Simulating Infant's Decision Tradeoffs Between Attachment and Exploration Using Reinforcement Learning Models

Xi Jia Zhou, Chris Doyle, Michael C. Frank, Nick Haber
Manuscript for 2024 CogSci Submission

Abstract

Infants must balance interacting with a caregiver and navigating the environment. This decision involves trading off proximity-seeking with the caregiver and reward-seeking in the surroundings, driven by infants' needs of attachment and exploration. Previous research in exploration neglected attachment – an important factor in infants' learning and exploration. Meanwhile, traditional attachment literature was qualitative and lacked generative models for how infants' attachment styles, together with their characteristic exploration behaviors, are learned through experience. Here we consider both attachment and exploration as adaptive reward-seeking strategies given different frequencies and scales of rewards from both caregiver and environment, including negative rewards due to distressing encounters. We implemented a simple environment with three different types of parents, and used Reinforcement Learning (RL) as the framework to implement an infant agent that learned to choose between staying with the caregiver and exploring. We were able to generate characteristic behaviors for secure and insecure-avoidant infants, but not the insecure-resistant infants. This study is a proof-of-concept that RL models can be used to simulate infants' exploratory behavior and learning in forms of attachment.

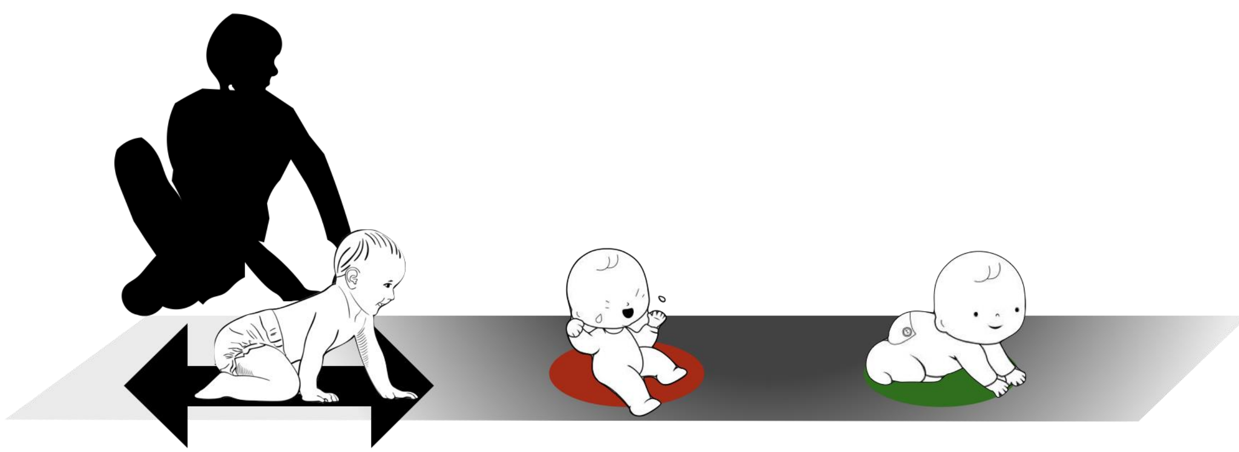


Figure 1: Problem setting. The environment is a 1-dimensional grid that contains an infant agent and a caregiver. The infant agent can move along the grid, while the caregiver is fixed at location 0. The infant agent can stay with the caregiver or explore the environment. The infant agent can become upset (red dot featuring the upset state) or become happy (green dot featuring their happy state).

Introduction

Attachment and exploration are two fundamental behaviors that infants engage in to fulfill their basic needs of optimal predictability, acceptance, and competence – seeing how the consequences of their actions bring either reward or distress from interacting with the physical world and their caregivers, and building skills to excel in both tasks (Dweck, C. S., 2017). Although we know much about infants' exploration and attachment in isolation, we have not studied them together in the same experiment, or have understood how they interact with one another. For example, can we simulate infants' various attachment styles that show different proximity-seeking and exploration patterns within the same environment?

Attachment and exploration are both important research topics that were typically studied separately using different frameworks, each with unique limitations. For attachment, lots of prior research from observational studies have shown the importance of attachment-development in infants which impacts their cognitive development in childhood, such as language acquisition and the development of Theory of Mind (Meins, E., 2013). In adulthood, the impacts of attachment-development are also observed with the quality of their other significant social relationships, such as friendship (Franco, 2022) and romantic relationships (Levine & Heller, 2010). A recent meta-analysis from Madigan et al. (2023) showed that only half of the global population tested from previous Strange Situation Procedure (SSP) in 20,000 parent-child dyads have secure attachment relationships, with the rest being anxious-insecurely attached (10%), avoidant-insecurely attached (15%), or disorganized-insecurely attached (25%). Although we have learned a lot about different attachment styles with their equivalent behaviors and how to classify these main categories of attachment styles from standard measurements, such as the SSP for infants and children and surveys like Attachment Q-sort for teenagers and adults (Farnfield & Holmes, 2014), we lack crucial understandings on how the 'internal world model' (Bowlby, 2018; Johnson et al., 2010) that governs different attachment styles was developed, represented, and deployed in navigating the social world, quantitatively and computationally. For inquiries on exploration, on the other hand, many empirical experiments were conducted in the Cognitive Science literature, and computational models were run in the Artificial Intelligence (AI) community to both study human exploration and advance the development of AI's exploration algorithms (See Appendix's Related Work section for histories of this tradition). Despite producing fruitful results, these lines of research did not capture the realistic and rich environments of infants in which their decisions of explorations are heavily influenced by caregivers (Du et al, 2023; Dahmani et al, 2023). These works typically studied or simulated infants' decision-making alone, omitting attachment, another important basic learning mechanism and behavior in early childhood, and its impact on learning and exploration.

