Computer-based Performance Assessment of Creativity Skills: A Pilot Study

Research Report

Yigal Rosen, Ph.D.
Maryam Tager

May 2013
About Pearson
Pearson, the global leader in education and education technology, provides innovative print and digital education materials for pre-K through college, student information systems and learning management systems, teacher licensure testing, teacher professional development, career certification programs, and testing and assessment products that set the standard for the industry. Pearson’s other primary businesses include the Financial Times Group and the Penguin Group. For more information about the Assessment & Information group of Pearson, visit http://www.pearsonassessments.com/.

About Pearson’s Research Reports
Pearson’s research report series provides preliminary dissemination of reports and articles prepared by Test, Measurement, and Research Services (TMRS) staff, usually prior to formal publication. Pearson’s publications in PDF format may be obtained at http://www.pearsonassessments.com/research.
Abstract

Creativity is an important skill set in college and career readiness. Many educational programs in school settings place great emphasis on promoting student ability to create original ideas, products, and solutions that will be both novel and valuable. However, there is lack research on potential tools that can be used for creativity assessment in the K-12 arena. The aim of this study was to explore patterns in student performance and motivation in creativity computer-based prototype assessment. Eighty seven 14-year-old students from the United States, United Kingdom, Turkey, and South Africa participated in the study. Students were asked to write different short stories based on given images or a video, create titles for a story, and propose captions for a cartoon. The student-created stories, titles, and captions were given a rating by simulated classmates. The findings showed that, while students were relatively successful in creating cognitively complex creative writing, they struggled with writing in an original manner, expressing emotions in their writing, descriptiveness, and humor. Male students outperformed female students in originality and use of humor. The findings indicated mostly significantly positive relationships between different dimensions of creativity, while the relationships between student creativity score and student school achievement was found as negative. Additionally, no significant correlations were found between student creativity and factors such as motivation and time-on-task. Directions for future research are discussed in terms of their implications to large-scale assessment programs, teaching, and learning.

Keywords: creativity, higher-order thinking skills, computer-based assessment
Computer-based Performance Assessment of Creativity Skills: A Pilot Study

**Introduction**

Creative people consistently look for ways to see problems that other people don’t look for, seek ways to overcome challenges, and propose novel solutions that are of value to an individual or a social group. Educators and researchers believe that it’s important to teach and foster creativity inside the classroom, but many of them claim that the school climate and curriculum guidelines discourage creativity (e.g. Andilou & Murphy, 2010; Baer & Kaufman, 2012; Benavides, Dumont, & Istance, 2008). Moreover, problems that require divergent thinking and novelty are inadvertently devalued by the use of standardized assessments (Plucker, & Makel, 2010; Runco, 2010). Traditionally, the more the student’s answer conforms to one or more pre-determined responses, the higher the grade. Thus, in many assessment programs, a culture that rewards uncreative thinking is cultivated, rather than a culture that encourages and rewards creative answers.

Environmental factors that foster creativity include external support or acceptance of novel ideas, products, solutions or settings in which people are exposed to or can express their creativity (Lubart, & Guignard, 2004; Runco, 2004; Sternberg, 2006). Research shows that computer-based constructivist environments can more effectively promote higher-order thinking skills in comparison to traditional settings (Rosen & Salomon, 2007). Just as technology and learning sciences play an essential role in helping to develop more effective learning practices, they also can provide key improvements in assessment (Bennett, 1999; Bennett et al., 2007; Pellegrino, Chudowsky, & Glaser, 2001; Tucker, 2009). Measuring complex skills such as critical thinking, creativity, and
collaborative problem solving requires designing and developing assessments that address the multiple facets implied by these skills. One of the possible ways to achieve these changes in educational assessment is providing visible sequences of actions taken by students using various tools within the contexts of relevant societal issues and problems that people care about in everyday life. Studying the role of innovative computer-based prototypes for assessing creativity is crucial to determining whether these types of assessments can bring a real added-value into large-scale computer-based assessment programs. The purpose of this study was to provide empirical evidence of what can be achieved by intertwining computer-based tools in a performance assessment of student creativity in terms of student achievement and motivation. This paper provides findings from an empirical pilot study conducted in four counties, and discusses implications of the findings on further research and development.

