Thinking Skills and Creativity 18 (2015) 43–52

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: http://www.elsevier.com/locate/tsc

The dark side of creativity revisited: Is students' creativity associated with subsequent decreases in their ethical decision making?☆

Christoph Niepel[a,*], Maida Mustafić[a], Samuel Greiff[a], Richard D. Roberts[b]

[a] Education, Culture, Cognition and Society (ECCS), University of Luxembourg, Luxembourg
[b] Professional Examination Service, New York City, NY, USA

ARTICLE INFO

Article history:
Received 19 October 2014
Received in revised form 21 April 2015
Accepted 23 April 2015
Available online 1 May 2015

Keywords:
Dark side of creativity
Ethical decision making
Self-ratings
Teacher-ratings
Longitudinal data

ABSTRACT

Both creativity and ethical decision making have been identified as 21st century skills that need to be facilitated in modern educational policy and practice. Prior research on the "dark side of creativity" suggests that creativity impacts ethical decision making adversely. This study is the first to study the reciprocity of students' creativity and ethical decision making longitudinally and to specifically investigate whether students' creativity is longitudinally associated with decreases in their ethical decision making. In addition, we investigated whether the observed longitudinal relations between creativity and ethical decision making hold after controlling for students' reasoning skills. Middle school students (overall $N = 1869$; 48.3% male) were assessed at two time points of measurement. To obtain measures of students' creativity and ethical decision making, we employed self-ratings as well as teacher-ratings. Reasoning test scores were available for a subsample of 417 students. No association between creativity assessments and changes in subsequent ethical decision-making assessments was found. This resulting pattern held after controlling for students' reasoning skills. Collectively, these findings suggest that creativity is not a general predictor of decreases in ethical decision making, indicating that being a creative student unlikely implies being an unethical decision maker.

1. Introduction

The continuing importance of information technology has led to a shift from routine toward novel, dynamic, and non-routine tasks at the average workplace in the 21st century (e.g., Autor, Levy, & Murnane, 2003). These changes provide an enormous challenge for modern education. With the ever-growing need for individuals to adapt to new situations, to generate and implement novel ideas, and to make decisions in an efficient and ethical manner, the need for transversal skills that enable today's students to successfully navigate through life in the 21st century similarly increases (e.g., Casner-Lotto & Barrington, 2006; Griffin, McGaw, & Care, 2012; National Research Council, 2012). Significant progress has already been

☆ This research was supported by joint funding from the Independent School Data Exchange (INDEX) and Educational Testing Service (ETS) to Richard D. Roberts (while he was at ETS), and the Fonds National de la Recherche Luxembourg (ATTRACT "ASKI21") to Samuel Greiff. All statements expressed in this article are the authors' and do not reflect the official opinions or policies of the authors' host affiliations or any of the supporting institutions. Finally, the authors would like to thank Kevin Petway for providing the data and assisting in the preparation of this manuscript.

* Corresponding author at: ECCS, University of Luxembourg, 11, Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg. Tel.: +352 466644 6747.
E-mail address: christoph.niepel@uni.lu (C. Niepel).

http://dx.doi.org/10.1016/j.tsc.2015.04.005
1871-1871/

made in identifying these so-called 21st century skills, and both creativity and ethical decision making are seen as crucial (e.g., Binkley et al., 2012; Casner-Lotto & Barrington, 2006). Creativity has been defined as the ability to move beyond what currently exists and to generate and implement new ideas (Ward, 2004), allowing the individual to remain flexible and to become a successful problem solver (cf. Runco, 2010). Accordingly, creativity has been identified as being of eminent importance not only for economic and technical development, healthy psychological functioning, and emotional growth but also for academic success (see Plucker, Beghetto, & Dow 2004). Ethical decision making, in contrast, can be described as the ability and willingness to be moral, for example, to consider others' needs, goals, and perspectives in one's own decisions (see, e.g., Moran, 2014). Managerial decisions in the 21st century, for instance, have a potential impact on the financial and health status of millions of people and it is therefore important that they are ethical (e.g., Lawton & Paez, 2014; Trevino, de Niewenboer, & Kish-Gephart, 2014). As ethical decision making is a precondition for cooperating, working, and living together in a society, teaching ethical decision making can therefore be considered one of the important educational goals (Perri, Callanana, Rotenberry, & Oehlers, 2009).

In contrast to many other 21st century skills (including ethical decision making), various scholars question whether creativity genuinely exerts only beneficial impacts (for overviews: Cropley, Cropley, Kaufman, & Runco, 2010; Glazer, 2009; Moran, Cropley, & Kaufman, 2014). For example, one might assume that creative students more successfully find innovative ways to cheat on their exams. In fact, prior research highlighting the "dark side of creativity" (Gino and Ariely, 2012; p. 445) suggests that creativity impacts ethical behavior adversely (e.g., Beaussart, Andrew, & Kaufman, 2013; Gino and Ariely, 2012). These findings raise questions as to whether enhancing students' creativity decreases their ethical decision making over the long term. To the best of our knowledge, there is no study investigating the reciprocity of students' creativity and ethical decision making in the long run and, thus, whether creativity is actually associated with decreases in ethical decision making over time. Furthermore, educational research on this topic is meager. To address this deficiency, the present study subjected the dark side of creativity to further empirical scrutiny by examining self-assessments and teacher-assessments of creativity and ethical decision making in middle-school students using a longitudinal design.

1.1. The dark side of creativity: does creativity impact ethical decision making negatively?

Due to its obvious beneficial characteristics, creativity has been proposed as a helpful tool for mastering modern societal demands. However, creativity is also said to have a dark side. Creative people, who are more inclusive in their thinking and who use broad conceptions, have been found to be more prone to developing schizophrenia and psychotic disorders (Eysenck, 1993). Moreover, creative thinkers have been found to be less conscientious and less conventional, to tolerate ambiguity, to interpret rules leniently, to be impulsive, more dominant, and more hostile; indeed, the least favored students are perceived as being the most creative (Feist, 1998; Gino & Wiltermuth, 2014; Tegano, 1990; Westby & Dawson, 1995). In line with this, Beaussart et al. (2013) suggested that fostering creativity potentially resulted in violating social norms and expectations as it encouraged people to think in different ways than others. In fact, the authors found some cross-sectional evidence supporting their suggestions in that creative performance (i.e., assessed by an association test) was found to be negatively and significantly related to integrity assessments (i.e., self- and observer-assessed integrity; Beaussart et al., 2013).

