The Cross-Cultural Invariance of Creative Cognition: A Case Study of Creative Writing in U.S. and Russian College Students

Sergey A. Kornilov, Tatiana V. Kornilova, Elena L. Grigorenko

Abstract

Unlike intelligence, creativity has rarely been investigated from the standpoint of cross-cultural invariance of the structure of the instruments used to measure it. In the study reported in this article, we investigated the cross-cultural invariance of expert ratings of creative stories written by undergraduate students from the Russian Federation and the United States. Analyses of differential rater and item functioning using Many-Facet Rasch Measurement and multiple levels of invariance using confirmatory factor analyses suggested partial measurement invariance of creative ability estimates obtained using this method in two cultures. Russian and U.S. students demonstrated similar overall levels of creativity; however, U.S. students received higher emotionality ratings than Russian students did. The findings are discussed in the context of viewing creativity as at least a partially culturally invariant trait whose manifestation is moderated by culture-specific semantic knowledge and patterns of linguistic behavior. © 2016 Wiley Periodicals, Inc.
Creativity is widely recognized as an inherently cultural phenomenon, embedded in a complex network of environmental influences. Thus, definitions of creativity often explicitly or implicitly mention culture as the key contextual element that both (a) shapes creativity as an ability that depends on the functioning of convergent and divergent processes and the acquisition of knowledge, and (b) serves as the necessary framework for the evaluation of the products resulting from the application of this ability in a specific task context. Underscoring the role of culture in the development and manifestation of creativity mandates the investigation of the complex interactions between cognitive processes that support creativity and symbolic, behavioral, ideational, and value patterns transmitted and internalized through exposure to a particular culture (Kroeber & Kluckhon, 1952) and its specific contexts, including family and school (Mourgues, Barbot, Tan, & Grigorenko, 2014). These interactions are hypothesized to result in the development of the ability to create novel, original, elaborated, and task as well as culturally appropriate ideas, solutions, and products. Crucially, the requirement for the creative product to be culturally appropriate (and effective; see Runco & Jaeger, 2012, for a discussion on the standard definition of creativity) reflects the existence of substantial differences in the previously mentioned symbolic and value patterns across different cultures, and has important implications for our understanding of the nature of this ability as well as our approaches to measuring it.

Studies of cross-cultural differences in creativity mainly revolve around the examination of conceptions of creativity across different cultures, cross-cultural variation in personality traits that are viewed as characteristic of a “creative person,” and direct evaluation of creative ability across diverse cultures that generally hinges on the evaluation of the characteristics of a particular creative product. With respect to conceptions of creativity, the general consensus is that there are properties or features of creativity that are recognized as key across multiple cultures—most notably specific cognitive processes (e.g., divergent thinking) and personality traits (e.g., risk taking and perseverance) that contribute to creativity (Erez & Nouri, 2010), and characteristics of the “final” creative product such as novelty/originality and its adaptive value/task appropriateness (Lubart, 2010). However, cultures have been found to differ with respect to the relative emphasis that they place on specific characteristics of the product (i.e., novelty vs. adaptation) and/or on the ways they regulate the specifics of the implementation of these characteristics (Rubera, Ordanini, & Griffith, 2011; Sternberg & Lubart, 1995).

Both the relative emphasis on different features of creativity across cultures and the cultural specificity of the actual manifestation of these features reflect the impact of the cultural expectations regarding what is considered creative and why. Although multiple accounts of how these expectations shape creativity have been developed, they mostly concern the so-called Big C or eminent creativity—the types of creative cognition and products
that reshape the culture from within. However, much less is known with respect to the extent that these cultural influences are manifested in the little c (i.e., in individual differences in creative ability estimates viewed through the lens of the rater evaluation of task-specific creative products).

