Big Five, Big Influence: Personality-Driven Word Spread in Conversations

Akaran Sivakumar

202208344

School of Communication and Culture, Aarhus University

Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark

202208344@post.au.dk

Supervisor: Riccardo Fusaroli

January 7, 2025

Summary:

Prior research highlights the critical role of personality traits in novelty-seeking and language use. Traits such as Extraversion and Openness have been linked to curiosity, exploration, and adaptability in novel situations. These traits also influence linguistic behaviour, fostering lexical entrainment and the formation of conceptual pacts during conversations. Building on these findings, this study examines how personality traits influence the adoption of a novel word, "Glowchum", in collaborative interaction. The experiment utilized a Generative Agent-Based Model (GABM) in which five agents, each assigned a dominant Big Five personality trait (Agreeableness, Conscientiousness, Extraversion, Neuroticism, or Openness), completed a group collaboration task. The study addressed two research questions: (1) whether the dominant personality trait of the agent affects their likelihood of adopting a novel word and (2) whether the interaction between receiving and introducing agents' personality traits influences this likelihood. Agents high in Extraversion and Openness were most likely to adopt the novel word. Pairwise comparisons highlighted significant differences in word adoption between Extraversion and Conscientiousness, and Extraversion and Neuroticism. Interaction patterns revealed combinations of Extraversion-Conscientiousness amplified word adoption probabilities. Model comparison showed that the receiving agent's personality has a stronger effect than the introducing agent's personality.

Keywords - Generative Agents, ABM, Big-Five personality traits, Language evolution

Table of Contents

1. Introduction

1.1. Personality Traits

The five-factor model (FFM) is the most widespread approach for representing the human trait structure of today (Roccas et al., 2002). The model asserts that five basic factors describe most personality traits: neuroticism, openness to experience, extraversion, agreeableness, and conscientiousness. The OCEAN-model made popular by (Barrick & Mount, 1991) is used in industrial and organizational settings, counselling, and more. Traits are ”dimensions of individual differences in tendencies to show consistent patterns of thoughts, feelings, and actions” (P. T. Costa & McCrae, 1990). The FFM emerged from data-driven analyses of self-reports, peer evaluations, and personality questionnaires, identifying five robust factors (Goldberg, 1990; John, 1990; Lanning, 1994; Tupes & Christal, 1992).

High scorers in extraversion are sociable, talkative, assertive, and energetic, while low scorers are reserved and cautious. Extraversion is characterized by novelty, excitement, and assertiveness. Agreeable individuals tend to be cooperative, gentle, and modest, while low scorers are irritable, suspicious, and inflexible. Openness to experience reflects intellectual curiosity, imagination, and open-mindedness, while low scorers are conventional and practical. It aligns with a desire for novelty and stimulation (Roccas et al., 2002). Conscientious individuals are organized, responsible, and goal-oriented, whereas low scorers are disorganized and impulsive. Conscientiousness includes a proactive drive to achieve and the ability to inhibit impulsive behaviour (McCrae & John, 1992). High neuroticism is linked to anxiety, insecurity, and mood instability, while low neuroticism indicates emotional stability. (Bilsky & Schwartz, 1994) suggest that depression in high neuroticism may stem from unmet personal values. You can measure the five personality factors with the NEO Personality Inventory, which is a questionnaire with 180 questions divided into six subscales (McCrae & John, 1992; Roccas et al., 2002).

Some researchers hypothesize that there are even more basic factors that may underlie the five factors. (Digman, 1997) suggested two higher-order factors. Factor α (Agreeableness, Conscientiousness, and Emotional Stability) linked to socialization success, and Factor β (Extraversion and Openness) related to self-actualization and growth. Personality traits are largely immune to influence by any environmental variable or personal strivings, values, and attitudes. This split of Factor α and Factor β can also be seen as novelty seeking emerges as another practical expression of Extraversion and Openness.

Novelty seeking, defined as "the tendency of humans and animals to explore novel and unfamiliar stimuli and environments" (A. Costa et al., 2014), is a behavioural tendency that plays a significant role in exploration and adaptability. This trait has been linked to greater sensation-seeking, impulsivity, risk-taking, and a strong orientation toward independence. Importantly, novelty seeking is also closely associated with the Big Five personality traits of Extraversion and Openness to Experience (Gordon & Luo, 2011). These two traits are central to creative and adaptive behaviours, as they are rooted in dopaminergic brain functioning, which supports exploration and engagement with new stimuli (Ebstein et al., 1996; Shiner & Deyoung, 2013). The connection between novelty seeking and the Big Five is further supported by research on plasticity, which shows a shared variance between Extraversion and Openness. Extraverts are more likely to engage with novelty in a physical or behavioural sense, while individuals high in Openness tend to explore abstractly, by altering mental categories and reconceptualizing existing ideas (DeYoung et al., 2002) Together, these findings underscore the unique and complementary ways in which Extraversion and Openness contribute to novelty-seeking behaviours, offering critical insights into how personality influences creative and adaptive responses to novel situations.

Language learning and evolution are driven by pressures such as communicative success, production effort, and learnability, which shape linguistic structure over time (Chater & Christiansen, 2010). In emergent communication, agents develop protocols through interaction, guided by inductive biases like “laziness” (penalizing long messages) and “impatience” (encouraging early inference) (Galke & Raviv, 2024b). When people repeatedly refer to the same object or concept, they establish conceptual pacts, a temporary agreement on how to conceptualize the object, which results in lexical entrainment, which is when the same or related terms are consistently reused (Clark & Wilkes-Gibbs, 1986). Lexical choices are shaped by recency, frequency of use, and partner specificity (Garrod & Doherty, 1994). People tend to rely on previously established pacts with familiar partners, while adopting simpler or modified references with new ones, highlighting the role of interaction history in shaping language use (Brennan & Clark, n.d.; Wilkes-Gibbs & Clark, 1992). Even a small, biased minority (10% of the population) can shift the language of the unbiased majority, resulting in a hybrid language. Network structure influences the speed of language stabilization (Raviv et al., 2020). If the network is structured so the highly influential agents are the biased minority, then linguistic differences will be amplified (Josserand et al., 2021).