In fact, attachment and exploration are two deeply intertwined behaviors. For attachment, the SSP assessments of infants' attachment styles are heavily dependent on coding infants' exploratory behaviors, such as their exploratory locomotion, exploratory manipulation of physical items, and visual exploration (Ainsworth et al., 2015, p.331-p.332). Moreover, the interpretation of characteristic behaviors and the classification of different attachment styles are essentially defined equally by infants' exploratory patterns between the caregiver and the physical environment, as well as their comfort-seeking behaviors towards the returning parent in SSP. According to Mary Ainsworth (2015), Group B, the secure babies, explore most freely according to their true interests, with "nonanxious exploration" (p.312), and explore the longest (p.304). Group A, the avoidant insecure babies, explore to avoid distress interacting with the caregiver and explore in a way that is "devoid of the true interest" (p.312). Group C, the anxious insecure babies, explore the least amount of time, or have no exploration, due to fear of inaccessibility and nonresponsiveness of the caregiver. They can show 'clinging and other manifestations of contact-maintaining behavior" (p.308). On the other hand, for exploration, recent research has shown that children's learning and exploration could be significantly impacted by caregivers, sometimes distorted in suboptimal ways. In one study, children who learned about an aversive stimulus with the presence of a caregiver not only developed more attentional bias towards the negatively rewarding stimulus, but also actively sought out and chose to explore it. Contrastingly, when children were alone in the learning phase, they would avoid the aversive stimulus (Tottenham et al., 2019). In another old study, researchers found that infants of four- and five-month olds favored strangers who exhibited similar contingent responses in terms of smiles and vocalizations as their mothers, more than strangers who were more responsive to them (Bigelow, 1998). These findings show that young humans' learning and exploration significantly interacts with the presence of the caregiver and their previous interaction histories. Therefore, it is important to study infants' learning and exploration in tandem with attachment. Analyzing both behaviors under the same computational framework could generate novel insights on both attachment and learning.

Here we proposed using Reinforcement Learning (RL) as the suitable framework to study infants' attachment and exploration together. Much prior research has used RL models to simulate human-like explorations (Also see Related Work). For example, Haber et al., (2018) deployed an RL infant agent that used curiosity-driven intrinsic motivation, a neural network that learned a world model while another neural network simultaneously challenged it. The infant agent developed complex exploratory behaviors such as ego motion prediction, object attention and object gathering. Recent research has also examined caregiver's presence and its impact on artificial infant agent's curiosity using RL. Doyle et al (2023) looked at how a contingent artificial parent supported an artificial infant's exploration in the virtual environment. For studying and simulating attachment, we argue that RL is not only a suitable choice, but also provides particular advantages. First, the 'internal working model' concept in attachment literature mirrors the 'world model' learning framework in RL. Attachment theory proposes an individual's internal world model, which is shaped by early attachment experiences, influences their expectations and actions in social interactions. In parallel, the world model in RL, which is formed from previous interaction histories, represents the agent's understanding and predictions about its environment, thereby guiding the agent's actions to achieve desired outcomes. Next, RL also provides unique advantages to study attachment. For example, the real-world parent-child interaction history over time can be challenging to capture and intervene on in observational research, while RL offers an unique opportunity to examine the learning history and even modify the dyad's interactions. Relatedly, we could generate different kinds of caregivers with different probabilities of reward or punishment, as well as changing the environment where the artificial infant agent learns and explores. For instance, we could ask questions like: How does changing the environment setting to be more risky or rewarding impact artificial infant's learning of exploration and attachment? How would the temperament of the artificial infant, being more risk-taking or easily scared, determine its exploration and attachment patterns? Lastly, RL is a simplified, yet powerful and flexible tool. We could add complexities into the model to see how each component adds to the interaction and agents' outcome behaviors.

In this work, we plan to make the following contributions:

1. We introduce an RL modeling approach for testing the learning of attachment and exploration from environmental rewards in an artificial infant.
2. We create a 1-dimensional grid environment that features distances from the caregiver which represent the real world where the the further away the infant is from the parent,

the higher exploration reward and higher probability of distress, an artificial infant who can crawl to explore or stay close to the parent, and different parents with probabilities of rewarding, ignoring, or punishing the infant.

3. We generate characteristic patterns of attachment and exploration in our artificial infants, as described by the literature.
4. We demonstrate that changing reward structures of the environment and artificial infant's self-regulation could change their attachment and exploration behaviors.

Method

Inspired by the attachment literature that describes infants with different attachment styles as balancing the caregiver on one end and exploring the environment on the other, we recreated such an environment with a 1-dimensional grid world*.* The world consists of two agents: an infant agent and a caregiver agent. The infant agent has two internal states: happy and upset, which the caregiver could help regulate. The infant agent gets rewarded (positively or negatively, in cases of it being upset) by both the caregiver and the environment, based on the infant's location and its internal state.

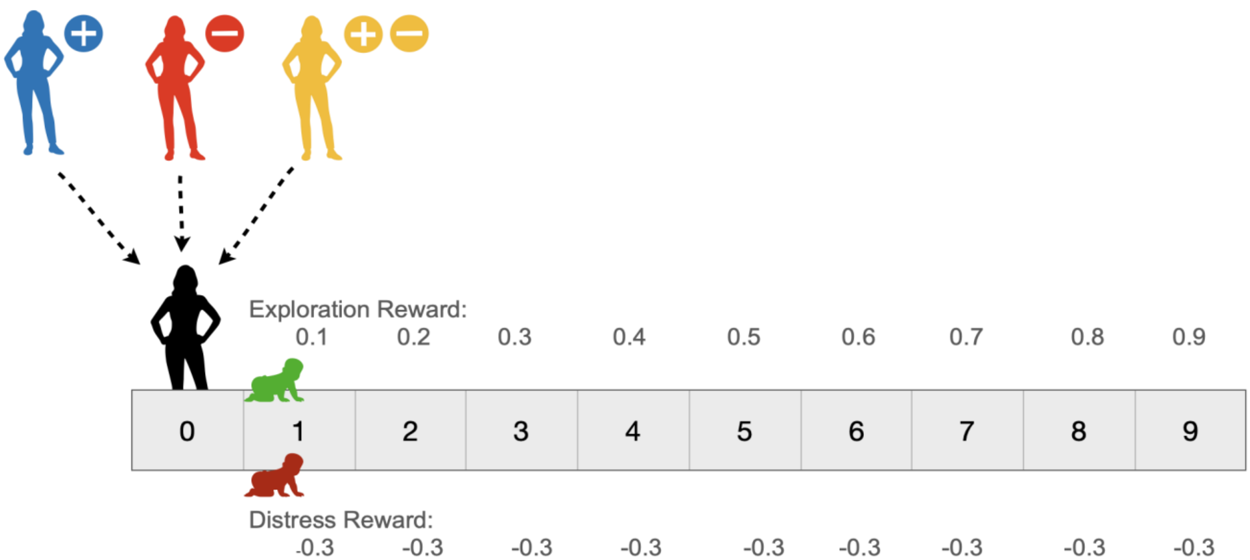


Figure 2. The environment setting. The infant agent explores the environment with fixed reward and distress values at each cell of the grid. The exploration rewards range from 0.1 to 0.9, with a 0.1 increment from cell 1 to cell 9. The distress rewards are fixed with -0.3 at each cell. The infant agent gets exploration rewards when it is exploring the environment and is in a happy state; it gets the distress rewards when it is exploring but is in an upset state. We also implement three types of caregivers (consistently rewarding caregiver, always punishing caregiver, and inconsistently rewarding caregiver) , which are hypothesized from the literature to cause children's different attachment styles.

