THE EFFECTS OF EXPLICIT INSTRUCTIONS ON THE PERFORMANCE OF ESSAY-WRITING TASK

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Arts

by

Qiaochu Ren

August 2019

ABSTRACT

This study examined the effects of instructions with emphasize to be creative, be analytical, be practical or with no emphasize on the creativity, analytical rigor, and practical implications of solutions produced in an essay-writing task. The sample consisted of 137 students at Cornell University (US). Consistent with our expectations, the "be creative" instruction resulted in a higher creative rating over the standard instruction with no specific focus, while the "be analytical" instruction facilitated analytical performance. In contrast, the practical instruction seemed to depress the subjects' practical performance. The analysis also revealed that instructions with a different focus affected students' performance asymmetrically for different genders and educational major groups. Implications of the findings are discussed.

BIOGRAPHICAL SKETCH

Born in China, Qiaochu Ren developed a curious mind in individuals' decision-making process and had a passion for incorporating different fields, including economics, psychology, and education. She has explored the relationship between citizens' sense of justice and their asset allocation during her study at Wuhan University. After she received her bachelor's degree in finance, she continues her studies in the field of developmental psychology at Cornell University. Having realized that creativity is an essential factor for academic success and could be facilitated or hindered under different conditions, she conducted in-depth research on the impact of "be creative" instruction on students' creative performance with professor Robert Sternberg. Besides, she employed her skills in data visualization to develop a dashboard for instructors and students using data from a business intelligence course. In the future, she hopes to dedicate her life to enhance the decision making of students and realize their full potential.

ACKNOWLEDGMENTS

The whole process of completing this thesis is hard but rewarding. I am eternally grateful to my parents, who tightened their belts to finance my study in the U.S and stood by me during all my ups and downs. None of this would be possible without their support.

To my advisor, professor Robert Sternberg, who gave me this meaningful topic and the confidence to independently conduct a psychological experiment where I have limited experience. He was always a responsible mentor, encouraging me to take the initiative and choose my pathway where I could dedicate my life.

A special thanks to my beloved, Shawn, who was the most patient listener that I can turn to whenever I was facing struggles and difficulties. Thank you for working behind the scenes, enabling me to be myself and bringing all those little happiness into my life. To my senior schoolmate Nan Zhou: for giving me the courage to follow my heart and fighting with me along the way.

Finally, to my four lovely research assistants, Ru Feng, Caiwei Zhu, Sierra Forester, Madeline Hanscom, who have been an indispensable part of my getting there. Thank you for helping me transcribe and grade all those responses conscientiously without asking for return.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH iii
ACKNOWLEDGEMENTS iv
Introduction 1
Method 8
Participants 8
Procedures and Instruction 9
Measures 11
Results 12
Reliability of Measures 12
Main Effect of Instructions on Creative Performance 13
Instructional Effect on Creative Performance by Gender 15
Instructional Effect on Creative Performance by Major 16
Instructional Effect on Analytical Performance 17
Instructional Effect on Practical Performance 18
Discussion 18
REFERENCES 22
APPENDIX 27

Introduction

Throughout history, creativity has served as a driving force in the evolution of society, in terms of both technology and the economy. Thus, creativity has been seen as an important element of success, especially in academic settings. Creative students are more likely to achieve better academic performance in school (Grigorenko, Jarvin, Diffley, Goodyear, Shanahan, & Sternberg, 2009; Sternberg & the Rainbow Project Collaborators, 2006). In addition to aiding the production of higher standardized test scores, creativity is also a vital factor for college success (Sternberg, 2010; Sternberg, Bonney, Gabora, & Merrifield, 2012). One valuable question in education is whether creativity performance can be enhanced. One school of scholars argue that creativity is a personality trait and cognitive ability that will be relatively stable across time (e.g., Runco, 2007). Others view being creative as a behavior, an attitude, or a decision that can be modified through formal schooling and life experience (e.g., Sternberg, 2012). Fortunately, there is a large body of empirical findings results suggesting that creativity is a set of processes, skills, and strategies, and that creative performance can be hindered or facilitated under various circumstances. For example, Sawyer's eight-stage model of explaining creativity demonstrates that creative performance involves the process of knowledge collection, incubation, idea formation, evaluation, and externalization. Sternberg and Lubart (1995), in their investment theory, suggest that six aspects of creativity—intellectual ability, knowledge, styles of thinking, personality, motivation, and environment—can interact with each other to promote a higher level of creativity. They also suggest that having strength in one component may or may not compensate for the inferiority of another element of creativity.

Sufficient attempts have been made to demonstrate that every individual has the potential to think creatively in a particular domain or context (e.g., Amabile, 1996). This serves as an

underlying assumption for the idea that creative thinking can be taught or influenced by a variety of practices. For instance, Brophy (2006) reports that individuals working in a collaborative environment are more likely to think creatively. Conforming to the belief that creativity is pliable and creative performance can be fostered, domain-specific techniques targeted at facilitating creativity have been explored in numerous disciplines. Dugosh and Paulus (2005) found that subjects who were exposed to plentiful ideas were capable of generating additional ideas on their own. Shalley (1991) revealed that there was a significant main effect of setting a creativity goal for problem-solving tasks. The study provides support for the priming effect of creativity goals, which involves directing the individuals' attention toward being creative. Numerous programs succeed in boosting creativity with a focus on the development of domain-specific cognitive skills and heuristics involved in skill application.

