Title
Dynamics of creative exploration: A study of signal space constraints in music and communication

Permalink
https://escholarship.org/uc/item/20h9h892

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Mudd, Katie
Youngblood, Mason
schedel, margaret

Publication Date
2026

Dynamics of creative exploration: A study of signal space constraints in music and communication

Katie Mudd (katherine.mudd@stonybrook.edu)[1], Mason Youngblood[2] & Margaret Schedel[2]
[1]Institute for Advanced Computational Science, Stony Brook University, USA

Abstract

Creative exploration is shaped by the functional pressures placed on signals. Yet we know little about how those pressures structure exploration, or how they connect to individual creativity and task constraints. Using a rhythm production task, we investigate how signal space exploration is affected by constraints in a communicative or musical act, as well as the presence of feedback. We quantify exploration using three measures: distance, novelty and syncopation, following our preregistered analysis. Rhythms produced in the music task were more novel and more syncopated than those produced in the communicative task, suggesting a communicative pressure towards more predictable structures. Feedback reduced novelty and syncopation, while increasing distance. These findings reveal how functional pressures shape signals and call for more research on how collective creativity and environmental constraints shape cultural acts across domains.

Keywords: creativity; music; communication; experimental semiotics; environmental constraints

Introduction

Creativity is considered to be a defining feature of human cognition, but its expression may depend on the cognitive domain in which it unfolds. In this paper, we experimentally investigate how different constraints on the production of musical and communicative acts shape creative behavior using a rhythm production experiment.

While early research on creativity focused on a handful of prominent individuals who excelled in particular domains (e.g., Van Gogh and Picasso, Ghiselin 1952), in reality, all of us engage in creative acts for everyday problem-solving (Lewis & Piekut, 2016). Here, we define a creative act as one that is novel and useful, following (Boden, 2004). For example, a new sound can be useful in that it is aesthetically pleasing and a new sentence can be useful in that it conveys a meaning.

Empirical research has long recognized creativity as being comprised of discrete abilities, and many different categorizations of creativity have been made (see Table 1 in Fogarty et al. 2015 for an overview). The often used classification by Guilford (1967) distinguishes divergent creativity from convergent creativity: divergent creativity refers to the generation of multiple, novel responses, while convergent creativity refers to the focused identification of the best solution to a problem (Guilford, 1967). Here, we focus on divergent creativity to investigate how individuals explore a given space in different domains, as this type of creativity can be conceptualized as exploring problem-solving options within a space (Newell et al., 1958).

Formal theories have articulated creativity as search under constraints, where novelty arises through the exploration of allowable transformations within a defined space (Wiggins, 2006). As such, the idea that specific constraints may actually enhance creativity is not new: classic work on problem solving demonstrates that creative breakthroughs often occur not through exhaustive search, but through changes in representation that relax or reinterpret constraints (Kaplan & Simon, 1990). Building on this, Ohlsson (1992) formalized this idea by proposing that creative impasses arise from self-imposed constraints embedded in problem representations, and that insights (aha moments!) emerge when these constraints are overcome, allowing previously unavailable solutions to be considered (e.g., as shown in a study of the success of constraint-based brainstorming Gould et al., 2023). Together, these findings support the idea that creativity may be enhanced by constraints.

A growing body of work in experimental semiotics has investigated how different constraints on a signal space affect structure formation. Across both communicative and non-communicative domains, these studies typically vary the size, dimensionality or topology of the available signal space in order to examine how constraints shape emergent structure (Little & de Boer, 2014; Roberts & Galantucci, 2012; Verhoef et al., 2014). Iterated learning and communication paradigms have shown that constraining the form or meaning space can promote regularity and compositionality in emergent signaling systems, even when individual productions are initially variable. For example, a model by Brighton and Kirby (2001) demonstrates that the structure of the meaning space itself strongly determines whether compositional languages emerge and persist. Similarly, Carr et al. (2017) show experimentally that when participants communicate in an open-ended and continuous signal space, structured languages can evolve through repeated interaction, with constraints shaping how signals are discretized and reused. These studies focused predominantly on explaining the emergence and evolution of linguistic structure, but how might constraints differ across domains?

