Older but not Younger Children Distinguish between Ability and Proximity-Based Close Counterfactuals

by

Amy Howard

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2026

© Amy Howard 2026

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public

Statement of Contributions

Amy Howard was the sole author of this thesis which was written under the supervision of Dr. Ori Friedman and Dr. Stephanie Denison. This research was conducted at the University of Waterloo by Amy Howard under the supervision of Dr. Friedman and Dr. Denison. Ariana Hadjiyianni assisted with data collection in Experiments 2 and 3.

Abstract

Counterfactuals are events that could have happened but did not. This study examined whether children (4-7-years-old, N = 408) understand close counterfactuals based on proximity (e.g., “He almost won the race”) and those based on ability (e.g., “He easily could have won the draw”). Children saw a race between three animals where the fastest lost (e.g., cheetah), one of the two slower animals won (e.g., sheep) and the other was just behind (e.g., pig). They were then asked which animal was counterfactually close to winning in terms of proximity or ability. In Experiment 1, they were either asked which animal almost or easily could have won and in Experiment 2 they were either asked which animal almost or should have won. In Experiment 3, we added a fourth animal which tied for last with the fastest animal and explained that the animals in last place suffered a mishap while racing. Children were asked which animal almost won, easily could have won, wants to race again most, and will still be sad tomorrow. Children showed an understanding of proximity-based counterfactual closeness starting at 4-years-old, consistently identifying the second-place animal as the one who almost won. In contrast, only older children showed an understanding of ability-based counterfactuals. Older children mostly identified the fastest animal as the one who easily could have or should have won. This suggests a developmental gap where younger children recognize proximity but not ability-based counterfactual closeness.

Acknowledgements

There are countless people who have supported and guided me throughout this journey. I'll begin by thanking my supervisors Ori Friedman and Stephanie Denison for their incredible mentorship. They recommended the Developmental area's Master's of Applied Science to me and welcomed me to continue my studies in the Masters of Arts program. Under their guidance I developed ideas into testable hypotheses and developed my research and writing skills.

Thank you to the members of the Child Cognition Lab (CCL) and Developmental Learning Lab (DLL) who provided feedback on my presentations and stimuli. With this feedback I was able to improve my presentations and the design of my study. One of the most important ways members of CCL and DLL helped was by participating when I asked adults to rate animals' skill. These ratings helped me decide which animals and competitions to feature in my stimuli. Without those ratings I may have stumbled into unexpected differences of opinion on animals' ability levels that could have skewed my results. A special thanks for Ariana Hadjiyianni for helping me collect data for Experiments 2 and 3.

My friends and family also participated in this survey of animals' ratings and provided feedback on stimuli. They provided me with a layperson's perspective on the design of my studies. By sharing their thoughts, they helped me understand different ways of thinking about the questions and provided an informal adult contrast to the way kids were responding.

I thank my parents for all they've done to support me. Their financial and emotional support made it possible for me to continue my studies at the graduate level. Their support at this stage of my life builds upon the incredible support I've received from them throughout my life that allowed me to reach my goals.

I'm thankful for the support and community I've found in the Developmental area at the University of Waterloo. From the questions and feedback I received during my brown bag presentations to the friendship I found in fellow graduate students, I'm grateful for the way the people in the Developmental area have supported my growth. In particular, thanks to Claudia Sehl and Sarah English for their mentorship as senior graduate students and Katherine Cost for her help thinking through the statistics of the final experiment.

Table of Contents

Author's Declaration ii

Statement of Contributions iii

Abstract iv

Acknowledgements v

List of Figures viii

List of Tables ix

Introduction 1

General Methods 4

Experiment 1 6

Methods 6

Participants 6

Materials and Procedure 6

Results 7

Discussion 9

Experiment 2 11

Methods 11

Participants 11

Materials and Procedure 11

Results 11

Discussion 13

Experiment 3 15

Methods 16

Participants 16

Materials and Procedure 16
Results 17
Discussion 20
General Discussion 22
Language-based Explanations 22
Ability-based Explanations 23
A Simulation-based Explanation 24
Persistence and Sadness 25
References 28
Appendix A Check of Children's Understanding of Animals' Abilities 33
Methods 33
Discussion 33
Appendix B Alternate Analyses 34
Experiment 1 Binomial GEE 34
Experiment 2 Binomial GEE 34
Appendix C Survey of Adults' Beliefs about Animals' Abilities 35
Methods 35
Discussion 35

List of Figures

Figure 1. Experiment 1 Sample Stimuli and Script 7

Figure 2. Children's Selection of the Second-place or High-Ability Animal in Experiment 1 8

Figure 3. Proportion of Animals Selected for Each Condition in Experiment 1 9

Figure 4. Children's Selection of the Second-place or High-Ability Animal in Experiment 2 12

Figure 5. Proportion of Animals Selected for Each Condition in Experiment 2 13

Figure 6. Experiment 3 Sample Stimuli and Script 17

Figure 7. Children's Selection of Expected Animals in Experiment 3 19

Figure 8. Proportion of Animals Selected for Each Question in Experiment 3 20

Figure C1. Histograms for Running Ability 38

Figure C2. Histograms for Climbing Ability 39

List of Tables

Table C1. Mean and Standard Deviation of the Animals in Running and Climbing Races..... 37

Introduction

In our daily lives we are constantly considering counterfactuals – ways events could have happened differently. They are important for reasoning about events. People consider counterfactuals when identifying causes (Levy, 2015; Mandel & Lehman, 1996; Roese, 1997) and how to prevent similar events (Mandel & Lehman, 1996). Comparisons between events and the counterfactuals we generate determine whether we feel regret or relief (Kahneman & Miller, 1986) and consider events lucky or unlucky (Teigen, 1995).

Some counterfactuals are far from reality, like what the world would be like if the wheel had never been invented. Others are so close that people talk about them as unrealized realities (Kahneman & Varey, 1990). For example, as you watch your bus pull away, you could imagine having caught it. These are close counterfactuals. People describe these events with words like "almost" (e.g., Beck & Guthrie, 2011; Kahneman & Varey, 1990; Lu & Ganea, 2025) and "easily could have" (e.g., Goodman, 2015; Levy, 2015; Teigen, 1995). For example, you might say a ball that bounced off a goal post *almost* went in. Alternatively, a long shot that made it in *easily could have* missed the net.

There are different kinds of counterfactual closeness. Sometimes counterfactual closeness comes down to proximity and distance. Someone with a lottery ticket that is just one digit off from the winning number might claim they almost won the jackpot. Other times, closeness comes down to ability. If someone's name was on 90 of 100 raffle tickets but still was not drawn, they may feel they *easily could have* won – their ability to win had been very high. People also express these two kinds of counterfactual closeness differently. They often express proximity-based closeness by saying an event "almost" happened but ability-based closeness by saying an event "easily could have" happened (Doan et al., 2024, 2020). In one study, adults saw stories about a contest where winning required getting six red balls from a machine with both blue and red balls (Doan et al., 2023). Participants judged that contestants almost won when their outcomes were proximal to the winning one (i.e., five of the six balls received were red). In contrast, they judged that contestants easily could have won when they used a machine with a high ability of giving a winning result – a machine containing mostly winning red balls and few losing blue ones.