Concerning the limitations on the feasibility and generalizability of those techniques, a plethora of empirical studies have been carried out to examine the impact of extrinsic motivation on creative performance. They have focused especially on the effect of rewards. However, research evidence regarding the impact of compensation failed to reach a consensus. One group of researchers who espouse the over-justification hypothesis argues that external incentives have a detrimental effect on creative performance by decreasing intrinsic interest. In particular, Amabile (1983) states that people are less likely to be creative if they do not enjoy the work they are doing or focus on the potential rewards instead of doing the work for its own sake. In contrast, another camp of scholars with a behavioral perspective has found otherwise. They suggest that creative behavior can be promoted with extrinsic rewards without undermining intrinsic motivation when the requirement of creative performance is specified clearly

(Eisenberger & Armeli 1997), or the reward is given under proper conditions (Eisenberger & Cameron, 1998; Eisenberger & Sekbst, 1994; Glover & Gary, 1976).

Despite the abundance of literature reviewing the impact of rewards, a majority of the research examines the effect of explicit instruction to "be creative" on the creative output of a heterogeneous sample, especially when using divergent thinking tests. Again, these investigations have resulted in contradictory evidence as to the effect of instructions to "be creative" on creativity outcomes. Chistensen, Guilford, and Wilson (1957) revealed that subjects produced more clever responses under instructions to be clever than those with no specific instructions. Similarly, Manske and Davis (1968) demonstrated that using the Guilford Unusual Uses Test, the responses generated by subjects being told to "be original" were significantly more original than answers produced by those who received neutral or practical instructions. Moreover, researchers found that visual art students enhanced the fluency, flexibility, and originality of their performance after receiving information about knowledge and skills involved in the creative process regulation and strategies for divergent thinking (Kamp, Admiraal, Drie, & Rijlaardam, 2015).

The literature has reviewed the effects of specific instructions on subjects along with other factors, including initial creative level, gender, and culture. Datta (1963), using the Lowenfeld Mosaic Test and two paper-and-pencil problem-solving tasks, found that scientists who fell in the upper third of the creativity score range as rated by their supervisors and were told to "be as creative as you can" had the most creative performance. Harrington (1975) discovered that male participants who were told to "be creative" generated significantly more creative responses in the Alternative Uses Test than those in the standard condition. Congruent with Datta's results, when using the Adjective Check List as a measure of the Creative

Personality Scale, subjects in the upper third of the scale had the best performance when told to "be creative." The performance of subjects in the lower third of the scale was depressed when placed in the creative condition. In contrast, Oziel, Oziel, and Cohen (1972) report that subjects who were predefined as initially highly creative using the Revised Art Scale of the Welsh Figure Preference Test did not significantly improve their performance on the scale when assigned the creative instructions.

Studies on the facilitation effect of instruction using other tests have emerged in recent years. For example, O'Hara and Sternberg (2001) used an essay-writing task to investigate the impact of four different instructions, three of which emphasized either being creative, being practical, or being analytical. The control group received no specific emphasis. Researchers found that the "be creative" instruction resulted in a significant positive effect on the creativity ratings compared with other groups. Moreover, Chen et al. (2002) found that drawings made under the "be creative" instruction had a higher score on creativity than those made with no explicit request to be creative, and the main effect of the instruction was universal among American and Chinese participants.

The literature has also revealed that the magnitude of the impact of instructions also varied across different tasks. A study by Chand and Runco (1992), for example, used a problem generation task to test for the effect of explicit instruction. The tasks not only required the subjects to list as many problems they have in school but also asked them to select one problem and generate solutions for it. The researchers suggested that the explicit instructions resulted in a significant effect on ideational performance. Research conducted by Chen (2005) demonstrates that explicit instruction to be creative showed a domain-specific facilitation effect across different cultural and ethnic groups. In particular, artistic creativity and mathematical creativity

were shown to have a greater magnitude of improvement, which benefited more from the instruction than verbal creativity. Furthermore, Runco (2005) found that responses under the creative instructions scored higher in their originality in the unrealistic divergent thinking task compared with those in the realistic divergent thinking task.

Niu and Liu (2009) also investigated the effect of instruction with a specific focus on the collage-making task and the story completion task. The study reports that students' creative performance was facilitated under detailed instructions which contained a task-specific method and a strategy as to how to be creative. However, the study failed to find the impact of creative instruction, which simply told the participants to be creative. This suggests that mere instruction was not enough to effectively enhance creative performance.

Some literature has examined gender differences in response to creativity instructions. Katz and Poag (1979), for instance, have extended Harrington's study by taking into account sex difference in "creative style" by including female subjects. They reported that instructions to list creative uses that are both unusual and worthwhile, enhance the performance on both the Alternative Uses Test and Things Category Tests for male subjects but not for females. Other studies showed a gender difference in creative performance under several external restraints. For example, Baer (1997, 1998) suggests that the expectation of evaluation, while remarkably decreasing the creative performance of middle-school girls in a collage-making activity, did not impact boys' creative performance. The researcher also reports that female subjects suffered a decline in their creative performance when being offered a reward (bonus points for their art grade), whereas male subjects did not.

Existing research evidence has limits as to its interpretability and generalizability of the facilitation effect of explicit instruction. Traditional tasks such as collage making, drawing, and

divergent thinking tests are not necessarily relevant to real-world problems. According to Okuda, Runco, and Berger (1991), the performance of a real-world problem-finding task has more strength in predicting creative behavior in real-world activities compared with traditional divergent thinking measures. In this study, we used an essay-writing task which was similar to the scientific inquiry process, starting with problem generating and hypothesis formulation and continuing on to experiment design and imagining of possible results.