Evidence for the domain-dependent role of constraints in creative cognition comes from Tchernichovski et al. (2025), who compared the effects of restricting creative action across

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

two domains: visual (creating a pixel drawing) and auditory (creating a melody on a step synthesizer). Participants were assigned to transmission chains: the first participant in the chain created a drawing or a melody and subsequent participants modified the artifact from the member of the chain before them. In one manipulation relevant to the present study, participants were restricted in the number of changes they could make to pixels in the drawing condition and to musical notes in the music condition. Crucially, melodies that were produced in this constrained condition were on average rated as more pleasant than their unrestricted counterparts. However, in the visual condition, restricting the number of allowed changes actually reduced pleasantness, suggesting that domain differences may exist with regards to creative constraints.

This body of work suggests that creativity may be enhanced by constraints that structure a signal space, and that may differ systematically across domains. However, no relationship has yet to be proposed between individual-level creativity and signal space constraints across domains that differ in their functional demands. Addressing this gap requires an experimental approach that allows creative behavior to be measured as structured exploration while holding task demands comparable across domains. In the current study, we adopt such an approach to examine how constraints on signal production shape exploratory behavior and creative outcomes in the domains of communication (where signals convey meaning) and music (where signals carry aesthetic value).

The current study: signal space constraints and exploration across domains

In this project, we adopt a functional definition of creativity as the production of acts that are both novel and useful (adapted from Fogarty et al. 2015). Under this definition, usefulness is domain specific; for example, a communicative signal is useful if it successfully conveys meaning, whereas a musical signal is useful if it elicits perceptual or aesthetic reward. Framing creativity in terms of signal space exploration enables the comparison of signals produced across domains. Music and communication are particularly well suited for such comparison as they both rely on structured signal systems and shared conventions, as well as learned mappings between form and meaning (communication) or aesthetics (music). Crucially, both domains allow individuals to exhibit creativity in signal production, but norms do exist.

Here, we present an empirical study in which we investigate how signal space exploration is affected by individual creativity, signal space constraints, the presence of feedback and differences across two domains: communication (where signals refer to a meaning) and music (where signals have no required meaning). The two domain conditions are meant to be a coarse approximation of the difference in pressures present for communication vs. music. We hypothesize that more creative participants will be more prone to signal space exploration, and that constraints will accentuate this exploration. By examining how constraints and feedback shape exploratory signal production across domains, this work aims to clarify how creative behavior emerges from the interaction of individual cognitive tendencies and the structured environments in which cultural artifacts are produced.

Method

This study received approval from the ethics committee of Stony Brook University. Prior to taking part in the study, participants received information about the experiment and gave their consent.

Participants Participants were recruited via the online platform Prolific and were eligible to participate if they had a Prolific approval rating of 95% and higher on previous experiments and reported being fluent in English, ensuring comprehension of the task instructions. Participants were compensated at a rate of $16 per hour.

A total of 690 participants were recruited and were assigned consecutively to the combination of these conditions (conditions are elaborated on in Section Conditions and materials): 1. The music domain with no feedback, 2. The communication domain with no feedback, 3. The communication condition with feedback and 4. The music domain with feedback. Each combination of domain and feedback included three different time constraints. Batches of ten participants were recruited until each of the three time constraint conditions per grouping reached at least 50 participants.

Conditions and materials Participants were asked to create rhythms on a drum machine. Each rhythm is a 1-bar loop with 8 steps corresponding to 1/8th notes. Users have a grid where they can turn on or off sounds, where each row is a different sound (hi-hat, snare or kick), as shown in Figure 1.

