POLITECNICO DI TORINO

Master's Degree in Computer Engineering



Master's Degree Thesis

Factors Influencing the Evolution of Culture in Population Encounters

Supervisors

Luca CAGLIERO

Jérôme EUZENAT

Candidate

Edoardo CECCHINI

October 2026

Summary

The encounter between different populations and their cultures is a particularly interesting factor influencing cultural change, yet most computational models of cultural dynamics focus on a single isolated population, and the few that address contact between groups typically consider the migration of individuals rather than the encounter of two whole, already-established cultures. This thesis investigates what happens when two artificial agent populations, each having independently developed its own culture, are brought into contact and allowed to interact.

Building on an interaction-driven ontology adaptation model, in which each agent holds a private ontology representing its individual view of the environment, agents interact by encountering objects and taking decisions on them until they reach a behavioral consensus on every object, a state we refer to as *stabilization*. We introduce a two-phase protocol in which two populations first stabilize in isolation and then merge into a single population that interacts and stabilizes together. We characterise the resulting culture through two complementary outcomes: *assimilation*, where one population's knowledge is adopted by the other, and *hybridization*, where the two combine into something new. These outcomes are assessed on two levels, a behavioral *Decision Distance* and a semantic *Ontology Distance*, which, to be turned into interpretable outcomes, are combined into a geometric classification scheme and a continuous distance-difference metric. A full factorial study then varies three factors as possible influences on cultural dynamics: the size of the two populations, the structure of their interactions, and how much knowledge they shared before meeting.

The encounter reliably drives the merged population to stabilization. Which culture prevails, however, is dictated almost entirely by population size: the larger population assimilates the smaller, with the strength of assimilation growing monotonically as the size asymmetry increases. Neither the structure of the interactions nor the amount of prior shared knowledge exerts a comparable effect, in contrast with single-population studies that tie network structure to the preservation of cultural diversity. The one notable exception is an observation on the internal semantic diversity of the merged population's ontologies: populations that stabilized on entirely non-overlapping knowledge produce the least diverse final culture, though

this effect is small and requires further investigation.

Acknowledgements

First of all, I would like to express my sincere gratitude to my two supervisors: Jérôme Euzenat, for his constant guidance and supervision throughout this project, and Luca Cagliero, for having agreed to follow my work from Italy.

I cannot forget to thank my family, who, despite the physical distance, always managed to make me feel close to them, supporting me and giving me the strength to keep going even in the most difficult moments.

Thank you, Dad: you taught me that kindness is not a weakness, but an incredible strength. I truly admire you as a person. Thank you, Mom: you have always been good at bringing me back down to earth and showing me the true size of things, reminding me that problems that seemed enormous often were not.

Thank you, Jaco: with you I share a magic energy, and it was fun to keep our passion for music alive, even over video calls.

Thank you, Giordy: your playful approach to life showed me that a balanced life, not made only of work, is possible, and reminded me to enjoy the journey.

A big thank you to my Italian friends, who never left me even though I was often busy, in particular my best friends Gimmi and Giovi: having you two in my life made me feel safer.

Antonio, thank you for keeping me updated on what was happening at the Poli while I was abroad, and for all the long calls about so many interesting topics. You are a good listener and friend.

Thanks also to all the friends I met during my Erasmus: we were all in the same situation, and we built a strong bond that helped us feel at home even far from the real one. Thank you for all the nights out, the laughs and the adventures: Giuse, Mattia, Matteo, Alex and all the others. Giulia, it was great to have you around: I love how kind and funny you are.

Thanks to my French friend Pedro, who helped me every time I got into some "French" trouble. He taught me that asking for help is not a sign of weakness, but something that can truly enrich you.

I also want to thank the entire mOeX team, who never hesitated to offer valuable advice whenever I needed it.

Finally, I want to thank myself. Very often my life has been driven by fear:

fear of not being enough, of falling behind, of making mistakes, of discovering that everything I had achieved was just luck. This fear made me forget what really matters and kept me from doing one of the things I love most: trying, playing. Yet, despite all this, I made it here. I have never been able to say it to myself, but I too have some courage. It is never too late to change, and I feel that the wind has finally shifted direction. From now on, I want to be driven by curiosity, by the beauty of discovering how things work, of getting my hands dirty. I love you, life. Thank you for this opportunity to live.

Table of Contents

List of Tables IX

List of Figures X

1 Introduction 1
- 1.1 Context and Motivation . . . 1
- 1.2 Problem Statement . . . 2
- 1.3 Contributions . . . 2
- 1.4 Report Structure . . . 3

2 State of the Art 4
- 2.1 From Social Learning to Cumulative Culture . . . 4
- 2.2 The Impact of Demographics and Network Architecture . . . 5
- 2.3 Transmission Biases and Knowledge Selection . . . 6
- 2.4 Multi-Agent Systems and Cultural Evolution . . . 7
  - 2.4.1 The Baseline Model: Interaction-Driven Ontology Adaptation 7
- 2.5 The Unstudied Gap: Cross-Cultural Encounters . . . 9

3 Experiment Design 11
- 3.1 Modelling Cross-Cultural Encounters . . . 11
- 3.2 Research Questions and Hypotheses . . . 12
  - 3.2.1 Hypotheses and Motivations . . . 12
- 3.3 The Two-Phase Protocol . . . 14
  - 3.3.1 Setup . . . 15
  - 3.3.2 Phase 1: Individual Stabilization . . . 15
  - 3.3.3 Phase 2: Encounter . . . 16
- 3.4 Experimental Methodology . . . 17
  - 3.4.1 Independent Variables . . . 17
  - 3.4.2 Measures . . . 18
  - 3.4.3 Experiment Plan . . . 25
- 3.5 Implementation . . . 26

3.5.1 Software Framework 27
3.5.2 The Population Abstraction 28
3.5.3 Implementing the Independent Variables 29
3.5.4 The Two-Phase Game Iterator 30
3.5.5 Detecting Stabilization and Recording Metrics 31

4 Results and Analysis 33
4.1 Hypothesis 1: Encounter Stabilization 33
4.1.1 Results 33
4.1.2 Discussion 34
4.2 Hypothesis 2: Population Asymmetry and the Direction of Assimilation 35
4.2.1 Results 35
4.2.2 Discussion 36
4.3 Hypotheses 3: Network Connectivity and Information Bottlenecks . 37
4.3.1 H3.1: Results 37
4.3.2 H3.1: Discussion 38
4.3.3 H3.2: Results 40
4.3.4 H3.2: Discussion 40
4.4 Hypotheses 4: Environmental Overlap and Cultural Dynamics . . . 41
4.4.1 H4.1: Results 41
4.4.2 H4.1: Discussion 42
4.4.3 H4.2: Results 43
4.4.4 H4.2: Discussion 43
4.5 Refining the Analysis: The Distance Difference 45
4.5.1 Limitations of the Geometric Scheme 45
4.5.2 The Distance Difference Metric 46
4.5.3 H2 Revisited: Population Asymmetry 47
4.5.4 H3.1 Revisited: Network Connectivity and Assimilation . . . 49
4.5.5 H4.1 Revisited: Environment Overlap and Assimilation Direction 50

5 Conclusion 54
5.1 Summary of Findings 55
5.2 Limitations and Perspective 56

A Data and Software Availability 58

Bibliography 59

List of Tables

3.1 Fixed experimental parameters applied across all simulation runs. . 26
3.2 Grid of the independent variables varied across experiments. 26
3.3 Derived absolute sizes for populations and contact groups based on the interaction of α and η (Total Agents $N = 30$). 27

4.1 Behavioral Decision Outcomes (%) across varying Population Distribution (α) values. 36
4.2 Semantic Ontology Outcomes (%) across varying Population Distribution (α) values. 36
4.3 Behavioral Decision Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Network Connectivity (η) levels. 38
4.4 Semantic Ontology Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Network Connectivity (η) levels. 38
4.5 Behavioral Decision Outcomes (%) under Asymmetric Population Distributions ($\alpha = 0.12, 0.25$) across varying Network Connectivity (η) levels. 39
4.6 Semantic Ontology Outcomes (%) under Asymmetric Population Distributions ($\alpha = 0.12, 0.25$) across varying Network Connectivity (η) levels. 39
4.7 Behavioral Decision Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Environment Overlap (ω) levels. 41
4.8 Semantic Ontology Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Environment Overlap (ω) levels. 41

List of Figures

3.1 Network connectivity (η) determines which agents are permitted to interact across populations during the encounter phase. Grey edges denote intra-population interactions, which are never constrained; orange edges denote the admissible inter-population interactions. In (a), $\eta = 1.0$ and every agent may interact with every member of the other population. In (b), $\eta < 1.0$ and each population can be reached from the outside only through its contact group (shaded): an agent may interact with the contact-group members of the other population, whereas pairs in which neither agent belongs to a contact group are excluded. To make the illustration clearer, both panels show a reduced configuration of four agents per population, with a single-agent contact group in (b). 19

3.2 Environment overlap (ω) fixes the fraction of the object space shared by the two populations' training sets during the individual training phase. In (a), $\omega = 0$ and each population is trained on a completely disjoint region of the object space. In (b), $0 < \omega < 1$ and a shared region (yellow) belongs to both training sets, while each population retains its own exclusive region. Region widths are for illustration purposes only and do not reflect the tested values of ω. 20

3.3 Two examples of the geometric classification scheme applied to cultural outcomes. In (a), the post-encounter culture C lies inside Population 2's sphere ($d(B, C) < \rho$) while remaining outside Population 1's sphere ($d(A, C) > \rho$), classifying the outcome as Assimilation by Population 2. In (b), C lies strictly outside both spheres ($d(A, C) > \rho$ and $d(B, C) > \rho$), indicating that the final culture cannot be attributed to either original population, a state of Hybridization. 24

3.4 Selection logic of `PopulationEncounterGameIterator`'s `next()` function, showing agent and object selection rules for both the stabilization and encounter phases. 32

4.1 Average success rate of the post-encounter iterations 34
4.2 Stacked column chart of the Behavioral Decision Outcomes across varying Population Distribution (α) levels. The progressive Assimilation by Population 2 as α decreases is clearly visible. 37
4.3 Stacked column chart illustrating the distribution of Behavioral Decision Outcomes across varying Environment Overlap (ω) levels. The chart visually highlights the non-monotonic spike in hybridization at full overlap ($\omega = 1.00$), though this shift does not reach statistical significance. 42
4.4 Line plot comparing the Pre-Encounter Semantic Distance ($d_o(\mathcal{O}_1, \mathcal{O}_2)$) and Final Semantic Diversity across varying Environment Overlap (ω) levels. 44
4.5 Distributions of the normalized Distance Difference Δd_d (behavioral, top) and Δd_o (semantic, below) across population distribution conditions. The dashed vertical line at 0 marks a perfectly balanced cultural outcome. Positive values indicate assimilation toward Population 2; negative values toward Population 1. Values approaching ± 1 represent near-complete cultural dominance. 48
4.6 Heatmaps illustrating the mean Distance Difference for decisions (Δd_d, left) and ontologies (Δd_o, right) across all combinations of Population Distribution (α) and Network Connectivity (η). 49
4.7 Line plot with standard deviation bands illustrating the mean continuous Distance Difference for both behavioral (Δd_d) and semantic (Δd_o) dimensions across varying Environment Overlap (ω) levels. . . 51
4.8 Line plot with standard deviation bands illustrating the mean continuous Behavioral Distance Difference not normalized (Δd_{raw}). . . 52

Chapter 1

Introduction

Throughout history, cultural evolution has been shaped by the meeting of different populations. The encounter between distinct cultures raises fundamental questions about how knowledge, behaviours, and conceptualisations of the world evolve under contact. While observing these dynamics in the real world is inherently difficult due to a lack of controlled conditions, artificial societies offer the capability to simulate and study these phenomena directly. This thesis investigates the cultural dynamics that emerge when two artificial populations, each possessing their own independently established knowledge, encounter one another.

1.1 Context and Motivation

Knowledge, Ontologies and Heterogeneity For artificial agents to act and cooperate effectively in an environment, they need to be able to represent their knowledge of the world. Ontologies serve as one such formal, structured representation. However, a fundamental problem arises because these ontologies are often built or learned independently. Because they are not standardised, agents end up with heterogeneous representations of the same environment. When these agents attempt to interact or perform shared tasks, this heterogeneity leads to disagreements. The need to reconcile these differing worldviews motivates the study of interaction-driven adaptation, which allows agents to dynamically align their knowledge to resolve such failures [1].

Cultural Evolution and Multi-Agent Systems This process of alignment can be viewed through the lens of Cultural Evolution, where culture is defined as the shared behaviours and accumulated knowledge of a population. Traditional observational methods in this field struggle to isolate specific transmission variables or control environmental conditions [2]. Multi-Agent Systems (MAS) serve as a

controlled experimental testbed to overcome this limitation [3]. In these simulations, a society of artificial agents adapts a shared culture through repeated interaction games. While historically applied to the evolution of language [4, 5], this approach has more recently been extended to address the evolution of complex formal ontologies, notably in the works of Bourahla et al. [1, 6].

From Single Populations to Encounters The majority of existing computational models focus almost exclusively on the dynamics within a single, isolated population. Yet, significant cultural change occurs at the boundary between distinct, already-formed cultures [7, 8]. The existing literature on cultural contact typically concerns the migration of single individuals into a new group, rather than the collision of two whole, established cultures [8, 2]. This highlights an interesting gap in the current research that has not yet been much explored.

1.2 Problem Statement

To investigate this less-studied phenomenon, this study introduces a scenario where two distinct populations first develop independent cultures through the interaction-driven adaptation model of Bourahla et al. [1], and subsequently are brought into contact and allowed to interact together.

The core question interrogated is: what becomes of the resulting shared culture? Conceptually, there are two candidate outcomes: *assimilation*, where one population's knowledge is behaviorally adopted by the other, and *hybridization*, where a novel combination emerges that cannot be attributed to either original population alone.

To understand what dictates the final cultural outcome, three key independent variables are systematically varied and evaluated: the population distribution (size asymmetry), the network connectivity (the structure and density of cross-cultural interactions), and the environment overlap (how much of the world the two populations shared prior to contact).

1.3 Contributions

The specific contributions of this thesis are:

1. A novel two-phase experimental protocol extending the baseline MAS experiment by Bourahla et al. [1] to explicitly model the encounter and merging of two stabilized populations.

2. The introduction of complementary behavioral (*Decision Distance*) and semantic (*Ontology Distance*) metrics across populations, alongside a geometric

classification scheme and a continuous distance-difference measure to characterise the outcomes of cross-cultural encounters.

3. A full factorial empirical study testing population distribution, network connectivity and environment overlap.

4. The empirical finding that the cultural outcome of such an encounter is dictated by the size of the two populations: the larger one assimilates the smaller, with assimilation strengthening as the size asymmetry grows, whereas neither the structure of their interactions nor the amount of prior shared knowledge has a comparable effect.

1.4 Report Structure

The remainder of this report is structured as follows. Chapter 2 reviews the state of the art regarding cultural evolution and introduces the baseline interaction-driven ontology adaptation model. Chapter 3 details the experimental design, formalises the research hypotheses, explains the parameters of the two-phase simulation protocol, and defines both the evaluation measurements and the comprehensive experimental plan. It also describes the implementation of the extension in the software framework used to run the simulation. Chapter 4 presents the analysis of the experimental results, assessing how each independent variable influenced the resulting cultural dynamics. Finally, Chapter 5 summarises the core findings, discusses the implications of the study, and outlines limitations and potential perspectives for future work.

Chapter 2

State of the Art

This chapter situates the present study within the literature on cultural evolution. It traces a path from the basic mechanisms of social learning to the accumulation of complex culture, examines how demographic and structural factors shape that accumulation, and reviews how multi-agent systems have been used to study these dynamics computationally. It then introduces the interaction-driven ontology adaptation model that serves as the baseline for this work, and closes by identifying the gap this research investigates: the encounter between two already-established cultures.

2.1 From Social Learning to Cumulative Culture

The foundation of cultural evolution lies in social learning, the universal phenomenon where individuals use information about others' decisions to make their own. By acting as a transmission mechanism, social learning allows behavioral traits to be passed from individual to individual; when coupled with the generation of novel variation and the selective adoption of advantageous traits, the result is cultural evolution [2]. While such transmission is widely documented across various non-human species, these socially learned traditions rarely increase in complexity over time.

In human societies, however, cultural traits are frequently refined and built upon across generations, with innovations progressively accumulating on top of existing ones. This continuous accumulation creates technologies and knowledge structures that are too complex to have been invented by a single individual from scratch. To specifically describe this high-complexity outcome of the cultural evolutionary process, some researchers use the term Cumulative Cultural Evolution (CCE) [2], a concept also central to experimental and observational work on cultural complexity [9, 7, 8].

The capacity for a population's culture to accumulate such complexity is fundamentally dependent on the difficulty of the task. Simple cultural traits can be easily reinvented through individual trial-and-error and do not rely heavily on transmission between different populations. In contrast, highly complex tasks are "causally opaque", meaning that the underlying mechanics or steps required to produce them are not obvious to a naive observer [8]. Because they cannot be easily deduced or reinvented, these complex traits rely on high-fidelity transmission mechanisms, such as active teaching or precise imitation, to be faithfully passed between members of the same generation and down across next generations [2]. Furthermore, beyond high-fidelity transmission within a group, the continued accumulation and recombination of such complex traits is favoured by physical migration and interaction between distinct populations [8, 7].

2.2 The Impact of Demographics and Network Architecture

Assuming a context of complex tasks, specific demographic characteristics have been shown to significantly influence a population's cultural evolution. The first critical feature is the population size: Cumulative Cultural Evolution (CCE) occurs faster, and more reliably, in larger populations, enabling them to reach more complex cultures. Larger populations also experience a reduced probability of cultural loss, which is particularly vital for preserving rare and complex knowledge [10, 9, 2].

