The Exploration Advantage:

Children's instinct to explore leads them to find information that adults miss.

Emily S. Sumner[a], Amy X. Li[b], Amy Perfors[c], Brett K. Hayes[b], Danielle J. Navarro[b], and Barbara W. Sarnecka[a]

University of California, Irvine[a], University of New South Wales[b], University of Melbourne[c]

To whom correspondence should be addressed. E-mail: sumnere@uci.edu

Abstract

Humans have a long childhood in comparison to all other species. Across disciplines, researchers agree that humans' prolonged immaturity is integral to our unique intelligence. The studies presented here support the hypothesis that human beings' extended childhood pays off in the form of an ability to learn more about changing environments. Across two studies ($n = 213$), children and adults played a game where they chose among four different cartoon monsters yielding different numbers of star rewards. Adults focused on maximizing reward, while children chose to explore longer, even at the cost of earning fewer stars. As a result, adults won significantly more stars than children did. However, in the 'dynamic' version of the task, the rewards given out by the monsters changed halfway through: the monster that had been giving out the fewest stars began giving out the most. Because children continued to explore whereas adults ignored the low-reward monster, children were much more likely than adults to detect the change. This illustrates that while exploration may be costly in the short term, it leads to a more flexible understanding of the world in the long term, particularly when that world is changing.

The Exploration Advantage:

Children's instinct to explore leads them to find information that adults miss.

Compared to other species, humans take a very long time to grow up. Whereas 24-hour-old blue wildebeests can outrun a hyena (Estes & Estes, 1979), and two-year-old wolves are ready to start their own families (Hayssen, Van Tienhoven, & Van Tienhoven, 1993), human children of those ages can barely feed and dress themselves. Humans' extended period of immaturity means that children require a huge investment of caregiving time and resources, this resource debt isn't paid off until around the age of fifty (Sterelny, 2012). This creates an evolutionary puzzle: What payoff could make this long period of neediness and vulnerability worthwhile? In this paper, we present research suggesting that one payoff of extended immaturity is that it allows children to a spend longer time exploring their environment, and this leads to improved learning, particularly when the environment is changing.

Across disciplines, researchers agree that humans' prolonged immaturity is integral to our unique intelligence (Deacon, 1998; Gentner & Goldin-Meadow, 2003; Gopnik et al., 2017; Papagianni & Morse, 2015; Piantadosi & Kidd, 2016; Riede, Johannsen, Högberg, Nowell, & Lombard, 2018; Sterelny, 2012; Tomasello, 2019). We tested a specific hypothesis about the relationship between prolonged childhood and human intelligence. When faced with a new situation, children spend significantly longer exploring the environment than adults do, even at the cost of fewer immediate rewards (Blanco & Sloutsky, 2019; Schulz, Wu, Ruggeri, & Meder, 2019; Sumner, Steyvers, & Sarnecka, 2019). We found that this childhood bias towards exploration, although disadvantageous in the short term or in stable environments, is particularly advantageous when an environment is changing. By exploring more, children detect changes that are missed by adults, leading to increased adaptability.

In two experiments, child and adult participants played a computer game where they could earn points (stars) by clicking on any of four cartoon monsters, each of which gave out a different number of stars (Figure 1A). Children were tested in a museum or school setting on a tablet device, and adults were tested on Amazon Mechanical Turk. Participants were randomly assigned to play either the static or the dynamic version of the game for 80 trials. Participants were not told that the game had two versions, and did not know which version they were playing. The arrangement of monsters on the screen was randomized across participants, but remained the same for all 80 trials by each participant. In the static version (Figure 1B), each monster gave out a certain number of stars, and the number remained constant for the duration of the task. In the dynamic version (Figure 1C), the monster that gave out 1 star (the lowest reward) during Trials 1-40 surreptitiously switched to giving out 8 stars (now the highest reward) during Trials 41-80. Players were not told about the switch but could discover it for themselves if they ever clicked on that monster after the 40th trial. Each time they earned $\frac{1}{6}$ of the total number of possible stars ($\frac{1}{6}$ of the total possible was 80 stars in the static version, or 93 stars in the dynamic version), child participants were rewarded with a sticker. Adult participants were given $2.00 at the end of 80 trials, regardless of their performance. At the end of the task, participants were asked how many stars each monster gave out. Both experiments were pre-registered (osf.io/4rzsa/registrations). More details on the procedure, method, and all of the data can be found on the Open Science Framework (osf.io/4rzsa/).

