

# Measuring Creativity: A Case Study Probing Rubric Effectiveness for Evaluation of Project-Based Learning Solutions

Renee M. Clary<sup>1</sup>, Robert F. Brzuszek<sup>2</sup>, C. Taze Fulford<sup>2</sup>

<sup>1</sup>Department of Geosciences, Mississippi State University, Starkville, USA;

<sup>2</sup>Cognitive Department of Landscape Architecture, Mississippi State University, Starkville, USA.

Email: rclary@geosci.msstate.edu

Received July 26<sup>th</sup>, 2011; revised August 28<sup>th</sup>, 2011; accepted September 6<sup>th</sup>, 2011.

This research investigation focused upon whether creativity in project outcomes can be consistently measured through assessment tools, such as rubrics. Our case study research involved student-development of landscape design solutions for the Tennessee Williams Visitors Center. Junior and senior level undergraduates (N = 40) in landscape architecture design classes were assigned into equitable groups (n = 11) by an educational psychologist. Groups were subsequently assigned into either a literary narrative or abstract treatment classroom. We investigated whether student groups who were guided in their project development with abstract treatments were more likely to produce creative abstract design solutions when compared to those student groups who were guided with literary narrative interpretations. Final design solutions were presented before an audience and a panel of jurors (n = 9), who determined the outstanding project solutions through the use of a rubric, custom-designed to assess the project outcomes. Although our assumption was that the measurement of the creativity of groups' designs would be consistent through the use of the rubric, we uncovered some discrepancies between rubric score sheets and jurors' top choices. We subjected jurors' score sheets and results to a thorough analysis, and four persistent themes emerged: 1) Most jurors did not fully understand the rubric's use, including the difference between dichotomous categories and scored topics; 2) Jurors were in agreement that 6 of the 11 projects scored were outstanding submissions; 3) Jurors who had directly worked with a classroom were more likely to score that class' groups higher; and 4) Most jurors, with the exception of two raters, scored the abstract treatment group projects as higher and more creative. We propose that while the rubric appeared to be effective in assessing creative solutions, a more thorough introduction to its use is warranted for jurors. More research is also needed as to whether prior interaction with student groups influences juror ratings.

*Keywords:* Creativity Assessments, Rubric, Creativity Measurement, Rubric Consistency, Problem Based Learning

## Introduction

A quick perusal of education resources attests to the importance of creativity: Researchers advocate creativity within art, literature, and even science classrooms (Yager, 2000; Taylor, Jones, & Broadwell, 2008; Csikszentmihalyi, 1996; Niaz, 1993). Incorporating creativity is not without challenges, however. Markham (2011) recently discussed some of the complexities and difficulties for teaching creativity. While most instructors recognize the importance of encouraging creativity in a classroom, they also realize that teaching creative processes involves significantly different techniques than content instruction.

However, as challenging as it may be to *foster* creativity in a classroom, *assessing* it can be even more difficult. How can an instructor anticipate creative solutions? Perhaps more problematic is whether instructors will *consistently* recognize and evaluate creative outcomes.

This case study project involved student-created landscape design solutions for the Visitor's Center in Columbus, Mississippi, which is the birthplace of noted playwright and Mississippi native, Tennessee Williams. In this quasi-experimental research design, an educational psychologist systematically grouped students to create equitable teams for a project based learning assignment. These teams were then assigned into one of two classrooms. While one classroom heard presentations on a literal narrative of Tennessee Williams' life and works, and was assisted by an outside designer recognized for his literal

interpretations, a theater professor discussed metaphorical meanings in Williams' work with the other class, and students were then assisted by an external designer known for his abstract solutions. Although we hypothesized differences between the groups' final projects, one research assumption was that the measurement of the creativity of groups' designs would be standardized and *consistent* through a customized rubric, an assessment tool developed specifically to evaluate this project's outcomes. This paper investigates whether the rubric was, in fact, an effective assessment tool, and whether the assumption of rater consistency was valid.

## Assessing Creativity

### Creative Problem Solving, Problem Based Learning, Divergent Thinking

The Creative Problem Solving (CPS) model identifies various stages of divergent and convergent actions, which in addition to describing the creative process, can also be used to facilitate it (Osborn, 1963; Parnes 1982; Isaksen & Treffinger, 1993). This combination of divergent and convergent processes is incorporated in the pedagogy of problem based learning (PBL), which provides students with complex real-life situations. Although originally implemented in the health sciences, PBL techniques are appropriated for general classroom use, and have researcher support for providing student opportunities for

divergent thinking and creativity (Delisle, 1997; Tan, 2008). Sternberg (2010) provided general instructor guidelines for promoting creative processes in a classroom—including project based learning and facilitation of student inquiry—and suggested an encouragement of idea generation, risk-taking, and tolerance of ambiguity.