**Defining Creativity Skills**

Although most researchers agree that creativity involves the development of a novel product, idea, or problem solution that is of value to the individual and/or the larger social group, researchers have had great difficulty finding consensus on a definition beyond these two criteria of novelty and value (Kaufman, 2003; Plucker, Beghetto, & Dow, 2004; Sawyer, 2012). Creativity may range from a low level, as in, for instance, solving a typical insight problem, to the very high level involved in the shift of paradigms or genres involved in science and art. The two-level definition of creativity is widely accepted by researchers, with the two levels referred to “Big C” and “little c.” Big-C creativity occurs when a person solves a problem or creates a product that has a major
impact on how other people think, feel, and live their lives. This level of creativity consists of clear-cut, eminent creative contributions. Little-c creativity, on the other hand, includes actions in which a non-expert may adapt to changes each day. Kaufman and Baghetto (2009), proposed “mini-c” and “Pro-c” in addition to the two-level definition of creativity. Mini-c refers to the creativity inherent in the learning process when students discover something for the first time. Pro-c indicates a professional expertise in a creative domain that does not attain the level of transforming the domain.

According to Assessment and Teaching of 21st Century Skills (Binkley et al., 2012) and the Partnership for 21st Century Skills (2009), a comprehensive set of creativity and innovation competencies include: creative thinking (e.g., using a wide range of idea-creation techniques, creating new and worthwhile ideas), creative work with others (e.g., communicating new ideas to others effectively, being open and responsive to new and diverse perspectives), and the ability to implement the innovation (e.g., implementing creative ideas to make a significant and useful contribution).

To be creative one must first decide to generate new ideas, analyze these ideas, and share the ideas with others. A person may have the creativity skills but may not apply them to situations that potentially involve creativity. For example, one may decide to follow other people’s ideas rather than create one’s own, or to decide not to try to persuade other people of the value of these ideas. Therefore, in order to encourage the decision to be creative, one should believe that he or she will be awarded for the attempt to be more creative rather than punished (O’Hara & Sternberg, 2000–2001). In order to promote creativity there is a need to construct opportunities to engage in it, encourage, and reward when people respond to such opportunities. However, most of the
conventional assessments penalize students if they try being creative (Beghetto, 2010; Smith & Smith, 2010; Sternberg, 1997). Student answers are often analyzed against prototype responses, while answers that reflect novel perspectives are discouraged. Thus, an educational and social atmosphere in which students feel free to play with ideas is essential in establishing optimal settings for creativity assessment, as well as teaching and learning.

There is a notable lack of consensus regarding the generalizability of creativity versus whether creativity is domain-specific. There is significant evidence that large portions of creative competency are domain-specific (e.g., Kaufman, & Baer, 2005; Sawyer, 2012; Treffinger et al., 2002). However, there are other perspectives, such as intermediate approaches (e.g., only some traits are domain-general) and developmental approaches (e.g., domain-general skills translate into domain-specific accomplishments). In this study we adopt the intermediate approach (e.g. Lubart, & Guignard, 2004), according to which some creativity skills apply to multiple domains (e.g., creating different ideas or taking risks in introducing new ideas), whereas others are unique to specific subject areas (e.g., the ability to create different solutions in science or to write a conceptually different essay in Language Arts). Although a certain level of domain knowledge is essential for creativity, too ingrained, traditional domain-specific thinking may prevent the individual from manipulating the concepts within a particular field in novel ways (Sternberg, 2006). This approach suggests involving both domain-general and domain-specific dimensions of creativity in creativity assessment.
In our research, an operational definition of creativity refers to the capacity of an individual to effectively engage in a process of developing a novel product, idea, or problem solution that is of value to the individual and/or the larger social group. Creativity skills identified by this research attempt to incorporate skills identified in other assessment frameworks, such as the Partnership for 21st Century Skills (2009) and Assessment and Teaching of 21st Century Skills (Binkley et al., 2012).

