THE INSTABILITY OF CREATIVITY OVER WEEKS THROUGH THE LENS OF DISTRIBUTIONAL MODELS

Atefeh Katrahmani[1], Alberta A. Ansah[2], Nabil Ch[2], Julia Burmeister[3], Andrew L. Kun[2], Caitlin Mills[2], Orit Shaer[3], John D. Lee[1]
[1]University of Wisconsin - Madison, Madison, WI
[2]University of New Hampshire, Durham, NH
[3]Wellesley College, Wellesley, MA

Remote work presents a challenge to workers' creativity, especially during the COVID-19 pandemic and the stay-at-home requirements. Individual differences in creativity, considered through the lens of distributional models, and their stability across different conditions are unknown. We assess the between-person variability in common metrics of creativity, despite sharing similar experiences of virtual reality and mindfulness. The paper also assesses the stability of an individual's creativity over time. We measured the creativity of 20 remote-workers daily, during a 9-week study. Creativity was measured with respect to divergent thinking and convergent thinking. Distributional models show significant individual differences in variability of creativity. Stability analyses also revealed that individuals' creativity is relatively unstable over time— both within and across conditions. Although one measure of divergent creativity was relatively stable, the other was not. We suggest more research should assess the extent of variability in creativity relative to individual differences and under different conditions.

INTRODUCTION

Remote work due to the COVID-19 pandemic has forced a physical separation between work teams thereby affecting team performance. Developing new ideas and novel procedures or products is essential for competitive working environments and creativity is proposed to be an essential part of team performance (Carroll, 1968). Understanding and supporting creativity is thus crucial for effective work from home. However, creativity, especially over time, is poorly understood. Creativity is typically defined as generating, creating, and discovering novel ideas or solutions (Rozenes & Cohen, 2017). Research has indicated that novel and diverse experiences enhance creativity (Peppercorn, 2020), but others have argued that it is a stable personality trait. Further, others suggest that creativity may even be trainable, so it might change under some conditions (Sternberg, 1999). Here we assess the "stability" of creativity over time, where stability is defined as similar behaviors and performance over time and across different conditions.

In this paper, we specifically ask: how does creativity vary over time while working from home, in the presence of different interventions (i.e., is it stable)? Are people more creative under one condition and less creative under another? Existing research has not considered the stability of creativity over weeks or months. Here we assess the creativity using two different measures over nine weeks.

Creative Thinking

Creativity is often considered in terms of divergent and convergent thinking. Divergent thinking represents the ability to generate novel ideas in a context where more than one solution exists. One common measure of divergent thinking creativity is the Alternate Uses Task (AUT) (Vartanian et al., 2019). In contrast, convergent thinking taps into the ability to produce quick and logical solutions to a particular problem and is commonly assessed with the Remote Associates Test (RAT)(Mednick, 1968). Both of these measures are well-validated, often uncorrelated, and are widely used in the creative problem-solving literature (Carroll, 1968; Colzato, Szapora, & Hommel, 2012; Mednick, 1968). Still, little is known about the stability within individuals in these measures over longer periods of time (Baer, 1994; Magnusson & Backteman, 1978).

Experiencing Nature Through Virtual Reality (VR)

Although creativity has been linked to stable personality traits, it might be affected by diverse practices (Eysenck, 1993). One promising practice is that of experiencing nature through virtual reality (VR). Prior work has established that daily walking in nature may be an effective way to increase creativity (Atchley, Strayer, & Atchley, 2012; Logan, Berman, Berman, & Prescott, 2020). One way to experience nature, especially during a pandemic or when working from home, is virtual reality (VR). In the experiment that we describe here, we explored how

VR nature experiences affect creativity. In the current paper, we specifically focus on the stability of daily assessments of divergent and convergent creative thinking.

Stability in behavior

There is a gap in understanding stability of creativity as a theoretical construct. Considering creativity as a personality trait suggests that creative capacity is consistent and stable (Cervone & Shoda, 1999; Diener & Larsen, 2009; Epstein, 1979). However, people show instability in behavior that personality traits might influence, such as moral feelings, across situations (Endler & Hunt, 1966; Mischel, 2013). Individual differences often account for less variance in behavior compared to the interaction of individuals and situations (Argyle & Little, 1972; Endler & Hunt, 1966).

This study contributes to understanding behavioral stability by examining the stability of creativity in individuals across nine weeks during the COVID-19 pandemic. During this time, participants worked from home. We exposed them to mindfulness and virtual reality nature interventions, and measured creativity daily. We fit a multi-level distributional model to understand individual differences in response to VR experiences and mindfulness training. This model shows whether an individual's creativity varies across conditions and the degree to which creativity is stable. We also calculated a form of within-person stability to describe each individual's stability over time.