1.2. Agent-Based Modelling

This paper uses a subcategory of a method of research called Agent-Based Modelling (ABM). ABMs study interactions between simulated beings (agents), by creating a simulated digital environment in which hyperparameters are set to most effectively reflect the real-world environment that is being investigated (Salgado & Gilbert, n.d.). ABMs, typically used for simulating highly complex environments with many agents, were famously used to study the effect of different regulations (hyperparameters) on the transmission rates between agents during the COVID-19 pandemic (Kerr et al., 2021). (Hammond, 2015) outlines three main strengths of the ABM: heterogeneity, spatial structure, adaption and coevolution. First, their ability to simulate diverse agents eliminates the need for aggregation or a "representative agent". Second, ABMs can incorporate detailed geographic and social network data, which enables the modelling of real-world dynamics like disease spread. Lastly, they can simulate individual changes (e.g. skill learning or weight changes) and the evolution of behaviours, used to understand feedback loops in complex systems.

Recent years have seen significant advancements in pre-trained language models (PLMs), which use deep-learning transformer architecture to generate natural language (Vaswani et al., 2017). When scaled up, these models become large language models (LLMs), capable of predicting words based on context (Zhao et al., 2023). LLMs have revolutionized natural language processing (NLP) and are increasingly used in social sciences to mimic human behaviour (Devlin et al., 2019; Grootendorst, 2022). This reduces reliance on human participants, cuts costs, and improves scalability(Wang et al., 2024).

By generating agents from these LLMs and placing them in an environment where they can communicate with each other, we are coming closer to generating autonomous agents who can act independently and solve highly complex tasks. At the current state of LLMs, these techniques are not automatically deployed but are still achievable by using an agentic workflow where LLM-based agents are assigned different roles in a team, performing together to iterate on the best possible solution for a given problem using complex reasoning (Wang et al., 2024). Though at a very early stage, people have combined the capabilities of these LLM agents and the strengths of ABMs to create Generative Agent-Based Modelling (GABM). A GABM is an

ABM that can produce natural language based on a set of hyperparameters which often involves prompting LLM agents as the hyperparameters for the model (Wang et al., 2024).

The most famous of these GABMs is the Generative Agents framework developed by a set of researchers at Stanford and Google (Park et al., 2023). This framework showed that LLM agents could be part of a simulated society with jobs, routines, and relationships. The agents were also able to form new relationships and remember interactions with other agents. Similar achievements were also seen in GABMs simulating World War scenarios, A Public Administration Crisis, and Epidemic Modelling (Hua et al., 2023; Williams et al., 2023; Xiao et al., 2023). People have also begun using generative agents as virtual assistants or study buddies (*CrewAI*, n.d.; *LangChain*, n.d.; Hong et al., 2023). Both humans and language models can learn from linguistic input, but humans are guided by cognitive pressures such as communicative success and efficiency (Galke & Raviv, 2024a). Unlike humans, who acquire language through interaction and social transmission, language models are exposed to vast amounts of textual input and lack multimodal learning experiences (Seyssel et al., 2024; Warstadt & Bowman, 2022).

1.3. Thesis Statement

This study aims to investigate the potential applications of GABM in understanding the impact of the "Big-Five" personalities on the adoption of novel words/ideas, and its potential outcome on the evolution of language, in cases where the minority has a different vocabulary than the majority.

RQ1: Does the dominant personality trait of an agent affect their propensity to pick up a new word?

RQ2: Does altering the dominant personality of the agent introducing the word influence the receiving agent's propensity to pick up a new word?

H_1: Agents with dominant personality traits characterized by adventure-seeking, excitement, and novelty will pick up the new word more often than other agents.

H_2: Agents characterized by being accommodating, cooperative, and ambitious will have more success at spreading novel words.

2. Methods

2.1. Participants & Experiment Design

This study is inspired by the generative agents simulation created at Stanford (Park et al., 2023). This study used simulated agents as participants, each embodying one of the Big Five personality traits: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness. The agents were implemented using OpenAI's GPT-4o-mini model (*GPT-4o Mini*, n.d.) and they interacted with each other in a controlled, turn-based environment. Each simulation consisted of five agents, one for each personality trait. These traits were randomly assigned for each simulation to avoid order effects or biases. This ensured that approximately 400 simulations were conducted for each receiving agent personality trait.

Each agent's dominant trait was operationalized through specific behavioural descriptors embedded in their prompts, as follows:

- **High Agreeableness**: Cooperative, trusting, and collaborative.
- **High Conscientiousness**: Organized, meticulous, and reliable.
- **High Extraversion**: Outgoing, engaging, and thriving in social interactions.
- **High Neuroticism**: Tendency toward anxiety, self-doubt, and worry.
- **High Openness**: Imaginative, Curious, and focused on exploring ideas.

Agents received one of these "high-level" descriptors and four baseline descriptors for the other personality traits. These prompts guided agent behaviour during the task, and agents received specific trait-based instructions when it was their turn.

The task designed for this study was a collaborative planning exercise, in which the five agents worked together to create a detailed six-hour birthday party schedule. The task served as a structured environment to observe conversational dynamics. The goal was to finalize a structured plan that included a seamless flow of activities. The roleplay-style of this task was used as prior research suggests LLM agents can immerse themselves better when the environment and their personality are properly defined (Giray, 2023).

One agent, referred to as the introducing agent, was tasked with incorporating the target word "Glowchum" into the conversation during their turn. The following agent, referred to as the receiving agent, was the first agent to encounter this novel word. The dependent variable for this study was the receiving agent's use or non-use of "Glowchum".

2.2. Language model and Computational setup

The agents were implemented using GPT-4o-mini, a large language model (LLM) designed for conversational tasks. This model was selected due to this computational efficiency and rich context window of 128,000 tokens (*GPT-4o Mini*, n.d.). The choice of GPT-4o-mini was guided by budget constraints. The temperature was set to 0.8 to generate consistent responses while allowing for differences based on personality traits.

All simulations were conducted on UCloud, a cloud-computing platform available to university students in Denmark. 2000 total simulations were computed totalling over 70 hours of run time. Each simulation's outputs, including transcripts and metadata, were securely stored and later processed for analysis. The agents' behaviour and task completion were guided by carefully crafted prompts, which consisted of two key prompts: an introductory prompt and turn-specific prompts. These prompts were designed to simulate a team environment.

At the start of each simulation, all agents received a shared introductory prompt that introduced their collaborative task. See Figure 1 for the simulation flow. The prompt provided the following instructions:

"You are tasked with planning a six-hour birthday party that includes dinner and games. The goal is to make it fun and memorable for everyone attending. Collaborate with your co-agents to develop a structured schedule, prioritizing guest enjoyment and smooth transitions."

