Curiosity Exploration Styles in Word Association Tasks

Stephen Steinle (ssteinle@usf.edu)

My Duong (myduong@usf.edu)

Rima El Brouzi (relbrouzi@usf.edu)

Suraj Kumar (kumar8@usf.edu)

John Licato (licato@usf.edu)

Abstract

Recent analyses of human creativity and curiosity have identified the existence of three styles of exploration: busybody, hunter, and dancer. These styles were recognized largely by observing participants' explorations within a task, converting those observations into networks, and measuring networks' properties. But do these exploration styles still appear across different tasks? And when graph-based descriptors of an individual's exploration style are identified, how well do they transfer to similar tasks? We study inter- and intra-individual differences in two similar, but distinct, word association tasks: Chain Free Association and Semantic Fluency. We demonstrate that in some cases, graph-theoretic features do seem to capture individual semantic exploration patterns across tasks. Furthermore, our results provide evidence supporting the existence of the dancer style and its relationship to the Busybody-Hunter score. These findings highlight the potential of graph analysis as a tool for characterizing and exploring individual cognitive styles in semantic tasks.

Keywords: busybody; hunter; dancer; creativity; curiosity; chain free association; semantic fluency; psychometrics; graphical analysis, forward flow

Introduction

Understanding human creativity and reasoning involves examining how we navigate our conceptual spaces while performing tasks—a process we refer to as *exploration styles*: individual patterns of idea generation, transitions, and search behavior. But to what extent are these styles unique to an individual, and how much do they vary across individuals?

Recent research has proposed two primary exploration styles in the domain of curiosity: the busybody-hunter and the dancer (Lydon-Staley, Zhou, Blevins, Zurn, & Bassett, 2021; Zhou et al., 2024). In that work, participants were free to navigate Wikipedia articles at their whim and researchers recorded the titles of articles and the hyperlinks that connected them. By interpreting the titles as nodes and the hyperlinks as edges, the researchers formed graphs and examined their underlying structures. Busybodies are said to "toss their heads, and frisk about" (Zurn, 2019), which was represented by Lydon-Staley et al. (2021) as graphs with looser interconnectivity between nodes. Hunters, as the opposite to busybodies, are stated to "track down hunches or leads" (Zurn, 2019), and were represented by tightly connected graphs. A third exploration style, the dancer (Zhou et al., 2024), is categorized by great leaps in the graph space, moving to seemingly unconnected topics (nodes) without the constant navigation of the busybody or depth of the hunter.

Busybodies and hunters form two ends of a spectrum, with the busybody forming loose networks and the hunter forming dense networks. Lydon-Staley et al. (2021) operationalized this concept of network density as the need to find answers, finding that it correlates with a type of curiosity known as deprivation curiosity (Litman & Silvia, 2006). Conversely, the busybody end of the spectrum has an inverse correlation with deprivation curiosity, which is appropriate for both the philosophical concept of Busybody-Hunter (BB-H), and the busybody relationship with network tightness. The dancer exists separately from the BB-H continuum and is measured with creativity instead of curiosity. These styles form the Busybody-Hunter-Dancer (BB-H-D) triad and are observable in tasks without specific emphasis on curiosity or creativity (Zhou et al., 2024). However, much is still unclear about these three exploration styles. For example, can they be observed within and across different cognitive tasks? Are they context-dependent, or do these exploration styles still exist as distinct patterns in other cognitive tasks, specifically word association? Motivated by these questions, we had participants perform both Chain Free Association (CFA) (Marron et al., 2018) and Semantic Fluency (SF) (Anderson & Bower, 2014) cognitive tasks. These tasks were chosen as they are very similar in structure to the Wikipedia exploration task studied in Lydon-Staley et al. (2021), while being more restrained with several pre-existing datasets to draw upon. Following Lydon-Staley et al. (2021), we converted word lists gathered from these tasks into semantic graphs, with the words as nodes and similarity between words as edges, and then calculated properties, "*graph features*", of those graphs. We then set out to address two research questions:

RQ 1: How stable are intra-individual graph features both within and between associative word tasks? We asked participants to perform the SF task twice, and the CFA task once. We then combined this data with existing CFA and SF data and studied the relationship between different graph features within the same individuals; i.e.: do two separate SF lists produced by the same individual have similar graph features? What about a SF list and CFA list produced by the same individual? We find that lists created by the same individual are more similar to each other than to those created by different participants. List similarity is negatively impacted

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

with more distinction between the task instructions, but remained consistent throughout all our testing.