Investigating this dark side of creativity in a series of studies, Gino and Ariely (2012) concluded that creativity increases unethical behavior by making people more cognitively flexible. In particular, the authors showed that creativity is associated with justifying immoral actions by generating reasons why immoral actions might be appropriate in a particular situation. Gino and Ariely (2012) demonstrated that creative people behaved more dishonestly and that the ability to successfully generate options to justify their immoral actions mediated the degree of dishonesty. For instance, the authors confronted participants with an ambiguous task, in which participants earned their profit only based on their responses and not based on their accuracy. Creative participants were more likely to intentionally generate responses to increase their payoff. In another problem solving task, participants had 5 min to solve 20 matrices anonymously, with 5 min not enough to solve all 20 matrices. After solving the 20 matrices, creative participants were more likely to overstate their performance to increase their payment for the task. These findings highlight a potential dark side of creativity and therewith eventually raise the question as to whether or not creativity should be promoted in students without reservation, as it may have undesired effects on students' ethical behavior. However, in Gino and Ariely's (2012) experimental studies, participants were confronted with ethical dilemmas that stimulated cheating in that participants were not only given the opportunity to behave dishonestly but were also tempted to do so (Gino and Ariely, 2012). In a situation in which cheating is a normative response, creativity may in fact be a tool that people who desire to cheat use to do so successfully. Thus, it remains an open question whether creativity really diminishes a global trait, such as ethical decision making, over time. Because this situation was largely artificial in Gino and Ariely's experiments as the authors used only tasks that enhance participants' desire to cheat, the current study investigates whether creativity is associated with subsequent decreases in students' ethical decision making in a sample of students not exposed to any experimental manipulation per se.

Another construct that should be taken into account when revisiting the dark side of creativity is intelligence. In fact, creativity and intelligence have not only been repeatedly suggested to be entangled (cf. Kim, Cramond, & VanTassel-Baska, 2010), prior research has also found a negative relation between intelligence and ethical decision making (Rayburn & Rayburn, 1996). Thus, an observed relation between creativity and ethical decision making in favor of a dark side of creativity (i.e., a negative relation between both constructs) might eventually only reflect a potential negative impact of intelligence on

ethical decision making – with intelligence as an underlying third variable that is positively related to creativity. Accordingly, Gino and Ariely (2012) have also controlled for participants' intelligence in one of their studies. In sum, they did not found intelligence to be related to creativity nor to dishonesty. After controlling for intelligence, creativity was still significantly related to dishonesty, supporting the notion of a dark side of creativity (Gino & Ariely, 2012). In the present study, we controlled for students' reasoning skills as a proxy for students' underlying general intelligence to examine the longitudinal relation between creativity assessments and subsequent ethical decision making assessments independent from students' reasoning skills.

1.2. The present study

Recent empirical findings suggest that creativity bears risks for unethical decision making and subsequent behavior (e.g., Gino & Ariely, 2012). However, longitudinal studies testing for an association between creativity and ethical decision making at school have yet to be conducted. To enlighten future research and practice, the present study aimed to test the relation between creativity and ethical decision making within the framework of a longitudinal design. In particular, the present study tackles the following research questions:

(1) Does creativity predict subsequent decreases in students' ethical decision making? Prior research suggests that students' creativity may predict decreases in their subsequent ethical decision making. In the present study, as self-ratings of personality are likely biased (e.g., Paunonen & O'Neill, 2010), we used self-ratings as well as teacher ratings as measures of creativity and ethical decision making.
(2) Are the observed longitudinal interrelations between creativity and ethical decision making independent from students' reasoning skills (i.e., testing whether the observed longitudinal interrelations were generalizable across different levels of students' reasoning skills)?

To explore the longitudinal interrelations between self-rated creativity, teacher-rated creativity, self-rated ethical decision making, and teacher-rated ethical decision making (our first research question), we applied a two-wave latent variable cross-lagged design (i.e., in specifying Model A). To examine our second research question, we additionally included students' reasoning skills (i.e., quantitative reasoning test scores) as predictor into Model A (i.e., in specifying Model B).

2. Method

2.1. Sample and procedure

Data were obtained from two measurement occasions of a larger, ongoing longitudinal study focusing on the development of noncognitive skills in students enrolled in private middle schools throughout the United States. Time 1 data was collected during the Fall Semester of 2012, and Time 2 data was collected during the Fall Semester of 2013. The sample was comprised of 1869 students (48.3% male) from 17 different U.S. middle schools in the 2012 wave and of 1143 students enrolled in 16 of these 17 schools (48.2% male) in the 2013 wave. With regard to the 2013 wave, we selected only those students who had participated in 2012. The students were rated by 245 teachers in 2012 and 181 teachers in 2013. Students completed the assessment using an online platform during normal classroom hours. School-assigned student identification numbers were used to identify students at both time points. This allowed the researchers to keep track of changes over time for each student. The average self-reported age was 12.26 years (SD = 0.94, range 10–14 years) in 2012 and 12.83 years (SD = 0.75, range 10–14 years) in 2013.

2.2. Variables and measures

2.2.1. Self- and teacher-rated creativity

The self- and teacher-rated creativity measures included six corresponding statements that utilized a four-point response scale ranging from 1 (never or rarely) to 4 (usually or always), respectively. The scale based on items obtained from the International Personality Item Pool (IPIP; ipip.ori.org). Higher scores on the response scale indicated higher levels of self- or teacher-rated creativity. The items asked students (and teachers, for students) to rate the quality of their ideas (e.g., "I generate novel ideas") or evaluate their thinking style (e.g., "I think outside the box"; see for validation studies: Petway, Rikoon, Brenneman, Burrus, & Roberts, 2015; Rikoon, 2013).

