UNIVERSITY OF CALIFORNIA,
IRVINE

Computational Models Applied to Various Philosophical Topics

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Philosophy

by

Nathan Lawrence Gabriel

Dissertation Committee:
Jeff Barrett, Co-Chair
Cailin O'Connor, Co-Chair
Simon Huttegger
Brian Skyrms

2023

DEDICATION

To my parents who continue to be exemplary role models, teaching me the value of diversity and compassion. Thank you for nurturing my love of learning.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank first and foremost my advisors Jeff Barrett and Cailin O'Connor. Their help and guidance has been essential at each step in my progression through the Logic and Philosophy of Science department's doctoral program. No finite number of pages could understate my appreciation. Cailin has a shrewd eye for diversity, equity and inclusion that extends well beyond her rigorous academic research. Jeff has an open minded patience that allows the most inarticulate students to express their best thoughts. I will never forget his bemused demeanor when I suggested that there was a straightforward way in which a Lewis-Skyms signaling game could satisfy the philosophical constraints I put on a compositional semantics for putty-nosed alarm calls. Cailin's positive guidance has been unrelenting regardless of the context. I hope the content of this dissertation attests to Jeff and Cailin's expert guidance.

Chapter 1 has been accepted for publication, after peer review, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s10670-023-00735-x.

The Version of Record, the published version of Chapter 1, unintentionally omitted some important acknowledgments. Jeff Barrett is responsible for the initial suggestion that I model putty-nosed monkey alarm calls. He listened and provided comments when I first white boarded the model architecture. After I generated simulation results, Jeff met with me weekly to discuss each section of the original write up as it was written. These meetings were essential in forming the definitions of different types of compositionality given in Chapter 1. Saira Khan also provided comments on early drafts of Chapter 1. After the initial write up, the putty-nosed monkey model was presented in Brian Skyrms and Louis Narens' Social Dynamics seminar in which I received helpful feedback from more people than I can name. However, it should be highlighted that Brian Skyrms, Louis Narens and Simon Huttegger pursued an incredibly helpful line of questioning about the non-Nash equilibria that obtain in the game when punishment is included in the learning dynamics. This substantially improved my understanding of not only the putty-nosed monkey model, but also many other signaling game models that I subsequently researched. Commentary from Cailin substantially benefited the content of Chapter 1, helping me to focus it more on explaining compositionality itself rather than compositionality as it obtains in putty-nosed monkey alarm calls. Finally, 1.4.2 benefited substantially from two anonymous reviewers who helped clarify some philosophical issues related to when characteristics of both the senders and receivers rather than just the receivers are important to attributing content to compositions. Thank you all!

Both Jeff and Cailin separately ran regular lab meetings in which many of my fellow grad students positively contributed to my understanding of the topics researched in this dissertation. Cailin's Networking for Modelers meeting were co-organized by Jim Weatherall, who provided extensive guidance and constructive criticism in my researching epistemic networks. While I cannot name everyone who participated in each of my advisiors' lab meetings, there are some whose influence on my research is undeniable. Jingyi Wu, David Freeborn, and

Matthew Coats were constants in Cailin and Jim's Networking for Modelers meetings. At sometime or another each of them has answered extensive questions on their own social epistemology models. My understanding of NK-landscapes or Bayesian networks would be meager without their contributions. Additionally, their feedback contributed to significant improvements to Chapter 2 of this dissertation. Chapter 2 was also influenced by a prior model of illusory truth developed by Saira Khan, who participated in many of the early Networking for Modelers meetings. In Jeff's lab meetings, Jack VanDrunen and Christian Torsell enduring companions. They are the only grad students who have taken the time to understand the dynamics of the models in Chapter 3. In the process, they have helped me more clearly present the models. Jack and Christian have also accompanied me in an ambitious exploration of spiking neural networks as well as sharing their own research on signaling games. This has substantially benefited my understanding of the broader context of the models presented in both Chapters 1 and 3.

I have also received constructive comments on a variety of research projects from several faculty and grad students who have participated in Brian Skyrms and Louis Narens' Social Dynamics seminars. Though too many to name, I would like to thank every person who participated in these seminars. They were essential to my forming a broad understanding of signaling game models, epistemic network models, and many other formal approaches to philosophical problems. The Social Dynamics seminars were essential in general to having time and flexibility to pursue interesting research projects. Brian and Louis always took time to thoroughly understand and comment on student research. They inspire everyone to pursue a higher level of rigor in their research.

In the spring of 2022, I was Louis Narens' only student in a class covering his book on probabilistic lattices. Given this, he choose to tailor the class to my needs and I benefited immensely. Not only did I learn about lattices and associated logics, I also broadened my understanding of the relationship between formal models and psychology. Many weeks the class simply consisted in conversation about a chapter from the probabilistic lattices book with discussion of potentially relevant psychological experiments extending through his walk to the parking lot and another 20 minutes outside of his car. One day, I shared some of my early thoughts about the models that appear in Chapter 3. Once he heard that a prior model of transitive inference given by Siemann and Delius (1998) was based on Duncan Luce's research, we had an enlightening conversation about how the Siemann and Delius (1998) model may have misinterpreted Luce. (Louis thought that, while the Siemann and Delius (1998) model might be confounded by the circular variant of the transitive inference task, a careful interpretation of Luce should be compatible with pigeons apparent success in the circular variant.) From then on, it was common for Louis to not only give his own take on how a formal model related to a psychological experiment, but also explain what he thought Luce would say. This is just one example of Louis' dynamic teaching, which substantially improved my understanding of the intersection of formal models and psychology.

Chapter 2's research on illusory truth in epistemic networks was supported by NSF grant 1922424: Consensus, Democracy, and the Public Understanding of Science.

I am also grateful to all of the University of Houston's philosophers who, during my master's, helped me transition from a background in mathematics into philosophy. In particular, James Garson gave me a thorough understanding of his model theoretic inferentialism and intuitionist logic. Cameron Buckner is responsible for my first introduction to philosophy of animal cognition. His discussions of anthropofabulation, corvid theory of mind, and content in connectionist networks have all influenced my understanding of how simple and precisely operationalized psychological experiments can be philosophically important.

VITA

Nathan Lawrence Gabriel

EDUCATION

Doctor of Philosophy in Logic and Philosophy of Science **2023**
University of California Irvine *Irvine, California*

Master of Arts in Social Science **2023**
University of California Irvine *Irvine, California*

Master of Arts in Philosophy **2018**
University of Houston *Houston, Texas*

Bachelor of Arts in Mathematics **2013**
Rice University *Houston, Texas*

RESEARCH EXPERIENCE

Graduate Student Researcher **2021–2022**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2018–2023**
University of California, Irvine *Irvine, California*

Instructor **summer 2023**
Introduction to Inductive Logic
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

On the Stability of Racial Capitalism **2022**
Ergo

Reinforcement with Iterative Punishment **2022**
Journal of Experimental & Theoretical Artificial Intelligence

Pyow-Hack: Ordered Compositions in Lewis-Skyrms Signaling Games **2023**
Erkenntnis

ABSTRACT OF THE DISSERTATION

Computational Models Applied to Various Philosophical Topics

By

Nathan Lawrence Gabriel

Doctor of Philosophy in Philosophy

University of California, Irvine, 2023

Jeff Barrett, Co-Chair
Cailin O'Connor, Co-Chair

This dissertation investigates some philosophical issues using computational models. Chapter 1 presents a Lewis-Skyrms signaling game that can exhibit a type of compositionality novel to the signaling game literature. The structure of the signaling game is motivated by an analogy to the alarm calls of putty-nosed monkeys (*Cercopithecus nictitans*). Putty-nosed monkeys display a compositional system of alarm calls with a semantics that is sensitive to the ordering of terms. This sensitivity to the ordering of terms has not been previously modeled with a Lewis-Skyrms signaling game literature. Signaling games are valued for showing how communicative systems can arise with minimal learning tools. Simulation results show that basic (Roth-Erev) reinforcement learning is sufficient for the acquisition of a compositional signaling system sensitive to the ordering of terms.

Chapter 2 investigates social epistemology in the context of an effect of cognitive biases called the illusory truth effect. The illusory truth effect is exhibited when repeated exposure to a statement increases an individual's credence in that statement. While most investigations of the illusory truth effect focus on individuals' belief formation, humans typically form beliefs within a social structure. This is particularly relevant because various social structures can give rise to repeated exposure to statements; e.g. a popular book might recurringly be

discussed in one's social circle, or one or two foundational papers might always be cited by a particular lab. So, how does the illusory truth effect influence learning and belief formation in a group? This chapter uses network models to investigate this question. These models show that the illusory truth effect can be very detrimental to a group's belief formation. The effect causes networks to prematurely settle on a belief, in part, through repeated exposure to data points that are not independently generated. Previous research has indicated that the probability of such failures is near zero when networks are large or scientists are forced to explore unpopular and risky science. However, simulation results show that the harmful consequences of the illusory truth effect are robust even in large networks with mandatory exploration.

Finally, Chapter 3 shows how a rudimentary type of abstraction can obtain in Lewis-Skyrms signaling games. Here, abstraction is understood as occurring when different particulars take on the same functional role. Some abstraction may be guided by innate biases, and the chapter develops an analogy of reasoning about strategic thinking in chess to highlight some epistemic concerns that are raised by the presentation of abstraction. Concretely, the signaling game models of this chapter are developed in the context of two tasks that are much simpler than strategizing in chess. The first task is a transitive inference task that has been substantially studied in both humans and non-human animals. General features of the models developed for the transitive inference task are then carried over to a model for the second task of learning rudimentary grammatical structures. This second task is based on studies of human infants and non-human primates' ability to learn "nonsense grammars". Closing discussion highlights some strengths of the abstraction exhibited in the Lewis-Skyrms signaling game models.

Chapter 1

Pyow-Hack: Ordered Compositions in Lewis-Skyrms Signaling Games

1.1 Introduction

Compositionality is exhibited when a full statement in a language is comprised of component parts that contribute to its meaning. E.g. "the apple is red" and "the traffic light is red" are two distinct statements that have "red" as a component term. There are a number of ways of precisifying informal notions of compositionality. It is thought to be a key feature of human language, differentiating it from other animal communication (Hauser et al., 2002a; Scott-Phillips and Blythe, 2013). Lewis initially developed signaling games as a way of showing how communicative conventions can arise without prior knowledge of a language (Lewis, 1969). Later, Skyrms showed how Lewis signaling games can be understood as evolutionary games accessible to agents with low-rationality learning dynamics (Skyrms, 2008, 2010). Recently, Sterelny and Planar have advocated the use of Lewis-Skyrms signaling games for modeling the early evolution of human language (Planer and Sterelny, 2021). There are

some signaling game models that have been invoked in explanations of how and why humans acquired compositional language (Franke, 2015; Steinert-Threlkeld, 2016a). However, the signaling game literature has yet to provide concrete models that exhibit important baseline types of compositionality, e.g. what Sterelney and Planar call linear syntax. This chapter shows how such compositionality can obtain.

The pyow-hack signaling game model exhibits meaningful compositions sensitive to the ordering of terms. However, some of the game's structural features have not yet been explicated in the established literature. Consequently, the chapter begins with describing two simpler signaling games along with some basic terms and diagrams for describing signaling games. Section 2.1 presents a trivially compositional game, which was initially developed by Barrett (Barrett, 2007) to describe the evolution of kind language. Next, Section 2.2 outlines the progression of three properties of signaling games that will be combined to produce ordered compositions in the pyow-hack game: sender-compositionality, receiver-compositionality, and sender independent terms. It also shows that the model from Section 2.1 already meets the definition of sender-compositionality. Section 2.3 further explicates receiver-compositionality using a hierarchical signaling game model, which was originally developed in a pair of papers by Barrett, Cochran, and Skyrms (Barrett et al., 2019, 2018). Then, Section 3.1 outlines the behavior of putty-nosed monkeys, which have an alarm call system with linear syntax, and then 3.2 presents the pyow-hack game, in which a signaling system analogous to the monkey's can obtain. Lastly, Section 4 reviews the three properties of the pyow-hack game which allow for ordered compositions to obtain.

1.2 Ancient Artisans: Compositionality in Prior Signaling Game Models

This section precisifies three types of compositionality in order of increasing complexity. Simultaneously, it introduces diagrams of the signaling game models. These diagrams represent *both* the structure of a game *and* the reinforced dispositions of its players. Understanding how game structure relates to reinforcement of dispositions is essential to understanding compositionality in signaling games. This is not immediately apparent section 2.1's *trivially compositional* game. However, section 2.2, on *sender-compositionality* and its extensions, shows how one can erroneously consider distinct terms as identical when attending to only game structure. Finally, Section 2.3 explicates *receiver-compositionality* with direct reference to the reinforcement of dispositions.

This chapter illustrates the signaling game models with a fictional story of prehistoric artisans, Devasena, Valli, and Narundi. Devasena is a logistics expert. She knows whether there is a better supply of wood or clay; when Devasena wears blue (**b**_) there is a better supply of wood, and when she wears red (**r**_) there is a better supply of clay. Valli is a market strategist. She knows whether pots or figurines are in greater demand; when Valli wears blue (_**b**) pots are in demand, and when she wears red (_**r**) figurines are in demand. Narundi manages acquisitions. She observes what colors Devasena and Valli are wearing and brings either wood or clay along with either tools for making pots or tools for making figurines. Thus when **rb** (Devasena wears red and Valli blue), Narundi brings clay and tools for making pots. Since the two term statements in this signaling system are merely conjunctive (i.e. the intersection of two properties), it is called *trivially compositional* (Steinert-Threlkeld, 2020). There are no overbearing reasons for considering **rb** as a single statement with two component parts rather than understanding it as two independent statements, one made by Devasena and the other by Valli.

1.2.1 Learning Trivial Compositionality

The ancient artisans story is called a two-term two-sender one-receiver game in the established literature (Barrett, 2007; Barrett et al., 2019, 2018). In this game there are four states of nature (needing wood pot, wood figurine, clay pot, or clay figurine supplies), four correspondingly appropriate actions (bringing the corresponding supplies), and two terms (**b** and **r**). How could Devasena, Valli, and Narundi acquire their signaling system without any established communicative conventions, without knowing whether Devasena or Valli is more attune to market demand or material supplies? The prior literature has shown that such signaling systems can arise through basic Roth-Erev reinforcement learning.

The reinforcement procedure is as follows. Suppose that Devasena and Valli each have four urns associating one and only one with the four states of nature: supply and demand for wood pots, wood figurines, clay pots, and clay figurines. In each urn, they place one red stone and one blue stone. Narundi has four urns associating each with one of the four signals she can receive: **bb**, **br**, **rb**, and **rr**. Narundi places four tokens in each of her urns corresponding to the four actions she can perform: bring wood pot, wood figurine, clay pot, or clay figurine supplies. Each day Devasena and Valli observe the state of nature and then draw a stone at random with equal probability from their corresponding urns to determine what color to wear. Narundi then observes Devasena and Valli's colors and likewise draws randomly from her corresponding urn to determine what action to perform. If the action matches the state of nature, the day is a success. Consequently, each player returns what was drawn to the urn from which it was drawn along with an additional stone or token of the same type; thus, it is more likely that, when the same state of nature occurs in the future, the same signals and then action will occur. If Narundi's action does not match the state of nature, then the day is a failure. Consequently, stones and tokens are returned to the urns that they were drawn from leaving the probabilities of signals and actions unchanged.

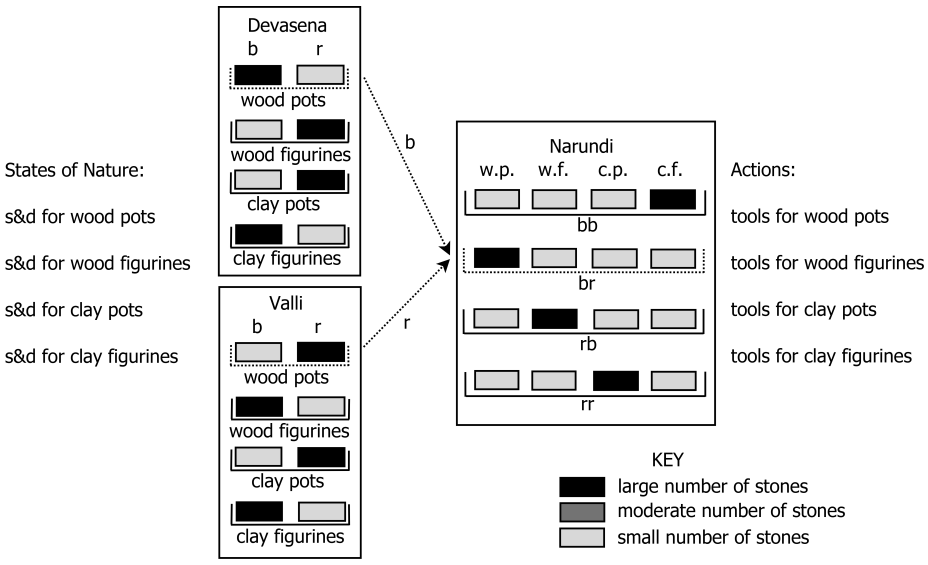


Figure 1.1: A Signaling System for the two-term two-sender one-receiver signaling game. To indicate the quantities of different types of stones, each urn is depicted with multiple boxes, one box for each type of stone/token in the urn. The most likely stones/tokens to be drawn from a urn are indicated the darker shaded in boxes.

The entirety of what occurs on a single day is called a single play of the game. A set of n repeated plays is called an n-run. The only way to know the outcome of a run is to actually perform all of the plays in a run. A strategy profile describes a sender or receiver's dispositions for all states of nature or signals that she can observe. Figure 1.1 shows both the structure of the game and a set of strategy profiles that could, in principle, be reached through the reinforcement procedure just described. When a set of strategy profiles describes dispositions such that for a state of nature the correct action is performed, it is called a signaling system. The set of strategy profiles in Figure 1.1 describe a signaling system in this sense. A run is successful if it converges to a signaling system. An easy and close approximate measure of a run's success is commonly performed by checking whether the run's cumulative success rate is above an appropriate cutoff. The cumulative success rate is defined as $\frac{\#\text{ of successful plays}}{\#\text{ of plays}}$. In the ancient artisans game there are four equally probable states of nature with unique corresponding actions. So if a strategy profile tends to lead to successful plays for all but one of the states of nature, we should expect a run with

that strategy profile to have a cumulative success rate near 0.75. Runs that tend to have successful plays for all states of nature have a cumulative success rate that converges towards 1. Consequently, in the ancient artisans game, it is reasonable to use a cutoff of 0.8 and measure only those runs that have a cumulative success rate greater than 0.8 as successful. Using this measure, after 10^6 plays per run under simple reinforcement learning, the run success rate in the ancient artisans game is approximately 73% (Barrett, 2007).

This game's dynamics are best described as involving two distinct senders, each with her own strategy profile. However, distinct senders in a signaling game model need not represent distinct organisms in the world. The reinforcement dynamics of drawing stones from an urn is intended to represent an organism's internal mechanisms for learning through reinforcement conditioning. Consequently, one might think of different senders in a game as representing different functional components of a single organism. In the ancient architects game, this might look like a single person using Devasena's draws to determine whether to wear a blue or a red top while using Valli's draws to determine whether to wear blue or red pants. Section 2.3 adds an executive send to the game that could perhaps be thought of as modeling a person's prefrontal cortex determining whether to attend to their Devasena dispositions, Valli dispositions, or both. Later, in modeling monkeys, different basic senders might represent different functional components responsible for the first or second term in an alarm call sequence.[1] Though it should also be noted that the game in section 3 succeeds in producing a type of compositionality novel to signaling games irrespective of whether or not it accurately reflects how the monkeys acquired their alarm calls.

[1]Certainly there is evidence that different areas of the human brain realize different functional roles in language production and comprehension (Dronkers et al., 2007; Naeser et al., 1987). But, it is also possible that neurons within a particular brain area could be arranged to realize different functional roles. Cao (Cao, 2012) has rightly shown that literature should be more careful about such claims. To that end, it might be noted that, supposing one were to model the players in the pyow-hack game with spiking neural networks and spike time dependent plasticity learning dynamics, the the contents of players urns would be most analogous to the strength of the synaptic connections between layers rather than the actual neurons themselves.

