Measuring Creativity in Exploratory Visual Analysis

ALISHA VARMA, University of Maryland, College Park
HANNAH BAKO, University of Maryland, College Park
LEILANI BATTLE, University of Washington, Seattle

1 INTRODUCTION

As famously said by Ben Shneiderman, "the purpose of visualization is insight, not pictures." The process of uncovering insights is often described in visualization as a series of "eureka" moments [4], where an analyst discovers unique ways to connect their own knowledge and experience with the information presented within a visualization. This process allows for the generation of ideas and insights. These insights allow for the discovery and advancement of knowledge and patterns that further our understanding as a society, no matter how big or small. It is also from this uncovering insights process, that it stands to reason that creativity must play a role in this specific insight generation process. In particular, there has been interest in exploring the intersection of data exploration and creativity in the literature [42].

The definition of an insight in the visualization community is a piece of information or advancement in knowledge[4]. Insights in the visualization community are viewed as objects that can be attained[4] and in turn quantified, rather than a series of events. Insights in the visualization community are attained through the exploratory data analysis of visualizations. When these connections, insights, are formed they result in ideas that can be considered unique and effective. This output can be classified as a result of creativity.

Creativity is a vague concept, and so far it has been unclear what creativity means in the visualization community. However, in psychology, creativity is composed of two main components; the first component being originality in which something is novel or unique and second being effectiveness in which something is valued and considered useful [38]. The unclear relationship between creativity and data analysis has prompted us to question what the possible effects (if any) that an analyst's creativity could have on their ability to uncover insights during the exploratory visual analysis process.

In this report, we present the design for an experiment to investigate potential relationships between creativity and insight generation during exploratory visual analysis. Specifically, we seek to determine whether creativity has a measurable effect on the total insights and quality of insights an analyst is able to make as they perform visual exploratory analysis. However, rather than focusing on the specific insights an analyst uncovers, we plan to measure the act of making these connections. We are considering the act of making the connections as a form of measuring of creativity, along with more conventional psychology-based creativity survey measures.

This project, in particular, has brought to light a variety of ways that creativity can be measured and lights in which it can be viewed. Thus far, we have revised our experiment protocol over five times and went through countless review cycles to refine how we are measuring creativity. Creativity has been and is a difficult concept to measure, therefore the

Authors' addresses: Alisha Varma, alishav@umd.edu, University of Maryland, College Park; Hannah Bako, hbako@umd.edu, University of Maryland, College Park; Leilani Battle, leibatt@cs.washington.edu, University of Washington, Seattle.

following work though interdisciplinary focuses on the visualization community's currently unsolidified measure of creativity.

2 RELATED WORK

2.1 Measuring Creativity

2.1.1 Measuring 'In-The-Moment' Creativity. 'In-the-moment' creativity is the current state of an individual's creativity, as there can be fluctuations in an individual's creativity during the execution of a creative task [3]. 'In-the-moment' creativity is typically measured through surveys and self-reporting, however there also has been a record of using external judges to rate an individual's creative work [3, 18]. Carroll and Latulipe used all such approaches when trying to capture 'in-the-moment' [3]. Though not the exact same approach, we look to "free-association tasks" to conduct a similar external judge approach in order to identify the current state of an individual's thinking process and creativity.

"Free-association tasks" have been used for decades to measure (in-the-moment) creativity [1, 30, 32]. The intuition behind free-association tasks is that they enable researchers to measure associative thinking processes, which have been known to correlate with divergent thinking, a core component of creativity [30]. Divergent thinking is also a focus of other well-known creative thinking tests [30], such as the Torrance Tests of Creative Thinking [48] and Guilford's Alternate Uses test [15]. Tests related to divergent thinking have also been used to measure creativity in the context of human-computer interaction [28]. Though not the exact same tests observed in prior HCI work, our selected measures target similar cognitive phenomena, and are also well-established in the psychology literature [1, 30, 32].

2.1.2 Measuring Everyday Creativity. Everyday creativity, also called "little c" creativity, is the creativity that individuals exhibit everyday through their daily activities [22, 43]. Everyday creativity is an important part of measuring creativity as captures a broad timeframe of creative activities rather than relying on the result of a single intelligence test, providing a more stable measure of an individual's creative experience. This type of creativity is typically measured using self-reporting instruments such as the application of the Biographical Inventory of Creative Behaviors (BCIB) survey and the Creative Behavior Inventory (CBI) survey [43, 44]. The CBI has previously been used in used in HCI literature [3], and in previous psychology literature [10, 44] as self-reporting assessments in measuring creativity. The more recent BCIB is also considered a reliable survey to measure everyday creativity [43], and has been shown to correlate with the CBI [44]. Using the BCIB allows us to identify the creative behaviors that an individual does throughout their everyday life without having to keep track of that individual's daily activities and it allows for a rolling time-frame to reduce any external influences pertaining to factors like age [43]. External factors like age could influence the amount of creative experiences one may have encountered during their time, therefore allowing a rolling window allows us to evaluate their most recent mindset.