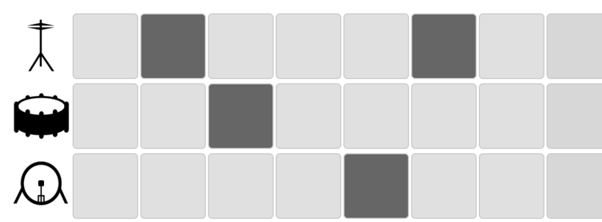

Figure 1: An example of a possible rhythm production on the drum machine.

In addition, participants are assigned to either the music or communication domain; in the music domain, participants are asked to "create an appealing rhythm" using the drum machine. In the communication condition, participants are shown a color and are asked to "create a rhythm to describe this color" using the drum machine. There are nine colors: blue, green, yellow, orange, red, pink, purple, brown, grey. Figure 2 shows examples of the music and communication task.

Further, participants are either given feedback on their rhythm production or not. In conditions where participants are given feedback, participants are told that they will be paired up with a partner and have the role of either the producer, producing rhythms on the drum machine, or the receiver, giving feedback on the rhythms produced by their partner. In the music condition, participants are told that the receiver will choose if the rhythm is appealing or not appealing. In the communication condition, participants are told that the receiver will guess which color they were referring to. In actuality, the participant is always the producer, and feedback is not given by another participant, but rather is determined by the following formula: probability of success = $\frac{1}{10-\text{trial number}}$. As such, the probability of success increases over the rounds, finally reaching 1.0 in the final round. In the music conditions, after having produced a rhythm on the drum machine, participants are told "Success! Your partner found your rhythm appealing" if the probability of success is larger than a randomly generated number between 0 and 1. Otherwise, participants are told "Unsuccessful! Your partner did not find your rhythm appealing." In the communication condition, after having produced a rhythm on the drum machine to refer to a color, participants are told "Success! Your partner guessed the color you were referring to" if the probability of success is larger than a randomly generated number between 0 and 1. Otherwise, participants are told "Unsuccessful! Your partner did not guess the color you were referring to."

Finally, participants were randomly assigned to one of three time constraint conditions: participants were given 15, 30 or 45 seconds to create rhythms.

In sum, there were 12 conditions total, combining domain (communication or music) with the presence of feedback (True or False) and time constraint (15, 30 or 45 seconds). Table 1 shows the number of participants assigned to each condition in the studies without feedback and with feedback.

Table 1: Number of participants by domain, feedback, and time constraint

Domain	Feedback	Time constraint (s)		
		15	30	45
Music	False	70	52	57
Communication	False	51	50	65
Music	True	56	57	51
Communication	True	62	49	70

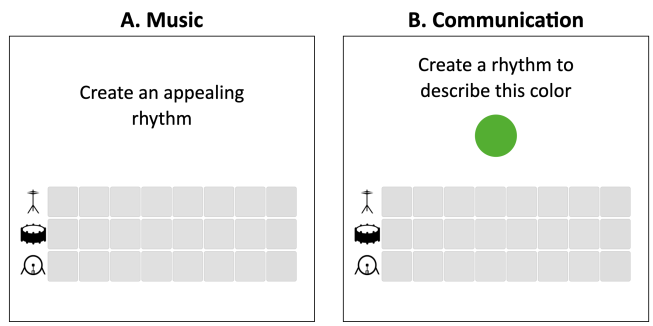

Figure 2: A. An example of the music task with three drum sounds. B. An example of the communication task with three drum sounds

Procedure The experiment was programmed in Python using the PsyNet package (Harrison et al., 2020).

After reading an information brochure and signing a consent form, participants completed a number of pre-tests: a color blindness test and an audio check. All participants passed these checks.

Following this, participants completed a measure of divergent creativity, the Divergent Association Task (DAT), in which participants are asked to "generate 10 nouns (single words) that are as different from each other as possible in all meanings and uses of the words" (Olson et al., 2021). This task has been shown to correlate with other measures of divergent creativity, for instance, with the Alternative Uses Task and the Bridge-the-Associative-Gap Task.