The study reported in this article attempts to address this gap in the literature by examining the cross-cultural differences in the ratings of creative products produced by Russian and U.S. college students. However, instead of focusing on the group differences in creative ability estimates obtained by the two groups of students, we focused our attention on the internal structure and psychometric properties of the writing task we used to elicit creative stories. Correspondingly, we examined the cross-cultural invariance of creative cognition in the domain of creative writing. Such an examination serves the purpose of illuminating the extent to which the cognitive architecture of verbal creativity is similar across the two cultures (in the context when the same process is realized through the means of distinct symbolic and linguistic systems the use of which is culturally regulated), and whether there are dimensions of creative cognition (measured through product) that are particularly amenable to cultural influences. We performed this investigation within the psychometric paradigm, focusing our attention on (a) construct invariance—the extent to which creative ability, as measured by the Creative Stories task—exhibits similar internal structure in the two samples of students and (b) differential item functioning (DIF), traditionally viewed as indicating the presence of a local and specific bias in particular elements of creative ability that cannot be attributed to overall group differences in creative ability.

Methods

Our quasi-experimental study focused on the cross-cultural invariance of verbal creativity ratings obtained through administering an open-ended creativity task (see below) in two samples of undergraduates attending colleges in Russian Federation and the United States.

Participants. One thousand undergraduate students participated in the study—500 from Moscow State University (MSU; in the age range from 18 to 60 years, with \( M = 21.11, SD = 4.78; 390 \) were females), Russian Federation, and 500 from the University of California in San Bernardino (UCSB) and San Francisco (UCSF), United States, in the age range from 17 to 58 years (\( M = 24.97, SD = 6.29; 325 \) females). Of these, after the removal of missing data, responses were provided by 318 Russian and 346 U.S. students, for the total effective sample of 664.

The analyses capitalized on two existing datasets: the U.S. sample constituted part of the Transitions in the Development of Giftedness project (Sternberg et al., 2007) data, whereas the Russian sample was recruited for the cross-cultural investigation of the psychometric properties of some of
the assessments used by Sternberg et al. (2007; the complete description of the sample can be found in Kornilov & Grigorenko, 2010).

**Creative Stories Task.** All participants were administered an open-ended creative writing task called Creative Stories. According to the task procedure, participants were asked to write a brief creative story in 15 minutes using one of the five offered titles: “Two Chatting Spiders,” “The Fishing Moose,” “A Spotted Creature,” “A Banana with Many Peels,” and “The Reading Dragon.” Participants were instructed to use only one of the titles in their story and to write as creatively as possible to maximize performance.

The written stories were rated for creativity according to four key dimensions (hereafter, items) using the comprehensive scoring rubrics manual developed by the Transitions in the Development of Giftedness project team (Sternberg et al., 2007) based on Sternberg’s investment theory of creativity (Sternberg, 2012). Thus, the stories were rated on **Originality** (novelty, uniqueness, and unexpectedness of the story plot and its elements, individualized for each title), **Complexity** (of story structure, form/writing, and imagery), **Emotionality** (incorporation of emotion in the story and elicitation of emotions from the reader), and **Task Appropriateness** (whether the story is sufficiently descriptive and related to the title in a clear and logical manner). Each story was rated on the four dimensions using a five-category polytomous rating scale. Each category was supplemented by a detailed description of the scoring criteria (e.g., to receive a score of 4 on Complexity, the story had to have a solid and elaborated structure with imagery that the rater could describe as “striking” in at least one instance) and examples.

The stories generated by U.S. students were scored by two raters using a rating plan with an overlap of \( k = 30 \) written stories. The stories generated by Russian students were scored by three raters with the overlap of \( k = 100 \) stories. A set of \( k = 60 \) stories from the U.S. sample was also scored by one bilingual rater who scored stories from the Russian sample. All raters underwent extensive training prior to scoring the stories to reach acceptable within-country agreement levels (Pearson’s \( r = .60 \) or Cohen’s \( K = .50 \) for each element). Within- and between-country interrater reliabilities are presented in the Results section (Table 4.1).