Here I aimed to discover whether children understand one kind of counterfactual closeness before the other. Children's capacity for counterfactual thinking develops gradually, with certain counterfactual concepts emerging before others (Gautam & McAuliffe, 2024; Nyhout & Ganea, 2019b). For example, children experience and recognize regret at about 5-years-old but not relief until about 7-years-old, which are both felt in response to counterfactuals (i.e., what could have happened; Jones et al., 2025; McCormack et al., 2016; Weisberg & Beck, 2010; see also Gautam et al., 2022; Jones et al., 2024). Children also reason counterfactually about events with a single cause as early as 4-years-old but not about doubly caused events until age 6 (McCormack et al., 2018; Rafetseder et al., 2013). However, one study found children could do this at age 4 for physical events, suggesting children may find it easier to reason counterfactually about physical events than social events (Nyhout & Ganea, 2019a).

Children might understand proximity-based closeness before ability-based closeness. Proximity-based close counterfactuals may be easier to simulate. For example, imagine someone who is standing far from a net kicks a ball and misses by mere inches. This shot entering the net is a proximity-based close counterfactual. To imagine this, you only need to shift the trajectory by a few inches; it is not necessary to simulate the whole event from the beginning. In contrast, if the ball missed by a large margin, you might need to simulate the event from the beginning, with the ball taking an entirely new trajectory. This mental effort is required even if the kicker is an excellent shot and was standing right in front of the net. Thus, a counterfactual that is close in terms of ability might be harder to simulate than one that is close in terms of proximity.

In line with this possibility, one study found that while 4-year-olds understand proximity-based closeness, only older children recognized ability-based closeness (Pawsey et al., 2024). Children saw a race where the fastest racer tripped on a rock and was passed by the other two racers, losing the race. In this race, the second-place racer was just behind the winner when the race ended. Even at age 4 most children recognized that the second-place racer almost won. It was not until age 7 that most children recognized the fastest racer easily could have won. This suggests children may recognize proximity-based counterfactual closeness before ability-based closeness.

However, these findings may underestimate children's understanding of ability-based close counterfactuals. When children answered the counterfactual question, they could see that

the second-place racer was just behind the winner. In contrast, they could not see that the racer who tripped had been the fastest. Instead, children had to recall that the racer who tripped had initially been in the lead. This may have been difficult for 4-year-olds. The visual support for proximity but not ability-based closeness could explain the gap, rather than a lack of understanding. In line with this, previous research shows working memory supports counterfactual reasoning (Guajardo et al., 2009; Stragà et al., 2023; Van Hoeck, 2015; see also Beck et al., 2009). Children with good working memory may have better recall of details like which racer was fastest. Reducing the need to rely on working memory may reveal children understand ability-based counterfactuals earlier than previously thought.[1]

I examined 4-7-year-olds' ability to reason about close counterfactuals that are based on proximity and ability in three experiments. In my experiments, I showed children stories about races. In each race, the fastest racer ended far from the finish line, one slower racer ended near the finish line, and another won. To provide a visual reminder of ability, the racers were all animals with well-known abilities. For instance, in stories with running races, the high-ability animal was a cheetah, the winner was a sheep, and the runner-up was a pig. With these stimuli, children did not need to think back to the beginning of the race to remember which racer showed the most ability. They only had to draw on their knowledge that cheetahs are fast. So, if difficulty recalling story details explains the gap between proximity and ability-based closeness, I should not find this gap in my study.

I contrasted children's responses to proximity and ability-based counterfactual questions. In the first experiment, I probed children's proximity-based counterfactual thinking by asking who "almost" won and probed their ability-based counterfactual thinking by asking who "easily could have" won. In the second experiment, I followed the same procedure, but changed the

[1] Alternatively, these findings may overestimate children's understanding of proximity-based close counterfactuals (Pawsey et al., 2024). Prior work suggests children are unable to identify close counterfactual events that "almost" happened until 5-years-old (Beck & Guthrie, 2011; see also Harris, 1997 as cited in Beck & Guthrie, 2011). Children saw toy horses run toward the edge of a table. One horse stopped near the edge and the other fell off. Children aged 3-4 misidentified the fallen horse as the one that "almost" fell. It was not until age 5 that children correctly identified the horse that stopped near the edge as the one that "almost" fell. If children only grasp proximity-based closeness at age 5, then the gap observed here may be spurious (Pawsey et al., 2024). It is possible that 4-year-olds struggle with closeness in their counterfactual reasoning.

ability-based counterfactual question from "easily could have" to "should have." In both experiments, selecting the low-ability racer who is just behind the winner would suggest children recognize proximity-based counterfactual closeness. Selecting the high-ability racer who was far from the finish line would suggest children recognize ability-based counterfactual closeness.

In the final experiment, I examined whether the developmental gap in children's reasoning about proximity and ability-based counterfactuals can be explained by difficulty reasoning about ability. To do this, I replicated the findings of the first experiment (i.e., almost and easily could have judgements) and expanded my question set to explore how children reason about ability. For example, I asked children which animal wanted to race again the most. If children struggle with ability-based counterfactuals because they have difficulty using ability in their reasoning, then they should also struggle with other questions that require reasoning about ability. If true, children's success on these questions and the ability-based counterfactual question should show similar developmental trajectories. That is, once children become capable of using ability to reason, they succeed in using ability to reason about close counterfactuals. Alternatively, if I found no change in children's reasoning about ability, then ability alone would not explain the gap between children's understanding of proximity and ability-based counterfactuals.

General Methods

Children were tested in-person at local schools and daycares or online via Zoom. When run in-person, I showed the children the stimuli on a laptop and when online, I shared my screen to show children the stimuli. I narrated two stories where animals raced. The first was a running race and the second was a climbing race. In Experiments 1 and 2, stimuli were presented using Power Point. In Experiment 3, the stimuli were presented using Qualtrics.

Children were first asked "Which animal won the race?" to check they were paying attention and understood the story. If a child said the wrong animal, the story was repeated and the question was asked again. If they responded wrong a second time, they were told which animal won and the experiment continued. The final responses of all but one child in Experiment 1 and three children in Experiment 3 were correct.

I was concerned younger children may not know which animal is the high-ability animal in my races. To check for this, I showed 4 and 5-year-olds (N = 20) the openings of each race

and asked them which animal was the fastest. Children correctly identified the cheetah and monkey as the high-ability animals (see Appendix A). This suggests lack of knowledge about the animals' abilities could not explain young children's difficulty with ability-based counterfactual closeness.

To analyze the data, I ran generalized estimating equation (GEE) models using the *geepack* (Højsgaard & Halekoh, 2006) and *emmeans* (Lenth, 2025) packages in R. I initially ran GEE models for binomial data in all experiments. However, for Experiments 1 and 2 I found ceiling effects because nearly all children in the Almost condition selected the second-place animal. For this reason, I report the results of poisson GEEs for these experiments. The results of the initial binomial GEEs are reported in Appendix B.

I aimed to recruit 20 children per age, per condition. Children's age and gender were recorded. Demographic information was not collected, in accordance with permissions of the University of Waterloo's IRB. The demographics for the area where children were tested is 64% White, and South Asians are the largest visible minority. Most schools were in suburban, middle-class neighborhoods.