In the present study, internal consistencies for the six-item self-report scale were good ($\alpha = .85$ in 2012 and $\alpha = .83$ in 2013). The observed scale means were $M = 3.06$ (SD = 0.65) in 2012 and $M = 3.01$ (SD = 0.61) in 2013. Internal consistencies for the six-item teacher-report scale were notably higher ($\alpha = .95$ in 2012 and $\alpha = .94$ in 2013), whereas the observed scale means were lower, $M = 2.79$ (SD = 0.76) in 2012 and $M = 2.81$ (SD = 0.74) in 2013.

2.2.2. Self- and teacher-rated ethical decision making

As with creativity, the self- and teacher-rated scales for ethical decision making incorporated six corresponding statements measured on a four-point response scale ranging from 1 (never or rarely) to 4 (usually or always), respectively. All

ethical decision-making items based on items derived from the IPIP, and included statements such as "I treat all people fairly" and "I am an honest person". In general, the ethical decision-making scales focused on a student's perceived fairness, honesty, and charity (see for validation studies: Petway et al., 2015; Rikoon, 2013). In the present study, internal consistencies for the six-item self-report scale were high (α = .88 in 2012 and α = .83 in 2013). The observed scale means were M = 3.30 (SD = 0.61) in 2012 and M = 3.29 (SD = 0.54) in 2013. Internal consistencies for the six-item teacher-report scale were even higher (α = .93 in 2012 and α = .93 in 2013), a pattern consistent with the creativity scale. The observed scale means were M = 3.36 (SD = 0.65) in 2012 and M = 3.37 (SD = 0.63) in 2013.

2.2.3. Quantitative reasoning

Most of the participating private middle schools use the Comprehensive Testing Program 4 (CTP 4), an assessment developed by the Educational Records Bureau (ERB, 2012), to evaluate various academic areas in high achievement students. Quantitative reasoning is a subtest administered to students in these schools that focuses on logical, algebraic, geometrical, probabilistic, and statistical reasoning, classification, and recognition. As part of participating in the longitudinal study, schools were asked to provide these scores for students; however, this was optional, and many schools chose not to provide this information. Thus, the sample size was considerably smaller than the original sample sizes when accounting for the availability of these scores in the sample: N = 417. The quantitative reasoning scores used in the present study originate from the CTP 4 testing session conducted in Spring 2013. The average internal consistency across the three middle school grades is α = .88 according to ERB. The observed test score percentile mean was M = 85.91 (SD = 18.31).

2.3. Data analysis

We applied structural equation modeling (SEM) in the statistical software Mplus 7.11 (Muthén & Muthén, 1998–2013) to investigate the potential dark side of creativity. As some distributions of item scores differed from multivariate normality, we used the Mplus option for maximum likelihood estimation to calculate standard errors and fit statistics that are robust (MLR) against mild violations of normality (e.g., Kaplan, 2009). In all longitudinal models, we allowed the residuals of corresponding indicators to correlate across the two time points of measurement (i.e., correlated uniqueness; Little, Preacher, Selig, & Card, 2007). To assess model fit, the following criteria were used to indicate adequate fit: (a) the root mean square error of approximation (RMSEA) should be below .08, (b) the comparative fit index (CFI) should be greater than .90, and (c) the standardized root mean square residual (SRMR) should be below .08 (cf. Browne & Cudeck, 1993; Hu & Bentler, 1999; see also Kline, 2011).

2.3.1. Handling missing data

Characteristic of longitudinal research, complete data on all measures were not available across both measurement time points. In our study, we had complete data in 2012 (N = 1869) and, in 2013, we had a percentage of missing data ranging from 41.57% to 42.86% (with an average of 42.31%) in relation to our total sample of N = 1869 in 2012. With regard to students' reasoning test score (i.e., implemented in Model B), we had a percentage of missing data ranging from 77.69% to 79.13% (with an average of 78.26%) in relation to our total sample of N = 1869, as data was only available for a subsample of 417 students. By applying full information maximum likelihood (FIML), the MLR estimator allowed us to deal efficiently with data that were missing in this way (e.g., Kaplan, 2009), thus yielding (a) higher statistical power to study our research questions (Collins, Schafer, & Kam, 2001) and (b) less biased parameter estimates than traditional pairwise or listwise deletion methods (e.g., Graham, 2009).

2.3.2. Handling the nested data structure

Typical for educational research, our examined sample featured a nested data structure; that is, students were nested within different teachers. We therefore had to control for possible confounds of individual student characteristics with teacher characteristics. Consequently, we mean-centered all items (except for the normed reasoning test score) within the examined teachers for both time points to remove clustering effects, resulting in a fixed-effects design (see Enders & Tofighi, 2007; Legewie, 2012).

3. Results

3.1. Preliminary analyses

To isolate potential local misfit (Tomarken & Waller, 2003), we first specified eight separate measurement models that included the measures of self-rated creativity, teacher-rated creativity, self-rated ethical decision making, and teacher-rated ethical decision making at each of the two measurement points. Results indicated that in each measurement model, three residual terms shared common variance over and above their respective creativity or ethical decision-making factors due to similar item wordings (i.e., for creativity, three residual terms of the items including the word "ideas" [self- and teacher-rated items] and, for ethical decision making, three residual terms of items including the word "others" [self- and teacher-rated

Table 1
Results of a longitudinal confirmatory factor analysis: correlations between self- and teacher-rated creativity (i.e., Latent Measures), self- and teacher-rated ethical decision making (i.e., Latent Measures), and quantitative reasoning (i.e., Manifest Measure).

Constructs	Correlations							
	1	2	3	4	5	6	7	8
1. Creativity 2012 (Self)	–							
2. Creativity 2012 (Teacher)	.22**	–						
3. Ethical decision making 2012 (Self)	.55**	.15**	–					
4. Ethical decision making 2012 (Teacher)	.06*	.29**	.28**	–				
5. Creativity 2013 (Self)	.57**	.13**	.28**	−.02	–			
6. Creativity 2013 (Teacher)	.15**	.36**	.07†	.11**	.19**	–		
7. Ethical decision making 2013 (Self)	.33**	.11**	.59**	.24**	.52**	.04	–	
8. Ethical decision making 2013 (Teacher)	.09*	.13**	.22**	.44**	.06†	.31**	.30**	–
9.Quantitative Reasoning	.09	.30**	−.03	.10*	.08	.31**	.01	.14*

Note: Creativity (Self) = self-rated creativity; Creativity (Teacher) = teacher-rated creativity; Ethical decision making (Self) = self-rated ethical decision making; Ethical decision making (Teacher) = teacher-rated ethical decision making.
† $p < .10$.
* $p < .05$.
** $p < .01$.

items]). We therefore allowed these residual terms to correlate in each measurement model in our subsequent analyses (cf. Kline, 2011).