2.2 Creativity Support Tools

Shneiderman originally proposed the idea of creativity support tools to elicit new ideas and allow for a greater number of individuals to be creative more often [41]. Since then creativity has become a major field in continued HCI research [12]. Creativity support tools are an important part in enabling different forms of expression and the discovery of new ideas [41, 42]. It is these tools that allow individuals to obtain the Aha! moments of insight that Shneiderman mentions [42]. There have been a number of creativity support tools in the literature that are equipped for aiding individuals to perform a greater degree of creative work [25, 35, 46, 47, 51]. One of these creativity support tools is a

Tarot-Based Narrative Generation system created by Sullivan et al. which allows users to draw tarot cards for a story spread and initiate new story ideas in the process [46].

We observe that foundational research in visualization often discusses parallel concepts to those investigated in creativity research but seems to avoid explicit references to creativity, for example emphasizing “innovation” and “discovery” over creativity [14, 39, 42, 50], or focusing on how analysts formulate “hypotheses” and collect “evidence” [23, 37, 50], which we point out is a common way that analysts create stories about the data they explore. Furthermore, visual storytelling is itself a growing area of visualization research [19, 27], further strengthening the link between creativity and visual analysis. We argue that rather than limiting our perception of the role of creativity to presentation purposes only, visualization researchers should also consider the influence of creativity on the process of visual exploration.

2.2.1 Measuring Creative Work. The concept of creative work is a vital part of understanding the effect and impact of creativity support tools that have been developed in HCI [24, 41]. Several studies have been conducted that evaluate the creativity of outputs from creativity support tools [24, 36, 49]. One of the studies is by Kerne et al. which looks into evaluating individuals generated insights and creativity due to their interaction with an information-based ideation platform[24]. Another study by Tripathi and Burleson looked into building computational models to assess creativity based on face-to-face interactions and movements[49]. Measuring the output of creative work to inform and assess creativity support tools has forged the way for new techniques and ideas to improve these tools, however the same measurement has not yet been assessed for exploratory visualization tools.

2.3 Evaluating Exploration Performance

Exploration performance has been measured using a wide range of metrics [2]. We focus pm the most common exploration performance measures for our own analysis: total exploration time [9, 11], total interactions performed by the individual [5, 9, 11, 13], exploration interaction rate [11, 29, 57], and exploration breadth [53] (i.e., total unique combinations of data attributes explored).

A big part of evaluating exploration performance in relation to our study is the collection and evaluation of insights. Discovering insights is generally considered the a critical part of the exploratory analysis process [2, 4, 16, 29, 39]. Two of the most common previous metrics involving insights collection and evaluation that have been used are insight characteristics[16, 39] and originality[16]. These preceding measures are what we specifically intend to use and will inform us regarding the depth and breadth of an individual’s exploration performance.

In this work, we are interested in assessing potential connections between measures of creativity and measures of exploration performance. Specifically, we are seeking to understand whether creativity may influence or be influenced by an individual’s exploration activity. However, many existing visualization insight metrics do not consider insights as measurable objects created by individuals, but rather as transient utterances [4, 16, 29, 39]. We define insights as explicit objects created and maintained by analysts, similar to the definitions of insight proposed by Gotz et al. [14], Smuc et al. [45], and others. Furthermore, we extend this definition of insight by integrating concepts from research on creativity support tools. Specifically, we consider the idea that analysts curate and refine their insights as explicit outputs of their creativity, which researchers in turn can measure similar to other objects produced and curated using creativity support tools (e.g., [24]).

3 MOTIVATION AND RESEARCH QUESTIONS

In the visualization community, tools are developed to aid individuals in creating visualizations and exploring datasets for the end result of insight generation. Insight generation in the visualization community is defined as a substance of information (i.e. an idea) that can be attained [4]. Alternatively, in the HCI community, creativity support tools have been developed to aid the generation of ideas across a wide array of field [17, 42], similar to those developed in the visualization community. The visualization community has established that visualization tools used for exploratory analysis have an impact on insight generation. However, there has been no previous work that illustrates the impact of exploratory visual analysis through the use of visualizations tools on creativity.

On the other hand, psychology literature has identified a relationship between creativity and the generation of ideas [8]. Though this relationship has been identified in psychology, it has not been formally recognized in the visualization community. As a result, we seek to formally investigate and recognize the relationship between creativity and insight generation in the visualization community. We acknowledge that the relationship between creativity and the generation of ideas may also be bidirectional.

This experiment will allow for us to better understand if visualization tools have an impact on an individual's creativity, through the insight generation process. In this experiment, we seek to identify and show the link between exploratory visual analysis and creativity, and argue that in some case exploratory visualization analysis tools can be classified as creativity support tools.

For this experiment, our main research question is: *How might a user's creativity relate to their data exploration performance?* In this experiment, we will measure creativity using two types of surveys: surveys that measure "everyday creativity" [43] (i.e., as an inherent property of a person) and surveys that measure "in-the-moment" creativity [3] (i.e., as a property that can fluctuate in response to a person's environment, emotions, etc.). Since it is unclear whether one of these two existing treatments of creativity will be more predictive of a person's data exploration performance, we intend on measuring both forms of creativity and will analyze potential correlations between them, as well as any potential correlations with established exploration performance measures.