There is general support for problem based learning and divergent thinking pedagogy for encouraging student creativity; assessment of creative products has paralleled an assessment of divergent thinking. Guilford's early research (1959, 1986, 1988) identified divergent thinking components which were quickly appropriated for creativity assessment. Fluency, flexibility, originality, and elaboration are Guilford categories commonly encountered for rating student creative performance. Likewise, Wiggins and McTighe's (2005) categories for metacognitive thinking—explanation, interpretation, application, perspective, empathy, and self-knowledge—have also been commandeered to assess creative products.

### Previous Research: Assessment Instruments for Creativity

Instruments designed to assist in evaluating creative products are intended to bring consistency to the process (Starko, 1995). However, the criteria for scoring creativity must be appropriate to the product being assessed. Rubrics have become common scoring guides for creative assessment, but taxonomy of creativity is necessary for effectiveness (Shepherd & Mullane, 2008). Johnson et al. (2000) warned that rating of performances required "considerable judgment" of the raters, and noted that reliability was often improved by using multiple assessors. In a meta-analysis of scoring rubrics, Jonsson and Svingby (2007) concluded that reliable scoring of performance assessment could be enhanced by rubrics—especially if accompanied by rater training. However, the researchers noted that simple use of rubrics did not necessarily facilitate valid judgment. Shores and Weseley (2007) discovered that educators' political views affected their perception of student performance, and concluded that a rubric was not an effective tool to prevent rater bias.

## Methods

### Tennessee Williams Project

In 2010, representatives from the town of Columbus, Mississippi, sought suggestions for landscape development in the space surrounding their Visitors Center (circa 1875), which also happens to be the birthplace of playwright Tennessee Williams. In this quasi-experimental research design, two landscape architecture design classes at a research university in the southern US were combined in a vertical studio project. We utilized Yin's (2008) case study research and analysis guidelines to organize and direct this project, and all protocols and procedures were approved by the university's Institutional Review Board prior to the start of the project.

The courses involved in this research are junior and senior level landscape architecture courses (N = 40, where Design I: n = 21; Design III: n = 19). Design I is a junior level course taught in Fall semesters that is open to landscape architecture students who have completed introductory coursework in design, computers, and graphics. Design III, also taught in Fall semesters, is the third design class in the curriculum sequence, and as a result, students enrolled in Design III have more experience than the incoming Design I students. (Design II, the

sequential course after Design I, is taught in Spring semesters and was not involved in this study.) Landscape architecture design courses utilize a project based learning system in which students investigate various instructor-chosen locations, and work either individually or within groups to produce a design solution.

### Use of TypeFocus™ for Assignment of Student Design Groups

Prior to the case study assignment, students in both courses were directed to access and complete a TypeFocus™ online survey. TypeFocus™ is available at the university as a personal assessment tool to assist students in identifying their interests for career planning. However, TypeFocus™ also measures individuals' potential creativity and divergent thinking, and we utilized this tool to ensure that the potential creative students were equitably distributed among groups. Previous research (Nassar & Johnson, 1990) suggested that landscape architects are more commonly intuitive (N), thinking (T), and judging (J), although there was quite a bit of variability among the sample. Regardless of our students' characteristics, we wanted to ensure that one group did not have an inherent advantage over another because of student characteristics.

An educational psychologist used the TypeFocus™ data, as well as the class standing (junior or senior level) to assemble project groups. In addition to the creativity and divergent thinking assessment, TypeFocus™ data also provide indications of students' perception of time and attendance to structure. Six students elected not to participate in the project, and were assigned to two non-research groups. The remaining students (n = 34) were grouped into nine treatment groups with 3 - 4 members. Each group had the benefit of a senior student, a student with higher divergent thinking scores (and potential creativity), and a student who had an awareness of time and deadlines. Once students were assigned to groups, we randomly assigned groups to either a literal (Design I) or abstract (Design III) treatment classroom. However, no one other than the researchers was aware of the research purpose. Additionally, neither instructor of Design I or Design III offered critiques of student projects in order to minimize potential influence. Instead, assistant instructors and outside landscape architects—who were unaware of the research design—guided the students' project development.

### Tennessee Williams Project Introduction

All students completed a pre-test prior to the project assignment. Questions probed knowledge of Tennessee Williams' life and career, and basic knowledge of Williams' famous play, *The Glass Menagerie*. Students were given their group assignment, their classroom assignment, and then handed the project statement. The project statement directed students to design a public space surrounding the Visitors Center that reflected Tennessee Williams' life and work, while integrating the project into the overall site.