Participants then had a one minute familiarization phase with the drum machine. Next, participants took part in the rhythm creation task over nine trials. In the communication conditions, one of the nine colors is shown per trial (example shown in Figure 2).

Finally, participants completed a brief questionnaire about their linguistic and musical experience.

Analysis Preregistration materials are available on the Open Science Framework: https://osf.io/nbgj7.

To quantitatively assess the divergent creativity of participants, we followed the analysis of Olson et al. (2021) using the list of words produced by participants in response to the DAT. With the first seven valid words (out of ten) that participants provided, semantic distances were computed using the GloVe algorithm (Pennington et al., 2014) between all pairs of the seven words for each participant. The DAT score is the transformed average of the semantic distances per participant.

From the rhythms that participants produced, we compute three measures: the Earth mover's distance, the inverse probability (novelty) and the syncopation index. First, to measure signal space exploration, distances between rhythms are quantified by calculating the Earth mover's distance between consecutive pairs of rhythms produced by each participant (e.g., calculating the distance between rhythm 1 and 2, 2 and 3, etc.), yielding 8 distances per participant. Each sound is calculated separately and then averaged across the two or three sounds, depending on the participant's condition. More precisely, the Earth mover's distance measures how much effort it takes to transform one signal into another. In this way, this distance measure is more appropriate for the study of signal distances than the commonly used Hamming distance, because it captures the similarity between phase shifts (rhythms shifted right or left in position).

Second, to measure the novelty of a rhythm, the inverse probability is measured, capturing how rare each produced rhythm is relative to the distribution of rhythms in the dataset. This was measured by first computing kernel density distributions of rhythmic patterns for each sound (hi-hat, snare and kick) across participants. The novelty of each produced rhythm was computed by summing the negative log-likelihoods of its drum patterns and normalizing by the number of keys, such that rare rhythms were assigned higher novelty values.

Third, to measure the degree of variation within a rhythm, the syncopation index of a rhythm is calculated. In this case, syncopation is a better measure than entropy because it reflects where rhythmic events occur relative to listeners' expectations, whereas entropy only measures overall unpredictability without regard to rhythmic structure (Witek et al., 2014). Syncopation is computed at the level of individually produced rhythms, with instrument-specific syncopation scores (hi-hat, snare and kick) averaged to yield a single trial-level score for each participant.

In addition to analyzing the rhythm produced by participants, we also collected all of the rhythms produced prior to the final rhythm. As an additional measure of exploration, we consider the number of clicks per trial that participants produced.

Models were fit in a Bayesian multivariate framework using Markov chain Monte Carlo (MCMC) sampling in the brms package in R (Bürkner, 2017). We jointly modeled three outcomes (distance, novelty and syncopation) as functions of domain (music vs. communication), time (15 vs. 30 vs. 45 seconds), feedback (no vs. yes), number of clicks per trial, creativity (DAT score) and musicality. In the same multivariate model, number of clicks per trial was included as an additional outcome and was modeled as a function of time, creativity (DAT score) and musicality. All outcome equations included random intercepts for participant and trial number. Model comparison was performed using approximate leave-one-out cross-validation (Vehtari et al., 2017), comparing models by expected log predictive density (ELPD), indicating the model fit. In the results section, the model with the highest ELPD is presented.