However, it is also possible that other factors may modulate or even overwhelm the effect of creativity on exploration performance. To account for these possibilities alongside to our initial question, we investigate the following secondary research questions:

- How do measures of everyday creativity and in-the-moment creativity compare in the context of data exploration performance?
- How does a user's creativity scores differ pre- and post-exploration?
- How does user's perception of her own creativity relate to her data exploration performance?
- How might a user's prior experience with data analytics modulate the effects of creativity on exploration performance?
- How might a user's perception of the chosen data exploration system modulate the effects of creativity on exploration performance?

These research question will inform us to whether or not there is a link between exploratory data visualization performance and creativity.

4 EXPERIMENT DESIGN

We designed a study to explore the relationship between creativity and insight generation. The study will be conducted online via a video conferencing platform.[1]

4.1 Participants

We will recruit 24 video participants for the main experiment through emails to professional organizations. The video participants will each be compensated with $20 Amazon gift cards.

Recruited participants will be at least 18 years of age with a minimum of 3 months visualization experience through data analysis tools (e.g., Microsoft Excel, Tableau Desktop, and Power BI). Also, the participants will not have any prior experience with the dataset selected for the experiment (Movies), as to be sure that the insights participants exhibited in the experiment was not from prior exploration. We purposefully chose not to place more selective criteria as to allow for greater diversity in the participants [54].

4.2 Experiment Overview

Participants will review and sign a consent form prior to starting the study, and will be notified that they can withdraw at any time. Participants will then complete one main activity broken into seven phases:

- **Phase 1:** complete a survey to measure the participant's baseline in-the-moment creativity level;
- **Phase 2:** explore a provided dataset using a visualization system;
- **Phase 3:** complete a post-task survey to measure the participant's current in-the-moment creativity level;
- **Phase 4:** complete a survey to measure the participants' experience with the visualization system;
- **Phase 5:** complete a survey to measure the participants' perception of their own creativity;
- **Phase 6:** complete a demographic survey to collect information regarding the participants' background and past experience; and
- **Phase 7:** complete a survey to measure everyday creativity.

Phases 1, 3, 5 and 7 are creativity measurement activities, while phase 2 is a data exploration activity. Phases 4 and 6 are system perception and background survey activities. Each of the activities is described below.

4.2.1 Phases 1, 3, 5 and 7: Creativity Measurement. In phases 1 and 3, participants will be asked to answer survey questions designed to measure each participants' creative abilities. We designed our survey questions based on well-known "free-association tasks" from psychological research in measuring creativity [1, 30, 32]. We use one well-known free association task in our study called "continuous association," where participants come up with the first *n* words that they can think of that are related to a specific stimulus word [7, 31]. In this experiment, we focus on color-based and shape-based word associations [20, 26]. Our survey questions are all of the form: *List the first five words that come to mind that are associated with the word [e.g., purple, orange, circle, square].* For each phase, participants will be asked about one color parameter and one shape parameter, resulting in eight possible ordered color and shape combinations:

- purple, circle, orange, square
- purple, circle, square, orange
- circle, purple, square, orange
- circle, purple, orange, square

[1]Tentative pre-registration draft here: https://docs.google.com/document/d/1xDqaHJVLpU1TbwHEZIX4-x3wBx4cubTglxKWrVP16Cw/edit?usp=sharing.

- orange, square, purple, circle
- orange, square, circle, purple
- square, orange, purple, circle
- square, orange, circle, purple

For each cue word (purple, orange, circle, square), participants will be asked to list five words associated with that cue word. We chose not to formally time the participants based on participant performance in our pilots (see subsection 4.4 for details).

The question parameters were selected such that the cue words (purple, orange, circle, square) would have similar numbers of associated words observed in The “Small World of Words” English Word Association Norms dataset [6](124, 106, 116, and 120 total words, respectively). To further preserve the integrity of our survey results, participants will be explicitly asked not to look up answers online for each of the surveys, to only write individual words, and to focus on only the cue word provided for each survey question.

Survey Scoring. Creativity will be measured based on how *statistically uncommon* a participants' responses were. Nemeth and Kwan [34] considered a word to be statistically uncommon if it was not reported in the 1952 Minnesota Word Association Norms by J.J. Jenkins [20], otherwise the word was given an association score relative to the frequency in which it appeared. We will use the same approach as Nemeth and Kwan of calculating a frequency score for each survey response and summing all resulting scores to produce a single creativity measure per participant [33]; however, we will considered a word to be statistically uncommon if it was not reported in the more comprehensive “Small World of Words” English Word Association Norms [6]. In the context of The “Small World of Words” English Word Association Norms, we will use cue-response association strength to measure frequency [6]. In this way, we cab measure the relative originality of ideas generated by each participant. Note that we will account for some cases where a participant did not answer a survey question, for example when a participant could not think of a word in time, or repeated an answer they already provided. We will adapt our calculations accordingly by substituting each null result with the max frequency score observed in The “Small World of Words” English Word Association Norms for words associated with the given color or shape. If a participant writes a word that did not appear in the thesaurus for the given cue (i.e., provided a completely original word), then we will substitute a frequency of zero, consistent with the approach of Nemeth and Kwan [34].