We assigned Tennessee Williams' play, *The Glass Menagerie*, to all groups to illustrate Williams' use of storytelling and provide a flavor of Williams' work. All students visited the case study site in Columbus, Mississippi, where the classes were then divided for presentations. In the literal treatment class, students heard highlights and milestones of Tennessee Williams' life in a lecture presentation from Williams' historians. Meanwhile, groups assigned to the abstract treatment classroom heard a presentation from a theater professor on

metaphorical elements in Williams' plays.

Later at the university, students were led in the design charrette process by guest landscape architects. The literal classroom's landscape architect was encouraged to focus on project and design elements, while the abstract classroom's landscape architect was requested to discuss concept and metaphorical elements.

### Evaluating Creativity: The Rubric Design

The university design instructors researched creative measures, and discussed what factors needed to be assessed to determine abstraction and creativity in the group projects. (The group projects were scored via a separate set of criteria for students' recorded grades. Therefore, participation in either the literal or abstract treatment class did not affect recorded student performance.) After discussion and compromise, characteristics of *explanation* (naïve, developed, or sophisticated, after Wiggins & McTighe, 2005), *interpretation*, *elaboration*, and *originality* were chosen as rubric categories. In order to ascertain whether groups effectively used narrative or storytelling as a guiding theme rather than metaphorical design, the rubric included both storytelling elements and abstract ideas for both Tennessee Williams' life and works. Patti Carr Black, former director of the Old Capitol Museum in Jackson, Mississippi, noted that Mississippi artists may have experimented with abstraction, but they were more comfortable with "representational art" (Black, 2007). However, we hypothesized that students exposed to abstraction in their design process might be more apt to produce abstract designs.

The Design I and Design III instructors worked with the educational psychologist and an educational researcher to develop a rubric by which an audience would score the project designs from each group. The individual juror's packet included the problem statement ("Groups were to consider the development of a small parcel of land adjacent to the Tennessee Williams home on Main Street for a park. Groups were to design a detailed public space that reflects Tennessee Williams' life and work"). A rubric and separate score sheets for each project ( $n = 11$ ; non-research groups were also scored although data were not used) were also provided (Appendix A).

### Project Culmination: Juried Group Presentations

After an intensive two week project, student groups presented their design solutions to a public audience. Serving as jurors were designers and literary scholars, including guest landscape architects, the theater professor, a literary scholar of Tennessee Williams, an architect, a floral design professor, the educational psychologist, and representatives of the local community, including a newspaper publisher and a representative from the tourist bureau (Figure 1). Before group presentations began, the educational psychologist met briefly with the jurors. Each juror was given a rating packet with 11 score sheets and the rubric (Appendix A). The psychologist overviewed the rubric, and discussed how the score sheets were to be used. Additionally, each juror was asked to provide his/her top three design choices, by group, at the end of the presentations.

Groups showcased their solution designs on project boards, and overviewed their projects in brief presentations to the audience (Figure 2). Following the 11 group presentations, jurors were given the opportunity to revisit the project boards, ask questions, and discuss the designs in more detail with group members.



Figure 1.  
Groups showcased their project designs in the form of project boards. Nine jurors reviewed and scored group projects.



Figure 2.  
Each group summarized their project design for the audience.

When finished, jurors turned in their rubric packet, and the psychologist tallied the votes. The top three groups were announced, and winning groups' members were given small prizes. The design instructors then announced the purpose behind the research project.

## Results

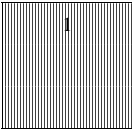
Our original research purpose for the Tennessee Williams' project was to determine whether student groups, who were presented with metaphorical and abstract presentations and project guidance, were more likely to produce abstract design solutions than groups who were guided through literary presentations and assistance. Our analysis of the abstract design solutions is published, and our results indicated that students who were exposed to abstract teaching methodologies had a greater tendency to produce abstract solutions, and that representational art was not necessarily the default position in the southern US (Fulford et al., in press). However, in the analysis of the groups' project solutions, we noticed that the rubric scores did not always coincide with some jurors' choices of the top three group designs.

We subjected the nine jurors' rater packets to a mixed methodology analysis, and examined each juror's individual score assignments for the rubric components: explanation (E); interpretation (I); storytelling (ST); abstraction (A); elaboration (elab); and originality (O) (Table 1). We next tallied each juror's rating sheet, and noted the top three projects *according to the scores*. Next, we compared each juror's *identified* top three projects with the top three rubric-scored projects (Table 2). Finally, we implemented Neuendorf's (2002) guidelines for

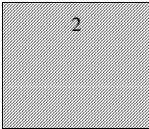
Table 1.