1.2.2 Sender-compositionality and its Extensions

The definition of trivial compositionality has mostly been used in critique. It does not define a type of compositionality that can easily be extended to yield a more sophisticated type of compositionality. Sender-compositionality is a type of compositionality that is useful for building up more sophisticated types of compositionality; it is also exhibited in the ancient artisans story. Call a set of strategy profiles in a signaling game *sender-compositional* if there is a term that is transmitted as a component of at least two distinct statements. E.g. Devasena wearing red (**r**_) is a component of statements **rb** and **rr**.

Franke (Franke, 2015) expresses dissatisfaction with the type of compositionality exhibited in the ancient artisans game. Some of this dissatisfaction comes from the fact that it can only generate statements that are composite with respect to the senders' dispositions. The receiver acquires its dispositions by reinforcing actions as if each composite statement is unitary; Narundi's reinforcement of an action picked from the **br** urn has no direct effect on the contents of the **bb** urn or the **rr** urn. Call set of strategy profiles *receiver-compositional* if there is a term that is a component of at least two distinct statements (sender-compositionality) *and* changing the receiver's dispositions for one of the statements directly results in changing the receiver's dispositions towards the other statement(s) containing the given term.[2] E.g.

[2]This definition is worthy of two clarifications. First, there is a claim, sometimes called the principle of compositionality, which asserts that in language the meaning of a complex expression is (fully) determined by its structure and the meaning of its component parts. Szabó (Szabó, 2020) as well as Pagin and Westerståhl (Pagin and Westerståhl, 2010a,b) have explicated this claim and refuted some objections to it, but their discussions are not exhaustive. Christiansen and Chater (Christiansen and Chater, 2016, 2022) primarily criticize accounts of compositionality emphasizing systematicity on the grounds that such accounts mischaracterize the phenomenon of language by making some sort of recursive cognitive capacity the essential capacity for acquiring language rather than being an emergent feature of other cognitive resources (though they do also give purported counterexamples). The definition of receiver-compositionality here is intended to be specific to Lewis-Skyrms signaling games and agnostic on the wider debate concerning compositionality. Any translation of the principle of compositionality into the language of signaling games would presumably entail something like receiver-compositionality, but the converse certainly does not hold. Receiver-compositionality says nothing about what other factors, in addition to the meaning of component parts, may be involved in determining the meaning of a complex expression.

Second, it must be acknowledged that there is reason to desire a further refinement of receiver-compositionality. Suppose a receiver is such that changing her dispositions towards assenting to "stop signs are red" directly causes her dispositions to move away from assenting to "roses are red". The strategy

the ancient artisans game is sender-compositional since **r**_ is a component of statements **rb** and **rr**; it is not receiver-compositional because the receiver's dispositions towards the statement **rb** are defined by the contents of the **rb** urn, her dispositions towards **rr** are defined by the contents of the **rr** urn, and when a play results in reinforcement of dispositions towards one of those statement (i.e. adding a stone to the urn that was drawn from) there is no change in the contents of the other urn. The model described in section 2.3 will exhibit receiver-compositionality by having a single urn that has an effect on the receiver's dispositions towards two statements that contain the same term (the statements **r**_ and **rb**); when a stone is added to the shared urn, it directly effects the receiver's dispositions towards both statements rather than only effecting dispositions towards the statement that was transmitted on the successful play.

Receiver-compositionality is intuitively desirable. Suppose someone, who already knows what "square" and "circle" mean, learns the appropriate response to hearing "bring the red circle". If she simultaneously fails to learn the appropriate response to hearing "bring the red square", then it seems doubtful that she has learned the meaning of "red" as a component part of the statement "bring the red circle". Conversely, if learning the appropriate response to "bring the red circle" causes a person to be disposed to act appropriately in response to "bring the red square", then this seems more reflective of "red" being treated as a component term. This is what receiver-compositionality allows.

It is worth attending to the reinforcment of dispositions when claiming that a term is a component of two distinct statements. If this is ignored, one might be tempted to claim that **r** is a component term of both **br** and **rb** in the ancient artisans story. However, there are strong reasons to reject this claim. One reason is that it is straightforwardly apparent

profile of such a receiver would satisfy the definition of receiver-compositionality, but not in a way that is desirable. Intuitively, the meaning of 'red' in one statement should be similar to its meaning in the other. Various schools of thought may readily have refinements at hand that resolve this issue. E.g. proponents of Smolensky et al. (Smolensky et al., 2013) will want to say something about the dispositions towards statements containing the same component term moving in the same direction of a state space. But again, this chapter remains agnostic on the issue of which refinement is best.

that the meaning of **r** in **br** is entirely different than its meaning in **rb**. However, it could be argued that there is some connection between the meaning of an **r** from Devasena and an **r** from Valli since they both disposed to transmitting **r** for clay pots. This is merely an accidental connection in the meaning of r from each sender. That this is merely an accidental connection can be seen in a second reason for rejecting the claim. Consider a signaling game identical to the ancient artisans game but with one alteration; Valli wears green and yellow instead of blue and red. This game is isomorphic to the original and in simulations will lead to the same 73% run success rate. It is difficult to defend the claim that statements **br** and **rb** share a component term when there exists an isomorphic signaling system containing statements **by** and **rg** which clearly do not share a component term.[3] So why not avoid confusion and describe Devasena and Valli as using different pairs of colors from the start?

In maintaining the superficial similarity between Devasena and Valli's terms, there is an immediately available extension of signaling game dynamics that creates a substantive identity between the terms used by distinct senders.[4] This extension is sender independence. An individual term X in a signaling game is *sender independent* if a receiver cannot condition her actions based on which sender transmitted X; this entails that the receiver typically performs the same action(s) irrespective of which sender transmitted the unitary X. In the pyow-hack game there will be two basic senders that can transmit a statement with a single P or H. Since the receiver will not be able to conditionalize on which basic sender transmitted the term, her dispositions towards a single P from basic sender A will be the same as her dispositions towards a single P from basic sender B; that is, the receiver will draw from (and on success reinforce) the same urns irrespective of which sender transmitted the single term. The power of sender independent terms will be more apparent after receiver-compositionality has been more thoroughly elaborated.

[3]See endnote 12 for a further discussion of why a naive information theoretic semantics would be mistaken in asserting that two distinct terms transmitted in the similar contexts have the same meaning.

[4]It is also the case that all prior descriptions of the model have maintained this superficial similarity between the terms of distinct senders (Barrett, 2007; Barrett et al., 2019, 2018).

1.2.3 Receiver-compositionality in a Hierarchical Game

Receiver-compositionality can be illustrated with an extension of the ancient artisans story. As trade networks expand, the three artisans are joined by industrialists Nekhbet and Hestia. Nekhbet sees a broader market context than Devasena and Valli; she determines whether only the material (clay or wood), only the form (pot or figurine), or both the material and form of the product are relevant. If only the material is relevant, Nekhbet only allows Devasena to observe the state of nature to determine what color to wear, while Valli wears an uninformative color; if only the from is relevant, then Nekhbet only allows Valli to observe nature to determine blue or red, and Devasena is uninformative; if both material and form are relevant she allows both Devasena and Valli to observe nature to determine what color to wear. When Narundi only sees a single term statement from Devasena and Valli (**b**_, **r**_, _**b**, or _**r**), then she draws at random with equal probability from either of the two corresponding urns; e.g. if Devasena wears red and Valli is uninformative, **r**_, then Narundi draws from either the **rb** or **rr** urn with equal probability. Hestia acquires supplies in bulk. She sees the signal from Devasena and Valli as well as Narundi's draw from the corresponding urn. If only Devasena wears an informative color, then Hestia only attends to the material of Narundi's draw and brings tools for both pots and figurines; e.g. if Narundi draws a wood pot token, Hestia brings wood and tools for both pots and figurines. If only Valli wears an informative color, then Hestia only attends to the form of Narundi's draw and brings the corresponding tools along with both wood and clay. If both Devasena and Valli are informative, then Hestia brings material and tools corresponding to Narundi's draw.

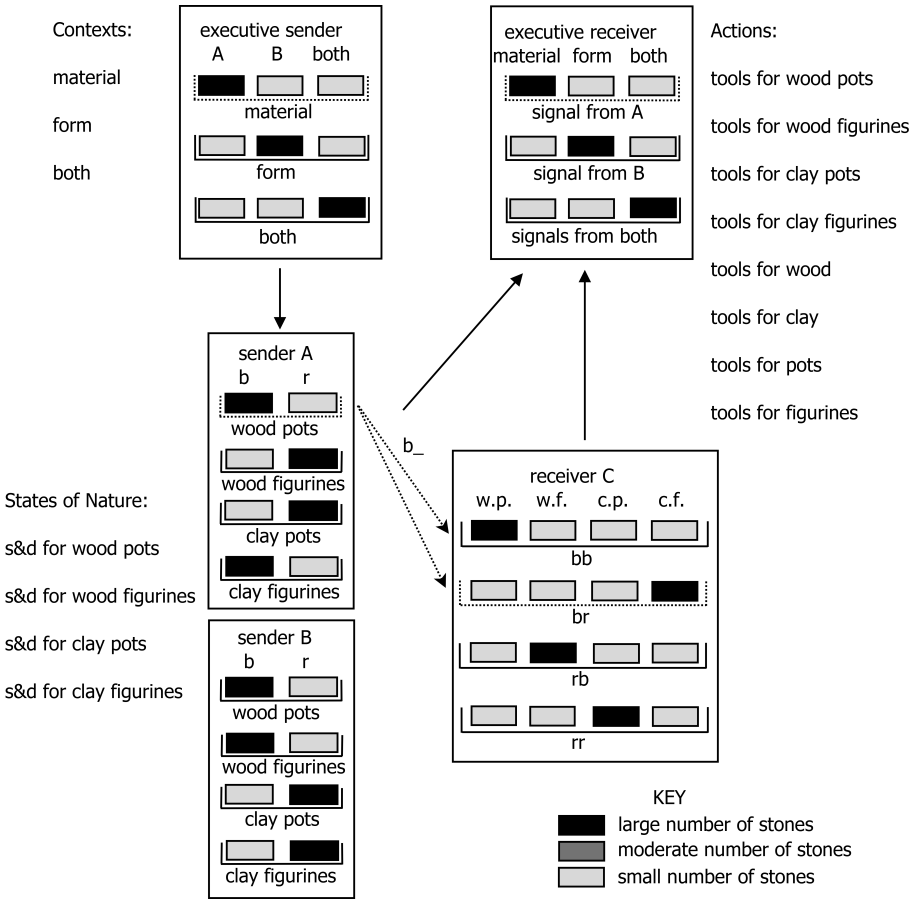


Figure 1.2: A Partial Pooling Equilibrium for the Hierarchical Signaling Game

Barrett, Cochran and Skyrms (Barrett et al., 2019, 2018) show how the extended ancient artisans signaling system can be acquired through reinforcement learning using a hierarchical extension of the two-term two-sender one-receiver signaling game that exhibits receiver-compositionality. The hierarchical signaling game has two basic senders (e.g. Devasena and Valli), one executive sender (e.g. Nekhbet), one basic receiver (e.g. Narundi) and one executive receiver (e.g. Hestia). A state of nature features two binary properties (e.g. material and form) and a context (e.g. only material, only form, or both are relevant). In the game the executive sender sees the context and the basic senders see the properties. Correspondingly, the executive sender has three urns and the basic senders have four urns each, as they did in the previous game. The executive sender determines which of basic senders' signals gets transmitted. She begins the game with three stones in each urn, a

sender A stone, sender B stone, and a both stone. The basic receiver sees what signal was transmitted and has four urns as she did in the previous game. The executive receiver sees both what was transmitted (a single term from sender A, a single term from sender B, or two terms), and correspondingly has three urns. The executive receiver sees whether sender A, B, or both transmitted a signal and determines whether the basic receiver's draw is interpreted as material, form, or both. Thus, the executive receiver begins the game with three urns, A, B, and Both, and each urn containing a material, form or both stone.

The set of strategy profiles depicted in Figure 2 is called a partial pooling equilibrium. A set of strategy profiles is a partial pooling equilibrium if the players perform better than chance, but worse than optimal, and their strategy profiles are an equilibrium in the sense that, for any given urn, changing which type of ball is most populous will not improve the success rate. The partial pooling equilibrium depicted in Figure 2 is worth highlighting because it is the first example of a signaling game exhibiting non-trivial compositionality; i.e. **br** indicates supply and demand for clay figurines, but this is not merely an intersection of what is indicated by **b**_ (which sometimes indicates wood) and _**r** (which sometimes indicates pots). In the hierarchical extension of the ancient artisans game, all optimal signaling systems are trivially compositional. It is not until the pyow-hack game that optimal signaling systems exhibit non-trivial compositionality.

Under basic Roth-Erev reinforcement learning, Barrett, Cochran and Skyrms' hierarchical signaling game model exhibits run success rates around 20%.[5] This run success rate can be increased to around 97% when the reinforcement dynamics are supplemented with punishment via costly signals (Barrett et al., 2019, 2018). Since the focus of this chapter is on how a particular type of compositionality can obtain, the details of stronger learning dynamics are omitted. However, it will be noted that the same type of reinforcement supplemented with punishment via costly signals produces similar gains in the pyow-hack signaling game

[5]Based on 1000 simulations of 10^8 plays per run and using a cutoff of 0.98 cumulative success rate for counting a run as a success.

model.

1.3 Sender Independent Terms in the Pyow-Hack Game

This section begins with a very brief description of putty-nosed monkey alarm calls, which were the inspiration for the pyow-hack game. This game is then described in Section 3.2. The monkeys provide a tangible motivation for what is otherwise a very abstract game. Their behavior also helps motivate some intuitions (discussed in Section 4.2) about how two terms can be composed together to generate different meanings based on their ordering. In turn, the pyow-hack game shows how a compositional signaling system with sensitivity to term ordering could be acquired by an organism that only has access to low rationality reinforcement learning dynamics. That said, this chapter is not concerned with defending any particular interpretation of putty-nosed monkey alarm calls. The Lewis-Skyrms pyow-hack signaling game presented in Section 3.2 is shown (in Section 4.2) to exhibit a type of compositionality novel to the signaling game literature irrespective of the extent to which it accurately describes putty-nosed monkey behavior.

1.3.1 A Brief Description of Putty-nosed Monkey Behavior

Cercopithecus nictitans martini, putty-nosed monkeys, are a West African species. They typically live in groups of 13-22 individuals comprised of one adult male with several females and dependent juveniles (Arnold and Zuberbühler, 2006). Their common predators are crowned eagles and leopards. Group leaders give different alarm calls that correlate fairly robustly with the presence of leopards and eagles. They also have a call associated with group movement (Arnold and Zuberbühler, 2006, 2008, 2013; Schlenker et al., 2016a,b).

Putty-nosed monkey alarm calls are comprised of two basic calls: a hack (H) and a pyow

(P). These basic calls are strung together in sequences of varying length. A sequence of repeated hacks, perhaps HHHHHHHH, is associated with aerial predators (eagles) and invokes the behavior of looking up. A sequence of repeated pyows, perhaps PPPPP, is associated with ground predators (leopards). Behaviorally, the pyow call sequences are associated with moving towards the caller since ground predators rely on stealth and the monkeys can collectively scare off the predator (Arnold and Zuberbühler, 2013). Sequences of pyows followed by hacks, perhaps PPPPHH, are associated with group movement. Sequences of hacks followed by pyows, perhaps HHHHPPP, occur when a nearby eagle moves away from the group. Longer call sequences seem to correlate with more urgent contexts when signaling for predators and increased distance traveled when signaling group movement; behaviorally, this correlates respectively with faster reaction times and potentially moving longer distances (Arnold and Zuberbühler, 2012, 2013; Schlenker et al., 2016a).

Schlenker et al. (Schlenker et al., 2016a) give a detailed overview of putty-nosed monkey alarm calls, and reasons for interpreting the calls as semantically compositional. Additionally, they propose some possible referential or imperative semantics for the alarm calls.[6] In developing a compositional semantics, Schlenker et al. make a particularly insightful obser-

[6]Whether the most appropriate semantics for the alarm calls is referential, with PPPP meaning there is a leopard nearby, or imperative, with PPPP being a command to move towards the caller, is just one of many dimensions of alarm call analysis that Schlenker et al. (Schlenker et al., 2016a) leave open ended. While they give substantive reasons for preferring a compositional analysis of alarm calls, Schlenker et al. also present a plausible non-compositional semantics for the alarm calls. Steinert-Threlkeld (Steinert-Threlkeld, 2016b) has also advocated understanding the alarm calls as non-compositional; though, one of Steinert-Threlkeld's objections, that a compositional analysis of the alarm calls presupposes the monkeys having a robust theory of mind, certainly does not apply to the model in this chapter which makes no such presupposition and is more or less compatible with Schlenker et al.'s semantics depending on how their reliance on pragmatic rules is precissified. Perhaps Schlenker et al.'s most intriguing open ended claim is that a compositional analysis of the alarm calls yields a semantics that strongly resembles Stalnaker's semantics for conditionals. Future debate on the semantics of putty-nosed monkey alarm calls should take seriously the possibility of uniting an imperative semantics, which provides a straightforward connection between group movement and response to a nearby leopard, with already available models of how humans process conditionals, such as those cataloged by Oaksford and Chater (Oaksford and Chater, 2012).

While noteworthy, the nuances of attributing a semantics to putty-nosed monkey alarm calls are beyond the scope of this chapter. The alarm calls, when interpreted as compositional, inspired the model that is presented and provide a tangible context for an abstract game. But should a divine source appear tomorrow an inform us definitively that the monkey's alarm calls are not compositional. It remains true that this chapter explains a type of compositionality that is novel to the signaling game literature and shows how it can obtain in an optimal signaling system.

vation about the relation between calls associated with ground predators and calls associated with group movement. Though over a shorter distance, the monkeys move towards the caller when the call for a ground predator is issued (so they can collectively mob the predator). This provides some reason for interpreting a "pyow" as contributing similar meaning to the ground predator call as a "pyow" contributes to a group movement call.

The pyow-hack signaling game will simplify things by only allowing six different statements in the game: P, PP, PH, H, HH, and HP. On this simplification, putty-nosed monkey behavior translates to the following call system. When a leopard is nearby, the group leader issues a P call, to which group members are disposed to move towards the group leader. When a leopard is very nearby, the group leader issues a PP call, to which group members are disposed to quickly move towards the group leader. When moving, the group leader issues a PH call, to which group members are disposed to move an extended distance towards the group leader. When an eagle is nearby, the group leader issues a H call, to which group members are disposed to look up. When an eagle is very nearby, the group leader issues a HH call, to which group members are disposed to quickly look up. When a nearby eagle is leaving, the group leader issues a HP call, to which group members are disposed to look up and then elsewhere.

1.3.2 The Pyow-Hack Game

The pyow-hack signaling game abstracts and simplifies away from several of the details of putty nose monkeys' environment and behavior. Most noticeably, it only allows for call sequences of at most two signals. Like the Barrett, Cochran and Skyrms's model (Barrett et al., 2019, 2018), it is a hierarchical signaling game consisting of an executive sender, two basic senders, an executive receiver, and a basic receiver.

In the pyow-hack game there are six states of nature: a leopard is nearby, a leopard is very

near (urgent), an eagle is nearby, an eagle is very near (urgent), a nearby eagle is moving away, and the group is moving. There are six corresponding appropriate actions: move towards caller, quickly move towards caller, look up, quickly look up, look up and elsewhere, and move an extended distance towards caller.

The executive sender as well as the basic senders can observe the state of nature. The executive sender determines whether just one or both of the basic senders will transmit a signal. This corresponds with the executive sender having six urns, one for each state of nature. These urns contain two types of balls, single transmission balls and dual transmission balls. As in the previous games, all of the players' urns start with one ball of each type. The basic senders each have six urns corresponding to the states of nature. The basic senders have two types of balls, P balls and H balls. On plays in which the executive sender draws a single transmission ball, it is determined at random with equal probability whether sender A or sender B transmits a signal.[7]

The basic receiver has four urns: PP, PH, HP, and HH. When a single P is transmitted, it is determined at random with equal probability whether the basic receiver draws from the PP or PH urn.[8] When a single H is transmitted, it is determined at random with equal probability whether the basic receiver draws from the HP or HH urn. As in the previous hierarchical game, the basic receiver draws balls that can be given multiple interpretations by the executive. The basic receiver's urns contain four types of balls labeled: (i) 'quickly move towards caller', (ii) 'move an extend distance towards caller', (iii) 'quickly look up', and (iv) 'look up and elsewhere'. These labels are the complex interpretations that the

[7]This is part of how the pyow-hack game avoids the oddity discussed at the end of section 2.2. In this game, executives are only sensitive to signal length. As will be seen shortly, the executive receiver cannot form dispositions relative to which basic sender transmitted a signal. She can only form dispositions relative to signal length.