Experiment 1 included 24 adults (range = 23 - 52, mean = 34) and 24 children (range = 6.0 - 12.2, mean = 8.9). As Figure 2 shows, most adult participants in both static and dynamic versions quickly determined which monster gave out the most stars and continued to choose that monster for the rest of the game. Most children, by contrast, never stopped exploring. They continued to click on different monsters throughout the game. We defined exploratory behavior as switching responses and non-maximizing responses. A switching response was defined as choosing a different monster than the one chosen on the previous trial. A non-maximizing response was defined as choosing a monster other than the one offering the highest payoff observed thus far. All statistical analyses involved Bayesian t-tests between children and adults, Bayesian linear models, and Bayesian tests of association using the BayesFactor package (Morey, Rouder, Jamil, & Morey, 2015), with default priors comparing a null model of no difference with the alternative model of a difference. Under both the 'switching' and the 'non-maximizing' definitions of exploration, children explored at a significantly higher rate than adults did (proportion of switch choices, Figure 2A, $BF_{10} = 6.20 \times 10^{10}$; proportion of non-maximizing choices, Figure 2B, $BF_{10} = 2.57 \times 10^{7}$). Consequently, children earned significantly fewer stars than adults (Figure 2C, $BF_{10} = 2.68 \times 10^{6}$). However, in the dynamic version of the task, children were much more likely to discover that the monster who had been giving out one star started giving out eight—the highest number of any monster. By the end of 80 trials in the dynamic version, 93.3% of children but only 33.3% of adults correctly identified the monster that gave out 8 stars (Figure 2D, Bayesian test of association, $BF_{10} = 276.43$ in favor of a relationship between age group and answers to this question). Of course, all of the adults who answered the posttest question correctly had discovered the 8-star option by exploring. Thus, while being a child was a good predictor of correctly identifying the 8-star monster (Bayesian linear model, $BF_{10} = 100.49$), exploratory behavior—defined as high levels of switching ($BF_{10} = 1705.60$) and non-maximizing ($BF_{10} = 1233.12$) were better predictors. Within child participants, there was inconclusive evidence about a relationship between age and exploratory behavior ($BF_{10} = 1.43$ for non-maximizing choices; $BF_{10} = 1.14$ for switching).

In Experiment 2, we replicated these findings with a new sample of 115 adults (range = 19 - 64, mean = 35.2) and 50 children (range = 4.9 - 12.1, mean = 8.6). Again, children switched between the monsters more often than adults did (Figure 2E, $BF_{10} = 7.62 \times 10^{36}$), chose non-maximizing monsters more than adults did (Figure 2F, $BF_{10} = 2.58 \times 10^{37}$), collected fewer stars than adults overall (Figure 2G, $BF_{10} = 6.41 \times 10^{29}$), and were more likely to detect the change in the dynamic condition than adults were (Figure 2H, $BF_{10}=558$). While the sample included a large age-range, within the child participants there was inconclusive evidence leaning towards the null regarding a relationship between age and exploratory behavior (BF_{10} = 0.39 for non-maximizing choices; BF_{10} = 0.45 for switching). Again, there was some variation in the data: some children maximized reward and some adults explored. The best predictor of correctly identifying the 8-star monster was a high proportion of switching responses (BF_{10} = 13216616), followed by a high proportion of non-maximizing responses (BF_{10} = 1632918), followed by being a child (BF_{10} = 188). In other words, children's tendency to explore led to lower payoffs in the short term, but greater chances of detecting important changes in the environment.