**Assessing Creativity**

Creativity assessments go beyond tests of analytical thinking in measuring performance on tasks that require individuals to deal with relatively novel situations. Creative performance can be observed and measured in ways that creative thinking cannot. Thus, creativity of products is typically the focus of assessments that vary the conditions under which individual’s creativity is measured (Runco, 2004). Participants are often asked to write stories, make collages, and do other tasks that result in some tangible product. Assessment of little c product creativity mainly relies on the consensual assessment of experts because of its relative simplicity and the consistently high levels of inter-rater agreements reached (e.g., Amabile, 1982; Getzels & Csikszentmihalyi, 1976; Hennessey, & Amabile, 1999; Kaufman et al., 2007). The most widely used tests of creativity are Torrance Tests of Creative Thinking (Torrance, 1974, 2008). The tests were designed to identify children with high creativity potential, and to transform education to fully realize the creative potential of every student. In these tests, Verbal and Figural forms include various subtests. The Verbal subtest consists of picture construction, picture completion, and sketching different objects by using lines and circles. The Figural
battery includes creating as many questions as possible about a picture and a common object, proposing unusual uses for a common object, guessing causes, guessing consequences, product improvement, and listing possible ramifications for an improbable situation. The revised tests report scores of fluency, originality, elaboration, resistance to premature closure, and abstractness of titles (Ball & Torrance, 1984; Hebert et al., 2002; Torrance, 2008).

Performance tasks are expected to tap into an important part of creativity that might not be measured using multiple-choice items alone because open-ended measures require more authentic and free-form responses. For example, in the Rainbow Project study, creativity was measured using open-ended, performance-based tasks among first-year college students in the United States (Sternberg & the Rainbow Project Collaborators, 2006). For each of the tasks, participants were given a choice of topic or stimuli on which to base their creative written stories, oral stories, or cartoon captions. Each of the creativity performance tasks was rated based on rubrics that were predetermined as measures of creativity. The score for creativity in cartoons captions was formed by summing the ratings given by judges on originality, cleverness, and humor on 5-point scales. Both written and oral stories were rated for originality, complexity, emotional evocativeness, and descriptiveness on 5-point scales. The study found that student creative performance in these tasks was separated from the other more conventional tests. It was also found that adding the creative measures to other measures of analytical and practical measures roughly doubled the predictive value of the SAT for the sample in predicting grades for first-year college students (Sternberg, & the Rainbow Project Collaborators, 2006).
Collaborators, 2006). The measures also served to decrease ethnic differences between groups.

This study adopted the assessment approach developed by Sternberg and the Rainbow Collaborators (2006) and extended it to the computer-based interactive performance tasks. As shown in the methodology section of this paper, students were asked to write different short stories based on given images or a video, create different titles for a story, and propose different captions for a cartoon. The student-created stories, titles, and captions were given a rating by simulated classmates. In addition, thinking tools, such as capturing bookmarks from the movie and using them to create the story, were embedded into the assessment. Embedding computer-based interactive tools in creativity performance assessment is one of the promising approaches that should be further explored.

In summary, technology offers opportunities for assessment in domains and contexts where assessment would otherwise not be possible or would not be scalable. One of the important premises brought by technology to educational assessment of creativity is the capacity to embed interactivity. These interactive components could be designed in such a way that students will be fully engaged in a task without introducing construct irrelevant difficulties and additional time allocations. Thus, it is necessary to understand the nature of the task and the creativity skills required while designing an appropriate computer-based assessment. This paper addresses these challenges by studying student performance in a computer-based assessment prototype of creativity.
Research Questions

The study addressed empirically the following questions regarding student performance and motivation in creativity assessment:

1. What are the differences in the student performance between different creativity dimensions?

2. How are a student’s abilities in different dimensions of creativity related to each other?

3. What are the differences in the student creativity performance between males and females?

4. How are a student’s GPA and ELA school achievement, as reported by the students, related to student performance in creativity assessment?

5. How is a student’s performance in creativity task related to student motivation and time-on-task?

Method

The study participants included 87 students, who were all 14 years old and from the United States, United Kingdom, Turkey, and South Africa. The results presented in the current article came from a larger study in which students from six countries were recruited to participate in a 21st Century Skills Assessment project study investigating the innovative ways of developing computer-based assessment in critical thinking, creativity, and collaborative problem solving (see Rosen & Tager, 2013a, for study of collaborative problem solving, and Rosen & Tager, 2013b, for study of critical thinking). The researchers collected data from November 2012 through January 2013. Recruitment of
participating schools was achieved through collaboration with local educational organizations based on the following criteria: (a) the school is actively involved in various 21st Century Skills projects, (b) population of 14-year-old students proficient in English, and (c) sufficient technology infrastructure (e.g., computers per student, high-speed Internet). Of the total students who participated, 47 were boys (54%) and 40 were girls (46%). Table 1 summarizes the country and gender distribution of participating students.