3.2. Confirmatory factor analyses

As a second step, we inspected the latent structure of each of our latent measures (i.e., self-rated creativity, teacher-rated creativity, self-rated ethical decision making, and teacher-rated ethical decision making). To this end, we conducted two confirmatory factor analyses (CFAs) including all latent constructs per time of measurement (i.e., 2012 and 2013). The two CFAs resulted in acceptable fit statistics (CFI = .96, RMSEA = .05, 90% CI [.04, .05], SRMR = .04, for 2012; CFI = .96, RMSEA = .05, 90% CI [.04, .05], SRMR = .04, for 2013), indicating that at each time point of measurement, self-rated creativity, teacher-rated creativity, self-rated ethical decision making, and teacher-rated ethical decision making compose distinct latent factors. The respective items can be considered good measures of all factors (standardized factor loadings ranged between .51 and .92).

To provide additional insight in our data, we further calculated a longitudinal CFA including all creativity and ethical decision-making assessments in 2012 and 2013 as well as students' reasoning test score (i.e., manifest measure) to obtain model-based correlations between all measures included in our analyses. The longitudinal CFA resulted in acceptable fit statistics (CFI = .96, RMSEA = .03, 90% CI [.03, .03], SRMR = .04); Table 1 depicts correlations between all measures. Notably, concerning the sixteen cross-construct correlations between creativity and ethical decision-making assessments, twelve were significantly positive (with $p < .05$), two marginally positively significant (with $p < .10$), and two failed to reach statistical significance; no correlation was found to be significantly negative. Students' reasoning test scores showed significant and positive associations with teacher-ratings of creativity and ethical decision making but not with correspondent self-ratings of both constructs (i.e., self-rated creativity and self-rated ethical decision making).

3.3. Measurement invariance tests

Third, we analyzed the measurement invariance of the examined latent measures across time to test whether they captured the same target constructs across measurement points. To examine our two research questions, at least metric invariance (i.e., invariance of the pattern of zero and nonzero loadings and of factor loadings across time) was required, as this would allow us to interpret and to identify correlations among latent constructs across time. To investigate measurement invariance, we first tested for configural invariance per construct separately (i.e., invariance of the pattern of zero and nonzero loadings). Then, we tested a baseline model (i.e., imposed configural invariance for all four constructs simultaneously) against a metric invariant model (i.e., test of invariance of the pattern of zero and nonzero loadings and of factor loadings for all four constructs simultaneously). Metric invariance was established if the metric invariant model showed similar fit indices as the baseline model (cf. Christ & Schlüter, 2012; for further information on measurement invariance see, e.g., Widaman & Reise, 1997). To compare the baseline and metric invariant model, we conducted a chi-square difference test using the Satorra–Bentler scaled chi-square (Satorra & Bentler, 2001). Table 2 documents the results of tests of measurement invariance for all four latent constructs from 2012 to 2013. An examination of model fit indices and the resulting nonsignificant difference test statistic between the baseline model and the metric invariant model suggested that the metric invariant model provided a reasonable approximation to the data. Thus, our subsequent analyses were conducted on models that imposed metric invariance.

Table 2
Results of measurement invariance tests for self- and teacher-rated creativity and self- and teacher-rated ethical decision making across two points of measurement.

Invariance	*N*	χ^2 (df)	CFI	RMSEA (90% CI)	SRMR	$\Delta\chi^2$ (df)
Configural creativity (Self)	1869	412.221 (41)**	.940	.070 (.064–.076)	.041	–
Configural creativity (Teacher)	1869	250.499 (41)**	.983	.052 (.046–.059)	.046	–
Configural Ethical decision making (Self)	1869	259.030 (41)**	.964	.053 (.047–.060)	.034	–
Configural Ethical decision making (Teacher)	1869	225.822 (41)**	.982	.049 (.043–.055)	.024	–
Baseline (Simultaneous)	1869	2868.686 (1004)**	.958	.032 (.030–.033)	.039	–
Metric	1869	2894.629 (1024)**	.958	.031 (.030–.033)	.039	26.670 (20), $p = .145$

Note: Creativity (Self) = self-rated creativity; Creativity (Teacher) = teacher-rated creativity; Ethical decision making (Self) = self-rated ethical decision making; Ethical decision making (Teacher) = teacher-rated ethical decision making.
** $p < .01$.

3.4. Main analyses: examining the dark side of creativity

Finally, we tackled our two research questions by examining the longitudinal relationship between self- and teacher-rated creativity assessments and changes in subsequent self- and teacher-rated ethical decision-making assessments. For this purpose, we specified a two-wave latent variable cross-lagged model (Little et al., 2007) in which all autoregressive and cross-lagged paths between all four latent constructs (i.e., self-rated creativity, teacher-rated creativity, self-rated ethical decision making, and teacher-rated ethical decision making) were simultaneously modeled across both measurement time points (i.e., Model A, see Fig. 1). By controlling for all measures in 2012, we tested whether changes in any of these constructs in 2013 (i.e., in self-rated and/or teacher-rated ethical decision making) could be explained by the characteristics of a different construct in 2012 (i.e., of the prior self-rated and/or teacher-rated creativity; see Finkel, 1995). As a second step, we specified Model B (see Fig. 2) by including students' reasoning test scores as predictor into Model A. In both models, we tested the dark side of creativity by examining the longitudinal relationship between prior creativity assessments and changes in subsequent ethical decision-making assessments (i.e., cross-construct relations). In Model B, we additionally examined