METHOD

We conducted a within-subject quasi-experiment to examine the stability of creativity across three conditions: baseline, virtual nature, and virtual nature with mindfulness training. We assessed participants' creativity in terms of convergent and divergent creativity. We used the RAT method to measure convergent thinking and the AUT method to measure divergent creativity.

Participants

Twenty participants were recruited for this study from major U.S. cities such as New York, Chicago, and Seattle during the COVID-19 pandemic. The participants were remote-workers, and they were employed for the nine weeks of the study They were between the ages of 25 to 68 (19 female, 1 male). Exclusion criteria included unwillingness to use a VR headset device, severe eye or neurological disorder, and general anxiety score higher than 50 on the State-Trait Anxiety Inventory (STAI) measure (Spielberger, 2010).

Materials

Each participant was given an Oculus Go VR headset, which they kept after the study as a form of compensation. The Oculus Go has a 538ppi 2560 x 1440 WQHD, fast-switch LCD. Participants installed the Guided Relaxation VR application for watching immersive nature videos. For the mindfulness practice, participants used a smartphone, and they downloaded and used the "Healthy Minds" application (Goldberg et al., 2020). The application can be installed on both Android and iOS systems.

Design

The study used a within-subjects quasi-experimental design. Participants were recruited for nine weeks, and they experienced three conditions each lasting three weeks. All participants completed the three conditions in the same order: baseline, VR nature experience, and VR nature and mindfulness. The dependent variables included two measures of divergent and convergent creativity.

Procedure

During the first three weeks (weeks 1 to 3), participants were asked to check-in daily (weekdays) and answer four questions including two creativity questions that they received via text messages. One of the questions measured divergent thinking and the other question measured convergent thinking.For the second three weeks (weeks 4 to 6), the participants were asked to experience a nature scene through VR headsets five days a week and answer similar creativity questions after the experience. During the third three weeks (weeks 7 to 9), the participants were asked to add about 10 minutes of mindfulness practice to their daily VR nature experience. They were asked not to do VR practice outside of the study during study period. A cloud communication platform was used to automate the daily text messages. Figure 1 shows the study process.

Analysis Method

Creativity measures. Convergent creativity was measured by using the Remote Associate Test (RAT), which included two metrics: response time and the number of correct answers. Divergent thinking is the capability of presenting several potential answers to a given problem. For example, the participants were given a word such as "toothbrush" and asked to think of as many uses for it in two minutes. The total number of answers shows fluency, and the answers' semantic distance indicates originality. These measures of divergent and convergent thinking do not guarantee creative achievement, but they are valid and reliable (Runco & Acar, 2012).

