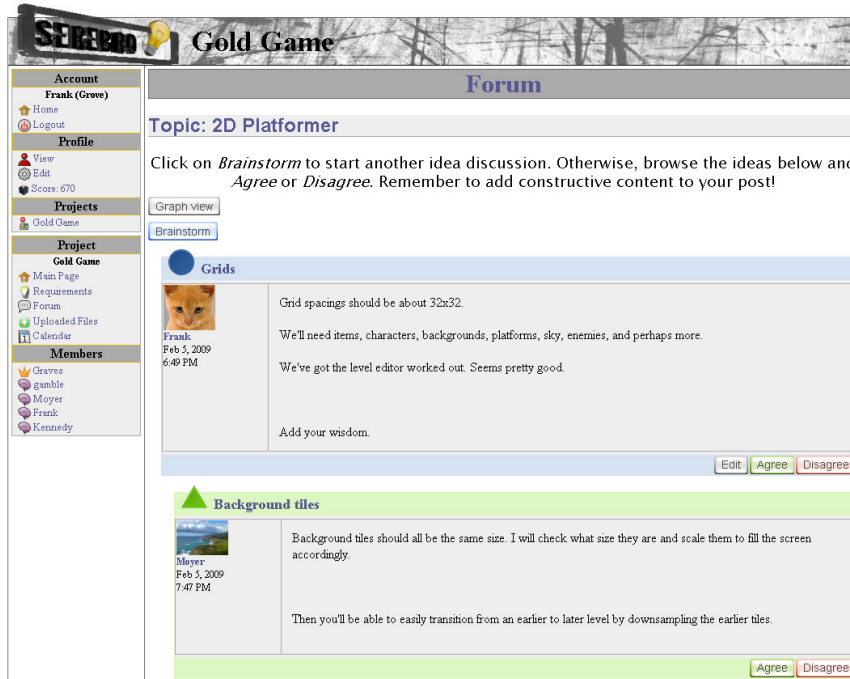
Figure 2: A Thread of Postings

To externally motivate creative contributions, SEREBRO implements a multi-agent system of rewards [10]. The reward system provides an egalitarian method of point distribution based on the structure of the idea network generated. Each user within the system is assigned a user-agent that distributes reward points to users based on their creative contribution following a specific protocol (see Section 3.1). In our current implementation, the agents reward according to the node type (*brainstorm*, *agree*, *disagree*, and *finalized*.) and its position in the network, but not the node content (contextual information) in performing these point allocations. When users post with a certain node type, they effectively classify the node with a positive or negative rating relative to its parent node. Based on these classifications, each user agent propagates reward points through each idea thread to produce cumulative rewards for each user. Idea nodes tagged during the finalization of a concept for the next milestone receive additional rewards. As points are incrementally allocated, the instructor and SEREBRO generate rewards for the top performing students. Since the nodes are classified by the participants depending on the length of the thread, we expect the resulting point allocations to correlate with expressed creativity. We hypothesized that such grounded and prudent external rewards that rely on peer feedback of creative contribution will motivate further creative expression from team members to attain a better team product. To ensure adequate participation within SEREBRO, students are encouraged to attain a certain level of reward points per milestone in the project.

The targeted class comprised six teams with a minimum of three and a maximum of four members. Each member had a role on the team, such as lead, analyst, and programmer, which gave him or her specific responsibilities. These roles were negotiated by the team members when the teams were formed in the first semester of the course [11]. Three non-trivial projects were defined and their requirements were passed to the teams by the product customers. Thus, two teams competed to create the best of each product to meet the requirements. The first project was a Chemistry Lab Creator and Submission software that allowed the instructor to create and post a lab assignment and students to complete the assignment and submit it online. The second was an Online Emergency Center that provided information on gas, food, and shelter during a disaster. The third project was a game to be used as part of a recruiting package for computer science students to choose a university. Initial requirements were sparse, forcing the team to work with the customer and delve into the domain of application. In addition to software artifacts produced, each team had to devise a novel product name and logo.

3   Methods of Creativity Measurement

We measure team creativity from multiple perspectives and sources of information about the software project as it is developed. We seek to determine the degree to which the discourse among the team is creative. This approach reveals unique information on the creativity of the overall software design process. We describe each of these sources and perspectives with respect to SEREBRO's contribution toward measuring creativity, thereby understanding certain differences in how creativity emerges in software design.

3.1   SEREBRO Threads

Team postings to SEREBRO are examined as consolidated threads related to a topic. Each thread stems from a single brainstorm node. A topic can have multiple evaluated threads, depending on the number of brainstorm nodes that start the various conversations within the topic. Therefore, threads form the basis for creativity analysis and are rated by experts at the end of the semester for the type and level of creativity.