### Table 4.1. Interrater Reliabilities for the Creative Stories Task Scores

<table>
<thead>
<tr>
<th>Item</th>
<th>Russian Raters</th>
<th>U.S. Raters</th>
<th>Russian and U.S. Raters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r )</td>
<td>Cohen’s ( K )</td>
<td>( r )</td>
</tr>
<tr>
<td>Originality</td>
<td>.95</td>
<td>.92</td>
<td>.63</td>
</tr>
<tr>
<td>Complexity</td>
<td>.90</td>
<td>.86</td>
<td>.94</td>
</tr>
<tr>
<td>Emotionality</td>
<td>.75</td>
<td>.71</td>
<td>.97</td>
</tr>
<tr>
<td>Descriptiveness</td>
<td>.95</td>
<td>.90</td>
<td>.87</td>
</tr>
</tbody>
</table>

*Note: Estimates represent averages of pairwise estimates obtained for each pair of raters who rated the same story; \( r = \) Pearson’s correlation coefficient; Cohen’s \( K = \) Cohen’s Kappa (chance-adjusted agreement coefficient).*
The raters were instructed to be lenient and assign the higher score when in doubt.

**Analytical Approach.** The analysis was performed in two distinct yet related frameworks. First, we analyzed the data using a Many-Facet Rasch Measurement (MFRM) model as implemented in Facets for Windows (Linacre, 2009). Briefly, this approach uses a unified conjoint measurement item response theory (IRT) model to obtain stable calibrations of multiple facets on a common scale, e.g., estimates of student ability (creativity), rater severity, and item difficulty. We analyzed the data using a five-facet MFRM rating scale model with students, items, raters, titles, and country designated as separate facets.

Crucially, MFRM enables the investigation of systematic biases in expert ratings by performing differential facet analysis (DFF; Bond & Fox, 2007). We used DFF analysis to establish the invariance of item difficulty (differential item functioning, DIF) and rater severity (differential rater functioning, DRF) parameters in the U.S. and the Russian samples within the same model, and we used the omnibus test for the presence of DFF elements as well as the t criterion to test each interaction parameter separately. First, the measures for all facets and the structure of the rating scale were calibrated and fixed; at the second step, the expected values for observations were subtracted from observed values, and residuals corresponding to each interaction term were integrated and transformed into DFF bias estimates. The DFF parameters were tested for statistical significance using the t statistic (using the standard Bonferroni procedure to correct for multiple statistical testing at $\alpha = .05$), and their overall impact was evaluated with the omnibus test for the presence of DFF items (Linacre, 2009) and percentage of variance in creativity ratings attributable to DFF.

We also analyzed the cross-cultural invariance of the Creative Stories task through multigroup maximum-likelihood confirmatory factor analysis (CFA) as implemented in EQS for Windows (Bentler, 1995) using the robust maximum likelihood estimation. Briefly, cross-cultural invariance testing within the CFA framework proceeds via the iterative comparison of nested CFA models with progressively stringent sets of constraints (Byrne, 2006) at each of the following levels of invariance: configural (equivalence of the number of latent factors and patterns of factor loadings in two groups), metric (magnitude of factor loadings in two groups), scalar (equivalence of indicators intercepts), and residual (equivalence of error variances). When more than one rater's scores were available for a particular participant, the scores were averaged prior to the analysis.

**Results**

As mentioned above, the analyses reported in this section focused on the interrater reliability of scores obtained for the Creative Stories task, its cross-cultural invariance, and the presence of differential item functioning.
**Interrater Reliability.** The interrater reliabilities are presented in Table 4.1. Overall, the raters demonstrated high interrater agreement levels within groups (mean $r = .89$ and .85 for Russian and U.S. raters, respectively). Although the interrater reliabilities between the Russian and the U.S. raters were lower ($r$ ranged from .54 to .64 for the four key dimensions), it was still at an acceptable level. Analyses of Cohen’s $K$ (Cohen, 1960) suggested that within-country chance-corrected reliabilities were in the moderate to substantial (Landis & Koch, 1977) range (from .71 to .92 for Russian raters; from .50 to .89 for U.S. raters); however, agreement was lower between the Russian and the U.S. raters (with the lowest estimate for Emotionality, .10).

**Many-Facet Rasch Modeling.** To further investigate whether there were systematic biases in the scores of the rater who scored the stories from both countries, we fitted a unidimensional five-facet MFRM model to the data.