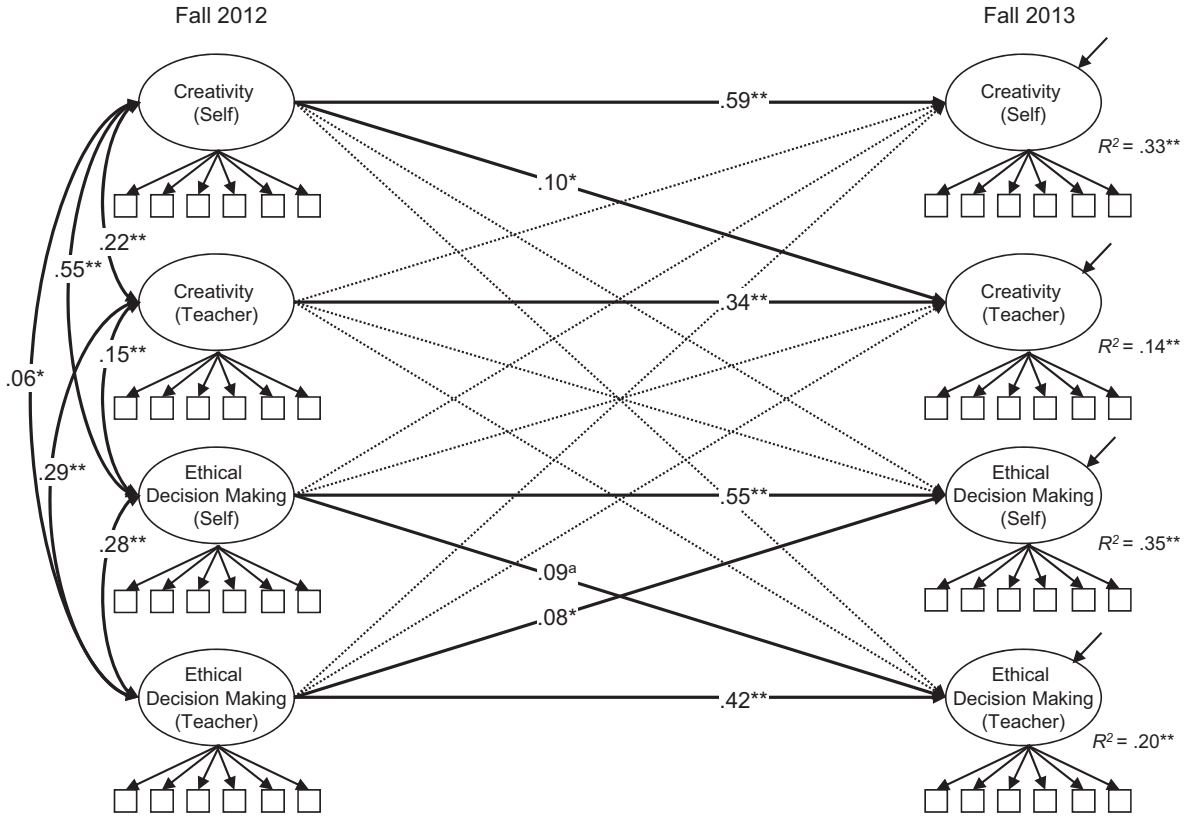

Fig. 1. Model A: Two-wave cross-lagged model between self-rated creativity (Creativity [Self]), teacher-rated creativity (Creativity [Teacher]), self-rated ethical decision making (Ethical Decision Making [Self]), and teacher-rated ethical decision making (Ethical Decision Making [Teacher]). For the sake of clarity, correlational paths between residuals were omitted from the path diagram and only standardized parameter estimates with $p < .10$ are shown. Positive associations are depicted by continuous lines, nonsignificant with $p > .10$ by dotted lines. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{a}p = .050$.