Even though a few studies have tried to incorporate problems relevant to life and work, researchers presented the problem directly to the participants instead of requesting them to raise a question by themselves. Since problem finding is also an indispensable process in the creative thinking, the present study integrated not only problem solving but also problem finding to elicit products that are more reliable and more predictive for real-world activities, at the meanwhile allowing for a wide variation in analytical thinking, creative thinking, and practical thinking. The current project requires subjects to firstly identify a phenomenon from their daily lives, which has the potential to lead to an analytical explanation, a practical solution, or a creative explication. They then need to adopt a way to obtain solutions to the problem. If students write more creatively, more analytically, or more practically when given different instructions, then their thinking mode might be different when completing a cognitive task. In other words, their tendency toward being creative, analytical, or practical might be hindered in an environment that favors one kind of instructional set or another. Specifically, schools' emphasis on analytical thinking might inadvertently inhibit students from thinking creatively.

Furthermore, a focus on explaining human behavior instead of other scientific phenomena would be appropriate in this research since participants have many opportunities, in their daily lives, to come into contact with examples of human behaviors that puzzle them. In

other words, research participants can choose topics that they have more knowledge of, which could have the potential to enlarge the instructional effect.

The primary goal of this study was to obtain further insight into the effect of instructions with different foci including "be creative," "be practical," and "be analytical" on three rated qualities (creativity, analytical rigor, and practical implication). Instead of using the controversial technique of rewarding creativity, our research alters a simple selection of words to make the subjects aware of the researchers' interest. The present study directed participants' attention to particular facets by gently introducing information about the researcher's interest ("We are most interested in the creativity/analytical rigor/practical implication of what you do") in the experiment groups. This avoids the potential dysfunctional consequence of reward or expected evaluation on intrinsic motivation (Amabile, 1979; Shalley & Oldham, 1985).

We used the following combinations for the research design: a 4 (way of instruction: creative-focus, practical-focus, analytical-focus, and no focus) $\times$ 2 (majors: Humanities, Art & Social Science and Natural Science & Engineering) mixed design and a 4 (instruction) $\times$ 2 (gender: male and female) mixed design. The study also controls for the confounding factor of age and education level since these may influence students' habitual thinking style. Therefore, we conducted data analysis using a 4 $\times$ 2 mixed design method.

There are three main hypotheses. First of all, we expected that instructions with a specific focus would enhance students' performance on the corresponding quality. Therefore, responses generated by students who received the instruction to be creative would score higher in overall creativity compared with responses generated under other conditions. Similarly, those produced by subjects under the analytical-focus instruction would have a higher rating on the analytical dimension, and students who were told to be practical would receive a higher grade on the

practical quality of their products. Based on the first hypothesis, it was also expected that students under the analytical instruction would have the most similar analytical rating to those in the standard instruction group because of the tacit emphasis on analytical thinking in school. The second hypothesis involved gender differences in the influence of explicit instructions. Research investigating gender difference in instruction shows that the effect of instructions on male subjects is more salient than for female subjects. Hence, it is hypothesized that there would be a significant interaction effect between instructions with a specific focus and gender. Last but not least, since according to investment theory creators need to obtain task-related knowledge in order to take a further step based on what already exists in the domain, the familiarity of the task might influence creative output. We assumed that those with humanities, art, or social science related majors would have more knowledge about the task, and thus we further hypothesized that the creative-oriented instruction would have a more substantial impact on them compared with students who are natural science or engineering majors.

Method

Participants

A total of 137 students at Cornell University participated in the experiment voluntarily in exchange for two credits for courses eligible in the SONA system. The average age was 20.24 years (S.D. = 1.53), and 67% were female students. Students came from diverse ethnic backgrounds: 39 (29%) of European American descent; 36 (27%) Asian American descent; 12 (9%) Asian; 8 (6%) African American descent; and 24 (18%) of other ethnic backgrounds (including mixed descent). This experiment was entitled “Explaining Human Behavior” and was

offered in the Spring semester in 2019 at Cornell University. Five participants were excluded due to missing data and incomplete responses, resulting in 132 products being analyzed.

Procedure and instruction

The Institutional Review Board's approval was obtained prior to any data collection for this study. The entire time spent by the participants was approximately one hour. Subjects were randomly assigned to one of the four groups. After being told to read and sign a consent form, all the participants were given basic instructions for a writing task with forty minutes as a time limit. The basic instructions were:

Think of some phenomenon that puzzles you about human behavior—any phenomenon at all. What is your hypothesis as to why people behave this way? How would you test your hypothesis through a scientific experiment? What set of results would be consistent with your hypothesis?

After the standard instructions, the creative instruction group (Condition C) was informed that researchers were most interested in the creativity of what they do and to be as creative as possible. Subjects in the analytical instruction group (Condition A) were told that researchers were most interested in their analytical rigor. Under the practical condition (Condition P), participants received the instruction that the researchers were most interested in the practical implications of what they do and to be as practical as possible. Under the standard condition (Condition S), no specific instructions indicated the researchers' interest. After they completed the task, participants turned in their essays to the experimenter, who then presented them with a questionnaire to collect their necessary demographic information including their gender, age,

major, education level, and ethnicity, as well as asking about their attitude toward being creative, being practical, or being analytical in their lives. Finally, they were debriefed and signed the debriefing form. In all, thirty-four participants were assigned to the creative instruction group, thirty-three were in the practical instruction group, thirty-one were in the analytical instruction group, and thirty-four were in the standard instruction group.