Results

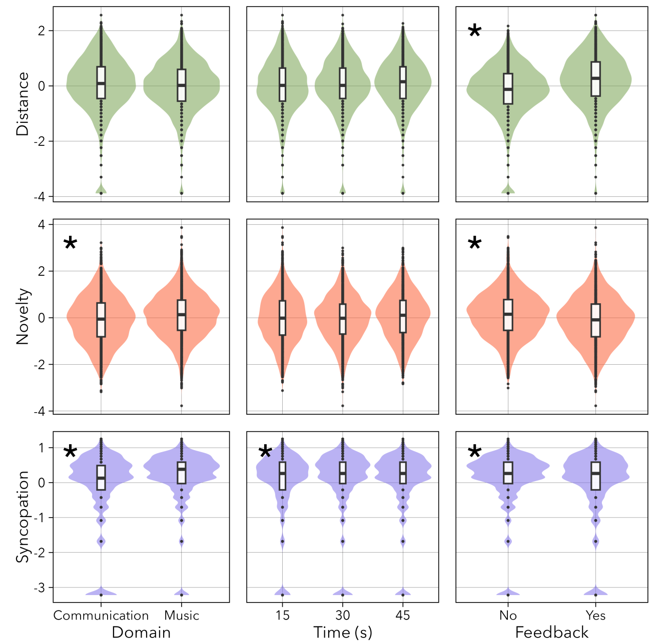

Figure 3: A dotplot showing the Earth mover's distance (in green), novelty (in red) and the syncopation index (in purple). Data is shown by domain (communication vs. music), time (15 vs. 30 vs. 45 seconds) and feedback (no vs. yes). Novelty and syncopation are significantly higher in the music conditions. Without feedback, the novelty and syncopation index are significantly higher, while distance is higher with feedback.

Thirteen participants were excluded because they did not provide at least seven valid words in response to the DAT. On average, the DAT score across participants is 77.6 (SD = 6.94). The lowest DAT score in the sample is 42.76, with these first seven valid words: *happy, cheerful, joyful, delighted, pleased, glad* and *thrilled*. The highest DAT score in the sample is 94.95 with the first seven valid words: *table, professional, headscarf, mangetout, cloud, highway* and *cruise*.

To examine how task constraints and domain shaped rhythm exploration and structure, we found that the model with the highest ELPD to predict the distance, novelty and syncopation included: the three outcomes (distance, novelty and syncopation) as functions of domain (music vs. communication), time (15 vs. 30 vs. 45 seconds), feedback (no vs. yes), number of clicks per trial and musicality. In the same multivariate model, number of clicks per trial was included as an additional outcome and was modeled as a function of time and musicality. All models included random intercepts for participant and trial number. The results of this model can be seen in Figure 3. Notably, including DAT score lowered model fit and as such it is not in the presented model.

For distance, the model revealed a significant effect of feedback: rhythms produced in conditions with feedback exhibited

Table 2: Model estimates for Earth mover's distance, novelty (inverse probability), syncopation and trial exploration

Outcome	Predictor	Estimate	L-95%	U-95%	Sig.
Distance	Domain (music)	-0.034	-0.129	0.061	
Distance	Time condition	0.038	-0.012	0.090	
Distance	Feedback (True)	0.397	0.301	0.492	*
Distance	Number of clicks per trial	-0.036	-0.102	0.028	
Distance	Musicality	-0.097	-0.144	-0.052	*
Novelty	Domain (music)	0.199	0.111	0.282	*
Novelty	Time condition	0.008	-0.035	0.053	
Novelty	Feedback (True)	-0.233	-0.315	-0.149	*
Novelty	Number of clicks per trial	0.185	0.127	0.243	*
Novelty	Musicality	0.050	0.009	0.092	*
Syncopation	Domain (music)	0.298	0.209	0.387	*
Syncopation	Time condition	0.098	0.051	0.145	*
Syncopation	Feedback (True)	-0.133	-0.224	-0.042	*
Syncopation	Number of clicks per trial	-0.231	-0.290	-0.170	*
Syncopation	Musicality	-0.052	-0.099	-0.007	*
Number of clicks per trial	Intercept	-0.017	-0.096	0.062	
Number of clicks per trial	Time condition	0.219	0.161	0.279	*
Number of clicks per trial	Musicality	0.028	-0.031	0.085	

greater distances between successive rhythms than those produced in conditions without feedback. For novelty, rhythms produced in the music domain were more novel than those produced in communication, the presence of feedback reduced novelty, novelty increased with the number of clicks per trial and novelty increased with musicality. For syncopation, rhythms produced in the music domain were more syncopated than in communication, syncopation increased with time, syncopation decreased with feedback, decreased with trial exploration and also decreased with musicality. Finally, time had a significant effect on the number of clicks per trial: more time leads to more clicks.