[8]One might wonder why the receiver is not equally likely to draw from the HP urn when receiving a single P. The intuition is that receiver C upon hearing P attends to the PP and PH urns, while ignoring the HH and HP urns, because the P is the first P in the transmission. Then, once a pause is long enough to determine that no second term is following the P in the transmission, C draws at random with equal probability from the PP and PH urns.

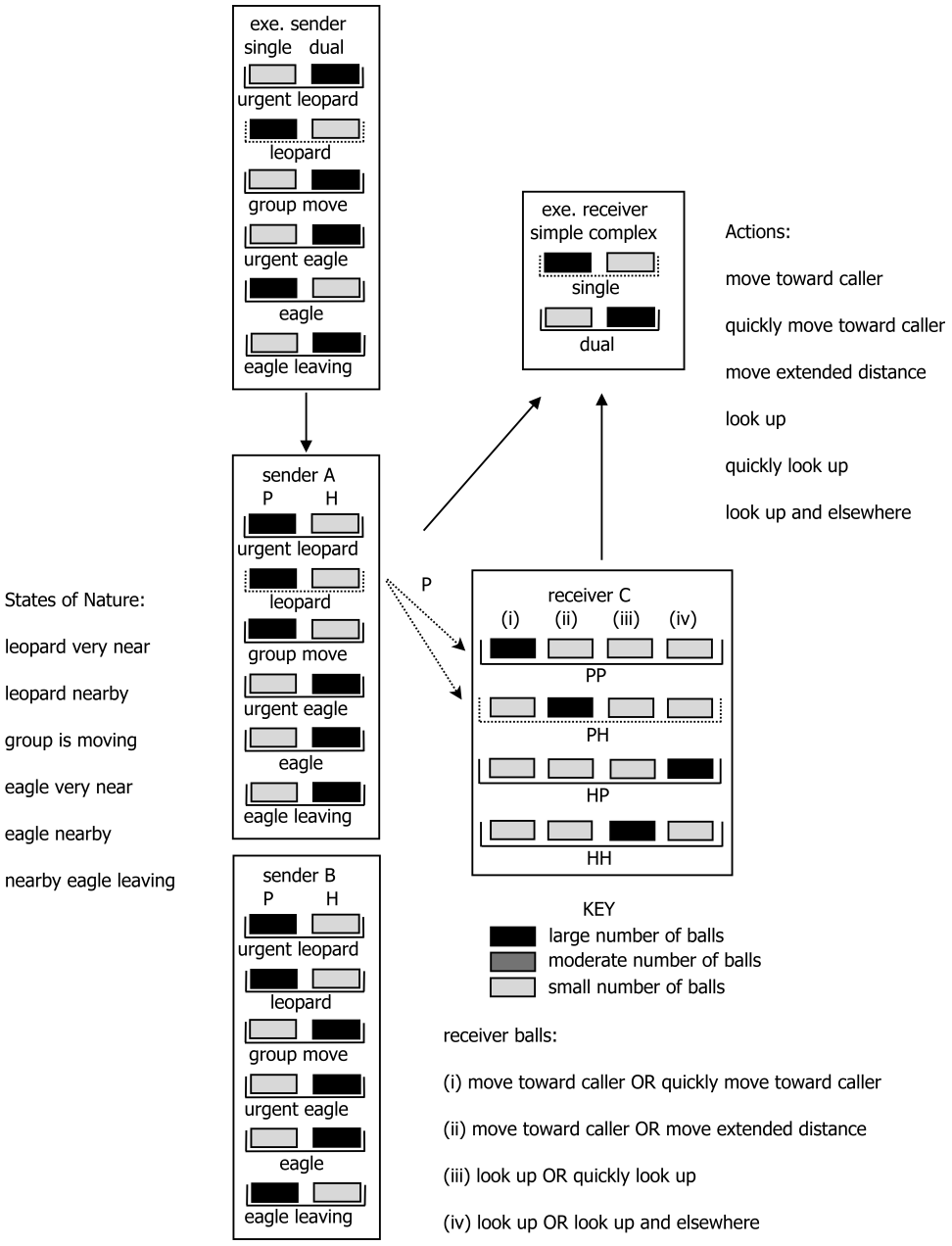


Figure 1.3: A Signaling System for the Pyow-Hack Game

executive can give to the balls. Type (i) and (ii) balls can be given the simple interpretation "move towards the caller". Type (iii) and (iv) balls can be given the simple interpretation "look up". Thus, the executive receiver has two urns, a single transmission urn and a dual transmission urn. Each of these urns has two types of balls, simple interpretation balls and complex interpretation balls.

Simple reinforcement learning (Roth-Erev) is the dynamic that was presented in Section 2.1. When a play is successful, drawn balls are returned to their urns and one additional ball of the type drawn is added to the urn that was drawn from for each player. On failures, balls are returned to the urns they were drawn from. Here's an example play for the equilibrium depicted in Figure 3:

1. Nature chooses a state at random with equal probability. Suppose the state of a nearby leopard is chosen.
2. The executive sender observes the state and chooses a ball at random with equal probability from her nearby leopard urn. Suppose the executive sender chooses a single transmission ball; this is the executive sender's most likely choice in the example equilibrium.
3. Since the single transmission ball was drawn, either sender A or sender B is chosen at random with equal probability to transmit a signal. Suppose sender A is chosen.
4. Sender A observes the state of a nearby leopard. So, she draws at random with equal probability from her nearby leopard urn. Suppose she draws a P ball. Again, this is the most likely choice in the example equilibrium.
5. Given her draw, sender A transmits 'P'.
6. Receiver C sees the 'P' and it is determined at random with equal probability whether she will draw from the PP urn or PH urn. Suppose it is determined that receiver C will draw from the PH urn.

7. Receiver C draws at random with equal probability from the PH urn. Suppose receiver C draws a (ii) ball. This is the ball that she is most likely to draw in the example equilibrium.
8. Receiver C's draw is now interpreted by the executive receiver.
9. The executive receiver sees that only a single signal was transmitted and draws from her single urn. Suppose the executive receiver draws a simple ball. In the example equilibrium, this is her most likely draw.
10. Given that the executive receiver drew a simple ball, she interprets receiver C's draw, (ii), as needing to move towards the caller. So this action is performed.
11. Since this is the correct action for the given state, this counts as a success.
12. Given the success, each player returns the ball that she drew along with an additional ball of the type that was drawn.
13. When a failure occurs, drawn balls are returned to the urns that they were drawn from.

This concludes a play of the game.

Simulating 1000 runs of the pyow-hack game, with 10^7 plays per run, produced the run success rate of 19.4%. This was calculated by measuring each run's cumulative success rate, the number of successful plays divided by the total number of plays. This was calculated by counting a run as successful if it had a cumulative success rate above 0.92. This was an appropriate cutoff for determining whether a run was successful as $0.92 > 5.5/6$. That is, a cumulative success rate greater than 0.92 is indicative of plays being successful for each of the six states of nature.

Under basic Roth-Erev reinforcement learning, the pyow-hack game has a run success rate of around 19%. This increases to around 58% when using costly signals analogous to Barrett, Cochran and Skyrms' reinforcement with punishment via costly signals (Barrett et al., 2019,

2018). An even stronger learning dynamics, described by Barrett and Gabriel(Barrett and Gabriel, 2022), can give a run success rate of 94.3% (of 1000 runs, 10^7 plays per run, and 10 iterations of [+2, -9] reinforcement with iterated punishment). However, even on the weakest learning dynamics, it remains the case that when an optimal signaling system obtains, the compositionality is novel in that it is sensitive to the ordering of terms.

1.4 Discussion

1.4.1 Review of Technical Terms

Two types of compositionality are emphasized in this chapter: (i) set of strategy profiles is *sender-compositional* if there is a term that is transmitted as a component of at least two distinct statements; (ii) a set of strategy profiles is *receiver-compositional* if there is a term that is a component of at least two distinct statements and changing the receiver's dispositions for one of the statements directly results in changing the receiver's dispositions towards the other statement(s) containing the given term. Receiver-compositionality allows a term to contribute similar dispositions to multiple statements that have the given term as a component part. Additionally, motivating some of the differences between the pyow-hack hierarchical game and the Barrett, Cochran, and Skyrms' (Barrett et al., 2019, 2018) hierarchical game, a term is *sender independent* if, upon transmission of the unitary term, the receiver cannot condition her actions on which sender transmitted the term; this entails that, for a given set of strategy profiles, transmission of the unitary term typically results in the same action regardless of which sender transmitted it. The signaling game literature discusses a third type of compositionality defined by Schlenker et al. (Schlenker et al., 2016b) and introduced to the signaling game literature by Steinert-Threlkeld (Steinert-Threlkeld,

2020)[9]: (iii) a set of strategy profiles is *trivially compositional* just in case complex expressions are always interpreted by intersection (generalized conjunction) of the meanings of the parts of the expression. It can be checked that optimal signaling systems for the Barrett, Cochran, and Skyrms hierarchical game are trivially compositional.[10]

1.4.2 Order Sensitive Compositionality

Signaling systems for the pyow-hack game are not trivially compositional. If they were, since the terms are sender independent, HP would be associated with the same dispositions, as PH. But this cannot occur in a signaling system since the game only allows for six possible statements and requires distinct actions to be performed for each of the six states of nature. For a given set of strategy profiles, if transmitting HP and transmitting PH typically results in the same action being performed in response to either statement, then at most only five of the six states of nature can be mapped to the correct action by the senders' and receiver's strategy profiles. This is a quick method of demonstrating that compositionality exhibited in the pyow-hack game is different from the compositionality exhibited in the Barrett, Cochran, and Skyrms hierarchical game. However, it does not show how compositionality with sensitivity to term ordering obtains in the pyow-hack game.[11]

[9]Although, it is not obvious that this is an appropriate definition for the sort of compositionality that Steinert-Threlkeld is concerned with. Steinert-Threlkeld constructs a game in which an artificial neural network is supposed to learn the function words "most" and "least" across different dimensions of properties that these function words can be applied to. In his analysis, Steinert-Threlkeld does not show (though it may be true) that the network's learning the correct output for "most blue" contributes to its learning the correct output for "most green". But, this is the case for the human language that Steinert-Threlkeld is attempting to model. When a human learns to use "most" appropriately for some small domain of properties, she then is able to use the term appropriately for novel properties. This is because the content of the term is not specific to particular properties. Furthermore, there is no obvious impediment to Steinert-Threlkeld showing that his model exhibits the desired behavior. To do this, Steinert-Threlkeld should show that first training the network on one or two dimensions of properties allows it to learn "most" and "least" for a second or third dimension at a faster rate than it would with no pre-training. This is exactly the sort of generalization that artificial neural networks are valued for.

[10]Though it should be noted that the hierarchical game can exhibit non-trivial compositionality in sub-optimal partial pooling equilibria (Barrett et al., 2018).

[11]This sensitivity to term order further highlights the differences between the pyow-hack game and the Barrett, Chochran, and Skyrms hierarchical game, which can exhibit non-trivial compositionality in suboptimal

Sensitivity to term ordering is allowed by the combination of both sender independent terms and receiver-compositionality. To see how this sensitivity is allowed, consider the signaling system diagrammed in Figure 1.3. It is easy to see that PH and HP are associated with different dispositions, actions. PH is typically transmitted when the state of nature is group movement and typically results in the action of moving an extended distance towards the caller. HP corresponds with the nearby eagle leaving state of nature and the look up & elsewhere action. However, this does not necessarily entail that the compositionality is sensitive to term ordering because of the worry described at the end of Section 2.2. Recall that this worry raises the concern that the P in PH is not the same term as the P in HP, perhaps one is an A-tone P and the other is a B-tone P making them functionally distinct terms. To establish that the P in the PH statement is the same term as the P in the HP statement, it must be shown that there is a connection between the dispositions associated with the P term in PH statements and the P term in HP statements. The remainder of this section shows that there is such a connection, but also highlights why it is desirable for future models to strengthen the connection.

The meaning of a term in a Lewis-Skyrms signaling game model is best understood as being determined by the dispositions associated with that term.[12] A P alarm call means "leopard", "move towards caller", or has some meaning in between the two (à la Milikan (Millikan, 1984,

pooling equilibria, but cannot exhibit sensitivity to term order.

[12]There is nuance to this claim and it substantively differs from a minimalist reading of the information theoretic semantics given by Skyrms (Skyrms, 2010) (though such a reading should question whether it is consistent with Skyrms and Barrett's later paper (Skyrms and Barrett, 2019)). If one does not include certain counterfactual information in a player's dispositions, then it is better to say that meaning is strongly correlated with dispositions but not entirely determined by them. Consider two signals A and B understood with a naive information theory. If the signal A moves probabilities in the same direction, and with the same magnitude, as signal B, then one might say they have the same informational content. However, if, in being composed with a third signal, C, AC moves probabilities in a different direction than BC, then intuitively one is not inclined to assert that A has the same meaning as B. So it seems a mistake to associate A's content with merely how it moves probabilities. Suppose further that AC and BC move probabilities in the same direction, but also that it is possible for future learning to cause AC and BC to move probabilities in divergent directions. Should we say that their present equivalence in informational content entails that they have the same meaning? Probably not. This is why sender independence is defined as the receiver not being able to condition her action on which sender transmitted the unitary term, which entails that her dispositions are the same irrespective of which sender the term came from.

1995)), both because the call is issued when a leopard is present and because the hearers respond as if a leopard is present when they hear the call. Suppose a group leader issues P calls when leopards are present and H calls when eagles are present; but, the monkeys hearing the calls look up at the sky when hearing P calls and move towards the leader in preparation for a leopard threat when hearing H calls. In this case, there is no signaling system and no communication. Neither the senders nor the receivers can unilaterally determine the meaning of a signal.

For the signaling system depicted in Figure 3, a dispositional connection between the P's in PH and HP can be understood as follows. A solitary P is correlated with the disposition to move towards the caller so the monkeys can collectively mob the leopard, and presumably this involves the monkeys looking in the direction they are moving or at the ground for a leopard, but not looking up at the sky. A PH call is correlated with group movement for an extended distance in the direction of the caller. So the dispositions associated with P and PH are similar in that they both involve movement towards the caller. An HP call is correlated with looking up and elsewhere. A solitary H is correlated with the disposition to look up, so it makes sense to take the P in an HP call with the disposition to look places other than the sky. So the P in an HP call also shares an overlap in associated dispositions with a solitary P. Schlenker et al. (Schlenker et al., 2016a) assert that it is plausible to think of putty-nosed monkey alarm calls as being analogous to semantics for conditionals. So, by loose analogy, one might think that P is associated with a set of possible worlds for which it is appropriate to move towards the caller or look towards the ground; H is associated with possible worlds for which it is appropriate to look up at the sky; and, if the nearest possible world to the center of the P set of worlds for which H is true is a different world than the nearest possible world to the center of the H set for which P is true, then this is why PH and HP have different meanings despite their component terms having the same meaning. This talk of possible worlds merely is intended to aid the reader's intuitions and it is not being claimed that some possible worlds semantics obtains in the signaling system from Figure 3.

While this line of reasoning might help one see a connection in the dispositions associated with the P in PH and the P in HP, Section 2.2 showed that such connections can be merely accidental.

Certainly, the argument from section 2.2 does not work for the Pyow-Hack game. If sender B's P and H terms were replaced with Y and G, then there would be no justification for the basic receiver choosing from the same urns when receiving a solitary P from A as when receiving a solitary Y from B, which is an essential feature of the game. We know that a solitary P from A means the same thing as a solitary P from B because the receivers cannot condition their actions on which sender transmitted the term. More explicitly, one can see that the connection in dispositions is not merely accidental with the following chain of reasoning:

- The P in PH statements is dispositionally connected to the P in solitary P statements from A by receiver-compositionality. In both the corresponding group movement and leopard states (when A happens to be the solitary transmitter) successful action reinforces the prevalence of (ii)-balls in receiver C's PH urn.
- Since terms are sender independent solitary P statements transmitted by sender A are associated with the same dispositions as solitary P statements transmitted by sender B.
- The P in solitary P statements from B is connected to the second P in PP statements by receiver-compositionality. In both the corresponding leopard (when B happens to be the solitary transmitter) and urgent leopard states, successful action reinforces the prevalence of (i)-balls in receiver C's PP urn.
- Finally, there is a connection in dispositions between the second P in PP statments and the P in HP statements since both are transmitted by the same functional component, sender B.

Now if it were possible for just the receivers or just the senders to unilaterally determine the meaning of statements in the game, then this chain of reasoning, relying on features of both the senders and receivers, would be problematic. However, since acquisition of a signaling system requires both senders and receivers to have dispositions consistent with each other, there is a substantive connection between the P in PH and the P in HP.

Still, the fact that features of both senders and receivers are necessary to trace a connection between the P in PH and the P in HP shows that neither the senders in isolation nor the receivers in isolation can be said to represent the connection between PH and HP. Worse, consider Figure 3 with the following changes: swap the contents of C's HP and HH urns, likewise swap the contents of B's urgent eagle and eagle leaving urns. This results in an optimal signaling system (which can and has obtained under the same learning dynamics) where HP means urgent eagle/look up quickly and HH means eagle leaving/look up and elsewhere. In this signaling system there is no longer an overlap in the behaviors associated with HP and those associated with a solitary P. However, it should be noted that it is not possible to produce an optimal signaling system that breaks the dispositional connection between P and PP nor the connection between H and HH. This is because the way in which the model implements receiver-compositionality (via C's balls that have two interpretations) guarantees that, in optimal signaling systems, a single term transmission X will always be dispositionally connected with XP and XH statements.

These considerations suggest at least two ways in which future models could attempt to improve on the pyow-hack game given in this chapter. First, a model could attempt to modify the senders to better represent the connection between states of nature associated with similar actions. For example, the game could be modified such that for any state X $\in$ {leopard very near, leopard nearby, group is moving, nearby eagle leaving} there is some small probability that the basic senders draw at random from one of the other three urns in the set rather than from the X urn. This would have a nominal negative effect on the

success rate of a signaling system and would make signaling systems where P means leopard and HP means eagle leaving more likely to obtain than signaling systems where P means leopard and HP means urgent eagle. Second, a model could focus on attempting to make a single player represent a connection in meaning between the P in PH and P in HP. For example, one could modify the game to have receiver C, upon receiving a solitary P, draw at random with equal probability from the PP, PH or HP urns. This would in some sense force a connection between the meaning of PH and HP.

However both of these naive examples have down sides. The first example still results in neither the senders in isolation nor the receivers in isolation capturing the connection in the meaning of P in PH and the P in HP; additionally, it is still possible (just less probable) for a signaling system to obtain where P means leopard and HP means urgent eagle. The second example results in a significant negative impact on every set of strategy profiles' success rate and, in the limit, the contents of the HP urn will only nominally overlap with the contents of the PH and PP urns since receiver C has a strictly greater probability of visiting the HP urn for either a solitary H or an HP than her probability of visiting the HP urn for a solitary P.

Despite the noted issues, it is clear that the pyow-hack game exhibits a novel type of compositionality that advances our understanding of signaling games. The conjunction of sender independent terms with receiver-compositionality does provide reason, in optimal signaling systems, to consider the P in PH as being the same term as the P in HP. The model's introduction of sender independent terms disallows the argument by isomorphism, given in section 2.2. In contrast with this, it is easy to see that the prior hierarchical model given by Barrett, Cochran and Skyrms (Barrett et al., 2019, 2018) is isomorphic with a model where B's terms are **y** and **g** (rather than **b** and **r**) since both receivers in that model condition their actions on which sender transmitted a unitary term. However, if one tried to do the same with the pyow-hack game it would be mysterious why the receivers' dispositions towards P

were the same as their dispositions towards Y. While this chapter borrows and explicates receiver-compositionality from the prior hierarchical signaling game, the pyow-hack game does show how a novel type of compositionality can obtain.