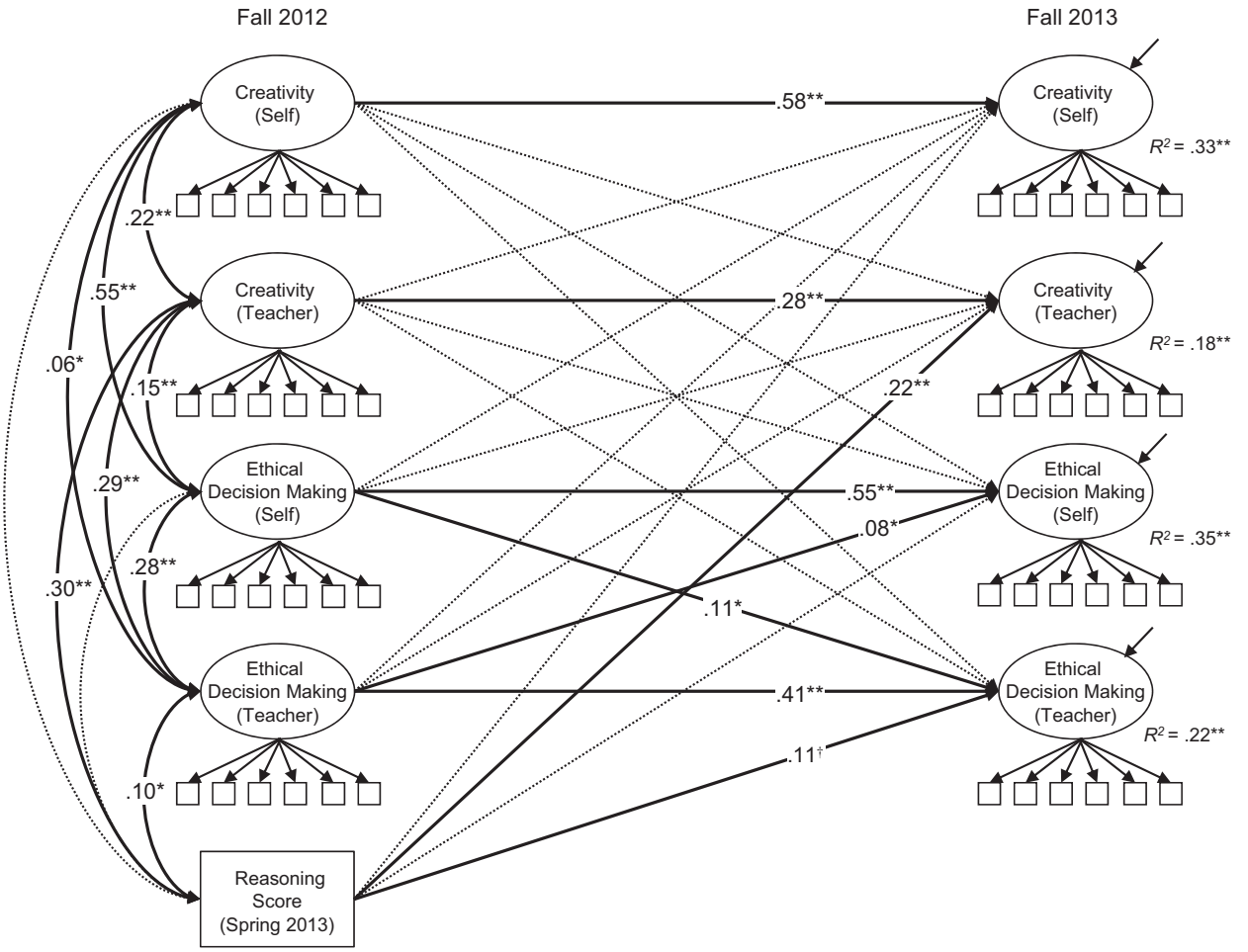

Fig. 2. Model B: Two-wave cross-lagged model between self-rated creativity (Creativity [Self]), teacher-rated creativity (Creativity [Teacher]), self-rated ethical decision making (Ethical Decision Making [Self]), and teacher-rated ethical decision making (Ethical Decision Making [Teacher]) while controlling for students' reasoning test scores. For the sake of clarity, correlational paths between residuals were omitted from the path diagram and only standardized parameter estimates with $p < .10$ are shown. Positive associations are depicted by continuous lines, negative by dashed lines, and nonsignificant with $p > .10$ by dotted lines. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$.

whether the longitudinal interplay between creativity and ethical decision making is generalizable across different levels of students' reasoning ability. Both models featured acceptable fit indices (CFI = .96, RMSEA = .03, 90% CI [.03, .03], SRMR = .04, for Model A; CFI = .96, RMSEA = .03, 90% CI [.03, .03], SRMR = .04, for Model B).

3.4.1. Cross-construct relations: testing the dark side of creativity

Without controlling for students' reasoning scores (Model A), prior creativity assessments (i.e., self-rated and teacher-rated) were not significantly related to changes in subsequent ethical decision-making (i.e., self-rated and teacher-rated) assessments. This pattern held after controlling for students' reasoning scores (Model B); we thus did not find any evidence for the dark side of creativity.

3.4.2. Further findings

3.4.2.1. Cross-construct relations: ethical decision making and changes in subsequent creativity. For the cross-construct relations between prior ethical decision making and changes in subsequent creativity, a similar pattern emerged in Models A and B in that no significant relations were found, indicating a null cross-lagged relation.

3.4.2.2. Creativity: within-construct relations between self-rated and teacher-rated measures. Without controlling for students' reasoning scores (Model A), prior self-rated creativity showed a positive association with changes in subsequent teacher-rated creativity. After controlling for students' reasoning scores (Model B), we did not find any significant relation between self-rated and teacher-rated creativity measures.

3.4.2.3. Ethical decision making: within-construct relations between self-rated and teacher-rated measures. In Models A and B, we found self-rated and teacher-rated ethical decision making to be positively and reciprocally related across time. Prior self-rated ethical decision making was significantly and positively related to changes in subsequent teacher-rated ethical decision making. Prior teacher-rated ethical decision making showed a positive and significant relationship to changes in subsequent self-rated ethical decision making.

3.4.2.4. Autoregressive paths. In both models, all autoregressive paths were positive and significant, indicating moderate to high temporal stabilities from 2012 to 2013. The absolute values were found to be slightly higher for self-rated measures than for teacher-rated measures.

3.4.2.5. Reasoning as predictor in Model B. In Model B, students' reasoning skills were found to be positively and significantly associated to subsequent teacher-rated creativity and ethical decision making but not to the respective self-rated measures.[1]

4. Discussion

The current study revisited the 'dark side of creativity' longitudinally, by examining whether self- and teacher-assessments of students' creativity are related to subsequent decreases in self- and teacher-assessments of ethical decision making. Taken together, we did not find any evidence for the dark side of creativity. This finding could be replicated after controlling for students' differences in their reasoning test scores. Moreover, our results showed that self- and teacher ratings of creativity and ethical decision making were concurrently positively correlated.

4.1. Strengths and limitations of the present study

To our knowledge, the present study was the first conceptualized to longitudinally examine the dark side of creativity in a large educational sample. As prior research has already gathered some evidence in support for a dark side of creativity (e.g., in using experimental designs), longitudinal research is a necessary complement to obtain a more comprehensive understanding of the relationship between creativity and ethical decision making. Furthermore, to get a more holistic picture of the longitudinal relation between creativity and ethical decision making, we implemented self-rated and teacher-rated measures of creativity and ethical decision making into our study. Although the observed self-rated and teacher-rated measures of creativity as well as the self-rated and teacher-rated measures of ethical decision making showed substantial unique variances and rather small shared variances across methods (i.e., self- and teacher ratings), the observed cross-sectional correlational pattern among all latent measures, however, supported their convergent validity. Overall, we found the concurrent latent correlations among self- and teacher ratings to range from .19 to .30. To classify these results accordingly, note that correlations among self- and observer ratings for Big Five variables were found—in meta-analytic research—to range from .46 to .62 (Connolly, Kavanagh, & Viswesvaran, 2007). However, as self-rated measures are likely to be influenced by participants' self-concepts and as teacher-rated measures may be biased to some extent (compare, e.g., halo effect; Thorndike, 1920), future research should incorporate the advantages of creativity tests (e.g., divergent thinking tests) and/or behavioral data (e.g., creative outcomes as expert-rated essays or ways of solving ill-defined complex problems) to measure students' creativity. Similarly, ethical decision making could be assessed by behavioral data in future research (e.g., number of observed dishonest behaviors or incidents of cheating as a proxy for students' ethics).