The model explained 60.2% of variance in the Creative Stories scores obtained by the students. The person reliability index (estimated reproducibility of relative measure location, interpreted similarly to Cronbach’s alpha; Linacre, 1997) was high at .85, and the model reliably differentiated between at least three levels of performance (separation index = 2.42). The analysis of local fit indices did not reveal items, raters, or story titles that violated model assumptions (all InfitMS < 1.37, all OutfitMS < 1.34; values between 0.50 and 1.50 generally suggest acceptable measurement quality; Wright & Masters, 1990; see Table 4.2) indicating overall satisfactory measurement properties of the task as applied to our samples.

**Table 4.2. Resulting Facet Element Parameters for the MFRM Model of Creative Stories**

<table>
<thead>
<tr>
<th>Facet</th>
<th>Average</th>
<th>Est.</th>
<th>SE</th>
<th>Infit</th>
<th>Outfit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>US</td>
<td>2.93</td>
<td>.51</td>
<td>.04</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>RUS</td>
<td>2.65</td>
<td>.51</td>
<td>.02</td>
<td>1.02</td>
</tr>
<tr>
<td>Item</td>
<td>Originality</td>
<td>2.81</td>
<td>-.10</td>
<td>.04</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>Complexity</td>
<td>2.96</td>
<td>-.66</td>
<td>.05</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td>Emotionality</td>
<td>2.17</td>
<td>1.19</td>
<td>.04</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Descriptiveness</td>
<td>2.95</td>
<td>-.43</td>
<td>.04</td>
<td>.85</td>
</tr>
<tr>
<td>Title</td>
<td>Title 1</td>
<td>2.66</td>
<td>.15</td>
<td>.04</td>
<td>.98</td>
</tr>
<tr>
<td></td>
<td>Title 2</td>
<td>2.87</td>
<td>-.24</td>
<td>.05</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>Title 3</td>
<td>2.74</td>
<td>-.03</td>
<td>.04</td>
<td>.96</td>
</tr>
<tr>
<td></td>
<td>Title 4</td>
<td>2.63</td>
<td>.05</td>
<td>.05</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>Title 5</td>
<td>2.70</td>
<td>.07</td>
<td>.05</td>
<td>1.04</td>
</tr>
<tr>
<td>Rater</td>
<td>Rater 1 (RUS/US)</td>
<td>2.57</td>
<td>.31</td>
<td>.03</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Rater 2 (RUS)</td>
<td>2.69</td>
<td>.30</td>
<td>.05</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Rater 3 (RUS)</td>
<td>2.78</td>
<td>.16</td>
<td>.04</td>
<td>.98</td>
</tr>
<tr>
<td></td>
<td>Rater 4 (US)</td>
<td>2.96</td>
<td>-.16</td>
<td>.06</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Rater 5 (US)</td>
<td>2.96</td>
<td>-.61</td>
<td>.06</td>
<td>1.01</td>
</tr>
</tbody>
</table>
The analysis of rater by country interaction suggested the absence of a significant DRF, $t_{(406)} = .05$, $p = .96$. Thus, the rater who was “linking” the U.S. and Russian samples applied the scoring rubrics consistently in the same fashion across the two samples. However, the analysis of DIF at the item level revealed significant item by country bias: out of the four items, two—Complexity and Emotionality—were differentially difficult for the two samples. Specifically, adjusting for the overall level of creativity and rater severity, it was more difficult for Russian students to obtain higher scores on Emotionality (item severity = 1.54), compared to the U.S. students (item severity = .80), $t_{(928)} = -8.06$, $p < .001$. On the other hand, Complexity was easier for Russian (–.72) compared to U.S. students (–.14), $t_{(909)} = 6.46$, $p < .001$.