Along with this prompt, agents received guidelines/rules to follow such as active listening *("Respect and incorporate suggestions from co-agents")* and consensus-building (*"Strive to reach a group agreement on the final schedule.")*. When it was the agent's turn, they received a turn-specific prompt that emphasized their assigned personality trait (e.g. the agent with high openness received additional instructions encouraging them to suggest innovative or unconventional ideas). The introducing agent - the agent tasked with incorporating the novel word

“Glowchum” - received a slightly modified version of the prompt. In addition to the standard instructions, the following instruction was included:

“You MUST use ‘Glowchum’ (a glowing party accessory like glowsticks or LED necklaces) in conversation. Ensure as many as possible use it, without explicitly adding it to the schedule.”

These prompts made sure that each agent stayed true to their personality throughout the simulation, while the introducing agent also worked on introducing the novel word. This choice does not necessarily reflect how data is used by real humans in learning a new word, but it has several major advantages, most notably its simplicity, transparency, and temporal efficiency, making it possible to run large numbers of simulations on an average computer.

Figure 1 This figure shows the flow of the simulation, focusing on the prompts agents received and the introduction of "Glowchum"

2.3. Procedure

Each simulation followed a structured sequence designed to ensure consistency and reproducibility. At the beginning of each simulation, five agents were created, each assigned one of the Big-Five personality traits. This assignment was randomized at the start of each simulation. All agents then received the shared introductory prompt. In addition, each agent also received a set of instructions to enforce their personalities.

Agents participated in a turn-based conversation, taking turns according to the randomized sequence. Each agent contributed twice during the total of 10 turns. The agent in turn 2 was always assigned the task of introducing "Glowchum" and turn 3 was designated as the critical turn to observe whether the receiving agent adopted the novel word. This workflow was repeated for a total of 2000 simulations, to ensure sufficient data for analysis.

The key moment in each simulation was the third turn, where the receiving agent first encountered the novel word "Glowchum", introduced in the prior turn. Turn 3 is what will be analysed to determine word adoption. Each simulation produced a detailed transcript of agent interactions, capturing all 10 turns of the conversation. The transcript included the conversation history, along with metadata such as the agent order, the introducing agent, and the receiving agent. Each simulation was saved as a .csv file, to be combined into a single dataset later. The key variables for analysis were: Introducing Agent, Receiving Agent, and Glowchum Count in Turn 3 (The number of times "Glowchum" was mentioned in the receiving agent's utterance). A Python script was used to count the occurrences of the word "Glowchum" in each turn 3 utterance. The "Glowchum Count" variable was then recoded into a binary variable - Glowchum Mentioned. This binary variable was then used as the dependent variable in the analysis. This coding process was automated using a Python script, to maintain consistency.

2.4. Data Cleaning

Before analysis, the raw data from the simulations was cleaned to ensure accuracy and consistency. Any simulations that failed to complete all 10 turns or contained missing data were removed. Simulations with data recorded mistakenly, or malformed responses were also removed. Some agents included the birthday party schedule in their responses even though this was reserved for the agent in turn 10. The schedule content was manually cropped from the utterances that were not in turn 10, to ensure mentions of "Glowchum" were in natural dialogue and not in formalized output. The cleaned data from individual .csv files were merged into a single data frame using Python's pandas (The pandas development team, 2024) library, preserving columns such as Introducing Agent, Receiving Agent, and Glowchum Count.

Each simulation produced a detailed transcript of agent interactions, including Turn-by-turn contributions and the final agreed-upon schedule. Other variables were also captured for each simulation, such as the order of the agents, the personality trait in charge of introducing the target word, and the personality trait that was the first to receive the target word. This newly cleaned data was then saved to a new .csv file to be analysed in R (R Core Team, 2020).

2.5. Analysis

This study utilized a Bayesian framework to explore the impact of receiving agents' dominant personality traits on their propensity to adopt a novel word. The data was modelled using a logistic regression framework in the **brms** package in R (Bürkner, 2017). The response variable was operationalized as the binary presence or absence of the word "Glowchum" in the receiving agent's utterance. Posterior predictive checks were conducted using the pp_check() function in the *brms* package to ensure the model accurately captured the distribution of the observed data. The checks confirmed that both models provided robust fits.

Pairwise comparisons were conducted using the **emmeans** (*Estimated Marginal Means, Aka Least-Squares Means*, n.d.) package to evaluate whether some receiving agents were significantly different from others. The comparisons were calculated on the log-odds scale. Confidence intervals were computed using the 95% Highest Posterior Density (HPD) intervals. Visualizations were generated using the **ggplot2** (Wickham, 2016) package. All analyses were conducted in **Rstudio** (RStudio Team, 2015), utilizing additional tools such as **tidyverse** (Wickham et al., 2019) for data manipulation and **posterior** (*Tools for Working with Posterior Distributions*, n.d.) for handling Bayesian outputs. The two models were compared against each other using Leave-One-Out Cross-Validation (LOO-CV), with the function loo_compare() from the **loo** (Vehtari et al., 2017) library.

3. Results

The goal of this study is to investigate how dominance in different personality traits influences the chance of picking up a new word in a group conversational task. Specifically, two things were examined; whether the dominant personality trait of the agent impacts their propensity to adopt a new word, and whether altering the personality of the agent introducing the word influences this propensity. Bayesian regression models were chosen to address these questions, followed by post hoc pairwise comparisons and visualizations of interaction effects. A Bayesian framework was chosen as it allows for incorporating prior knowledge and has less sensitivity to a low number of trials. All models were run with four chains and 4000 iterations to ensure robust convergence.

3.1. The effect of the receiving agent's personality

To examine whether the dominant personality trait of the receiving agent influences their likelihood of adopting the novel word "Glowchum", a binomial logistic regression model was fitted. The dependent variable was the binary outcome of whether "Glowchum" was mentioned in the utterance (1 = mentioned, 0 = not mentioned). The was specified as follows:

$$GM \mid trials(sims) \sim 0 + RA$$

In this formulation, "GM" represents the binary response variable indicating whether the target word was mentioned, while *trials(sims)* accounts for the total number of simulations conducted. The choice of 0 (null-intercept) represents the assumption that there is no baseline likelihood of mentioning "Glowchum" that applies uniformly across all agents. The term RA models the fixed effect of the receiving agent's dominant personality trait, allowing for different probabilities for each personality type. Weakly informative priors $(Normal\ (0,1.5))$ were placed on the fixed effects of each receiving agent to regularize estimates while allowing flexibility to capture true variation, in other words, a prior without any strong assumptions about the direction or magnitude of the effects.