Table 1

<table>
<thead>
<tr>
<th>Country</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Turkey</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>South Africa</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td><strong>40</strong></td>
<td><strong>47</strong></td>
</tr>
</tbody>
</table>

**Creativity Assessment**

In this creativity computer-based assessment task, the student was asked: (a) to create two different short stories by using given images; (b) write three different captions for a single cartoon; and (c) write a story based on a video and create three different titles for a story. The student’s writing was accompanied by a rating given by simulated classmates.

Due to the exploratory nature of the study, the students were not limited in time-on-task. The task was checked by teachers from the four participating countries to ensure
that students would be able to work on the task, and that the task could differentiate between high and low levels of creativity ability. Interviews were conducted with students representing the target population to validate the assessment approach.

The following information was presented interactively to the student during the task:

Episode #1: The task starts by asking the student to provide name, background information, and to select the preferred avatar. Then the tool panel and the story panel were introduced to the student as presented in Figures 1-2.

![Tool panel](image)

*Figure 1. Introducing the tool panel to the student*
Episode #2: The task was initiated by asking a student to write a short story (up to 40 words) by using the images presented on the tool panel (Figure 3).

**Figure 2.** Introducing the story panel to the student

**Figure 3.** Writing a short story using given images
In order to increase student motivation and provide a reason for asking the student to create a different story later based on the same images, the submitted story was followed by a simulated rating given by classmates (Figure 4).

![Simulated story rating given by classmates](image)

*Figure 4. Simulated story rating given by classmates*

The rating was pre-determined to indicate 5 out of 10 stars and was given a few seconds after the submission of the story. Then the student was asked to write a different story based on the same images. The second story submission was followed by a simulated rating given by classmates and was pre-determined to show 9 out of 10 stars.

Episode #3: Figure 5 shows an example of the task screen during the cartoon captions episode. In this assessment item, the student was asked to generate three different ideas for a cartoon caption (up to 10 words each). Similarly to the previous task, the submitted caption was followed by a simulated rating given by classmates.
Figure 5. Writing a cartoon caption

Episode #4: Figures 6-7 show examples of the task screens during the video-based story episode. The student was asked to create a story based on a YouTube theme video (up to 60 words).
Figure 6. Capturing video snapshots for the story
Figure 7. Writing a video-based story

Episode #5: In the last assessment item (Figure 8), the student was asked to generate three different ideas for a story title (up to 10 words each). A simulated rating given by classmates was presented for each created title.
Creativity Scoring Criteria

Scoring of the student responses was provided independently by two teachers from participating schools in the United States. Each judge rated the stories for originality, cognitive complexity, emotional expressiveness, and descriptiveness on 3-point scales. The captions for the cartoon and the titles for the written story were rated for originality and humor on 3-point scales. For the purposes of more meaningful interpretation of student scores for the teachers, the 0-3 scale was later converted into 0-100% scale. Because of the exploratory nature of this pilot study, a combined creativity score was formed by averaging the individual ratings on each dimension. Inter-coded agreement of scoring was 87% for originality, 92% for cognitive complexity, 90% for

Figure 8. Writing a title for a video-based story
emotional expressiveness, 94% for descriptiveness, and 92% for humor. It should be noted that student responses were scored based on the criteria presented above, while spelling and grammar issues did not affect the student score.

**Motivation Questionnaire**

The questionnaire included 4 items to assess the extent to which students were motivated to work on the task. Participants reported the degree of their agreement with each item on a 4-point Likert scale (1 = strongly disagree, 4 = strongly agree). The items, adopted from motivation questionnaires used in previous studies, included (Rosen, 2009; Rosen & Beck-Hill, 2012): I felt interested in the task; The task was fun; The task was attractive; I continued to work on this task out of curiosity. The reliability (internal consistency) of the questionnaire was .84.

Students were also asked to indicate their background information, including gender, Grade Point Average (GPA), and English Language Arts (ELA) average score. It should be noted that this informant was provided directly by the students and was not validated with the teachers. The background information was collected because of potential interaction with study variables.

**Results**

All results are presented on an aggregative level beyond the countries because no interaction with country was found. First, the descriptive results of student performance in a creativity assessment are presented to determine whether there is a difference in student scores in different dimensions of creativity. Next, the results regarding the
relationship between different dimensions of creativity are shown. Then, gender-related results are presented to indicate possible differences in student performance in creativity, as well as the relationship with a student’s school achievement. Last, the relationships between student performance in creativity and student motivation, as well as time-on-task, are demonstrated.