In addition, we explored how participants respond to positive and negative feedback: do participants explore more (i.e., increase in distance) following negative feedback? To explore this, we subsetted the six conditions with feedback. For each trial, we noted feedback success (True or False depending on probability of success) and the distance between the trial and the next trial. In line with our prediction, we found that the distance after negative feedback was higher (M = 1.07, SD = 0.607) compared to positive feedback (M = 1.04, SD = 0.609). We fit a Bayesian model predicting distance to the next trial from feedback success (True vs. False), domain (music vs. communication), and time (15 vs. 30. vs. 45), with random intercepts for participant and trial number. We found a negative relationship between feedback success and distance ($b = -0.070$, 95% CI $[-0.126 - 0.015]$), confirming that participants exhibit larger distances with negative feedback. Domain and time do not significantly predict distance.

In a second exploratory analysis, we investigated the correlation between rhythm distances and color distances. In designing the experiment, all participants were assigned a color per trial; in the communication domain, participants were shown this color and in the music domain participants were not. We hypothesized that for the communication domain, there would be a stronger correlation between colors and rhythms, reflecting a structured mapping, compared to the music domain where no correlation is expected. Pairwise color distances were calculated using LAB coordinates and rhythm Earth mover's distances were calculated between all pairs of rhythms produced per participant. A Mantel test between the color and rhythm distances was then calculated per participant determining the Pearson correlation between the distance matrices. We find a high degree of variation in correlation across participants, with the mean for music higher (M = -0.016, SD = 0.191) than for communication (M = -0.063, SD = 0.20), though both reflect a slightly negative relationship between color and rhythm distances. However, it's worth noting that significant correlations were always positive, even if they only occurred in a subset of fourteen participants (9 in the communication domain and 5 in the music domain).

Discussion

In our study of constraints on a rhythm task, exploration differed by domain: communicative contexts reduced novelty and syncopation relative to music. Previous work suggests that communicative contexts favor distinguishable signals for referring to meanings (Gibson et al., 2019), leading to movement through signal space but not per se complex or highly novel signals. Related to this, Carr et al. (2020) showed that when linguistic signals must refer reliably, systems compress and regularize, perhaps offering an explanation for the lack of novelty in our experiment's communication signals. In contrast, there are different pressures on music, such as to be engaging compared to other signals and have a favorable structure. This result supports the view that aesthetic domains tolerate (or reward) structural complexity and deviation.

Interestingly, (Tchernichovski et al., 2025) found an interaction between experimental constraints and domain (visual vs. music) in the pleasantness ratings of produced drawings and melodies in their study: melodies produced in a constrained condition were rated as more pleasant, but in the visual condition, restricting the number of allowed changes actually reduced pleasantness. In the experiment presented here, an effect of the time constraint was only observed in relation to syncopation: the more time that participants had, the more syncopated rhythms were. At present, we have not investigated interactions between time and domains, but plan to do so in future work. We did not observe an effect of time on distance or novelty; one possibility is that varying time in this task did not impose the kind of constraint that reliably promotes creative restructuring, as suggested in work on change in creative acts.

Surprisingly, participant's divergent creativity scores (measured by DAT) reduced model fit. Although future work will further investigate the relationship between divergent creativity and signal space measurements, several explanations may account for the lack of influence of DAT score on signal space exploration: first, DAT measures ideational fluency and flexibility, correlating with open-ended semantic tasks (Olson et al., 2021). In contrast, our experiment had a limited signal space and a goal (to produce an appealing rhythm in the music condition or to have a partner choose the color the director referred to in the communication condition). It is possible that a relationship between divergent creativity and signal space measures will be apparent when participants need to explore ways to solve a problem, for example in a communication task with a real partner who needs to guess a color until correct.