Environment and Environment Dynamics

Our environment consists of a simple 1-d grid world with 10 cells, ranging from cell 0 to cell 9. Both the infant agent and the caregiver are initialized at cell 0, where the caregiver is also fixed at. The infant agent has an action space of three: it can decide to turn left, turn right, or stay still. The caregiver can reward, ignore, or punish the infant agent.

The infant agent is initiated as happy. When the infant agent is exploring, it could change states as it encounters distress. The probability of the infant agent transitioning to being upset from being happy ranges from 0.01 to 0.09, with an increment of 0.01, from grid 0 to 9. The probability of the infant agent transitioning to being happy from being upset is fixed at 0.04 across all grids. The rationale of this design is that the infant agent has a probability of self-regulation to explore in the world and receive exploration rewards. However, if the infant agent becomes upset, it could travel back to the caregiver for comfort. The infant agent receives exploration rewards when it is exploring in cell 1 to cell 9 and is in a happy state; it receives distress rewards when it is exploring in cell 1 to cell 9 and is in an upset state. The exploration rewards are fixed at each cell of the grid, ranging from 0.1 to 0.9 with a 0.1 increment from cell 1 to cell 9. The distress rewards, however, are fixed at -0.3 at all cells (See Figure 2). The overall rationale behind these values is that the further the infant agent travels away from the caregiver, the higher the exploration reward and the more costly it becomes when the infant agent is in distress, which we think matches the reality. It takes longer for the infant agent to travel back to the caregiver to avoid incrementing negative rewards. We also try to find reasonable values for both environment and infant parameters to be able to reproduce desired attachments.

Here we summarize what happens when the infant agent is at grid 0 with the caregiver. If the caregiver comforts the infant agent, the infant agent becomes happy. But, if the caregiver punishes the infant agent, it becomes upset. Nothing happens when the caregiver ignores the infant agent, and nothing happens when the infant agent is happy. But if the infant agent is upset, due to either environment distress or parent punishment from the earlier step, it receives the reward from the caregiver – a positive reward of 1 when the caregiver comforts the infant agent, 0 when the caregiver ignores the infant agent, and a negative reward of 1 when the caregiver punishes the infant agent.

Caregiver Implementations

We implemented three different kinds of caregivers that were hypothesized from the attachment literature to cause differing attachment styles in infants (Isabella, 1993; Ainsworth et al., 2015). Consistently rewarding caregivers (CaregiverC) are associated with securely attached babies. CaregiverC always rewards the infant agent with a positive reward of 1. Always punishing caregivers (CaregiverP) are associated with avoidant insecurely attached babies. CaregiverP always punishes the infant agent with a negative reward of 1. Inconsistently rewarding caregivers (CaregiverIC) are associated with resistant insecurely attached babies. CaregiverIC comforts the infant agent with a probability of 0.5 with a positive reward of 1, and ignores the infant agent with a probability of 0.5 with a reward of 0.

Infant Agent Implementations

The infant agent consists of an epsilon greedy Q-learning algorithm for learning and action selection, a standard approach in RL that has garnered support from developmental literature (Gopnik, 2020; Giron et al., 2023). We deploy a standardized Q-learning algorithm from the RL literature (Sutton & Barto, 2018). The Q-learning table consists of three dimensions: position space (grid length of 10), state space (agent state of 2), and action space (agent action of 3). The values in the table are initialized as 0. From each step the infant agent interacts with the environment, including when the infant agent is at grid 0 with the caregiver, it collects a reward and updates the estimated value of the expected reward, which maps to the particular position (e.g, grid 1), infant state (e.g, happy), and action it took (e.g stay), for future action selection. The infant agent uses a standard epsilon greedy strategy for choosing the current action. First, we draw a random number from a uniform distribution from 0 to 1. Then, if the random number is smaller or equal to epsilon (a small value), the infant agent randomly selects an action from the action space (left, right, stay). Otherwise, it chooses the action (given the location and the infant state) that has the highest expected value from the Q-table. With time, epsilon decreases from close to 1 to close to 0. Thereby, the infant agent experiences smaller probabilities of selecting random actions.

Experiment & Preliminary Results

We run the environment for 10,000 episodes, with 500 steps (the amount of actions the infant agent could take) in each episode. We log infants' positions taken with all steps in each episode, average positions taken across all episodes (the frequency of each position taken averaged), and total episode return (total rewards the infant agent collects) across all episodes.

From preliminary results, we were able to reproduce secure and insecure avoidant attachment styles with equivalent exploration patterns described by the literature. For example, in Figure 3 and Figure 6, we can see that the infant agent with CaregiverC, the consistent rewarding caregiver, learns to stay on average at grid 5 and 6, which replicates the literature's claim that the secure babies are able to balance staying close with the caregiver and exploring the environment. In contrast, with time, the infant agent with CaregiverP, the consistently punishing caregiver, learns to stay on average at grid 7 or 8, while avoiding the caregiver. This behavior also matches the insecure avoidant children from the literature. In Figure 4, we see that for the last 2500 steps (the last 5 episodes), the infant agent with CaregiverC travels back and forth between the caregiver and the far end of the grid. From the infant state, It is likely that they can travel back to the caregiver when they experience distress from the environment. In comparison, the infant agent with CaregiverP stayed away from the caregiver on grid 8 and showed distress. This also matches the literature in that infants with insecure avoidant attachments still experience distress, even while avoiding the caregiver.