As an example, suppose two participants, Mary and Rosa, complete the color association task for the color blue. Mary writes *sea*, *sky*, *cheese*, *calm*, and *bird*. Rosa writes *sky*, *sadness*, *sea*, *water*, and *eyes*. Mary and Rosa each had five total responses. If Mary and/or Rosa had not given five answers their null answers are substituted with the max cue strength score for the color blue. Between the two participants, Mary produced a final strength score of 0.254based on The “Small World of Words” English Word Association Norms [6], while Rosa produced a final strength score of 0.294.To calculate creativity, we compare the final strength scores of both Mary and Rosa; Mary is considered to have more creative answers as she has a lower final strength score.

Phase 5: Short Scale Creative Self Survey. In this phase, participants will fill out the Short Scale Creative Self survey, a well known survey used to collect data regarding participants' perceptions of their own creativity [21]. These results will allow us to measure whether participants' perceptions of their own creativity could be used as a reliable proxy for established creativity measures.

Phase 7: Creative Behaviors Survey. In this last phase, participants will fill out a Biographical Inventory of Creative Behaviors survey so as to collect measures of participants' "everyday" creativity and be able to compare the results of a well-known creativity survey to our word association results [43].

4.2.2 Phase 2: Data Exploration Task. Participants will be asked to perform data exploration using a simplified version of the Voyager visualization system [53]; this simplified version was created by Zeng et al. [55]. During our study, participants will explore an existing dataset from the visualization literature, specifically the movies dataset used by Wongsuphasawat et al. to test the Voyager system [53].

Participants will be introduced to the system and given 5 minutes to familiarize themselves with the environment. Then, participants will then be asked to perform an open-ended data exploration task using the system, similar to the open-ended exploration tasks used in existing studies [29, 52, 53, 58]. All participants will be given 15 minutes to perform the data exploration task, which is in line with prior studies of data exploration [2, 56, 58].

For each insight that participants discovers during their exploration, we will ask participants to bookmark the visualization most closely associated with their insight and to annotate the bookmark with a short comment about their insight. Participants can write multiple comments for the same bookmark if this visualization elicited multiple insights. This approach is based in part on the insight diaries protocol originally proposed by Saraiya et al. [40]. Our approach is also based on prior work in measuring information-based ideation, which involves participants curating an ideas board to record and share insights [24].

4.2.3 Phases 4 and 6: Tool and Demographic Survey. In phase 4, the participants will be given a tool survey to complete after the main exploration task. The participants will be presented with Likert-scale rating questions to measure the participant's perception of the tool used in the study.

The participants will also given a demographic survey to complete after phase 5 of the study, to provide context for our measures of creativity and exploration performance. The survey will collect data on the background of each participant (age, gender, education level, race, and occupation), and Likert-scale ratings pertaining to the participant's perception of her own creativity.

4.3 Data Collection

During the study, we will collect three types of study data: survey responses, creativity measures, and exploration breadth and depth measures. The survey data we collect, as described above, will be transcribed, labeled (in the case of insights), and anonymized for later analysis and sharing.

For the creativity measures, we will measure the participants' creativity scores via their answers to the surveys completed during the study (see Section 4.2.1). For the exploration breadth and depth measures, we will collect known measures from the literature [2, 16, 29, 39, 53]: (1) the total number of insights a participant has during each exploration task; (2) the time between insights; (3) the task completion time; (4) the insight generation rate of the participants (total insights divided by task completion time); (5) the number of visualizations created; (6) the time taken to produce each visualization; (7) the total unique combinations of attributes explored; and (8) the logs of participants' activities within each tool (e.g., mouse clicks, cell changes, visualizations created). The exploration breadth and depth data we will collect will be used to measure exploration pacing and overall interaction rates. The participants' session recordings, and the researchers sessions notes, will be used to calculate the above measures. We will also apply the protocol proposed by Liu and Heer to label participants' insights based on the type of insight observed [29].

We will excluded all data for participant that met any of the following exclusion criteria: (1) the completion time of any one of the exploration tasks is too short (less than 10 mins) or too long (no interactions for more than 4 mins); (2) the participant did not complete the whole study for any reason; (3) the participant repeats the study multiple times (in which only the first trial was included, if the other other criteria are met); (4) the participant did not understand one or more of the tasks or surveys.

4.4 Pilot Study

The objective of the pilot was to test the validity and flow of our experiment design prior to running the full experiment. Our key concerns going into the pilot were experimental flow and Voyager tool comprehension and execution. In conducting the pilot study with four participants, we affirmed most of our perceptions whilst designing the study. However, the pilot did reveal some minor changes that needed to be made in order for us to successfully measure and address our objectives for this study.

During the pilot, we realized that we may have overestimated the time we believed it would take the participants to complete the survey sections of the study. With this realization, we decided not to time participants for the survey questions during Phases 1 and 3 of the experiment (See Section 4.2.1). All pilot participants took less than half of that time to formulate responses.