However, research indicates that population size alone is insufficient to predict the ultimate complexity and structural diversity of a culture; the architectural connectivity of the population is an equally fundamental factor. In fully-connected networks, where each individual can interact with and learn from any other, information diffuses rapidly and efficiently. While this might seem beneficial, it leads to the rapid homogenization of the population, potentially destroying cultural diversity. When combined with transmission biases, high connectivity drives the entire population to converge prematurely on a single solution within a rugged fitness landscape, effectively halting the exploration of alternatives [7, 11].

Conversely, in partially-connected (or fragmented) networks, individuals interact only with specific subsets of the total population. This structural fragmentation causes information to diffuse in a much slower, clustered manner, which preserves the population's cultural diversity for longer periods. Because cultural recombination fundamentally requires the coexistence of multiple distinct traits to act as building blocks, this process is hindered in fully-connected networks, where the rapid homogenization can eliminate alternative lineages before they have the opportunity to be combined. By buffering distinct lineages from immediate, overarching competition, partially-connected networks allow different subgroups to explore

alternative, path-dependent solutions independently. When these partially isolated subgroups eventually interact, this preserved diversity allows for the novel recombination of distinct, well-developed traits [7, 11, 12]. Ultimately, this recombination can lead to more diverse and complex solutions that perform better in the long run [7, 8].

2.3 Transmission Biases and Knowledge Selection

When acquiring new knowledge, individuals face the fundamental decision of whether to rely on personal trial-and-error or social learning (copying others). Individuals are generally more likely to rely on social learning when environments are uncertain or unfamiliar, and when independent learning is costly or inaccurate [2, 13].

Once social learning is employed, the process of cultural transmission is guided by cognitive filters known as transmission biases. These can be broadly categorized into **content biases**, which dictate *what* information is most likely to be copied or transmitted, and **context biases**, which learners use to determine *whom* to copy. Two primary context biases are particularly relevant to cultural adaptation:

- **Conformity bias:** This involves adopting the behavior of the local majority. This strategy is highly adaptive when an individual migrates into a new, unfamiliar population (spatial variation), allowing them to quickly acquire locally "working" behaviors from residents. However, it is less effective when environments change over time (temporal variation) and no individual yet possesses a more effective behavior.

- **Prestige (or Payoff) bias:** This involves preferentially copying the most successful individuals within a population, and is typically utilized when cues of success are reliable.

While migration and connectivity between distinct groups are key for sharing complex culture, accelerating cultural recombination, and preventing the erosion of cultural diversity [7, 8], this transmission is not guaranteed. The transmission of complex knowledge may fail due to the low skill level of the migrant or the absence of an appropriate ecological context for the behavior in the new environment. Most commonly, transmission fails because low-ranking migrants fall into a "conformity trap", wherein they simply adopt the established behaviors of their new community rather than transferring their own unique, natal knowledge.

2.4 Multi-Agent Systems and Cultural Evolution

As discussed in previous sections, the standard study of cultural evolution has traditionally relied on observational data. While foundational, these methods are limited by their inability to tightly control environmental conditions or isolate specific transmission variables. To overcome these limitations, Multi-Agent Systems (MAS) have long been used to simulate cultural evolution.

By embodying artificial agents with cognitive representations and having them interact through well-defined protocols, researchers can isolate variables and directly observe the macro-level impact of micro-level transmission rules. A key advantage of MAS is its ability to demonstrate how complex, collective intelligence emerges from simple individual behaviors. For example, computational models have shown that when individual agents employ simple heuristic rules, such as payoff bias, the population as a whole functions as a distributed algorithm capable of highly efficient information aggregation [13].

Researchers have successfully used artificial agents to model various foundational aspects of cultural evolution. Early computational models explored the emergence of shared culture through agent interaction [3], and subsequently focused on the transmission and selection of basic spatial behaviors across generations using artificial neural networks [14]. Another prominent line of research focused on the evolution of natural languages, using repeated interactions known as "language games" or "naming games", in which populations of agents were shown to converge on shared vocabularies and grammatical structures [4, 5]. Other agent-based frameworks have similarly explored the emergence of social norms, conventions, and collective behavior through peer-to-peer interactions [15].

Building directly upon this foundational state of the art, recent MAS simulations have advanced beyond the exchange of simple vocabulary or basic behaviors to address the evolution of complex, symbolic knowledge. Researchers have successfully applied interaction-driven adaptation to formal ontologies, demonstrating that agents can dynamically adapt their logical classes to repair communication failures [1, 6].

2.4.1 The Baseline Model: Interaction-Driven Ontology Adaptation

The study presented in this thesis builds upon the experimental setting introduced by Bourahla et al. [1]. In this baseline simulation, a population of agents interacts with objects in a shared environment. The environment comprises a set of object types I, each described by a set of binary properties P. For the sake of simplicity, an object type either possesses a specific property $p \in P$ or it does not (represented logically as $\neg p$). Because each unique combination of properties corresponds to

exactly one object type, the total size of the space I is $2^{|P|}$. A concrete object encountered by an agent is then an instance of one of these object types.

Furthermore, there exists a set of possible decisions $D = \{d_1, \ldots, d_k\}$. To each object type there corresponds only one correct decision, governed by a global objective function $h^* : I \to D$.

The baseline experimental framework consists of two main phases:

- **Learning Phase:** Each agent is trained independently on a random subset of the object types. Using the ID3 algorithm, the agent generates a decision tree based on its specific training set. This decision tree is then converted into a formal ontology: each node in the tree represents a logical class (a group of object types sharing specific features), branches create subclasses defined by combining the features of the parent classes, and leaves map to logical correspondences for decisions. These relationships are expressed using the Description Logic $\mathcal{ALC}$. Therefore, each agent stores a private ontology, a structured, hierarchical knowledge map that it uses to categorize and understand the environment.

- **Interaction Phase:** In the second phase, agents repeatedly interact with their peers over a defined number of iterations. The core of this phase is the interaction: two agents, a and b, are presented with a single, randomly selected object o and use their respective decision functions h^a and h^b to make a decision on it. Because their ontologies were learned independently, they may either agree ($h^a(o) = h^b(o)$) or disagree. In the event of a disagreement, a resolution mechanism is required to unlock the interaction and enable cooperation.

  The baseline model resolves disagreements through a mechanism inspired by the *prestige bias*, grounded in the agents' performance on a decision task. Before each interaction, the two agents are evaluated on this task: a random subset of objects $S \subseteq I$ is presented to each agent, which takes a decision on every object using its personal decision function. Each agent then receives a payoff $r_{S,a}$, defined as the number of objects for which it took the correct decision according to the global objective function h^*:

$$r_{S,a} = |\{o \in S \mid h^a(o) = h^*(o)\}| \tag{2.1}$$

  where o denotes an object drawn from the random subset S. When the two agents disagree, they compare these payoffs ($r_{S,a}$ and $r_{S,b}$), and the agent with the lower score adapts its ontology to implement the decision of the more successful one.

  This adaptation is achieved through a specific split operator: the lower-performing agent modifies its ontology by dividing the logical class it originally used to classify the object into two distinct subclasses. The first subclass is

formed by the intersection of the split class with the successful agent's class, and it is updated to adopt the successful agent's decision. The second subclass contains the rest of the original class and retains the less successful agent's initial decision.

This interaction-driven adaptation forms the core of cultural evolution within the system. Knowledge evolves through decentralized updates induced entirely by interaction failures, rather than through the direct fusion or merging of knowledge bases.

To evaluate the system, Bourahla et al. [1] introduced several metrics, the most important for the present study being the *Ontology Distance* (δ). This metric mathematically quantifies the semantic dissimilarity between two ontologies, O_a and O_b, and is computed as:

$$\delta(O_a, O_b) = 1 - \frac{eq(O_a, O_b)}{\max(|O_a|, |O_b|)} \tag{2.2}$$

where $eq(O_a, O_b)$ represents the number of equivalent classes between the two ontologies.

The central finding of this study was that, given sufficient time, peer-to-peer communication always reaches a 100% success rate, along with a higher quality of knowledge compared to the initial stage. Most importantly, agents achieve this state of global consensus without necessarily possessing identical ontologies, proving that a population can align its behavior while preserving the structural diversity of its individual knowledge.

2.5 The Unstudied Gap: Cross-Cultural Encounters

As established in the literature, the impact of network architecture on cultural evolution is primarily visible when tasks are highly complex and feature rugged fitness landscapes [7], while population size more broadly shapes the rate and reliability of cultural accumulation [9]. We assume that the multi-agent framework introduced by Bourahla et al. [1] provides a sufficiently complex environment, requiring agents to independently learn decision trees, classify objects, and continuously adapt their ontologies to repair interaction failures. This makes it a suitable ground on which to evaluate how demographic and structural variables influence cultural evolution.

However, this study, like much of the computational and experimental work in the field, has so far been applied mainly to the dynamics within a single group of agents. The encounter between two whole, already-established cultures remains comparatively little explored.

The present study is designed to address this gap by asking: what happens when two distinct agent populations, each with their own shared knowledge, encounter each other? Additionally, this research aims to assess whether the features demonstrated to be relevant within single populations (e.g., population size and network structure) will preserve their effect on knowledge evolution during these cross-cultural encounters [10, 9, 7, 11, 2, 12, 8].

Chapter 3

Experiment Design

This chapter details the experimental methodology designed to address the research questions of this thesis. It first establishes the conceptual definitions of populations and cross-cultural encounters, outlines the formal hypotheses, describes the simulation protocol, and finally details the experimental parameters, evaluation metrics, and the implementation of the experiment in the software.

3.1 Modelling Cross-Cultural Encounters

Before detailing the specific experimental protocol, it is necessary to conceptually define the primary entities of the simulation: populations and cross-cultural encounters.

In the context of this study, a **population** is defined as a set of artificial agents, each possessing an independent formal ontology representing its individual view of the world. Through repeated social interactions and adaptations, these agents gradually establish a consensus on how to behave within their environment. This shared behavioral consensus on every object type, reached while still preserving the structural diversity of their individual ontologies, constitutes the **culture** of that population, a state we refer to as *stabilization*.

A **cross-cultural encounter** occurs when two or more previously isolated populations, each with its own distinct culture, are brought into contact. In this setting, agents are no longer restricted to interacting solely with peers from their native group; they may also interact with members of the other populations. Because these distinct groups have typically evolved different solutions to environmental tasks, their encounter is likely to generate new interaction failures. These disagreements act as a catalyst, the variation that breathes new life into the cultures of the previous populations by pushing agents into new cycles of learning and adaptation.

Within the scope of the present study, we specifically model the encounter

between exactly two populations, which, following their initial contact, are subsequently treated and evaluated as a single, unified merged population.

3.2 Research Questions and Hypotheses

The aim of this study is to assess the following research question: *When distinct agent populations with their own pre-established cultures encounter one another, how do specific demographic and structural variables dictate the resulting cultural dynamics?*

This research investigates whether the encounter results in one of two macro-states:

- **Assimilation:** Does one population's knowledge get adopted by the other?
- **Hybridization:** Does the final culture represent a combined, newly adapted knowledge that does not show a clear dominance from either original culture?

To systematically evaluate these potential outcomes, we formulate specific hypotheses based on the three key independent variables manipulated in our simulation. Conceptually, these variables are:

- **Population Distribution (α):** The relative sizes of the two encountering populations (e.g., whether they are of equal size or if one is a minority).
- **Network Connectivity (η):** The density of social connections that allow agents from one population to interact with members of the other. Conceptually, this variable dictates what proportion of each population is reachable by the other, ranging from partially-connected networks with limited cultural bridges, to fully-connected populations where inter-population interactions are entirely unconstrained.
- **Environment Overlap (ω):** The degree to which the two populations were exposed to the same environmental objects during their initial, independent training phases. It represents the extent of their knowledge "common ground" prior to contact.

A rigorous technical formalization of these variables, along with the specific values tested in our experiments, is detailed subsequently in Section 3.4.

3.2.1 Hypotheses and Motivations

Based on the defined macro-states and independent variables, this study tests four formal hypotheses. Grounded in both the cultural evolution literature and the specific dynamics of our multi-agent framework (detailed in Section 2.4.1), these propositions are discussed below alongside their respective motivations:

Hypothesis 1: The encounter leads to stabilization, i.e., the success rate of the interactions converges to 1. This implies that the populations inevitably reach a global behavioral consensus.

Motivation for Hypothesis 1: The convergence to a global success rate of 1 directly follows from the failure-driven adaptation mechanism of the simulation. When two interacting agents disagree on a decision, the agent with the lower payoff adapts its ontology to align with the decision of the more successful agent. Because both populations interact within the exact same environment space (even if trained on different parts of it) and are evaluated against the same global objective function $h^*(o)$, the continuous pressure to repair interaction failures forces the populations to converge on a single behavioral consensus. This systemic requirement to resolve all disagreements renders long-term cultural coexistence impossible.

Hypothesis 2: The larger population assimilates the smaller one. This corresponds to the fact that the pre-established knowledge of the majority will behaviorally prevail in the final culture.

Motivation for Hypothesis 2: In encounters with skewed *population distributions* (α), interaction pairs are selected randomly across the permitted network. Because the minority population is numerically disadvantaged, its agents are disproportionately drawn into inter-population interactions where their pre-established knowledge conflicts with the consensus of the majority. Having fewer agents acting as a "backup" to reinforce their specific knowledge, the minority's unique trait lineages are susceptible to being overwritten during interaction-driven adaptations. The minority thus faces persistent demographic pressure from the larger population, leading to behavioral assimilation rather than the preservation of their original culture. Specifically, we predict this dominance primarily at the *behavioral* level. While the minority agents will adapt their outward decisions to align with the majority, their internal semantic structures (ontologies) only expand to accommodate these new rules rather than being entirely erased, making a situation of full assimilation difficult to assess at the semantic level.

Hypothesis 3:

- **Hypothesis 3.1:** Lower network connectivity results in more hybridization. In encounters with asymmetric population sizes, this implies a reduction in the absolute assimilation driven by the majority.

  Motivation for H3.1: In a fully-connected network, all agents can interact with one another, maximizing the immediate influence of the majority population and driving the system to a rapid convergence on the majority's pre-established behavior. Conversely, a restricted contact group bottlenecks inter-population interactions. This partial connectivity buffers the minority

culture from overwhelming majority pressure, slowing the rate of assimilation and preserving minority correct knowledge long enough to allow for more cultural hybridization.

- **Hypothesis 3.2:** Lower network connectivity yields a more semantic diverse final culture.

  Motivation for H3.2: The theoretical and experimental literature demonstrates that partially-connected (or fragmented) networks prevent the rapid homogenization of populations, albeit in settings different from the one studied here. By maintaining diverse cultural lineages in parallel for longer periods, partial connectivity provides the necessary time for distinct traits to percolate and recombine, leading to a more diverse final culture. This effect has been observed within single populations; we nonetheless extend this reasoning to the cross-population setting, treating it as an assumption to be tested.

Hypothesis 4:

- **Hypothesis 4.1:** Higher environment overlap increases hybridization.

  Motivation for H4.1: When two populations share a large portion of their training sets, they possess highly similar knowledge and they likely already agree on many decisions prior to the encounter. This shared baseline prevents the absolute behavioral dominance of either original culture, shifting the outcome toward a balanced hybridization where it becomes difficult to attribute the final knowledge to either pre-encounter population.

- **Hypothesis 4.2:** Higher environment overlap results in lower semantic diversity within the final culture.

  Motivation for H4.2: By establishing this substantial epistemic baseline during the initial learning phase, higher environment overlap causes the two populations to reach a more similar cultural state before the encounter begins. Consequently, owing to this prior similarity, the subsequent encounter phase will inherently trigger significantly fewer inter-population interaction failures. Since ontological adaptation in this framework is strictly failure-driven, a lack of disagreement means that the populations will undergo far fewer class-splitting events, reducing their exploration of the semantic space and producing a more uniform final culture (lower semantic diversity).

3.3 The Two-Phase Protocol

The underlying architecture described in Section 2.4.1 was originally designed to study the cultural dynamics and stabilization of a single, isolated population of

agents. To test the hypotheses of the present study, this baseline framework is extended to support multiple, distinct populations. This extension introduces mechanisms to define population sizes, control their inter-population network connectivity and manipulate the overlap of their initial training environments. Temporally, the extended simulation is structured as a two-phase protocol: an initial individual *stabilization phase* followed by a cross-cultural *encounter phase*.

3.3.1 Setup

Prior to the first phase of the simulation, the system and the agents are initialized according to the specific independent variables governing the run:

- **Population Partitioning:** The total number of agents N in the simulation is partitioned into two distinct groups, Population 1 and Population 2. Their respective sizes are determined by the *population distribution* parameter (α).
- **Network Initialization:** The inter-population interaction structure is established based on the *network connectivity* parameter (η). If $\eta < 1.0$, a subset of randomly selected agents from each population is designated to form its contact group. Two agents of different populations can be paired for an interaction only if at least one of them belongs to its population's contact group. The absolute size of each contact group is calculated proportionally based on η and the total number of agents in the respective population.
- **Environment Pool Generation:** The global environment of object types I is used to generate two sub-pools, I_1 and I_2, which serve as the base training environments for Population 1 and Population 2, respectively. The size and composition of these pools are generated such that the fraction of shared objects between them matches the specified *environment overlap* parameter (ω).
- **Agent Training:** During the initial learning phase, agents are not trained on the entire global environment I. Instead, each agent is trained independently on a random subset comprising approximately 50% of its respective population's object pool I_1 or I_2 (a training ratio of $r = 0.5$).

3.3.2 Phase 1: Individual Stabilization

During the first phase, each population undergoes its own stabilization process in isolation, using the interaction-driven adaptation protocol described in Section 2.4.1. This phase does not run for a fixed number of iterations; rather, because adaptation is strictly failure-driven, it terminates dynamically only when the population reaches full internal decision agreement on each object.