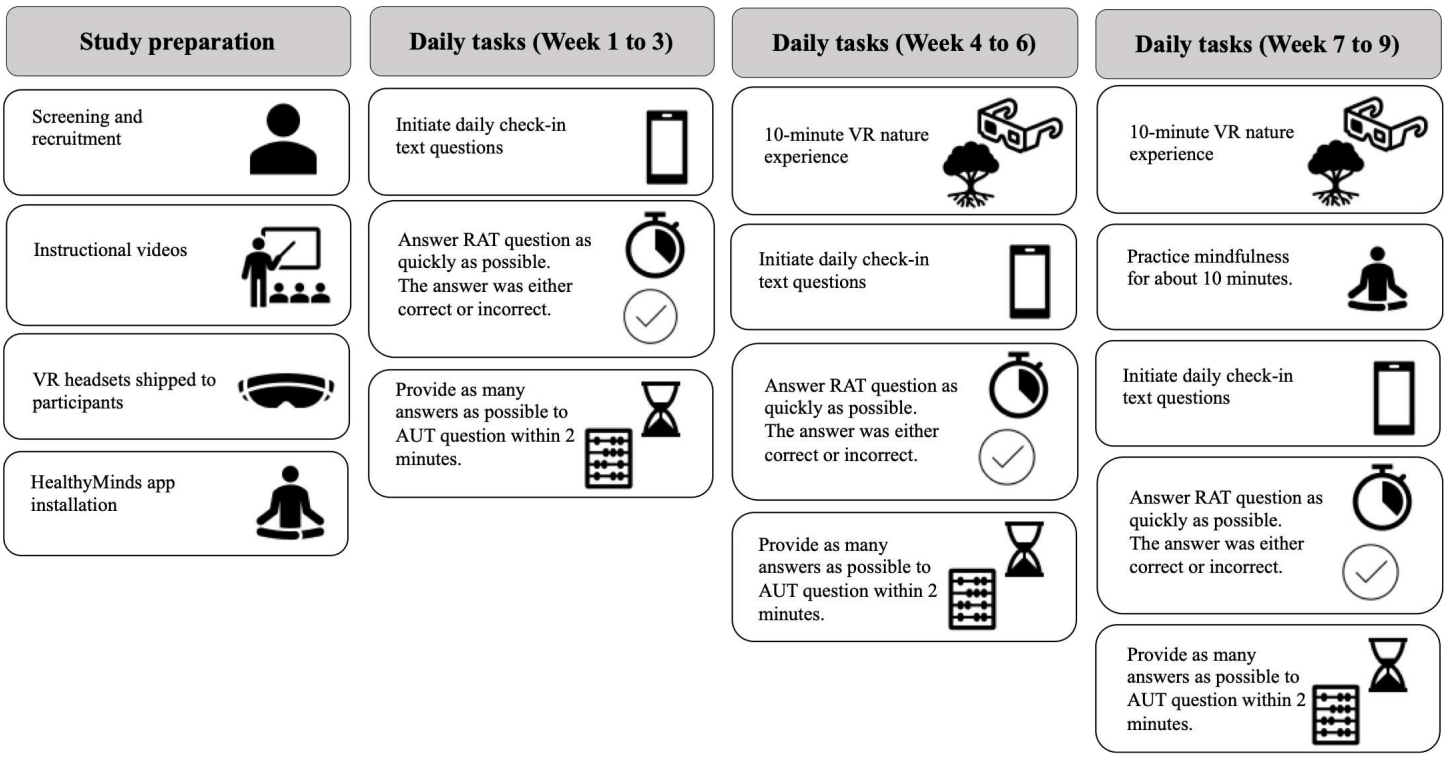

Figure 1. Experiment process for measuring stability in creativity.

Multi-level distributional model. Multi-level models represent data that have been measured in multiple levels or groups (Demidenko, 2013; Gelman & Hill, 2006). Such models consider dependency in the measurements that occur with repeated measures from the same person. Multi-level models can estimate the degree to which stable individual differences influence outcomes. Multi-level models are limited because they estimate the mean of the response distribution and the variance is considered to be constant across groups and conditions (Bürkner, 2017). Yet, in many cases, this variance is not constant. A distributional model estimates the variance of the response distribution across groups and conditions (Rigby & Stasinopoulos, 2005). Distributional models consider within person variability. In this study, we fitted distributional models to estimate individual differences in creativity. Individuals were considered as a random effect.

Stability analysis. Stability is the variance accounted for by a given "cluster," or in our case the individual. This is often calculated as the Intra-class Correlation Coefficient (ICC) and it describes the degree of similarity over time in a single person's response. ICC is the variance associated random effect divided by the random effect plus the residual. For example, if the ICC is zero, then participants responses show very low similarity, or stability. If the ICC is nearly 1, there is a very high similarity between individual scores indicating high stability. ICC is commonly used to assess strength of individual differences.

RESULTS

We fitted Bayesian generalized multi-level distributional models to four measures of creativity: response time and the number of correct answers in convergent thinking, and the total number of answers and semantic distance in divergent thinking. Also, we calculated the ICC of each of the four measures to assess the stability of creativity. Table 1 shows a generally similar pattern emerged for all dependent variables. Participants varied in the overall creativity—the random intercept—and in their response to the conditions—random slope. The distributional model shows that the variability of response shows a similar pattern, except for the RAT correctness.

Convergent thinking

Between subject variability. Response time data and the number of correct answers in the RAT questions were used in the Bayesian models. The random effect was the participants, and the fixed effect was the VR nature and VR+Mindfulness interventions. A model for the response time that included intercept and slope random effects fit the data best. This means that participants show variability in their response time and they differed in their response to the conditions. Also, the influence of the fixed effect was significant. The model's explanatory power is moderate (R^2 = 0.19 , CI [0.12, 0.25]). For the number of correct answers, the model included random effect for intercept, meaning that individuals vary in the rate of cor-

rect response. Still, in total, people did not show variability in the correctness of their response across different conditions. The model's explanatory power is substantial (R^2 = 0.29, CI [0.28, 0.33]).

Within person variability. The Intra-class Correlation Coefficient (ICC) ratio is 0.27, showing a low similarity between the same group's values, and low stability in the RAT response time. A similarly low ICC of 0.27 was found for correctness, showing a low similarity between the same group's values, and low stability in the RAT answers' correctness. Table 1 compares the four creativity measures models.