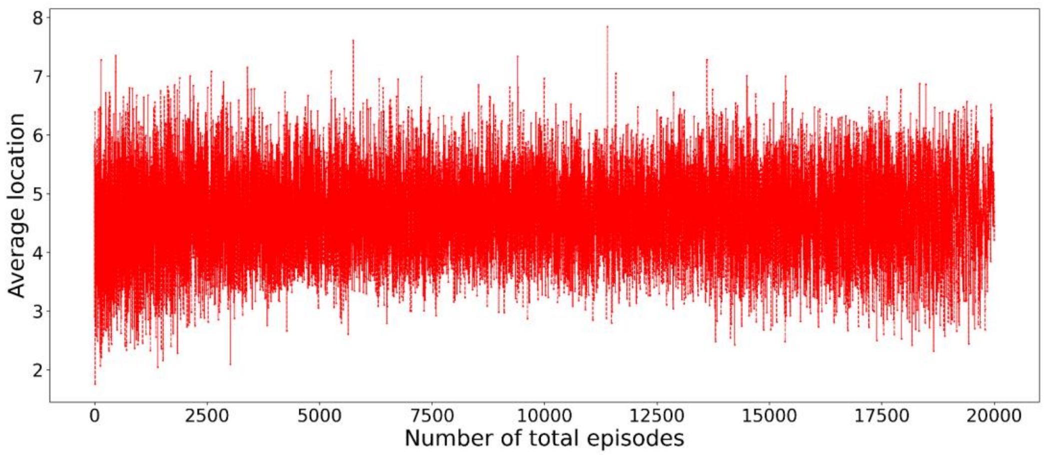
However, we failed to regenerate insecure resistant attachment infants' behaviors, with our simulation results resembling the patterns of secure attachment infants, which is not what we expected. From Figure 3, Figure 4, and Figure 6, you can see that the first and third plots look alike. Figure 5 showed the total episode returns of infant agents.

(Figures are shown in the next pages)

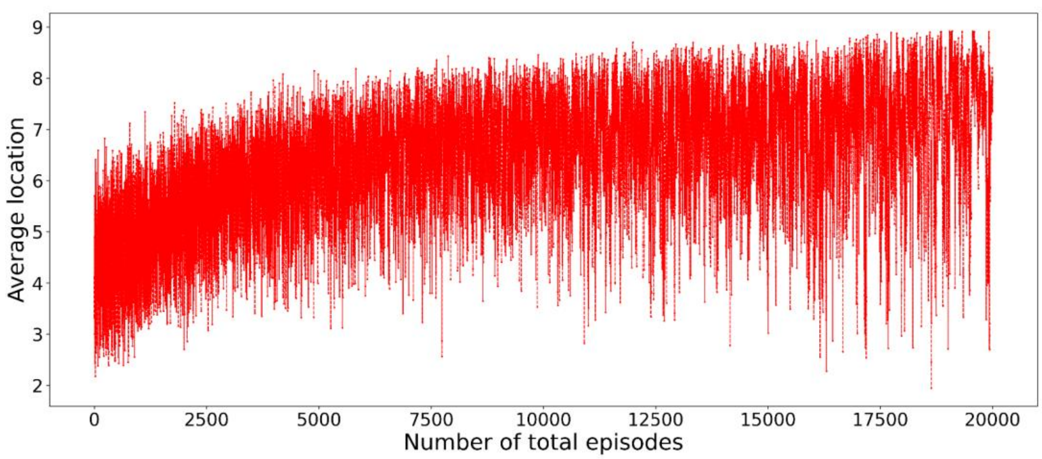
Figure 3. Average Position Across Episodes. Ordered in CaregiverC, CaregiverP, and CaregiverIC. We averaged the locations the infant agent took across 500 steps in each episode, and showed it on they-axis across the 20,000 episodes we ran and displayed on the x-axis. The infant agent with CaregiverC on average stays at grid 5 in the end; the infant agent with CaregiverP on average stays at grid 7 or 8 in the end; the infant agent with CaregiverIC on average stays at grid 4 or 5 in the end.

(Plots are shown in the next page)

CaregiverC: Average Position Across Episodes



CaregiverP: Average Position Across Episodes



CaregiverIC: Average Position Across Episodes

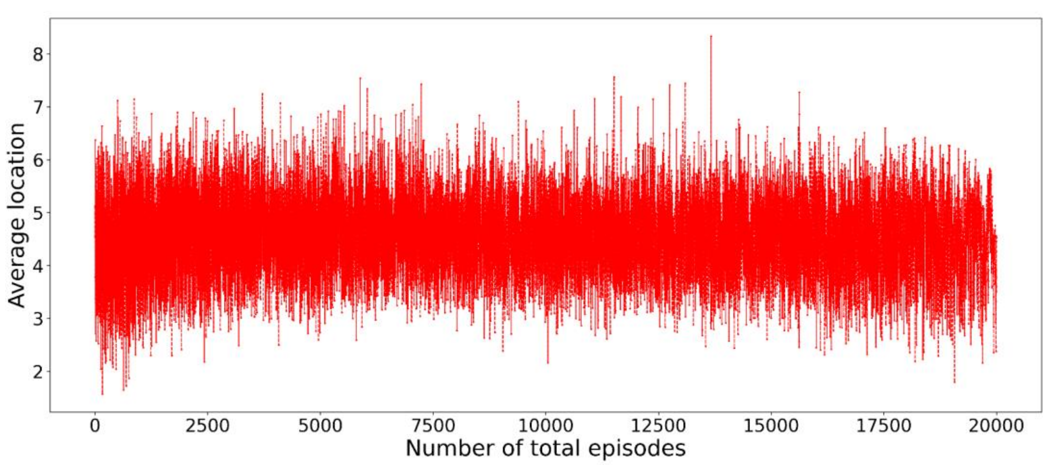


Figure 4. Child Location in Last 2500 Steps (Last 5 Episodes). Ordered in CaregiverC (top), CaregiverP (bottom left), and CaregiverIC (bottom right). The infant agent with CaregiverC travels back and forth between the caregiver and cell 6. From the infant state, they are equally split in happy and distress states when away from the caregiver. In comparison, the infant agent with CaregiverP stays away from the caregiver on cell 8, and presents as predominantly distressed. The Infant agent with CaregiverIC resembles the patterns of secure attachment infants with CaregiverC.