Chapter 2

Three People Make a Tiger: Illusory Truth in Epistemic Networks

2.1 Introduction

It has long been known that repeated exposure to a statement can cause someone to perceive the statement as more plausible. A very early example of this comes from the ancient Chinese *Annals of the Warring States*. In the book of Wei, an advisor to the king of Wei is ordered to accompany the king's son as a hostage to Handan. The advisor, knowing political rivals will slander him while he is absent, tells a parable. The advisor asks the king whether he would believe a man making the implausible claim that there is a tiger in the marketplace. The king says he would not. If two people made the claim, the king says he would not immediately dismiss it. Finally, when the advisor inquires about three people making the absurd claim, the king says he would believe there is a tiger in the marketplace. The advisor tells the king that there is without doubt no tiger presently in the marketplace; simple hearsay is not evidence. The advisor reminds the king that he has more than three political rivals

and Handan is much further away than the marketplace. The king should expect to receive disinformation about the advisor's loyalty. Of course, when the advisor eventually returns to the court, the king, having believed the expected rumors, ostracizes the advisor. It is from this story that we get the idiom "three people make a tiger" (Crump, 1970; Goldin, 2005).

More recently in the psychological literature, the effect of repetition on beliefs has been called the truth effect or the illusory truth effect (Dechêne et al., 2009; Pennycook et al., 2017; Fazio et al., 2019; Fazio, 2020; Ransom et al., 2021). The basic experimental paradigm demonstrating this effect consists in participants ranking statements on a a scale from 0 (definitely false) to N (definitely true). Participants are then exposed to some form of information affirming a subset of the ranked statements. Finally, participants rank for a second time the truthfulness of the statements. Researchers can then measure the change in truthfulness ranking for statements that were repeated relative to those that were not. Experimental results show that participants believe the repeated statements are more likely to be true. The illusory truth effect has been replicated many time for several variations of this basic experimental paradigm (Dechêne et al., 2009). Though the effect is more substantial when statements are repeated word for word, it persists when the wording of or reasons for affirming the statement are varied (Dechêne et al., 2009; Ransom et al., 2021). Strikingly, the illusory truth effect persists for implausible statements, statement repeated by an unreliable source, and statements repeated in articles labeled with a warning that third party sources dispute the credibility of the article (Pennycook et al., 2017; Fazio et al., 2019; Fazio, 2020).

This chapter introduces the illusory truth effect to the epistemic network literature. In formal social epistemology and philosophy of science, epistemic network models provide a computational simulation based approach to studying how a network of agents gather and share information. These network models have been used to study the effects of a variety of biases on scientific inquiry and consensus formation (Holman and Bruner, 2017; O'Connor

and Weatherall, 2018; Wu, 2022; O'Connor and Gabriel, 2023). Previous explorations have found that some biases are beneficial to group learning. However, simulation results reported in this chapter show that the illusory truth effect is detrimental to a network's likelihood of reaching true beliefs. Some prior research has indicated that the probability of false consensus formation is near zero when networks are large or scientists are forced to explore unpopular and risky science (Zollman, 2010; Kummerfeld and Zollman, 2015; Rosenstock et al., 2017). However, simulation results in this chapter show the harms of the illusory truth effect persist for large networks with mandatory exploration. Section 2 overviews some prior literature and motivations for modeling the illusory truth effect in the context of scientific inquiry. Then, section 3 describes a common epistemic network model along with an extension of the model to incorporate the illusory truth effect. Section 4 presents the simulation results form the model. Finally, section 5 discusses those results.

2.2 Fluency in Social Epistemology

As mentioned in the intro, the illusory truth effect has been widely documented. A common explanation of the illusory truth effect is that it is a consequence of processing fluency (Dechêne et al., 2009; Pennycook et al., 2017; Fazio et al., 2019).[1] A premise of this explanation is that processing fluency, the ease with which we process information, is used in a heuristic for determining the veracity of a statement. More fluent information is judged as more likely to be true. Additionally, repeated exposure to the same or similar information makes that information easier to process. Consequently, repeated exposure to a statement makes it more likely to be judged as true. This explanation of the illusory truth effect makes

[1]While receptive to the explanation, Ransom et al. (Ransom et al., 2021) seem inclined to wait for a better understanding of the extent to which the number of repetitions correlates with the degree of change in truthfulness judgments before emphasizing processing fluency as a primary explanation of their experimental results. Others have criticized processing fluency in a related context, the mere exposure effect (Montoya et al., 2017). However, these critiques do not attack the feature of processing fluency explanations that is relevant to this chapter, which is that the illusory truth effect seems to be a product of some subconscious mechanism rather than being easily accessible to executive control.

it easy to bridge lab experiments with other contexts. In *Annals of the Warring States*, the king of Wei was repeatedly exposed to slanderous rumors about the loyalty of his advisor. By the premises of the processing fluency explanation, repeated exposure makes information from the rumors easier to process.

Recently, van der Linden has used the processing fluency explanation to motivate a contagion-based model of the illusory truth effect in social networks (van der Linden, 2022). Others have developed contagion based models of repeated information in social networks, but have done so without reference to lab experiments on the illusory truth effect (Piedrahita et al., 2018; de Oliveira et al., 2022). In contagion based models, an agent's beliefs are merely a function of whichever beliefs are most prevalent among those to whom the agent is connected to in a social network. If an agent is surrounded by people who believe a particular claim, then that agent will also come to believe that claim. In contagion models, agents have no means of gathering information independently of what is shared with them. So even if an agent is surrounded by people who believe something wildly implausible, say that Ding Liren is a terrible chess player, that agent will still adopt the implausible belief. This type of model is well suited social contexts in which people cannot collect their own evidence for evaluating claims.[2] **However, contagion models are not necessarily appropriate for contexts in which people are able to independently collect their own evidence about the truth of a claim.**[3]

In contrast with contagion models, this chapter develops a model in which agents actively gather and share information that provides reasons for their beliefs. It builds on a prior model, a social network solving a two armed bandit problem, introduced to formal social epistemology by Zollman (Zollman, 2007, 2010).[4] This base model is taken to be a good

[2]We also know that misinformation and disinformation on social media has gravitated towards being particularly repetitive (Horne and Adali, 2017; Tandoc, 2019). So the illusory truth effect might explain why that content is so repetitive.

[3]Nathan Gabriel revised the concluding sentence of this paragraph to correct an ambiguous pronoun.

[4]Prior to Zollman, this model was developed in the economics literature by Bala and Goyal (Bala and Goyal, 1998).

representation of scientific theory development and more general belief development about scientific topics (Zollman, 2007, 2010; Holman and Bruner, 2017; O'Connor and Weatherall, 2018; Wu, 2022; O'Connor and Gabriel, 2023). In a two arm bandit problem, each arm has a characteristic rate of payout, or success. Agents collect information about this characteristic rate of payout by conducting trials of successive pulls of an arm, each pull either succeeding or failing. Agents share their trial results with those to whom they are socially connected. Then, each agents' beliefs are updated, according to Bayes' rule, based on both the trial results that were self generated and the results that were shared with them. Since agents gather their own information about the world, this model represents agents who use semi-rational strategies to find out about the world, rather than passively adopting information from peers. This base model is then extended to include an illusory truth effect by allowing old information, that agents have already seen, to continue being shared and to continue impacting their beliefs.

It is easy to see why it is realistic to extend the base model with illusory truth dynamics. Scientists' publications constantly cite prior research creating repeated exposure to those articles which are most often cited. In accordance with the processing fluency theory, as scientists are repeatedly exposed to the cited information it becomes easier to process. Consequently, they will be more likely to judge the repeatedly cited information as true. Given that number of citations is easily quantifiable, the documentation of dissemination and citation biases is fairly robust. It indicates that media coverage and even the number of Twitter followers a scientist has can be predictive of how many citations a paper receives (Kiernan, 2003; Liang et al., 2014; Akella et al., 2021; Song et al., 2010; Borenstein et al., 2021). This imbalance in citations can be thought of as both a vehicle for repeated exposure to information, and, perhaps, an indicator of repeated exposure to information. Regardless of whether repeat exposure to information manifests itself in increased citations of that information, there are a variety of plausible scenarios, in which scientists might be repeatedly exposed to the same information; perhaps some research lab inducts new members with a repetition of

certain early results produced by the lab, or perhaps some scientists' form a social circle that frequently discusses a particular set of papers because those papers report results that are accessible to interlocutors from outside the scientists' area of specialization.[5] So it is realistic to model scientists as being exposed to repeated information. Then, the processing fluency theory explains why this repeated exposure would continue to influence scientists beliefs.[6]

Of course it would be absurd to suggest that one model reflects the behavior of all scientists, using any methodology, across all domains of scientific inquiry. As already stated, in the base model agents are attempting to determine which of two bandit arms has the higher characteristic rate of success. So the model is most representative of research in which scientists are directly trying to determine which of two options has the higher characteristic rate of success; for example, scientists might be attempting to determine which of two drugs is more effective for treating a disease, or which of two photoresists leads to fewer defects in photolithography. Still, experiments are rarely so simple; the effectiveness of a drug might be modulated by the presence or absence of a particular gene, or a photoresist's performance might vary by transistor size. The details of different real world scientific inquiries will determine how analogous they are to a bandit problem. However, it should also noted that, in the context of social networks, bandit models behave in a way that is qualitatively very similar to a broader class of models indicating that they might be appropriate for understanding a broader range of scientific inquiry than was just described (Smaldino et al.,

[5]The problem could be even worse. There are indicators that scientists become less critical of heavily cited publications. LaCroix et al. note several instances in which a heavily cited article continues to be cited after the article has been retracted (LaCroix et al., 2021). Furthermore, we can worry about scenarios in which a heavily cited article, while itself has not retracted, relies substantially on misleading information. For example, Google Scholar search shows well over 100 citations in the past year of Hauser et al.'s *The faculty of language* despite the article's argument relying substantially on some, circa 2012, retracted publications by Hauser (Hauser et al., 2002a; of Health and Services, 2012). Since the model given in section 3 does not involve falsified information nor restrict the repetition of information to only data points that are unrepresentative of the truth, it can be thought of as modeling a best case scenario.

[6]As already noted, there is some debate about whether the processing fluency theory is the best explanation of the illusory truth effect (Ransom et al., 2021; Montoya et al., 2017). But it is not contentious that the illusory truth effect seems to be a consequence of some unconscious bias outside the domain of executive control; this is all that is needed to support the claim that scientists' beliefs will continue to be influenced by repeated information.

2023).

2.3 Model

Section 3.1 gives a description of a base model. Then, section 3.2 describes its extension with the illusory truth dynamics. A few amendments to these dynamics, such as forced exploration, are described at the end of section 3.2.

2.3.1 Base Model

Zollman (Zollman, 2010) models scientific inquiry as a two armed bandit problem. Agents in a network collect and share evidence about which of two arms has higher payout rates. At each time step in the model, agents conduct a trial of 1000 pulls from the arm they believe to have higher rate of payout and share the results of the action with everyone they are connected to in the network. Agents then use Bayes' theorem to update their beliefs based on their own trial results along with the results shared to them by other agents.

Agents' beliefs about the probability of payout from a given arm is represented as a Beta distribution.

Definition (*Beta Distribution*) A function on [0, 1], $f(\cdot)$, is a beta distribution iff for some $\alpha > 0$ and $\beta > 0$

$$f(x) = \frac{x^{(\alpha-1)}(1-x)^{(\beta-1)}}{B(\alpha,\beta)}$$

where $B(\alpha,\beta) = \int_0^1 u^{(\alpha-1)}(1-u)^{(\beta-1)}du$. This follows an established method of cognitive modeling and makes it easy for agents to update their beliefs according to Bayes theorem (Lee and Wagenmakers, 2013). Beta distributions are a convienent method of representing belief since, for a beta distribution with parameters α and β as an agent's prior, n pulls

from a bandit arm and s successes, then Bayes' theorem gives the agent's posterior as a beta distribution with parameters $\alpha + s$ and $\beta + n - s$. Likewise, expectation is also easily calculated for beta distributions as $\frac{\alpha}{\alpha+\beta}$. Intuitively, in this context, beta distributions are just being used to track the history of success (α) and failures(β). If there is a history of many successes and few failures, this results in a large α and small β. Consequently, one's expectation that the next outcome will be a success is close to 1. Or if, say, there is a history of about equal numbers of successes and failures, the expectation will be $\frac{\alpha}{\alpha+\beta} \approx \frac{\alpha}{\alpha+\alpha} = 0.5$.

For all simulations, arm A's characteristic rate of payout is fixed at $p_a = 0.499$, arm B's characteristic rate is fixed at $p_b = 0.5$, and agents perform trials of 1000 pulls at each timestep.[7] Agents' initial beta distribution parameters are random values between 0 and 4.[8] Each simulation uses an Erdős-Rényi random network to represent the connections between agents (Erdős and Rényi, 1960).[9] Networks are deemed successful if they reach a stable consensus in which all agents in the network think the bandit arm with the higher payout is the one that objectively has higher payout (arm B).

[7]These are the values that were used in the original presentation of the base model (Zollman, 2010). They make the problem of determining which arm has a higher success rate difficult enough to generate interesting results.

[8]Using small values for agents' initial beliefs allows them to be very flexible in the first time step of the simulation. For example, consider $\alpha = \beta = 2$, then expectation is 0.5. But this expectation can easily be shifted. If the next three pulls are failures, expectation drops to $\frac{2}{2+5} \approx 0.286$. However, if larger values were used, say $\alpha = \beta = 500$, expectation is more rigid. In this case, three repeated failures results in an expectation, $\frac{500}{500+503} \approx 0.499$, that is still close to 0.5

[9]Erdős-Rényi random networks are networks constructed by creating a connection between agents with probability p_{er} for each pair of agents. Results reported in this paper use $p_{er} = 0.5$. Simulations for other p_{er} values were run but did not qualitatively change results (smaller values do produce slightly higher success rates). Only fully connected Erdős-Rényi random networks are used in simulations; that is only those networks for which for any pair of agents in the network there is some path of connections between agents in the network connecting the given pair. This precludes networks from becoming polarized merely because there is a strict subset of agents that are not connected to the rest of the agents in the network. After a network structure was randomly determined at the beginning of a simulation it was not changed for the remainder of the simulation.

Using random network structures makes it clear that the phenomenon explored are not merely a consequence of a specific network structure. It should be noted that Zollman's 2010 model explored three specific network structures, cycle, complete, and a wheel structure (Zollman, 2010). While this chapter does not explicitly report results for these network structures, simulations were run for all three network structures. Using different network structures does produce different specific values for simulation results, but qualitatively the results were unchanged; i.e. for all network structures more sharing of old data causes lower correct consensus rates and the blunted benefit of exploration or increased network size can be observed in these structures as well.

2.3.2 Illusory Truth Dynamics

In the base model just described, agents only ever update on new evidence gathered that round. This section describes an extension of that model for studying the illusory truth effect. In this extension, information can be repeatedly shared throughout the network, including to an individual who has already updated on it. This is accomplished by giving agents a memory in which they store trial results from previous timesteps and allowing them to share some of those results from memory at each timestep in addition to sharing results from the current timestep.

Agents have memory of up to M many previously viewed trial results. At the beginning of a simulation agents have no data in memory. On each time step, agents conduct trials and share results with those they are connected to as they do in the base model. Additionaly, on each timestep, agents randomly choose up to S many trial results from memory to share with agents to whom they are connected.[10] In this first extension of the base model, agents do not differentiate between new and old trial results that are shared with them.[11] They update their beliefs according to Bayes' theorem, using both information from trials performed on the current time step as well as information shared from memory, as if each set of trial results is new information. This would reflect all shared information having the same degree of impact on processing fluency. Finally, agents randomly choose at most D many data points from among those that were received in the current round to replace a random data point in memory.

Simply varying parameters M, S, and D across simulations generates a lot of information about how the illusory truth dynamics can effect a network of agents' ability to arrive at true

[10]Specifically agents randomly select S of the M locations in memory to pull trial results from for sharing. However, in early timesteps of a simulation some locations in memory might not yet be populated with trial results. Hence, the wording of "up to S many trial results".

[11]In a further amendment of the dynamics to be discussed, agents discount the information shared with them from other agents' memory. This blunts the impact of the illusory truth effect on agents' belief formation.

beliefs. However, one might worry that it is unfair to use the simplifying assumption of agents allowing new and old trial results to have the the same degree of impact on their beliefs. Intuitively, when reading a paper, new information from the paper has a more significant degree of impact on one's beliefs than information cited in the paper. A further extension of the illusory truth dynamics has agents decreasing the impact of old information, i.e. any information shared from memory,[12] on their beliefs by some discount factor. For example, suppose an agent's prior belief about an arm's rate of payout is given by $\alpha = \beta = 0.5$, new information of 5 of 10 pulls paying out is shared with them, additionally old information from three prior trials is shared as 2 of 10, 4 of 10, and 4 of 10 payouts. On the simple dynamics this gives the agent expectation:

$$\frac{5+5+2+4+4}{10+10+10+10+10} = 0.4$$

On the dynamics decreasing the impact of old data with a discount factor of 0.1, the agent's expectation for the arm is:

$$\frac{5+5+0.2+0.4+0.4}{10+10+1+1+1} \approx 0.478$$

With a discount factor of 0.01, the agent's expectation for the arm is:

$$\frac{5+5+0.02+0.04+0.04}{10+10+0.1+0.1+0.1} \approx 0.498$$

So, this amended version of the dynamics will lessen the effect of the illusory truth dynamics.

This chapter explores two additional variations of the dynamics. One variation forces agents to explore using the ϵ-greedy strategy from prior literature (Sutton et al., 2018; Kummerfeld

[12]Here, "old information" means any information that was shared to an agent from memory. It is not necessarily information that the agent has seen before. Suppose agent B is connected to both A and C, but that A and C are not connected to each other. Then A will discount information shared shared to her from memory by B, even if that information was initially generated by C and A has never seen it before.

and Zollman, 2015). Rather than always best responding, agents using the epsilon greedy strategy conduct trials for the arm they think is better with probability $1 - \epsilon$, and they conduct trials for the arm they believe to be worse with probability ϵ. Secondly, a variant of the illusory truth dynamic reserves half of each agents memory for information about each arm. As in the prior illusory truth dynamics, for up to D many data points, each agent randomly selects a location in memory to store that data point, but, in this variant of the dynamics, the information is only written to that location in memory if it is about the arm that that location in memory is reserved for. Other than the changes just explicitly stated, both variants preserve all aspect of the illusory truth dynamics that were first described.

2.4 Results

In the following figures, each point represents an average over 2000 simulations of the model with each simulation lasting 50,000 timesteps. This was generally sufficient for $> 99\%$ of the simulations to reach a stable consensus. The only exception to this was simulations where exploration was forced, in which case simulations were run for 100,000 timesteps to guarantee that $> 99\%$ of simulations reached consensus. A network has reached a stable consensus if all agents in the network agree on which arm has the higher rate of payout and running the simulation for additional timesteps is unlikely to change which arm this is.[13] Results show data points for networks of 4, 6, 9, 12, 15, and 25 agents. Figure 2.1, clearly shows that increasing how much agents share old information correlates with decreasing a network's probability of success. Holding everything else fixed, varying the amount of information that is stored to memory has minimal impact on correct consensus rates.

[13]Since agents always pull the arm that they think has the higher rate of payout and their beliefs are relatively rigid after 50,000 timesteps, a consensus is stable if the objective rate of payout for the preferred arm is greater than each agent's expectation for the alternative arm.