Divergent thinking

Between subject variability. The total number of answers and the semantic distance of the answers to the AUT questions were used for modeling the data. The random effect was the participants, and the fixed effect was the VR nature and VR+Mindfulness interventions. Also, the influence of the fixed effect was significant. The total number of answers is an indicator of fluency and the semantic distance is an indicator of the originality of the answers. The fluency model showed a random effect for variance intercept and slope. This means that participants show variability in their response time and the participants' total number of answers they responded to the question differed from each other. The total number of answers is an indicator of fluency. The model's explanatory power is substantial (R^2 = 0.58, CI [0.55, 0.61]).

The semantic distance measures the similarity between words. We measured the semantic distance as a metric of divergent thinking, by calculating the distance between the given object in the AUT question, and each use given by the participant. The average distance was calculated using spaCy, a library for natural language processing. spaCy compares word vectors and returns a number between 0 and 1. The results showed that the semantic distance model has a random effect for variance intercept and slope, proving individuals' variability of the semantic distance of the total number of responses. Also, participants' originality (semantic distance), varied from each other. The model's explanatory power is weak (R^2 = 0.08, CI [0.05, 0.11]).

Within person variability. The Intra-class Correlation Coefficient (ICC) for the total number of answers is 0.56, showing a moderate similarity between the same group's values, and low stability in a total number of AUT answers. A low ICC of 0.07 was found for semantic distance, showing a low similarity between the same group's values, and low stability in the semantic distance of AUT answers. Table 1 shows the models for each creativity measure.

Table 1. *Distributional model and ICC for creativity measures*

Dependent Variable	Mean Response	Variance Response	R^2	ICC
RAT response time	Random: Intercept +Slope Fixed: Intervention	Random: Intercept +Slope Fixed: No	0.19	0.27
RAT correctness	Random: Intercept Fixed: Intervention	Random: Intercept Fixed: No	0.29	0.27
AUT total answers	Random: Intercept +Slope Fixed:Intervention	Random: Intercept +Slope Fixed: No	0.58	0.56
AUT semantic distance	Random: Intercept +Slope Fixed: Intervention	Random: Intercept +Slope Fixed: No	0.08	0.07

DISCUSSION

We investigated (a) individual differences in the level of creativity through multi-level distributional models, and (b) the stability of creativity in remote workers during the COVID-19 pandemic.

To assess creativity, we measured the convergent and divergent thinking of remote workers. Convergent thinking was measured by the response time and the number of correct answers to the RAT questions. Divergent thinking was measured by the total number of answers to the AUT questions and their semantic distance. The convergent thinking distributional models showed individuals differ in how fast they respond to convergent thinking questions, and the interventions affect people differently. Measured in terms of the number of correct answers, individuals differed, but they responded similarly to interventions. In terms of individual variability, the results showed that people are not very stable in convergent thinking; this was true for both the response times and the number of correct answers.

The divergent thinking distributional models revealed that individuals differed in terms of the total number of responses (fluency) and semantic distance of the answers (originality). Also, we found low stability in the AUT-semantic distance and similar stability for AUT-fluency. That is not entirely surprising for fluency because is likely a more stable individual difference driven by vocabulary (Beaty & Johnson, 2021). These individual differences, particularly those indicating differences in response to the creativity interventions, suggests a need to personalize the virtual reality experience.

CONCLUSIONS AND FUTURE RESEARCH

In general, the current study showed that individuals differ in the variance and mean of their creativity level. Their creativity level also varies in different conditions, such as exposure to VR and mindfulness. Also, our par-

ticipants showed instability in most measures of creativity.

The results of the current study indicate the usefulness of adding daily VR and mindfulness practice to remote workers' daily schedules to enhance their creativity. However, more research is needed to assess the magnitude of this effect and inform the design of such interventions.

One limitation of the study was that we only used RAT and AUT methods to measure creativity, even though many other exist. Our estimate of creativity and the associated stability calculations were only based on RAT and AUT measures. A more robust approach would adopt a multi-trait multi-method and operationalize creativity with multiple other tasks (Campbell & Fiske, 1959; Hammond, Hamm, & Grassia, 1986). If we measured convergent and divergent thinking with multiple measures, we could differentiate method from trait variance. Another limitation is that we collected creativity data daily across a nine-week period, but but we did not systematically sample creativity across the day. Our measures showed substantial daily variation, and some of this daily variability might reflect when creativity changes from morning to evening. Whether VR and mindfulness interventions have an immediate effect on creativity and how these compare to daily variability should be considered in future studies.