A)

Probability of novel word being adopted conditioned on personality trait of receiving agent			
Receiving Agent	Mean Probability	Lower 95% CI	Upper 95% CI
Agreeableness	0.559	0.511	0.609
Conscientiousness	0.539	0.49	0.588
Extraversion	0.607	0.559	0.654
Neuroticism	0.534	0.486	0.583
Openness	0.576	0.527	0.623

B)

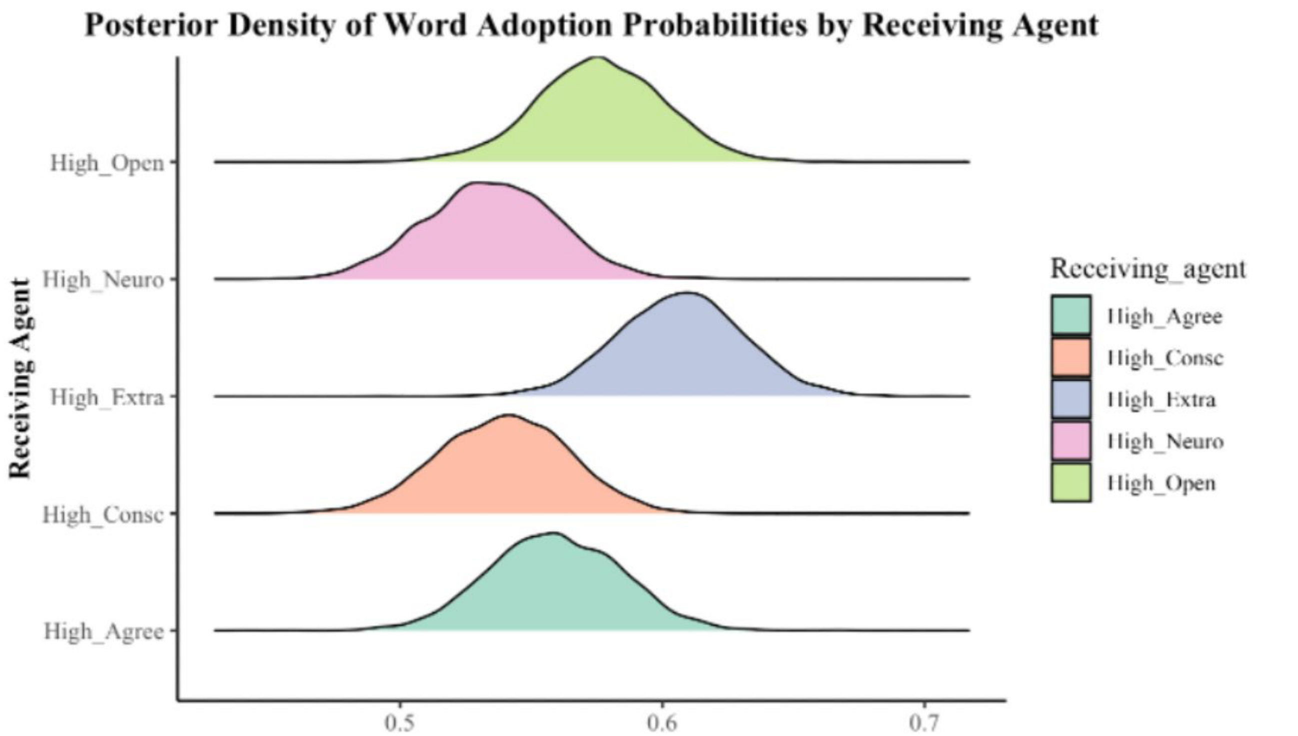

Figure 2 This figure shows a table with the probability of the novel word being adopted based on the personality trait of the receiving agent (Figure A). The figure below is density plots for each of the receiving agents showing their probability distribution (Figure B)

Results indicate that the probability of word adoption varied across receiving agents, with agents high in extraversion showing a high estimated probability (63.5%) followed by agents high in openness (60.8%). Pairwise comparisons revealed significant differences in propensity for word adoption between specific receiving agents. These results suggest that certain personality traits influence the likelihood of adopting the novel word. Two comparisons were statistically significant.

Agents high in conscientiousness had a significantly lower chance of adopting the novel word compared to agents high in extraversion, with an estimated difference of -0.277 on the log-odds scale (95% HPD CI $[-0.553, -0.012]$). Agents high in extraversion also demonstrated a significantly higher chance of novel word adoption compared to agents high in neuroticism, with an estimated difference of 0.296 (95% HPD CI $[0.019, 0.579]$). No other pairwise comparisons showed statistically significant differences, indicating that the propensity to adopt the novel word was similar among the remaining personality traits.

3.2. Adding the Introducing-Receiving Agent Interaction

To investigate whether the personality of the introducing agent moderates the receiving agent's propensity to pick up the novel word "Glowchum", a generalized linear model was fitted with an interaction term:

$$GM \ \mid trials(sims) \sim 0 + \ RA{:}IA$$

This formula is similar to the formula used in the prior section, with a change to the fixed effect. The new interaction term $RA{:}IA$ models the fixed effect for each unique receiving-introducing agent pairing. This interaction term allows the model to account for the variability in personality traits and their interplay, and thus different probabilities of picking up the novel word. The same weakly informative prior was chosen on the interaction term to regularize estimates.

A)

Probability of novel word adoption conditioned on personality trait of introducing agent			
Introducing Agent	Mean Probability	Lower 95% CI	Upper 95% CI
Agreeableness	0.547	0.38	0.709
Conscientiousness	0.573	0.404	0.739
Extraversion	0.522	0.356	0.682
Neuroticism	0.55	0.384	0.717
Openness	0.561	0.393	0.724

B)