Student Creativity Performance

The results of the creativity scores indicated that while students were relatively successful in cognitively complex writing (M=70.0, SD=26.0), they struggled with writing in an original manner (M=13.9, SD=14.8), expressing emotions in their writing (M=21.0, SD=26.0), and descriptiveness (M=25.0, SD=20.2). Of special difficulty for the students was the ability to use humor in their writing (M=2.1, SD=5.1). Overall, the average creativity score was 26.4 (SD=9.8), while the maximum and the minimum scores were 51.8 and 4.4, respectively.

Relationships between Creativity Dimensions

To better understand the relationship between the dimensions of creativity, analysis of intercorrelations between the variables was conducted. Table 2 reports the intercorrelations between the dimensions of creativity. The findings showed significantly positive relationships between student originality score and the ability to use humor (r=.63, p < .01), express emotions, and descriptiveness (r=.53, p < .01, for both dimensions). Emotional expression was also positively correlated with descriptiveness (r=.77, p < .01). Low negative statistically significant correlations were found between
the complexity dimension and both emotionality and descriptiveness ($r=-.23$, $p < .05$, and $r=-.26$, $p < .05$, respectively).

Table 2

*Intercorrelations between dimensions of creativity*

<table>
<thead>
<tr>
<th></th>
<th>Originality</th>
<th>Complexity</th>
<th>Emotionality</th>
<th>Descriptiveness</th>
<th>Humor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Originality</td>
<td></td>
<td>-.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complexity</td>
<td></td>
<td></td>
<td>-.23*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotionality</td>
<td>.53**</td>
<td></td>
<td></td>
<td>.77**</td>
<td></td>
</tr>
<tr>
<td>Descriptiveness</td>
<td>.53**</td>
<td>-.26*</td>
<td></td>
<td>.77**</td>
<td></td>
</tr>
<tr>
<td>Humor</td>
<td>.63**</td>
<td>-.19</td>
<td>.18</td>
<td>.18</td>
<td></td>
</tr>
</tbody>
</table>

** $p < .01$, * $p < .05$.

**Gender and Creativity Performance**

The findings indicated that males outperformed females in originality of their writing ($M=18.0$, $SD=17.9$, in male students, compared to $M=9.0$, $SD=7.5$ in females; $ES=.7$, $t(\text{df}=85)=3.1$, $p < .01$), and the use of humor ($M=3.3$, $SD=6.3$, in male students, compared to $M=3.6$, $SD=2.3$ in females; $ES=.6$, $t(\text{df}=85)=2.7$, $p < .01$). However, no significant difference between female and male student was found in the overall creativity score ($M=26.3$, $SD=9.3$, in female students, compared to $M=26.5$, $SD=10.3$ in males; $t(\text{df}=85)=-.1$, $p=.93$).

**Student School Achievement and Creativity Performance**

Correlations between the variables were conducted in order to determine potential relationships between student-reported GPA and ELA achievement as measured by
traditional school assessments and student performance in creativity assessment. The findings showed significantly negative correlation between student creativity score and student school achievement as reflected by GPA and ELA ($r=-.34$, $p < .01$ and $r=-.35$, $p < .01$, respectively). Student GPA was found in high negative correlation with student originality ($r=-.40$, $p < .01$), humor ($r=-.25$, $p < .05$), and cognitive complexity in writing ($r=-.21$, $p < .01$). Similar patterns were found regarding the relationships between student ELA score and creativity dimensions of originality ($r=-.43$, $p < .01$) and humor ($r=-.30$, $p < .01$).

**Student Motivation and Time-on-Task**

Data was analyzed to determine the possible relationship between student motivation to participate in creativity assessment and student performance. The results demonstrated no significant relationship between student motivation and student performance in creativity ($r=.11$, $p=.74$). In addition, no significant correlation was found between time-on-task and student performance in creativity ($r=.19$, $p=.08$). On average, time-on-task in the assessment was 20.8 minutes (SD=6.0), while the maximum and the minimum time-on-task were 48.0 and 8.1, respectively.