Table 1

Distribution of the Sample for Each Condition (Gender, Major, Age, Education Level)

		Instruction				
		Creative	Practical	Analytical	Standard	Total
Gender	Male	13	9	9	13	44
	Female	21	24	22	21	88
Majors	Humanities, Art & Social science	27	25	20	27	89
	Natural Science & Engineering	7	8	21	7	43
Age	Mean	20.24	20.36	20.35	20.03	20.24
	SD	1.39	1.85	1.52	1.34	1.52
Education level	Freshman	7	5	4	10	26
	Sophomore	9	14	11	8	42
	Junior	10	6	8	10	34
	Senior	8	7	7	5	27
	Graduate	0	1	1	1	3

Measures

Altogether, three primary measures were used, each containing several sub-components. The first dependent variable is creativity quality. Based on the widely accepted criteria for judging creativity, which include the dimension of novelty and usefulness (Sternberg & Lubart, 1999), judges were required to rate the novelty and usefulness of the responses on a 7-point scale. We describe the usefulness of the responses as having foreseeable value in future research and should not be deemed as inappropriate (illegal or unethical) by other people. In addition to these two aspects, we also included the third aspect of creativity elaboration to evaluate the quality of creative ideas. If the responses produced novel ideas but were too vague to contain any details, then they would have a low score on "creativity elaboration," which is on a 3-point scale. In the end, we used the sum of the scores of the three components as the creative measure.

The second dependent variable is analytical quality. In particular, three components in this measure, including the structure and organization, analytical elaboration, and logical quality, were rated for this measure. Structure and organization, rated on a 4-point scale, evaluated the connection between the phenomenon, the hypothesis, and the experiment, as well as the possible results supporting the hypothesis. The analytical elaboration (5-point scale) determined the elaboration of responses, especially the details of the experiment and its strength and weakness. Logical quality (7-point scale), which is an essential component in analytical quality, reflected the analytical strength. A response could contain plenty of detailed information but would not be able to stand up under closer scrutiny. Similarly, we summed up the score of each element in the analytical facet as the analytical quality rating.

The last measure is practical quality. Specifically, we took into consideration the practical implications (7-point scale) and the practicality of the response (5-point scale). On the one hand,

the practical implications represent the attempt of the responses to address the real-world problem and its potential impact. On the other hand, the practicality of the responses measures the feasibility of the experiment. Again, the practical rating is the total score of the two facets.

All the products were judged by four volunteer undergraduate research assistants and the principal researcher. They were given the description of the task, grading criteria containing the definition, and the rating scale of each dimension. Undergraduate judges received basic training before they graded the responses. In order to minimize estimation biases, all the judges were told to independently evaluate each response according to the dimensions mentioned above without knowing the instructions that each participant received. Raters were also encouraged to share their opinions about a particular response on the shared document containing each product so that less experienced raters might be more objective when judging products' creative quality.

Results

Reliability of Measures

A Cronbach alpha was calculated for each performance quality in order to test the inter-rater reliability. The reliability coefficient is 0.60 for the creative performance, 0.76 for the analytical performance, and 0.66 for the practical performance. Since the reliability coefficient for the creative performance and the practical performance below the fairly acceptable level according to Cicchetti and Sparrow (1998), we take the rater as a random factor. We also conducted an additional analysis using only the score from the goal rater, which may be more reliable and avoid any possible results discovered unintentionally resulting from a misunderstanding of the grading rubric. Although the halo effect was likely to exist when each judge needed to rate all the dimensions of each response, the judges were able to clearly

distinguish the three performance types. The correlation between the creativity rating and the practical rating is 0.38, whereas the correlation between the creativity rating and the analytical rating is 0.67. The practical rating correlated 0.48 with the analytical rating. The difference between the correlation coefficients, 0.48 and 0.67 was significant using Steiger Tests, $z = -2.61$, $p = 0.009$.

Main Effect of Instructions on Creative Performance

In this 4×2 mixed design, a marginal facilitation effect of instruction of "be creative" on creative performance, $F(3, 637) = 2.495$, $p = 0.059$ was observed. Overall, subjects' creative rating is the highest under the creative instruction ($M = 11.14$, $SD = 0.72$) compared with other groups. The difference in the creative performance between the creative group and the standard group is statistically significant to a 90% confidence level (mean difference = 0.80, $p = 0.08$), whereas the difference in the creative performance between those under creative instruction and those under practical instruction or analytical instruction is not significant. When we analyze the data with the analysis of covariance method using the score from only the most experienced judge, the facilitation effect of the instruction to "be creative" compared with the standard group is more salient (mean difference = 2.04, $p = 0.05$). Supplementary analysis also showed that students who were told to "be creative" were more likely to generate responses that are novel compared with other students ($F(3,637) = 2.91$, $p = 0.044$), especially students with no specific instruction (mean difference = 0.5, $p = 0.039$). Students under the "be creative" instruction also had the highest performance in evaluating their creative idea to be useful and task appropriate, but the difference is not statistically significant.