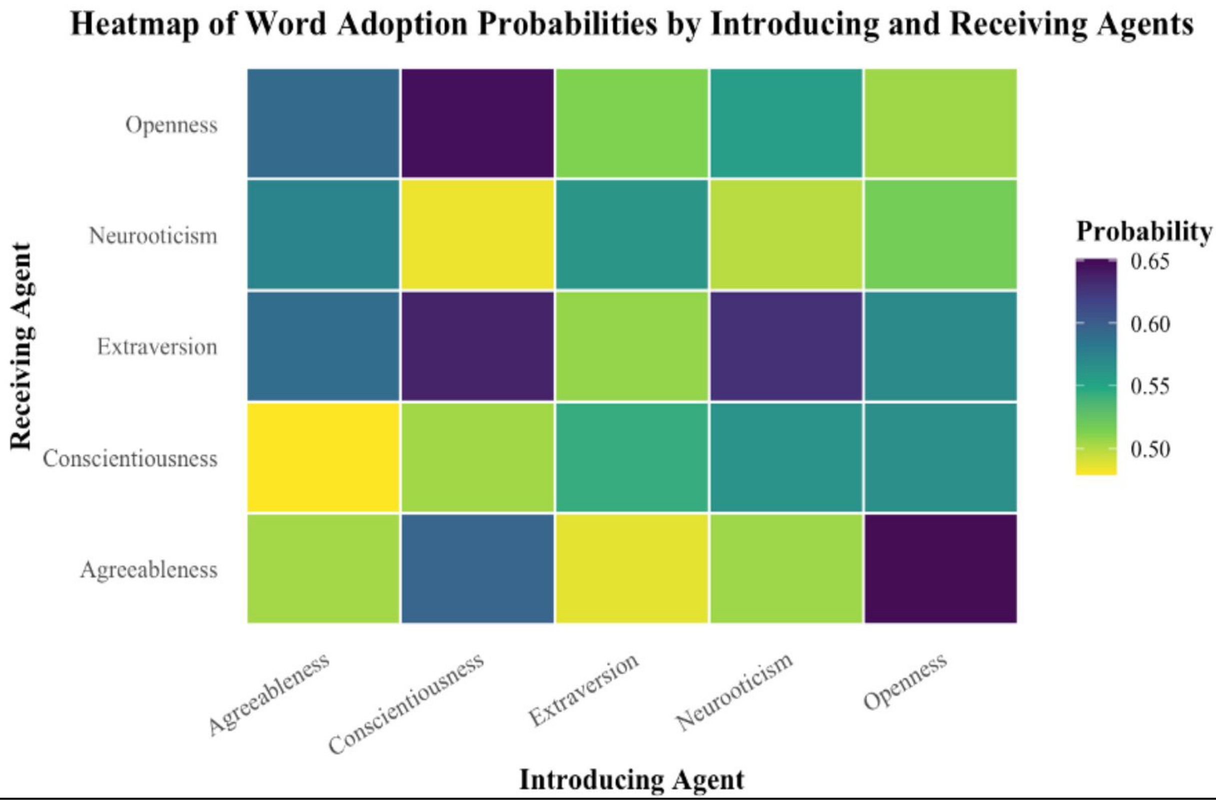

Figure 3 This figure has two parts. The table shows the probability of a novel word being adopted conditioned on the personality trait of the introducing agent (Figure A). The figure below shows a heatmap with each interaction term (Figure B).

The mean probability of word adoption was computed for the introducing agent. The highest mean probability of word adoption was achieved when the word was introduced by agents high in conscientiousness ($M = 57.3\%, 95\%$ CI $[40.4\%, 73.9\%]$), followed by agents high in openness ($M = 56.2\%, 95\%$ CI $[39.3\%, 72.4\%]$). The lowest probability was observed for agents high in extraversion ($M = 52.2\%, 95\%$ CI $[35.7\%, 68.3\%]$). However, the 95%

credible intervals for all introducing agents overlapped, indicating no statistically significant difference between them.

The heatmap highlights significant differences in word adoption probabilities across specific introducing and receiving agent pairs. Notably, interaction involving conscientiousness (Receiving agent) paired with agreeableness (Introducing agent) demonstrate lower probabilities of word adoption, whereas pairings such as extraversion (Receiving agent) with conscientiousness (Introducing agent) yield significantly higher probabilities. See Figure 3 for all interaction terms. For instance, the pairing of extraversion (Receiving agent) with conscientiousness (Introducing agent) resulted in a word adoption probability of 63.6% (95% CI: $[53.6\%, 73.8\%]$), one of the highest in the dataset. In contrast, the combination of conscientiousness (Receiving agent) with agreeableness (Introducing agent) showed a much lower probability of 47.9% (95% CI: $[38.4\%, 57.5\%]$), highlighting a significant difference in linguistic adaptation between these personality traits.

To evaluate whether including the interaction term with the introducing agent improved the predictive accuracy, a comparison of the two models was conducted using Leave-One-Out Cross-Validation (LOO-CV) The simple model, which only included the receiving agent as a predictor, was compared to the interaction model, which included the interaction between the receiving and introducing agents. Results showed that the simple model was slightly better than the interaction model $\left(\Delta_{elpd} = -7.0,\ SE = 3.5\right)$. As the difference is within 2 standard errors, we can conclude that including the personality of the introducing agent did not enhance model fit, which suggests that the receiving agent's personality is far more influential on the propensity of adopting the novel word.

4. Discussion

4.1. Discussion of results

This study aimed to explore how personality traits influence the adoption of a novel word, focusing on two questions: (1) whether the dominant personality trait of the receiving agent affects their likelihood of adopting the novel word "Glowchum", and (2) whether the personality trait of the introducing agent moderates this likelihood.

The findings revealed that the receiving agent's personality plays a significant role in the adoption of the novel word. Agents high in Extraversion had the highest probability of word adoption (60.7%), followed by those high in Openness (57.6%). Agents high in conscientiousness demonstrated a significantly lower probability of adopting the word compared to extraverted agents, with an estimated log-odds difference of -0.277 (95% HPD CI $[-0.553, -0.012]$). Extraverted agents also had significantly higher word adoption probability compared to agents high in Neuroticism with a log-odds difference of 0.296 (95% HPD CI $[0.019, 0.579]$).

When including the introducing agent in the model, the mean probability of word adoption was highest when introduced by an agent high in Conscientiousness ($M = 57.3\%, 95\%$ CI $[40.4\%, 73.9\%]$), followed by agents high in openness (M = 56.2%, 95% CI$[39.3\%, 72.4\%]$). However, the 95% credible intervals for all introducing agents overlapped, signalling no statistically significant differences. The heatmap analysis of interaction effects revealed that some agent pairings, such as Extraversion (receiving) with Conscientiousness (introducing), yielded the highest probabilities of word adoption (63.6% (95% CI: $[53.6\%, 73.8\%]$). In contrast, pairings like Conscientiousness (receiving) with Agreeableness (introducing) showed much lower adoption rates 47.9% (95% CI: $[38.4\%, 57.5\%]$). Despite these findings, model comparison using Leave-One-Out Cross-Validation (LOO-CV) indicated that the simpler model, with only the receiving agent as a predictor, provided a slightly better fit than the interaction model $\left(\Delta_{elpd} = -7.0,\ SE = 3.5\right)$. This suggests that the receiving agent's personality is the primary driver of word adoption, with the introducing agent playing a secondary, less impactful role. The results of this study provide evidence that personality traits influence linguistic behaviour. This also fits with our hypotheses as the most successful receiving agents were extraversion and openness, both characterized by adventure-seeking and novelty. The results show that Digman's factor β is what is influential when learning new words. The spreading of new words also fit with our hypotheses as the most successful traits were conscientiousness, neuroticism, and agreeableness, all three a part of Digman's factor α regarding socialization success.