Summary of jurors' scores for group projects. Abstract group treatments are represented in pink, while yellow groups were exposed to literary treatment. Judges 5 and 8 were involved with the abstract groups' design process, while judge 6 was involved with the literary groups. The first, second, and third place choices of each judge are noted by grid designs within the table.

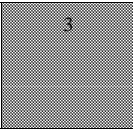
JUROR	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9
1	e = 3, i = 3 ST = 2/3, A = 2 elab = 3 orig = 3	e = 3, i = 1 ST = 3, A = 3 elab = 3 orig = 2	e = 3, i = 3 ST = 3, i = 3 elab = 3 orig = 3	e = 2, i = 2 ST = 2, A = 2 elab = 2 orig = 2	e = 3, i = 3 ST = 3, A = 3 elab = 3 orig = 3	e = 2, i = 2 ST = 2, A = 2 elab = 2 orig = 2	e = 3, i = 3 ST/A n/r elab = 2 Orig = 2	e = 3, i = 3 ST = 2/3, A = 3 elab = 3 orig n/r	e = 3, i = 3 ST = 3, A = 3 elab = 3 orig = 3
2	e = n/r, i = 1 ST = 1, A = 1 elab = 1 orig = 1	e n/r, i = n/r st = 2, A = 2 elab = 2 orig = 2/3	3 = 2/3, i = 3 ST = 3, A = 2/3 elab = 3 orig = 3	e = 1, i = 1 ST = 1, A = 1 elab = 1 orig n/r	e = 3, i = 3 ST = 3, A = 3 elab = 2 orig = 3	n/r n/r n/r	e = 2, i = 1 ST = 1, A = 1 elab = 1 orig = 1/2	e = 3, i = 3 ST = 2/3, A = 2 elab = 3 orig n/r	e = 2, i = 2 ST = 1, A = 1 elab = 2 orig = 1
3	e = yes, i = 3 ST = 2/3, A = 2 elab = 1 orig = 3	e = 2, i = 2 ST 3, An/r elab = 2 orig 2	e = 3, i = 2, ST = 2, A = 3 elab = 2 orig n/r	e = 1, i = 1 ST = 1, A = 1 n/r orig n/r	e = 2, i = 2 ST = 2, A = 2 elab = 2 orig n/r	e = 1, i = 1 St = 1, A = 1/2 n/r orig n/r	e = 3, i = 3 ST = 2, A = 2 n/r orig n/r	e = 3, i = 3 ST = 3, A = 3 elab = 3 orig n/r	e = 2, i = 2 ST = 2, A = 2 n/r orig n/r
4	e = 2, i = 2 y, y elab = 1 orig = 2	e = 2, i = 2 ST = 2, A = 2 elab = 2 orig = 2	e = 2, i = 2 ST = 2, A = 3 elab = 2 orig = 3	e = 2, i = 2 ST = 2, A = 3 elab = 2 orig n/r	e = 2, i = 2 ST = 2, A = 2/3 elab = 2 orig n/r	e = 2, i = 3 St = 2, A = 2 elab = 2 orig n/r	e = 2, i = 1 ST = 1, A = 1 elab = 2 orig = 2	e = 3, i = 3 ST = 2, A = 2/3 elab = 3 orig = 3	e = 2, i = 2 n/r n/r n/r
5	i = 2, i = 2 ST = 2, A = 1 elab = 2 e = 1, i = 2	e = 1, i = 1 ST = 1, A = 1 elab = 1 Orig = 1	e = 3, i = 3 ST = 2, A = 2 elab = 3 Orig = 3	e = 2, i = 2 ST = 2, A = 2 elab = 2 Orig = 2	e = 2, i = 3 ST = 2, A = 1 elab = 2 Orig = 2	e = 1, i = 1 ST = 2, A = 3 elab = 3 orig = 3	e = 2, i = 2 St = 2, A = 2 elab = 2 orig = 2	e = 2, i = 2 St = 2, A = 3 elab = 3 orig = 3	e = 3, i = 3 ST = 2, A = 3 elab = 2 orig = 3
6	e = 1, i = 2 ST = 1, A = 2 elab = 1 orig 2 = 3	e = 2, i = 2 St = 2, A = 2 elab = 2 orig = 2	e = 3, i = 3 ST = 3, A = 3 elab = 3 orig = 3	e = 2, i = 2 ST = 3, A = 3 elab = 3 orig = 2	e = 2, i = 3 ST = 2/3, A = 3 elab = 2 orig = 2	e = 1, i = 2 ST = 2/3, A = 3 elab = 3 orig = 3	e = 2, i = 3 St = 1/2, A = 3 elab = 3 orig = 3	e = 3, i = 3 ST = 2/3, A = 3 elab = 3 orig = 3	e = 2, i = 3 ST = 1, A = 2 elab = 3 orig = 2
7	e = 1, i = 1 ST = 1, A = 1 elab = 1 orig = 1	e = 3, i = 1 ST = 2/1, A = 1 elab = 1 orig = 1	3 = 3, i = 3 ST = 2/3, A = 2/3 elab = 3 orig = 3	e = 1, i = 1 ST = 1, A = 1 elab = 1 orig = 1	n/r ST = 2, A = 2 elab = 2 orig = 2	nr n/r n/r	e = 2, i = 2 st = 1, A = 1 elab = 2 orig = 2	e = 3, i = 3 ST = 3, A = 2/3 elab = 3 orig = 3	e = 2, i = 2 St = 1/2, A = 1/2 elab = 2 orig = 2
8	e = 2, i = 2 ST = 2, A = 1 elab = 3 orig 2	e = 2, i = 3 ST = nr, A = 3 elab = 3 orig = 2	e = 2, i = 3 ST = 2, A = 2 elab = 3 orig 3	e = 2, i = 2 ST = 2, A = 2 elab = 2 orig = 2	e = 3, i = 3 ST = 2/3, A = 3 elab = 3 orig = 2	e = 2, i = 3 St = 2, A = 3 elab = 2 orig = 3	e = 3, i = 2 ST = 2, A = 2 elab = 3 orig = 3	e = 3, i = 3 ST = 2, A = 3 elab = 3 orig = 3	e = 2, i = 2 ST = 2, A = 1/2 elab = 2 orig = 2
9	e = 1, i = 1 ST = 1, A = 1 elab = 2 orig = 1	e = 2, i = 3 St = 1/2, A = 1/3 elab = 2 orig = 2/3	e = 3, i = 3 ST = 3, a = 3 elab = 3 orig = 3	e = 2, i = 2 St = 1/2, A = 2 elab = 2 orig = 2	e = 3, i = 3 ST = 2, A = 3 elab = 2 orig = 2	e = 3, i = 3 ST = 3, A = 2/3 elab = 2 orig = 1	e = 2, i = 3 ST = 1/2, A = 1 e:lab = 1 orig = 2	e, I n/r n/r n/r n/r	n/r n/r n/r n/r