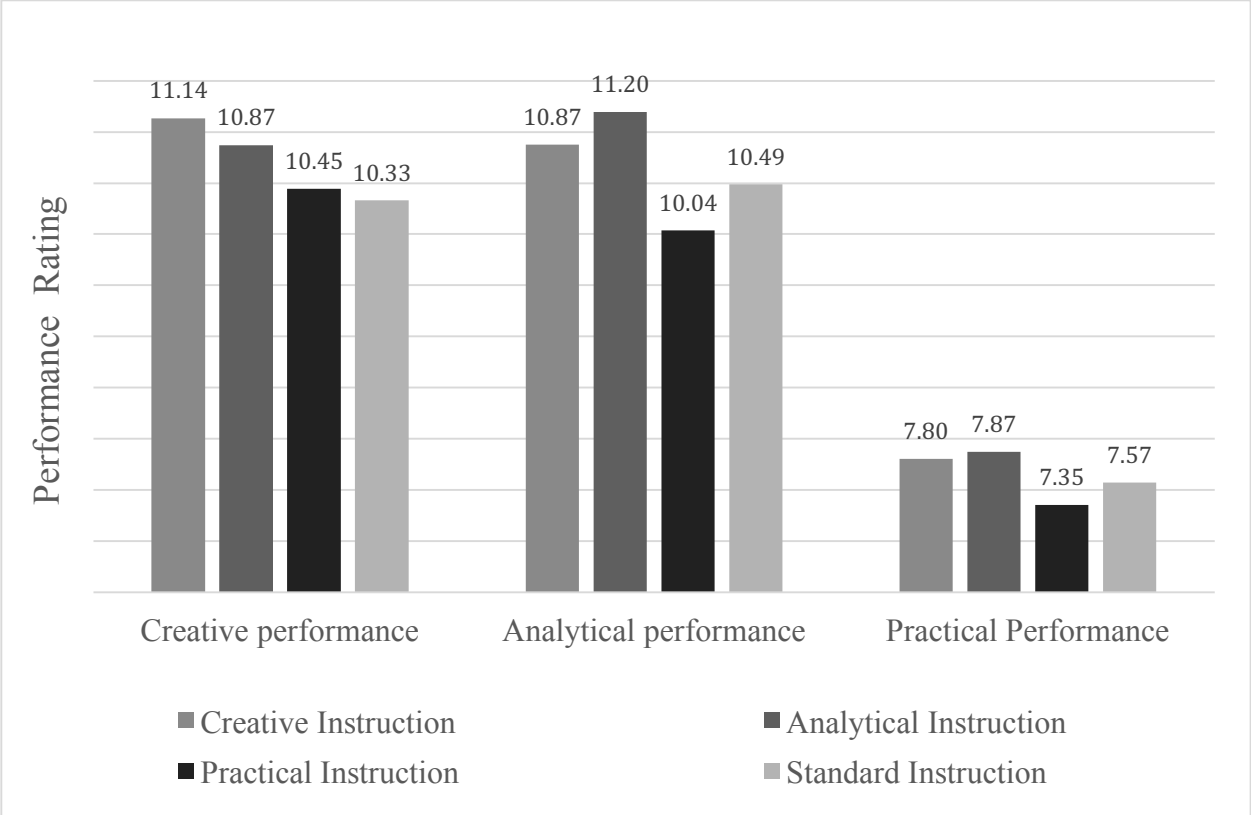

Figure 1. Creative, analytical, practical performance ratings by way of the instructions to "be creative," "be analytical," and "be practical."

Instructional Effect on Creative Performance by Gender

Table 2

Mean Score of Creative Performance and Standard Deviations as a Function of Instruction Condition and Gender

	Gender			
	Male		Female	
Condition	Mean	SD	Mean	SD
Creativity	11.51	0.70	10.76	0.70
Practical	10.27	0.70	10.62	0.70
Analytical	10.92	0.70	10.82	0.69
Control	10.29	0.70	10.38	0.70

The data did not yield a significant main effect of gender, nor was an interaction between instruction type and gender observed. Nevertheless, the separate analysis revealed that the instructions influenced the creative output of the male subjects differently, in line with our second hypothesis. For male subjects, those who were under the creative instruction scored higher in the creative rating compared with those under the standard instruction (mean difference = 1.22, $p = 0.048$), and others under the practical instruction (mean difference = 1.24, $p = 0.069$). For female students, however, no difference in the creative rating among the four instruction groups was found to be significant.

Instructional Effect on Creative Performance by Major

Table 3

Mean Score of Creative Performance and Standard Deviations as a Function of Instruction Condition and Majors

	Major			
	Humanities, Art & Social Science		Natural Science & Engineering	
Condition	Mean	SD	Mean	SD
Creative	11.28	0.71	10.99	0.80
Practical	10.18	0.74	10.71	0.80
Analytical	10.58	0.72	11.19	0.77
Standard	10.14	0.71	10.53	0.80

There is no main effect of major type on creative performance ($F(1,637) = 1.67$, $p = 0.20$) and no significant interaction effect between instruction and major type, either. We thus make a further investigation into the instruction effect on each major. For subjects who majored in the humanities, art, or social science, those under the "be creative" condition have a higher creative rating in comparison with those under standard instruction (mean difference = 1.14, $p = 0.0015$), and with those under "be practical" condition (mean difference = 1.10, $p = 0.019$). In other words, those students with humanities, art, and social science-related majors, and who were told to "be creative," outperformed in creative quality compared with their counterparts under practical instructions or the standard instruction. However, there was no significant mean difference in creative rating for those with natural science or engineering-related majors, which

supports our third hypothesis that the effect of instructions would be more salient for students who are more familiar with the task.

Instructional Effect on Analytical Performance

Consistent with our first hypothesis, students' analytical performance benefited from the "be analytical" instruction ($F(3,637) = 5.168$, $p = 0.0016$). The most notable differences were observed between the subjects under the analytical instruction and the practical instruction (mean difference = 1.16, $p = 0.0012$). In addition to our hypothesis, the data also showed that there was no significant difference between students being told to "be analytical" and students with the instructions to "be creative." One possible explanation could be that the rating of creativity contained the element of creative elaboration, which is highly correlated with the analytical elaboration in the analytical rating ($r = 0.8$). Nevertheless, when we conducted a separate analysis using only the score of logical quality, subjects told to "be creative" still had a similar outcome to the subjects under the analytical instruction (mean difference = 0.04, $p = 0.99$). When we took a further step into examining the gender difference, results showed that females had a stronger reaction to the "be analytical" instruction. For the female participants, analytical ratings were the closest between those under the "be analytical" instruction and the standard instruction, indicating that the female students tended to be analytical even under no special instructions.