From an adult perspective, children's exploration can seem inefficient because it does not lead to easily measurable achievements. For example, many parents believe that their children learn more from organized, adult-supervised activities than from unstructured free play, alone or with other children. Our research suggests that behavior seen by adults as inefficient (e.g., choosing any monster other than the one expected to give the most stars) actually allows children to learn more about their environments (in this case, the game environment) and importantly, to detect changes in the environment that adults miss. In other words, even when exploration is costly in the short term (resulting in fewer stars earned in 80 trials of the game), children's exploration leads to a better understanding of the world in the long term, particularly when that world is changing. Humans' extended period of immaturity means that we spend longer exploring the world before narrowing our focus to exploit the resources that we have discovered. Childhood exploration pays off in the form of a deeper and more flexible understanding of the world around us.

From our perspective as developmental psychologists, this research suggests that children's exploration and unstructured free play are not a waste of time, but an important way that children learn about the dynamic and changing world around them. This has implications not only for parenting and education but also for artificial intelligence. An algorithm that acts like an adult by immediately maximizing reward is likely to perform worse in a dynamic environment than one that acts like a child by exploring longer. While many researchers in artificial intelligence and reinforcement learning are beginning to build these ideas into their algorithms (Botvinick et al., 2019; Haarnoja, Tang, Abbeel, & Levine, 2017; Ritter et al., 2018), our study is among the first to empirically demonstrate the benefits of extended exploration for learning in dynamic environments.

To summarize, a childhood bias towards exploration leads to greater long-term learning. Thus, humans' long childhood enables us to learn much more about our complex world than we would if we behaved like adults, maximizing immediate payout at the cost of exploration. Exploration is not wasteful or inefficient; it equips children with the knowledge they need to master an environment that they will soon need to independently navigate.

Acknowledgements

This work was funded by a UCI Department of Cognitive Sciences Jean-Claude Flamagne Research Award awarded to ESS, a UCI Graduate Division Public Impact Fellowship awarded to ESS, a UCI School of Social Sciences Fellowship in Honor of Christian Werner awarded to ESS, ARC Discovery Project DP150104206 awarded to DJN, and ARC Discovery Project DP180103600 awarded to AP. We thank Sarnecka Lab Research Assistants Emily Garcia, Daria Lasc, Silvia Navarro Hernandez, Heaven Howse, Madison Salcado for their assistance with data collection. We thank the Discovery Cube of Orange County and all of the families who participated. We thank Aaron Bornstein, Rena B. Goldstien, Candy Liu, Steve Piantadosi, Jenny Richmond, Mac Strelioff and Darby Vickers for useful discussion.

References

Blanco, N. J., & Sloutsky, V. (2019, Aug). *Systematic exploration and uncertainty dominate young children's choices.* PsyArXiv. Retrieved from psyarxiv.com/72sfx doi: 10.31234/osf.io/72sfx

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement learning, fast and slow. *Trends in cognitive sciences*.

Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain.* WW Norton & Company.

Estes, R. D., & Estes, R. K. (1979). The birth and survival of wildebeest calves. *Zeitschrift für Tierpsychologie*, *50*(1), 45–95.

Gentner, D., & Goldin-Meadow, S. (2003). *Language in mind: Advances in the study of language and thought.* MIT press.

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., … Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences*, *114*(30), 7892–7899.

Haarnoja, T., Tang, H., Abbeel, P., & Levine, S. (2017). Reinforcement learning with deep energy-based policies. In *Proceedings of the 34th international conference on machine learning-volume 70* (pp. 1352–1361).

Hayssen, V. D., Van Tienhoven, A., & Van Tienhoven, A. (1993). *Asdell's patterns of mammalian reproduction: a compendium of species-specific data.* Cornell University Press.