Points. SEREBRO assigns reward points to each individual and team for their creative input based on the number and type of node of each thread posting. The nodes types are designed to align with the elements in the creative process. Fluency can be measured by the number of brainstorming nodes, flexibility by the fan-out of the brainstorming nodes, and originality and elaboration by the fan-in to spinning nodes. An in depth examination of the idea management tool can be found in [10]. Points are allocated through the agent based protocol, which has the following four basic rules. When a user replies with an Agree node to another post, the agent allocates k points to the author of the *parent* node. Similarly, when a user disagrees with its parent node, the parent node's author is charged k points. In order to reward

the progenitors of the thread tree, when an agent receives points at a node, it passes ($\frac{1}{2}$ * k) points to its parent node. This process effectively propagates the points throughout local areas of the idea network, yet discounts the reward as the distance increases from the node being considered. This rule also applies to the reward distributed at a Finalized node. Finalized nodes represent some accumulation of ideas that have been implemented within the project or identified as creative and useful to the project. Therefore, rewards at Finalized nodes are magnified by a factor m. This variable factor allows the instructor or project lead to adjust the importance of the Finalized node within the reward system. This $(k * m)$ points are also distributed to other nodes that are tagged with the same tag as the Finalized node. Thus, users directly correlate nodes that are related in content but not necessarily by distance in the network.

In combination, these rules distribute points throughout an idea thread, concentrating point totals on nodes that are well received by other users on the team. The user that authors positive nodes that are accepted by the team gains a higher point total than those who do not participate or do not create novel content. We hypothesize that the users with higher reward point totals have contributed more to the creativity of the group. For our experiments, $k = 1$ and $m = 2$. In our reported experiment, team point totals ranged from 1520 to 2210 with an average of 1857.

Expert Ratings. Each SEREBRO thread for each team was rated by five subject matter experts (SMEs) on the fluency, flexibility, originality, elaboration, overall creativity, and the relevancy of thread to the development of the project. The subject matter experts were faculty members and graduate students involved with the development of the SEREBRO program. Threads which were indicated as irrelevant by at least one of the raters were removed from further rating. These threads included reposts of the same content and tests of the system.

Though the sample size of the teams was small, there were a significant number of threads generated for rating. Overall, the SMEs showed sufficient levels of agreement across raters with internal consistencies (Cronbach's Alpha) on all ratings of each area of creativity ranging from 0.83 for Originality to 0.89 for fluency. Consistency ratings (ICC single measure scores) for each thread ranged from 0.49 on flexibility to 0.63 on fluency. This statistic indicates that the proportion of variance on each of the ratings across the five raters can be attributed to the variance on the construct being rated. The scores for each thread were averaged across the teams and the raters to calculate scores assigned to each team.

The second step in the analyses was to examine the correlations among different types of creativity expressed in the threads. The correlations between specific creativity metrics and overall creativity were quite high. These correlations indicated that threads were rarely rated highly on one metric of creativity and not on other levels of creativity. This result was expected and may have been influenced by the prevalence of threads with few entries which were typically rated low on all components of creativity. The highest correlation was between originality and overall creativity ($r = 0.95$). Essentially, the more original a thread was rated, the higher it was rated on creativity. This value was followed by examining elaboration ($r = 0.91$). Fluency and flexibility ($r = 0.88$) had the same correlation with rated creativity, and they predicted creativity to a lesser extent when compared to the other two elements. The lack of differentiation between types of creativity within threads indicated that

the raters typically judged threads to be creative or not rather than more fluent than flexible. Due to these results, further analyses relied only on the overall SME ratings of creativity. Team overall creativity ratings ranged from 1.99 to 3.09 on a five point scale from *Poor* to *Excellent*.

3.2 Final Projects and Demonstrations

At the final demonstrations of the team projects, each team was asked to evaluate their own team and the other teams on the overall creativity and overall quality of the projects. Additional survey questions addressed the perception of the creative process within each software team. Figure 3 indicates the different evaluations performed.

Team Colleague Ratings. Teams rated both the overall quality and the creativity of their own and other project teams based on their perceptions and understanding of the demonstrations of the final projects. The course instructor also performed the same ratings at the final demonstrations. Creativity ratings ranged from 3.18 to 4.84 and quality ratings ranged from 3.41 to 4.76. In general, teams with higher quality ratings received high creativity ratings as well. The highest rated quality team, however, was second on creativity.

Team grades. The course instructor graded teams at each milestone for a particular set of artifacts to provide feedback and assess progress. Final team grades ranged from 78 to 94 with a mean grade of 89.

Survey questions. Each team's own experiences were further investigated through a final survey asking about team performance and experiences with the SEREBRO program. This survey included questions about each of the creativity metrics with respect to the team's final project. The team members evaluated their own team on a 1 to 5 scale with anchors of *Poor, Neutral, Good, Very Good*, and *Excellent*. These ratings constitute a reference shift approach to team level variables in which individual scores about the team are combined to represent a team level construct [12].
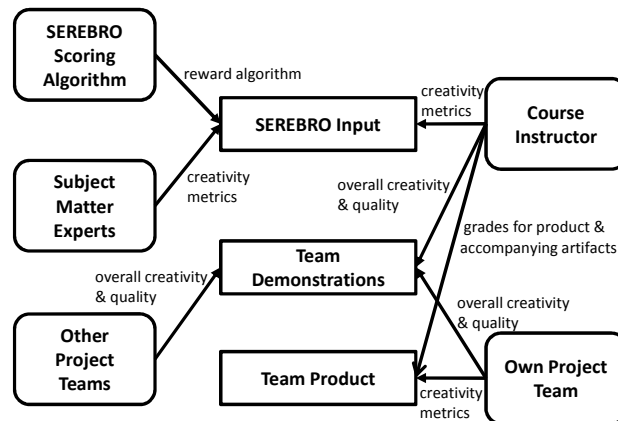