Although the overall magnitude of bias in both cases was moderate, and differential item functioning in the two samples explained only .77% of the total variance in Creative Stories scores, it had important implications for the possible comparisons of the creativity scores in the two groups. Thus, without adjusting for bias, Russian students overall received higher creative ability estimates (Est. = .73, $SE = .02$) compared to the U.S. students (Est. = .59, $SE = .04$), for test of resulting ability estimates, the fixed all-same chi-square; $\chi^2 (1) = 8.60$, $p < .01$ (Linacre, 2009). In order to examine whether these differences could be attributable to DIF, we reanalyzed the data while respecifying the rating scale structure in a way that permitted the two samples to have unique item difficulty estimates for Emotionality and Complexity. After the adjustment for bias, the two samples were no longer significantly different on the overall creative ability estimates, $M = .51$, $SE = .02$, and $M = .51$, $SE = .04$ for Russian and U.S. samples, respectively, $\chi^2 (1) = 0.05$, $p = .96$.

**Confirmatory Multigroup Factor Analysis.** Using confirmatory factor analysis, we examined the cross-cultural invariance of the Creative Stories task by evaluating a set of multigroup models that imposed constraints on the overall structure of the model and patterns of item-factor loadings (configural invariance), magnitude of item-factor loadings (metric invariance), item intercept values (scalar invariance), and error variances (residual invariance). The baseline one-factor model (Figure 4.1) included four indicator variables that had loadings on the common factor of creative ability. One of the loadings has been set to 1 for the purpose of model identification and latent variable scaling. The model displayed satisfactory fit in the U.S. sample, $\chi^2 (2) = .11$, $p = .943$, comparative fit index (CFI) = 1.00, Root Mean Square Error of Approximation (RMSEA) < .001, and, although somewhat worse, satisfactory fit in the Russian sample, $\chi^2 (2) = 9.86$, $p = .007$, CFI = 0.96, RMSEA = 0.10. The two-sample configural invariance model displayed satisfactory fit, $\chi^2 (4) = 10.08$, $p = .038$, CFI = 0.98, RMSEA = 0.08, suggesting that the configural invariance of the one-factor solution holds across both samples.
At the next step, we constrained the freely estimated item-factor loadings to be equal across the two groups: this metric invariance model displayed satisfactory fit ($\chi^2 (7) = 12.84$, $p = .076$, CFI = 0.98, RMSEA = 0.06), and the difference between the levels of the fit of the configural and metric invariance models was not significant, ($\Delta \chi^2 (3) = 2.76$, $p = .43$, $\Delta$CFI < 0.01). The examination of the results of the Lagrange Multiplier (LM) constraint testing statistic suggested that, across the two groups, the factor loadings were similar in magnitude, and all $p$s were >.05, indicating full metric invariance of the Creative Stories task in the samples of Russian and U.S. students.

To examine scalar invariance, in addition to imposing the set of constraints of the metric model, we constrained item intercepts to be equal in both groups of students. The model displayed worse fit, $\chi^2 (10) = 52.45$, $p < .001$, CFI = 0.97, RMSEA = .080, than the metric invariance model, $\Delta \chi^2 (3) = 39.61$, $p < .001$, $\Delta$CFI = 0.01. The examination of parameter estimates and results of LM testing suggested that the intercepts were non-invariant for the Emotionality item (at $p < .001$), corroborating the results from the MFRM analysis. Removing the constraint of item variance equality for Emotionality led to improved model fit, and the partial scalar invariance was reached, $\chi^2 (8) = 16.50$, $p = .036$, CFI = 0.99, RMSEA = .059, $\Delta \chi^2 (2) = 3.66$, $p = .16$, $\Delta$CFI < 0.01.
Finally, when we constrained item residuals to be equal in both groups, in addition to imposing the partial scalar invariance constraints, none of the items demonstrated noninvariant error variances, and the model displayed satisfactory fit, $\chi^2 (12) = 19.37, p = .08$, CFI = 0.99, RMSEA = .041. This model showed a fit similar to that of the model with partial scalar invariance constraints only, $\Delta \chi^2 (4) = 2.88, p = .58$, $\Delta$CFI < .01, suggesting that residual error invariance holds for the Creative Stories task.