**Discussion**

In order to understand how students perform on creativity computer-based assessment with interactive components, it is necessary to empirically examine student performance in these environments. The goal of this study was to explore patterns in student creativity performance and motivation in computer-based interactive prototype of
assessment. While technology tools can promote fundamental improvements in assessment of higher-order thinking skills (Bennett, 1999; Bennett et al., 2007; Pellegrino, Chudowsky & Glaser, 2001; Tucker, 2009), assessment of foundational knowledge, skills, and abilities can rely on more traditional assessment approaches. Interactive components that were used in this study can enable scaffolding and visibility in student thinking process while working on complex problem solving or decision-making situations that require creative thinking.

Students assessed in creativity computer-based performance task showed relatively high performance level in cognitive component of creativity, but their performance on all other components of creativity was significantly lower. Findings showed low performance levels in terms of originality, emotional expressiveness, and humor across the countries. It should be noted that the assessment was introduced to students as a general test of 21st Century skills and there are no right or wrong answers. The assessment criteria (e.g. “be original,” “add humor”) or any other information about the assessment were not included in the introduction (see Method section for more details). Similarly to other types of assessment, providing more information about the assessment measures could potentially affect student performance in a creativity test (e.g., Barron & Harrington, 1981, in the context of Torrance Test of Creative Thinking). Thus, further investigations are needed to examine the effect of introducing the assessment measures (e.g. originality, humor) on student performance.

Analysis of the intercorrelations between the creativity dimensions showed that overall there are mid-level significantly positive relationships between the dimensions. These results suggest that originality, emotionality, descriptiveness, and humor are
distinctive dimensions of the creativity construct. The only exception was found in the context of the relationships between a complexity dimension and both emotionality and descriptiveness, in which low-level significantly negative correlations were found. Therefore, the complexity dimension can be potentially considered as a construct irrelevant component for the creativity assessment. Further investigations should examine the generalizability of this pattern across different creativity assessment tasks.

Additional analysis examined the extent to which student GPA and ELA achievements, as reported by the students, are related to student creativity performance. We found evidence for a negative relationship between GPA, ELA, and creativity student scores. These results suggest that the current creativity approaches for assessment are distinctive from measurement of the conventional domains in schools. The requests to create different stories based on a given theme or propose different titles for a cartoon were new to many students, but allowed each of them to show creativity relatively regardless of his or her reading, writing, and other skills measured in school settings. Creating different ideas under identical conditions is a cognitively challenging task that requires divergent thinking and novelty, which are often discouraged in school climate and curriculum (Andilou & Murphy, 2010; Benavides, Dumont & Istance, 2008). More often a culture of rewarding uncreative thinking is cultivated, rather than encouraging and rewarding creative answers. Again, it should be noted that this informant was provided directly by the students and was not validated with the teachers.

With respect to possible differences in creativity performance between female and male students, we found that males were more successful in originality and humor dimensions. However, no significant differences were found in the overall creativity
score. Although previous creativity research revealed no gender differences (see Baer & Kaufman, 2008, for a meta-analysis), some studies found that males outperform females in originality (e.g., Jiliang & Baoguo, 2007). In addition, it is possible that the boys and the girls reacted differently to the simulated rating that was embedded into the assessment and created a competitive context. Similar to more conventional assessments with no immediate feedback to the student, students may benefit differently from qualitatively different types of assessment items types or environments. In this assessment the pre-defined classmates’ rating tool was introduced to increase student motivation and the authenticity of the task. One may consider adding real-time true rating capacity to such an assessment, as well as group contest settings. Contests and awards are widely used to foster creativity among teams through a sense of collaborative mission (e.g., Baer et al., 2010; Birkinshaw, & Lingblad, 2001).

The current study had several limitations. First, it is based on a relatively small and non-representative sample of 14-year-old students in four countries and self-reported school achievement. However, due to a lack of empirical research in the field of computer-based assessment of creativity skills with embedded interactive tools, it is necessary to conduct small-scale pilot studies in order to inform more comprehensive approaches of creativity assessment. Further studies could consider including a representative sample of students with a wider range of ages and backgrounds and collecting school achievement data a more controlled manner. Second, the study operationalized the interactive tools in critical thinking assessment through a simulated rating and video-capturing functionality, while other approaches could be considered, including semantic organization tools, dynamic modeling tools, information
interpretation tools, knowledge construction tools, microwords, and conversation and collaboration tools (Jonassen, 2006; Jonassen, & Reeves, 1996). Finally, it is important to conduct comparability studies between different modes of creativity assessment with and without interactive tools to explore similarities and differences in student performance. Moreover, future studies could consider exploring differences in student performance in a wide range of problems and situations that require creativity.
References