Creative acts do not exist in a vacuum and culture is rarely the product of one individual (Mesoudi & Thornton, 2018), supporting the idea that collective creativity is fundamental to human culture (Giuffre, 2016). Music is often made for an audience and communicative acts usually include at least two individuals. As a step in this direction, our feedback manipulation simulated the presence of an interlocutor. Interestingly, we found that the presence of feedback on rhythm production reduced novelty and syncopation, and increased distance. One open question is how participants interpreted the feedback: did they believe they were receiving feedback from another participant and did they adapt to it? Our findings show that participants modified their behavior depending on the feedback they received: directly following negative feedback, participants explore the rhythm space more, suggesting that they are moving away from their 'failed' rhythm towards a rhythm that their partner will find appealing (in music) or will lead to communicative success (in communication).

A natural next step is to include a human interlocutor, who gives the director feedback. Experimental semiotics has shown that different director–matcher setups affect how signaling systems are grounded (Delliponti et al., 2023; Galantucci & Garrod, 2011); this pattern aligns with grounding theory, which emphasizes the role of interactional constraints (Clark & Brennan, 1991). With regards to the evolution of linguistic signals, it has been shown that single participants alone tend to produce redundant, inefficient signaling systems (Motamedi et al., 2019). In our experiment, participants in the communication domain did not exhibit a stronger mapping between color and rhythm distances; it is possible that real interaction is necessary for this type of signal-meaning space structure to emerge. Building on dyadic interactions, computational models and experiments with participants provide evidence for within-generation interaction as a key driver in the emergence of systematic structure (Kirby et al., 2015; Raviv et al., 2019). Crucially, certain pressures promote the emergence of structure, such as inter-generational transmission (Kirby et al., 2015) or larger group sizes (Raviv et al., 2019). Finally, size and connectedness of populations may affect the diversity and complexity of cultural traits (Derex et al., 2013; Lutzenberger et al., 2023).

Conclusion

Our findings demonstrate that exploration in signal space is shaped more strongly by domain-level functional pressures than by individual creativity traits. In musical contexts, greater novelty and syncopation were found, supporting the view that aesthetic domains tolerate or reward structural deviation. Moreover, the increase of distance in the presence of feedback as well as the reduction in novelty and syncopation observed when feedback was introduced highlights how even minimal social input can alter creative behavior. Future work will move beyond testing isolated individuals towards richer interactive settings and modified task constraints, crucial for understanding how signal space exploration arises from the interplay between individual action and collective cultural processes.

Acknowledgments

Thanks to Manuel Anglada-Tort, Nori Jacoby and Susan Brennan for feedback on the experimental design. Thanks also to Thomas Oosterhuis, Nori Jacoby and Peter Harrison for their generous help with (PsyNet) programming questions.

References

Boden, M. A. (2004). *The creative mind: Myths and mechanisms* (2nd ed.). Routledge.

Brighton, H., & Kirby, S. (2001). Meaning Space Structure Determines the Stability of Culturally Evolved Compositional Language. *Cognitive Science Society conference*.

Bürkner, P.-C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Carr, J. W., Smith, K., Cornish, H., & Kirby, S. (2017). The Cultural Evolution of Structured Languages in an Open-Ended, Continuous World. *Cognitive Science*, *41*(4), 892–923. https://doi.org/10.1111/cogs.12371

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication.

Delliponti, A., Raia, R., Sanguedolce, G., Gutowski, A., Pleyer, M., Sibierska, M., Placiński, M., Żywiczyński, P., & Wacewicz, S. (2023). Experimental Semiotics: A Systematic Categorization of Experimental Studies on the Bootstrapping of Communication Systems. *Biosemiotics*, *16*(2), 291–310. https://doi.org/10.1007/s12304-023-09534-x

Derex, M., Beugin, M.-P., Godelle, B., & Raymond, M. (2013). Experimental evidence for the influence of group size on cultural complexity. *Nature*, *503*(7476), 389–391. https://doi.org/10.1038/nature12774

Fogarty, L., Creanza, N., & Feldman, M. W. (2015). Cultural Evolutionary Perspectives on Creativity and Human Innovation. *Trends in Ecology & Evolution*, *30*(12), 736–754. https://doi.org/10.1016/j.tree.2015.10.004

Galantucci, B., & Garrod, S. (2011). Experimental Semiotics: A Review. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00011

Ghiselin, B. (1952). *The Creative Process: A Symposium*. University of California Press.