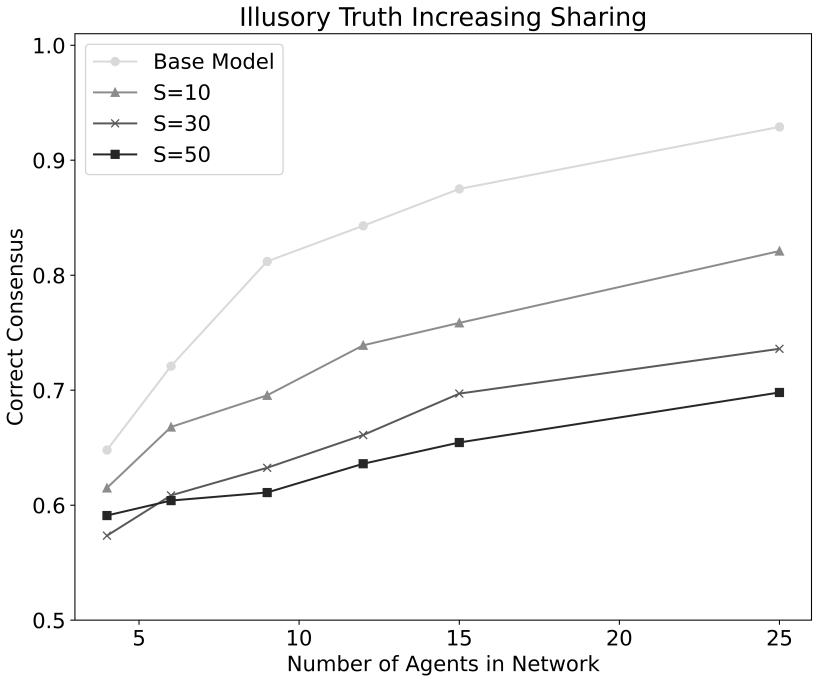


Figure 2.1: Average correct consensus rates for 2000 simulations with $M = 100$, and $D = 5$.

Simulation results did show that increasing the agents' memory size decreases the impact of the illusory truth effect. However, this was an artifact of a resulting decrease in sharing. Since agents are blank slates at timestep 0, with no trial results in memory, and S is the number of, potentially unpopulated, locations in memory from which trial results are shared, increasing the agents memory size, M, while holding S and D fixed decreases the likelihood of old trial results being shared in the initial stages of the game. Consequently, it is not increased memory that are correlated with the changes in correct consensus rates. Rather, it is increased exposure to repeated information in early stages of a simulation that correlates with decreased probability of a network reaching correct consensus.

A natural question to ask is how much old information is persisting in the networks with the illusory truth effect. The table below shows this data for networks of 15 agents. For three different sets of parameters, it shows at the 10th timestep what percentage of agents memory is older than 5 timesteps, 100th timestep what percentage of agents memory is older

than 10 timesteps, and 10,000th timestep what percentage of agents memory is older than 100 timesteps. For these later two metrics we see very high percentages showing that old information can become very dominant in agents's memories. However, the table highlights the negative impact of old information during initial stages of agents' belief formation. The parameters with the least old information in memory when measured at 10th and 100th timestep, $M = 100$, $S = 20$, and $D = 10$, had the highest correct consensus rate; this occurred despite those parameters generating the highest percentage of old information at the 10,000th timestep.

	older than 5 after 10 timesteps	older than 10 after 100 timesteps	older than 100 after 10^4 timesteps	corect consensus rate
M= 50 S=20 D=10	9.2%	90.1%	56.9%	0.661
M= 100 S=20 D=10	2.7%	83.7%	75.5%	0.708
M= 100 S=40 D=20	5.5%	94.3%	74.8%	0.677

Table 2.1: Presence of Old Information in Networks

On the dynamics amended to include a discount factor, Figure 2.2 shows results in which the impact of old information on beliefs is reduced by a factor of 0.1, 0.01, or 0.001. Predictably the negative effects of the illusory truth dynamics are reduced. When old repeated information is discounted by a factor of 0.001, networks have correct consensus rates close to the rates of the base model without any illusory truth. When repeated trial results are discounted only by a factor of 0.1, this measurably improves correct consensus rates, but not substantially. Not depicted are results for discount factors greater than 1. Simulations with a discount factor of 10 produced correct consensus rates about the same or only nominally lower than a discount factor of 1. It turns out there's a floor on how low the illusory truth dynamics can push success rates. Perhaps that is a silver lining.

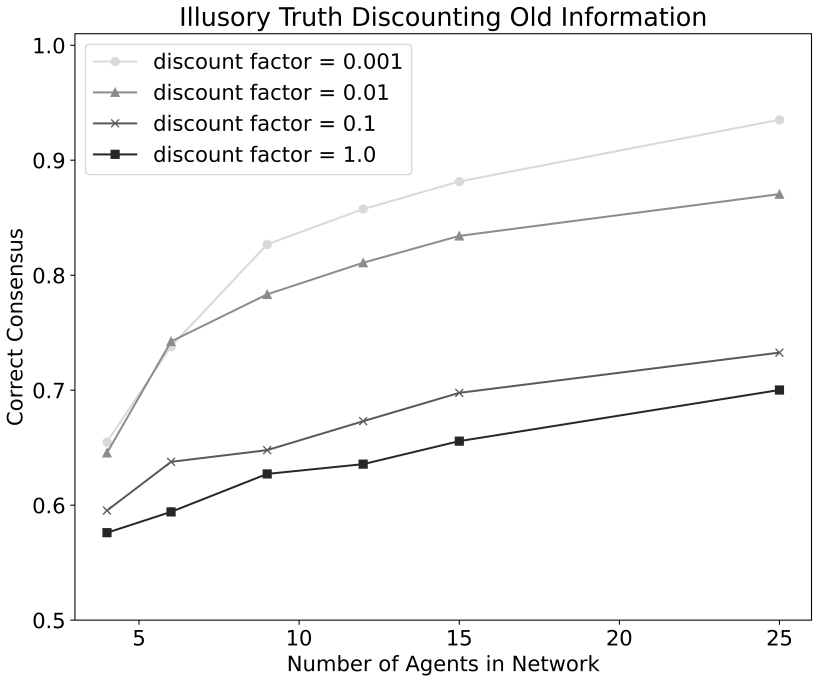


Figure 2.2: Average correct consensus rates for 2000 simulations with $M = 100$, $S = 40$, and $D = 20$.

As stated in the introduction, the negative impact of the illusory truth dynamics is robust for large networks with forced exploration. This can be seen in Figure 2.3, which reports results for ϵ-greedy agents ($\epsilon = 0.001$). Networks being "large" is a relative claim. Zollman's original presentation of the base model did not explore networks larger than 11 agents, and previously forced exploration was explored with networks of 8 agents (Zollman, 2010; Kummerfeld and Zollman, 2015).[14] So the results for networks of 25 agents, shown in Figure

[14]This section presents results for the parameters chosen to best illustrate the phenomenon being investigated. Consequently, there are a variety of superficial ways in which the parameters used here differ from some of the prior literature being referenced. Several prior papers have use fixed network structure (e.g. a wheel, cycle, or complete) network structure rather than using Erdős–Rényi random network structures that are used in the results reported here. While these different structures result in different particular correct consensus rates, simulations have been run for all of these network structures and they did not produce any qualitative change in the results. Using Erdős–Rényi random network structures makes it clear that what is being reported is not a consequence of a particular network structure. Similarly, this section shows results for a relatively large number of timesteps compared to other literature because this is what is required under the illusory truth dynamics for networks to reach a stable equilibrium. Most prior literature including Zollman's 2010 paper use several thousand timesteps, but Kummerfeld and Zollman's 2015 paper inexplicably restricts simulations to only 50 timesteps. This has the effect of conflating simulations in which there is a stable equilibrium of all agents having an incorrect belief about which bandit arm is better and simulations in which the network has merely failed to reach a consensus given the small number of timesteps simulated.

2.3, are for networks substantially larger than what is typically reported. It is true that correct consensus rate continues to increase as network size grows. For networks of 100 agents using ϵ-greedy strategy ($\epsilon = 0.001$, $M = 100$, $S = 30$, and $D = 5$), 1000 simulations resulted in a correct consensus rate of 0.881. While this is a significantly higher rate of success, it is still a notably low rate given that Rosenstock et al. report a correct consensus rate of 1 for networks larger than 40 agents (Rosenstock et al., 2017).[15] Similarly, holding all other parameters fixed, increasing the frequency of forced exploration increases the correct consensus rate, but not as much as one might expect based on the behavior of the base model. For $\epsilon = 0.001$-greedy agents, simulations of networks of 25 agents with the illusory truth dynamics resulted in a correct consensus rate of 0.824; in comparison, for 0.01-greedy agents without the illusory truth dynamics, the correct consensus rate was 1 for all networks composed of 9 or more agents.

Such a restriction actually produces results that are even more strongly in support of the qualitative result reported here, but the restriction is unwarranted. Lastly, showing results for 0.001-greedy agents highlights how little exploration is required to produce high correct consensus rates in the base model compared to the model with the illusory truth dynamics. But it can be noted that, given their modeling choices, Kummerfeld and Zollman suggest that 0.17-greedy agents are optimal a fully connected network of 6 agents. Restricted to only 50 timesteps, this generates a correct consensus rate of 0.668 in the base model and 0.492 with the illusory truth dynamics ($M = 100$, $S = 30$, and $D = 5$). This would further support what is reported here if the data points weren't worthless. When a more reasonable number of timesteps is used (i.e. 10^5 timesteps), 0.17-greedy agents in a fully connected network of 6 agents have a correct consensus rate of 1.00 in the base model and 0.896 with the illusory truth dynamics ($M = 100$, $S = 30$, and $D = 5$), which is a significant reduction in correct consensus rates given that significantly lower exploration rates generate correct consensus rates of 1.00 in the base mode. The negative impacts of the illusory truth dynamics are robust.

[15] Rosenstock et al.'s results are for the base model without forced exploration. As can be seen in Figure 2.3, 0.001-greedy agents have a correct consensus rate of 1 for networks larger than 15 agents.

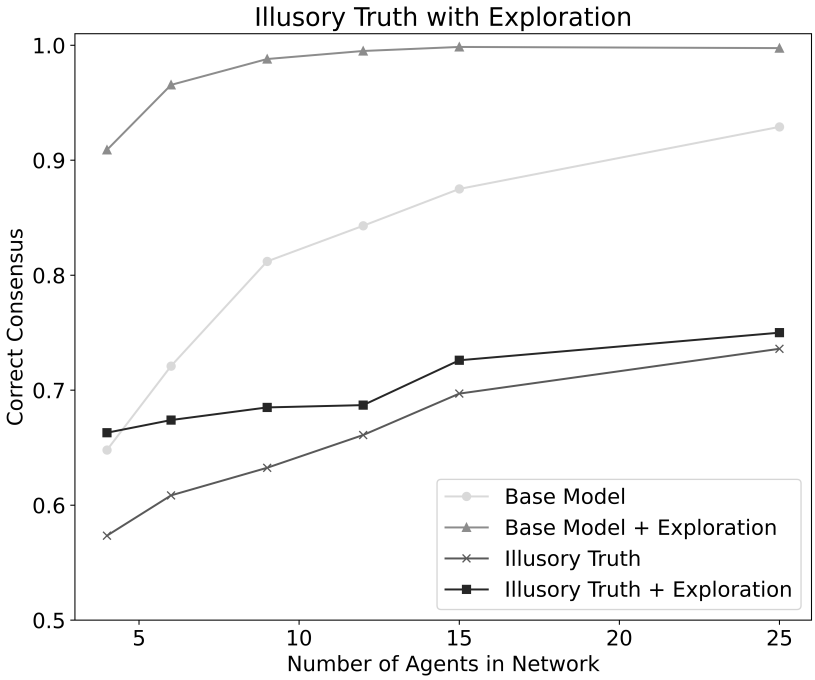


Figure 2.3: Average correct consensus rates for 2000 simulations of ϵ-greedy agents with $\epsilon = 0.001$, $M = 100$, $S = 30$, and $D = 5$.

An intuition one might have about the illusory truth dynamics is that failures are a result of agents' memories becoming flooded with information about just one of the bandit arms. This is not unlike what happens in the base model. Since agents always pull the arm they think has the highest rate of payout, their beliefs about the rate of payout for that arm actually converge towards the correct rate for that arm. Thus, in the base model, a network's failure to arrive at the correct belief about which arm has the higher rate of payout always means the agents have mistakenly come to believe the objectively better arm has a lower rate of payout than it actually has. So running with intuitions from the base model, one might think that the harms of the illusory truth effect might be overcome if half of agent's memory is reserved for data about each arm. Surprisingly, the variation of the illusory truth dynamics with half of agents' memory reserved for each arm resulted in even lower correct consensus rates.[16] It seems that reserving agents half of agents memory for each bandit arm simply

[16]For example, with parameters: $M = 100$, $S = 40$, $D = 20$, and network size = 15, the original illusory truth dynamics had a correct consensus rate of 0.677 and the variation with half of agents memory reserved

allows aberrant information about both arms to persist in the network.

2.5 Discussion

Scientists are not immune to cognitive biases and the bias exhibited in the illusory truth effect seems to be outside the domain of conscious executive control. Consequently, the harmful effects of the simulated illusory truth dynamics should be taken seriously. Of the different variations of the dynamics explored, small discount factors produced the highest correct consensus rates. However, the discount factor essentially amounts to a parameter that directly modulates the extent to which the illusory truth dynamics impacts agents' belief formation. It does not reflect a real world variable that we could intervene upon in order to improve the outcomes of scientific research. Some variants of the illusory truth dynamics do reflect interventions that could actually be performed. The ϵ-greedy dynamics could reflect an institution forcing scientists to explore unpopular and risky hypotheses(Kummerfeld and Zollman, 2015). The reserved memory dynamics could reflect scientists intentionally equalizing their exposure to publications about competing hypotheses. However, these variants respectively produced a minimal and a negative impact of correct consensus rates. So, further research on possible interventions would be valuable.

Currently, there is no empirical evidence that would allow the model to be tuned to the degree that repeated information impacts scientists beliefs, perhaps by tuning the discount factor to reflect empirical evidence. So one could hope that future research indicates that the illusory truth effect is less significant than the reported simulation results suggest. However, there are a couple of ways in which the simulation results could be characterized as a best

for each arm had a correct consensus rate of 0.607. Worse, one could argue that, since the reserved memory dynamics result in a 0.5 probability a selected data point not being written to a memory location (there is a 0.5 chance the randomly selected location is reserved for data from the other arm), then it is more comparable to use parameters $M = 200$, $S = 80$, $D = 40$ for the reserved memory dynamics; but, this generated an even lower correct consensus rate of 0.587 for networks of 15 agents.

case scenario. First, the model does not incorporate low quality information. Information could be low quality for subtle reasons, such as scientists only reporting positive results, or for overt reasons, such as falsifying results. While journals do have mechanisms for retracting low quality information, there is evidence that retracted papers continue to be cited after retraction (LaCroix et al., 2021). Second, it is plausible that something like processing fluency results in repeated information impacting scientists' beliefs more than novel information. This would correspond to a model with a discount factor greater than 1, which produces equally bad results.[17]

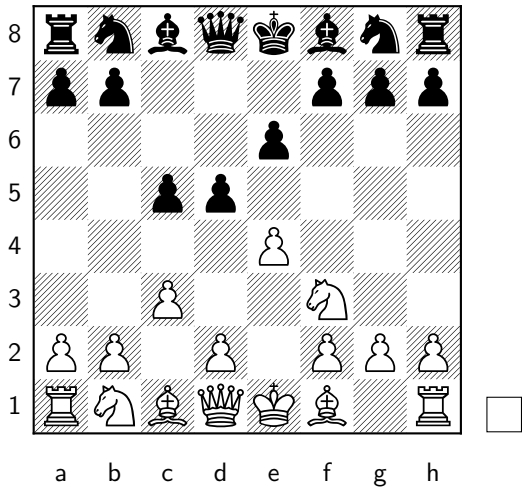
In conclusion, results showed that the illusory truth dynamics negatively impacted correct consensus rates. This negative impact was relatively robust when compared to known methods of improving network success rates in the base model, such as forcing exploration or increasing network size. In the base model large networks with forced exploration essentially guarantee a simulation will reach the correct consensus (Rosenstock et al., 2017; Kummerfeld and Zollman, 2015). The robustness of the results from the illusory truth dynamics is surprising and highlights the importance of investigating how different network dynamics interact. There are ways future research can further investigate consistent the illusory truth effect is in negatively impacting scientists ability to arrive at true beliefs. For example, Smaldino et al. have identified several different social network models that behave qualitatively similar to models based on a bandit arm problem (Smaldino et al., 2023). If translating the illusory truth dynamics into those models yields similar results, then it might be time to start believing there is a tiger in the marketplace.

[17]Additionally, one could worry about even more subtle and indirect ways in which scientists are exposed to repeated information such as prima facie independent lab results being produced using the same flawed legacy software.

Chapter 3

Franco-Sicilian Abstraction: transpositions, transitivity and transfer learning with Lewis-Skyrms signaling games

3.1 Introduction

1 e4 c5 2 ♘f3 e6 3 c3 d5 

In the above chess position, computer engines suggest that capturing the d pawn, **4 e×d5**, is the best move. However, about 25% of the time, chess masters instead advance the e pawn, **4 e5**. Why? "White is hoping a **1...c5** [Sicilian Defense] player won't be comfortable in a French [Defense]. And that's also why Black often avoids the objectively best **4...♘c6 5 d4** and prefers a gambit with **4...d4!? 5 c×d4 c×d4 6 ♗b5+ ♘c6 7 ♗c6+ b×c6 8 ♕a4** or **6...♗d7 7 ♘×d4!?**" (Soltis, 2007). In chess, a transposition occurs when a sequence of moves begins with the moves of a particular opening, say a Sicilian Defense, but results in a board position typical of a different opening, say a French defense. When White sub-optimally advances her e pawn, she attempts to transpose the game into a French Defense. If Black sub-optimally advances her d pawn, she avoids the transposition at the cost of a pawn.

This story of chess transpositions serves dual analogical roles for this chapter. First, it demonstrates how we might think about the type of abstraction to be described. While this chapter explicates a specific and relatively simple example of abstraction using transitive inference, the chess story highlights our ability to use this type of abstraction in more complicated contexts. Consider Frege's conception of abstraction. According to him, we know the abstract form of a sum by knowing that a particular formula is a sum, say 2+3=5, and seeing that replacing any subset of the numbers in the formula with different numbers will also yield a formula that is a sum (though, it is certainly not guaranteed that the resulting formula's reference is True). In this conception of abstraction we see different particulars, numbers, can satisfy the same roles, an addend or result, to make a summation formula. It is also the case that in Frege's conception of abstraction, we know exactly what particulars can be substituted into the formula, any number and nothing that is not a number. While this may be a good concept of abstraction in the context of formal languages, this chapter allows that there may be vagueness in the boundaries around what particulars can satisfy the same role in an abstract category.

Chess masters use the abstract category of an opening to narrow the calculations that they have to make. While a particular board position might be novel, their familiarity with the abstract category of positions associated with a particular opening guides what strategic ideas they consider (de Groot, 1978; Gobet et al., 2004). The fact that transpositions are possible entails that some board positions might occupy a vague boundary between different categories of opening and associated strategic ideas. So when not discussing the concrete Lewis-Skyrms signaling game to be presented, this chapter's discussion of abstraction simply refers to contexts where different particulars take on the same functional role. That is, just as two different numbers can play the role of addends, two different board positions can play the role of reminding a chess master of strategic ideas in the French Defense; but while we know exactly what can and ought to play the role of an addend, there may be board positions occupying a vague boundary around the set of positions that are or ought to be associated with strategic ideas in the French Defense.[1]

Second, the story points us towards a philosophical problem that accompanies abstraction. Let's suppose that we think beliefs that lead to success in action are true beliefs. The chess story immediately shows why we might doubt this. A player may believe **4 e5** is a stronger move than **4 e×d5** due to her having more success, i.e. more wins, with that continuation. However, this may merely be a consequence of her having more successfully abstracted the strategic dynamics of board positions associated with the French Defense than the Sicilian; i.e. if your only tool is a hammer, then every problem looks like a nail. As another example, one might think that we come to represent the world as having three dimensional Euclidean structure by abstracting over experiences and that since we successfully interact with the world using this representation, then it is true that the world has this structure. But just because we successfully navigate the world representing it as Euclidean does not mean there is not alternative representation with which we would have more success. Worse, sections 3.3.3

[1]See De Marzo and Servedio for a statistical clustering of board positions that coheres with theoretical openings (De Marzo and Servedio, 2023).

and 3.3.4 will suggest that some abstractions may be a consequence of (possibly arbitrary) innate dispositions rather than being wholly empirical.[2]

Some research has investigated how humans store abstract information about chess positions (Chassy, 2013). However, this chapter focuses on simpler cognitive tasks. This allows us to examine correspondingly simple models that are more easily interpreted; in turn, this allows us to develop a clearer understanding of the philosophical implications of the models rather then getting lost in the details of continually evolving complex models. Accordingly, this chapter models two tasks taken from the comparative psychology literature since we take nonhuman animals to have less sophisticated cognitive resources available for learning and inference. Of course, these tasks have also been studied with human subjects who behave similarly to the nonhuman animal subjects that succeed at the tasks.