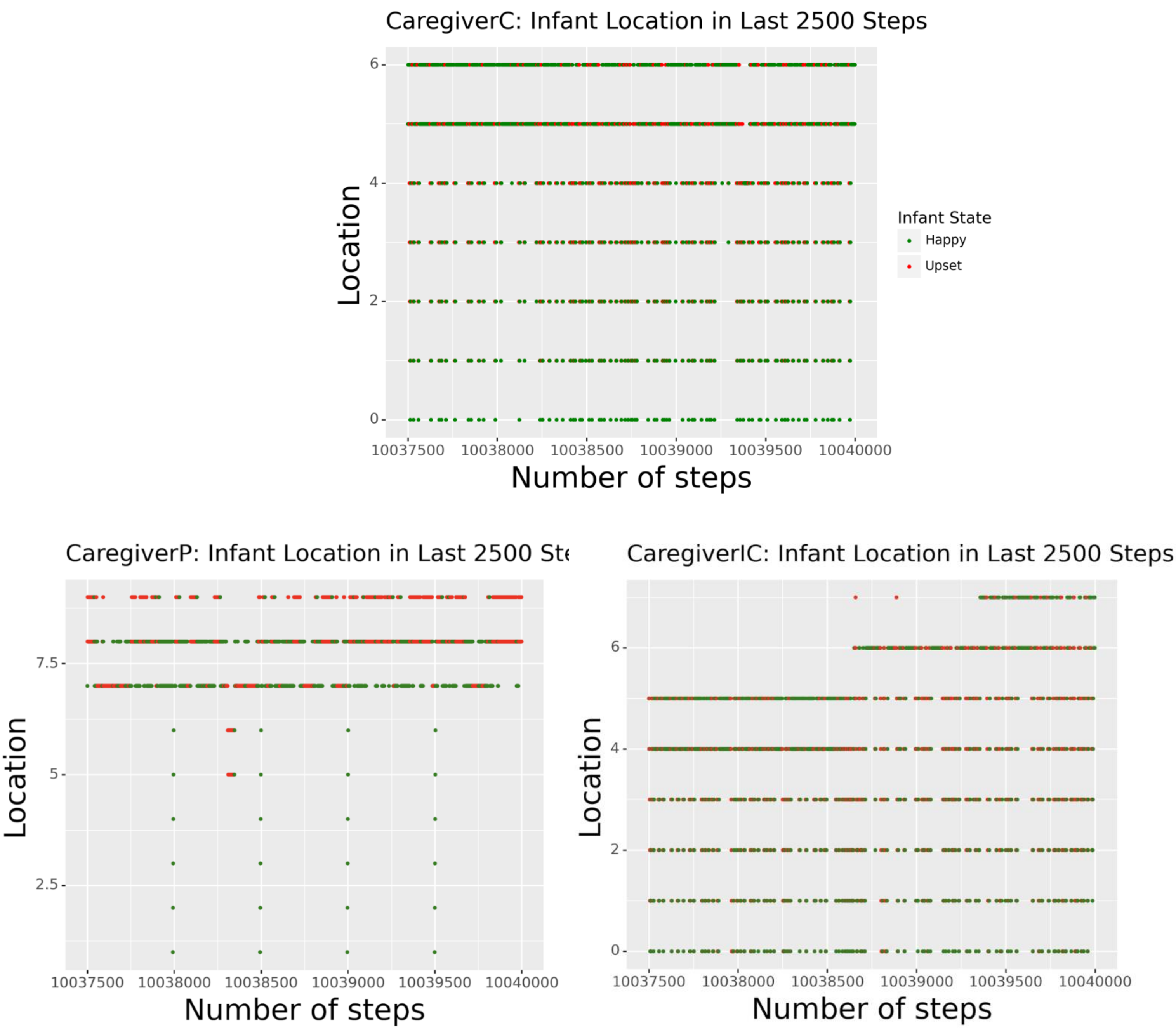


Figure 5. Total Episode Return. Ordered in CaregiverC (top), CaregiverP (bottom left), and CaregiverIC (bottom right). We calculate the total reward the infant agent collected for each episode (displayed on the y-axis) across 20,000 episodes (displayed on the x-axis). In the end, the infant agent with CaregiverC collects on average 150 rewards; the infant agent with CaregiverP collects on average 0 rewards; and the infant agent with CaregiverIC collects on average 150 rewards, resembling the patterns of secure attachment infants with CaregiverC.