We also decided not to prompt participants about the task before every survey succeeding the post-task survey (Phase 3 of 4.2.1). Instead, we will prompt participants following the post-task survey about the remainder of the study, and the following execution will be self-directed.

Another alteration the we decided to make as a result of the pilot was to the demo section of the study. Right before the task we allowed participants to interact with the demo Voyager interface for up to 5 minutes. We allocated this time to allow for the participant to get familiar with the environment with an undirected mini exploration. However, it seemed that those that did not spend the full length of time exploring the demo were a bit overwhelmed when it came time to completing the task. Therefore, we decided to require participants to explore the demo for the full 5 minutes and execute a smaller version of the task that was to follow. During this demo process, we will also guided the participants a bit about all the features of the Voyager system that they may have failed to explore, despite the participants watching a tutorial video on the Voyager system prior to the demo.

5 ANALYSIS

The intention of this study is to understand whether there is a relationship between a person's creativity and their exploration performance. In this section, we discuss how we plan on analyzing our study data to answer the research questions stated in Section 3.

5.1 How might a user's everyday creativity relate to their data exploration performance?

To assess if there was any correlation between the participants' baseline creativity and their exploration performance, we will use our creativity survey results to calculate a creativity score for each participant (see subsubsection 4.2.1), and treat these scores as an independent variable for creativity. We will use the following four measures from our study as dependent variables (see subsection 4.3): insight generation rate, interaction rate, total visualizations created, and total unique attribute combinations explored. For each dependent variable, we will train a separate mixed effects model to predict this variable, with creativity score as a fixed effect and participant as a random effect, resulting in four different mixed effects models.

5.2 How might a user's in-the-moment creativity relate to their data exploration performance?

To assess if there was any correlation between the participants' in-the-moment creativity and their exploration performance, we will use our continuous association results to calculate a creativity score for each participant (see subsubsection 4.2.1), and treat these scores as an independent variable for creativity. In this analysis, we will consider only the first measure of in-the-moment creativity (Phase 1 of subsubsection 4.2.1). We will use the following four measures from our study as dependent variables (see subsection 4.3): insight generation rate, interaction rate, total visualizations created, and total unique attribute combinations explored. For each dependent variable, we will train a separate mixed effects model to predict this variable, with creativity score as a fixed effect and participant as a random effect, resulting in four different mixed effects models.

To see if there is a correlation between everyday and in-the-moment creativity, we will also train a linear regression model using the scores from the BICB and our pre-exploration word association task scores.

5.3 Does a user's in-the-moment creativity differ pre- and post-exploration?

To assess whether there was a significant difference between participants' in-the-moment creativity scores before and after exploration, we will train a mixed effects model with survey order as a fixed effect (pre- or post-exploration), and participant as a random effect. The in-the-moment creativity score calculated in subsubsection 4.2.1 will be used as the dependent variable to be predicted.

5.4 How might a user's perception of their own creativity affect their data exploration performance?

To assess whether a participants' perception of her own creativity was indicative of her "true" baseline creativity, we will use the survey results from our study as an independent variable for perceived creativity. We will use our everyday creativity scores variable from subsection 5.1 as a dependent variable. We will then train a mixed effects model to predict the dependent variable, using perceived creativity as a fixed effect, and participants as a random effect. We will also train a mixed effects model for each of the four exploration performance measures from **??**. We will use perceived creativity as the fixed effect instead of creativity scores, and included participants as a random effect.

5.5 How might a user's prior analysis experience modulate the effects of creativity?

To assess how a participant's perception of Voyager influenced their exploration performance, we will incorporate the demographic survey results from our study as an additional independent variable (i.e., an additional fixed effect) in our mixed effects models in subsection 5.2 and subsection 5.1. Specifically, we will categorize participants into discrete categories of analysis expertise based on their survey results, and use these categories as an additional fixed effect in our models.

5.6 How might a user's experience with a given data exploration system modulate the effects of creativity?

To assess how a participant's perception of Voyager influenced their exploration performance, we will incorporate the Likert-scale survey results from our study as an additional independent variable (i.e., an additional fixed effect) in our mixed effects models in subsection 5.2 and subsection 5.1. Specifically, we will categorize participants into discrete preference groups based on their survey responses, and use these groups as an additional fixed effect in our models.

6 DISCUSSION AND FUTURE WORK

Through this project, we are seeking to measure whether there is a relationship between creativity and exploratory data analysis through the insight generation process. We intend for this research to lead to the formal recognition of a relationship between creativity and exploratory visual analysis in the visualization community.

During this project, one of the major lessons that we learned was creativity is not an easy concept to measure. In this project, we focused on two types of creativity: “in-the-moment" creativity and “everyday creativity." When designing the study, we discovered that it is possible that an individual’s perception of their own creativity may play a role in their exploration experience and end performance. This individualized unknown variable has brought about additional questions in our project and majorly changed the way we intend to conduct our analysis. In addition, knowing that one’s perception of their own creativity plays a role in data exploration performance, it stands to reasons that we could call into question the impact of directed exploration or scientific exploration techniques on creativity.