Experiment 1

Methods

Participants

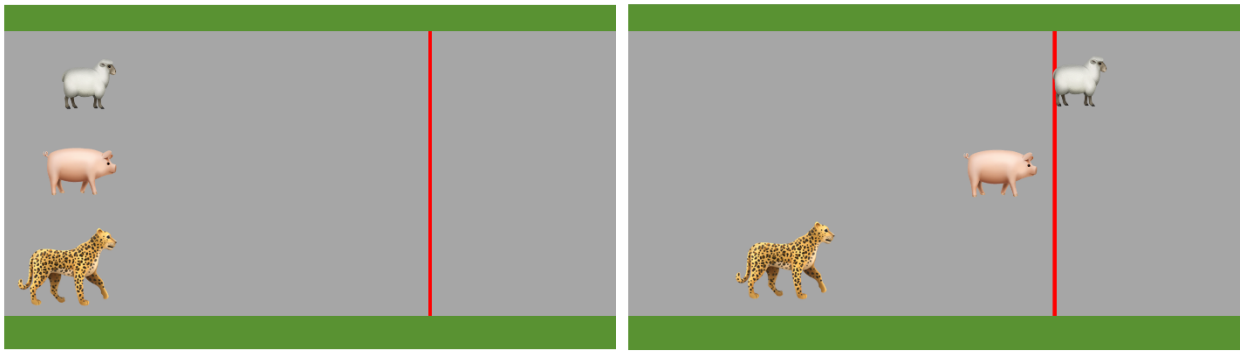
My final sample included 161 children aged 4-7 (M_{age} = 5;11, range = 4;0 - 7;11, 84 girls and 77 boys). Three additional children were seen but excluded: one did not wish to participate, one did not respond to any questions, and there was external interference for one.

Materials and Procedure

Children heard stories where animals raced. The first story was about a running race, and the second story was about a climbing race; see Figure 1. In each race, one animal was well known for its ability, while the other two were not. In the running race, a cheetah (high-ability) competed against a pig and a sheep; in the climbing race, a monkey (high-ability) competed against a cat and raccoon.[2] In both races the high-ability animal lost the race, one of the slower animals won and the other was just behind. For example, in the running race the cheetah was in last place, the sheep won, and the pig was just behind. At the end of each story children were asked a close counterfactual question which differed across two between-subjects conditions. In the Almost condition children were asked which animal almost won and in the Easily condition they were asked which animal easily could have won.

[2] Animals were selected based on a small poll where adults rated 22 animals on their ability to run and 11 animals on their ability to climb (see Appendix C). For each race, we selected the highest rated animal and selected two similarly rated animals with lower ratings.

Figure 1
Experiment 1 Sample Stimuli and Script



Do you see the animals? The sheep, pig, and cheetah are going to have a running race. The animal that crosses the red line first wins the race. Here are the animals at the start of the race.

And look! Here they are at the end of the race. Which animal won the race? *Which animal [almost/easily could have] won the race?*

Note: The words in the square bracket indicate differences in script across between-subject conditions. All participants also saw a climbing race featuring a monkey, raccoon, and cat.

Results

If children recognize proximity-based counterfactual closeness, they should choose the second-place animal in the Almost condition. Likewise, if they recognize ability-based counterfactual closeness, they should select the high-ability animal in the Easily condition. To examine children's understanding, selections of the second-place animal in the Almost condition and selections of the high-ability animal in the Easily condition were scored 1. All other selections were scored 0.

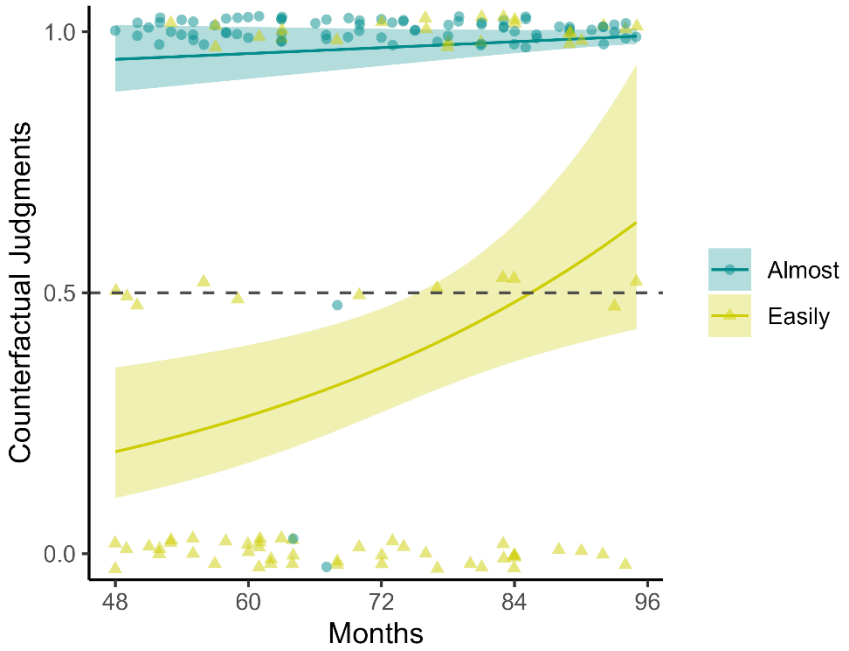
I entered scores into a GEE with condition (Almost, Easily) and age as predictors; see Figure 2. There was a main effect of condition, $F(1, 318) = 50.00$, $p < .001$ and age, $F(1, 318) = 8.05$, $p = .005$ as well as an interaction, $F(1, 318) = 6.89$, $p = .009$. Overall, children in the Almost condition scored higher than those in the Easily condition. The interaction occurred because whereas scores did not differ by age in the Almost condition, $F(1, 318) = 1.99$, $p = .159$ they improved with age in the Easily condition, $F(1, 318) = 7.50$, $p = .007$.

To understand the development of children's recognition of the kinds of counterfactual closeness, I examined confidence intervals. I looked at the age children mostly picked the

expected animal (i.e., at a rate greater than 0.50). This is not a test against chance, since there were three animals children could choose. It would be difficult to determine a "chance" rate since I did not know whether children would pick between the three animals equally (e.g., children may ignore the winner). Nearly all children selected either the second-place or high-ability animal.[3] In the Almost condition, children mostly selected the second-place animal $CI_{95\%}$[0.93, 1.01]. In contrast, in the Easily condition even the oldest children (7;11) only chose the high-ability animal about half the time $CI_{95\%}$[0.43, 0.94].

Figure 2

Children's Selection of the Second-place or High-Ability Animal in Experiment 1



Note: For the Almost condition the second-place animal (i.e., pig, racoon) was coded as 1 and for the Easily condition the high-ability animal was coded as 1 (i.e., cheetah, monkey). All other responses were coded as 0. This graph is based on a poisson GEE.