RQ 2: Is there evidence for three distinct exploration styles in associative word tasks? The BB-H-D model indicates that there should be three groups of individual exploration styles when using our specific set of graph features. The busybody and hunter styles exist on a spectrum, while the dancer style lies orthogonal to that spectrum. However, it has not yet been studied whether this pattern exists in associative word tasks like CFA and SF. Again using graph features, we analyze whether there is evidence of these three exploration styles in our associative word task data.

Novel Contributions This work is the first to explore whether graph features can identify inter-individual (IRP) similarities and intra-individual (IAP) similarities in associative word tasks. We also contribute to ongoing literature surrounding the previously-proposed BB-H-D categorization of semantic exploration styles, extending it to a new category of tasks. Finally, we present a new cross-task dataset, with the same participants performing both the CFA and SF tasks, that can be used to study individual differences in semantic exploration.

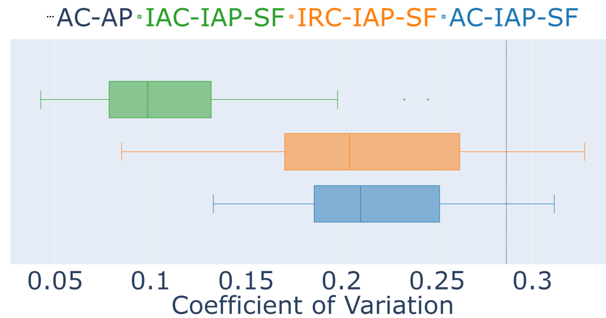
Background

Word association tasks have a long history in cognitive psychology as a way of gaining insight into semantic exploration and retrieval processes.

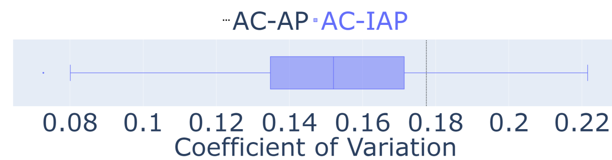
Chain Free Association is an associative word task where participants are given a cue word and instructed to produce a series of words, where each subsequent word is associated with the previous one. CFA has been used to measure creative thinking: metrics such as *Forward Flow* (FF) have been developed to quantify how far people travel in semantic space when freely associating (Gray et al., 2019). It has also been used to study cognitive processes involved in creative ideation (Kenett, 2024; Koivisto & Toivanen, 2024) and investigate the neural correlations of spontaneous associative thinking (TR et al., 2018; Beaty & Kenett, 2023).

Semantic Fluency, or ideational fluency, is an associative word task in which participants are asked to produce as many words as possible in a given category (e.g., fruit, animal, furniture, etc.) within a time limit (Shao, Janse, Visser, & Meyer, 2014). According to Hills, Jones, and Todd (2012), we access our semantic memory in a structured way similar to how animals forage for food. When completing an SF task, people retrieve words in clusters of related terms before moving to a different theme. Several cognitive processes may be involved in the SF task, including perception, parallel and serial processing, memory retrieval, working memory, decision-making, and functional lateralization (Paschoal et al., 2021).

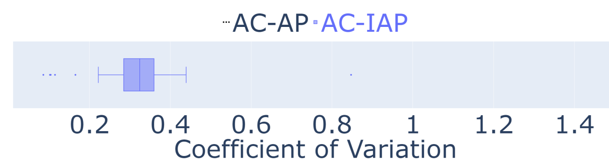
Given its ability to provide insight into how individuals search through semantic spaces, the SF task has also been



(a) Coefficient of Variation of graph features from the SF task.



(b) Coefficient of Variation of graph features from the CFA task.



(c) Coefficient of Variation of graph features across the CFA and SF tasks.