Figure 3: Perspectives and Sources of Information in the Creativity Metrics

4   Overall Results and Discussion

Table 1 displays the average scores and standard deviations of the various ratings. Table 2 displays the correlations between the creativity metrics

Table 1: Average Scores and Standard Deviations

	Serebro Points	Expert Ratings	Other Team Creativity	Other Team Quality	Own Team Creativity	Own Team Quality	Team Grades
Mean	1857	2.43	3.85	3.08	3.98	4.11	89.10
SD	270	0.48	0.41	1.09	0.60	0.52	5.52

The correlations among the various creativity metrics reveal both substantial agreement across the metrics and differentiation between project creativity and quality. Because there were only 6 teams involved in this project, none of the correlations at the team level are statistically significant at $p < 0.05$ level, but they do reflect a general level of agreement on the rank order of team creativity and quality across the various metrics. These correlations are still the best estimate of the relationships between these measures or team creativity and quality. Specifically, SEREBRO point totals were related to all other measures of both team creativity and quality. Expert ratings were able to distinguish between creativity and quality as only the correlations with other measures of creativity (including team grade) were positive and of medium size. Creativity demonstrated by the process metrics of SEREBRO points, Expert Ratings, and Team Creativity was also demonstrated in the final project as indicated by the other team's ratings of Creativity, as well as the Team Grades as shown in Table 2. Overall these results provide promising directions for using these metrics to continue to develop and measure creativity in computer science courses and software development teams.

Table 2: Correlations among team-level creativity metrics

		1.	2.	3.	4.	5.	6.
1.	SEREBRO Points						
2.	Expert Ratings	0.43					
3.	Presentation Creativity*	0.53	0.47				
4.	Presentation Quality*	0.60	-0.04	0.81			
5.	Team Grades	0.60	.043	0.53	0.33		
6.	Team Creativity	0.74	0.32	-0.02	0.25	0.42	
7.	Team Quality	0.64	-0.11	-0.03	0.54	0.86	0.62

*Includes results for only 5 teams. One team did not demonstrate their project on the assigned day.

5   Conclusion and Future Work

Ideas and novel software properties can take many forms in software engineering. Students who are not accustomed to conversing over creative material have difficulty

expressing their own ideas in a group setting, such as the one chosen for experimentation. SEREBRO is a software tool that includes both idea management and motivational rewards to foster creativity within a team developing a non-trivial software product. The results discussed in this paper indicate that SEREBRO performed well with respect to providing insight into creativity in a Senior Software Projects class, as well the correlation of product and team criteria with creativity metrics. The positive outcomes of the research are driving the expansion of SEREBRO with additional software process tools. We are examining different reward mechanisms to determine if certain protocols may be more suitable than others with respect to an individual's role on the team, his or her skill set, team size, and general attitudes about class and exposing their creative ideas. Further research into the creative process and team's perception of that process is currently taking place within a new set of students in the Software Projects Class setting. In addition, content based assessment of SEREBRO nodes is being undertaken during the software development process.

References

1. Sternberg, R.J., *The Nature of Creativity.* Creativity Research Journal, 2006. **18**(1): p. 87-98.
2. Sternberg, R.J. and T.I. Lubart, *Defying the Crowd: Cultivating Creativity in a Culture of Conformity.* 1995, New York: Free Press.
3. Amabile, T.M., *The Social Psychology of Creativity.* 1983, New York: Springer-Verlag.
4. Guilford, J.P., *Way beyond IQ: Guide to improving intelligence and creativity.* Buffalo, NY: Creative Education Foundation, 1977.
5. Guilford, J.P., *Varieties of Divergent Production.* Journal of Creative Behavior, 18,1--10, 1977.
6. Pressman, R.S., *Software Engineering: A Practitioner's Approach.* 7th ed. 2009: McGraw Hill Higher Education.
7. Ambler, S., *Agile Modeling: Effective Practices for eXtreme Programming and the Unified Process.* 1st ed. 2002: Wiley.
8. *Spigit.* http://www.spigit.com/index.html, 2009.
9. *BrainBank.* http://www.brainbankinc.com/Default.aspx, 2009.
10. R. Gamble, S.Sen, et. al., *The SEREBRO Project: Fostering Creativity through Collaboration and Rewards.* Presented at EDUMas Workshop, part of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009), 2009.
11. R. Gamble, M.Smith, *Moving toward "reality" in team selection for software engineering.* Frontiers in Education, 2008.
12. Chan, D., *Functional relations among constructs in the same content domain at different levels of analysis: A typology of compositional models.* Journal of Applied Psychology, 83(2): p. 234-246, 1998