**Discussion**

Using two different analytical approaches, we demonstrated the partial cross-cultural invariance of the Creative Stories task, an open-ended measure of verbal creativity that relied on ratings provided by multiple raters who scored creative stories on four different dimensions. Our many-facet Rasch analysis suggested that, despite higher within-compared to between-country levels of interrater reliability, the raters were able to apply the scoring rubrics consistently when scoring stories written by U.S. and Russian students. The analysis of differential item functioning, however, suggested that two items exhibited significant bias—it was more difficult for Russian students to obtain higher ratings on Emotionality, compared to the U.S. students, and the situation was reversed for Complexity. The results held when, in a follow-up analysis within a generalized linear model (GLM) framework, the linking bilingual expert was removed from the data. Although these biases explained only a minor fraction of the variance in students’ creativity ratings, they were sufficient to drive the statistically significant difference in latent creativity estimates in the two samples. When adjusted for this bias, the group differences in overall creativity levels disappeared, suggesting that, overall, Russian and U.S. undergraduates display similar levels of verbal creativity in the written domain.

We obtained nearly identical results using multigroup confirmatory factor analysis for cross-cultural invariance testing. Thus, the measurement model was invariant in the two samples at the configural level and the metric level but was only partially invariant at the scalar and residual level. The partial scalar noninvariance was driven by the significant difference in the Emotionality item intercept between the two samples in the direction parallel to the findings from the MFRM analysis. Although the two methods differ in assumptions, properties, and interpretations of measurement invariance analyses, they converged on Emotionality as the particularly noninvariant element of the Creative Stories task. The advantage of the MFRM analysis is its ability to handle complex rater plans and take into account nonlinearity of the scores as manifested in category thresholds; the advantage of the CFA, on the other hand, is its ability to model differential item discrimination. The two approaches can be viewed as complementary and tend to produce similar results (Salzberger & Sinkovics, 2006), although
the research on the comparative power of these approaches in detecting measurement noninvariance is scarce.

All other parameters held equal, Russian students tended to receive lower scores on Emotionality, a Creative Stories item that requires the raters to evaluate the presence and efficiency of the emotional component in written stories. Specifically, the raters were asked to evaluate whether the story mentioned emotional states and used mental language, the richness of these emotional representations, and whether the story was effective in eliciting an emotion from the raters. Emotions have long been recognized as at least partially culturally mediated despite showing substantial evidence for the presence of certain cross-cultural universals (Russell, 1991). Creative writing places heavy demands on the use of the symbolic system that might reflect the different ways in which different cultures label, conceptualize, and describe emotions—that is, language (Kovecses, 2003; Wierzbicka, 1986). This point of view has received substantial support in linguistic anthropology (Goddard & Zhengdao, 2014; Wilce, 2014) and suggests the need for a detailed semantic analysis of cross-cultural representations of emotions but also of the patterns of the use of language to describe mental states in general and emotions in particular. These patterns can be identified in both conversational discourse and writing (Edwards, 1999). In fact, they are likely more pronounced and institutionalized in writing because the latter is a cornerstone of the transmission of cultural knowledge and one of its key communicative channels.

In our study, the raters scored the use of emotional language rather than performed a semantic analysis of the emotion representations in Russian or English. Why would Russian students score lower specifically on Emotionality? Pending the detailed semantic analysis of the emotional descriptors in English and Russian used by the students in our sample, we would like to suggest that these findings might reflect the bias originating from the interaction between task demands (i.e., creative writing under a time limit) and the patterns of use of emotional language specific to each culture and language. Although both Russian and English have emotion nouns, adjectives, pseudoparticiples, and (in)transitive verbs, the distribution of emotion terms across these categories differs significantly (e.g., Russian has more intransitive verbs than English does). In addition to these distributional differences, Russian speakers display a pattern of use of emotion words that is slightly different from that of English speakers—that is, Russian speakers favor intransitive and reflexive emotion verbs, and English speakers prefer emotion adjectives; this difference has been suggested to originate from the conceptualization of emotion as a more voluntary inner activity in Russian but as more of a reactive state in English (Wierzbicka, 1998). A corpus study by Pavlenko and Driagina (2007) partially corroborated this differentiation of verbal versus adjectival pattern of use of mental lexicon. Thus, one of the possible explanations for our findings is that Russian students used less of the mental lexicon given
the limited time resources: in this case, emotion verbs would be used less frequently to save “space” for action verbs that have a greater potential for propelling the story plot, thereby reducing the overall use of emotion words.