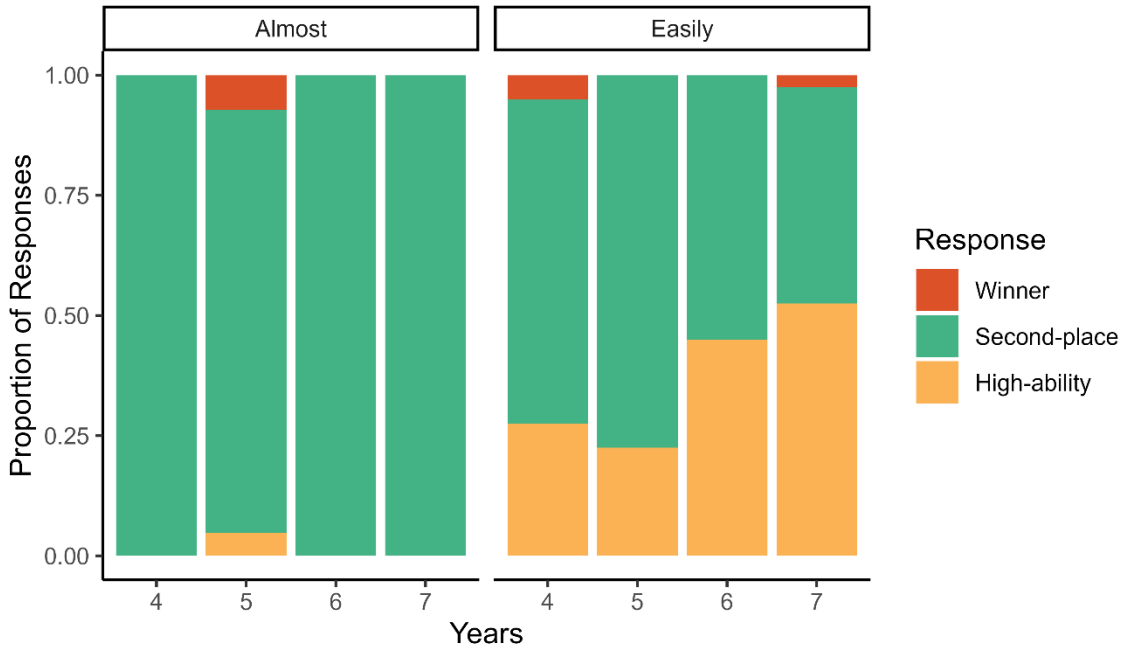
To further understand the development of children's close counterfactual reasoning, I examined the proportion of responses where children selected each animal at each age; see Figure 3. In the Almost condition, children across ages almost always selected the second-

[3] Only 2 children in the Almost condition and 2 children in the Easily condition selected the winner on at least one trial.

place animal: all responses of children aged 4, 6, and 7 and 88% of 5-year-olds' responses were the second-place animal. In the Easily condition, the modal response was the second-place animal at ages 4 (68%), 5 (78%), and 6 (55%). Only 7-year-olds selected the high-ability animal most often (53%).

Figure 3

Proportion of Animals Selected for Each Condition in Experiment 1



Note: Stacked bar graph showing the proportion of trials where children selected each racer at each age in years, with the y-axis displaying the proportion as a score of 0 to 1.

Discussion

Children nearly always selected the second-place animal in the Almost condition, showing an understanding of proximity-based close counterfactuals as early as age 4. In the Easily condition, it was only at 7-years-old that children mostly selected the high-ability

animal and even then, it was only approximately half the time. This suggests children may find proximity-based counterfactuals easier to recognize than ability-based counterfactuals.

However, this task may underestimate children's understanding of ability-based closeness. Children may have interpreted the outcome of the race as indicative of each animal's skill, concluding that the second-place animal had sufficient skill to "easily" win the race. Alternatively, children may have found the wording of the Easily question harder to understand. To address this, I ran a second experiment where I asked children which animal "should have won" instead of which animal "easily could have won." If the first explanation is correct and children think the second-place animal nearly winning means it is fast, then children should again choose the second-place and high-ability animals at a similar rate in response to the ability-based counterfactual question (i.e., who should have won). Alternatively, if children found the "easily could have" question difficult, then they should pick the high-ability animal more often in response to the easier question.

Experiment 2

Methods

Participants

My final sample included 124 children aged 4-6 (M_{age} = 5;5, range = 4;0 - 6;11, 63 girls and 61 boys). Four additional children were seen but excluded: one did not wish to participate, one did not respond to any questions, and one was out of the age range planned for the study.

Materials and Procedure

Experiment 2 used the same materials and procedures as Experiment 1, but instead of the Easily condition there was a Should condition where children were asked "Which animal should have won the race?"

Results

I examined how often children selected the second-place animal in the Almost condition and the high-ability animal in the Should condition, demonstrating understanding of proximity and ability-based counterfactual closeness respectively. These responses were scored 1 and all other responses were scored 0.[4]

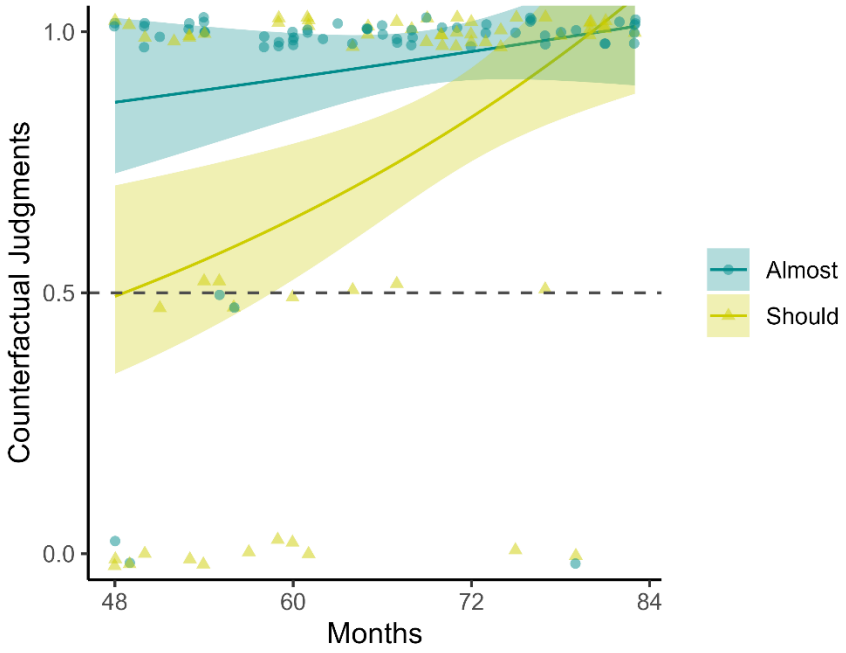
Scores were entered into a GEE with condition (Almost, Should) and age as predictors; see Figure 4. This analysis found a main effect of condition, $F(1, 242) = 10.58$, $p = .001$, a main effect of age, $F(1, 242)$ 10.37, $p = .002$, and an interaction between condition and age, $F(1, 242) = 4.58$, $p = .033$. The main effect occurred because overall children in the Almost condition scored higher than those in the Should condition. The interaction occurred because scores in the Almost condition did not differ by age, $F(1, 242) = 1.30$, $p = .255$, whereas scores in the Should condition did increase with age, $F(1, 242) = 9.26$, $p = .003$.

[4] Children almost always selected either the second-place or the high-ability animal. Only 1 child in the Almost condition and 4 children in the Should condition selected the winner on at least one trial.

I also examined confidence intervals to determine when children mostly selected the expected animal as an indication of when children begin recognizing each kind of close counterfactual. In the Almost condition, children at all ages mostly selected the second-place animal $CI_{95\%}$[0.88, 0.99]. Children in the Should condition mostly selected the high-ability animal at 4;11 years $CI_{95\%}$[0.51, 0.78].