Giuffre, K. (2016). *Collective creativity: Art and society in the south pacific*. Routledge.

Gould, R. K., Saito, T., Allen, K. E., Bonn, A., Chapman, M., Droz, L., Herrmann, T. M., Himes, A., Ishihara, H., Coelho-Junior, M. G., et al. (2023). Constraint breeds creativity: A brainstorming method to jumpstart out-of-the-box thinking for sustainability science.

Guilford, J. (1967). *The nature of human intelligence*. McGraw-Hill.

Harrison, P., Marjieh, R., Adolfi, F., van Rijn, P., Anglada-Tort, M., Tchernichovski, O., Larrouy-Maestri, P., & Jacoby, N. (2020). Gibbs Sampling with People. *Advances in Neural Information Processing Systems*, *33*, 10659–10671.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive psychology*, *22*(3), 374–419.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102. https://doi.org/10.1016/j.cognition.2015.03.016

Lewis, G., & Piekut, B. (2016). *The Oxford Handbook of Critical Improvisation Studies* [Google-Books-ID: mXzADAAAQBAJ]. Oxford University Press.

Little, H., & de Boer, B. (2014). The effect of size of articulation space on the emergence of combinatorial structure. *The Evolution of Language*, 479–480. https://doi.org/10.1142/9789814603638_0103

Lutzenberger, H., Mudd, K., Stamp, R., & Schembri, A. C. (2023). The social structure of signing communities and lexical variation: A cross-linguistic comparison of three unrelated sign languages. *Glossa: a journal of general linguistics*, *8*(1). https://doi.org/10.16995/glossa.10229

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B*, *285*(1880), 20180712.

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J., & Kirby, S. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition*, *192*, 103964. https://doi.org/10.1016/j.cognition.2019.05.001

Newell, A., Shaw, J. C., & Simon, H. A. (1958). The processes of creative thinking. *Contemporary Approaches to Creative Thinking*.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. *Advances in the psychology of thinking*, 1–44.

Olson, J. A., Nahas, J., Chmoulevitch, D., Cropper, S. J., & Webb, M. E. (2021). Naming unrelated words predicts creativity. *Proceedings of the National Academy of Sciences*, *118*(25), e2022340118. https://doi.org/10.1073/pnas.2022340118

Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global Vectors for Word Representation. In A. Moschitti, B. Pang, & W. Daelemans (Eds.), *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543). Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1162

Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Compositional structure can emerge without generational transmission. *Cognition*, *182*, 151–164. https://doi.org/10.1016/j.cognition.2018.09.010

Roberts, G., & Galantucci, B. (2012). The emergence of duality of patterning: Insights from the laboratory. *Language and Cognition*, *4*(4), 297–318. https://doi.org/10.1515/langcog-2012-0017

Tchernichovski, O., Globerson, E., Harrison, P., & Jacoby, N. (2025, February). How constraints on editing affects cultural evolution [arXiv:2502.16694 [q-bio]]. https://doi.org/10.48550/arXiv.2502.16694

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC [arXiv:1507.04544 [stat]]. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Verhoef, T., Kirby, S., & de Boer, B. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, *43*, 57–68. https://doi.org/10.1016/j.wocn.2014.02.005

Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, *19*(7), 449–458. https://doi.org/10.1016/j.knosys.2006.04.009

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, Body-Movement and Pleasure in Groove Music. *PLOS ONE*, *9*(4), e94446. https://doi.org/10.1371/journal.pone.0094446