Figure 1: Coefficient of Variation of graph features from the CFA, SF, and between CFA-SF tasks. For the SF task, median values increase in variation as list groups become more generalized (1a), variation across all participants is 50% higher than inter or all category groupings, and nearly three times higher than intra-category. Variation in CFA data is more modest, but still larger across participants. Variation increases dramatically across tasks as expected. Acronyms are discussed in detail in the Analysis section.

of use in studying creativity and curiosity as both are linked to one's semantic memory (Kenett, Humphries, & Chatterjee, 2023). The study of individual differences in creativity and curiosity has spawned a variety of measures that are applicable to word association tasks. For example, *Forward Flow* is a measure that quantifies the semantic evolution of thoughts over time, capturing how much present thoughts diverge from past thoughts. The instantaneous FF of a word describes the semantic distance between words or concepts in a sequence of thoughts using the equation from Gray et al. (2019):

$$\frac{\sum_{i=1}^{n-1} D_{n,i}}{n-1}$$

Where D is the semantic distance between thoughts and n is the numerical location of thought within a stream. The *dynamic forward flow* of an entire thought sequence is calculated as the average of this equation across all thoughts in a sequence (Gray et al., 2019):

$$\frac{\sum_{i=2}^{n}\sum_{j=1}^{i-1} D_{i,j}}{\frac{n(n-1)}{2}}$$

FF strongly correlates with creativity and quantifies a person's ability to generate a continuous, diverse stream of ideas, enabling cognitive flexibility, divergent thinking, and novel associations crucial for creative problem-solving. Higher FF suggests more extraordinary cognitive leaps, showcasing the capacity for divergent thinking and the ability to link unrelated ideas. Several elements contribute to creativity, such as semantic diversity, where individuals engage with a wide range of knowledge domains, and non-linear exploration, which challenges traditional thinking patterns.

Exploration Styles Zurn (2019) observed that interpretations of curiosity usually leaned toward a categorization of curiosity types based on "modes", or actions performed during exploration, rather than the drives behind curiosity. Thus, they proposed three models: busybody, hunter, and dancer, all of which embody typical movements of knowledge exploration while representing different kinds of curiosity. People who follow the busybody style make quick associations and seek sparse coverage (Zurn, 2019). They tend to spontaneously traverse long distances among topics that have little relationships with each other, forming loosely connected networks (Lydon-Staley et al., 2021). Hunters pursue ideas that deepen their knowledge on particular topics (Zurn, 2019). Their networks tend to be tight and structured as they dive into specific problems and explore related ideas (Lydon-Staley et al., 2021). The dancer style emphasizes creative leaps across semantic space. Dancers build expansive knowledge networks that connect otherwise isolated domains, nurturing new associations and producing innovative insights (Zurn, 2019). This style stands apart due to its capacity to integrate diverse ideas connected in surprising ways, resulting in a unique and separate knowledge network representation.

The BB-H score quantifies the connectivity and structure of a person's knowledge network. This score is generated by subtracting the busybody metric from the hunter metric, where H_i is the hunter score and BB_i is the busybody score as seen below (Zhou et al., 2024):

$$H_i = \frac{E_i + \rho_i + C_i + k_i + eff_i + cp_i + modq_i}{n_m^2}$$

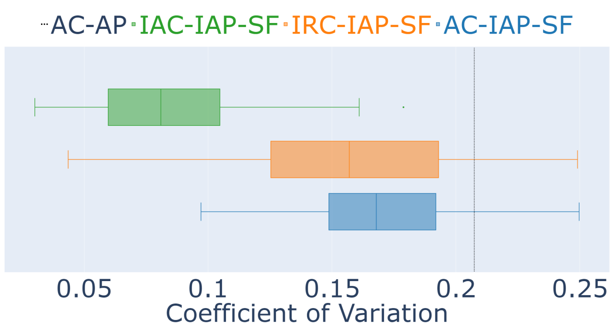
$$BB_i = \frac{N_i + L_i + B_i + mpl_i}{n_m^2}$$

$$\text{BB-H} = \sum H_i - BB_i$$

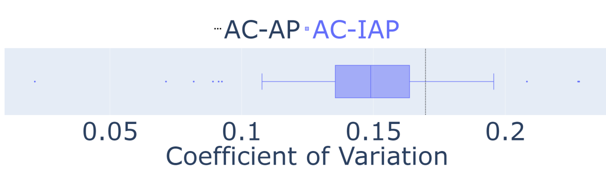
These equations are comprised of many graph features: N_i is the number of nodes in a single network, E_i is the number of edges in the network, ρ_i is the network density, C_i is the clustering coefficient, k_i is the average degree, L_i is the characteristic path length, eff_i is the global efficiency, cp_i is the core-periphery structure calculated with the Borgatti-Everett method, mdl_i is the minimum description length of the hierarchical degree-corrected stochastic block model (SBM), B_i is the number of blocks in the SBM on the lowest level of the hierarchy, $modq_i$ is the modularity of the partition obtained from the SBM, and n_m is the number of graph features used in each formula. As the networks formed by CFA and SF word lists are fully connected and weighted we employed a weighted version of the BB-H score. We utilized average degree, characteristic path length, average clustering coefficient, and average edge weight to calculate the weighted BB-H score, defined in (Lydon-Staley et al., 2021) as:

$$\text{Weighted BB-H} = \bar{w} + \bar{C}_i - L\ ,$$

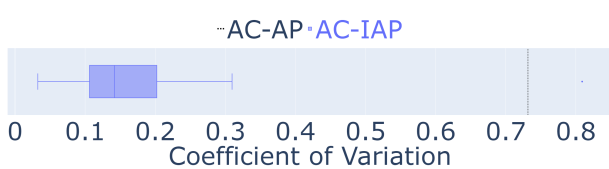
where $\bar{w}$ is the average edge weight, $\bar{C}_i$ is the average clustering coefficient, and L is the characteristic path length. In our case, N_i is the number of words in an individual list. Zhou et al. (2024) applied the BB-H score to knowledge networks constructed from the engagement of a naturalistic set of readers with different Wikipedia articles, finding that the busybody and hunter styles are universal and that the score can represent to which style a particular network belongs.



(a) Coefficient of Variation of BB-H and FF from the SF Task.



(b) Coefficient of Variation of BB-H and FF from the CFA Task.



(c) Coefficient of Variation of BB-H and FF across the CFA and SF Task.

Figure 2: Coefficient of Variation of BB-H and FF from the CFA, SF, and across CFA-SF Tasks. Patterns are similar to those in Figure 1. Acronyms are discussed in detail in the Analysis section.

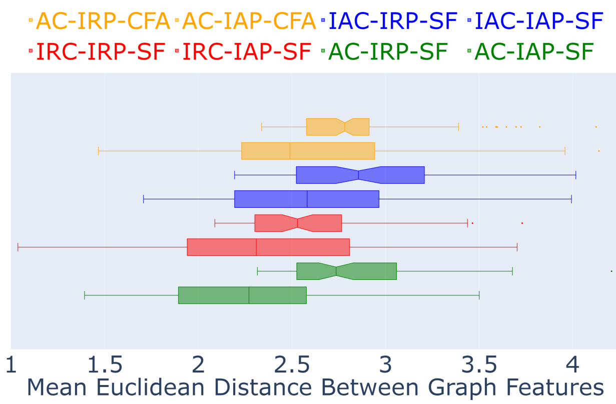


Figure 3: Pairwise and Cross Distance of graph features from the CFA and SF tasks. Colors indicate measurement pairs, with the notched boxes showing distances from one participant's list features to all others (IRP). In contrast, the flat boxes show distances between list features internal to a participant (IAP). All IAP distances are smaller than IRP distances, supporting the existence of IRP consistency in both tasks. Acronyms are discussed in detail in the Analysis section.

Siew and Guru (2023) observed that semantic networks constructed from the SF tasks of college students were tight and had short characteristic path lengths compared to high school students, leading them to speculate a hunter-like exploration pattern in college students. However, no work to our knowledge has studied in detail whether these three types of exploration styles exist in associative word tasks nor whether the graph features that make up the FF and BB-H scores are consistent within individuals across different tasks. We now turn to our study to address this gap.

Methods

This work progressed across three phases: data collection, graphical transformation, and analysis[1] . This section will discuss the first two phases and detail the selected graph features.