Morey, R. D., Rouder, J. N., Jamil, T., & Morey, M. R. D. (2015). Package 'bayesfactor'. *URLh http://cran/r-projectorg/web/packages/BayesFactor/BayesFactor pdf i (accessed 1006 15).*

Papagianni, D., & Morse, M. A. (2015). *Neanderthals rediscovered: How modern science is rewriting their story (revised and updated edition).* Thames & Hudson.

Piantadosi, S. T., & Kidd, C. (2016). Extraordinary intelligence and the care of infants. *Proceedings of the National Academy of Sciences*, *113*(25), 6874–6879.

Riede, F., Johannsen, N. N., Högberg, A., Nowell, A., & Lombard, M. (2018). The role of play objects and object play in human cognitive evolution and innovation. *Evolutionary Anthropology: Issues, News, and Reviews*, *27*(1), 46–59.

Ritter, S., Wang, J. X., Kurth-Nelson, Z., Jayakumar, S. M., Blundell, C., Pascanu, R., & Botvinick, M. (2018). Been there, done that: Meta-learning with episodic recall. *arXiv preprint arXiv:1805.09692*.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *BioRxiv*, 327593.

Sterelny, K. (2012). *The evolved apprentice*. MIT press.

Sumner, E., Steyvers, M., & Sarnecka, B. W. (2019). It's not the treasure, it's the hunt: Children are more explorative on an explore/exploit task than adults. In *Proceedings of the cognitive sciences society 2019* (Vol. 41).

Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. Belknap Press.

Figure 1. (A) Illustration of the task. (B) Summary of payouts in the static version. (C) Summary of payouts in the dynamic version. The onscreen positions of the monsters and the pairing of specific monsters with payout values was randomized across participants.

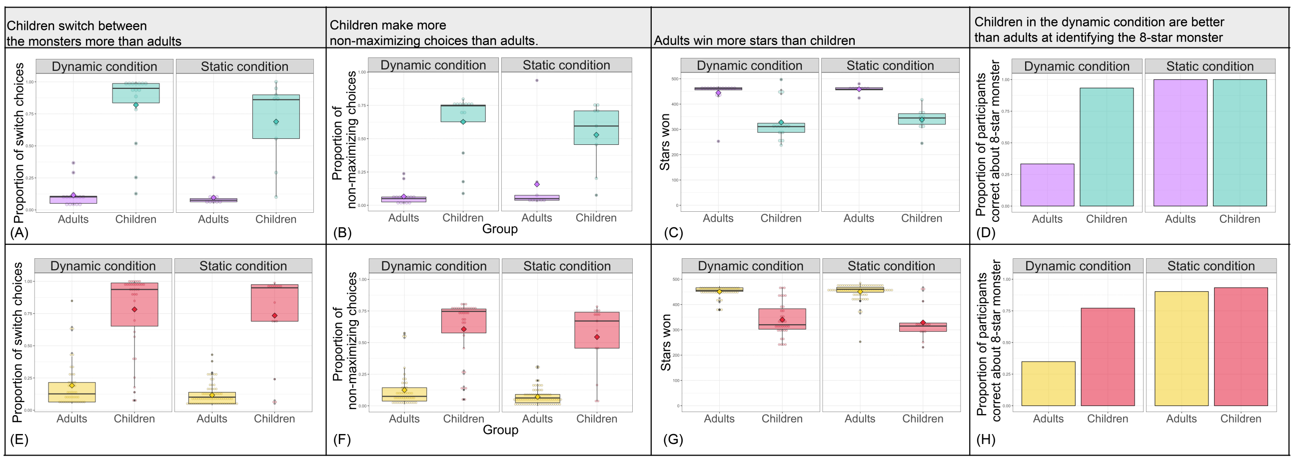

Figure 2. Circles represent individual participants' performance, diamonds represent mean performance. Panels A-D show the data from Experiment 1, Panels E-H show the data from Experiment 2. D & H show post-test performance to the question, "Did any of the monsters give you 8 stars?" If the participant said yes, they were asked to point to the one that did. (For the static version, the correct answer to this question was that none of the monsters gave 8 stars.)