The first task that is modeled is a transitive inference task, which involves testing whether a subject trained on contiguous pairs of linearly ordered stimuli can abstract information about that relation to apply it two a novel non-contiguous pair of stimuli. Psychology experiments have shown children as well as a variety of nonhuman animals succeed at the transitive inference task. Later discussion of this task, in accordance with the second analogical role of the chess transpositions story, notes that expected behavior for the task should not strictly be called "success" as the inference is not warranted.

The second task is a nonsense grammar task. It is a variation of experiments habituating human infants and nonhuman primate infants to sound patterns, called nonsense grammars, and testing whether the habituation persists for novel sounds presented in the same patterns. While this is an entirely distinct task from the transitive inference task, it can be similarly described: it tests whether a subject trained on a relation between triplets of stimuli can

[2]Explicitly, the worry is that if our representation of something in the world is reflective of an arbitrary disposition, then this may entail that in some relevant way we understand the world as having a particular structure when that structure is merely a consequence of our arbitrary innate disposition. Some have, perhaps wrongly, interpreted Kant as having a concern along these lines (Janiak, 2022; Warren, 1998).

abstract information about that relation to apply it to a novel triplet of stimuli. This second task is important for showing that the themes abstracted from the transitive inference task are not merely consequences of particular details of that task.

3.2 Transitive Inference

Abstractly, transitive inference task is as follows: Five objects/stimuli are arranged in a serial order: A, B, C, D, E. The agent to be tested is conditioned on adjacent pairs in the serial ordering through rewards for selecting the first object in a presented pair. E.g. If presented with A-B the agent is rewarded for choosing the item on the left. If presented with D-C the agent is rewarded for choosing the item on the right. After the agent has learned to respond correctly to pairs of objects that are adjacent in the serial order, she is then tested on the non adjacent pairs B-D and D-B. (A and E are not tested since choosing A is always rewarded and choosing E is never rewarded; thus, choosing the left object in the pair A-C might reflect choosing A always having been rewarded rather than reflect the transitive ordering having been represented.) Choosing the correct item in the non-adjacent pairing is then understood as the agent representing the transitive ordering.

Concretely, Bryant and Trabasso showed that children ages 4-6 succeeded at the transitive inference task. The stimuli used were rods of different colors and varying lengths. The children were shown adjacent adjacent pairs, in a serial ordering by length, were shown to the children, with the rods protruding through holes such that one inch of the rod was visible. Thus the children could observe the color of the rods, but not the length. After being asked which rod in a pair was taller or shorter, the rods were removed and laid flat on a table in front of the child so that they could examine whether they had answered correctly. After the children learned to answer correctly for pairs that were adjacent in the ordering by length, they were then test on non adjacent pairs and tended to answer correctly. This experiment

was taken to disprove earlier work by Piaget who claimed that children did not understand or represent transitive relations (Bryant and Trabasso, 1971).

Concretely, Gillan used colored plastic containers to show that chimpanzees succeed at the transitive inference task. The colored containers were assigned an arbitrary serial ordering and then adjacent pairs were presented to the chimps with a food reward in whichever container was first in the serial ordering. After the chimps were conditioned to almost always choose the container with the food, they were then tested on the nonadjacent pairs. The result was that the chimpanzees tended to chose the container that had food in it, i.e. the one that was first in the serial ordering (Gillan, 1981). In a similar experiment, von Fersen et al. showed Pigeons succeed at the transitive inference task using black and white ink blots as the stimuli and and using a contraption that rewarded the pigeons with food when they pecked at the correct ink blot in a pair (von Fersen et al., 1991). Likewise, Davis showed that rats succeed at the transitive inference task using the stimuli of distinct odors to mark pairs of doors that the rats could open to retrieve a food reward (Davis, 1992). Bond et al. showed that scrub-jays succeed at the transitive inference task using a mechanism similar to von Feresen et al., but with projections of different colors of light rather than ink blots as the stimuli (Bond et al., 2003).

It should be noted that, in the nonhuman animal experiments just listed, the transitive inference is not warranted; e.g. food always being in the green container when a green and yellow pair is presented and always being the the yellow container when a yellow and red pair is presented does not entail that the food will be in the green container when a green and red pair is presented. This contrasts with Bryant and Trabasso's study involving different colored rods of varying lengths because transitivity does hold for length: i.e. if a green rod is longer than a yellow rod and that yellow rod is longer than a red rod, then this entails that the green rod is longer than the red rod. That the transitive inference is warranted in the human experiment and unwarranted in the nonhuman animal experiments is merely a coincidence

of the experiments that happened to comprise the most pertinent background information. There have been experiments showing successful transitive inference in nonhuman animals for which transitivity does hold for the stimuli, such as physical dominance, and there have been experiments showing "successful" transitive inference in human participants for which transitivity does not hold for the stimuli, such as characters from a written language that the participants do not know (Grosenick et al., 2007; Frank et al., 2005).

There have also been many computation models of transitive inference. Barrett's (Barrett, 2014) model is the most transparent. In Barrett's model, agents are first conditioned on a full serial ordering of stimuli including non-adjacent pairs. The agents are then presented with a new set of serially ordered stimuli, this time conditioned on only adjacent pairs. Modeling a type of transfer learning, the agents are allowed to transfer some of their dispositions from the first set of stimuli to their dispositions for the new stimuli by allowing initial signals transmitted for the new set of stimuli to be processed as if the signals were the old stimuli. That is, the agents learn to map the new stimuli to their evolved dispositions for the old stimuli, which now serve as an intermediary representation in an agent's evolution of dispositions for the new stimuli. This allows the agents to infer the appropriate response to a novel pairing of nonadjacent stimuli from the new ordering most of the time.

Many computational models of transitive inference do not explicitly invoke transfer learning. However, given that the correct action for the test pair is underdetermined by the contiguous pairs that a model is trained on, some type of prior bias must be present to explain successful action for the test pair. For example, Siemann and Delius develop a model that makes use of lateral inhibition, whereby, on the transmission of two signals, only the signal of larger magnitude is acted upon. When presented with a stimulus pair from a serial ordering, an agent might initially transmit both a signal to choose the left object and a signal to choose the right object, but only the signal of greater magnitude will go through (Siemann and Delius, 1998). On reflection, it is clear that this is a type of transfer learning since

the lateral inhibition mechanism is itself a presupposed set of dispositions representing the serial ordering of signal magnitudes. Other models are more opaque. De Lillo et al. give a three layer neural network model, which, although not explicitly invoking lateral inhibition, is trained using a cost function that always moves the weights in the direction of hidden layer nodes always associating the rewarded stimulus with a greater magnitude than the magnitude associated with the other stimulus in a pair (or always with a smaller magnitude depending on the randomly determined initial state of the network) (De Lillo et al., 2001). So this model might also be thought of as presupposing a serially ordered representation.[3]

3.3 Transitive Inference with Transfer Learning

This section presents three models of transitive inference by means of types of learning that can be characterized as transfer learning. The first model is a straightforward example of transfer learning. In this model, an agent first learns to choose the preferred stimuli in a pair, being trained on both adjacent and non-adjacent pairs in a serial ordering. Then, the dispositions from this first phase of training are co-opted in a second phase. In this second phase of training an agent is trained on only adjacent pairs in a new set of serially ordered stimuli. Since this second phase of training involves mapping the new stimuli to the dispositions formed in the first phase of training, agents subsequently succeed when tested on the non-adjacent pair for the new stimuli. This first model is essentially the same as Barrett's 2014 model, albeit with some negligible differences (Barrett, 2014). Call this the "diachronic model".

The primary purpose of the diachronic model is to simplify the presentation and explanation of the other two models. Like the diachronic model, both of these models make use of two sets

[3]It is possible that this sort of prior bias could have evolved for an organism to navigate an environment in which there is often a linear relation between objects. However, section 3.3.4 will explain that such a bias could be a result of relatively abstract combinatorial properties. This makes it plausible that the bias facillitating the inference is a spandrel rather than reflecting environmental evolutionary pressures.

of serially ordered stimuli. However, agents are trained on both sets of stimuli simultaneously. Additionally, agents are trained on pairings between the two sets of stimuli; e.g., let A, B, C, D, E be the first set of stimuli and a, b, c, d, e be the second set of stimuli, then agents are trained to choose B over c in a B-c pair. They are only trained on non-adjacent pairings when both components of a pairing are from the first set of stimuli. One of these models is a simple extension of the Diachronic Model in its first phase of training while the other makes use of signal magnitudes. Call these models the "synchronous model" and "synchronous magnitudes model" respectively.

3.3.1 Diachronic Transfer

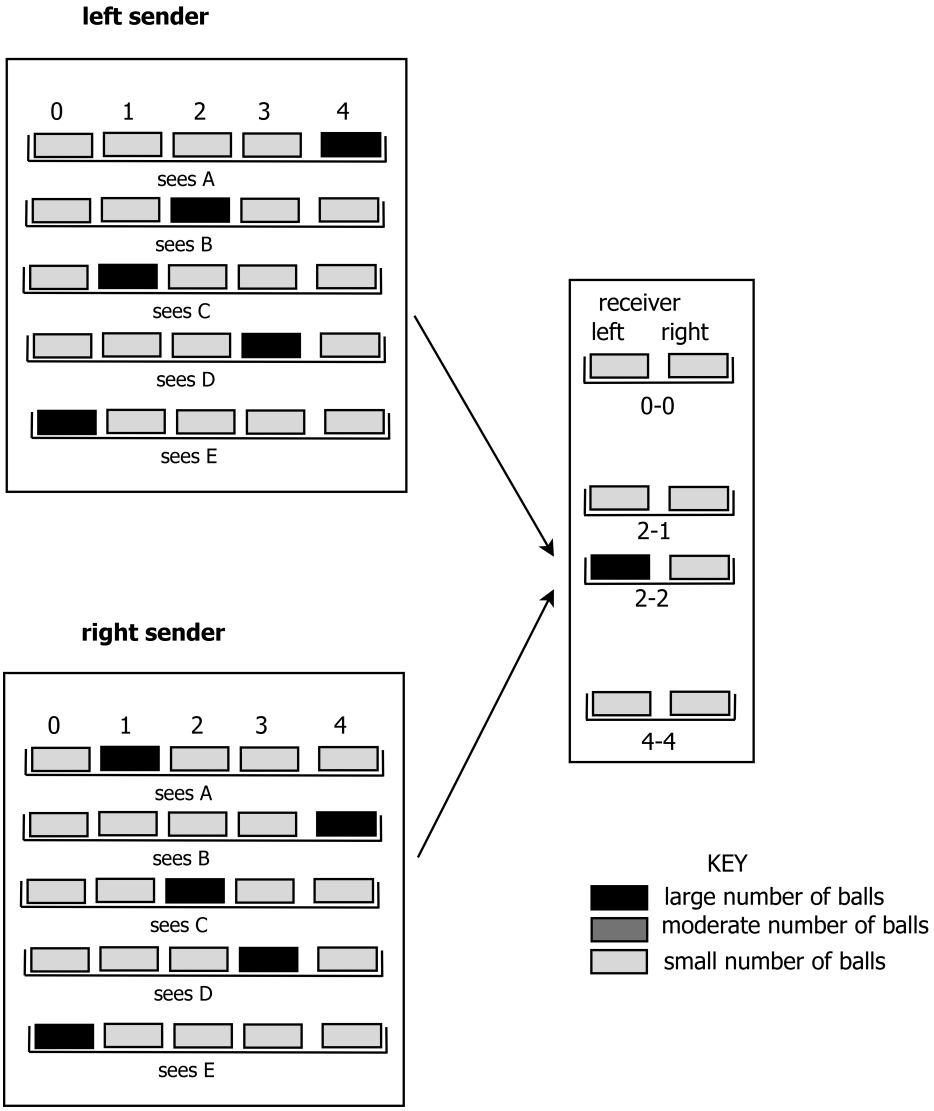


Figure 3.1: First Training Phase of Diachronic Model

The diachronic transfer learning model is a two sender one receiver game during the first phase of training. Each sender see one of the two stimuli in a pair composed of either adjacent or non-adjacent stimuli in the serial ordering. That is, letting A, B, C, D, E be the stimuli in their serial order. Then every pairing of two distinct stimuli is in the training set. The senders each have one urn for each of the five stimuli that can be observed, and each urn begins the game with a single ball of each type. Figure 3.1, depicts each sender

with five types of balls, but this is parameter in the model that can be varied. The receiver has one urn for each of the possible combinations of signals that can be received from the senders; each of these urns begins the game with two balls in it, a left ball and a right ball. During the training, pairs are selected at random with equal probability. Suppose, the pair B-D is selected. Then the left sender will draw from her B urn and the right from her D urn. Suppose the left sender draws a 2-ball and the right a 1-ball, then the receiver will draw from her 2-1 urn. The receivers draw then determines whether the left or the right stimuli in the pairing will be chosen. If the correct stimmuli is chosen, in this case B which is on the left, then the urns that were drawn from are reinforce with additional balls of the type that were drawn. If the incorrect stimuli is chosen than the urns are punished by removing balls of the type that were drawn, so long as at least one ball of each type remains in each urn. This is the same basic Roth-Erev reinforcement with punishment learning that has been described in detail elsewhere and represented as $(+1, -1)$ (Barrett and Gabriel, 2022). The notation $(+r, -p)$ represents on a success adding r-many balls of the type drawn to each of the urns drawn from, and on a failure removing p-many balls of the type drawn from each of the urns drawn from, so long as at least one ball of each type remains in each urn.

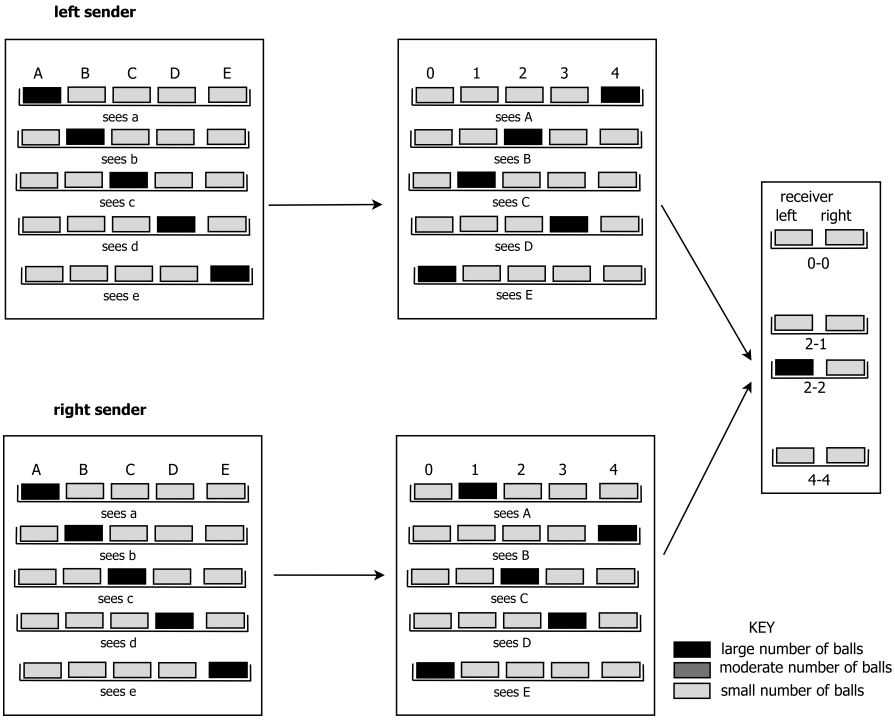


Figure 3.2: Second Training Phase of Diachronic Model

In the second phase of training (see Figure 3.2), the left and right senders are given new urns, one for each of the new set of stimuli, a, b, c, d, e. Each of these urns have ball types A, B, C, D, and E which then determine which urn from the first phase of training will be drawn from. Draws from the urns the A, B, C, D, and E urns then determine what is transmitted to the receiver who retains her urns from the first phase of training. It is in this way that dispositions from the first phase of training are co-opted in learning correct actions for the second set of stimuli. Learning during the second phase of training also proceeds according to Roth-Erev reinforcement with punishment learning dynamics. The testing phase is then performed with the b-d and d-b pairs. No learning occurs during the testing phase.

3.3.2 Synchronous Transfer

In the synchronous transfer model (see Figure 3.3), there is only one training phase. The dynamics of the model are the same as the model in the first phase of the training for the diachronic model. However, the training consists in pairings from both the A, B, C, D, and E stimulus set and the a, b, c, d, and e stimulus set. Pairings across the stimulus sets are allow, e.g. B-c. But non adjacent pairs are only trained for stimuli from the first set, e.g. C-E. There are no pairings in which both elements are occupy the same position in the serial ordering, e.g. D-d is disallowed. So, A-D, B-A, d-C, and e-d are all members of the training set and neither B-d nor a-d are member of the training set. During training pairs from the training set are selected at random with equal probability and the learning dynamics are the same as in the diachronic model, i.e. Roth-Erev reinforcement with punishment. After the training phase is complete, the testing phase consists of the b-d and d-b pairs and no learning occurs during testing.

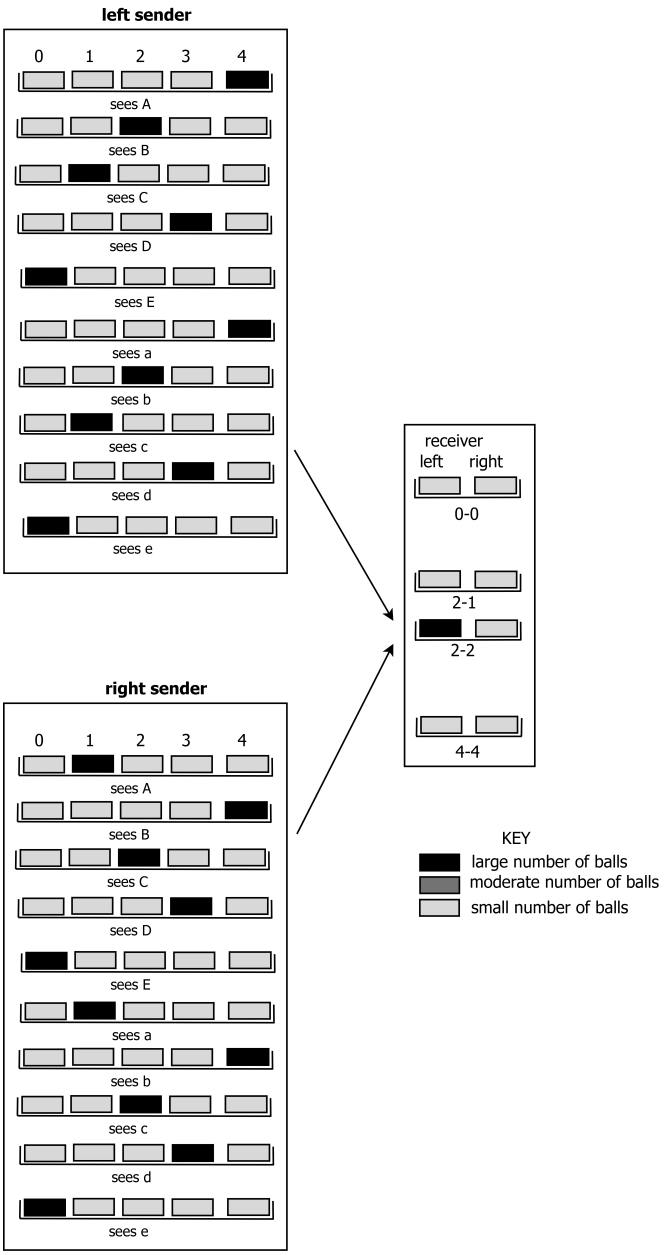


Figure 3.3: Synchronous Model. For simplicity, this model is depicted with each of a sender's urns containing five ball types. However this is a parameter that can be varied. The results section focuses on simulation results for which each urn contained ten ball types.