The time required to reach this state of total consensus can vary significantly between populations. This duration is influenced by stochastic interaction pairings and the initial similarity of the agents' knowledge. Consequently, even when populations share highly similar initial configurations, this inherent noise leads to distinct, population-specific cultures. Upon reaching global consensus, a snapshot of all agent ontologies ($\mathcal{O}_1$ and $\mathcal{O}_2$) and the resulting population-wide consensus decision functions (h_1 and h_2) are recorded. These baselines are later used to compute the cultural distance metrics.

It is important to note that, during this *stabilization phase*, agents interact over the entire global object space I, rather than being restricted to their respective training subsets (I_1 or I_2). While this aligns directly with the initial interaction protocol, it also serves as a methodological design choice for the current study. The primary objective of this experiment is to isolate and measure the cultural dynamics triggered exclusively by the cross-cultural encounter. Restricting populations to their training subsets during stabilization and only exposing them to the full environment upon contact would introduce a severe confounding variable: it would become impossible to distinguish whether post-encounter adaptations were driven by cultural contact or by exposure to environmental novelty.

Furthermore, by allowing agents to explore the full environment based on their partial initial training, the simulation naturally models a scenario where different populations possess unique, partial views of the same shared world, ensuring that any subsequent adaptation is purely a product of cultural interaction.

3.3.3 Phase 2: Encounter

The second phase starts immediately after both Population 1 and Population 2 have stabilized. As mentioned in Section 3.1, at this stage, the two previously isolated populations are conceptually merged into a single, global population. The interaction dynamics within this "newly formed" population are strictly governed by the *network connectivity* parameter (η) established during the setup phase.

The underlying adaptation protocol during the encounter remains identical to that of the individual stabilization phase: agents are paired stochastically, evaluate objects from the full environment, and use the split operator to resolve disagreements. The critical difference lies in the interaction structure, which explicitly restricts partner selection based on η. If the network is fully connected ($\eta = 1.0$), partner selection is globally uniform, meaning any agent has an equal probability of interacting with any other agent regardless of their original population. Conversely, if the network is partially connected ($\eta < 1.0$), a restriction applies only to interactions between the two populations. Two agents of the same population may always interact. Two agents of different populations may interact only if at least one of them belongs to a contact group. Contact groups are therefore

the bridges through which each population can be reached from the outside, and the knowledge they acquire then spreads inside their population through ordinary interactions.

Another fundamental methodological distinction in Phase 2 is its termination condition. Unlike Phase 1, which terminates dynamically upon reaching total behavioral consensus, the encounter phase is executed for a strictly fixed number of iterations. This fixed temporal budget aligns with the base experiment explained in Section 2.4.1. While the sole purpose of Phase 1 is to generate a stable cultural baseline, justifying a dynamic stopping criterion, Phase 2 aims to isolate the specific impact of the independent variables (e.g. *population distribution* and *network connectivity*) on the cultural merging process. If Phase 2 were allowed to terminate dynamically, the total number of inter-agent interactions would fluctuate drastically across different experimental conditions. Consequently, any observed differences in the final cultural states could be conflated with the volume of interaction opportunities. By resetting the iteration counter to zero at the exact moment of the encounter and enforcing a strictly fixed interaction budget, we guarantee that all populations across all experimental conditions are granted a mathematically identical temporal window to interact and merge.

This fixed-iteration design ensures that the final evaluation is statistically fair and that evaluation metrics can be reliably averaged across multiple independent runs without introducing missing values or temporal confounding variables.

At the conclusion of this phase, the final consensus decision function (h') and the structural snapshot of all agents' ontologies ($\mathcal{O}'$) are recorded to compute the final distance metrics.

3.4 Experimental Methodology

3.4.1 Independent Variables

Building upon the conceptual descriptions established previously, we now detail the rigorous mathematical and structural definitions for the three manipulated variables:

1. **Population distribution** (α): The proportional distribution of the total number of agents between the two populations prior to the encounter. This variable effectively manipulates the absolute size of each individual population. By convention, when $\alpha < 0.5$, Population 1 is the minority group. For example, if the total number of agents in the simulation is $N = 30$ and the distribution parameter is set to $\alpha = 0.25$, Population 1 will be formed of $30 \times 0.25 = 8$ agents (rounded up), while Population 2 will consist of the remaining 22 agents. This allows ranging from perfectly balanced configurations ($\alpha = 0.5$) to highly

unbalanced majority-minority encounters.

2. **Network Connectivity (η):** This variable defines the structural density of the inter-population possible interactions during the encounter phase. Specifically, it dictates the proportion of agents from each population that the other population can interact with.

   - When $\eta = 1.0$, the network is **fully connected**, meaning that every agent in Population 1 has the same probability of interacting with every agent in Population 2.
   - When $\eta < 1.0$, the network is **partially connected**, meaning that the inter-population interactions are restricted through a specifically sized contact group. For example, if $\eta = 0.2$, a randomly selected 20% of the agents from each population form its contact group and act as bridges: they are the only agents of their population that can be paired for an interaction with an agent of the other population outside its contact group.

   Figure 3.1 provides a visual representation of both configurations.

3. **Environment Overlap (ω):** This variable quantifies the fraction of shared objects between the training sets of the two populations during their initial, independent learning phases. Let $I_1 \subset I$ be the subset of objects on which Population 1 is trained, and $I_2 \subset I$ be the subset on which Population 2 is trained, where $I = I_1 \cup I_2$. The environment overlap is set such that:

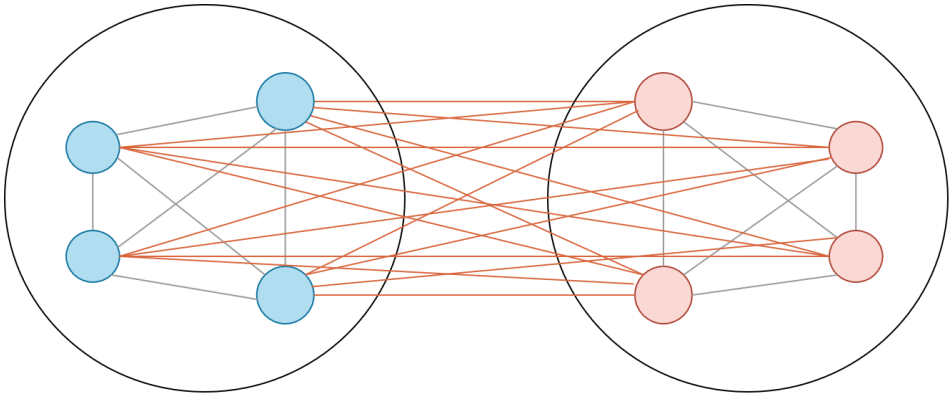
$$|I_1 \cap I_2| = \omega \cdot |I| \tag{3.1}$$

   The value of ω ranges from 0 (the populations are trained on completely disjoint subsets of the environment) to 1 (both populations are trained on the whole environment's objects). Figure 3.2 provides a visual representation of disjoint and partially overlapping training sets.

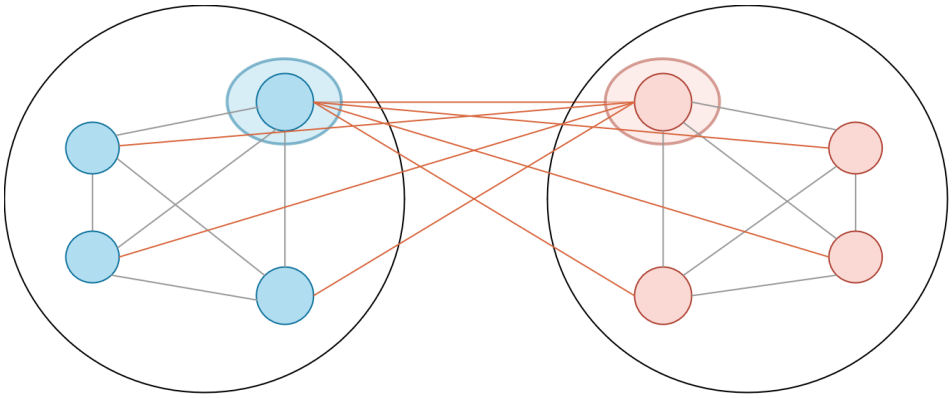
   It is important to remember that ω defines the overlap strictly at the *population* level. The actual training set provided to any individual agent within that population is further governed by the training ratio parameter (r). As a consequence, even in scenarios with full environment overlap ($\omega = 1$), individual agents do not typically possess identical knowledge; they each learn from uniquely sampled subsets of their population's available pool, preserving internal diversity.

3.4.2 Measures

We assess cultural change through two distinct measures to explicitly separate external actions from internal reasoning.

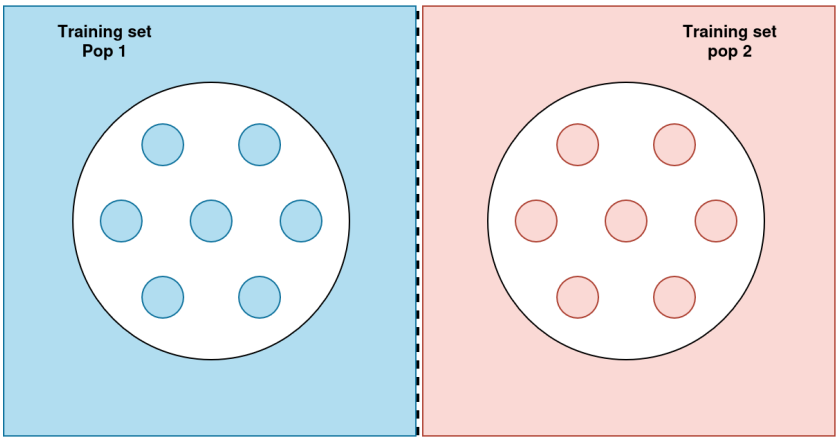


(a) Fully connected network ($\eta = 1.0$)

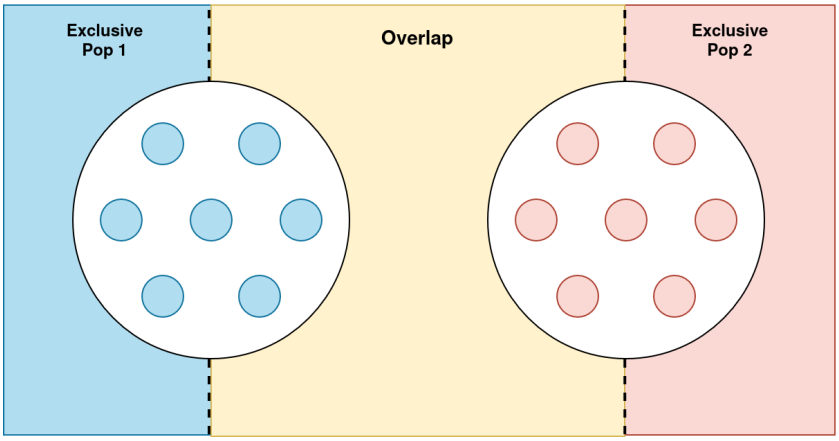


(b) Partially connected network ($\eta < 1.0$)

Figure 3.1: Network connectivity (η) determines which agents are permitted to interact across populations during the encounter phase. Grey edges denote intra-population interactions, which are never constrained; orange edges denote the admissible inter-population interactions. In (a), $\eta = 1.0$ and every agent may interact with every member of the other population. In (b), $\eta < 1.0$ and each population can be reached from the outside only through its contact group (shaded): an agent may interact with the contact-group members of the other population, whereas pairs in which neither agent belongs to a contact group are excluded. To make the illustration clearer, both panels show a reduced configuration of four agents per population, with a single-agent contact group in (b).



(a) Disjoint training sets ($\omega = 0$)



(b) Partial overlap ($0 < \omega < 1$)

Figure 3.2: Environment overlap (ω) fixes the fraction of the object space shared by the two populations' training sets during the individual training phase. In (a), $\omega = 0$ and each population is trained on a completely disjoint region of the object space. In (b), $0 < \omega < 1$ and a shared region (yellow) belongs to both training sets, while each population retains its own exclusive region. Region widths are for illustration purposes only and do not reflect the tested values of ω.

Decision Distance captures the behavioral level, identifying to which of the two pre-encounter cultures the final merged population's decisions are most similar.

Ontology Distance captures, instead, the semantic level, evaluating the structural diversity of the agents' knowledge representations.

Motivation

This dual approach is required because knowledge within this multi-agent system is represented as formal ontologies. As demonstrated in previous studies using this framework, agents within a population can reach a state of total behavioral consensus without necessarily converging on the same underlying ontological structure. Therefore, cultural change operates on two distinct levels that a single metric cannot fully capture: the **behavioral level** (the outward decisions agents make) and the **semantic level** (the internal structural logic used to reach those decisions).

This distinction defines exactly what each measure captures that the other misses. Relying solely on behavioral output might falsely indicate perfect cultural assimilation, completely missing the novel structural recombination happening internally. Contrarily, measuring only ontological changes might obscure the fact that one population's practical decision-making has entirely dominated the other.

Together, these two measures provide a comprehensive view of both the external actions and the internal logic of the post-encounter culture.

Decision Distance

To assess how culture has changed without averaging complex ontologies, we measure the behavioral output of the populations through the **Decision Distance** (d_d). Given any two generic decision functions, h_x and h_y, the distance between them is defined as:

$$d_d(h_x, h_y) = 1 - \frac{|\{o \in I \mid h_x(o) = h_y(o)\}|}{|I|} \tag{3.2}$$

Following the notation of Bourahla et al. [1], h denotes the function mapping an object type o to a decision d. Since each population converges to a total consensus after stabilization, a single h represents an entire population's behavior.

In simple terms, this formula calculates the ratio of behavioral disagreement between two cultures. The numerator counts the exact number of object types(o) for which both populations make the identical decision ($h_x(o) = h_y(o)$). We then divide this count by the total number of object types in the overall environment ($|I|$) to compute the proportion of total agreement. Finally, taking the complement to 1 of this agreement ratio yields the proportion of disagreement, which acts as our absolute behavioral distance.

Ontology Distance

To complement the behavioral measure, we assess cultural change at the knowledge representation level through **Ontology Distance** (d_o), which reflects the semantic dissimilarity between cultures. As demonstrated in the baseline study [1], agents within a population can reach behavioral consensus without necessarily converging on the same ontology. Therefore, unlike Decision Distance (d_d), a single representative function is not sufficient here; the culture of a population must be represented as the full set of its individual agents' ontologies.

When computing the distance between sets of ontologies, one might consider finding a single "representative" ontology for the whole population, such as a centroid or a medoid, and applying the individual ontology distance formula (Formula 2.2). However, this approach is problematic in this context. First, because ontologies are complex logical structures rather than numerical vectors in a Euclidean space, calculating an exact mathematical "centroid" (an average ontology) is not well-defined. Second, while calculating a medoid (the most centrally located existing ontology in the set) is possible, selecting a single representative would risk discarding the rich structural diversity present within the population, which is precisely the phenomenon this metric aims to capture.

To overcome this, we treat each population's culture as a discrete cluster of ontologies and measure the semantic divergence between them using **average linkage**. In standard clustering methodology, average linkage defines the distance between two clusters as the arithmetic mean of all pairwise distances between the individual data points of the first cluster and those of the second. Applying this principle to our context, given two generic sets of ontologies, $\mathcal{O}_x$ and $\mathcal{O}_y$, the average linkage distance is formally defined as:

$$d_o(\mathcal{O}_x, \mathcal{O}_y) = \frac{1}{|\mathcal{O}_x||\mathcal{O}_y|} \sum_{O_i \in \mathcal{O}_x} \sum_{O_j \in \mathcal{O}_y} \delta(O_i, O_j) \tag{3.3}$$

In simple terms, this formula calculates the mean semantic distance between every possible pair of agents across the two populations. It takes the base ontology distance δ (Formula 2.2) between every individual ontology O_i in the first set and every ontology O_j in the second set, sums all these distances together, and divides by the total number of pairs. This method ensures that every single agent's unique knowledge structure contributes equally to the final population-level distance.

The Geometric Classification Scheme

Because the metrics used to evaluate cultural change are raw continuous distances, there is no immediate, natural way to map them directly onto the two discrete cultural macro-states of *assimilation* and *hybridization*. To resolve this, we introduce

a categorical *Geometric Classification Scheme* designed to convert these raw distances into one of the specific macro-state outcomes. This analytical methodology is defined in its general mathematical form below, and is subsequently instantiated for both the behavioral and semantic levels.