Figure 4

Children's Selection of the Second-place or High-Ability Animal in Experiment 2



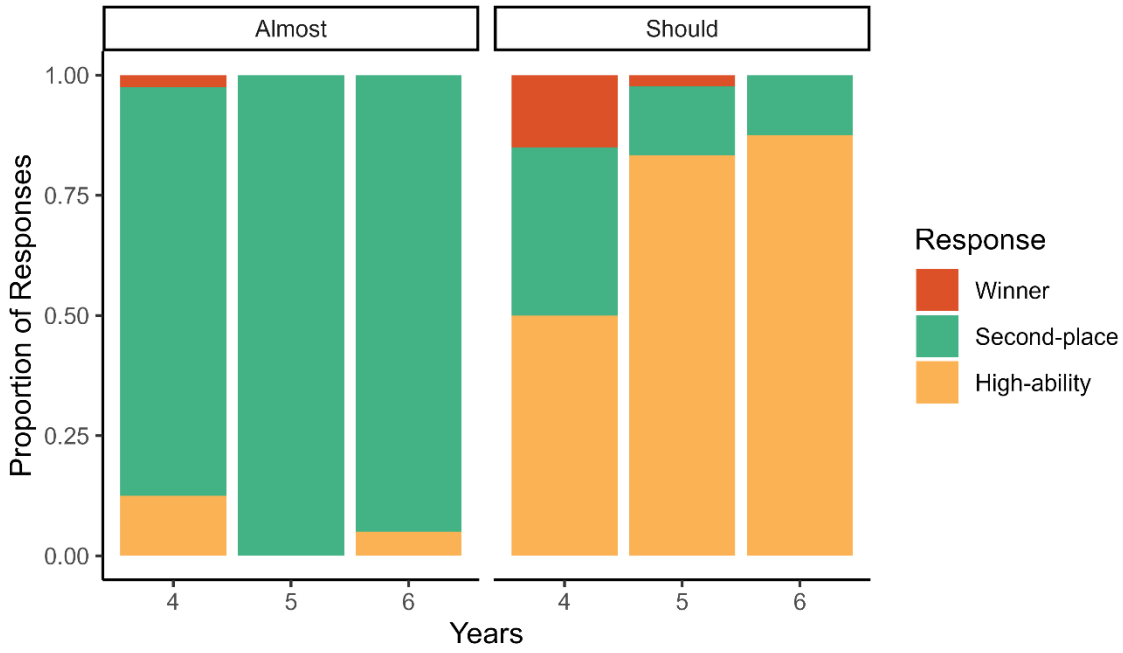
Note: For the Almost condition the second-place animal (i.e., pig, racoon) was coded as 1 and for the Easily condition the high-ability animal was coded as 1 (i.e., cheetah, monkey). All other responses were coded as 0. This graph is based on the poisson GEE.

To better understand the development of children's counterfactual reasoning, I examined the proportion of responses where children selected each animal at each age; see Figure 5. In the Almost condition, the second-place animal was the modal response across

ages: 85% at age 4, 100% at age 5, and 95% at age 6. In the Should condition, the modal response for 4-year-olds was the high-ability animal (50%) and at 5-years-old children started selecting the high-ability animal most of the time (83%).

Figure 5

Proportion of Animals Selected for Each Condition in Experiment 2



Note: Stacked bar graph showing the proportion of trials where children selected each racer at each age in years, with the y-axis displaying the proportion as a score of 0 to 1.

Discussion

Compared to Experiment 1, children showed an earlier understanding of ability-based closeness. When they were asked which animal "should have" won the race children aged 4;11 years mostly selected the high-ability animal. Still, the gap between children's understanding of proximity and ability-based close counterfactuals persisted. This suggests

that the difficulty with the phrase “easily could have” cannot fully explain the gap between children’s performance on proximity and ability-based counterfactual questions.

It is possible that the “should have” question was easier for children because it led them to be less concerned with what happened in this specific race and more concerned with what is generally true. For example, if a child assumed there was a mishap that caused the cheetah to lose, they may feel uncertain about whether the cheetah could have avoided this mishap to easily win this race when simulating who easily could have won. In contrast, since “should” can be prescriptive, the child may be less concerned about whether this cheetah could avoid a mishap and conclude that since cheetahs are generally very fast, the cheetah should have won. If this is true, not knowing what happened during the race may have hampered children’s performance when asked who “easily could have” won but not who “should have” won as simulating an alternate version of this race would be difficult with so few details.

Additionally, children may struggle with ability-based counterfactual questions because they struggle to reason based on ability. For example, adults understand that someone high in ability is more likely to try again after a failure than someone who is low in ability. Children may not recognize this if they struggle to reason about ability.

In Experiment 3 I told children the high-ability animal had an accident to explain why they lost the race. By reducing ambiguity, I intended to make the Easily question easier to answer. I asked two counterfactual questions and two additional questions which I expected children to answer by reasoning about ability.

Experiment 3

I again asked children, “Which animal almost won the race?” and “Which animal easily could have won the race?” to examine their counterfactual thinking. I decided to use “easily could have” over “should have” as the ability-based counterfactual question. I wanted to avoid the prescriptive associations inherent in the word “should” because I was concerned children may be less likely to simulate an alternate version of the race compared to when they are asked who easily could have won. Additionally, I thought children may find the Easily question easier to answer if they knew why the fastest animal did not win the race, so I told them this animal tripped or slipped.

I also asked two additional questions to gauge children’s understanding of ability: “Which animal wants to race again the most?” and “Which animal will still be sad about losing the race tomorrow?”. If children recognize that the fastest animal had a high ability to win the race, then they should recognize that this animal would be motivated to race again because it is likely to win subsequent races.

Children should also expect the fastest animal to be saddest. This animal may have expected to win, and so might be especially disappointed about losing. Understanding this requires children to recognize that the animals’ ability impacts their emotional response. To differentiate between children selecting the high-ability animal because it was in last place and selecting this high-ability losing animal because it expected to win, I added a fourth slow animal to the race. This animal was tied with the high-ability loser and suffered the same accident during the race. If children judged the saddest racer based on whether they were in last place, then they should select the two losing animals at similar rates. If they do recognize the high-ability animal expected to win, then they should say the high-ability animal is saddest more often.

Methods

Participants

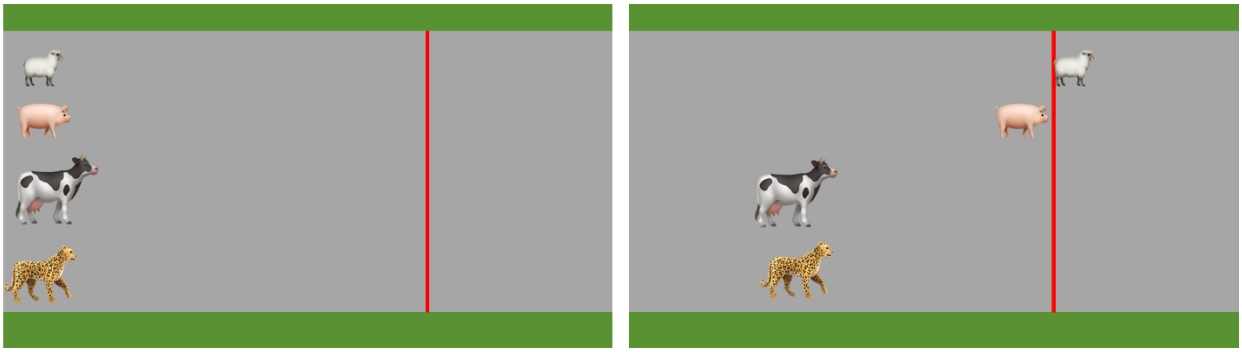
My final sample included 123 children aged 4-7 (M_{age} = 5;11, range = 4;0 - 7;11, 64 girls and 59 boys). Four additional children were seen but excluded: two had previously participated in a similar experiment, one due to interference, and one due to technical issues.