Another main effect on analytical performance was the major type ($F(3,637) = 5.538$, $p = 0.0189$). The analytical performance of students with humanities-related majors was inferior in comparison to those with a natural science or engineering major. For the humanities students, the "be analytical" instruction boosted analytical performance compared with their counterparts in the standard instruction group (mean difference = 0.82, $p = 0.05$), whereas for the students with a

natural science or engineering major, the most salient effect of analytical instruction was observed when compared with the practical group (mean difference = 1.96, p = 0.005).

Instructional Effect on Practical Performance

Unlike the positive instructional effect of "be creative" and "be analytical," students in the practical group did not benefit from the "be practical" instruction. Instead, they seemed to react against the instruction, which was the case for both males and females. The practical instruction, while having a mild but not statistically significant positive effect on practical performance for those with humanities, art or social science majors, did not have a facilitation effect on students with a natural science or engineering major. For students with a natural science or engineering major, those who were told to "be practical" had a lower practical performance compared with those who were under "be creative" instruction (mean difference = 1.18, p = 0.03), and with those who were under "be analytical" instruction (mean difference = 1.23, p = 0.01), which are both contradictory to our expectation.

Discussion

The most essential take-home message for the educator is that prompting students to be creative, even without teaching them how to be creative or listing synonyms for being creative, can nudge them to think more creatively compared with the standard condition with no focus. Compared with other techniques that are time-consuming and require substantial training and effort for teachers, setting a creativity goal for students is sufficient to make a difference even in an education system that favors analytical thinking over other modes. Similarly, when we

tailored the instruction to focus on analytical rigor, students' analytical performance was enhanced compared with that of students given practical instruction.

In this study, contrary to the research evidence from O'Hara and Sternberg (2001), the instruction to "be practical" had a dysfunctional effect on practical performance, which was consistent even using the rating of the expert rater. One possible explanation might be that when the participants were confused about whether to spend more time thinking about the practical implication or the feasibility of the product, they just tried to stick with their regular cognitive mode. Moreover, participants and raters might differ in their opinions for the practical implication of a specific topic, leading to the unexpected lower practical performance for those given practical instruction.

The results showed that the "be creative" instruction has facilitated the creative performance for male participants but did not do so for female participants in line with our second hypothesis. Previous study from Kogan (1974) suggested that subjects tend to rely on their habitual thinking mode when the task requirement is confusing. Instructions, on the other hand, have the power of reducing the ambiguity of the situation. Female students, even under control condition with no specific focus, tend to produce task-appropriate response. In other word, the "be creative" instruction, while prompting the male participants to generate task-appropriate response, did not have much effect on female participants. Thus, only the male subjects had an increase in their creative performance when they were told to be creative.

Moreover, in this human behavior-explaining task, the creative instruction was stronger for those with humanities-related majors. One explanation for this difference, in line with investment theory, is that knowledge plays a critical part in creative performance. Although subjects who had more conceived task-related knowledge seemed to have more restrictions on

writing creatively, they had an advantage over novices in that they could view the phenomenon with a unique perspective without replicating others' findings. However, on the one hand, the disproportionately large number of students with natural science or engineering major in the analytical group might lead to biased results. On the other hand, the present study is not able to rule out the possibility that such difference is a result of the different habitual thinking styles of students, which might be an influential covariate affecting not only what type of major students choose, but also their performance in the task. According to Sternberg's mental self-government theory (O'Hara and Sternberg, 2001), there are three types of thinking styles: legislative style, executive style, and judicial style. Those with a legislative thinking style prefer to think more creatively, whereas people with a judicial thinking style have a predilection to think analytically. Students who are used to thinking creatively might react more to the "be creative" instruction but fail to have an adequate arena when told to "be analytical." Thus, future research could include the thinking style of being creative, being practical, or being analytical when investigating the interaction of instruction and major type using a similar experimental design task. Another expansion of this study could be to use more major-relevant topic. For instance, we could ask students with a physics major to think about any physics phenomenon that puzzle them and design an experiment to test their hypothesis.

In the present study, the inter-rater reliability did not reach the expected level, partly due to raters' lack of training before the grading. We failed to gather all of the judges together, grading the first few responses and explore more on the grading criteria and the definition of the measures, which are very helpful for them to reach a higher level of consensus. Thus, for the future research, we could use more experienced judges and provide sufficient training for them before they evaluate the responses.

In a nutshell, it is crucial to bear in mind that creative performance could be enhanced if it is appropriately encouraged. Even when no evaluation is expected, and no reward is directly linked to creative performance, students still made up their mind to employ their creative thinking when they were told to be creative. Educators have the responsibility to create an environment where not only logical thinking is emphasized. Otherwise, students will lock up their creative thinking, struggling so hard to fit themselves in an analytical system, and failing to realize their full potential.

REFERENCES

AMABILE, T. M. (1983). The Social Psychology of Creativity: A Componential Conceptualization. *Journal of Personality and Social Psychology*, 45(2), 357–377.

AMABILE, T. M. (1990). Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology*, 37, 221–233.

AMABILE, T. M. (1996). *Creativity in context*: *Update to the social psychology of creativity*. Boulder, CO: Westview Press.

ANNA ABRAHAM. (2015). Gender and creativity: An overview of psychological and neuroscientific literature. *Brain Imaging and Behavior.* DOI 10.1007/s11682-015-9410-8

BAER, J. (1997). Gender differences in the effects of anticipated evaluation on creativity. *Creativity Research Journal*, 10, 25–31.