In the near future, we intend on executing the remainder of the full study. In order to collect quality data, we will take our time during recruiting and study session execution stages of this study. The data from each participant is highly important as we have chosen a smaller participant pool. Afterwards, we intend on conducting our analysis mentioned in section 5, and examine the results. From this analysis, we will decide whether or not to conduct a crowdsourced replication study. If we decided to conduct the replication study, we will collect the data using Prolific and repeat the analysis process. Once the analysis is complete, we intend to summarize our findings and submit a paper to a conference.

While we would like to further investigate the relationship between creativity and a individual’s exploration performance after this study, we will not be able to see the direction a subsequent study could take until after the study is complete.

7 ACKNOWLEDGEMENTS

We thank all the reviewers, study participants, and members of both the Human-Computer Interaction Lab and the Battle Data Lab for their valuable feedback.

This work is also supported in part by the Distributed Research Experiences for Undergraduates (DREU) program, a joint project of the CRA Committee on the Status of Women in Computing Research (CRA-W) and the Coalition to Diversify Computing (CDC), which is funded in part by the NSF Broadening Participation in Computing program.

REFERENCES

[1] Teresa M Amabile. 1983. The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology* 45, 2 (1983), 357.

[2] Leilani Battle and Jeffrey Heer. 2019. Characterizing exploratory visual analysis: A literature review and evaluation of analytic provenance in tableau. In *Computer Graphics Forum*, Vol. 38. Wiley Online Library, 145–159.

[3] Erin A. Carroll and Celine Latulipe. 2011. Capturing ’in the Moment’ Creativity through Data Triangulation. In *Proceedings of the 8th ACM Conference on Creativity and Cognition* (Atlanta, Georgia, USA) *(Camp;C ’11)*. Association for Computing Machinery, New York, NY, USA, 321–322. https://doi.org/10.1145/2069618.2069677

[4] Remco Chang, Caroline Ziemkiewicz, Tera Marie Green, and William Ribarsky. 2009. Defining Insight for Visual Analytics. *IEEE Computer Graphics and Applications* 29, 2 (2009), 14–17. https://doi.org/10.1109/MCG.2009.22

[5] Filip Dabek and Jesus J Caban. 2017. A Grammar-based Approach for Modeling User Interactions and Generating Suggestions During the Data Exploration Process. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 41–50. https://doi.org/10.1109/TVCG.2016.2598471

[6] Simon De Deyne, Danielle J Navarro, Amy Perfors, Marc Brysbaert, and Gert Storms. 2019. The “Small World of Words” English word association norms for over 12,000 cue words. *Behavior research methods* 51, 3 (2019), 987–1006.

[7] Simon De Deyne, Daniel J Navarro, and Gert Storms. 2013. Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behavior research methods* 45, 2 (2013), 480–498.

[8] Carsten KW De Dreu, Bernard A Nijstad, Myriam N Bechtoldt, and Matthijs Baas. 2011. Group creativity and innovation: A motivated information processing perspective. *Psychology of Aesthetics, Creativity, and the Arts* 5, 1 (2011), 81.

[9] Kyriaki Dimitriadou, Olga Papaemmanouil, and Yanlei Diao. 2014. Explore-by-Example: An Automatic Query Steering Framework for Interactive Data Exploration. In *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data* (Snowbird, Utah, USA) *(SIGMOD '14)*. Association for Computing Machinery, New York, NY, USA, 517–528. https://doi.org/10.1145/2588555.2610523

[10] STEPHEN J. DOLLINGER. 2003. Need for Uniqueness, Need for Cognition, and Creativity. *The Journal of Creative Behavior* 37, 2 (2003), 99–116. https://doi.org/10.1002/j.2162-6057.2003.tb00828.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2162-6057.2003.tb00828.x

[11] Mi Feng, Evan Peck, and Lane Harrison. 2019. Patterns and Pace: Quantifying Diverse Exploration Behavior with Visualizations on the Web. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan. 2019), 501–511. https://doi.org/10.1109/TVCG.2018.2865117

[12] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. *Mapping the Landscape of Creativity Support Tools in HCI*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3290605.3300619

[13] David Gotz and Zhen Wen. 2009. Behavior-Driven Visualization Recommendation. In *Proceedings of the 14th International Conference on Intelligent User Interfaces* (Sanibel Island, Florida, USA) *(IUI '09)*. Association for Computing Machinery, New York, NY, USA, 315–324. https://doi.org/10.1145/1502650.1502695

[14] David Gotz, Michelle X. Zhou, and Vikram Aggarwal. 2006. Interactive Visual Synthesis of Analytic Knowledge. In *2006 IEEE Symposium On Visual Analytics Science And Technology*. 51–58. https://doi.org/10.1109/VAST.2006.261430

[15] Joy Paul Guilford. 1967. The nature of human intelligence. (1967).