Data Collection This work used three datasets: one with CFA data, one with SF data, and one where the same participants performed both tasks. The SF-only dataset used in this work, SNAFU (Zemla, Cao, Mueller, & Austerweil, 2020), contains 807 lists from 82 participants across six categories: 'Fruits,' 'Vegetables,' 'Animals,' 'Supermarket Items,' 'Tools,' and 'Foods.' Each participant gave responses to three categories with eight lists on average. Participants provided an average of 30.4 words per list with a standard deviation (SD) of 16.4 after processing. The CFA-only dataset used in this work, PSU-FF (Beaty, Zeitlen, Baker, & Kenett, 2021), contains 2017 lists from 151 participants across 16 categories: 'Coat,' 'Costume,' 'Dock,' 'Drum,' 'Ladder,' 'Leather,' 'Mail,' 'Purse', 'Robot,' 'Tea,' 'Bench,' 'Engine,' 'Rocket,' 'Train,' 'Gum,' and 'Screen.' Participants gave six to fifteen responses, with one list per category word and provided an average of 8.4 words per list with an SD of 2.7. However, neither the SNAFU nor PSU-FF datasets contain data on how the same individuals perform across both tasks. We therefore collected a new CFA-SF dataset, which utilized anonymized participants drawn from the student body at a public university and via Prolific.co. The student participants were selected from a pool who were involved in a previous multi-task study. Prolific participants responded to a study offer on the website, where they can view and apply for any of the hundreds of studies for which they qualify. Prolific participants had not joined our previous studies or sent individual invitations. Both groups of participants were provided links to a Qualtrics survey and compensated via gift cards or the Prolific system. We selected participants above 18 years of age, located in the United States, and who have access to a desktop computer.

The Qualtrics link led participants to a landing page with a consent form which informed participants of their involvement in a study and the tasks involved. They were instructed to have a distraction-free environment and that their progress would be timed. Timers were used to avoid abuse of the Prolific payment system, which is based on completion time, as well as to encourage more naturalistic list generation. CFA and SF tasks rely on the participant responding with words "off the top of their head" rather than searching for "correct" answers. Participants were then presented with the survey in the following order: SF 1, Break 1, CFA, Break 2, SF 2, End. SF Tasks 1 and 2 were conducted with a two-minute timer following common practices (Harrison, Buxton, Husain, & Wise, 2000; Gourovitch, Goldberg, & Weinberger, 1996; Zarino, Crespi, Launi, & Casarotti, 2014). Participants were given unlimited time to read the instructions; then they needed to press a start button to see the word and start the timer for both task types. Participants were instructed to list as many words belonging to a category, Animals for SF 1 and Fruits for SF 2, that came to their mind within two minutes. Repeating words would not be allowed, and failure to enter a minimum of six words would result in rejection of the response. After SF 1, participants were instructed to relax their minds for two minutes. The CFA task was administered similarly to the SF task, except for the instructions and the timer. The participants were instructed to list the first word that came to mind upon seeing the cue word "table". As they entered words, the previous cue would be replaced with the new word to reinforce the sequential nature of the task. A 30-second timer accompanied each word, and the task would only end when 19 words had been entered; this requirement was not shared with participants. 30 seconds was chosen to conform to previous work (Marron, Berant, Axelrod, & Faust, 2020) and provide additional time to compensate for internet latency. Once the CFA task was completed,

[1] https://github.com/Advancing-Machine-Human-Reasoning-Lab/curiosity-exploration

participants were given another two-minute break, followed by SF 2. Of the Prolific.io participants, 31 were male and 61 were female, with a mean age of 36 and an SD of 12. Of the 108 responses, ten were rejected for various reasons, including non-participation or rejection of the consent form, and six were excluded due to partial completion. Four additional lists were provided by university students. The final CFA-SF dataset held 96 cross-task responses, which provided a mean of 18 words for the CFA task with an SD of 1.9, and a mean of 44.5 for the SF task with an SD of 13.8.

Analysis

Graph Generation

The data gathered from our participants is in the form of word lists. After pre-processing lists to account for spelling mistakes, vulgar language, and low quality (such as repeated words, failure to follow tasks, etc.), we set each word as a graph node. We then used Word2Vec (W2V) (Mikolov, Chen, Corrado, & Dean, 2013) to calculate cosine similarity scores (*CS*) between each pair of words in a participant's list to act as the edge between those nodes. Words that did not appear in the W2V corpus were replaced with '_'. We then used the NetworkX package (Hagberg, Schult, & Swart, 2008) to create a 2-D fully connected, undirected, weighted graph for each list with words as nodes and semantic distance ($w_{i,j}$) as the weight of the edge connecting node i and j, where:

$$w_{i,j} = 1 - CS(i, j)$$

Finally, we assembled our task datasets, now augmented with network features and scores.

Graph Features

The original graph features used in Zhou et al. (2024) were selected due to their applicability for unweighted graphs. As our work is concerned with weighted graphs, due to edge generation in word association tasks, we chose to use a subset of these features to reduce the possibility of confounds or spurious correlations post graph transformation. We selected the component features used to calculate the BB-H and dancer scores as well as an additional equation, average triangle area, that is commonly used in the field of topology to help interpret high dimensional datasets (Clarke, 1986).