ACKNOWLEDGEMENT

We thank the members of the University of Wisconsin-Madison Cognitive Systems Laboratory for their helpful comments and discussions. This work was partially supported by NSF grants CMMI- (University of New Hampshire: 1840085; Wellesley College: 1840031; University of Wisconsin-Madison: 1839484).

REFERENCES

Argyle, M., & Little, B. R. (1972). Do personality traits apply to social behaviour? *Journal for the Theory of Social Behaviour*.

Atchley, R. A., Strayer, D. L., & Atchley, P. (2012). Creativity in the wild: Improving creative reasoning through immersion in natural settings. *PloS one*, *7*(12), e51474.

Baer, J. (1994). Performance assessments of creativity: Do they have long-term stability? *Roeper Review*, *17*(1), 7–11.

Beaty, R. E., & Johnson, D. R. (2021). Automating creativity assessment with semdis: An open platform for computing semantic distance. *Behavior research methods*, *53*(2), 757–780.

Bürkner, P.-C. (2017). Advanced bayesian multilevel modeling with the r package brms. *arXiv preprint arXiv:1705.11123*.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, *56*(2), 81.

Carroll, J. B. (1968). Reviews: Guilford, jp the nature of human intelligence. new york: Mcgraw-hill, 1967. 538+ xii pp. $14.75. *American Educational Research Journal*, *5*(2), 249–256.

Cervone, D., & Shoda, Y. (1999). Beyond traits in the study of personality coherence. *Current Directions in Psychological Science*, *8*(1), 27–32.

Colzato, L. S., Szapora, A., & Hommel, B. (2012). Meditate to create: the impact of focused-attention and open-monitoring training on convergent and divergent thinking. *Frontiers in psychology*, *3*, 116.

Demidenko, E. (2013). *Mixed models: theory and applications with r*. John Wiley & Sons.

Diener, E., & Larsen, R. J. (2009). Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses. In *Assessing well-being* (pp. 7–24). Springer.

Endler, N. S., & Hunt, J. M. (1966). Sources of behavioral variance as measured by the sr inventory of anxiousness. *Psychological Bulletin*, *65*(6), 336.

Epstein, S. (1979). The stability of behavior: I. on predicting most of the people much of the time. *Journal of personality and social psychology*, *37*(7), 1097.

Eysenck, H. J. (1993). Creativity and personality: Suggestions for a theory. *Psychological inquiry*, *4*(3), 147–178.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Goldberg, S. B., Imhoff-Smith, T., Bolt, D. M., Wilson-Mendenhall, C. D., Dahl, C. J., Davidson, R. J., & Rosenkranz, M. A. (2020). Testing the efficacy of a multicomponent, self-guided, smartphone-based meditation app: Three-armed randomized controlled trial. *JMIR mental health*, *7*(11), e23825.

Hammond, K. R., Hamm, R. M., & Grassia, J. (1986). Generalizing over conditions by combining the multitrait-multimethod matrix and the representative design of experiments. *Psychological Bulletin*, *100*(2), 257.

Logan, A. C., Berman, S. H., Berman, B. M., & Prescott, S. L. (2020). Project earthrise: Inspiring creativity, kindness and imagination in planetary health. *Challenges*, *11*(2), 19.

Magnusson, D., & Backteman, G. (1978). Longitudinal stability of person characteristics: Intelligence and creativity. *Applied Psychological Measurement*, *2*(4), 481–490.

Mednick, S. A. (1968). The remote associates test. *The Journal of Creative Behavior*.

Mischel, W. (2013). *Personality and assessment*. Psychology Press.

Peppercorn, S. (2020). *How to stay creative when life feels monotonous. harvard business review.* Retrieved from `https://hbr.org/2020/11/how-to-stay-creative-when-life-feels-monotonous`

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *54*(3), 507–554.

Rozenes, S., & Cohen, Y. (2017). *Handbook of research on strategic alliances and value co-creation in the service industry*. IGI Global.

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity research journal*, *24*(1), 66–75.

Spielberger, C. D. (2010). State-trait anxiety inventory. *The Corsini encyclopedia of psychology*, 1–1.

Sternberg, R. J. (1999). *Handbook of creativity*. Cambridge University Press.

Vartanian, O., Beatty, E. L., Smith, I., Forbes, S., Rice, E., & Crocker, J. (2019). Measurement matters: the relationship between methods of scoring the alternate uses task and brain activation. *Current Opinion in Behavioral Sciences*, *27*, 109–115.