[16] Hua Guo, Steven R. Gomez, Caroline Ziemkiewicz, and David H. Laidlaw. 2016. A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 51–60. https://doi.org/10.1109/TVCG.2015.2467613

[17] S.R. Herring, B.R. Jones, and B.P. Bailey. 2009. Idea Generation Techniques among Creative Professionals. In *2009 42nd Hawaii International Conference on System Sciences*. 1–10. https://doi.org/10.1109/HICSS.2009.241

[18] Dennis Hocevar. 1981. Measurement of Creativity: Review and Critique. *Journal of Personality Assessment* 45, 5 (1981), 450–464. https://doi.org/10.1207/s15327752jpa4505_1 arXiv:https://doi.org/10.1207/s15327752jpa4505$_1$ PMID: 16370669.

[19] Jessica Hullman and Nick Diakopoulos. 2011. Visualization Rhetoric: Framing Effects in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2231–2240. https://doi.org/10.1109/TVCG.2011.255

[20] James J. Jenkins. 1970. CHAPTER 1 - THE 1952 MINNESOTA WORD ASSOCIATION NORMS. In *Norms of Word Association*, LEO POSTMAN and GEOFFREY KEPPEL (Eds.). Academic Press, 1–38. https://doi.org/10.1016/B978-0-12-563050-4.50004-2

[21] Maciej Karwowski, Izabela Lebuda, and Ewa Wiśniewska. 2018. Measuring creative self-efficacy and creative personal identity. *The International Journal of Creativity & Problem Solving* (2018).

[22] James C. Kaufman and Ronald A. Beghetto. 2009. Beyond Big and Little: The Four C Model of Creativity. *Review of General Psychology* 13, 1 (2009), 1–12. https://doi.org/10.1037/a0013688 arXiv:https://doi.org/10.1037/a0013688

[23] Daniel A. Keim. 2001. Visual Exploration of Large Data Sets. *Commun. ACM* 44, 8 (Aug. 2001), 38–44. https://doi.org/10.1145/381641.381656

[24] Andruid Kerne, Andrew M. Webb, Steven M. Smith, Rhema Linder, Nic Lupfer, Yin Qu, Jon Moeller, and Sashikanth Damaraju. 2014. Using Metrics of Curation to Evaluate Information-Based Ideation. *ACM Trans. Comput.-Hum. Interact.* 21, 3, Article 14 (June 2014), 48 pages. https://doi.org/10.1145/2591677

[25] Joy Kim, Mira Dontcheva, Wilmot Li, Michael S. Bernstein, and Daniela Steinsapir. 2015. *Motif: Supporting Novice Creativity through Expert Patterns*. Association for Computing Machinery, New York, NY, USA, 1211–1220. https://doi.org/10.1145/2702123.2702507

[26] G. R. Kiss, C. Armstrong, R. Milroy, and J. Piper. 1973. An associative thesaurus of English and its computer analysis. In *The computer and literary studies*, A. J. Aitkin, R. W. Bailey, and N. Hamilton-Smith (Eds.). University Press, Edinburgh, UK.

[27] Robert Kosara and Jock Mackinlay. 2013. Storytelling: The Next Step for Visualization. *Computer* 46, 5 (2013), 44–50. https://doi.org/10.1109/MC.2013.36

[28] Sheena Lewis, Mira Dontcheva, and Elizabeth Gerber. 2011. *Affective Computational Priming and Creativity*. Association for Computing Machinery, New York, NY, USA, 735–744. https://doi.org/10.1145/1978942.1979048

[29] Zhicheng Liu and Jeffrey Heer. 2014. The Effects of Interactive Latency on Exploratory Visual Analysis. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2122–2131. https://doi.org/10.1109/TVCG.2014.2346452

[30] Tali R Marron and Miriam Faust. 2018. 15 Free Association, Divergent Thinking, and Creativity: Cognitive and Neural Perspectives. *The Cambridge handbook of the neuroscience of creativity* (2018), 261.

[31] Martha T Mednick, Sarnoff A Mednick, and Charles C Jung. 1964. Continual association as a function of level of creativity and type of verbal stimulus. *The Journal of Abnormal and Social Psychology* 69, 5 (1964), 511.

[32] Sarnoff Mednick. 1962. The associative basis of the creative process. *Psychological review* 69, 3 (1962), 220.

[33] Douglas L Nelson, Cathy L McEvoy, and Thomas A Schreiber. 2004. The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers* 36, 3 (2004), 402–407.

[34] C. Nemeth and Julianne L. Kwan. 1985. Originality of word associations as a function of majority vs. minority influence. *Social Psychology Quarterly* 48 (1985), 277–282.

[35] Tricia J. Ngoon, C. Ailie Fraser, Ariel S. Weingarten, Mira Dontcheva, and Scott Klemmer. 2018. *Interactive Guidance Techniques for Improving Creative Feedback*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3173574.3173629

[36] Adam Perer and Ben Shneiderman. 2008. Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 265–274. https://doi.org/10.1145/1357054.1357101

[37] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.