Given three elements in a metric space where relative distances between elements are well-defined scalars, let A and B represent the baselines of the two pre-encounter cultures (Population 1 and Population 2, respectively), and let C represent the post-encounter culture. We first compute the distance between the two original cultures, $d(A, B)$, and define a radius:

$$\rho = \frac{d(A, B)}{2} \tag{3.4}$$

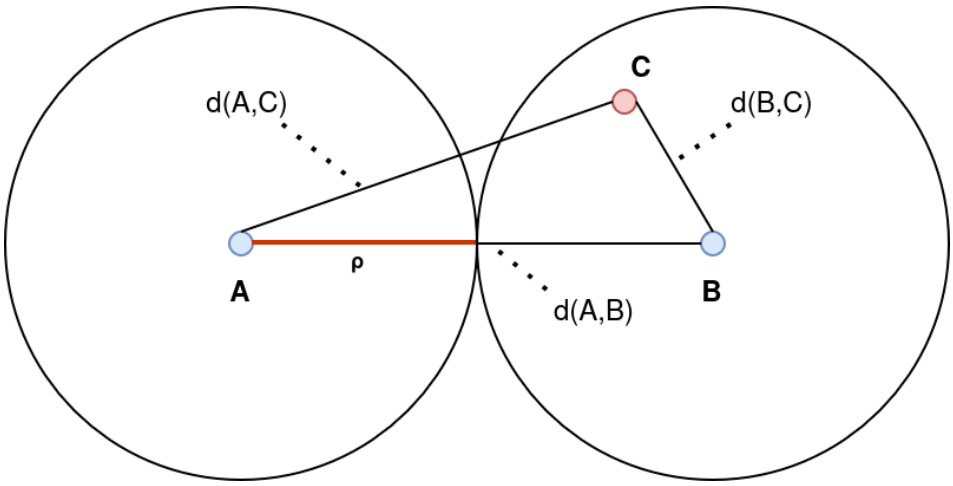
This establishes two spheres of radius ρ centered on A and B, such that they touch at exactly one point. The decision distance d_d is a proper metric and therefore satisfies the triangle inequality ($d(A, B) \leq d(A, C) + d(C, B)$). For the average-linkage ontology distance d_o, no such guarantee exists in general for set-based distances; we nonetheless assume the triangle inequality holds and treat the resulting semantic classification as a geometric interpretation rather than a strict metric guarantee. Under this assumption, the spatial positioning of C relative to A and B is well-defined, and we can classify the macro-state of C into four distinct geometric cases:

1. $d(A, C) < \rho \wedge d(B, C) > \rho$: C falls inside A's sphere. $\rightarrow$ **Assimilation** by Population 1.
2. $d(A, C) > \rho \wedge d(B, C) < \rho$: C falls inside B's sphere. $\rightarrow$ **Assimilation** by Population 2.
3. $d(A, C) < \rho \wedge d(B, C) < \rho$: Impossible in a valid metric space unless $A = B$.
4. $d(A, C) > \rho \wedge d(B, C) > \rho$: C falls outside both spheres. $\rightarrow$ **Hybridization**.

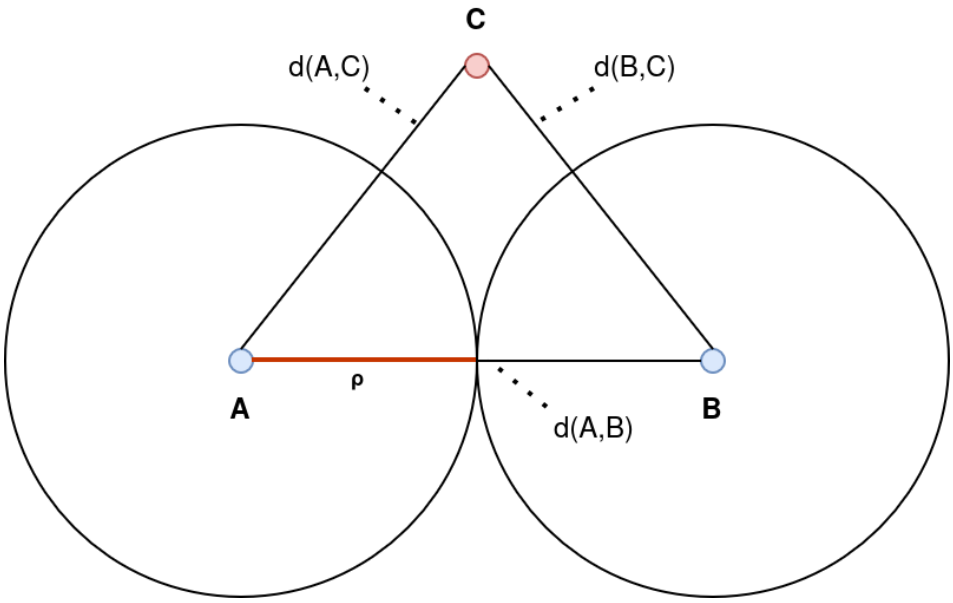
Figure 3.3 illustrates two concrete applications of the scheme. In the assimilation example (a), the post-encounter culture C is positioned closer to Population 2 than to Population 1: $d(B, C) < \rho$ places it inside B's sphere while $d(A, C) > \rho$ confirms it has drifted far from A's original culture, classifying the outcome as assimilation by Population 2. In the hybridization example (b), C instead falls outside both spheres: neither original culture can claim the final outcome as its own, and C represents a genuinely novel culture that emerged from the encounter.

Instantiating the Classification Scheme

Behavioral Classification: At the behavioral level, the scheme is instantiated using the Decision Distance (d_d) with the consensus decision functions h:



(a) Example of Assimilation



(b) Example of Hybridization

Figure 3.3: Two examples of the geometric classification scheme applied to cultural outcomes. In (a), the post-encounter culture C lies inside Population 2's sphere ($d(B, C) < \rho$) while remaining outside Population 1's sphere ($d(A, C) > \rho$), classifying the outcome as Assimilation by Population 2. In (b), C lies strictly outside both spheres ($d(A, C) > \rho$ and $d(B, C) > \rho$), indicating that the final culture cannot be attributed to either original population, a state of Hybridization.

- $A = h_1$: consensus decision function of Population 1 **before** the encounter.
- $B = h_2$: consensus decision function of Population 2 **before** the encounter.
- $C = h'$: consensus decision function of the merged population **after** the encounter.

Using these specific parameters, we compute the threshold $\rho_D = \frac{d_d(h_1,h_2)}{2}$ to classify the categorical outcome based on the four cases.

Semantic Classification: At the semantic level, the schema is instantiated using the Ontology Distance (d_o) computed via average linkage across the sets of agent ontologies:

- $A = \mathcal{O}_1 = \{O_1, \ldots, O_n\}$: set of agents' ontologies of Population 1 **before** the encounter.
- $B = \mathcal{O}_2 = \{O_1, \ldots, O_m\}$: set of agents' ontologies of Population 2 **before** the encounter.
- $C = \mathcal{O}' = \{O_1, \ldots, O_k\}$: set of agents' ontologies of the merged population **after** the encounter.

Using these specific parameters, we compute the threshold $\rho_O = \frac{d_o(\mathcal{O}_1,\mathcal{O}_2)}{2}$ to classify the categorical outcome based on the four geometric cases.

3.4.3 Experiment Plan

With the simulation phases, evaluation metrics, and analytical tools formally defined, this section outlines the specific parameter configurations used to execute the experiments.

Fixed Experimental Parameters

To ensure that any observed cultural shifts are strictly the result of the cross-cultural encounter, we prevent the introduction of confounding variables by holding the baseline simulation parameters constant. As detailed in Table 3.1, these specific values were adopted based on the results established by Bourahla et al. [1].

The **Number of features** dictates the boolean properties describing an object, while the **Number of decisions** defines the total possible classifications an agent can map to them. Regarding the learning phases, the **Training ratio** (r) determines the proportion of the environment provided to agents as initial labeled samples to build their starting ontologies, and the **Task ratio** (t) sets the proportion of objects randomly presented during the game phase to evaluate performance and calculate payoffs.

Parameter	Value	Justification
Number of agents	30	Balances sufficient population scale with manageable computation.
Number of features	4	Generates a diverse object space ($2^4 = 16$ objects).
Number of decisions	3	Ensures sufficient semantic diversity.
Training ratio (r)	0.5	Provides adequate initial knowledge coverage from a reduced pool.
Task ratio (t)	0.2	Sufficient evaluation subset for stable payoff computation.
Encounter iterations	150,000	Sufficient to reach near-complete stabilization.
Number of runs	10	Provides adequate statistical reliability across stochastic variation.

Table 3.1: Fixed experimental parameters applied across all simulation runs.

Variable Parameters and Factorial Design

To evaluate our specific hypotheses, the independent variables introduced in the current study are systematically varied to explore both balanced states and extreme edge cases. The complete parameter space explored in this study is detailed in Table 3.2. To ensure a rigorous and comprehensive analysis, we employ a full factorial plan, executing every possible combination of these parameters exactly 10 times (defined as the *Number of runs*).

Parameter	Tested Values
Population distribution (α)	$\{0.12, 0.25, 0.5\}$
Network connectivity (η)	$\{0.2, 0.5, 1.0\}$
Environment overlap (ω)	$\{0.0, 0.33, 0.66, 1.0\}$

Table 3.2: Grid of the independent variables varied across experiments.

Because the parameters α and η are proportional ratios, they interact to define the absolute, discrete size of the populations and their respective contact groups. Table 3.3 explicitly details these derived sizes for a total population of $N = 30$ agents.

3.5 Implementation

Realizing the protocol described in Section 3.3 under the variables defined in Section 3.4 required extending the framework used to implement Bourahla et al.'s interaction-driven adaptation model (Section 2.4.1). This extension proceeds along four axes: (1) introducing a notion of population, grouping agents that

Configurations		Absolute Pop. Size		Contact Group Size	
α	η	Population 1	Population 2	CG Pop 1	CG Pop 2
0.12	0.2	4	26	1	5
	0.5			2	13
	1.0			/	/
0.25	0.2	8	22	2	4
	0.5			4	11
	1.0			/	/
0.50	0.2	15	15	3	3
	0.5			8	8
	1.0			/	/

Table 3.3: Derived absolute sizes for populations and contact groups based on the interaction of α and η (Total Agents $N = 30$).

the baseline treated as a single undifferentiated set; (2) splitting agents into populations proportionally to the *population distribution* (α) parameter, optionally assigning each population a contact group according to the *network connectivity* (η) parameter, and partitioning the environment into different training sets according to the *environment overlap* parameter (ω); (3) dividing the simulation into two phases: an individual *stabilization phase*, in which each population evolves independently, followed by the merging of the populations and the start of their *encounter*; (4) computing and saving the metrics defined in Section 3.4.2.

The remainder of this section first describes the baseline framework and then details each of these extensions.

3.5.1 Software Framework

The simulation is implemented in *Lazy lavender*, a Java framework for cultural knowledge evolution [16]. The framework provides the following core components: `LLExperiment`, an interface defining the life-cycle of a run through two main functions, `init()`, which sets up the environment and trains the agents, and `process()`, which repeatedly alternates, over a certain number of iterations, between agents performing tasks against part of the environment and playing pairwise interaction games; `Environment`, which defines the object space: it generates object instances over a number of binary features, assigns each instance one of a fixed number of decisions, and provides a function to sample instances from this space; `Agent`, which holds the agent's ontology, represented as a decision tree induced via ID3, together with the split-based adaptation operator used to resolve interaction failures; and `GameIterator`, which implements the scheduling policy of the simulation, selecting at every iteration the two interacting agents and the object presented to

them. `LLExperiment` is implemented by `Experiment`, the baseline class realizing the interaction-driven adaptation model (Section 2.4.1); the present work extends it through the `PopulationEncounter` experiment class, described in the following subsections.

Beyond these main components, the framework decouples run parameters from code: a run is configured through a set of Java properties, stored in a `params.sh` file, passed as command-line arguments, and dispatched by a dedicated `ExperimentalPlan` engine class that instantiates and runs the requested experiment with the given parameters. Configuring runs this way allows every combination of the independent variables to be executed through a single script.

This thesis reuses the agent-level adaptation mechanism and the object/decision model of the baseline unchanged, and builds on them to implement the extensions introduced above.

3.5.2 The Population Abstraction

Prior to this work, *Lazy lavender* had no notion of a population: every existing experiment operated on a single, undifferentiated set of agents. Supporting the encounter of two independently stabilizing populations therefore required a new class, `DecisionPopulation`, designed to be reusable by any future experiment involving this type of decision-taking agent. It is deliberately decoupled from any concrete experiment: the agents it manages and the set of objects it operates on are always passed in as arguments by the caller, rather than held or determined by the class itself, so it carries no dependency on the environment or on `PopulationEncounter`. A population is represented as a list of member agents, each identified by a unique identifier. It may also include an optional contact group, a subset of these identifiers used during the *encounter phase*. Around this representation, the class provides the following responsibilities:

- **Membership.** Agents are added up to the population's target size, and a contact group of a given size can be drawn at random from its members.
- **Stabilization check.** Given a set of object types, `isTotalAgreement()` returns whether every member agent reaches the same decision on every object type, i.e. whether the population has reached full internal behavioral agreement; `computeConsensus()` relaxes this into a continuous measure, returning, for each object type, the fraction of agents sharing the majority decision, averaged over the object set.
- **Behavior extraction.** `getUnanimityDecision()` extracts the population's decision function h as a mapping from object types to decisions. Under total agreement this is simply the decision of an arbitrary member, since by

definition all members agree; a fallback mode also allows extracting h without total agreement, in which case the majority (more popular) decision is taken independently for each object type.

The same class is used to represent both of the two pre-encounter populations and the merged population resulting from their encounter, so that all of the functionalities above are reused identically for both.

3.5.3 Implementing the Independent Variables

The three independent variables of this study are realized inside the `init()` function of `PopulationEncounter`, once their values for the current run have been read and assigned to the corresponding fields by `loadParams()`.

Population Distribution

The *population distribution* (α) determines the size of the first population, computed as round($\alpha \cdot N$), where N is the total number of agents in the simulation; the second population is assigned all remaining agents. The implementation generalizes beyond two populations: once the first population has been assigned its share, any remaining agents are divided equally among additional populations. However, only the case of two populations is exercised by the experiments reported in this study, and it is the one described throughout this section. Consequently, given the values of α considered in this study, all ≤ 0.5 (Table 3.2), population 1 is always the minority, or, when $\alpha = 0.5$, of equal size to population 2.

Network Connectivity

When the *network connectivity* (η) is lower than 1, the two populations are no longer fully connected during the *encounter phase*: each instead interacts through a contact group, a subset of its members reachable by the other population. The size of population i's contact group is computed as round($\eta \cdot n_i$), where n_i is the total number of agents of that population. This allows each contact group to scale proportionally with the size of its population. The members of the contact group are then drawn at random using the corresponding method of `DecisionPopulation` (Section 3.5.2), which takes this computed size as a parameter.

Environment Overlap

Also within `init()`, two further functions are called in sequence: `loadTrainingSets()`, which builds each population's training set, and `trainAgents()`, which trains the agents on it. The training set of each population is built by combining two

parts: a shared *overlap block*, common to every population, whose size is computed as round($\omega \cdot |I|$), where I is the full object space; and a disjoint, population-specific *exclusive block*, over the remaining objects. Each population's training set is the union of the *overlap block* and its own *exclusive block*. Every agent of a population is then trained on a subset of its own population's training set, drawn through Bernoulli sampling at the training ratio (r), fixed to 0.5 throughout this study.

3.5.4 The Two-Phase Game Iterator

The simulation is organized into two phases: individual *stabilization*, in which the two populations evolve independently, and *encounter*, in which they are merged and interact as a whole. The interaction game played in both phases is the same one used by the baseline model; what changes between phases is which agents are eligible to be paired at a given iteration. This is governed by a single iterator, `PopulationEncounterGameIterator`, driving the main loop of `PopulationEncounter`'s `process()` method, which runs until the iterator signals that no further iterations remain. Figure 3.4 visually summarizes its selection logic.

At each iteration during the *stabilization phase*, the iterator selects a population at random, then two distinct agents at random from its members. The object presented to them is drawn from the full object space I. Before the pair plays, every agent of the selected population has its payoff recomputed against a freshly drawn sample of I, taken at the task ratio (t); the two selected agents then play the interaction game, which may trigger an ontology adaptation. The *stabilization phase* has no fixed iteration budget: it continues until every population has independently reached total agreement.

During the *encounter phase*, partner selection depends on the *network connectivity* (η). When the two populations are fully connected ($\eta = 1$), the iterator draws two agents uniformly at random from the entire merged population. When they are not, one population is selected at random and a pool is formed from all of its members together with the contact group of the other population; the two agents are then drawn uniformly from that pool. A cross-population pairing therefore always involves a contact-group member of the non-selected population, while its partner may be any member of the selected population. In both cases, the object presented is drawn from the full object space, and payoffs are recomputed for every agent before each game, exactly as in the *stabilization phase*. The *encounter phase* runs for a fixed number of iterations.

How the transition from the *stabilization phase* to the *encounter phase* is triggered, and how the *encounter phase* terminates, are detailed in Section 3.5.5.

3.5.5 Detecting Stabilization and Recording Metrics

Every 500 iterations, the simulation checks whether the *stabilization phase* should end. This check requires every population to have reached total agreement, verified over the full object space through `isTotalAgreement()` (Section 3.5.2). Once triggered, the transition performs three actions. First, the simulation is marked as being in the *encounter phase*, and the iterator is reset and given a fixed budget of 150 000 iterations. Second, the decision function of each population, h_1 and h_2, is extracted through `getUnanimityDecision()` (Section 3.5.2). Third, since an ontology cannot summarize an entire population, the ontology of every individual agent in both populations is deep-copied at this exact moment, forming the snapshots O_1 and O_2 used later for computing the ontology distance. Only interactions occurring during the *encounter phase* are recorded through the framework's built-in logging utilities, which is what allows the evolution of the interaction success rate during the *encounter phase* to be traced. When the *encounter phase*'s fixed budget is exhausted, a final routine computes all remaining metrics. The population's final decision function h' is extracted from the merged population: if it happened to reach exact total agreement within the fixed budget, the same total-agreement shortcut used for h_1 and h_2 applies; otherwise h' falls back to the majority-vote extraction, since, unlike the *stabilization phase*, the *encounter phase* is not guaranteed to end in full agreement. From h_1, h_2, and h', the three decision distances $d_d(h_1, h_2)$, $d_d(h', h_1)$, and $d_d(h', h_2)$ are computed (Equation 3.2). Analogously, O' is taken as the set of the final ontologies of all agents, and the three ontology distances $d_o(O_1, O_2)$, $d_o(O', O_1)$, and $d_o(O', O_2)$ are computed via average linkage (Equation 3.3). All computed metrics are written to a single `metrics.tsv` file per run, stored in an output folder structured according to the run's combination of independent variables.

Outlook. Having defined the encounter scenario, the formal hypotheses, the two-phase protocol, the evaluation measures, the full factorial plan, and their implementation, the experimental setup is now complete. The next chapter presents and analyses the results hypothesis by hypothesis.