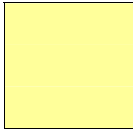
1  
First Place



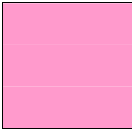
2  
Second Place



3  
Third Place



Literal

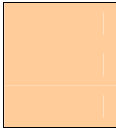


Abstract

**e = Elaboration; I = Interpretation, ST = storytelling; A = abstract; Elab = Elaboration; Orig = Originality**

Table 2.

Analysis of jurors' scores for creative elements, and selection of top three awards. The peach color represents the jurors' top three project choices. The Summary column notes whether top choices are supported by rubric data (YES) or not (X).

Juror	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8	Group 9	SUMMARY
1	iiii ST>	ii ST = 3	iiii ST = 3	ST = 2	iiii ST = 3	ST = 2	ii n/r	iii n/r ST < A	iiii ST = A	XX
2	ST = 1	ST = 2	iiii ST > A	ST = 1	iii ST = 3	n/t	ST = 1	iii n/r ST >	ST = 1	X
3	ii ST >	n/r	i ST < A	ST = 1	ST = 2	ii ST <	ST = 2	iiii ST = 3	ST = 2	XX
4	n/r	ST = 2	i ST < A	ST <	ST <	i ST = 2	ST = 1	iiii ST < A	n/r	YES
5	ST >	ST = 1	iiii ST = 2	ST = 2	i ST >	ii ST <	ST = 2	ii ST < A	iii ST < A	YES, not 1st
6	ST <	ST = 2	iiii ST = 3	i ST = 3	i ST <	ii ST <	iii ST <	iiii ST < A	ii ST < A	YES
7	ST = 1	i ST >	iiii ST = 2/3	ST = 1	ST = 2	n/r	ST = 1	iiii ST >	iiii ST + 1/2	YES, not 3rd
8	i ST >	ii n/r	iii ST = 2	ST = 2	iii ST <	ii ST <	iii ST = 2	iiii ST < A	ST >	Yes, not 3/1
9	ST = 1	i ST <	iiii ST = 3	ST <	ii ST <	ii ST >	i ST >	n/r	n/r	YES, not 2nd
 <p><b>i = Creative element</b> <b>X = place not justified</b></p> <p><i>Juror Choice</i></p>										

content analysis of jurors' comments within the rater's packets.