- **Average degree** is the mean sum of weighted edges adjacent to each node in the graph (G)

  $$\bar{k} = \frac{1}{n}\sum\sum_{i\in G} w_{i,j},$$

  where $w_{i,j}$ is the edge weight between node i and j (Hagberg et al., 2008).

- **Characteristic path length** is the average of all the shortest paths in the network between two nodes i and j, denoting the minimum distance $d_{i,j}$ for traversing a graph (Hagberg et al., 2008).

  $$L = \frac{\sum_{i\in G}\sum_{j\in G} d_{i,j}}{n\cdot(n-1)}$$

- **Average weighted clustering coefficient** measures the mean extent to which nodes cluster together, given the edge weights $w_{i,j}$, $w_{j,k}$, $w_{k,i}$ among any three nodes that form a triangle (Hagberg et al., 2008).

  $$\bar{C}_i = \frac{2t_i}{k_i(k_i-1)}\sum_{j,k}\left(w_{i,j}w_{j,k}w_{k,i}\right)^{1/3},$$

- **Average triangle area** is calculated using Heron's formula, where $w_{i,j}$, $w_{j,k}$, $w_{k,i}$ are the sides of the triangle, and s is the semiperimeter of the triangle. The resulting areas are then averaged by dividing the total area by the number of triangles t formed in the graph (Weisstein, 2003).

  $$A_{ijk} = \frac{1}{\mathrm{t}}\sum\sqrt{s(s-w_{i,j})(s-w_{j,k})(s-w_{k,i})}$$

- **Average edge weight** is the mean edge weight of the network, computed by averaging all edge weights (Hagberg et al., 2008).

  $$\bar{w} = \frac{1}{n}\sum_{i,j} w_{i,j}$$

Once participant lists were transformed into these graph features, we set to answer our proposed research questions.

RQ 1: How stable are *intra-participant* graph features within and between associative word tasks? If graph features show IAP similarities, we expect that graph features of lists generated by the same individual should be more similar than those generated by different individuals. In other words, IAP graph feature variances should be lower than that of *inter-participant* (or participant-agnostic) graph feature variances. In testing this experimentally, we want to be careful to identify the possible effect of category words—lists generated for the same category word will be more similar than those generated for different category words. We then define the following:

- **Intra-Category, Intra-Participant (IAC-IAP)** for a participant p and function F, denoted $IACP_F(p)$, is computed by first calculating $F(g,c)$ for each graph feature g in the five graph features G within a category c, then averaging these values within each category and again averaged across all categories C belonging to p:

  $$IACP_F(p) = \frac{1}{|C|}\sum_{c\in C}\left(\frac{1}{|G|}\sum_{g\in G} F(g,c)\right)$$

- **Inter-Category, Intra-Participant (IRC-IAP)** for a participant p and function F, denoted $IRCP_F(p)$, is computed by first averaging each feature's values within each category, resulting in five feature-wise scalars per category. F is then computed for each feature across all categories and averaged:

  $$IRCP_F(p) = \frac{1}{|G|}\sum_{g\in G} F\left(\left\{\frac{1}{|S_g^c|}\sum_{s\in S_g^c} s\right\}_{c\in C}\right)$$

- **All-Category, Intra-Participant (AC-IAP)** for a participant p and function F, denoted $ACP_F(p)$, is the mean of $F(g)$ computed for each feature g in the set of graph features G:
$$ACP_F(p) = \frac{1}{|G|} \sum_{g \in G} F(g)$$
- **All-Category, All-Participant (AC-AP)** for all participants P and function F, denoted $ACAP_F(P)$, is identical to $ACP_F(p)$ except that no distinction between participants is made.

We start by defining $CV(S)$ as the *Coefficient of Variation* of a set of graph features S, which is simply the standard deviation of S divided by its mean (Abdi, 2010). The Coefficient of Variation provides a convenient, automatically normalized way to capture the amount of variability in a population and can be calculated as follows:

$$CV(S) = \frac{\sigma_S}{\mu_S}$$

CV is then then applied as the function F for $IACP_F(p)$, $IRCP_F(p)$, $ACP_F(p)$, and $ACAP_F(P)$. This set of functions was used to generate Figures 1 and 2.