Materials and Procedure

Children again saw two stories about races between animals: a running race and a climbing race. Each race had one high-ability animal and three lower-ability animals. For example, in the running race a cheetah raced against a sheep, pig and cow. Children were shown the animals lined up in their starting positions, then in their final positions at the end of the race; see Figure 6. One of the lower-ability animals won the race (i.e., sheep, cat), with another just behind (i.e., pig, raccoon). The high-ability animal (i.e., cheetah, monkey) and one of the lower-ability animals (i.e., cow, mouse) tied for last place. During the race, both of these animals had a minor accident (e.g., tripped). Children were then asked 5 questions. First I asked "Which animal won the race?" to check children's understanding of the story. If children answered wrong the story was repeated and if they answered wrong a second time they were corrected. The 4 test questions were asked in randomized order. Two were counterfactual questions: "Which animal almost won the race?" and "Which animal easily could have won the race?" (i.e., Almost and Easily questions). The two additional questions were: "Which animal wants to race again the most?" and "Which animal will still be sad about losing the race tomorrow?" (i.e., Again and Sad questions).

Figure 6

Experiment 3 Sample Stimuli and Script



Do you see the animals? The sheep, pig, cow and cheetah are going to have a running race. The animal that crosses the red line first wins the race. Here are the animals at the start of the race.

And look! Here they are at the end of the race. While they were racing, the cow and the cheetah tripped on rocks. So now I have some questions for you.

Note: The first question was always "Which animal won the race?". The four test questions asked in randomized order were: "Which animal almost won the race?", "Which animal easily could have won the race?", "Which animal wants to race again the most?", and "Which animal will still be sad about losing the race tomorrow?"

Results

Most children correctly identified the winner. Two children answered incorrectly on one of two trials and one child answered incorrectly on both trials. Children who answered incorrectly were told which animal won the race.

To begin my analyses, I examined how often children selected the animal that I expected them to if they responded based on close counterfactuals for the counterfactual questions and high-ability for the additional questions. Responses that reflected the expected response were scored 1 and all other responses scored 0. For "Which animal almost won the race?" selections of the second-place animal were scored 1 because this response is expected if children use proximity-based counterfactual thinking. For all other questions selections of

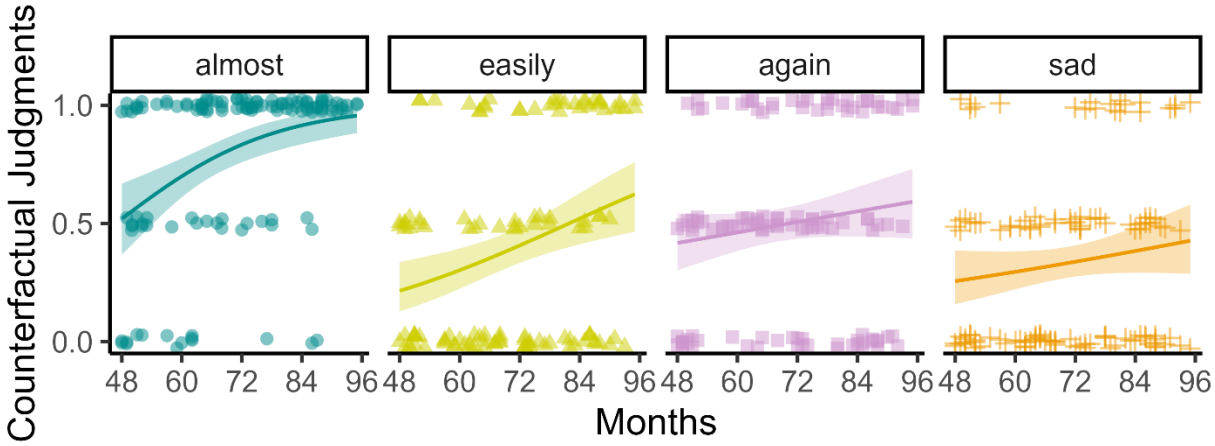
the high-ability animal scored 1 because this response is expected if children base their response on which animal had the highest ability.

I entered scores into a GEE with question and age as predictors; see Figure 7. This found a main effect of question, $F(3, 972) = 28.35$, $p < .001$ and age, $F(1, 972) = 19.94$, $p < .001$, as well as an interaction, $F(3, 972) = 2.79$, $p = .040$. I did 6 pairwise comparisons Holm corrected to explore the main effect of question. Children scored significantly higher on Almost than all other questions, ps $< .001$. Children also scored higher on Again than Sad, $t(972) = 3.38$, $p = .002$. Easily did not differ significantly from Again, $t(972) = -1.95$, $p = .103$ or Sad, $t(972) = 1.17$, $p = .224$.

The interaction occurred because although scores increased with age for both counterfactual questions (Almost, $F(1, 972) = 15.62$, $p < .001$; Easily, $F(1, 972) = 10.45$, $p = .001$), age did not affect performance for the additional questions (Again, $F(1, 972) = 1.98$, p = .160; Sad, $F(1, 972) = 2.03$, p = .155). Examining 95% confidence intervals showed that whereas children mostly selected the second-place animal for Almost from age 4;7 years $CI_{95\%}[0.52, 0.73]$, they never chose the high-ability animal more than half the time for Easily (at age 7;11 $CI_{95\%}[0.47, 0.76]$)

Figure 7

Children's Selection of Expected Animals in Experiment 3



Note: For Almost selections of the second-place animal were scored 1; for all other questions selections of the high-ability animal were scored 1. All other selections were scored 0. CI 95%.

I also examined the proportion of responses where children selected each animal for each question; see Figure 8. For Almost, the second-place animal was the modal response at all ages: 58% at 4-years-old, 78%, at 5-years-old, 89% at 6-years-old, and 90% at 7-years-old. This aligns with my previous findings, suggesting that even 4-year-olds understand proximity-based counterfactual closeness.