[38] Mark A. Runco and Garrett J. Jaeger. 2012. The Standard Definition of Creativity. *Creativity Research Journal* 24, 1 (2012), 92–96. https://doi.org/10.1080/10400419.2012.650092 arXiv:https://doi.org/10.1080/10400419.2012.650092

[39] P. Saraiya, C. North, and K. Duca. 2005. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics* 11, 4 (2005), 443–456. https://doi.org/10.1109/TVCG.2005.53

[40] P. Saraiya, C. North, Vy Lam, and K.A. Duca. 2006. An Insight-Based Longitudinal Study of Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 12, 6 (2006), 1511–1522. https://doi.org/10.1109/TVCG.2006.85

[41] Ben Shneiderman. 2002. Creativity Support Tools. *Commun. ACM* 45, 10 (Oct. 2002), 116–120. https://doi.org/10.1145/570907.570945

[42] Ben Shneiderman. 2007. Creativity Support Tools: Accelerating Discovery and Innovation. *Commun. ACM* 50, 12 (Dec. 2007), 20–32. https://doi.org/10.1145/1323688.1323689

[43] Paul J. Silvia, Rebekah M. Rodriguez, Roger E. Beaty, Emily Frith, James C. Kaufman, Paul Loprinzi, and Roni Reiter-Palmon. 2021. Measuring everyday creativity: A Rasch model analysis of the Biographical Inventory of Creative Behaviors (BICB) scale. *Thinking Skills and Creativity* 39 (2021), 100797. https://doi.org/10.1016/j.tsc.2021.100797

[44] Paul J Silvia, Benjamin Wigert, Roni Reiter-Palmon, and James C Kaufman. 2012. Assessing creativity with self-report scales: A review and empirical evaluation. *Psychology of Aesthetics, Creativity, and the Arts* 6, 1 (2012), 19.

[45] Michael Smuc, Eva Mayr, Tim Lammarsch, Wolfgang Aigner, Silvia Miksch, and Johannes Gärtner. 2009. To Score or Not to Score? Tripling Insights for Participatory Design. *IEEE Computer Graphics and Applications* 29, 3 (2009), 29–38. https://doi.org/10.1109/MCG.2009.53

[46] Anne Sullivan, Mirjam Palosaari Eladhari, and Michael Cook. 2018. Tarot-Based Narrative Generation. In *Proceedings of the 13th International Conference on the Foundations of Digital Games* (Malmö, Sweden) *(FDG '18)*. Association for Computing Machinery, New York, NY, USA, Article 54, 7 pages. https://doi.org/10.1145/3235765.3235819

[47] Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. 2012. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on human factors in computing systems*. 1461–1470.

[48] Ellis Paul Torrance. 1998. *Torrance tests of creative thinking: Norms-technical manual: Figural (streamlined) forms A & B*. Scholastic Testing Service.

[49] Priyamvada Tripathi and Winslow Burleson. 2012. Predicting Creativity in the Wild: Experience Sample and Sociometric Modeling of Teams. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA) *(CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 1203–1212. https://doi.org/10.1145/2145204.2145386

[50] J.J. van Wijk. 2006. Views on Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12, 4 (2006), 421–432. https://doi.org/10.1109/TVCG.2006.80

[51] Hao-Chuan Wang, Dan Cosley, and Susan R. Fussell. 2010. Idea Expander: Supporting Group Brainstorming with Conversationally Triggered Visual Thinking Stimuli. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA) *(CSCW '10)*. Association for Computing Machinery, New York, NY, USA, 103–106. https://doi.org/10.1145/1718918.1718938

[52] Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. 2016. Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 649–658. https://doi.org/10.1109/TVCG.2015.2467191

[53] Kanit Wongsuphasawat, Zening Qu, Dominik Moritz, Riley Chang, Felix Ouk, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. 2017. Voyager 2: Augmenting Visual Analysis with Partial View Specifications. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2648–2659. https://doi.org/10.1145/3025453.3025768

[54] Rachael Zehrung, Astha Singhal, Michael Correll, and Leilani Battle. 2021. Vis Ex Machina: An Analysis of Trust in Human versus Algorithmically Generated Visualization Recommendations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 602, 12 pages. https://doi.org/10.1145/3411764.3445195

[55] Zehua Zeng et al. 2021. An Evaluation-Focused Framework for Visualization Recommendation Algorithms. *(to appear in) IEEE VIS 2021* (2021).

[56] Emanuel Zgraggen, Alex Galakatos, Andrew Crotty, Jean-Daniel Fekete, and Tim Kraska. 2016. How progressive visualizations affect exploratory analysis. *IEEE transactions on visualization and computer graphics* 23, 8 (2016), 1977–1987.

[57] Emanuel Zgraggen, Alex Galakatos, Andrew Crotty, Jean-Daniel Fekete, and Tim Kraska. 2017. How Progressive Visualizations Affect Exploratory Analysis. *IEEE Transactions on Visualization and Computer Graphics* 23, 8 (2017), 1977–1987. https://doi.org/10.1109/TVCG.2016.2607714

[58] Emanuel Zgraggen, Zheguang Zhao, Robert Zeleznik, and Tim Kraska. 2018. Investigating the Effect of the Multiple Comparisons Problem in Visual Analysis. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3174053