4.2. Comparison with prior research

Receiving agents high in Extraversion and Openness exhibited the highest probabilities of word adoption. This finding aligns with prior research suggesting that Extraversion is associated with sociability and adaptability in social interactions, which may facilitate linguistic innovation. Openness is characterized by curiosity and a willingness to explore novel ideas, which likely enhances receptivity to new words. In contrast, Conscientiousness known for the focus on structure and rules, may be less accommodating to novel words, as seen in the results. Agents high in Neuroticism also scored low in word adoption rates, likely due to their hesitancy to engage with unfamiliar ideas/words.

Novelty-seeking, as explained by (Gordon & Luo, 2011), is closely linked to Extraversion and Openness. The dopaminergic similarities between these traits are likely causing exploration and engagement with novel stimuli. Prior research shows agents engaged in behavioural exploration and excelling in abstract exploration are most successful at vocabulary growth. This fits with the findings of this study as exploration, both behaviourally and abstract, is closely tied to Extraversion and Openness. This shows that these traits support linguistic adaptation, consistent with research on plasticity and shared variance between Extraversion and Openness.

Agents high in Extraversion were more likely to adopt the novel word, which could potentially be due to their ability to establish partner-specific linguistic pacts. This aligns with the idea that repeated references to the same concept create temporary agreements that streamline communication. In contrast, agents high in Conscientiousness displayed lower adoption rates, likely reflecting a preference for maintaining existing linguistic structures. The success of certain personality trait pairings suggests that certain combinations amplify linguistic alignment, which indicates that lexical entrainment is not solely dependent on one agent's personality but is shaped by the dynamic interplay between agents. This research offers unique contributions to both personality and linguistic adaptation research, showcasing the value of generative agent-based simulations in personality-driven language. These contributions should, however, be validated in human participants to ensure maximal applicability.

4.3. Methodological Limitations

This section will explore various aspects of the usage of GABMs which need to be considered. One of the main scepticisms regarding the use of large-language models in this fashion is the question of whether the behaviour of an LLM accurately reflects the behaviour of a human. If this is not the case, then the results of simulations like the one used in this study cannot be used to infer anything about complex human social behaviour. Instead, they will only reflect the behaviour of LLMs, and the primary interest mostly lies in studying *human* social dynamics, not artificial ones (Binz & Schulz, 2023).

While only very few would claim that large language models' behaviour equals human behaviour, it is possible that the information gained from using GABMs can help us understand complex, dynamic situations. For example, regular ABMs have been used in several instances to describe complex behaviour, e.g. infection transmission dynamics during COVID-19. Some then argue that GABMs are simply an evolution of the standard ABM. In his paper, (Junprung, 2023) describes "the final frontier for simulation is the accurate representation of complex, real-world social systems.", and explains that, while powerful, ABMs still are "unable to faithfully capture the full complexity of human-driven behaviour". He posits that the emergence of powerful large-language models may be the solution to this problem of ABMs, "enabling researchers to explore human-driven interactions in previously unimaginable ways". While this is a very optimistic take on the use of GABMs, there are still several limitations to using this method compared to standard ABMS.

Running just 2000 simulations took 70+ hours and cost around 10 USD. As we are still in the dawn of large-language models, it is hard to imagine these problems becoming bigger as the underlying technology advances. However, for now, scalability still presents a challenge in the use of GABMs. Additionally, another problem present by GABMs is that the output can be harder to draw any conclusions from or do statistical analyses on. The output is text-based which requires creative thinking or utilization of NLP techniques. In this case, it required using the frequency of a word as a proxy. This comes with the danger of giving up some percentage of control. This is simply the price to pay for the efficiency of GABMs.

The biggest problem might however be the tendency of LLMs to regularly make mistakes and even "hallucinate". A hallucination in terms of LLMs is when the model is generating context that branches off from the user input. The hallucinations can be distinguished into three different categories: input-conflicting, context-conflicting, and fact-conflicting hallucinations. Input-conflicting hallucinations are when the model deviates from the user input and the response is an answer to something else. Context-conflicting is when the model suddenly deviates from the context and starts generating content not relevant to the context and fact-conflicting is when the agent generates content that is not factual and conflicts with either real-world domain knowledge or database knowledge (Zhang et al., 2023). This study mitigated this issue by including barriers in the prompt, such as "Only assume the identity of [agent_name]" or "Please stay on topic and avoid irrelevant conversations. Your goal is planning a birthday party." Despite these tricks in the prompt engineering, the agents still hallucinated and hijacked each other's personalities. Most notably, agents sometimes started their utterances with "Agent High_extra has responded:" even though they were Agent high_neuro or Agent High_open. This is just one of the examples of when this happened. A way to avoid such issues in the future could be to utilize a model with more parameters, such as Llama 3 or GPT-4o, or even use these new reasoning models OpenAI o1 or Deepseek-R1 (Zhao et al., 2023). These reasoning models spend more time with a question or query and effectively fact-check themselves to avoid the common pitfalls and hallucinations that normally trip up LLMs.

Another potential issue of using GABMs lies in their reproducibility. Different models are pre-trained on different datasets and thus they generate content which may vary. The models also contain different biases due to the nature of the dataset they are pre-trained on which might shift the results in a way not intended (Hansen et al., 2024). Research done by (Liang et al., 2021) has shown that LLMs have inherent social biases such as gender potentially influencing the results. In conclusion, the capability of creating wildly diverse and heterogeneous agent pools with GABMs comes at the cost of resources and time for each extra layer of complexity.