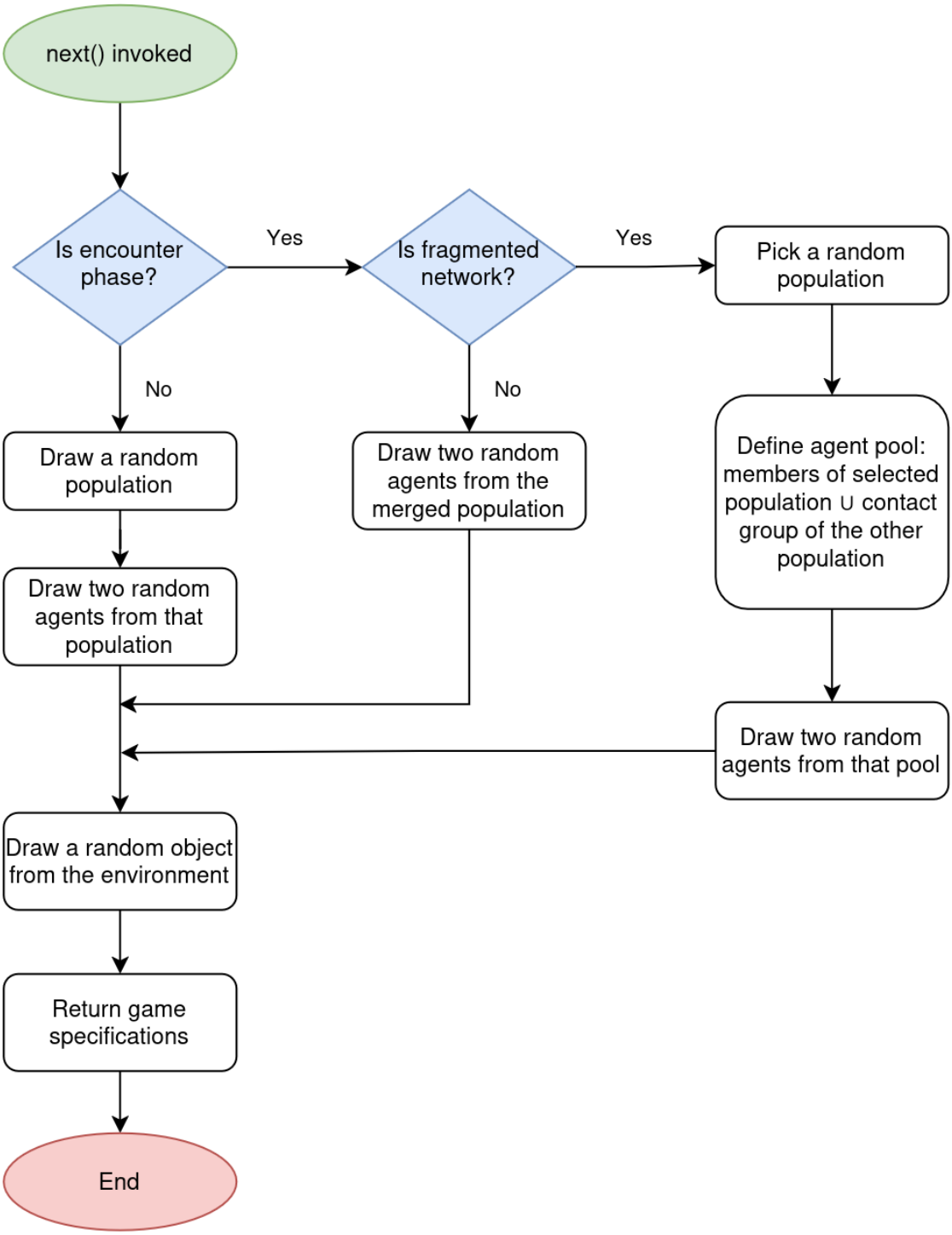


Figure 3.4: Selection logic of `PopulationEncounterGameIterator`'s `next()` function, showing agent and object selection rules for both the stabilization and encounter phases.

Chapter 4

Results and Analysis

This chapter presents the results and analyses them in light of the four hypotheses formulated in Chapter 3. Each hypothesis is examined in turn, first presenting the results, then discussing their interpretation. The analysis initially relies on the geometric classification scheme (Section 3.4.2); its limitations are then exposed, motivating a further analysis based on a new continuous metric (Section 4.5.2).

4.1 Hypothesis 1: Encounter Stabilization

4.1.1 Results

Figure 4.1 illustrates the average success rate in agent interactions across all runs during the encounter phase. The curve trends upward throughout the 150,000 iterations with decreasing standard deviation, indicating progressive convergence toward unanimous decision-making.

To assess the final stabilization of the merged population independently of the cumulative success rate, we compute the *final consensus* metric at the end of each simulation run. This metric directly calculates the average percentage of agents agreeing on the most popular decision for each object across the entire environment, where a value of 1.0 indicates perfect behavioral unanimity.

Of the 360 runs (36 parameter combinations $\times$ 10 runs), 351 reached perfect consensus (final consensus = 1.0). The remaining 9 runs achieved a final consensus between 0.98 and 0.99, distributed across different combinations of parameters. Despite these marginal deviations, the average final consensus across all runs was 0.9996.

Hypothesis 1 is supported: the cross-cultural encounter consistently drives the merged population toward absolute behavioral consensus, and so stabilization.

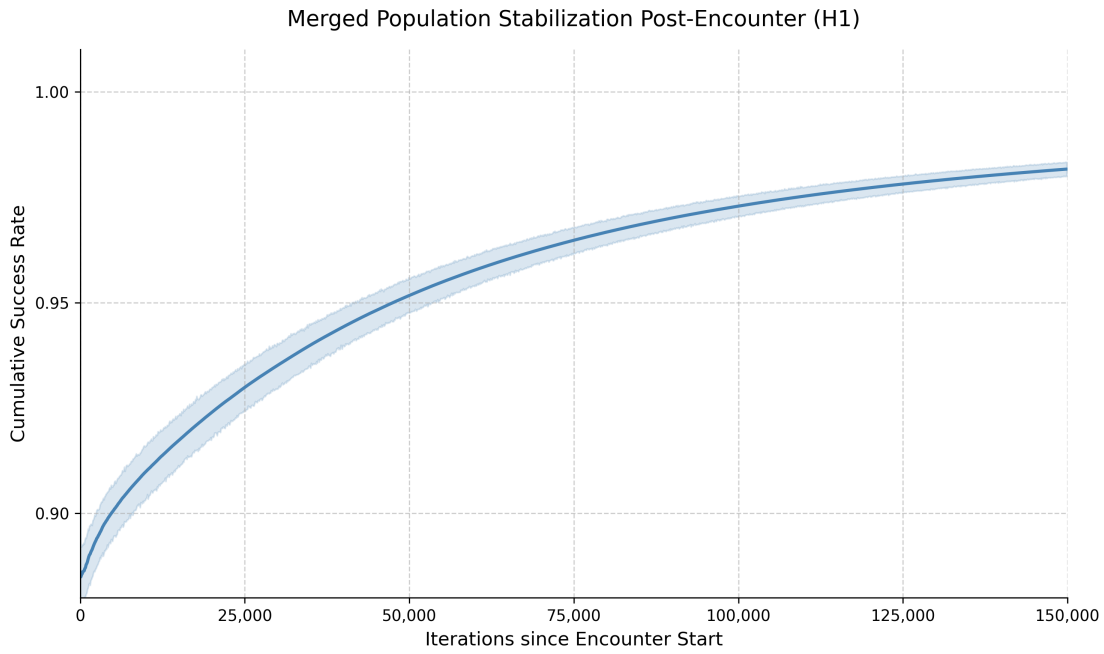


Figure 4.1: Average success rate of the post-encounter iterations

4.1.2 Discussion

The convergence to universal agreement observed within the population is a direct structural consequence of the adaptation mechanism. Whenever an interaction fails, the adaptation operator forces the less successful agent to adjust its internal ontology to align with its partner. Given sufficient time, this mechanism guarantees convergence toward population-wide decision unanimity.

The apparent discrepancy between the visual success rate curve, which reaches only about 0.98 by the end of the encounter phase rather than 1.0, and the final consensus values requires clarification. The success rate is a cumulative metric, computing the ratio of all successful games to all games played since the start of the encounter. Hence, it reflects a historical log of past interactions rather than the actual, real-time internal state of the agents. Even if the population reaches total consensus, meaning the agents' ontologies, at that time, are perfectly aligned on the decisions and all subsequent games succeed, the denominator perpetually carries the burden of the unsuccessful games played during the turbulent early phase of the encounter. This "weight of the past" causes the metric to asymptotically approach, but never reach, 1.0 within the fixed iteration budget. The success rate is therefore a useful visualization of the convergence trajectory, but an inadequate judge of the agents' true stabilization state.

Initially, one might expect that the 9 runs failing to reach perfect consensus were

exclusively the result of problematic, extreme parameter configurations. Instead, their spread across a diverse set of parameter combinations reveals that this phenomenon is a general property of the system's inherent stochasticity combined with the fixed iteration budget.

To understand why, we can consider a scenario where a single agent holds a divergent decision for one specific object in the environment. For the agent to successfully align with the majority, a highly specific sequence of random events must occur simultaneously, specifically: first, the agent must be randomly selected for an interaction; second, the randomly chosen object for that specific game must be the exact point of disagreement; and third, the agent must happen to hold a lower payoff than its randomly assigned partner. While connectivity limitations or highly asymmetric populations can further decrease the likelihood of these aligning events, the purely random nature of agent pairing means that, occasionally, 150,000 iterations are simply insufficient to satisfy all these probabilistic conditions for the final few dissenting agents in the simulation. This highlights a boundary condition of the fixed iteration limit rather than a fundamental violation of Hypothesis 1.

4.2 Hypothesis 2: Population Asymmetry and the Direction of Assimilation

4.2.1 Results

Tables 4.1 and 4.2 present the percentage of simulation runs falling into each classification category under the discrete geometric scheme, grouped by the *population distribution* parameter (α). Recall that, by convention, when $\alpha < 0.5$, Population 1 represents the minority group.

Chi-square tests confirm that the differences across population distributions are highly statistically significant. For the behavioral decision outcomes, the effect is robust ($p = 6.62 \times 10^{-29}$), demonstrating a medium-high effect size (Cramér's $V = 0.44$). Similarly, the ontology outcomes exhibit significant categorical shifts ($p = 2.25 \times 10^{-5}$), although with a smaller effect size (Cramér's $V = 0.24$).

Figure 4.2 visualizes the behavioral decision outcomes, making the shift toward Assimilation by Population 2 as α decreases immediately apparent.

Hypothesis 2 is supported: demographic size asymmetry directly dictates the direction of cultural assimilation, with the resulting culture consistently skewing toward the pre-encounter majority.

α	**Assimilation Pop 1**	**Assimilation Pop 2**	**Hybridization**
0.12	0.0%	99.2%	0.8%
0.25	0.8%	89.2%	10.0%
0.50	28.3%	40.0%	31.7%

Table 4.1: Behavioral Decision Outcomes (%) across varying Population Distribution (α) values.

α	**Assimilation Pop 1**	**Assimilation Pop 2**	**Hybridization**
0.12	0.0%	12.5%	87.5%
0.25	0.0%	1.7%	98.3%
0.50	0.0%	0.8%	99.2%

Table 4.2: Semantic Ontology Outcomes (%) across varying Population Distribution (α) values.

4.2.2 Discussion

The behavioral results are clear and consistent across the categorical scheme: as α decreases, assimilation by Population 2 (the majority) becomes increasingly dominant, rising from a near-even split at $\alpha = 0.50$ to near-total dominance at $\alpha = 0.12$. Notably, the scheme reports a slight apparent advantage for Population 2 at $\alpha = 0.50$ (40.0% vs. 28.3%). However, this is consistent with a balanced encounter: in symmetric conditions, minor stochastic drifts will occasionally tip the categorical threshold in either direction, and the near-equal split of assimilation cases confirms there is no genuine directional tendency.

Semantically, the scheme classifies the overwhelming majority of runs as Hybridization across all α values. This is tied to a structural property of the adaptation mechanism: ontologies only grow via the split operator and never shrink, so the post-encounter ontologies are substantially richer than the pre-encounter baselines. The resulting $d_o(\mathcal{O}_1, \mathcal{O}')$ and $d_o(\mathcal{O}_2, \mathcal{O}')$ are both naturally large, systematically placing $\mathcal{O}'$ outside both assimilation radii and forcing the Hybridization classification.

This structural artifact limits the utility of the geometric scheme on the semantic dimension, a limitation that is addressed in Section 4.5.

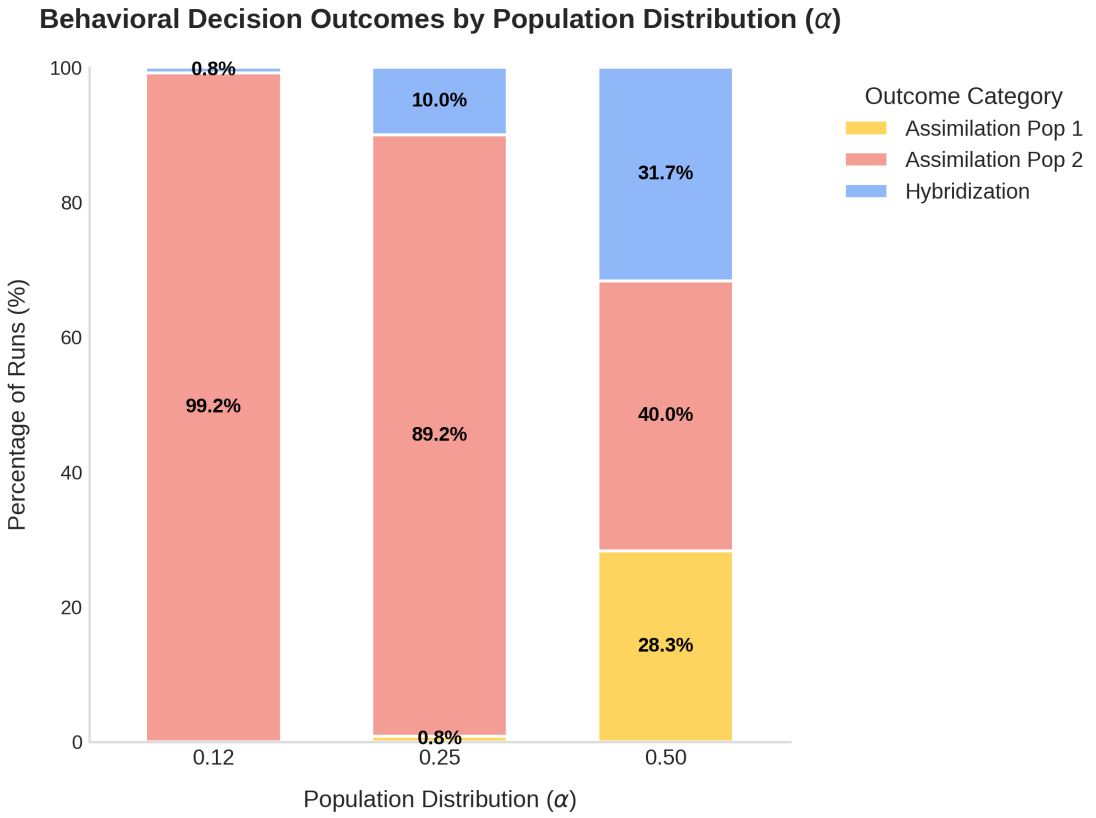


Figure 4.2: Stacked column chart of the Behavioral Decision Outcomes across varying Population Distribution (α) levels. The progressive Assimilation by Population 2 as α decreases is clearly visible.

4.3 Hypotheses 3: Network Connectivity and Information Bottlenecks

4.3.1 H3.1: Results

Taking into account the dominant effect of *population distribution* established by Hypothesis 2, the analysis first isolates the effect of *network connectivity* (η) under balanced conditions ($\alpha = 0.50$), where the size effect is neutralized.

Tables 4.3 and 4.4 present the geometric scheme classifications. No meaningful pattern is observed across connectivity levels, and Chi-square tests detect no statistically significant difference in outcome distributions (both $p > 0.05$).

Since the hypothesis predicts that the effect of connectivity is stronger when populations are asymmetric, implying a reduction in the absolute assimilation driven by the majority when the connectivity is lower, the analysis was extended to

η	Assimilation Pop 1	Assimilation Pop 2	Hybridization
0.2	27.5%	42.5%	30.0%
0.5	32.5%	37.5%	30.0%
1.0	25.0%	40.0%	35.0%

Table 4.3: Behavioral Decision Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Network Connectivity (η) levels.

η	Assimilation Pop 1	Assimilation Pop 2	Hybridization
0.2	0.0%	0.0%	100.0%
0.5	0.0%	0.0%	100.0%
1.0	0.0%	2.5%	97.5%

Table 4.4: Semantic Ontology Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Network Connectivity (η) levels.

the unbalanced configurations ($\alpha = 0.12$ and $\alpha = 0.25$). Tables 4.5 and 4.6 present these results. A marginal increase in hybridization is observable at $\eta = 0.2$ in the behavioral outcome (8.8% vs. 3.8% at higher connectivity), but this trend is not statistically confirmed by either Chi-square tests or Kruskal-Wallis tests on the overall distributions.

Hypothesis 3.1 is not supported.

4.3.2 H3.1: Discussion

The categorical results show no structured pattern attributable to connectivity: the distribution of outcomes is essentially flat across η values at both balanced and unbalanced population sizes. As already established for H2, the semantic dimension is dominated by the Hybridization artifact caused by ontology growth, which renders the scheme virtually insensitive to any connectivity effect on that dimension.