Additionally, in investigating U.S. middle school students enrolled in private schools, we examined a very selective sample. Our results need therefore to be replicated in differing samples, for example, different types of schools (e.g., public schools), different age groups (e.g., children in kindergarten), and other educational systems (outside the United States).

To control for potential clustering effects, we mean-centered all variables (except the normed reasoning test scores) within teachers. Mean-centering may have led to an underestimation of the underlying effects between the observed variables as the variances of the variables were slightly restricted (Legewie, 2012). Thus, the observed relations can be considered as conservative estimates of the true relations between creativity and ethical decision making.

Furthermore, as we had only access to the quantitative reasoning test scores of 417 students, the amount of missing data was considerably higher in Model B. However, the results of Model B based on FIML (as presented in the article) and on listwise deletion (without any missing values) were comparable, illustrating the overall robustness of our results.

4.2. No support for a dark side of creativity

Although previous research has shown that creativity affects unethical behavior (i.e., the dark side of creativity; see Section 1), our longitudinal results do not support the hypothesis that creativity is a general predictor of ethical decision

[1] To test the robustness of our results, the analyses of Model B were performed a second time using the listwise deletion technique without any missing values. Virtually no differences emerged between Model B as presented in the article (using FIML; see Results section) and Model B using listwise deletion (with $N = 356$; CFI = .94, RMSEA = .04, 90% CI [.04, .05], SRMR = .05); only the path between teacher-rated ethical decision making and subsequent self-rated ethical decision making failed to reach statistical significance.

making. We found that prior creativity assessments (based on self- and teacher ratings) did not predict any changes in students' subsequent ethical decision-making assessments (based on self- and teacher ratings). This resulting patterns was generalizable across different levels of reasoning skills: When controlling for differences in students' reasoning skills, the resulting pattern remains similar indicating no negative associations between creativity assessments and subsequent ethical decision-making assessments. Notably, the arguably strongest evidence in favor for a dark side of creativity to date derived from experimental studies using tasks that stimulated cheating (Gino & Ariely, 2012; see Section 1). Therefore, one may assume that creativity may in fact facilitate dishonest behavior in ethical dilemma situations, where dishonest behavior might indeed be appropriate sometimes (cf. Kohlberg dilemmas; e.g., stealing medicine to save somebody's life; Kohlberg, 1981). However, the results of the present study contradict the notion that creativity genuinely impacts ethical behavior adversely. To sum up, creativity may thus be sometimes misused as a tool to behave dishonestly in certain settings, though, the fact of being creative unlikely implies being an unethical decision maker per se. Furthermore, as only the fact of perceiving themself as creative (i.e., creative self-concept) has been discussed to affect unethical behavior (as creative people tend to develop an exaggerated sense of entitlement; Vincent & Goncalo, 2014), our study provide strong evidence that students who perceive themselves as more creative are concurrently and subsequently not perceived by their teachers as more unethical (and do not perceive themselves as more unethical).

4.3. Further findings of the present study

We consistently found in both models that self-rated and teacher-rated ethical decision making were positively and reciprocally related across time. Accordingly, students who describe themselves as ethical may more likely show ethical behaviors in classroom that are observed by teachers; students who are perceived as ethical by their teachers may describe themselves as more ethical across time.

Furthermore, we found quantitative reasoning to be positively related to teacher-ratings of creativity and ethical decision making. Thus, students with stronger reasoning skills are perceived as more ethical and more creative by their teachers, which may be tentatively explained by potential observer biases (compare, e.g., halo effect; Thorndike, 1920) as well as by underlying shared variances between creativity and reasoning (cf. Kim et al., 2010).

Additionally, we did not find any association between ethical decision-making assessments (self- and teacher-rated) and subsequent creativity assessments (self- and teacher-rated). This result is of eminent interest as Gino and Wiltermuth (2014), in a way, recently extended the dark side of creativity hypothesis in suggesting that acting dishonestly leads to greater creativity (see Gino & Wiltermuth 2014, for more information). In fact, the authors found some first experimental evidence for their hypothesis. However, the results of the present study disagree with Gino and Wiltermuth's (2014) conclusions and, having the limitations of the present study in mind (as discussed earlier), highlight the need for further research on this topic.

4.4. Conclusion

Previous literature suggests that creativity has a negative influence on ethical behavior in that it helps people to generate innovative ways to justify their immoral actions (e.g., Gino & Ariely, 2012). The results of the current study provide some first longitudinal evidence that students' self- and teacher-perceived creativity does not have undesirable effects on students' self- and teacher-perceived ethical decision making across time. Thus, we did not find any evidence for a dark side of creativity in that creativity impacts ethical decision making adversely. As creativity and ethical decision making can be considered as key 21st century skills (e.g., Binkley et al., 2012), the need to foster both in contemporary education has been repeatedly articulated (for creativity: see, e.g., Plucker et al., 2004; for ethical decision making: see, e.g., Perri et al., 2009). Importantly, the results of the present study suggest that the goals of simultaneously fostering students' creativity and ethical decision making are compatible with each other. However, in acknowledging the need for further research, we hope that our findings will serve as a vehicle for continued empirical attempts on the relation between creativity and ethical decision making.

References

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, *118*, 1279–1333. http://dx.doi.org/10.1162/003355303322552801

Beaussart, M., Andrews, C., & Kaufman, J. (2013). Creative liars: The relationship between creativity and integrity. *Thinking Skills and Creativity*, 129–134. http://dx.doi.org/10.1016/j.tsc.2012.10.003

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills*. Dordrecht, New York: Springer.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century US workforce*. Washington, DC: Partnership for 21st Century Skills.

Christ, O., & Schlüter, E. (2012). *Strukturgleichungsmodelle mit Mplus: Eine praktische Einführung [Structural equation modeling with Mplus: An introduction]*. München, Germany: Oldenbourg Verlag.