The second way we capture population variability is by interpreting graph features for a list as a point in five-dimensional space and then calculating the distance between these points. This is done with the *pairwise distance* of a set S, $PD(S)$, where each point is compared to each other point in the same set. We also calculate the *cross distance* $CD(S,T)$ between a set S and all other sets of lists T.

PD for a set of graph features S, denoted $PD(S)$ is applied $(s_i, s_j) \in S \times S$:

$$PD(s_i, s_j) = \sqrt{\sum_{k=1}^{5} (s_{ik} - s_{jk})^2}$$

CD for a set of graph features S and another set of graph features T, denoted $CD(S,T)$ is applied to $(s_i, t_j) \in S \times T$:

$$CD(s_i, t_j) = \sqrt{\sum_{k=1}^{5} (s_{ik} - t_{jk})^2}$$

PD and *CD* are then applied as the function F for $IACP_F(p)$, $IRCP_F(p)$, $ACP_F(p)$, and $ACAP_F(P)$. This set of functions was used to generate the box plots in Figure 3.

We first consider the SNAFU dataset. As expected, the IAP variation is significantly smaller than the IRP variation for both *CV* in Figure 1a and when comparing *PD* to *CD* in Figure 3. As shown in the figure, the median CV for each sample increases as we remove category distinctions. However, all IAP observations retain a significantly smaller CV than the IRP dotted line. This observation is present in all three category groupings (intra, inter, and all) and conclusively demonstrates that individual differences are retained through graph transformation for the SF task. Next, we consider the PSU-FF dataset. While the lack of intra-category data prevented us from performing the first two tests, the All-Category results for both Figure 1b and Figure 3 display the expected behavior of IAP variation being significantly smaller than the IRP. Finally, examining our cross-task CFA-SF dataset, we can see from Figure 1c that the difference is quite extreme. This is expected as we compare results from one task to another with different participants, but the confirmation is still important. Returning to **RQ1**, our analyses shows that IAP graph features are significantly more similar to other features generated by that participant than to other participants. This shows that not only are graph features able to be successfully transformed from lists of words for the CFA and SF tasks, but that individual differences are retained as well.

RQ 2: Is there evidence for three distinctive exploration styles in associative word tasks? Recall that the three exploration styles (busybody, hunter, and dancer) are differentiated by two measures: BB-H and FF scores. We can use the earlier measurement for cross-task, IAP variance (Figure 1c), except instead of performing the calculation over graph features, we use BB-H and FF-scores. If cross-task, IAP variance here is lower than cross-task, all-participant variance, it would suggest that individuals' exploration style do have some consistency across the CFA and SF tasks. Indeed, Figure 2c shows this is the case. The IAP CV for the SF task ranges from 25% to 60% smaller than for IRP, with a more modest 12% for the CFA task and a 73% difference across the CFA and SF tasks. These are all strong evidence for the individual nature of the BB-H-D exploration styles that participants utilize when performing the CFA or SF tasks.

Conclusion

This work explored how graph transformations of the CFA and SF tasks retain individual participant information and the application of the BB-H-D score to those tasks, and we collected a new cross-task dataset to do so. We confirmed that graph features preserve IAP similarities and IRP differences on the CFA and SF tasks. We also showed the existence of IAP similarities across these tasks.Our paper strengthens the argument for the existence of the three exploration styles—busybody, hunter, and dancer—by demonstrating their IAP persistence across tasks. However, as the CFA and SF tasks are similar word association tasks, future work should expand on these findings by exploring additional tasks and examining how external factors influence semantic exploration patterns. Additionally, datasets with large cross-task lists from the same participants would enable more detailed cross-task graph analyses, including comparisons of global structure (e.g., density, modularity), local motifs (e.g., triads, bridges), and node-level roles. Such datasets would also allow for embedding-based similarity measures to evaluate whether individuals' semantic exploration graphs retain consistent topological features across different task types.

Acknowledgments

Part of this research was sponsored by the DEVCOM Analysis Center and was accomplished under Cooperative Agreement Number W911NF-22-2-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

Abdi, H. (2010). Coefficient of variation. *Encyclopedia of research design*, *1*(5), 169–171.

Anderson, J. R., & Bower, G. H. (2014). *Human associative memory*. Psychology press.

Beaty, R. E., & Kenett, Y. N. (2023). Associative thinking at the core of creativity. *Penn State*.

Beaty, R. E., Zeitlen, D. C., Baker, B. S., & Kenett, Y. N. (2021). Forward flow and creative thought: Assessing associative cognition and its role in divergent thinking. *Thinking Skills and Creativity*, *41*, 100859.