Before interpreting this null result, however, it is necessary to consider whether the manipulation reduced inter-population interaction as intended. Because the encounter pool is built by selecting one population uniformly at random and adding the contact group of the other (Section 3.5.4), the realised rate of cross-population pairings depends jointly on η and α. Under balanced conditions the intended reduction holds, the rate falling from 51.7% at $\eta = 1.0$ to 29.4% at $\eta = 0.2$; at $\alpha = 0.25$ it still holds, though far more weakly, falling only from 40.5% to 32.2%. Examining this trend further across the tested configurations shows that, under strong asymmetry, it does not hold at all. At $\alpha = 0.12$ with $\eta = 0.2$, the two branches of the selection behave very differently: when Population 2 is selected, the

η	Assimilation Pop 1	Assimilation Pop 2	Hybridization
0.2	1.2%	90.0%	8.8%
0.5	0.0%	96.2%	3.8%
1.0	0.0%	96.2%	3.8%

Table 4.5: Behavioral Decision Outcomes (%) under Asymmetric Population Distributions ($\alpha = 0.12, 0.25$) across varying Network Connectivity (η) levels.

η	Assimilation Pop 1	Assimilation Pop 2	Hybridization
0.2	0.0%	5.0%	95.0%
0.5	0.0%	11.2%	88.8%
1.0	0.0%	5.0%	95.0%

Table 4.6: Semantic Ontology Outcomes (%) under Asymmetric Population Distributions ($\alpha = 0.12, 0.25$) across varying Network Connectivity (η) levels.

pool contains its 26 members together with the single contact agent of Population 1, so cross-population pairs are rare; but when Population 1 is selected, the pool consists of its four members together with the five contact agents of Population 2, and more than half of the pairs drawn from it are cross-population. Since each branch occurs half of the time, the average rate is 31.5%, which is higher than the 23.9% obtained at $\eta = 1.0$, where pairs are drawn from the whole merged population and the minority accounts for only four agents out of thirty. Reducing connectivity therefore shrinks the pool in a way that concentrates the minority within it, and in the most asymmetric configuration of this study ($\alpha = 0.12$) restricted connectivity exposed the minority to the majority more, rather than shielding it.

This has direct consequences for how the null result should be read, since the buffering effect postulated by H3.1 concerns the protection of a minority and can therefore only be observed where a minority exists. At $\alpha = 0.50$ the manipulation is strongest, but the two populations are of equal size and there is no minority to protect. At $\alpha = 0.12$ a clear minority exists, but the manipulation is inverted and no reduction in contact was achieved.

The only configuration in which both conditions are satisfied is $\alpha = 0.25$, where the reduction in contact is real but modest; however, there too the outcome distribution is flat across connectivity levels. Because the implementation artifact invalidates the most asymmetric case, the absence of a connectivity effect under asymmetric population sizes cannot be attributed with certainty to a genuine lack of influence of network structure in this encounter scenario: the network fragmentation tested may be too loose. Alternatively, the null result may reflect a disagreement resolution mechanism that is too strong to allow the fragmented network to have a buffering effect.

4.3.3 H3.2: Results

To assess whether network connectivity (η) has an impact on the final structural heterogeneity of the merged population, we evaluated the **internal semantic diversity** at the end of each simulation run by computing the average pairwise Ontology Distance (Formula 2.2) between all individual agents within the post-encounter population.

The mean internal semantic diversity remained remarkably consistent across all connectivity levels: 0.4894 for strict connectivity ($\eta = 0.2$), 0.4760 for intermediate connectivity ($\eta = 0.5$), and 0.4800 for fully connected networks ($\eta = 1.0$). A Kruskal-Wallis H-test across the three connectivity groups detects no statistically significant difference in final semantic diversity ($p = 0.148$).

Hypothesis 3.2 is not supported: altering the network connectivity does not significantly impact the final internal semantic diversity of the merged population.

4.3.4 H3.2: Discussion

While a marginal increase in the mean final semantic distance is observed under the smallest contact-group size ($\eta = 0.2$, mean = 0.4894) compared to intermediate ($\eta = 0.5$, mean = 0.4760) and fully connected ($\eta = 1.0$, mean = 0.4800) configurations, the lack of statistical significance precludes any broad generalization.

This result is consistent with the findings of Hypothesis 3.1: in the tested configurations, network connectivity showed no significant influence on the final merged culture, in terms of neither the direction of assimilation nor the internal semantic diversity of the resulting population.

This null effect contrasts with existing cultural evolution literature, which has demonstrated that fragmented network structures preserve cultural diversity and foster the accumulation of more complex knowledge [11, 7]. However, this discrepancy does not invalidate those findings; rather, it highlights the specificity of the present context. Given sufficient iterations, the constraint to agree imposes too strong a pressure, overriding, at this level of fragmentation, the buffering effect that network fragmentation can provide.

Whether these conclusions hold in systems with more limited network connectivity, or where the adaptation mechanism is less strict, allowing agents to retain or revert decisions rather than always converging on a single one, remains an open question; here, we attribute the ineffectiveness of connectivity to the particular constraints of the knowledge representation and disagreement-resolution mechanism. This also points to a possible limitation of the experimental setting: the decision task may not be complex enough to produce the conditions under which network fragmentation has been shown to matter [7].

4.4 Hypotheses 4: Environmental Overlap and Cultural Dynamics

4.4.1 H4.1: Results

Following the same analytical approach used for Hypothesis 3, the analysis of H4.1 was conducted under balanced population conditions ($\alpha = 0.50$) to isolate the effect of *environment overlap* (ω).

Tables 4.7 and 4.8 present the geometric scheme classifications for behavioral and semantic outcomes, respectively. The outcomes reveal no clear monotonic pattern across the overlap levels. Chi-square tests detect no statistically significant difference in the distribution of outcomes across the four *environment overlap* configurations, for both the behavioral ($p = 0.197$) and semantic ($p = 0.388$) dimensions.

ω	**Assimilation Pop 1**	**Assimilation Pop 2**	**Hybridization**
0.00	30.0%	43.3%	26.7%
0.33	26.7%	46.7%	26.7%
0.66	26.7%	50.0%	23.3%
1.00	30.0%	20.0%	50.0%

Table 4.7: Behavioral Decision Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Environment Overlap (ω) levels.

ω	**Assimilation Pop 1**	**Assimilation Pop 2**	**Hybridization**
0.00	0.0%	3.3%	96.7%
0.33	0.0%	0.0%	100.0%
0.66	0.0%	0.0%	100.0%
1.00	0.0%	0.0%	100.0%

Table 4.8: Semantic Ontology Outcomes (%) under Balanced Population Distribution ($\alpha = 0.50$) across varying Environment Overlap (ω) levels.

Hypothesis 4.1 specifically predicted that higher environment overlap would shift the cultural evolution toward hybridization. Prompted by the observation that behavioral hybridization notably spiked to 50.0% under full overlap ($\omega = 1.00$), as clearly visualized in the stacked column chart in Figure 4.3, an exploratory comparison between the two extreme overlap conditions (completely disjoint $\omega = 0.00$ vs. full overlap $\omega = 1.00$) was conducted to test whether any signal exists at the absolute boundaries of this parameter. While behavioral hybridization increased from 26.7% of the runs at $\omega = 0.00$ to 50.0% at $\omega = 1.00$, a Chi-square

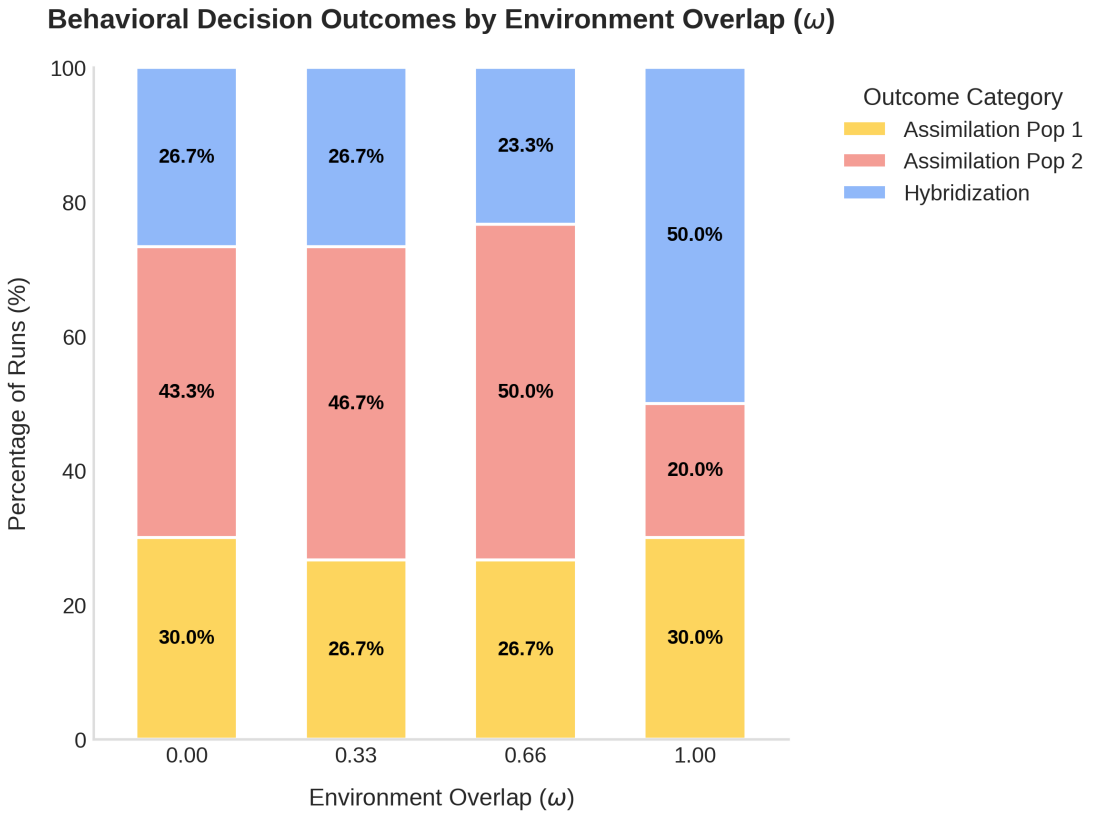


Figure 4.3: Stacked column chart illustrating the distribution of Behavioral Decision Outcomes across varying Environment Overlap (ω) levels. The chart visually highlights the non-monotonic spike in hybridization at full overlap ($\omega = 1.00$), though this shift does not reach statistical significance.

test indicates that this shift falls just short of statistical significance ($p = 0.095$), exhibiting a small-to-medium effect size (Cramér's $V = 0.28$).

Hypothesis 4.1 is not supported: altering the environment overlap does not systematically dictate the overall rate of cultural hybridization.

4.4.2 H4.1: Discussion

The categorical results do not reveal a systematic hybridization trend. The spike at $\omega = 1.00$ is the most notable observation, but it falls short of significance and, as the continuous analysis in Section 4.5 will show, is partly a boundary artifact driven by the near-zero pre-encounter behavioral distance that arises when both populations are trained on identical environments.

4.4.3 H4.2: Results

Consistent with the analytical approach used previously, the evaluation of Hypothesis 4.2 is conducted under balanced population conditions ($\alpha = 0.50$). To assess whether the degree of shared environment context impacts the final structural heterogeneity of the merged population, we evaluated the *internal semantic diversity* of the agents at the end of the simulation, following the same approach as Hypothesis 3.2.

A Kruskal-Wallis H-test conducted on the final internal semantic diversity across the different environment overlap configurations revealed a statistically significant difference ($H = 8.00$, $p = 0.046$), demonstrating a small-to-medium effect size ($\eta^2 = 0.043$).

An examination of the group means highlights the source of this statistical variance. While partial and full overlap configurations produced remarkably consistent levels of internal diversity (0.4901 for $\omega = 0.33$; 0.4917 for $\omega = 0.66$; and 0.4879 for $\omega = 1.00$), the completely disjoint configuration ($\omega = 0.00$) resulted in a noticeably lower mean semantic distance of 0.4562.

Hypothesis 4.2 is not supported: higher environment overlap did not decrease the final semantic diversity as predicted. Conversely, completely disjoint environments ($\omega = 0.00$) exhibited a lower diversity; however, given the marginal significance and small effect size, we treat this as an observation rather than a robust effect.

4.4.4 H4.2: Discussion

The observation that completely disjoint environments ($\omega = 0.00$) appear to produce the *least* diverse final culture is counterintuitive and requires careful examination. The initial hypothesis assumed a simple causal chain: higher overlap $\rightarrow$ more similar pre-encounter cultures $\rightarrow$ fewer adaptations $\rightarrow$ less diverse final culture.

Building upon the structural indications observed at the conclusion of Hypothesis 4.1, we first verified whether the foundational link in this causal chain holds true. To explicitly test if higher environmental overlap produces semantically more similar populations prior to contact, the baseline ontology distance $d_o(\mathcal{O}_1, \mathcal{O}_2)$ was evaluated. A Kruskal-Wallis test confirmed a highly significant monotonic decrease in this pre-encounter distance ($H = 64.34$, $p = 6.93 \times 10^{-14}$, $\eta^2 = 0.53$). The group means perfectly follow the predicted trend, dropping from 0.740 ($\omega = 0.00$) to 0.685 ($\omega = 0.33$), 0.557 ($\omega = 0.66$), and finally 0.517 ($\omega = 1.00$).

This validates the first step of the assumption: increased environmental overlap definitively yields populations with more uniform pre-encounter structural baselines. However, the subsequent assumption, that this pre-encounter similarity leads to fewer adaptations and a less diverse final outcome, is not supported by the results: the final diversity does not decrease with increasing overlap.

Figure 4.4 illustrates this phenomenon directly. The dashed line shows, as demonstrated above, the pre-encounter distance decreasing monotonically with increasing overlap, confirming greater structural similarity before contact. The solid red line, representing the final internal diversity of the merged population, initially moves in the opposite direction: it is lowest precisely where the pre-encounter distance was greatest.

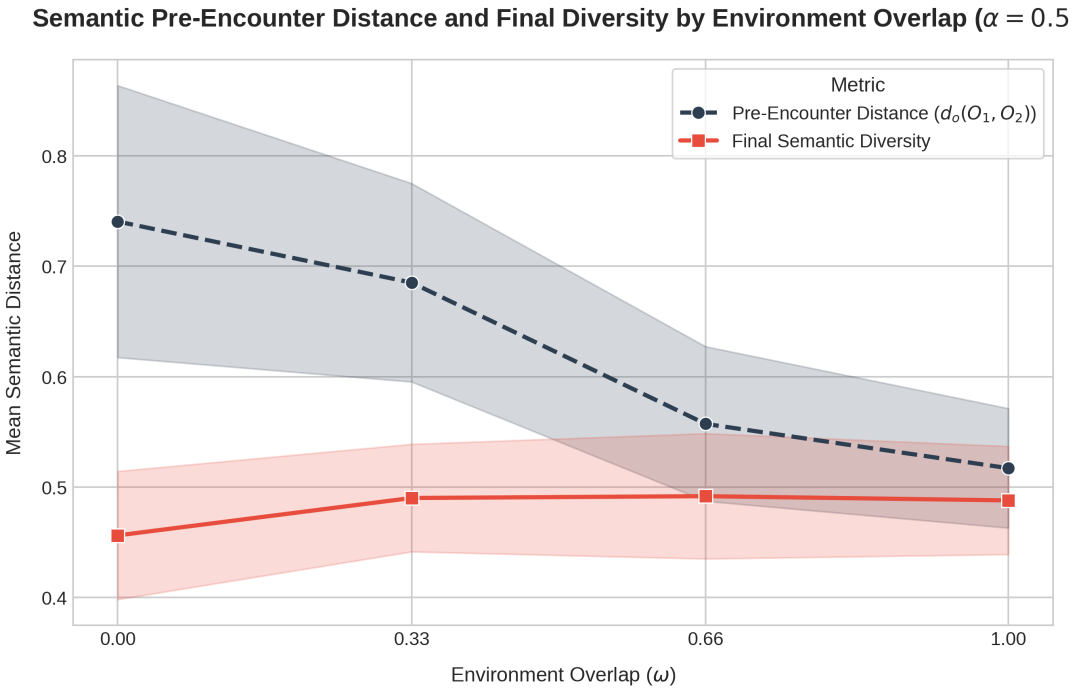


Figure 4.4: Line plot comparing the Pre-Encounter Semantic Distance ($d_o(\mathcal{O}_1, \mathcal{O}_2)$) and Final Semantic Diversity across varying Environment Overlap (ω) levels.

The key to resolving this apparent "paradox" is to recall precisely what the final-diversity metric captures: the internal heterogeneity *among* the agents of the merged population, not its distance from the two pre-encounter baselines. A more precise interpretation therefore emerges when the focus shifts from *how many* adaptations occur to *what kind* of adaptation process unfolds. In the completely disjoint configuration ($\omega = 0.00$), each population's initial classifiers are built from entirely separate training examples. Even though both populations stabilize their internal consensus by interacting over the full environment, the decision functions they converge to, h_1 and h_2, are built on incompatible conceptual foundations and therefore disagree on a large fraction of objects. This high pre-encounter disagreement rate triggers an extensive, system-wide adaptation process during the encounter phase that touches virtually all agents across both populations. Rather

than making incremental adjustments to an already mostly shared structure, the agents must collectively negotiate and reconstruct functionally compatible knowledge states from two structurally incompatible starting points. This "real" stabilization process drives all agents toward a more shared structural solution, ultimately producing a less diverse but more uniformly integrated final culture.

In the other configurations, with any degree of overlap ($\omega > 0$), both populations train on partially or fully shared object sets, producing pre-encounter decision functions that already agree on a larger portion of the environment. The encounter phase is therefore less disruptive: agents make only targeted adjustments to resolve the remaining disagreements, largely preserving the structural diversity accumulated during their independent stabilization trajectories.