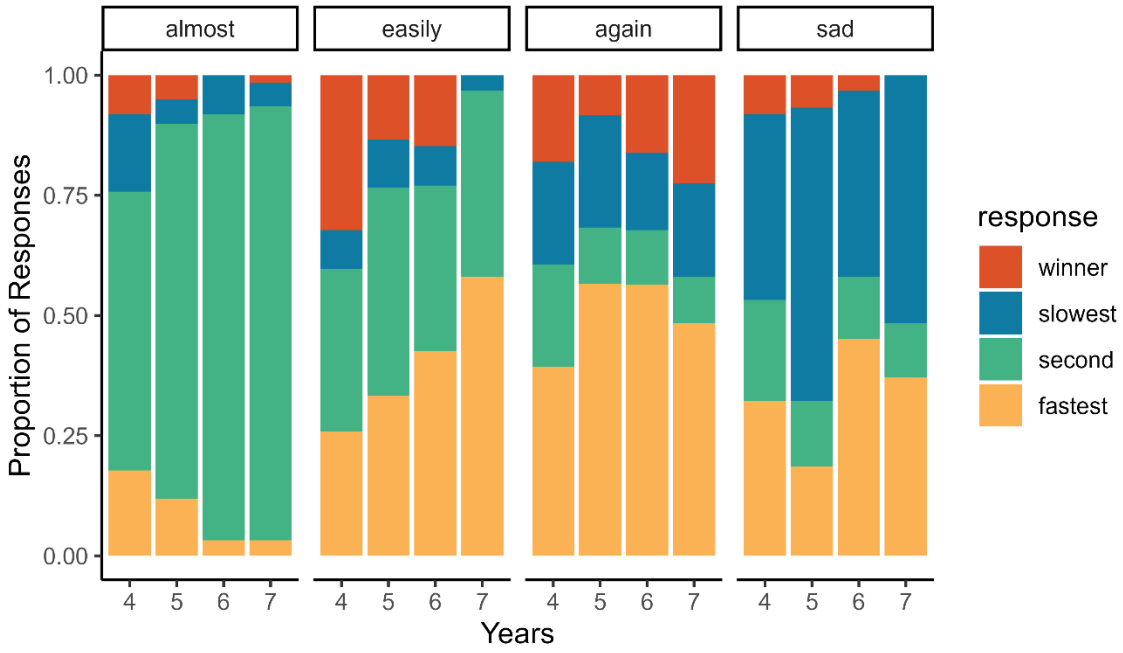
For Easily, the modal response for younger children was the second-place racer: 34% at 4-years-old and 43% at 5-years-old. The modal response was the high-ability racer at a rate of 43% at age 6 and 58% at age 7.

For Again, the modal response across ages was the high-ability animal at a rate of 39% at 4-years-old, 57% at 5-years-old, 57% at 6-years-old, and 48% at 7-years-old. Even the youngest children seemed to understand the high-ability racer would be most motivated to race again.

For Sad, the modal response overall was the slowest animal: 39% at age 4, 61% at age 5, and 52% at age 7. The exception was 6-year-olds, for whom the modal response was the high-ability animal at a rate of 45%.

Figure 8

Proportion of Animals Selected for Each Question in Experiment 3



Note: Stacked bar graph showing the proportion of trials where children selected each racer at each age in years, with the y-axis displaying the proportion as a score of 0 to 1.

Discussion

In contrast to Experiments 1 and 2, with age children's understanding of both proximity and ability-based counterfactuals improved. Previously, children across my age range were at ceiling, showing their understanding of proximity-based closeness. Elements of task design such as the additional questions and an additional distractor animal may have made the task harder for the youngest children. Despite this, 4-year-olds did show an understanding of proximity-based close counterfactuals. As in my previous studies, children

only later showed an understanding of ability-based close counterfactuals. Overall, children showed an earlier and stronger understanding of proximity-based than ability-based counterfactual closeness.

Children's responses to the additional questions did not improve significantly with age. Even 4-year-olds said the high-ability animal wanted to race again the most. Children seemed to recognize that ability influences whether someone will try again. They also picked the low-ability animal as the one that would be saddest, suggesting they considered low-ability a disadvantage that would lead to more sadness upon losing. As I consider below, these findings also show that the gap between children's understanding of proximity and ability-based counterfactual closeness is not explained by children's understanding of ability.

General Discussion

In three experiments I investigated whether there is a gap between when children understand proximity and ability-based counterfactual closeness. In each experiment children saw races and judged which losing racer was counterfactually closest in terms of proximity or ability. Children showed an understanding of proximity-based counterfactual closeness before ability-based closeness. When asked which racer *almost* won, children as young as 4-years-old consistently selected the second-place racer. In contrast, younger children did not consistently select the high-ability racer in response to ability-based counterfactual questions. Children mostly selected the high-ability racer as the one who *should have* won starting at age 5. They did not *mostly* select the high-ability racer as the one who *easily could have* won even at age 7, but 7-year-olds did select the high-ability racer more often than the other racers.

Together, these findings suggest recognition of proximity-based closeness occurs earlier than ability-based closeness. While older children recognize the two distinct kinds of counterfactual closeness, younger children only recognize proximity-based closeness. These findings contribute to our overall understanding of children's counterfactual reasoning and the mistakes they make. I discuss competing explanations for these results, limitations, and future directions below.

Language-based Explanations

Why do children find ability-based closeness harder to understand than proximity-based closeness? This difference could come down to language. One possibility is that young children have difficulty understanding the phrase 'easily could have.' Consistent with this, children selected the high-ability racer more often, and from an earlier age, when asked who *should* have won (Experiment 2). However, this is unlikely to be the full explanation, since children still showed an understanding of proximity-based counterfactual closeness before ability-based closeness.

Another language-based explanation is that children understand proximity-based closeness before ability-based closeness because ‘almost’ communicates proximity-based counterfactual closeness more clearly than ‘easily could have’ or ‘should have’ communicate ability-based closeness. Adults connect ‘almost’ to proximity-based closeness more strongly than they connect ‘easily could have’ to ability-based closeness (Doan et al., 2023). ‘Almost’ may uniquely tap into proximity-based closeness whereas ‘easily could have’ may tap into both. The difference in clarity may make it easier for children to learn about proximity-based than ability-based closeness.

Cross cultural studies suggest that language does influence when children gain counterfactual concepts (Wu, 2025). In other languages the phrases for ability-based closeness may communicate the concept more clearly than either of the English phrases. Future work may explore how differences in how clarity of phrases for close counterfactuals impacts when these concepts develop. Knowing whether or not the gap between proximity and ability-based closeness persists in other languages would clarify the role of language in explaining why this developmental gap exists.

The reliance on language in my study presents a limitation. It is possible that children have a better understanding of ability-based closeness than my results suggest, but that some children did not connect the concept to the language used. Non-linguistic ways of measuring children’s understanding of counterfactual closeness could show that children understand counterfactual closeness even earlier.

Ability-based Explanations

Children’s difficulty with ability-based closeness could also stem from broader difficulties understanding ability. Although children have a good understanding of the properties of ability (Muradoglu et al., 2025), they nonetheless make some errors with regard to others’ ability. Some previous work suggests that children under age 5 judge ability based on outcomes, even when that outcome was caused by a mishap (Yang & Frye, 2016). A few accounts are possible here. On one account, it could be that younger children do not

recognize the cheetah and monkey as the high-ability animals. However, this is unlikely because 4- and 5-year-olds identify the cheetah and monkey as the fastest animals in my races (see Appendix A). Furthermore, children's responses to the additional questions in my final experiment showed an understanding of ability and did not improve significantly with age.