3.3.3 Synchronous Magnitudes

In the synchronous magnitudes model (see Figure 3.4), the training and testing sets are the same as in the synchronous model. Likewise there is only a single training phase before the

testing phase. The only difference in the synchronous magnitudes model is in the senders' urns. Rather than the senders each having a single urn for each of the stimuli that can be observed, each stimuli is associated with M-many urns with each of these urns beginning the game with one 0 ball and one 1 ball. Figure 3.4 depicts the model with M=6, though later results are reported for M=10. When a stimuli is observed by one of the senders, that sender draws from every urn associated with that stimulus and then transmits the sum of her draws. For example, if the left sender, upon seeing A draws five 1 balls and one 0 ball, then she will transmit 5; if the right sender upon seeing B draws two 1 balls and four 0 balls, she will transmit 2; then the receiver will draw from the 5-2 urn just as she would in the prior models. Again, the learning dynamics is urn based reinforcement with punishment. Thus if, in the example just given, the correct stimuli was chosen, then the left sender would reinforce 1 balls in the five urns that she drew 1 balls from and reinforce 0 balls in the urn that she drew the 0 ball from. Likewise the right sender would reinforce 1 balls in the two urns that she drew 1 balls from and reinforce 0 balls in the four urns that she drew zero balls from.

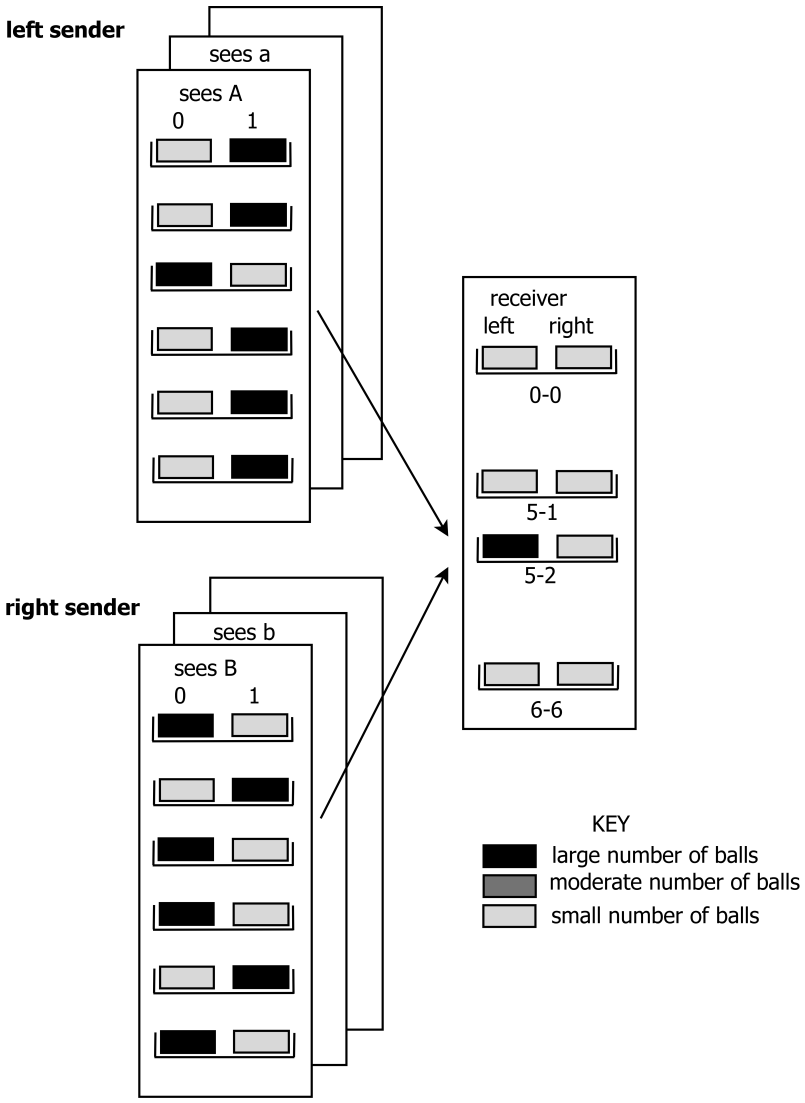


Figure 3.4: Synchronous Magnitudes Model. For simplicity, this model is depicted with each sender having six urns for each stimuli that can be observed. However, this is a parameter that can be varied. The results section focuses on simulation results for which senders have ten urns for each stimuli that can be observed.

3.3.4 Transitive Inference Simulation Results

The diachronic model performs very poorly. Running simulations for a variety of parameters never yielded a significantly high mean success rate during just the first phase of training. As stated earlier, the primary purpose of the diachronic model is as a stepping stone towards the

synchronous and synchronous magnitudes models. Barrett (Barrett, 2014) achieves significantly high success rates with a similar model by utilizing invention in the learning dynamics. This limits the terms available to the senders to only those that have led to success in action, and this limitation makes learning much easier. As will be seen, the synchronous magnitudes model will yield a similar increase in success rates by making the middle magnitude the most likely magnitude to be transmitted during the early stages of a simulation.

For the synchronous model, with ten ball types in each of the senders' urns, using (+4, -11) reinforcement with punishment, and 2×10^9 training plays before testing yeilded a mean success rate of 0.7228 during training and a mean success rate of 0.66293 during testing. However, the simulations also yeilded a very odd result when training was restricted to just the second set of stimuli and only training on adjacent pairs. The model still performed above chance during the testing phase. Simulations had a mean success rate of 0.7256 during training and 0.6217 for the testing phase.[4]

What is even more confusing is that, prima face, it seems that combinatorially we should expect performance during testing to be worse than chance when restricting training to just the second set of stimuli with only adjacent pairings in the training set. For the five stimuli in the second training set and ten terms in each of the senders urns there are essentially 831,847,680 strategy profiles for which every pairing in the training set gets mapped to the correct action and pairings in the test set tend to get mapped to the *incorrect* action, but there are only 766,177,920 strategy profiles for which every pairing in the training set gets mapped to the correct action and pairings in the test set tend to get mapped to the *correct* action.[5] Thus, it seems like mean success rates for the test pairs should be less than

[4]This is averaging over just 300 simulations as 2×10^9 training plays takes a decent amount of time to simulate.

[5]This is brushing over some details. By using the term 'essentially', what is indicated that this count is of pure strategy profiles rather than mixed strategy profiles. So whats really being counted are the pure strategy profiles that a simulation could converge to rather than the mixed profiles that necessarily obtain (since the learning dynamics require that there is always at least one ball of each type in each urn). The count of 831,847,680 is of pure strategy profiles such that every adjacent pairing is mapped to the correct action and where at least one of the b-d or d-b pair is mapped to the incorrect action and neither is mapped

0.5 when the training phase only consists of the second set of stimuli and the training set only consists of adjacent pairings. The resolution to this problem lies in observing that for strategy profiles that map at least two distinct pairings to the same receiver urn, there are only 78,943,680 strategy profiles for which every pairing in the training set gets mapped to the correct action and pairings in the test set tend to get mapped to the *incorrect* action, there are 137,649,600 strategy profiles for which every pairing in the training set gets mapped to the correct action and pairings in the test set tend to get mapped to the *correct* action, and the learning dynamics makes these strategy profiles much more likely to obtain (since it is easy for the senders dispositions towards a pairing, for which successful dispositions have not yet been reinforced, to evolve a mapping to a receiver urn for which successful dispositions have already been reinforced for a distinct pairing).[6]

For the synchronous magnitudes model, with senders having ten urns for each of the stimuli, using (+4, -11) reinforcement with punishment, and 10^7 training plays before testing yielded a mean success rate of 0.999 during training and a mean success rate of 0.845 during testing; 635 simulations of 1000 had a success rate greater than 0.9 for the testing phase. When restricting training to only the second set of stimuli and only training on adjacent pairs, simulations had a mean success rate of 0.999 during training and 0.734 during testing. Again, the explanation for this better than chance performance when only training adjacent pairings

to the correct action. Likewise, the count of 766,177,920 is of pure strategy profiles such that every adjacent pairing is mapped to the correct action and where at least one of the b-d or d-b pair is mapped to the correct action and neither is mapped to the incorrect action. There are 203,400,360 pure strategy profiles such that every adjacent pairing is mapped to the correct action and exactly one of the b-d or d-b pair is mapped to the correct action and the other is mapped to the incorrect action.

[6]Precisely, this explanation needs to be more nuanced for the synchronous model since it was rare for strategy profiles with training phase success rates near 1.0 to obtain. However, this nuance is unnecessary when considering the synchronous magnitudes model for which the learning parameters reported in the next paragraph yielded training phase success rates near 1.0 for every simulation. Here, combinatorially, only 12%, of strategy profiles for which every pairing in the training set gets mapped to the correct action and at least one of the b-d or d-b pairs gets mapped to the same urn as one of the pairs from the training set, are such that at least two pairs from the training set get mapped to the same receiver urn. However, in simulations about 40% of the final strategy profiles have this property. Moreover, for the synchronous model, only 13.6% of strategy profiles for which every pairing in the training set gets mapped to the correct action and at least one of the b-d or d-b pairs gets mapped to the same urn as one of the pairs from the training set, are such that at least two pairs from the training set get mapped to the same receiver urn. However, in simulations 60% of the final strategy profiles have this property.

from the second set of stimuli seems to be the learning dynamics tendency to produce strategy profiles in which multiple pairings get mapped to the same receiver urn.[7]

The synchronous magnitudes model can be thought of as baking some ordered structure into the model since there is an ordered relation between the magnitudes that can be transmitted: in the initial state of the model, for any stimuli, if there are ten urns per stimuli, then the most likely magnitude to be transmitted by each sender is 5 and the least likely transmissions are 0 and 10; furthermore, once a disposition to transmit some magnitude, say 7, has been reinforced, then the most likely deviations from this transmission are 6 and 8 rather than all alternative transmissions be equally likely as in the synchronous model. These properties of magnitudes can be thought of as reflecting neurons having a base firing rate that is then increased or decreased by learning events or can even be thought of as a potential spandrel, a la Gould and Lewontin, simply resulting from the connectivity of neurons (Gould and Lewontin, 1979).

Two variations of the synchronous magnitudes mode were simulated to ensure undue structure was not being baked into the model. First, it was trained on a single set of stimuli, a, b, c, d, and e being conditioned to choose according to the regular ordering for adjacent pairs, but the reverse order for non-adjacent pairs; e.g. for the pair c-b, choosing right was reinforced and choosing left was punished, and for the pair a-d, choosing right was reinforce and choosing left was punished. This variant was successfully learned with a mean success rate of 0.998 across simulations. Second, a variant was explored where for single set of stimuli, a, b, c, d, and e training on only adjacent pairs reinforced choosing according to the serial

[7]As stated in a prior footnote, only 12%, of strategy profiles for which every pairing in the training set gets mapped to the correct action and at least one of the b-d or d-b pairs gets mapped to the same urn as one of the pairs from the training set, are such that at least two pairs from the training set get mapped to the same receiver urn. However, in simulations about 40% of the final strategy profiles have this property. This is close, but not entirely sufficient for explaining the 0.734 mean success rate during testing. It could be that some amount of transitivity is baked into the model or simply that further investigation of the combinatorial properties of the game reveals further insight. While the number of strategy profiles that have at least two pairs being mapped to the same receiver urn were counted, it could be that strategy profiles with strictly more than two pairs being mapped to the same receiver urn obtain and that these profiles are even more biased towards the testing pair being mapped to the correct action.

ordering, but training also included an additional stimuli x, such that choosing left for e-x, right for x-e, right for a-x, and left for x-a were reinforced. That is, the model was trained on adjacent pairs in a circular rather than linear ordering. This yielded a mean success rate of 0.999. So, it seems clear that the model is perfectly capable of representing relations other than linear transitive orderings.

The circular ordering variant of the synchronous magnitudes model is notable because it actually reflects an experiment that was performed with pigeons. Von Fersen et al. in their black and white ink blots version of the transitive inference task hypothesized that the pigeons were relying on a type of analogue magnitude representation, called value transfer, during training and that this explained their success during the testing phase. Thus they predicted that it would not be possible to condition the pigeons on adjacent pairs according to a circular ordering. Contrary to their prediction two of three pigeons did learn the circular ordering (von Fersen et al., 1991). So the synchronous magnitudes model shows how the pigeons may have relied on magnitudes for learning yet still been able to represent a circular ordering.

3.4 Transferring to Abstraction

As previously stated, this chapter's discussion of abstraction simply refers to contexts where different particulars take on the same functional role. In the transitive inference task, we can see a simulation has learned to attribute the same functional role to stimuli from each serial ordering that occupy the same place in their respective ordering if the final dispositions are such that the correct stimuli are chosen in the b-d and d-b test pairs (as already reported, this occurred roughly 63% of the time for the simultaneous magnitudes model). To make this claim concrete, suppose there are five ancient coins each with a head and tail side. Suppose further that the coins are serially ordered by value and that a subject is unfamiliar with the

coins, i.e. upon seeing the heads side of a coin she cannot tell you what the tails side looks like and vice versa. In this scenario, the stimuli of a coin's head or tail is equally capable of fulfilling the functional role of indicating where in the serial ordering the coin is. Now, we could describe a deductive process by which the subject upon learning the ordering of the coins according the head sides and having more limited exposure to information about the tail sides deduces which tail stimuli occupies the same place in the ordering as each head stimuli. However, what the models show is how simple associative processes can also lead to different particular stimuli occupying the same functional role. The fact that the model has no prior exposure to the b-d and d-b pairs shows that it is actually dispositions from the first serial ordering to the second rather than merely independently learning the appropriate dispositions for each set of stimuli.

It is also clear that the different particular transitive inference models presented accomplish this abstraction in the same way. The receiver urns are organized in the same way for all three models. This in conjunction with having separate left and right senders who condition their action on just a single stimulus in a pair allows signaling systems to obtain that exhibit already reinforced dispositions towards novel stimulus pairings. As will be seen in the following section, this type of model architecture can accomplish abstraction in tasks other than transitive inference.

3.5 Nonsense Grammar

In a series of papers documenting experiments on cotton-top tamarins (*Saguinus oedipus*) and human infants, Hauser et al. claimed to have shown that the cognitive mechanisms for processing complex grammatical structures was uniquely human (Hauser et al., 2002b, 2001; Saffran et al., 2008). Critically, the content of these papers was used to support arguments in the prolific "Faculty of Language" paper he co-authored with Chomsky and Fitch

(Hauser et al., 2002a). It must be noted that Hauser falsified data in at least one of these papers, making human infants appear to be more competent than the cotton-top monkeys at processing particularly complex grammatical structures. However, data on simpler grammatical structures seems to be accurate and has been independently replicated by Neiworth et al. (Neiworth et al., 2017). The most relevant of these experiments by Neiworth et al. is summarized in the following paragraph.

The ability to learn simple grammatical patterns was assessed using a habituation paradigm with 16 adult cotton-tops as subjects. During a habituation phase, cotton-tops were exposed to consonant vowel sequences in an 001 or 011 pattern, e.g. "pupuki" or "pukiki". Seven different consonants (p, b, d, k, t, m, and n) and two vowels (i and u) were used to make the consonant vowel sounds that composed the 001 and 011 sequences. After a group of two or three monkeys was exposed consonant vowel sequence in only the 001 pattern or only the 011 pattern, a testing phase consisted in novel consonant vowel sounds (e.g. "wa", "ji", "la", "ri") arranged in sequences of both the same pattern that had been habituated and the alternative pattern. For example, if the 001 pattern was habituated, the monkeys might be tested on "lalari" as the same pattern and "wajiji" as the alternative pattern. Whether the same or alternative pattern was presented first was varied from group to group. The result was that the monkeys spent significantly more time looking towards the novel sounds in the alternative pattern than the novel sounds in the same pattern. This was interpreted as the monkeys becoming habituated to the pattern from the habituation phase and being capable of detecting the differences between the two nonsense grammars.

The details of the nonsense grammar model that follow differ from the comparative psychology experiment that inspired the model, but follow the experiment in spirit. In the nonsense grammar model, four nonsense grammars are considered: (i) 000, (ii) 001, (iii) 011, and (iv) 010. Sequences are composed of five different stimuli (A, B, C, D, E) arranged in one of the four nonsense grammar patterns. During training, sequences are selected at random with

equal probability to be classified as (i), (ii), (iii), or (iv). The training set contains all sequences except those that contain stimuli D or E in pattern (iv). Then, in the testing phase, it is checked whether the sequences DED and EDE are successfully classified as pattern (iv).

The model's architecture, depicted in Figure 3.5, is essentially the same as the synchronous model. The only real difference is that there are three senders rather than two, the receiver correspondingly has urns indexed by three terms rather than two, and the receiver's urns begin the game with one ball of each of four types (one for each grammar type that a sequence can be classified as). While depicted with each sender's urns containing five different ball types, results are reported for each sender's urns containing ten different ball types.

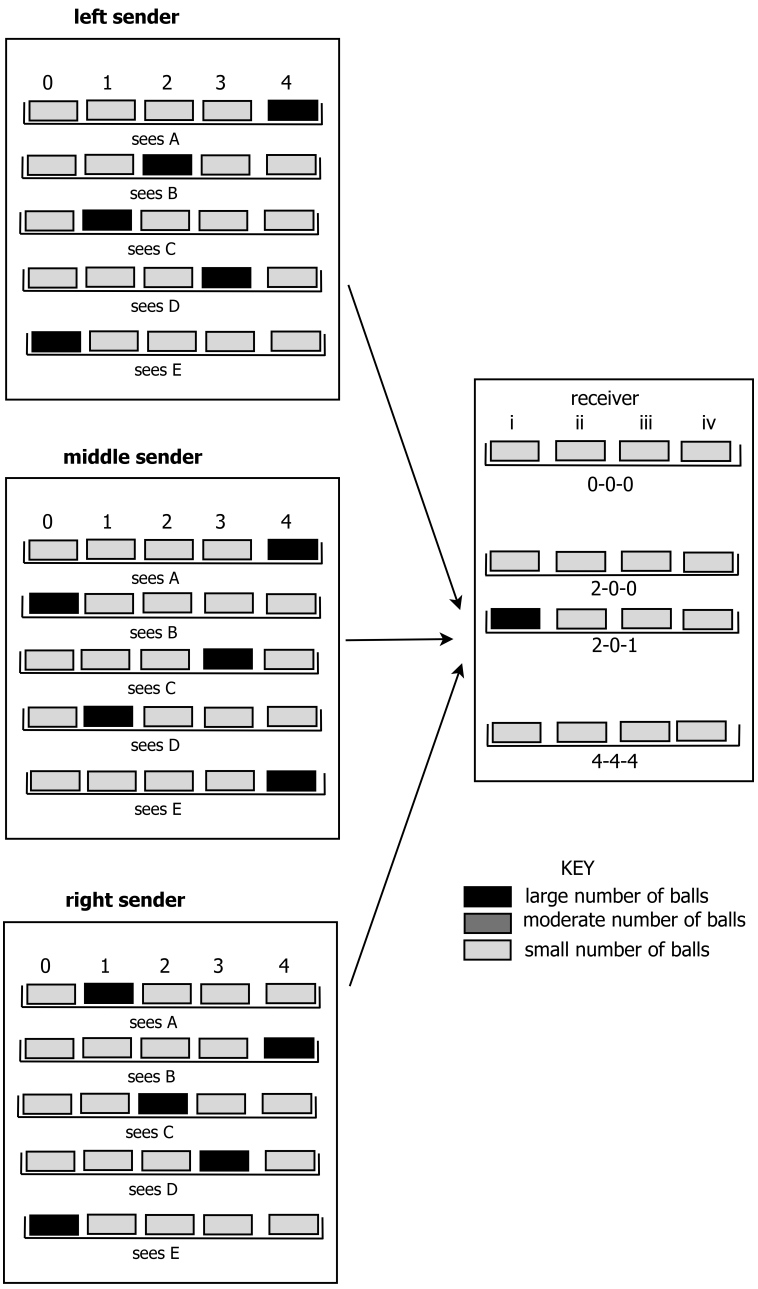


Figure 3.5: Synchronous Nonsense Grammar Model. For simplicity, this model is depicted with each of a sender's urns containing five ball types. However this is a parameter that can be varied. The results section focuses on simulation results for which each urn contained ten ball types.