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*, 330–351. http://dx.doi.org/10.1037/1082-989X.6.4.330

Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, *15*, 110–117. http://dx.doi.org/10.1111/j.1468-2389.2007.00371.x
Cropley, D. H., Cropley, D. H., Cropley, A. J., Kaufman, J. C., & Runco, M. A. (Eds.). (2010). *The dark side of creativity*. Cambridge University Press.
Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121–138. http://dx.doi.org/10.1037/1082-989X.12.2.121
ERB. (2012). *Comprehensive testing program 4. Technical report*. New York, NY: Educational Records Bureau.
Eysenck, H.-J. (1993). Creativity and personality: Suggestions for a theory. *Psychological Inquiry*, *4*, 147–178. http://dx.doi.org/10.1207/s15327965pli0403_1
Feist, G. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, *2*, 290–309. http://dx.doi.org/10.1207/s15327957pspr0204_5
Finkel, S. E. (1995). *Causal analysis with panel data. Sage university papers series on quantitative applications in the social sciences*. Thousand Oaks, CA: Sage.
Gino, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, *102*, 445–459. http://dx.doi.org/10.1037/a0026406
Gino, F., & Wiltermuth, S. (2014). Evil genius? How dishonesty can lead to greater creativity. *Psychological Science*, *25*, 973–981. http://dx.doi.org/10.1177/0956797614520714
Glazer, E. (2009). Rephrasing the madness and creativity debate: What is the nature of the creativity construct? *Personality and Individual Differences*, *46*, 755–764. http://dx.doi.org/10.1016/j.paid.2009.01.021
Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549–576. http://dx.doi.org/10.1146/annurev.psych.58.110405.085530
Griffin, P. E., McGaw, B., & Care, E. (Eds.). (2012). *Assessment and teaching of 21st century skills*. Dordrecht, New York: Springer.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. http://dx.doi.org/10.1080/10705519909540118
Kaplan, D. (2009). *Structural equation modeling: Foundations and extensions*. Los Angeles, CA: Sage.
Kim, K. H., Cramond, B., & VanTassel-Baska, J. (2010). The relationship between creativity and intelligence. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 48–73). Cambridge, New York: Cambridge University Press.
Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.
Kohlberg, L. (1981). *Essays on moral development. The philosophy of moral development* (vol. I) San Francisco, CA: Harper & Row.
Lawton, A., & Paez, I. (2014). Developing a framework for ethical leadership. *Journal of Business Ethics*, http://dx.doi.org/10.1007/s10551-014-2244-2
Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule [The estimation of causal effects: An introduction to methods of causal inference based on peer effects in education]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *64*, 123–153. http://dx.doi.org/10.1007/s11577-012-0158-5
Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, *31*, 357–365. http://dx.doi.org/10.1177/0165025407077757
Moran, S. (2014). Introduction: The crossroads of creativity and ethics. In S. Moran, D. Cropley, & J. C. Kaufman (Eds.), *The ethics of creativity* (pp. 1–22). Houndsmill, UK: Palgrave Macmillan.
Moran, S., Cropley, D., & Kaufman, J. C. (Eds.). (2014). *The ethics of creativity*. Houndsmill, UK: Palgrave Macmillan.
Muthén, L. K., & Muthén, B. O. (1998–2013). *Mplus version 7.11 [statistical software]*. Los Angeles, CA: Muthén & Muthén.
National Research Council. (2012). Education for life and work. Developing transferable knowledge and skills in the 21st century. In J. W. Pellegrino, & M. L. Hilton (Eds.), *Committee on defining deeper learning and 21st century skills*. Washington, DC: The National Academies Press.
Paunonen, S. V., & O'Neill, T. A. (2010). Self-reports, peer ratings and construct validity. *European Journal of Personality*, *24*, 189–206. http://dx.doi.org/10.1002/per.751
Perri, D., Callanan, G., Rotenberry, P., & Oehlers, P. (2009). Education and training in ethical decision making: Comparing context and orientation. *Education & Training*, *51*, 70–83. http://dx.doi.org/10.1108/00400910910931841
Petway, K. T., II, Rikoon, S., Brenneman, M. W., Burrus, J., & Roberts, R. D. (2015). *Development and validation of the Mission Skills Assessment. [Technical report]*. Report submitted for publication.
Plucker, J. A., Beghetto, R., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, *39*, 83–96. http://dx.doi.org/10.1207/s1532698Sep3902_1
Rayburn, J. M., & Rayburn, L. G. (1996). Relationship between Machiavellianism and type A personality and ethical-orientation. *Journal of Business Ethics*, *15*, 1209–1219. http://dx.doi.org/10.1007/BF00412819
Rikoon, S. H. (2013). *Toward an omnibus assessment of noncognitive skills: A longitudinal multitrait-multimethod application in middle schools (Doctoral dissertation)*. Retrieved from ProQuest Dissertations & Theses Full Text (Order No. 3594846).
Runco, M. A. (2010). Divergent thinking, creativity, and ideation. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 413–446). Cambridge, New York: Cambridge University Press.
Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. http://dx.doi.org/10.1007/BF02296192
Tegano, D. (1990). Relationship of tolerance of ambiguity and playfulness to creativity. *Psychological Reports*, *66*, 1047–1056. http://dx.doi.org/10.2466/PR0.66.3.1047-1056
Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, *4*, 25–29. http://dx.doi.org/10.1037/h0071663
Tomarken, A. J., & Waller, N. G. (2003). Potential problems with "well fitting" models. *Journal of Abnormal Psychology*, *112*, 578–598. http://dx.doi.org/10.1037/0021-843X.112.4.578
Trevino, L., den Nieuwenboer, N., & Kish-Gephart, J. (2014). (Un)ethical behavior in organizations. *Annual Review of Psychology*, *65*, 635–660. http://dx.doi.org/10.1146/annurev-psych-113011-143745
Vincent, L. C., & Goncalo, J. A. (2014). License to steal: How creative identity entitles dishonesty. In S. Moran, D. Cropley, & J. C. Kaufman (Eds.), *The ethics of creativity* (pp. 137–151). Houndsmill, UK: Palgrave Macmillan.
Ward, T. B. (2004). Cognition, creativity, and entrepreneurship. *Journal of Business Venturing*, *19*, 173–188. http://dx.doi.org/10.1016/S0883-9026(03)00005-3
Westby, E., & Dawson, V. (1995). Creativity: Asset or burden in the classroom? *Creativity Research Journal*, *8*, 1–10. http://dx.doi.org/10.1207/s15326934crj0801_1
Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). http://dx.doi.org/10.1037/10222-009