Three of the jurors had also been directly involved in the classroom prior to their assessment of group projects: The external landscape architects each directed a classroom (abstract or literal) charette, and the theater professor had led the discussion and presentation on metaphorical elements in Tennessee Williams' life and works. (The literature professor who led the literary group presentation on the milestones in Williams' life had a conflicting engagement and was unable to attend the juried presentation. We selected another literary scholar to replace him.) Therefore, we also conducted a detailed analysis to see whether professors and instructors with previous group involvement had a tendency to rate their groups higher.

**Rubric and Score Sheet Analysis**

One of the first observations we made with the rater score sheets was that they were often incomplete. Only 2 of the 9 jurors turned in completed score sheets for all groups; interestingly, these jurors were the guest landscape architects who had worked with the student groups prior to the juried presentation. Many jurors did not fully rate certain groups' projects, and three jurors turned in empty score sheets for some groups. Jurors were inconsistent in their scoring of individual elements as

well. While some projects were scored 1 - 3 on abstract and storytelling elements, other projects were scored by the same juror as "yes/no" for these elements. In fact, two jurors defaulted to yes/no responses in categories which required a 1 - 3 rating.

The two elements which were meant to distinguish the treatment groups, storytelling (ST) and abstract (A) components, did not discriminate between projects as we anticipated. Most jurors rated these numbers as equivalent in the same project. When there was a difference between them, it was not a consistent measure.

**Congruence of Rubric and Juror Selections**

When we compared each juror's identified group winners against his/her score sheets, we saw that not all jurors' project rubric scores *matched* their top three choices (Tables 1 and 2). Only two jurors, a landscape architect and a professor of floral design, had rubric score sheets that justified their first, second, and third design choices. Four of the jurors partially justified their choices through rubric score sheets: one juror's first place choice was not supported by rubric scores, one juror's second place choice was not supported, one juror's third place choice was not supported, and one juror's first and third choices were

not supported in their placement (the first and third scored projects were switched in the juror's preference). Of the three remaining jurors, none of their top three project choices was supported by rubric score sheets. One juror's third place selection corresponded to a blank rubric score sheet.

When we investigated jurors' assignments for groups' elaboration, interpretation, and originality—those elements that indicate divergent thinking and creativity (Guildford 1959, 1986, 1988; Wiggins & McTighe 2005)—we found that three jurors who scored group projects' high in these categories were in complete agreement with their identification of the overall projects as exceptional (Table 2). Three jurors' scores on these elements were in primary agreement for the projects they scored as exceptional, but three jurors' scores were in disagreement with their identifications of exceptional projects. Therefore, the inclusion of these rubric elements for measuring creativity and divergent thinking appears to be a discriminating one. Although we did not observe complete agreement among jurors within scoring and/or interpreting these elements, the trend appears to confirm the usefulness of the rubric for scoring creative project solutions. Undoubtedly, the reliability of the rubric was increased by the use of multiple jurors (Johnson et al., 2000).

### Potential Impact of Juror Direct Involvement

One third of the jurors was previously involved with the classroom groups: the landscape architects were each involved with a classroom (literal and abstract), and the theater professor was involved with the abstract treatment classroom (Table 1). For the landscape architect involved with the abstract classroom, two of his top three group choices emerged from the classroom he was assisting: Both his first and third place group choices were participants in the abstract classroom. For the landscape architect involved with the literary classroom, two of his top three choices also emerged from the classroom he assisted. (His first and third place choices were literary treatment groups.) The theater professor's top three choices *all* came from within the abstract treatment classroom. Although our population is small, our case study research hints that perhaps jurors who are involved with treatment groups tend to score these groups higher. However, it also appears that jurors recognized a good project solution, regardless of what their design emphasis might have been.

## Discussion and Implications

When we combined our rubric analysis with content analysis of jurors' comments, four persistent themes emerged: 1) Most jurors did not fully understand the rubric's use, including the difference between dichotomous categories and scored topics; 2) Jurors were in agreement that 6 of the 11 projects scored were outstanding submissions; 3) Jurors who had directly worked with a classroom were more likely to score that class' groups higher; and 4) Most jurors, with the exception of two raters, scored the abstract treatment group projects as higher and more creative.