BAER, J. (1998). Gender differences in the effects of extrinsic motivation on creativity. *Journal of Creative Behavior*, 32, 18–37.

BROPHY, D. R. (2006). A comparison of individual and group efforts to creatively solve contrasting types of problems. *Creativity Research Journal, 18*(3), 293–315.

CHAND, I., & RUNCO, M. A., (1992). Problem finding skills as components in the creative process. *Personality and Individual Differences, 14*(1), 155–162

CICCHETTI, D. V., & SPARROW, S. S. (1981). Development of criteria for establishing the interrater reliability of specific items in a given inventory: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, 86, 127–137.

DATTA, L. E (1963). Test instructions and identification of creative scientific talent. *Psychological Reports*, 13, 495-500.

DUGOSH, K. L., & Paulus, P. B. (2005). Cognitive and social comparison processes in brainstorming. *Journal of Experimental Social Psychology, 41*(3), 313–320.

EISENBERGER, R., & CAMERON, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist*, 51, 1153–1166.

EISENBERGER, R., & ARMELI, S. (1997). Can salient reward increase creative performance without reducing intrinsic creative interest? *Journal of Personality and Social Psychology*, 72, 652–663.

EISENBERGER, R., ARMELI, S., & PRETZ, J. (1998). Can the promise of reward increase creativity? *Journal of Personality and Social Psychology*, 74, 702–714.

EISENBERGER, R., & CAMERON, J. (1998). Reward, intrinsic interest, and creativity: New findings. *American Psychologist*, 53, 676–679.

EISENBERGER, R., & SHANOCK, L. (2003). Rewards, intrinsic motivation, and creativity: A case study of conceptual and methodological isolation. *Creativity Research Journal*, 15, 121–130.

GRIGORENKO, E. L., JARVIN, L., DIFFLEY, R. III, GOODYEAR, J., SHANAHAN, E. J., & STERNBERG, R. J. (2009). Are SSATS and GPA enough? A theory-based approach to predicting academic success in secondary school. *Journal of Educational Psychology, 101*(4), 964-981. http://dx.doi.org/10.1037/a0015906.

HARRINGTON, D.M (1975). Effects of explicit instructions to "be creative" on the psychological meaning of divergent thinking test scores. *Journal of Personality*, 43, 434 454.

KAMP, M., ADMIRAAL, W., DRIE, J., AND RIJLAARSDAM, G. (2015), Enhancing divergent thinking in visual arts education: Effects of explicit instruction of meta cognition. *Br Journal of Education Psychology*, 85, 47-58.

KATZ, A. N., & POAG, J. R. (1979). Sex differences in instructions to "be creative" on divergent and non-divergent test scores. *Journal of Personality*, 47, 518–530.

KOGAN, N. (1974). Creativity and sex differences. *The Journal of Creative Behavior*, 8(1), 1 14. http://dx.doi.org/10.1002/j.2162-6057.1974.tb01103.x.

MANSKE, M.E., & DAVIS, G.A. (1968). Effects of simple instructional biases upon performance in the unusual uses test. *Journal of General Psychology*, 79, 25–33.

MUHAMMAD ABDUR RAHMAN MALIK & ARIF N. BUTT. (2017). Rewards and creativity: Past, present, and future. *Applied Psychology,* 66, 290-325.

O'HARA, L. A., & STERNBERG, R. J. (2000-2001). It doesn't hurt to ask: Effects of instructions to be creative, practical, or analytical on essay writing performance and their interaction with students' thinking style. *Creativity Research Journal,* 13:197–210.

OKUDA, S. M., RUNCO, M. A., & BERGER. D. E. (1991). Creativity and the finding and solving of real-world problems. *Journal of Psychoeducational Assessment*, 9, 45–53.

OZIEL, L. J., OZIEL, L. D., & COHEN, R. H. (1972). Effects of instructional set on production of creative responses. *Psychological Reports*, 31, 93–94.

RUNCO, M. A., & OKUDA, S. M. (1991). The instructional enhancement of the flexibility and originality scores of divergent thinking tests. *Applied Cognitive Psychology*, 5, 435–441.

RUNCO, M. A., ILLIES, J. J., & EISENMAN, R. (2005). Creativity, originality, and appropriateness: What do explicit instructions tell us about their relationship? *Journal of Creative Behavior*, 39, 137–148.

RUNCO, M. (2007). Correcting the research on creativity. *Creativity Research Journal*, 19(4), 321-327. http://dx.doi.org/10.1080/10400410701756724

SAWYER, R. K. (2012). *Explaining creativity: The science of human innovation.* New York, NY: Oxford University Press.

SHALLEY, C. E., & OLDHAM, G. R. (1985). Effects of goal difficulty and expected evaluation on intrinsic motivation: A laboratory study. *Academy of Management Journal*, 28, 628–640.

SHALLEY, C. E. (1991). Effects of productivity goals, creativity goals, and personal discretion on individual creativity. *Journal of Applied Psychology,* 76, 179–185.

SHALLEY, C. E. (1995). Effects of coaction, expected evaluation, and goal setting on creativity and productivity. *Academy of Management Journal,* 38, 483–503.

STERNBERG, R. J., & LUBART, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York, NY: Free Press.

STERNBERG, R. J. (2002). Creativity as a Decision. *American Psychologist*, 57, 376–376.

STERNBERG, R. J., & The Rainbow Project Collaborators (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical and creative skills. *Intelligence*, 34, 321–350.

STERNBERG, R. J. (2010). *College admissions for the 21st century*. Cambridge, MA: Harvard University Press.