Clarke, K. C. (1986). Computation of the fractal dimension of topographic surfaces using the triangular prism surface area method. *Computers & Geosciences*, *12*(5), 713–722.

Gourovitch, M. L., Goldberg, T. E., & Weinberger, D. R. (1996). Verbal fluency deficits in patients with schizophrenia: Semantic fluency is differentially impaired as compared with phonologic fluency. *Neuropsychology*, *10*(4), 573.

Gray, K., Anderson, S., Chen, E. E., Kelly, J. M., Christian, M. S., Patrick, J., ... Lewis, K. (2019). "forward flow": A new measure to quantify free thought and predict creativity. *American Psychologist*, *74*(5), 539.

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using networkx. In *Proceedings of the python in science conference* (pp. 11–15).

Harrison, J. E., Buxton, P., Husain, M., & Wise, R. (2000). Short test of semantic and phonological fluency: Normal performance, validity and test-retest reliability. *British Journal of Clinical Psychology*, *39*(2), 181–191.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological review*, *119*(2), 431.

Kenett, Y. N. (2024). The role of knowledge in creative thinking. *Creativity Research Journal*, 1–8.

Kenett, Y. N., Humphries, S., & Chatterjee, A. (2023). A thirst for knowledge: Grounding curiosity, creativity, and aesthetics in memory and reward neural systems. *Creativity Research Journal*, *35*(3), 412–426.

Koivisto, M., & Toivanen, H. (2024). Schizotypy and creativity: Divergent thinking, inhibitory control, and the spontaneous flow of thought. *Creativity Research Journal*, 1–21.

Litman, J. A., & Silvia, P. J. (2006). The latent structure of trait curiosity: Evidence for interest and deprivation curiosity dimensions. *Journal of personality assessment*, *86*(3), 318–328.

Lydon-Staley, D. M., Zhou, D., Blevins, A. S., Zurn, P., & Bassett, D. S. (2021). Hunters, busybodies and the knowledge network building associated with deprivation curiosity. *Nature human behaviour*, *5*(3), 327–336.

Marron, T. R., Berant, E., Axelrod, V., & Faust, M. (2020). Spontaneous cognition and its relationship to human creativity: A functional connectivity study involving a chain free association task. *NeuroImage*, *220*, 117064.

Marron, T. R., Lerner, Y., Berant, E., Kinreich, S., Shapira-Lichter, I., Hendler, T., & Faust, M. (2018). Chain free association, creativity, and the default mode network. *Neuropsychologia*, *118*, 40–58.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Paschoal, A., da Silva, P. H. R., Rondinoni, C., Arrigo, I. V., Paiva, F. F., & Leoni, R. (2021). Semantic verbal fluency brain network: delineating a physiological basis for the functional hubs using dual-echo asl and graph theory approach. *Journal of Neural Engineering*, *18*. doi: 10.1088/1741-2552/ac0864

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? predictors of verbal fluency performance in older adults. *Frontiers in psychology*, *5*, 772.

Siew, C. S., & Guru, A. (2023). Investigating the network structure of domain-specific knowledge using the semantic fluency task. *Memory & cognition*, *51*(3), 623–646.

TR, M., Y, L., Berant E, K. S., I, S.-L., T, H., & M., F. (2018). Chain free association, creativity, and the default mode network. *Neuropsychologia*.

Weisstein, E. W. (2003). Heron's formula. *https://mathworld.wolfram.com/*.

Zarino, B., Crespi, M., Launi, M., & Casarotti, A. (2014). A new standardization of semantic verbal fluency test. *Neurological Sciences*, *35*, 1405–1411.

Zemla, J. C., Cao, K., Mueller, K. D., & Austerweil, J. L. (2020). Snafu: The semantic network and fluency utility. *Behavior research methods*, *52*, 1681–1699.

Zhou, D., Patankar, S., Lydon-Staley, D. M., Zurn, P., Gerlach, M., & Bassett, D. S. (2024). Architectural styles of curiosity in global wikipedia mobile app readership. *Science advances*, *10*(43), eadn3268.

Zurn, P. (2019). Busybody, hunter, dancer: Three historical models of curiosity. In M. Papastefanou (Ed.), *Toward new philosophical explorations of the desire to know: Just curious about curiosity* (pp. 26–49).