Stated from a higher-level perspective: when two cultures disagree fundamentally on how to classify the same shared world, the reconciliation process is more transformative and ultimately more homogenizing than when they differ only marginally.

It should be noted, however, that this remains a tentative interpretation of a marginally significant observation, and further investigation would be needed to confirm it.

4.5 Refining the Analysis: The Distance Difference

4.5.1 Limitations of the Geometric Scheme

The preceding analyses have revealed two systematic limitations of the geometric classification scheme that constrain its ability to fully characterise cultural outcomes.

The first limitation concerns the **semantic dimension**. As detailed in Section 4.2, ontologies in this framework never shrink; they only adapt by splitting classes and continually growing. After 150,000 iterations, the post-encounter ontologies are substantially richer than the pre-encounter baselines $\mathcal{O}_1$ and $\mathcal{O}_2$. Consequently, $d_o(\mathcal{O}_1, \mathcal{O}')$ and $d_o(\mathcal{O}_2, \mathcal{O}')$ are both naturally large, systematically forcing classification into the Hybridization outcome regardless of the actual cultural dynamics. This renders the geometric scheme largely uninformative on the semantic dimension, except in strongly asymmetric configurations where the demographic effect is large enough to remain visible despite the noise.

The second limitation is the **rigid granularity** of the three classification bins. The discrete nature of the scheme leaves no room for intermediate states or degrees of assimilation. A run is classified as full assimilation even if the difference in attraction between the two cultures is minuscule, a nuance that becomes particularly evident in the H2 results, where the balanced condition ($\alpha = 0.50$) shows an apparent

Population 2 advantage (40.0% vs. 28.3%) that is, in reality, a product of stochastic drift around a symmetric equilibrium.

4.5.2 The Distance Difference Metric

To address both limitations, a complementary continuous metric is introduced: the **Distance Difference** (Δd). Given two pre-encounter cultures A and B and the post-encounter culture C, it is defined as:

$$\Delta d = \frac{d(A, C) - d(B, C)}{d(A, B)} \tag{4.1}$$

Dividing the raw distance difference by $d(A, B)$ normalises the cultural shift relative to the initial separation between the two populations, removing the scale bias that would otherwise make comparisons across experimental conditions unfair.

Intuitively, Δd can be understood as projecting the post-encounter culture onto the line connecting the two original populations, indicating how far the final state sits from the midpoint and in which direction. The metric is evaluated as follows:

- $\Delta d > 0$: assimilation toward Population 2 ($d(A, C) > d(B, C)$, i.e., the final culture is closer to Population 2).
- $\Delta d < 0$: assimilation toward Population 1.
- $\Delta d \to 0$: a balanced outcome with no clear directional pull, indicating hybridization.

The closer the value is to ± 1, the stronger the respective assimilation; the closer to 0, the more balanced and hybrid the outcome.

Following the definitions in Section 3.4.2, the metric is instantiated for the two dependent variables as:

$$\Delta d_d = \frac{dd(h_1, h') - dd(h_2, h')}{dd(h_1, h_2)} \tag{4.2}$$

$$\Delta d_o = \frac{d_o(\mathcal{O}_1, \mathcal{O}') - d_o(\mathcal{O}_2, \mathcal{O}')}{d_o(\mathcal{O}_1, \mathcal{O}_2)} \tag{4.3}$$

By replacing strict categorical boundaries with a continuous relative scale, Δd_d resolves the granularity limitation of the geometric scheme, and Δd_o mitigates the threshold issue caused by ontology growth, rendering the semantic dimension genuinely informative for the first time.

Only Hypotheses 2, 3.1 and 4.1 are revisited below, since these are the ones evaluated through the geometric classification scheme and thus affected by its limitations. Hypotheses 3.2 and 4.2 instead rely on the internal semantic diversity, so they require no re-analysis.

4.5.3 H2 Revisited: Population Asymmetry

The continuous distance difference metric confirms and strengthens the findings of the geometric scheme. Figure 4.5 presents the full empirical distributions of Δd_d (behavioral, top) and Δd_o (semantic, bottom) across the three population distribution conditions. A Kruskal-Wallis H-test was conducted to compare distributions across groups.

For the behavioral dimension (Δd_d), the test yielded highly significant results ($H = 138.13$, $p = 1.01 \times 10^{-30}$) with a substantial effect size ($\eta^2 = 0.39$). The mean assimilation strength toward Population 2 decreased monotonically with increasing minority size: 0.859 for $\alpha = 0.12$, 0.631 for $\alpha = 0.25$, and 0.045 for $\alpha = 0.50$.

For the semantic dimension (Δd_o), the test confirmed overwhelming statistical significance ($H = 203.96$, $p = 5.14 \times 10^{-45}$, $\eta^2 = 0.57$). The mean structural assimilation strength followed the same monotonic trend: 0.346 for $\alpha = 0.12$, 0.196 for $\alpha = 0.25$, and 0.021 for $\alpha = 0.50$.

The ridge plot distributions reveal the nature of this effect with a precision that summary statistics alone cannot communicate. At $\alpha = 0.12$, virtually all probability mass concentrates near $\Delta d_d = 1.0$, forming a sharp spike indicating near-complete and highly consistent assimilation by Population 2 across all runs. As size asymmetry narrows at $\alpha = 0.25$, this spike weakens and the distribution broadens leftward, showing that while majority assimilation remains the prevailing outcome, a meaningful fraction of runs produce weaker or intermediate results. This progressive flattening is precisely what the scheme's transition from 99.2% to 89.2% assimilation does not capture: the shift is not merely a drop in frequency, but a fundamental decrease in the *intensity* of the assimilation itself.

In the balanced configuration ($\alpha = 0.50$), the continuous metric accurately resolves the apparent Population 2 advantage reported by the geometric scheme. The distribution of Δd_d values centers symmetrically around zero (mean $\Delta d_d = 0.045$), confirming that the categorical asymmetry was a stochastic artifact.

Surprisingly, the Kruskal-Wallis effect size for the semantic dimension ($\eta^2 = 0.57$) exceeds that of the behavioral dimension ($\eta^2 = 0.39$), despite the smaller absolute values of Δd_o. This stronger semantic anchoring can be explained algorithmically: because ontologies only grow, every agent permanently retains the structural baseline of its original pre-encounter population. In highly asymmetric encounters, the vast majority of agents in the final merged population originate from Population 2, mathematically forcing the aggregate ontology distance to remain heavily anchored to the majority.

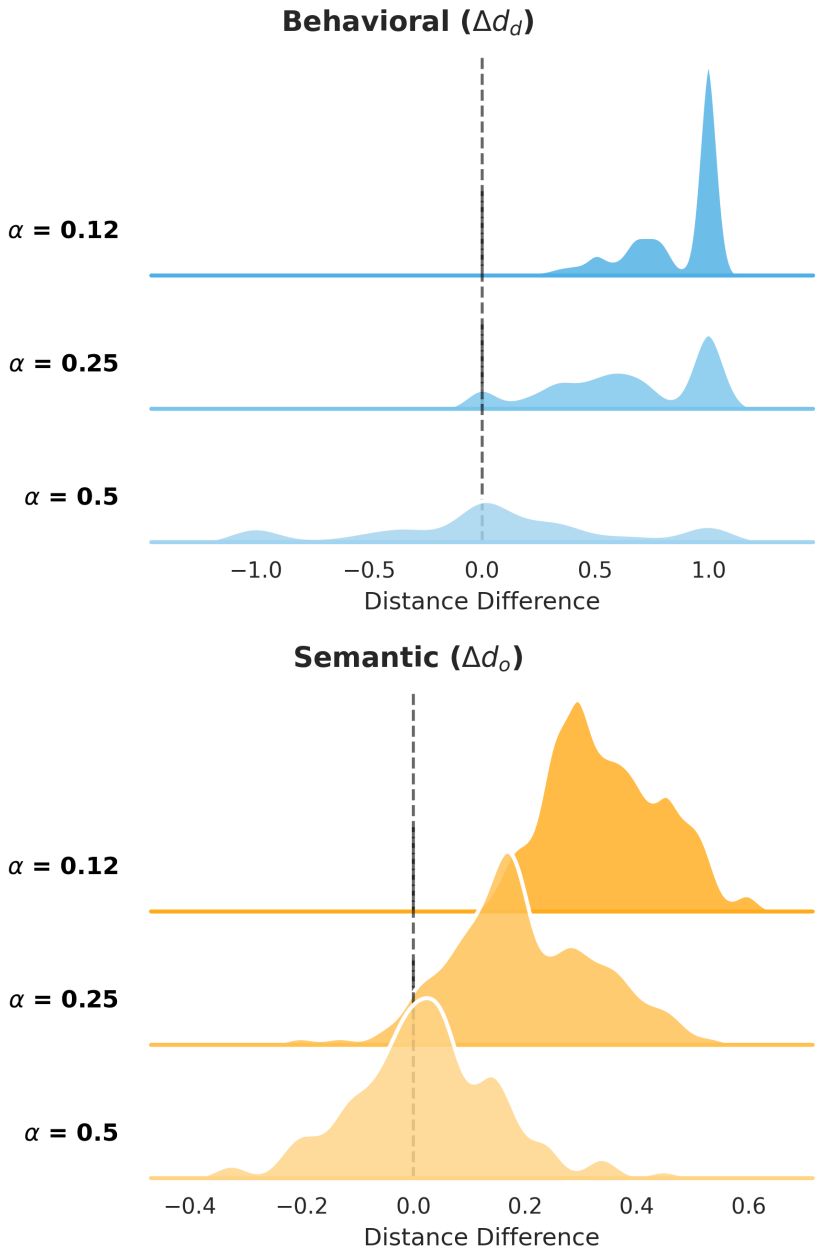


Figure 4.5: Distributions of the normalized Distance Difference Δd_d (behavioral, top) and Δd_o (semantic, below) across population distribution conditions. The dashed vertical line at 0 marks a perfectly balanced cultural outcome. Positive values indicate assimilation toward Population 2; negative values toward Population 1. Values approaching ± 1 represent near-complete cultural dominance.

4.5.4 H3.1 Revisited: Network Connectivity and Assimilation

The heatmaps in Figure 4.6 show the mean Δd_d (left) and Δd_o (right) across all combinations of *population distribution* (α) and *network connectivity* (η).

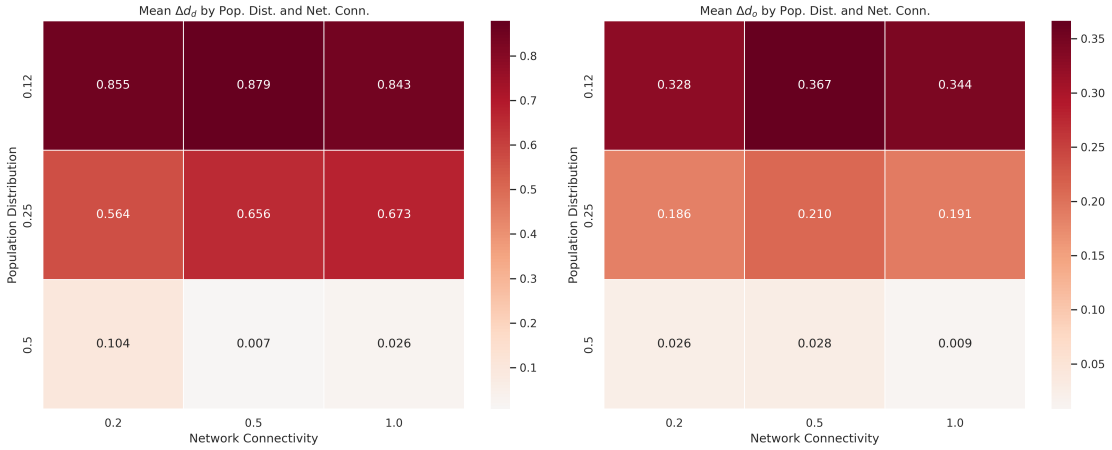


Figure 4.6: Heatmaps illustrating the mean Distance Difference for decisions (Δd_d, left) and ontologies (Δd_o, right) across all combinations of Population Distribution (α) and Network Connectivity (η).

The distinct horizontal bands visually confirm that the outcome is dominated by *population distribution*: the shift between rows is far larger than any variation within them, confirming that α is the principal driver of the cultural outcome. Some within-row variation is nonetheless visible at $\alpha = 0.25$, where the mean Δd_d rises monotonically from 0.564 at $\eta = 0.2$ to 0.673 at $\eta = 1.0$, that is, assimilation by the larger population increases as the network becomes more connected, exactly as predicted by H3.1.

As discussed in Section 4.3.2, $\alpha = 0.25$ is also the only asymmetric configuration in which a minority exists and the manipulation behaves as originally intended, and it is therefore the only setting in which such a trend would be expected. However, a Kruskal-Wallis test restricted to this configuration finds the trend to be non-significant ($H = 0.975$, $p = 0.614$).

One reason this apparent trend does not translate into a measurable effect can lie in the temporal nature of the metrics. Because these metrics are computed at the very end of the simulation, they represent the definitive, stabilized end-state of the newly formed population. While it is probable that a fragmented network slows down the rate of convergence, the primary goal of the system's inherent disagreement resolution mechanism is to force agents to converge on the same decisions (as demonstrated by Hypothesis 1). Given a sufficient iteration budget,

the minority will therefore inevitably be forced to adapt to the majority's decisions, and any temporary buffering effect provided by restricted connectivity is ultimately overridden by the end of the simulation.

Initially, it was expected that the semantic ontology metric would yield different results. Unlike behavioral decisions, which purely reflect the final external choice, ontologies retain the structural history of how agents arrived at those decisions, making Δd_o potentially more sensitive to the effect of restricted connectivity, even when external behaviors have already aligned. To understand what Δd_o actually captures in this system, it is important to recall that agents adapt exclusively by splitting classes, never by pruning them. This means that $d_o(\mathcal{O}_1, \mathcal{O}')$ does not measure what the post-encounter ontologies look like in absolute terms, but rather how much the agents of Population 1 were structurally forced to deviate from their pre-encounter baseline. A large $d_o(\mathcal{O}_1, \mathcal{O}')$ relative to $d_o(\mathcal{O}_2, \mathcal{O}')$ therefore indicates that Population 1's agents underwent extensive splitting to accommodate the other population's decisions, a structural signature of assimilation by Population 2. The metric thus remains a valid and meaningful indicator of the direction and intensity of adaptation pressure.

However, this same growth-only property explains why Δd_o mirrors the behavioral results rather than revealing any protective effect of restricted connectivity. Given sufficient iterations, the minority agents accumulate the same structural adaptations regardless of how far the realised cross-population contact was reduced, in those configurations where such a reduction was actually achieved. Under the adaptation mechanism of this system, and given the way inter-population contact is realised in its implementation, the final outcome shows no detectable dependence on network topology, and restricted connectivity did not preserve minority cultural traits in the long term.

4.5.5 H4.1 Revisited: Environment Overlap and Assimilation Direction

Although the difference between the groups is not statistically significant, a deeper inspection of the continuous metrics reveals a critical interaction between the specific dynamics emerging at the boundary condition of full overlap ($\omega = 1.00$) and the mathematical formulation of the normalized metric (Δd).

As visualized in Figure 4.7, the semantic dimension (Δd_o) remains anchored near zero with consistently narrow standard deviation bands across all conditions. The behavioral dimension (Δd_d), however, exhibits a dramatically expanding standard deviation band as overlap increases, culminating in a massive spread at $\omega = 1.00$. Concurrently, the discrete geometric scheme reports an apparent spike in behavioral hybridization (50.0%) at the same condition. While these two observations appear contradictory, wild swings in assimilation versus high hybridization, they are

actually symptomatic of the same underlying mathematical artifact: a near-zero initial baseline distance.

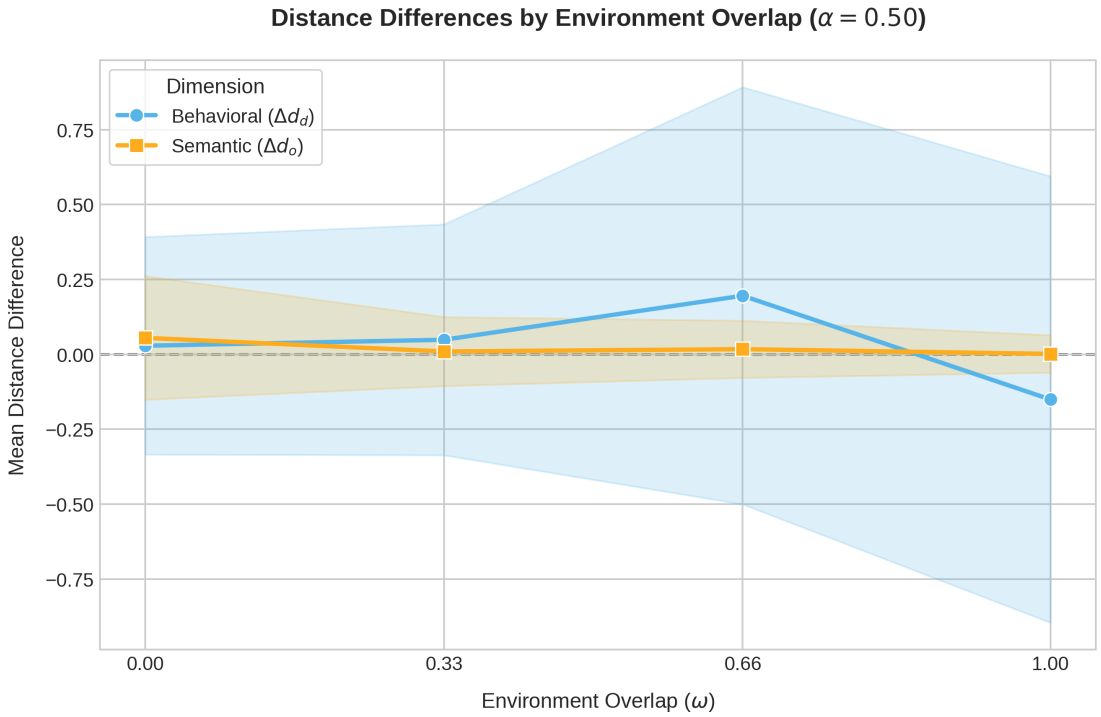


Figure 4.7: Line plot with standard deviation bands illustrating the mean continuous Distance Difference for both behavioral (Δd_d) and semantic (Δd_o) dimensions across varying Environment Overlap (ω) levels.