### Rubric Effectiveness

The design class instructors and the educational researcher worked closely with the educational psychologist to design the rubric that would effectively measure the creativity and divergent elements of the submitted group projects. Although there was not perfect agreement among jurors' scores, the creative

and divergent elements that were measured in the rubric aligned completely (33%) and primarily (33%) with the majority of jurors' selected top projects. Only with one-third of jurors did the rubric fail to measure the projects with highest creativity (that it was designed to do). Therefore, given that rubrics typically do not produce consistently aligned scores among all raters (Shores & Weseley, 2007; Johnson et al., 2000), our findings indicate that the rubric designed for this project was still an effective measuring device.

### Consistent Use of Rubric as a Creativity Assessment Tool

Our analysis also indicates that the majority of jurors did not fully understand the rubric, or were inconsistent in their use of it as an assessment tool. Although we trained jurors with the rubric prior to its use, our efforts did not appear to sufficient for consistency among all jurors. Time does not appear to be a factor in our case study, as jurors were provided time after the presentations to meet with individual groups and clarify their understanding (and rating) of a specific project. It is also puzzling that some jurors completely abandoned their rubric score sheets when deciding their top three projects. There appear to be additional criteria for choosing these projects that were not made evident in the rubric, or by jurors' additional comments on score sheets.

### Implications for Assessing Creative Outcomes and Rubrics

We think that our results indicate that a more intensive introduction of the rubric is warranted before its use as an assessment tool. It might be an effective use of time to expose jurors to a test design example, and then have jurors score the project in a "trial run". This may help clarify the intent of the categories, and whether or not rubric elements require a scaled (1 - 3) ranking, or a dichotomous response. Although our case study population is small, the tendency of jurors to rate higher the groups with whom they have previously interacted may indicate potential bias (Shores & Weseley, 2007). Conversely, these tentative results may support the choices of the researchers: The jurors may not have been scoring their student *groups* higher as much as they were scoring a design position that was congruent with their professional *worldview*. More research is needed to determine whether previous exposure to student groups significantly influences project scores.

In this case study, the rubric appears to be an appropriate tool for scoring creative projects, although it was not completely utilized as it was intended. The storytelling and abstract element categories, designed to separate abstract and narrative products, had little effect on jurors' scores. However, the use of the rubric, coupled with multiple assessors, resulted in the fairly consistent identification of superior design solutions. The majority of jurors' scores on creative elements were within perfect or primary agreement with their exceptional project identification, and jurors effectively identified the same six projects as outstanding. Moreover, 7 of the 9 jurors scored the abstract group's projects as higher in creativity; our previous analysis (Fulford et al., in press) suggested that these projects did, in fact, have a greater concentration of abstract and metaphorical elements than the group projects that emerged from the literal classroom. This indicates that the rubric overall did what it was designed to do, and helped in the identification of creative elements within the projects.

## Acknowledgements

The researchers wish to thank Dr. Donna Gainer for analyzing students' TypeFocus™ results and assigning equitable groups for this project. We also appreciate Dr. Gainer's expertise and efforts in working with us to customize the rubric for assessment of creative solutions. We also thank Mr. Gordon Lackey for assisting with this project, which included the transcription of the literal and abstract presentations to the student groups.