For the synchronous nonsense grammars model, simulations were run using [+4, -10] reinforcement with iterated punishment and 10^7 training plays before testing (see Barrett and Gabriel for an explanation of reinforcement with iterated punishment (Barrett and Gabriel,

2022)). During the training phase, all simulations achieved a success rate above 0.85 and 91.1% of runs had a success rate greater than 0.99. During testing, nearly all simulations (96.9%) had a success rate better than chance (0.25), 58.1% of simulations had a success rate greater than 0.75, and 33.5% of simulations had a success rate greater than 0.9. So, a significant number of simulations correctly classified the test sequences.

3.6 Discussion

The models presented in this chapter show how a low level associative learning dynamics, specifically Roth-Erev reinforcement with punishment and reinforcement with iterated punishment, can be employed to succeed at both the transitive inference task and the nonsense grammar task. In all of the models, success involved different stimuli getting mapped to the same intermediary variable, i.e. the same receiver urn.[8] For example, the stimuli pairs B-D and b-d can be mapped to the same receiver urn, transferring the successfully conditioned dispositions towards B-D to b-d. This only happens if the left sender transmits the same signal for B as she does for b, and likewise the right sender sends the same signal for D and d. Concretely, this may look like an agent learning to associate the head side of a coin, call it B, with the same dispositions as the tail side of the same coin, b. In this way, a rudimentary type of abstraction is realized in the models when they evolve successful dispositions; different particular stimuli play the same functional role of eliciting the appropriate action. This is analogous to different particular board positions guiding chess masters to the same strategic ideas in virtue of the, potentially novel, board positions being associated with the same opening.

Abstraction provides significant heuristic utility to finite boundedly rational agents. Chess masters would have no use for abstract categories of openings if they had the capacity to

[8]The term "intermediary variable" is used here in accordance with Marr's account of the algorithmic level of explanation (Marr, 2010).

calculate every possible continuation of a board position. In fact, reliance on this sort of abstraction and chunking of common piece configurations, while common amongst human players, is virtually non-existent in computer chess engine methods of position evaluation (Gobet et al., 2004). However useful, abstraction's role as a heuristic carries with it epistemic concerns. This chapter's introduction gave an example of how a person's familiarity with success playing the against the French Defense may lead to a suboptimal transposition from the Sicilian into a French.[9] If humans had unlimited cognitive resources, a chess master could take into account strategic ideas for all openings in a given board position rather than only those associated with a particular opening. Masters can also fail to notice a transposition leading to evaluating a position according to a suboptimal set of strategic considerations (Soltis, 2007).[10]

Building on the chess analogy, the synchronous magnitudes model illustrates how abstraction can be facilitated by innate biases and we can worry that such biases lead us to systematically misrepresent the world. Now it could be that an organism is innately (and possibly for arbitrary reasons) biased towards representing things in its environment as linearly ordered because it evolved in an environment where it was adaptive to represent as being linearly ordered. But it seems equally plausible that such a bias could be a spandrel of neuron

[9]Specifically, this is objectively a suboptimal transposition for white when black is equally capable of playing a French or Sicilian. It is only in service of illustrating the epistemic concern that this paragraph ignores the possibility that white is making a strategic bet that black is less comfortable in French positions than Sicilian positions.

[10]Similar epistemic concerns can be raised for chunking of piece configurations. While there may be principles that are generally true of some common piece configurations, those principles need not hold for all positions. It is said that "a knight on the rim is dim", referencing the general principle that it is typically a mistake to move a knight to the edge of a board, say a5, because on the edge the knight can only control four squares rather than the eight it can control in the center of the board and also because, compared to a bishop, rook or queen, a knight requires substantially more moves to transition from attacking one side of the board to the other. But this principle might be violated if it allows the knight to be exchanged for a more valuable piece. Nimzovich dedicates the entire second part of his iconic book, *My System*, to analyzing common piece configurations (Nimzovich, 1964). What me might worry is that we are innately disposed only condition our judgments on a strict subset of relevant piece configurations. For example, novice chess players often find it easier to analyze the linear movement of the bishop than the non-linear movement of the knight and consequently systematically miss important strategic considerations involving knights. This might reflect a general bias towards representing things as being linearly related. It is an open empirical question whether chess masters' chunking of piece configurations is such that they systematically make suboptimal judgments when certain piece configurations occur.

connectivity.[11] Regardless of whether an innate bias is a result of environment pressures or an accidental consequence of some other adaptive trait, it is an open question whether that bias leads to systematic misrepresentation of the world. If objects in an environment are typically linearly related, an organism biased towards representing things as linearly related might represent those few things that are not linearly related as being linearly related. If an organism is innately biased towards represent things as linearly related for arbitrary reasons, it might still be able to represent things as not being linearly related; e.g. results showed that the simultaneous magnitudes model could learn a non-linear relation between the stimuli when conditioned to choose according to one ordering for contiguous stimuli but according to the reverse ordering for noncontiguous stimuli.

It seems that there are a couple viable responses to a generic worry that abstraction facilitated by innate biases puts us in an epistemically precarious position. First, it is true that the synchronous magnitudes model demonstrates how abstraction can be facilitated by some, perhaps arbitrary, innate bias. However, it is also the case that the model demonstrated an ability to represent a non linear relation despite its bias towards representing things as linearly ordered. Recall that the model was successfully trained to chose according to the regular ordering when presented with adjacent pairs, but also conditioned to chose according to the reverse ordering when presented with nonadjacent pairs; for example, choosing left was rewarded when presented with the pair a-b or b-c, but choosing right was rewarded when presented with the pair a-c.

By analogy, suppose we found some species that succeeded on the transitive inference task with minimal training, but was so innately biased towards linear representation that the species could not be conditioned to correctly choose the winner in rock paper scissors (i.e. conditioned to chose left for the pairs rock-scissors, scissors-paper, and paper-rock as well as

[11] In defense of this claim, it might be noted that the architecture of the synchronous magnitudes model was not intended to exhibit a bias towards linearly ordered representations. Upon discovery of this bias, the author's first response was to look for an error in the model's code.

conditioned to chose right for the pairs scissors-rock, paper-scissors, and rock-paper). The species continued survival is indicative of it generally representing the world well enough sustain itself and procreate. But in the context of rock paper scissors, the species is a miserable failure. Upon discovering that some way in which we abstract information is facilitated by some innate bias, we can worry that perhaps we are like this species there is some niche context in which we are miserable failures. Given our generally proficient navigation of the world we might be entirely blind to our failure in this niche context. We might simply represent such contexts as being governed by random chance rather than a recognizable pattern.

What the synchronous magnitudes model shows is that such worries can be too hasty. Abstraction can be facilitated by an innate bias in conjunction with enough flexibility in the learning mechanism to represent the world as contrary to the bias if necessary. Novice chess players often find it easier to understand tactics involving the bishop which moves linearly compared to tactics involving the knight which does not move linearly. It could be that this reflects some innate bias towards linear representation, but it would be wrong to conclude that we are incapable of appropriately representing tactics involving knights.[12]

Beyond highlighting the fact that the synchronous magnitudes model is able to represent things as being contrary to its innate bias given the appropriate training, it is worth considering the power of accomplishing abstraction through low level associative learning. What the models in this paper show is how minimal the preconditions for abstracting from experience can be. If we were to give a linguistic description of how one might abstract a serial ordering from pairs or a grammatical relation from triplets, it would be difficult to do so without using language that appears to presuppose some prior representation of the type of

[12]Anecdotally, it is well known that novices struggle more with the knight than the bishop. As far as academic research goes, Gobet et al. (2004) does make explicit claims that the power of the bishop pair in open positions is more easily understood by humans than queen and knight vs queen and bishop citing Sturman (1996) in support of this claim. But, the evidence is weak, essentially amounting to anecdotes from world champion Capablanca and Grand Master Timoshchenko.

relation that is being learned.[13] But the dynamics of the models previously described need not operate at the coarseness of the language used to describe them. The model does not need to first reliably choose left for an A-B pair before those dispositions can be co-opted for learning to choose left in response to observing the a-b pair. Those joint dispositions can grow gradually and simultaneously.

To conclude, this discussion has not argued that our inferences guided by abstraction and biases are always warranted. We can always worry that we are suboptimally advancing a pawn. However, there are some reasons to be optimistic about our use of abstraction. Abstraction can be a useful heuristic tool for navigating a complex world (or complex chess positions). Furthermore, while it may be facilitated by some biases, there is also evidence that those biases can be overcome. Finally, it was suggested that abstraction through associative learning has the benefit of requiring minimal preconditions for learning when compared to explicit symbolic reasoning.

[13]This mirrors an interpretation of one of Kant's arguments that our representation of space cannot be empirical because in order to abstract the concept of space from relations between objects occupying space we first have to represent them as such (Janiak, 2022; Warren, 1998).

Bibliography

Akella, A. P., Alhoori, H., Kondamudi, P. R., Freeman, C., and Zhou, H. (2021). Early indicators of scientific impact: Predicting citations with altmetrics. *Journal of Informetrics*, 15(2).

Arnold, K. and Zuberbühler, K. (2006). The alarm-calling system of adult male putty-nosed monkeys, cercopithecus nictitans martini. *Animal Behaviour*, 72(3):643–653.

Arnold, K. and Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. *Current Biology*, 18(5).

Arnold, K. and Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or idiomatic expressions? *Brain & Language*, 120(3):303–309.

Arnold, K. and Zuberbühler, K. (2013). Female putty-nosed monkeys use experimentally altered contextual information to disambiguate the cause of male alarm calls. *PLoS ONE*, 8(6):e65660.

Bala, V. and Goyal, S. (1998). Learning from neighbors. *Review of Economic Studies*, 65:595–621.

Barrett, J. A. (2007). Dynamic partitioning and the conventionality of kinds. *Philosophy of Science*, 74(4):527–546.

Barrett, J. A. (2014). Rule-following and the evolution of basic concepts. *Philosophy of Science*, 81(5):829–839.

Barrett, J. A. and Gabriel, N. (2022). Reinforcement with iterative punishment. *Journal of Experimental & Theoretical Artificial Intelligence*.

Barrett, J. A., Skyrms, B., and Cochran, C. (2018). Hierarchical models for the evolution of compositional language. *IMBS Technical Report*, MBS(18).

Barrett, J. A., Skyrms, B., and Cochran, C. (2019). On the evolution of compositional language. *Philosophy of Science*.

Bond, A. B., Kamil, A. C., and Balda, R. P. (2003). Social complexity and transitive inference in corvids. *Animal Behaviour*, 65(3):479–487.

Borenstein, M., Hedges, L. V., Higgins, J. P., and Rothstein, H. R. (2021). *Introduction to meta-analysis*. Wiley.

Bryant, P. E. and Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature*, 232:456–458.

Cao, R. (2012). A teleosemantic approach to information in the brain. *Biology & Philosophy*, 27:49–71.

Chassy, P. (2013). The role of memory templates in experts' strategic thinking. *Journal of Psychology Research*, 3(5):276–289.

Christiansen, M. H. and Chater, N. (2016). *Creating language integrating evolution, acquisition, and processing*. The MIT Press.

Christiansen, M. H. and Chater, N. (2022). *Language game: How improvisation created language and changed the world*. Basic Books.

Crump, J. (1970). *Chan-Kuo ts'e*, pages 120–122, 377–378. Clarendon.

Davis, H. (1992). Transitive inference in rats (rattus norvegicus). *Journal of Comparative Psychology*, 106(4):342–349.

de Groot, A. D. (1978). *Thought and choice in chess*. Mouton, 2nd edition.

De Lillo, C., Floreano, D., and Antinucci, F. (2001). Transitive choices by a simple, fully connected, backpropogation neural network: Implications for the comparative study of transitive inference. *Animal Cognition*, 4(1):61–68.

De Marzo, G. and Servedio, V. D. (2023). Quantifying the complexity and similarity of chess openings using online chess community data. *Scientific Reports*, 13.

de Oliveira, J. F., Marques-Neto, H. T., and Karsai, M. (2022). Measuring the effects of repeated and diversified influence mechanism for information adoption on twitter. *Social Network Analysis and Mining*, 12(16).

Dechêne, A., Stahl, C., Hansen, J., and Wänke, M. (2009). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14(2):238–257.

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., and Cabanis, E. A. (2007). Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130(5):1432–1441.

Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60.

Fazio, L., Rand, D. G., and Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychon Bull Rev*, 26:1705–1710.

Fazio, L. K. (2020). Repetition increases perceived truth even for known falsehoods. *Collabra: Psychology*, 6(1). 38.

Frank, M. J., Rudy, J. W., Levy, W. B., and O'Reilly, R. C. (2005). When logic fails: Implicit transitive inference in humans. *Memory & Cognition*, 33(4):742–750.

Franke, M. (2015). The evolution of compositionality in signaling games. *Journal of Logic, Language and Information*, 25(3-4):355–377.

Gillan, D. J. (1981). Reasoning in the chimpanzee: Ii. transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes*, 7(2):150–164.

Gobet, F., de Voogt, A., and Retschitzki, J. (2004). *Moves in mind: The Psychology of Board Games*. Psychology Press.

Goldin, P. R. (2005). The theme of the primacy of the situation in classical chinese philosophy and rhetoric. *Asia Major*, 18(2):1–25.

Gould, S. J. and Lewontin, R. C. (1979). The spandrels of san marco and the panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161):581–598.

Grosenick, L., Clement, T. S., and Fernald, R. D. (2007). Fish can infer social rank by observation alone. *Nature*, 445(7126):429–432.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002a). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598):1569–1579.

Hauser, M. D., Newport, E. L., and Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, 78(3):B53–B64.

Hauser, M. D., Weiss, D., and Marcus, G. (2002b). Rule learning by cotton-top tamarins. *Cognition*, 86:B15–B22.

Holman, B. and Bruner, J. (2017). Experimentation by industrial selection. *Philosophy of Science*, 84(5):1008–1019.

Horne, B. D. and Adali, S. (2017). This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. *The Workshops of the Eleventh International AAAI Conference on Web and Social Media AAAI Technical Report*, WS(17):759–766.

Janiak, A. (2022). Kant's views on space and time.

Kiernan, V. (2003). Diffusion of news about research. *Science Communication*, 25(1):3–13.

Kummerfeld, E. and Zollman, K. J. (2015). Conservatism and the scientific state of nature. *The British Journal for the Philosophy of Science*, 67(4):1057–1076.

LaCroix, T., Geil, A., and O'Connor, C. (2021). The dynamics of retraction in epistemic networks. *Philosophy of Science*, 88(3):415–438.

Lee, M. D. and Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Wiley-Blackwell.

Liang, X., Su, L. Y.-F., Yeo, S. K., Scheufele, D. A., Brossard, D., Xenos, M., Nealey, P., and Corley, E. A. (2014). Building buzz. *Journalism & Mass Communication Quarterly*, 91(4):772–791.

Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.

Millikan, R. G. (1984). *Language, thought and other biological categories: New Foundations for Realism*. Mit Press.

Millikan, R. G. (1995). Pushmi-pullyu representations. *Philosophical Perspectives*, 9:185–200.

Montoya, R. M., Horton, R. S., Vevea, J. L., Citkowicz, M., and Lauber, E. A. (2017). A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking. *Psychological Bulletin*, 143(5):459–498.

Naeser, M., Helm-Estabrooks, N., Haas, G., Auerbach, S., and Srinivasan, M. (1987). Relationship between lesion extent in 'wernicke's area' on computed tomographic scan and predicting recovery of comprehension in wernicke's aphasia. *Archives of neurology*, 44:73–82.

Neiworth, J. J., London, J. M., Flynn, M. J., Rupert, D. D., Alldritt, O., and Hyde, C. (2017). Artificial grammar learning in tamarins (saguinus oedipus) in varying stimulus contexts. *Journal of Comparative Psychology*, 131(2):128–138.

Nimzovich, A. (1964). *My system, a treatise on chess*. D. McKay Co.

Oaksford, M. and Chater, N. (2012). *Cognition and conditionals: probability and logic in human thinking*. Oxford University Press.

O'Connor, C. and Gabriel, N. (forthcomming 2023). Can confirmation bias improve group learning?

O'Connor, C. and Weatherall, J. (2018). Scientific polarization. *European Journal for Philosophy of Science*, 8(3):855–875.

of Health, U. D. and Services, H. (2012). Case summaries. *Office of Research Integrity Newsletter*, 20(4):14–17.

Pagin, P. and Westerståhl, D. (2010a). Compositionality i: Definitions and variants. *Philosophy Compass*, 5(3):250–264.

Pagin, P. and Westerståhl, D. (2010b). Compositionality ii: Arguments and problems. *Philosophy Compass*, 5(3):265–282.

Pennycook, G., Cannon, T. D., and Rand, D. G. (2017). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12):1865–1880.

Piedrahita, P., Borge-Holthoefer, J., Moreno, Y., and González-Bailón, S. (2018). The contagion effects of repeated activation in social networks. *Social Networks*, 54:326–335.

Planer, R. J. and Sterelny, K. (2021). *From signal to symbol the evolution of language*. The MIT Press.

Ransom, K., Perfors, A., and Stephens, R. G. (2021). Social meta-inference and the evidentiary value of consensus. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43.

Rosenstock, S., Bruner, J., and O'Connor, C. (2017). In epistemic networks, is less really more? *Philosophy of Science*, 84(2):234–252.

Saffran, J., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., and Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, 107(2):479–500.

Schlenker, P., Chemla, E., Arnold, K., and Zuberbühler, K. (2016a). Pyow-hack revisited: Two analyses of putty-nosed monkey alarm calls. *Lingua*, 171:1–23.

Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinović, D., Arnold, K., Cäsar, C., Keenan, S., and et al. (2016b). Formal monkey linguistics: The debate. *Theoretical Linguistics*, 42(1-2).

Scott-Phillips, T. C. and Blythe, R. A. (2013). Why is combinatorial communication rare in the natural world, and why is language an exception to this trend? *Journal of The Royal Society Interface*, 10(88):20130520.

Siemann, M. and Delius, J. D. (1998). Algebraic learning and neural network models for transitive and non-transitive responding. *European Journal of Cognitive Psychology*, 10(3):307–334.

Skyrms, B. (2008). Signals. *Philosophy of Science*, 75(5):489–500.

Skyrms, B. (2010). *Signals evolution, learning, and information*. Oxford University Press.

Skyrms, B. and Barrett, J. A. (2019). Propositional content in signals. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 74(C):34–39.

Smaldino, P. E., Moser, C. J., Velilla, A. P., and Werling, M. (forthcomming 2023). Maintaining transient diversity is a general principle for improving collective problem solving. *Perspectives on Psychological Science*.

Smolensky, P., Goldrick, M., and Mathis, D. (2013). Optimization and quantization in gradient symbol systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science*, 38(6):1102–1138.

Soltis, A. (2007). *Transpo tricks in chess*. Batsford.

Song, F., Parekh, S., Hooper, L., Loke, Y., Ryder, J., Sutton, A., Hing, C., Kwok, C., Pang, C., and Harvey, I. (2010). Dissemination and publication of research findings: An updated review of related biases. *Health Technology Assessment*, 14(8).

Steinert-Threlkeld, S. (2016a). Compositional signaling in a complex world. *Journal of Logic, Language and Information*, 25(3-4):379–397.

Steinert-Threlkeld, S. (2016b). Compositionality and competition in monkey alert calls. *Theoretical Linguistics*, 42(1-2):159–171.

Steinert-Threlkeld, S. (2020). Toward the emergence of nontrivial compositionality. *Philosophy of Science*, 87(5):897–909.

Sturman, M. (1996). Beware the bishop pair. *ICGA Journal*, 19(2):83–93.

Sutton, R. S., Bach, F., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press Ltd.

Szabó, Z. G. (2020). Compositionality. *Stanford Encyclopedia of Philosophy*.

Tandoc, E. C. (2019). The facts of fake news: A research review. *Sociology Compass*, 13.

van der Linden, S. (2022). Misinformation: Susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, 28:460–467.

von Fersen, L., Wynne, C. D., Delius, J. D., and Staddon, J. E. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 17(3):334–341.

Warren, D. (1998). Kant and the apriority of space. *Philosophical Review*, 107(2):179–224.

Wu, J. (2022). Epistemic advantage on the margin: A network standpoint epistemology. *Philosophy and Phenomenological Research*, n/a(n/a).

Zollman, K. (2007). The communication structure of epistemic communities. *Philosophy of science*, 74(5):574–587.

Zollman, K. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72(1):17–35.