Our tentative explanation for the pattern of results obtained in this study can be tested empirically through further detailed lexical-semantic and frequency analysis of the stories written by the students in our sample. Yet, we would also like to acknowledge that there exist several other potential sources of variance (or, for that matter, lack of invariance) that we did not examine. These include, but are not limited to, students’ gender, age, academic major and area of concentration, academic achievement, and adaptation to college life, as well as the interactions among them and between them and other unidentified factors that mediate the effects of culture on creative cognition and behavior. We also found that Russian students tended to receive higher ratings on Complexity, a measure of the elaboration and complexity of their creative writing. This is not surprising given the heavy emphasis on composition in the standard Russian language and literature curriculum but could also reflect the selective nature of the studied population (Kornilova, Kornilov, & Chumakova, 2009).

As pointed out by the anonymous reviewer, the results of this study are interpreted as indicating partial cross-cultural invariance of creativity as measured with Creative Stories. Within this approach, expert ratings are viewed as indicators of creativity, and the results of the analysis suggest that while the application of the rating scores seems to be consistent in two samples, they are partially locally biased, favoring the U.S. sample. We would like to note that such a bias could indeed reflect cultural bias in expert ratings rather than be driven by the culturally moderated properties of the product (i.e., the story) itself. Interestingly, Emotionality was the dimension of creativity for which the lowest between-country reliability estimates were obtained, suggesting that the same stories were somewhat differently rated by the Russian compared to the U.S. raters. A follow-up analysis performed in the GLM framework suggested that even excluding the linking bilingual rater and regressing out scores on three other dimensions did not eliminate the group difference in Emotionality estimates between the two samples; this suggests that the Russian students consistently received lower scores on the Emotionality item. This bias could be also driven by culturally mediated conceptions of emotion and the role of emotionality in creativity, a possibility that should be explored in future studies of implicit conceptions of creativity and creative writing in Russian.

The results obtained in this study suggest that creative writing can be at least partially considered as culturally invariant, thereby indicating the commonalities in the overall architecture of cognitive processes that contribute to creativity in the written domain. Yet, we identified a dimension of creative writing, namely, emotionality, that seems to be particularly susceptible to cultural influences, necessitating further research into the connection between the use of mental lexicon, imagination, and verbal creativity.
across cultures. Such an examination is necessary for the field to bridge the gap between its methodological and conceptual approaches to creativity. As mentioned in the beginning of this paper, culture guides the development of creativity and patterns of its manifestation; yet, it also serves as the key contextual element for the evaluation of creative products. The increasing globalization of the world poses new challenges for the field, particularly related to the issues of cultural specificity and cultural universality of creative cognition and the methods used to measure it and also to the issues of multiculturality and the blurring of cultural boundaries (for a review, see Mourguès et al., 2014) and its implications for our understanding of the very concept of culture and “cross-cultural” differences or similarities. Our study illustrates one of the possible methodological frameworks for addressing these issues by linking the properties of a measurement instrument with those of the linguistic and cultural systems that shape creativity and the way it manifests in specific creative products evaluated with that particular instrument.

Acknowledgments

This research and the preparation of this manuscript was supported by the College Board and the Russian Foundation for Humanities (RGNF), project No. 13-06-00049.

References


**Sergey A. Kornilov** is a postdoctoral associate at Yale University. He received his PhD in experimental psychology from University of Connecticut.

**Tatiana V. Kornilova** is a professor of general psychology at Moscow State University, Moscow, Russian Federation. She received her PhD and completed requirements for habilitation at Moscow State University.

**Elena L. Grigorenko** is Emily Fraser Beede professor at Yale University. She received her PhD in psychology from Moscow State University and PhD in developmental psychology and genetics from Yale University.