## References

- Black, P. C. (2007). *The mississippi story*. MS: University of Mississippi Press.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. New York: Harper Collins.
- Delisle, R. (1997). *Use problem-based learning in the classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Fulford, C. T., Brzuszek, R., Clary, R., Gainer, D. & Lackey, B. (in press). Influencing the story: Literal versus abstract instruction. *Proceedings of the Council of Educators in Landscape Architecture (CELA) 2011 Urban Nature*.
- Guildford, J. P. (1959). Three faces of intellect. *American Psychologist*, 14, 469-479. doi:10.1037/h0046827
- Guildford, J. P. (1986). *Creative talents: Their nature, use and development*. Buffalo, NY: Bearly Ltd.
- Guildford, J. P. (1988). Some changes in the structure-of-intellect model. *Educational and Psychological Measurement*, 48, 1-6. doi:10.1177/001316448804800102
- Isaksen, S. G., & Treffinger, D. J. (1985). *Creative problem solving: The basic course*. Buffalo, NY: Bearly, Ltd.
- Johnson, R. L., Penny, J., & Gordon, B. (2000). The relation between score resolution methods and interrater reliability: An empirical study of an analytic scoring rubric. *Applied Measurement in Education*, 13, 121-138. doi:10.1207/S15324818AME1302\_1
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity, and educational consequences. *Educational Research Review*, 2, 130-144. doi:10.1016/j.edurev.2007.05.002
- Markham, T. (2011). Can we really teach creativity? ASCD Edge. Association for Supervision and Curriculum Development. [http://edge.ascd.org/\\_Can-we-really-teach-creativity/blog/3577828/127586.html](http://edge.ascd.org/_Can-we-really-teach-creativity/blog/3577828/127586.html)
- Nasar, J. L., & Johnson, E. (1990). The personality of the profession. *Landscape Journal*, 9, 102-108.
- Niaz, M. (1993). Research and teaching: Problem solving in science. *Journal of College Science Teaching*, 23, 18-23.
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.
- Osborn, A. F. (1963). *Applied Imagination* (3rd ed.). New York: Scribner's.
- Parnes, S. J. (1981). *Magic of your mind*. Buffalo, NY: Bearly Ltd.
- Shepherd, C. M., & Mullane, A. M. (2008). Rubrics: The key to fairness in performance based assessments. *Journal of College Teaching and Learning*, 5, 27-32.
- Shores, M., & Weseley, A. J. (2007). When the A is for agreement: Factors that affect educators' evaluations of student essays. *Action in Teacher Education*, 29, 4-11.
- Starko, A. J. (1995). *Developing creativity in the classroom: Schools of curious delight*. White Plains, NY: Longman Publishers.
- Sternberg, R. J. (2010). Teach creativity, not memorization. *The Chronicle of Higher Education*. <http://chronicle.com/article/Teach-Creativity-Not/124879/>
- Tan, O.-S. (2008). *Problem-based learning and creativity*. Singapore: Cengage Learning Asia.
- Taylor, A. R., Jones, M., & Broadwell, B. (2008). Creativity, inquiry, or accountability? Scientists' and teachers' perceptions of science education. *Science Education*, 92, 1056-1075.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Yager, R. E. (2000). A vision for what science education should be like for the first 25 years of a new millennium. *School Science and Mathematics*, 100, 327-341. doi:10.1111/j.1949-8594.2000.tb17327.x
- Yin, R. K. (2008). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage Publications.

### Appendix A. Rubric and Score Sheet for the Tennessee Williams Park

	Level one	Level two	Level three
<b>Explanation</b> (Wiggins & McTighe)	Naïve: a superficial account, more implicit than analytical or explanatory, sketchy account of experience; less a theory than an unexamined hunch or borrowed ideas.	Developed: an account that reflects some in-depth and personalized reflection; making a thinking process that is their own; going beyond the given.	Sophisticated: an unusually thorough, explanatory, and inventive account; fully supported, verified, and justified; deep and broad.
<b>Design Concept</b>	Guiding idea is basically explained	Guiding idea is well written and conceived with some allusion to form	Guiding idea is rich and novel, compelling statement that leads to strong forms.
<b>Interpretation</b> (hybrid) <b>Storytelling</b>  Or <b>Abstract</b>	Simplistic or superficial; no interpretation  A decoding with no interpretation; no sense of wider significance	A plausible storyline with clear details  A helpful interpretation of analysis of the significance or meaning of cognitive strategies	A well structured storyline with rich details and imagery; provides a detailed history  A powerful and illuminating interpretation and analysis; tells a rich and insightful account of cognition through reflection; sees deeply
<b>Elaboration</b>	Some details or ideas	Expanded details or ideas	Rich imagery and elaborate details
<b>Forms/Structures</b>	Forms have basic expression for selection; little expansion	Forms chosen for design are well selected to reinforce original concept	Forms are complex or novel and excellently reflect the guiding idea
<b>Originality/Novelty</b>	Commonplace ideas and expected usage	Unusual ideas and elements	Sophisticated: an unusually complex and rich approach, far outside the ordinary

Rating Scale Score Sheet

Name \_\_\_\_\_

Group Number	Adq 1	Good 2	Superior 3	Yes√	No √
<b>Explanation</b>					
<b>Interpretation</b>					
<b>TN Williams Life details included</b>					
<b>Story telling elements</b>					
<b>Abstract ideas</b>					
<b>TN Williams Works details included</b>					
<b>Story telling elements</b>					
<b>Abstract ideas</b>					
<b>Elaboration</b>					
<b>Site Details:</b>					
Use of site is appropriate					
Vegetation and Faunal Elements					
Infrastructure and maintenance					
Access and circulation issues					
<b>Forms/Structures</b>					
Aesthetic					
Appropriate					
Express main idea					
Kinetic elements					
<b>Originality/Novelty of Design</b>					
<b>Total</b>					
<b>Comments</b>					
<b>Non-shaded items rate 1, 2, 3 using rubric</b>					
<b>Shaded items check Yes or No</b>					