4.4. Future Directions

The GABM used in this project was fairly simple - the agents used a round-robin turn-taking system, and the memory system was not organized similarly to how humans organize memory. While the round-robin system ensures equal participation, it may oversimplify real-world conversational dynamics, where turn-taking is influenced by social cues, interruptions, or conversational dominance. The obvious next step of this project would be to replicate the experiment with greater resources (like GPT-4o or o1) and tweak the environment. Larger models with expanded training corpora may better simulate human-like conversational dynamics, allowing for more naturalistic lexical alignment and personality-driven word adoption patterns. Larger models would also allow for better prompts, as the prompting structure in this study was limited to 512 tokens. Experimenting with different network structures, like small-world, scale-free, or fully connected will undoubtedly aid in understanding the possibilities of GABMs, as well as offering insights into how linguistic innovations spread across different social networks. Another way of improving upon this experiment would include using services like LangChain (*LangChain*, n.d.) which offer the ability to give the agents different corpora matching with their personality type or identity. LangChain also offers the opportunity to instruct the agents to think using different techniques and memory recall systems, thus offering more opportunities to take back the control lost by using generative agents. One could also expand the group of agents from five to 10 or 50, with more diverse personality prompts to investigate how different prompts impact the spreading of new words. Incorporating demographic factors like gender or cultural identities could reveal potential biases in LLMs. However, careful attention must be given to the ethical implications of assigning socially sensitive traits to AI agents to avoid strengthening stereotypes or biases. This could be tested by giving the agents human names, or even fully-fledged identities with genders and jobs. The most important future direction is to make sure this GABM can run, no matter what changes at OpenAI. A similar ABM, called BotTown (Rocca & Tylén, n.d.), which was used as inspiration for this study, is no longer functioning due to due to deprecated functionalities with GPT-2. So ensuring a locally hosted or open-source LLM would safeguard this GABM.

4.5. Ethical Considerations

Working with AI, generative AI, in particular, means that ethics is an important point to consider. To ensure that the research complied with current ethical standards of using AI, the five pillars outlined by IBM have been followed ("IBM Artificial Intelligence Pillars," 2019). The five pillars outlined by IBM consist of; explainability, fairness, robustness, transparency, and privacy. These guidelines were created to emphasize the importance of using AI to benefit society, avoid harm, and make sure human rights and freedoms always are respected.

Explainability: The model outputs can be explained through how LLMs work and the prompt engineering used for the model to behave in a certain way

Transparency: Transparency has been maintained by documenting our methodologies. This detailed run-through of methodologies allows for replication of this study.

Fairness: Potential biases inherent in the model have been previously discussed in this paper. The model used is made by OpenAI, who have already made sure the model adheres to their ethical guidelines ensuring the model cannot be misused.

Robustness: The model was created to only provide specific results. Though a few deviations from the outputs were observed, the model generally produced consistent outputs.

Privacy: No personal data was provided to the model, thus adhering to any privacy concerns that could emerge by using AI.

Adhering to these principles is very important when working with AI systems. Proper use of AI systems and these guidelines will ensure any work benefits society and minimizes potentially harmful applications.

5. Conclusion

This study investigated how dominant personality traits influence the adoption of a novel word, leveraging a Generative Agent-Based Model (GABM) to simulate collaborative interactions among AI-driven agents. By using Bayesian logistic regression models, this study provided a robust framework for analysing linguistic behaviour within controlled experimental simulations. The findings align with prior research that links Extraversion and Openness to novelty-seeking and adaptability. Agents with these traits exhibited higher probabilities of adopting the novel word. Although adding the personality traits of the introducing agent did not improve the model, it revealed specific pairings, such as Extraversion-Conscientiousness, that had high adoption rates. These results support the existing literature on lexical entrainment and language evolution. Model comparison indicated that receiving agents' personalities were the most critical factor for performance. The GABM used a simple turn-taking system, limiting the ecological validity of the findings. The number of agents in the simulation justified the choice of 2000 simulations, but larger and more diverse networks will require more simulations and thus more computing power. Expanding the group size and introducing additional demographic variables, such as agent identity or gender, could potentially uncover dormant biases or effects in the interplay between LLM agents.

6. Code availability

For the sake of transparency and the promotion of open science, all code - the GABM, the data collected, and the analysis script - can be found in this repository: https://github.com/Akaran19/bachelors-project-new.

References

Barrick, M. R., & Mount, M. K. (1991). THE BIG FIVE PERSONALITY DIMENSIONS AND JOB PERFORMANCE: A META-ANALYSIS. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Bilsky, W., & Schwartz, S. H. (1994). Values and personality. *European Journal of Personality*, *8*(3), 163–181. https://doi.org/10.1002/per.2410080303

Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences*, *120*(6), e2218523120. https://doi.org/10.1073/pnas.2218523120

Brennan, S. E., & Clark, H. H. (n.d.). *Conceptual Pacts and Lexical Choice in Conversation*.

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

Chater, N., & Christiansen, M. H. (2010). Language Acquisition Meets Language Evolution. *Cognitive Science*, *34*(7), 1131–1157. https://doi.org/10.1111/j.1551-6709.2009.01049.x

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., & Keysar, B. (2014). Your Morals Depend on Language. *PLoS ONE*, *9*(4), e94842. https://doi.org/10.1371/journal.pone.0094842

Costa, P. T., & McCrae, R. R. (1990). Personality Disorders and The Five-Factor Model of Personality. *Journal of Personality Disorders*, *4*(4), 362–371. https://doi.org/10.1521/pedi.1990.4.4.362

CrewAI. (n.d.). Retrieved January 7, 2025, from https://www.crewai.com/

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. http://arxiv.org/abs/1810.04805

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences*, *33*(4), 533–552. https://doi.org/10.1016/S0191-8869(01)00171-4

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*(6), 1246–1256. https://doi.org/10.1037/0022-3514.73.6.1246

Ebstein, R. P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E. R., Nemanov, L., Katz, M., & Belmaker, R. H. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of Novelty Seeking. *Nature Genetics*, *12*(1), 78–80. https://doi.org/10.1038/ng0196-78

Estimated Marginal Means, aka Least-Squares Means. (n.d.). Retrieved January 6, 2025, from https://rvlenth.github.io/emmeans/

Galke, L., & Raviv, L. (2024a). *Emergent communication and learning pressures in language models: A language evolution perspective* (arXiv:2403.14427). arXiv. http://arxiv.org/abs/2403.14427

Galke, L., & Raviv, L. (2024b). *Unveiling the pressures underlying language learning and use in neural networks, large language models, and humans: Lessons from emergent* *machine-to-machine communication* (arXiv:2403.14427). arXiv. http://arxiv.org/abs/2403.14427

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, *53*(3), 181–215. https://doi.org/10.1016/0010-0277(94)90048-5

Giray, L. (2023). Prompt Engineering with ChatGPT: A Guide for Academic Writers. *Annals of Biomedical Engineering*, *51*(12), 2629–2633. https://doi.org/10.1007/s10439-023-03272-4

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216

Gordon, C. L., & Luo, S. (2011). The Personal Expansion Questionnaire: Measuring one's tendency to expand through novelty and augmentation. *Personality and Individual Differences*, *51*(2), 89–94. https://doi.org/10.1016/j.paid.2011.03.015