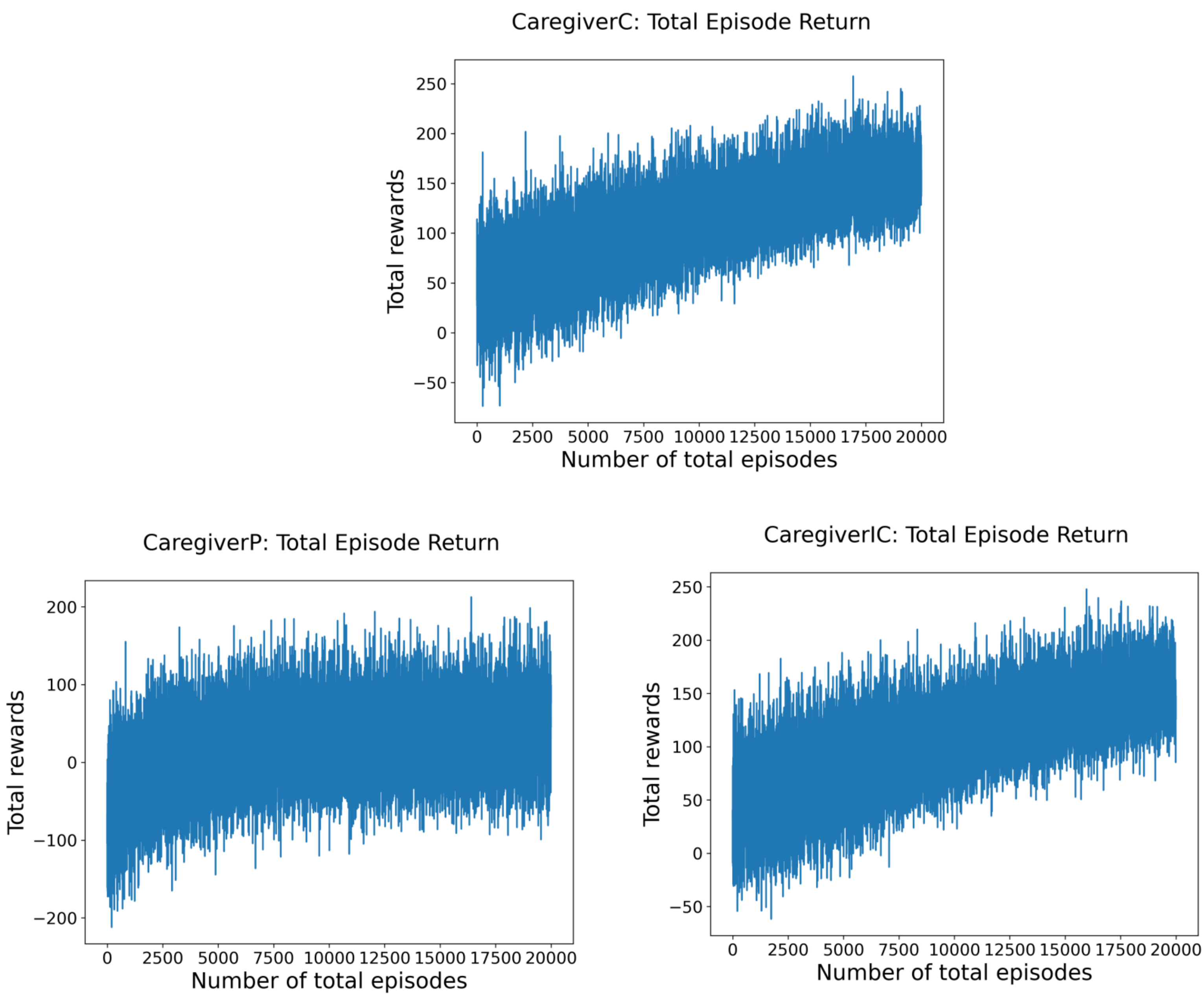
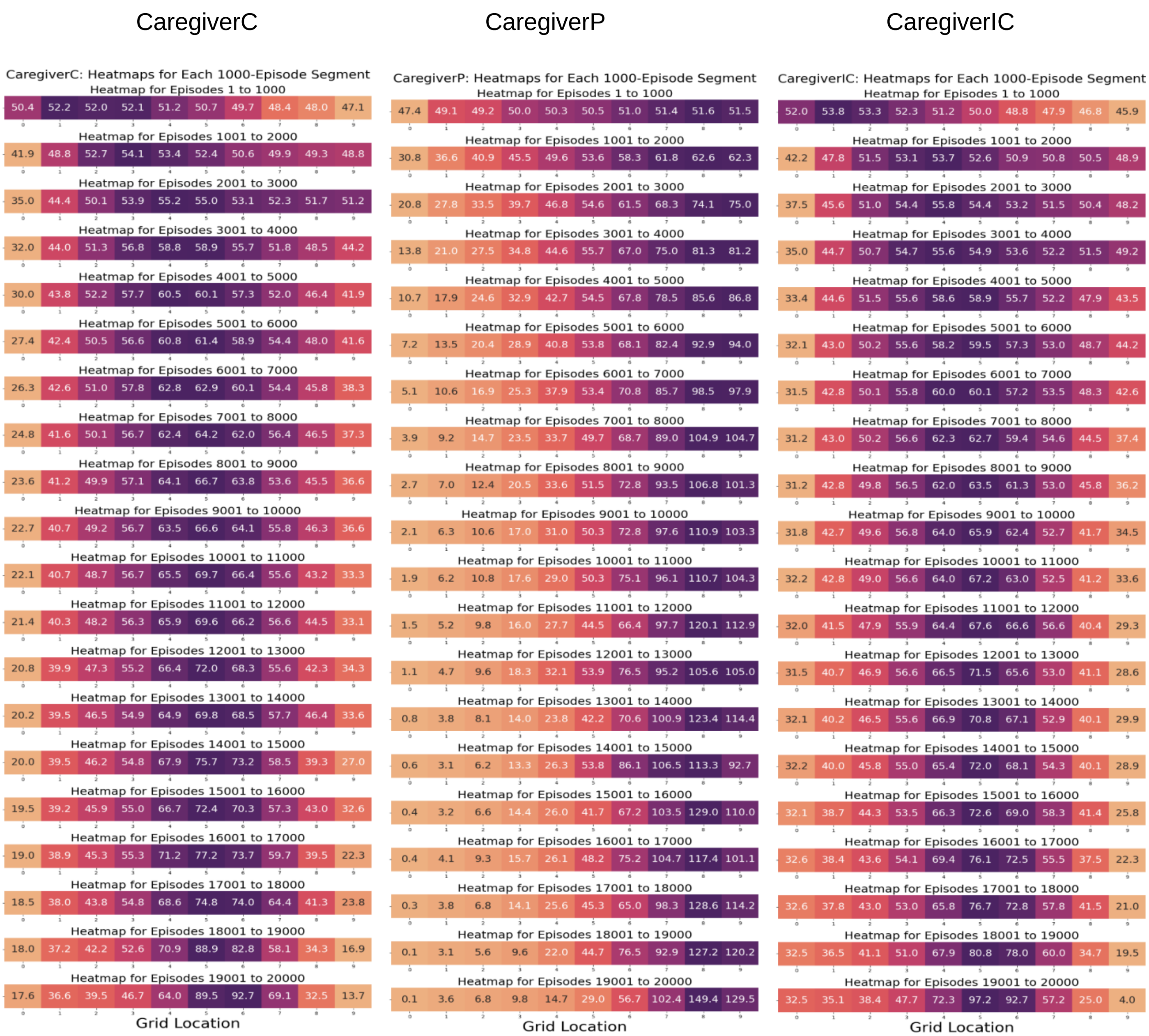


Figure 6. Location Heatmaps. Ordered in CaregiverC (left), CaregiverP (middle), and CaregiverIC (right). We plot the locations the infant agent stayed at across 20,000 episodes in a heatmap with 1000 episodes as the unit for each heatmap. In the end,the infant agent with CaregiverC stays most frequently at grid 5 and 6; the infant agent with CaregiverP stays most frequently at grid 8, avoiding the caregiver; and the infant agent with CaregiverP stays most frequently at grid 5 and 6.



Discussions and Future Work

We find that secure attachment and insecure avoidant attachment styles can arise from simple reinforcement learning set-ups assuming infants are reward maximizers in their surroundings. Our results show that without other complexities, having different types of parents ,who would reward or punish infants on different probabilities and scales, could account for secure and insecure avoidant attachment and exploration patterns.