STERNBERG, R.J., BONNEY, C. R., GABORA, L., & MERRIFIELD, M. (2012). WICS: A Model for College and University Admissions, *Educational Psychologist,* 47:1, 30–41, DOI: 10.1080/00461520.2011.638882

STERNBERG, R. J., & STERNBERG, K. (2017). Measuring scientific reasoning for graduate admissions in psychology and related disciplines. *Journal of Intelligence*, http://www.mdpi.com/2079-3200/5/3/29/pdf.

STERNBERG, R. J. (2017). Creativity from start to finish: A “straight-A” model of creative process and its relation to intelligence. *The Journal of Creativity Behavior*. doi:10.1002/jocb.223.

STERNBERG, R. J. (2018). Yes, creativity can predict academic success! *Creativity. Theories – Research – Applications*, 5, 142–145.

APPENDIX

Table 1

Rubric for Rating Responses to the Experiment Design for Creative Quality

Value	Description
Understanding of the task	
0	The answer is completely wrong or off-topic. For example, the responses are not about the phenomenon of human behavior but focus on the traits, unobservable emotions or cognitive process of human being.
1	The answer is focused on the phenomenon of human behavior, which is defined as the response of individuals or groups of humans to internal and external stimuli. It refers to the array of physical actions and observable emotions associated with individuals, as well as the human race in general.
Novelty	
0~1	The response does not show any evidence at all of being creative. The hypothesis and experiment are very similar to those of other studies.
2~4	The answer demonstrates a minimal attempt to be creative. There are one or two innovative ideas in the response but in general, the answer is ordinary.
5~7	The response, in general, is very creative with several surprising ideas. For

example, the phenomenon may be common but the participant has thought about it in a different and unusual way. The hypothesis and the experiment are both innovative. It might be drawn from other studies but the participant has improved the experiment in a distinctive way. Common methods used in the experiment are combined in an unexpected way. The possible results given by the participants are imaginative.

Usefulness / Appropriateness

0~7 **Description of the highest point**:

The response does not simply generate an unexpected solution to test the hypothesis of human behavior. It may have foreseeable value to research, policy design, and society. It may be helpful in provoking new thoughts, facilitate the understanding of long-standing human inquiries and improve decision making. The solutions should not be deemed as inappropriate (illegal or unethical) by other people.

Creativity elaboration

0~3 **Description of the lowest point**:

The response contains little information. The response might come up with a creative experiment but does not explain it in enough detail.

Description of the highest point:

The response dwells on details and may have more than one hypothesis. Also, the participant may have tested each of them accordingly. Several possible results

are listed which are imaginative.

Table 2

Rubric for Rating Responses to the Experiment Design for Analytical Quality

Value	Description
Structure and connection	
0	The rationale of the response is not clear. The hypothesis is not relevant to the phenomenon. The experiment given is completely unrelated to what is needed to test the hypothesis.
1~2	The phenomenon, the hypothesis, the experiment, and the results are basically related to each other.
3~4	The answer is well-organized with the phenomenon, the hypothesis, the experiment, and the results tightly related to each other. The experiment is organized step by step and is explained clearly.
Analytical elaboration	
0~1	The response is too vague and contains little information. For example, the answer does not mention information such as the method involved in the study, how the experiment will be implemented, who will take part in the experiment, or what they will be asked to do. Or, the response has missing parts. For example, the participant didn't illustrate how the experiment is to

	be conducted, or what the possible results could be.
2~3	Though the response might have a missing part, it is elaborated in some detail. It contains information regarding the subjects to be investigated, the method to measure the dependent variable and the independent variables, and also ways to control for other covariates.
4~5	The participant may have discussed the strengths and weaknesses of the experiment and also listed possible alternative explanations for the results. The response illustrates clearly the dependent variables, independent variables, and different treatment groups, and may also take into consideration several control variables and compound effects.

Logical quality and analytical strength

0	The idea may appear to have been memorized from other studies, indicating that the participant did not think through the hypothesis, experiment, and the results.
1~4	The response demonstrates a basic logical thinking process but may still contain some logical errors. For example, the experiment might not be able to measure the psychological construct accurately and the results of the experiment might not be consistent over time. There might be bias in the experiment. Overall, the answer arrives at a satisfactory result.
5~7	The response exhibits sophisticated and thorough reasoning processes. The

	response has formed a clear interpretation of the phenomenon and may explain why the participant has come up with the hypothesis. He/she offers in-depth ideas for the hypothesis and the experiment as well as how the result may or may not support the hypothesis. The experimenter can accurately observe the behavior or any psychological constructs that are involved without inducing bias in the experiment. Other researchers could repeat the study and get similar results.

Table 3

Rubric for Rating Responses to the Experiment Design for Practical Quality

Practical implications of the response	
0~7	**Description for the highest point**: The response may be influential for policy making, practice, theory, and other subsequent research. The participant has made a clear effort to solve realistic problem in real life that could make society better off. Results from the sample population could be generalized to other populations.
The difficulty of conducting the experiment:	
0~5	**Description of the lowest point**: **Participant recruitment and data collection**: The study may have difficulty in reaching out to potential subjects; people may be unwilling to take part in the study,

or the researcher for the experiment may have difficulty in obtaining access to the secondary data.

Budget constraints: The implementation of the experiment may require a large amount of money.

Time and place constraints: It may take a long time to conduct the experiment; the procedure may be too complex and time-consuming, such that the participants may drop out; participants may become bored and spend less effort to complete the experiment in the end. The experiment might need to take place in a specific place.

Ethics: The study may contain several risks that could affect the physical/psychological health of human participants.

It is difficult to observe the behavior or measure the psychological construct accurately.