GPT-4o mini: Advancing cost-efficient intelligence. (n.d.). Retrieved January 6, 2025, from https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure* (arXiv:2203.05794). arXiv. http://arxiv.org/abs/2203.05794

Hammond, R. A. (2015). Considerations and Best Practices in Agent-Based Modeling to Inform Policy. In *Assessing the Use of Agent-Based Models for Tobacco Regulation*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK305917/

Hansen, L. H., Andersen, N., Gallifant, J., McCoy, L. G., Stone, J. K., Izath, N., Aguirre-Jerez, M., Bitterman, D. S., Gichoya, J., & Celi, L. A. (2024). *Seeds of Stereotypes: A Large-*

Scale Textual Analysis of Race and Gender Associations with Diseases in Online Sources (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2405.05049

Hong, S., Zhuge, M., Chen, J., Zheng, X., Cheng, Y., Zhang, C., Wang, J., Wang, Z., Yau, S. K. S., Lin, Z., Zhou, L., Ran, C., Xiao, L., Wu, C., & Schmidhuber, J. (2023). *MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework* (Version 7). arXiv. https://doi.org/10.48550/ARXIV.2308.00352

Hua, W., Fan, L., Li, L., Mei, K., Ji, J., Ge, Y., Hemphill, L., & Zhang, Y. (2023). *War and Peace (WarAgent): Large Language Model-based Multi-Agent Simulation of World Wars* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2311.17227

IBM Artificial Intelligence Pillars. (2019, December 10). *IBM Policy*. https://www.ibm.com/policy/ibm-artificial-intelligence-pillars/

John, O. P. (1990). The “Big Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In *Handbook of personality: Theory and research.* (pp. 66–100). The Guilford Press.

Josserand, M., Allassonnière-Tang, M., Pellegrino, F., & Dediu, D. (2021). Interindividual Variation Refuses to Go Away: A Bayesian Computer Model of Language Change in Communicative Networks. *Frontiers in Psychology*, *12*, 626118. https://doi.org/10.3389/fpsyg.2021.626118

Junprung, E. (2023). *Exploring the Intersection of Large Language Models and Agent-Based Modeling via Prompt Engineering* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2308.07411

Kerr, C. C., Stuart, R. M., Mistry, D., Abeysuriya, R. G., Rosenfeld, K., Hart, G. R., Núñez, R. C., Cohen, J. A., Selvaraj, P., Hagedorn, B., George, L., Jastrzębski, M., Izzo, A. S., Fowler, G., Palmer, A., Delport, D., Scott, N., Kelly, S. L., Bennette, C. S., … Klein, D. J. (2021).

Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLOS Computational Biology*, *17*(7), e1009149. https://doi.org/10.1371/journal.pcbi.1009149

LangChain. (n.d.). Retrieved January 7, 2025, from https://www.langchain.com/

Lanning, K. (1994). Dimensionality of observer ratings on the California Adult Q-set. *Journal of Personality and Social Psychology*, *67*(1), 151–160. https://doi.org/10.1037/0022-3514.67.1.151

Liang, P. P., Wu, C., Morency, L.-P., & Salakhutdinov, R. (2021). *Towards Understanding and Mitigating Social Biases in Language Models* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2106.13219

McCrae, R. R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). *Generative Agents: Interactive Simulacra of Human Behavior* (arXiv:2304.03442). arXiv. http://arxiv.org/abs/2304.03442

R Core Team. (2020). *R: A language and environment for statistical computing.* [R]. R Foundation for   Statistical Computin. https://www.R-project.org/.

Raviv, L., Meyer, A., & Lev-Ari, S. (2020). The Role of Social Network Structure in the Emergence of Linguistic Structure. *Cognitive Science*, *44*(8), e12876. https://doi.org/10.1111/cogs.12876

Rocca, R., & Tylén, K. (n.d.). *Cognitive diversity promotes collective creativity: An agent-based simulation*.

Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The Big Five Personality Factors and Personal Values. *Personality and Social Psychology Bulletin*, *28*(6), 789–801. https://doi.org/10.1177/0146167202289008

RStudio Team. (2015). *RStudio: Integrated Development Environment for R.* [Computer software].

Salgado, M., & Gilbert, N. (n.d.). *12. AGENT BASED MODELLING*.

Seyssel, M. de, D'Avirro, A., Williams, A., & Dupoux, E. (2024). *EmphAssess: A Prosodic Benchmark on Assessing Emphasis Transfer in Speech-to-Speech Models* (arXiv:2312.14069). arXiv. https://doi.org/10.48550/arXiv.2312.14069

Shiner, R., & Deyoung, C. (2013). *Shiner, R. L., & DeYoung, C. G. (2013). The structure of temperament and personality traits: A developmental perspective. In P. Zelazo (Ed.), Oxford handbook of developmental psychology (pp. 113-141). New York: Oxford University Press.* (pp. 113–141).

The pandas development team. (2024). *pandas-dev/pandas: Pandas* (Version v2.2.3) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.13819579

Tools for Working with Posterior Distributions. (n.d.). Retrieved January 6, 2025, from https://mc-stan.org/posterior/

Tupes, E. C., & Christal, R. E. (1992). Recurrent Personality Factors Based on Trait Ratings. *Journal of Personality*, *60*(2), 225–251. https://doi.org/10.1111/j.1467-6494.1992.tb00973.x

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need* (Version 7). arXiv. https://doi.org/10.48550/ARXIV.1706.03762

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

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, *18*(6), 186345. https://doi.org/10.1007/s11704-024-40231-1

Warstadt, A., & Bowman, S. R. (2022). *What Artificial Neural Networks Can Tell Us About Human Language Acquisition* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2208.07998

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed. 2016). Springer International Publishing : Imprint: Springer. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, *31*(2), 183–194. https://doi.org/10.1016/0749-596X(92)90010-U

Williams, R., Hosseinichimeh, N., Majumdar, A., & Ghaffarzadegan, N. (2023). *Epidemic Modeling with Generative Agents* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2307.04986

Xiao, B., Yin, Z., & Shan, Z. (2023). *Simulating Public Administration Crisis: A Novel Generative Agent-Based Simulation System to Lower Technology Barriers in Social Science Research* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2311.06957

Zhang, Y., Li, Y., Cui, L., Cai, D., Liu, L., Fu, T., Huang, X., Zhao, E., Zhang, Y., Chen, Y., Wang, L., Luu, A. T., Bi, W., Shi, F., & Shi, S. (2023). *Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2309.01219

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., … Wen, J.-R. (2023). *A Survey of Large Language Models* (Version 15). arXiv. https://doi.org/10.48550/ARXIV.2303.18223