We failed to reproduce insecure resistant attachment infants' patterns. This means that our current model varying only the caregiver's probability of reward (specifically, less consistently rewarding) alone cannot explain the emergence of insecure resistant attachment. In fact, previous literature has suggested that parental sensitivity could only account for part of insecure attachments' development, while other important factors account for learning and development, such as the child's temperament and emotion regulation ability (Calkins & Fox,1992), specifically due to their cortisol stress response (Houbrechts et al., 2021). The way we implemented the infant's reward learning could also be incorrect in accounting for insecure resistant attachments. For example, insecure resistant infants could have some intrinsic reward functions that make them seek inconsistency and thus cling to the caregiver. Alternatively, we may need to design different environments with varying reward structures. Luckily, our model would be flexible enough to further test these possibilities.

Future work could extend our models in multiple ways. First, the current 'internal world model' is represented simply in a Q-learning table that maps the infant state, infant location, and expected future rewards estimated from previous reward histories to actions. In the future, researchers could incorporate more complex world models, such as ones with binary, mutual theory of mind abilities for both the caregiver and infant. Next, current work looks at the intersection between attachment and exploration, without considering intrinsic motivations in the agent. Future research could build on intrinsic motivations in the infant agent, and explore the interplays between intrinsic and extrinsic rewards as well as attachment's intersection with curiosity.

Beyond the current study that looks at the intersection of attachment and exploration, a few exciting directions could be further explored. Researchers can ask questions such as: How would the attachment representation of a caregiver's sensitivity and responsiveness in distress situations link to their daily communicative sensitivity and responsiveness, such as in serve and return interactions (Fisher et al. 2016)? Could attention be the mediating factor in between attachment styles and cognitive development in children, and can we study it computationally using RL that looks at the agent's attention (Kuno et al., 2020)? Can we simulate culture in multi-agent RL environments and quantify attachment differences in virtual communities? Can we implement attachment theory in the development of novel AI and robotics, especially in dangerous or highly unpredictable environments that are costly to explore alone? How can we simulate generative agents in an 'Interactive Simulacra', (Park et al., 2023) that represent different people's attachment styles, for attachment coaching and therapy? Can we clarify what is the cut-off point for an individual to shift from insecure to secure attachments? As one of the authors is attached to the study of attachment, this would be a rewarding research journey with exploration and learnings.

References:

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. N. (2015). Patterns of attachment: A psychological study of the strange situation. Psychology Press.

Bigelow, A. E. (1998). Infants' sensitivity to familiar imperfect contingencies in social interaction. Infant Behavior and Development, 21(1), 149-162.

Bowlby, J. (2018). The nature of the child’s tie to his mother 1. In Influential Papers from the 1950s (pp. 222-273). Routledge.

Calkins, S. D., & Fox, N. A. (1992). The relations among infant temperament, security of attachment, and behavioral inhibition at twenty-four months. Child development, 63(6), 1456-1472.

Dahmani, A., Amir, D., Gopnik, A. (2023 November). Caregiver presence influences the explore-exploit tradeoff [Conference Poster #182]. Causality in Minds and Machines: Probing the mathematical and computational foundations.

Doyle, C., Shader, S., Lau, M., Sano, M., Yamins, D. L., & Haber, N. (2023). Developmental Curiosity and Social Interaction in Virtual Agents. arXiv preprint arXiv:2305.13396.

Du, Y., Kosoy, E., Dayan, A., Rufova, M., Abbeel, P., & Gopnik, A. (2023, November). What can AI Learn from Human Exploration? Intrinsically-Motivated Humans and Agents in Open-World Exploration. In NeurIPS 2023 workshop: Information-Theoretic Principles in Cognitive Systems.

Dweck, C. S. (2017). From needs to goals and representations: Foundations for a unified theory of motivation, personality, and development. Psychological review, 124(6), 689.

Farnfield, S., & Holmes, P. (Eds.). (2014). The Routledge handbook of attachment: Assessment (Vol. 4324, p. 9781315770666). Abingdon: Routledge.

Fisher, P. A., Frenkel, T. I., Noll, L. K., Berry, M., & Yockelson, M. (2016). Promoting healthy child development via a two-generation translational neuroscience framework: The Filming Interactions to Nurture Development video coaching program. Child Development Perspectives, 4, 251-256. PMC5603284

Franco, M. G. (2022). Platonic: How the Science of Attachment Can Help You Make--and Keep--Friends. Penguin.

Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2023). Developmental changes in exploration resemble stochastic optimization. Nature Human Behaviour, 1-13.

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. Philosophical Transactions of the Royal Society B, 375(1803), 20190502.

Houbrechts, M., Cuyvers, B., Goossens, L., Bijttebier, P., Bröhl, A. S., Calders, F., ... & Bosmans, G. (2023). Parental support and insecure attachment development: The cortisol stress response as a moderator. Attachment & Human Development, 25(1), 104-116.

Ho, M. K., MacGlashan, J., Littman, M. L., & Cushman, F. (2017). Social is special: A normative framework for teaching with and learning from evaluative feedback. Cognition, 167, 91-106.

Isabella, R. A. (1993). Origins of attachment: Maternal interactive behavior across the first year. Child development, 64(2), 605-621.

Johnson, S. C., Dweck, C. S., Chen, F. S., Stern, H. L., Ok, S. J., & Barth, M. (2010). At the intersection of social and cognitive development: Internal working models of attachment in infancy. Cognitive science, 34(5), 807-825.

Kuno, K., Sano, M., Freitas, J. D., Yamins, D. L. K., & Haber, N. (2020). Towards modeling the developmental variability of human attention. In Proceedings of the International Conference on Learning Representations Workshops.

Levine, A., & Heller, R. (2010). Attached: the new science of adult attachment and how it can help you find--and keep--love. Penguin.

Madigan, S., Fearon, R. M., van IJzendoorn, M. H., Duschinsky, R., Schuengel, C., Bakermans-Kranenburg, M. J., ... & Verhage, M. L. (2023). The first 20,000 strange situation procedures: A meta-analytic review. Psychological Bulletin, 149(1-2), 99.

Meins, E. (2013). Security of attachment and the social development of cognition. Psychology press.

Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023, October). Generative agents: Interactive simulacra of human behavior. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (pp. 1-22).

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

Tottenham, N., Shapiro, M., Flannery, J., Caldera, C., & Sullivan, R. M. (2019). Parental presence switches avoidance to attraction learning in children. Nature human behaviour, 3(10), 1070-1077.

Appendix:

Related Work

1. History and Current Works Studying Learning, Exploration, and Curiosity in Human and AI.

There is a long history of studying learning, exploration and curiosity in humans, and more recently in AI. Here I summarize the relevant literature which I have learned about. It is not an exclusive list of studies, but it is a start to understanding this rich space of inquiry in learning, exploration and curiosity.

<u>a. Research in young children:</u>

Learning and Play. Works by Gopnik et al. (1999) and Schulz (2015) emphasize how young children learn through playful exploration and experimentation, just like scientists, in the way that children test hypotheses, make sense of statistical data (Gweon & Schulz, 2011), and explore the unexpected events (Stahl & Feigenson, 2015) in their surroundings. This research shows that children's learning and play are active, structured, and sophisticated.

Visual Exploration: Infants Preferential Looking. There's a long history of this line of research. Robert Fantz (1964) in his study reveals that infants show selective attention to novel or complex stimuli, while showing decreased attention to familiar ones. This finding demonstrates that infants' innate curiosity drives their preferential attention to visual patterns in their surroundings. Colombo and Mitchell (2009) summarize this history of studying infant's visual habituation.

Locomotion Exploration. Karen Adolph's research looks into how infants and toddlers fine-tune their movements in response to the varying challenges presented by their surroundings, highlighting the adaptive and selective nature of infants' motor exploration (Adolph et al, 1997). Justine Hoch's research continues this line of investigation, showing how environmental affordances support infants' locomotion (Hoch et al., 2019) and that infants' locomotion exploration is both embedded in their own body and environmental constraints and enables new learning opportunities (Adolph & Hoch, 2019).

Curiosity. Kidd et al. (2012) shows that infants' attention is selectively directed towards objects and events that are optimally novel and informative, not too simple (already known) nor too complex (unlearnable). In short, infants' preferential-looking follows an inverse U shape. More related work along this line is summarized by Kidd & Hayden (2015).

<u>b. Research in adults:</u>

In parallel, researchers from psychology and cognitive science have looked into curiosity and learning in adults. Recently, we have learned the importance of perceived novelty, uncertainty, information gain, expected learning progress, and empowerment in human learning and exploration (Du et al., 2023). We also know that adults are curious about things that are learnable, important, or socially preferred (Dubey & Griffiths, 2020; Dubey et al., 2020; Dubey et al., 2021).

<u>c. Research in AI:</u>

(i) History of modeling curiosity in artificial intelligence. The history of modeling curiosity in AI is a captivating journey that reflects the evolving understanding of what drives learning and exploration in both biological and artificial systems. This exploration began with the recognition that curiosity, a key element in human learning, could be modeled and incorporated into AI systems to enhance their learning capabilities.

One of the seminal works in this area is by Oudeyer et al. (2007), who introduced a computational model of curiosity-driven learning. This model, grounded in the concept of intrinsic motivation, posited that AI systems could be designed to seek out novel or surprising information, thereby facilitating autonomous learning. This approach marked a significant shift from traditional AI models that were primarily driven by external rewards. Andrew Barto (2013) further advanced this field by elaborating on the role of intrinsic rewards in reinforcement learning. Barto's research emphasizes the importance of internal drives, such as curiosity, in enabling AI systems to explore their environments more effectively, leading to the development of more robust and adaptable learning algorithms. Most recently, Schmidhuber (2021) has continued to push the boundaries of this field. Schmidhuber has been a prominent figure in this area for decades, advocating for the importance of intrinsic motivation in AI. His latest contributions elaborate on how curiosity and a drive for novelty can be encoded within the learning processes of AI, leading to systems that not only learn more efficiently but are also capable of generating new, creative solutions to complex problems.

Together, these works form the backbone of the current understanding of curiosity in AI, highlighting the shift from reward-driven to intrinsically motivated learning processes. This body of research not only provides insights into the development of more advanced AI systems but also offers a deeper understanding of the fundamental mechanisms of learning and intelligence.

Additional References:

Adolph, K. E., Bertenthal, B. I., Boker, S. M., Goldfield, E. C., & Gibson, E. J. (1997). Learning in the development of infant locomotion. Monographs of the society for research in child development, i-162.

Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. Annual review of psychology, 70, 141-164.

Barto, A. G. (2013). Intrinsic motivation and reinforcement learning. Intrinsically motivated learning in natural and artificial systems, 17-47.

Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. Neurobiology of learning and memory, 92(2), 225-234.

Du, Y., Kosoy, E., Dayan, A., Rufova, M., Abbeel, P., & Gopnik, A. (2023, November). What can AI Learn from Human Exploration? Intrinsically-Motivated Humans and Agents in Open-World Exploration. In NeurIPS 2023 workshop: Information-Theoretic Principles in Cognitive Systems.

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. Psychological Review, 127(3), 455–476.

Dubey, R., Griffiths, T., & Lombrozo, T. (2020). If it’s important, then I’m curious: Increasing perceived usefulness stimulates curiosity.

Dubey, R., Mehta, H., & Lombrozo, T. (2021). Curiosity Is contagious: A social influence intervention to induce curiosity. Cognitive science, 45(2), e12937.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. Science, 146(3644), 668-670.

Hoch, J. E., O'Grady, S. M., & Adolph, K. E. (2019). It's the journey, not the destination: Locomotor exploration in infants. Developmental science, 22(2), e12740.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). The scientist in the crib: Minds, brains, and how children learn. William Morrow & Co.

Gweon, H., & Schulz, L. (2011). 16-month-olds rationally infer causes of failed actions. Science, 332(6037), 1524-1524.

Jürgen Schmidhuber. Artificial Curiosity & Creativity Since 1990-91. 2021, https://people.idsia.ch/~juergen/artificial-curiosity-since-1990.html.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. Neuron, 88(3), 449-460.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. PloS one, 7(5), e36399

Oudeyer, P. Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. IEEE transactions on evolutionary computation, 11(2), 265-286.

Schulz, L. (2015). Infants explore the unexpected. Science, 348(6230), 42-43.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants’ learning and exploration. Science, 348(6230), 91-94.