To verify this, the pre-encounter behavioral distance $d_d(h_1, h_2)$ was analyzed across overlap conditions. A Kruskal-Wallis test confirmed a highly significant monotonic decrease ($H = 79.81$, $p = 3.37 \times 10^{-17}$, $\eta^2 = 0.66$), with mean values of 0.481 ($\omega = 0.00$), 0.323 ($\omega = 0.33$), 0.163 ($\omega = 0.66$), and 0.104 ($\omega = 1.00$). As environmental overlap increases, the two populations train on increasingly similar object pools, and their pre-encounter behaviors converge toward near-identity. In some runs at $\omega = 1.00$, this baseline distance $d_d(h_1, h_2)$ reaches exactly zero, indicating that the two populations had already reached total behavioral consensus prior to the encounter, which renders the normalized metric undefined.

This vanishing denominator has two cascading effects. First, it causes the normalized Δd_d metric to mathematically amplify any minor post-encounter behavioral drift, producing the illusion of massive directional variance even when the absolute shifts are negligible. Second, it causes the geometric scheme's radius ρ to become microscopic, such that even the slightest post-encounter drift pushes

the final culture outside both assimilation boundaries, triggering the hybridization classification.

To confirm that this variance explosion is a metric artifact rather than a genuine dynamical phenomenon, the raw, non-normalized behavioral distance difference $\Delta d_{raw} = d_d(h_1, h') - d_d(h_2, h')$ was also evaluated, as shown in Figure 4.8. The Kruskal-Wallis test on the raw difference yielded no statistically significant difference across overlap levels ($p = 0.228$), with all group means hovering within 0.03 units of zero. Critically, the standard deviation of the raw difference *decreases* monotonically as overlap increases (0.169 at $\omega = 0.00$ down to 0.085 at $\omega = 1.00$), which is the exact inverse of the normalized metric's behavior. This directly confirms that the variance explosion observed in Δd_d is driven entirely by the shrinking denominator, not by any genuine increase in the variability of the encounter dynamics.

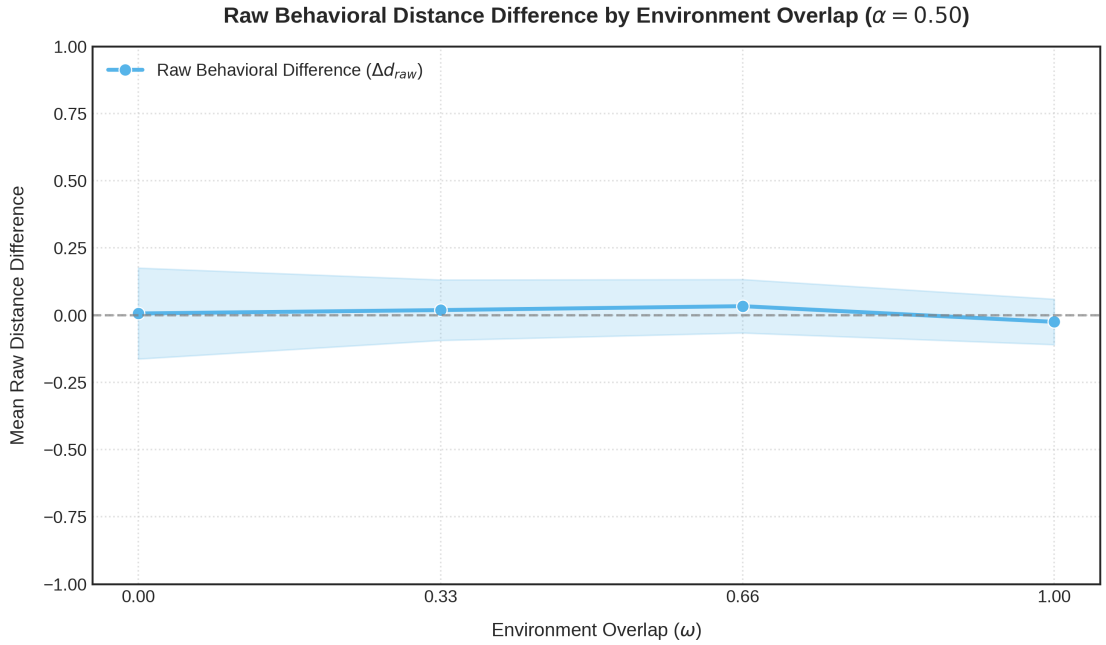


Figure 4.8: Line plot with standard deviation bands illustrating the mean continuous Behavioral Distance Difference not normalized (Δd_{raw}).

Beyond the mathematical mechanics, this boundary condition exposes a conceptual limitation of the macro-states defined at the outset of this study. When two populations are already practically identical prior to contact, the notion of cultural hybridization loses its meaning. The post-encounter state does not represent a novel mixture of distinct cultural traits, but rather the minor resolution of the last fractional disagreements between two already aligned groups.

This phenomenon does not affect the semantic dimension (Δd_o) in the same way.

Even when two populations share identical training environments and produce nearly identical behavioral decision functions, the structural paths by which they arrive at those decisions, dictated by the specific random sequence of interactions during independent stabilization, remain distinct. The *ontology distance* $d_o(\mathcal{O}_1, \mathcal{O}_2)$ does not approach zero as ω increases, preserving the metric's stability and explaining the consistently narrow standard deviation bands observed for Δd_o across all overlap conditions.

Outlook. Taken together, these results converge on a single conclusion: the outcome of an encounter between two established cultures in this system is governed almost entirely by their relative population size. The imbalance between the two populations (α) alone systematically shaped the result, determining both the direction and the strength of assimilation, whereas neither the topology of contact nor the extent of shared prior experience had any comparable effect.

Chapter 5

Conclusion

This study aimed to extend the current understanding of cultural evolution in multi-agent systems beyond the dynamics of a single population. The majority of the existing literature focuses on the internal dynamics of an isolated group, examining how its culture forms and evolves over time. This work broadens that scope by introducing the notion of multiple populations and, in particular, by studying what happens when they encounter one another.

As a starting point, we built upon a well-established framework that simulates a single set of agents, each endowed with its own formal ontology representing an individual view of the world, which we treat conceptually as a single *population* (Section 3.1). Within this framework, agents take decisions on objects and progressively learn from one another through interaction-driven adaptation. We extended this system by introducing the concept of distinct populations, an initial phase in which each population stabilizes independently, thereby developing its own distinct culture, and, as the central novelty of this work, a subsequent encounter phase in which the two populations merge and stabilize together.

To characterize the culture resulting from such an encounter, we defined two macro-states. *Assimilation*, in which the knowledge of one population is adopted by the other, and *hybridization*, in which the final culture cannot be clearly attributed to either original population but instead combines their traits into something new. Beyond defining these outcomes, a central part of the work is to assess whether factors that the literature associates with cultural dynamics within single populations would carry over to this new cross-cultural setting. Specifically, we tested different combinations of three variables: the relative size of the encountering populations (population distribution), the structure of their interactions (network connectivity), and the degree to which the two populations shared knowledge prior to the encounter (environment overlap), expressed as the overlap between their training sets.

Because agents represent their knowledge as formal ontologies, a population's

culture exists on two distinct levels: the behavioral level, i.e. the decisions its agents make, and the semantic level, i.e. the internal ontological structure that produces them. A single measure does not fully capture both, since two agents may agree on every decision while organizing their knowledge in structurally different ways. We therefore studied the outcomes through two complementary distance metrics: the *decision distance*, which captures the external behavior of a population, and the *ontology distance*, which captures the internal semantic structure of its knowledge representation. Building upon these two distances, we designed a way to use them to concretely assess the macro-states of *assimilation* and *hybridization*. To this end, we proposed a geometric schema able to classify each run into one of three categories: *assimilation* by Population 1, *assimilation* by Population 2, or *hybridization*. Running the analysis, however, revealed certain limitations of this schema, particularly for the semantic metric. Therefore, we introduced a complementary analytical methodology: the *distance difference*, normalized over the distance between the two pre-encounter populations. On the semantic dimension, moving from a categorical threshold to a continuous relative scale resolved the systematic bias toward hybridization caused by using the inter-population distance as the classification radius, recovering a measurable direction and intensity of assimilation that the strict classification bins could not capture.

5.1 Summary of Findings

Hypothesis 1 examined whether the central finding of the baseline study, namely that a single population converges to total behavioral consensus given sufficient time, would still hold in the new multi-population setting, under the same knowledge representation and disagreement-resolution mechanism. This hypothesis was **supported**: almost all runs reached total final consensus over the entire object space by the end of the encounter.

Hypothesis 2 concerned the *population distribution* (α), predicting that the pre-encounter knowledge of the larger population would behaviorally dominate the final merged culture. This hypothesis was also **supported**, with the strength of assimilation toward the majority increasing monotonically as the size asymmetry between the two populations grew. The continuous *distance difference* metric proved essential in establishing this trend, as it revealed gradations of assimilation strength that the categorical schema alone obscured.

Hypothesis 3 advanced two claims regarding *network connectivity* (η): first, that lower connectivity would increase hybridization in general, which in unbalanced configurations would specifically translate into reduced assimilation by the larger population (H3.1); and second, that it would yield a more internally diverse final culture (H3.2). **Neither was supported**. Under the most asymmetric

distribution ($\alpha = 0.12$), the restricted network did not reduce inter-population contact as intended, due to an implementation artifact of this study; in the remaining configurations, *network connectivity* had no significant effect on either the direction of assimilation or the final internal diversity of the merged population.

Hypothesis 4 likewise advanced two claims regarding *environment overlap* (ω): that higher overlap would favor hybridization (H4.1) and produce a less diverse final culture (H4.2). Both claims rested on the premise that greater overlap in the training sets would lead the two populations to stabilize more similarly, and thus to be more alike before the encounter. This underlying premise was *confirmed*: the pre-encounter distance between the populations, both behavioral and semantic, decreased monotonically as overlap increased. However, the subsequent inferential step proved unfounded, as higher overlap produced no significant shift toward hybridization, leaving Hypothesis 4.1 **not supported**. Hypothesis 4.2 was likewise **not supported**, though in an informative way: while higher overlap did make the two populations more similar before contact, the predicted decrease in final diversity did not follow. Instead, completely disjoint training sets ($\omega = 0$) were observed to produce the least diverse final culture, a threshold effect distinguishing only the disjoint condition from all overlapping ones. This observation also raises a conceptual concern about the meaning of hybridization when the two pre-encounter populations are nearly identical, examined in Section 4.5.5.

5.2 Limitations and Perspective

Two practical constraints merit acknowledgement. The fixed-iteration budget, while justified, left a small number of runs short of full stabilization, and the scale was modest: thirty agents in total, which becomes a very small minority population under the most unbalanced distributions. Both were concessions to the computational cost of the framework, and larger populations over longer runs would strengthen the results.

A second point to consider is the disagreement-resolution mechanism itself. The current system forces the two populations onto a single shared decision for each object, and therefore cannot represent a genuine state of coexistence in which distinct but equally valid behaviours persist side by side. While this is a deliberate design choice inherited from the baseline experiment [1], exploring resolution mechanisms that allow agents to retain their own perspective on an object when it is locally appropriate, rather than always converging on a single one, would be a natural and interesting extension in the specific context of cross-cultural encounters.

A further open question concerns whether the central finding of this work, the dominance of population size and the irrelevance of network structure and prior

similarity, is specific to the symbolic-ontology representation and disagreement-resolution mechanism used here, or holds more generally. A naming-game study in populations of LLM agents, although not framed as an encounter between distinct cultures, found an interesting related phenomenon: a small minority of agents that never adapted their behaviour was able, once a critical mass was reached, to impose its preferred convention on a majority initially settled on a different one [15]. The cultural-evolution literature has also repeatedly tied network structure and partial isolation to the preservation of cultural diversity [7, 11], and a related experiment simulating cultural evolution within a single population of large language models likewise finds that network structure shapes the diversity of the culture produced [12]. As discussed in Section 4.3.2, the network fragmentation tested here did not achieve the intended buffering of the minority under the most asymmetric population distribution ($\alpha = 0.12$, $\eta = 0.2$), so testing more tightly restricted connectivity levels, for instance by allowing inter-population interaction only between members of the two contact groups, would help establish whether these findings reflect general properties of cultural encounters or are specific to the implementation used here.

Knowledge accuracy Finally, a central finding of Bourahla et al. [1] is that this mechanism, most of the time, raises the population's accuracy above its initial level, even though, after the initial training, agents no longer receive the correct decision for individual objects. A closely related study using the same framework introduced successive generations of agents and showed that this parent-to-child mechanism supplies the variation needed to push a population's accuracy beyond the ceiling reached by within-generation interaction alone [6]. The open question is whether an encounter between two established populations can play an analogous role, though this is far from obvious: there, the variation came from new agents born with fresh, incomplete ontologies, whereas an encounter merges populations whose ontology trees have already grown and stabilized. What remains to be seen is whether the "shock" of merging two distinct cultures provides enough variation to raise the merged population's accuracy above what either achieved in isolation. This would also sharpen the interpretation of assimilation itself, distinguishing the case where a population assimilates toward genuinely better knowledge from the case where it assimilates merely toward the more numerous group.

Appendix A

Data and Software Availability

All experiments were performed in the Lazy Lavender software environment [16]. Settings, output and data analysis notebooks are made available at `https://sake.re/20260504-DOTP`.

Bibliography

[1] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. «Knowledge Improvement and Diversity under Interaction-Driven Adaptation of Learned Ontologies». In: *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021)*. 2021, pp. 242–250 (cit. on pp. 1, 2, 7, 9, 21, 22, 25, 56, 57).

[2] Alex Mesoudi. «Experimental Studies of Cultural Evolution». In: *The Oxford Handbook of Cultural Evolution*. Ed. by Jamshid J. Tehrani et al. Oxford University Press, 2023. DOI: 10.1093/oxfordhb/9780198869252.013.6 (cit. on pp. 1, 2, 4–6, 10).

[3] Robert Axelrod. «The dissemination of culture: A model with local convergence and global polarization». In: *Journal of conflict resolution* 41.2 (1997), pp. 203–226 (cit. on pp. 2, 7).

[4] Luc Steels. «Experiments in cultural language evolution». In: (2012) (cit. on pp. 2, 7).

[5] Andrea Baronchelli, Maddalena Felici, Vittorio Loreto, Emanuele Caglioti, and Luc Steels. «Sharp transition towards shared vocabularies in multi-agent systems». In: *Journal of Statistical Mechanics: Theory and Experiment* 2006.06 (2006), P06014–P06014 (cit. on pp. 2, 7).

[6] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. «Knowledge Transmission and Improvement Across Generations do Not Need Strong Selection». In: *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022)*. 2022, pp. 163–171 (cit. on pp. 2, 7, 57).

[7] Maxime Derex and Robert Boyd. «Partial connectivity increases cultural accumulation within groups». In: *Proceedings of the National Academy of Sciences* 113.11 (2016), pp. 2982–2987 (cit. on pp. 2, 4–6, 9, 10, 40, 57).

[8] Cassandra Gunasekaram et al. «Population connectivity shapes the distribution and complexity of chimpanzee cumulative culture». In: *Science* 386 (2024), pp. 920–925 (cit. on pp. 2, 4–6, 10).

[9] Marius Kempe and Alex Mesoudi. «An experimental demonstration of the effect of group size on cultural accumulation». In: *Evolution and Human Behavior* 35 (2014), pp. 285–290 (cit. on pp. 4, 5, 9, 10).

[10] Maxime Derex, Marie-Pauline Beugin, Bernard Godelle, and Michel Raymond. «Experimental evidence for the influence of group size on cultural complexity». In: *Nature* 503.7476 (2013), pp. 389–391 (cit. on pp. 5, 10).

[11] Mauricio Cantor, Michael Chimento, Simeon Q. Smeele, Peng He, Danai Papageorgiou, Lucy M. Aplin, and Damien R. Farine. «Social network architecture and the tempo of cumulative cultural evolution». In: *Proceedings of the Royal Society B* 288 (2021), p. 20203107 (cit. on pp. 5, 6, 10, 40, 57).

[12] Jérémy Perez, Corentin Léger, Marcela Ovando-Tellez, Chris Foulon, Joan Dussauld, Pierre-Yves Oudeyer, and Clément Moulin-Frier. «Cultural evolution in populations of Large Language Models». In: *arXiv preprint arXiv:2403.08882* (2024) (cit. on pp. 6, 10, 57).

[13] P. M. Krafft, Erez Shmueli, Thomas L. Griffiths, Joshua B. Tenenbaum, and Alex "Sandy" Pentland. «Bayesian collective learning emerges from heuristic social learning». In: *Cognition* 212 (2021), p. 104469 (cit. on pp. 6, 7).

[14] Alberto Acerbi and Domenico Parisi. «Cultural transmission between and within generations». In: (2006) (cit. on p. 7).

[15] Ariel Flint Ashery, Luca Maria Aiello, and Andrea Baronchelli. «Emergent social conventions and collective bias in LLM populations». In: *Science Advances* 11.20 (2025), eadu9368 (cit. on pp. 7, 57).

[16] Jérôme Euzenat. *Lazy Lavender*. `https://gitlab.inria.fr/moex/lazylav`. Git commit: `2bf30aa31d3ce78a4ff1f39b665588093fec056c`. 2026 (cit. on pp. 27, 58).