On another account, after seeing the results of the race, children may conclude that the winner had the highest ability, even though members of its species are typically slow. My findings suggest this is not the case as even 4-year-olds thought the high-ability racer would be most motivated to race again. Children seemed to recognize that ability influences whether someone will try again. Furthermore, although children unexpectedly said the lowest-ability racer would be saddest, they still showed an understanding of ability. Children seemed to draw on ability to infer that the animal disadvantaged by their low ability would be sadder about losing. Altogether, this suggests ability alone cannot account for the gap in children's understanding of proximity and ability-based closeness.

A Simulation-based Explanation

I close with a simulation-based explanation. Children may find it easier to simulate the counterfactual outcome for proximity-based closeness than ability-based closeness. To simulate the second-place racer winning, children would only have to imagine them going a bit faster at the very end. There is a visual reminder that this could have happened since they were right behind the winner. In contrast, with ability the whole race has to go differently, because the fastest animal did not make it very far. Furthermore, children may find proximity easier to reason about than ability because it is more intuitive. Children understand distance and proximity even as infants (Uller et al., 2013; Yonas et al., 1985), but their reasoning about ability in general is still developing at age 4 (Yang & Frye, 2016). Altogether this means children may find the ability-based counterfactual more difficult to simulate than the proximity-based counterfactual.

One study examining children's ability to simulate counterfactuals found even 4-year-olds simulate alternate versions of events, but that children under age 6 make key mistakes (Kominsky et al., 2021). Perhaps children can accurately simulate close counterfactuals that require very few or recent changes because there are fewer opportunities for mistakes. If so, at age 4 children may still be making key mistakes when simulating ability but not proximity-based counterfactuals. However, in the study by Kominsky et al. (2021) children were able to simulate counterfactuals at age 6, but I found children were not successful in their ability-based counterfactual judgements until at least age 7. The difference could be caused by features of the task. For instance, in the Kominsky et al. (2021) study children viewed the path of the ball and picked the most correct path from a set of alternate paths whereas in my study children did not see the progression of the race and were not shown alternate outcomes. This may have made my task more difficult, particularly in the case of ability. In contrast, proximity-based counterfactuals like those in my stories may not require simulation at all.

This explanation is somewhat speculative. It is also possible that the gap is best explained by a combination of the explanations discussed here. Future work can test the simulation-based account by showing children the progression of an event where the ability of a close counterfactual to occur is visually depicted by probability and children are shown alternate versions of the event that differ in whether they follow from simulating the ability-based close counterfactual.

Persistence and Sadness

In my final experiment I added two questions with the goal of examining whether children can use ability in their reasoning. The results showed this to be true, but also yielded results relevant to children's concepts of persistence and sadness.

To my knowledge, when children expect others to persist has not been directly examined. We have rich knowledge of when children persist and what adults can do to encourage persistence (for review see Leonard et al., 2021), but we know little about

children's understanding of persistence in others. Children as young as 5-years-old make inferences about preferences based on whether someone was persistent (Amemiya et al., 2023), suggesting they recognize liking an item as a motivation to persist. By age 4 they also make predictions about their own future persistence: children say they would not try an activity again after receiving outcome-focused praise followed by criticism (Cimpian et al., 2007).

In my final experiment, children said the high-ability racer wanted to compete again the most. This suggests they thought high ability was the strongest motivator for persistence. Notably, they did not think nearly winning the race was an especially strong motivator for persistence, selecting the second-place animal at a similar rate as the other low-ability animals. However, it remains possible that ability is simply a more salient motivator in my race and that children would recognize experiencing a near miss as a motivator for persistence if it were more salient. For example, if the racers had similar ability levels, children may think the racer who experienced a near miss would be most motivated to try again.

The findings of the final experiment are also informative about children's attributions of emotions and on how children base these attributions on dimensions of cognitive appraisal, such as control (for a review see Doan et al., 2025). When children were asked which animal would feel saddest, they mostly selected one of the two losing animals. However, whereas I expected children to say the animal from the high-ability species would be saddest, they actually said the slower animal would be saddest. Children thought the disadvantaged racer, who was unlikely to win, would be sadder about losing than the advantaged racer who expected to win. This may have occurred because children considered control: whether the racers could have changed the outcome. While the high-ability animal would have won if not for an avoidable accident, the slower animal had no chance of winning, and therefore less control. Children under 8-years-old feel worse about uncontrollable causes of failure than controllable ones (Jones et al., 2024). So, when the high and low-ability animals tied for last place, instead of believing the high-ability animal would

be shocked and devastated, children may have thought the low-ability animal would despair over their inability to win. Though the line of reasoning outlined here uses counterfactual reasoning, it is not the only way to reach this conclusion. Future work can investigate the circumstances and reasons younger children believe someone who is disadvantaged would be sadder about losing.

Other dimensions children may have considered could be relevant to this. For my racing story children may also consider the *fairness* of the race, animals' *belief* that they would win, or whether the outcome was *consistent* with the animals' self-image. For example, children may have felt the race was unfair for the slower animals so the low-ability animal that also tripped would be doubly devastated by the mishap and their inability to win. The design of my study cannot determine which dimension of emotional appraisal children are drawing on. Future work may manipulate details of the situation to untangle how belief and fairness impact children's appraisal of who would be saddest.

Additionally, I specifically asked children who would be saddest *tomorrow*, but the results of this study cannot tell us whether or how this detail impacted children's responses. Adults understand that thought processes like rumination can change how people feel about an event. Usually, people feel less sad about negative events over time, but if someone ruminates on an event, their feelings may become more intense. My study showed that children thought the disadvantaged loser would be saddest the following day, but it is possible they believe a different racer would be saddest immediately after losing. Whether children understand that how people feel about an event changes over time is a question for future work.

References

Amemiya, J., Heyman, G. D., & Walker, C. M. (2023). Emphasizing others' persistence can promote unwarranted social inferences in children and adults. *Journal of Experimental Psychology: General*, *152*(10), 2977–2988. https://doi.org/10.1037/xge0001428

Beck, S. R., & Guthrie, C. (2011). Almost thinking counterfactually: Children's understanding of close counterfactuals. *Child Development*, *82*(4), 1189–1198. https://doi.org/10.1111/j.1467-8624.2011.01590.x

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Thinking & Reasoning*, *15*(4), 337–354. https://doi.org/10.1080/13546780903135904

Cimpian, A., Arce, H.-M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, *18*(4), 314–316. https://doi.org/10.1111/j.1467-9280.2007.01896.x

Doan, T., Denison, S., & Friedman, O. (2023). Two kinds of counterfactual closeness. *Journal of Experimental Psychology: General*, *152*(6), 1787–1796. https://doi.org/https://doi.org/10.1037/xge0001369

Doan, T., Denison, S., & Friedman, O. (2024). Close counterfactuals and almost doing the impossible. *Psychonomic Bulletin & Review*, *31*(1), 187–195. https://doi.org/10.3758/s13423-023-02335-w

Doan, T., Friedman, O., & Denison, S. (2020). Young children use probability to infer happiness and the quality of outcomes. *Psychological Science*, *31*(2), 149–159. https://doi.org/https://doi.org/10.1177/0956797619895282

Doan, T., Ong, D. C., & Wu, Y. (2025). Emotion understanding as third-person appraisals: Integrating appraisal theories with developmental theories of emotion. *Psychological Review*, *132*(1), 130–153. https://doi.org/10.1037/rev0000507

Gautam, S., & McAuliffe, K. (2024). Why imagining what could have happened matters for children's social cognition. *WIREs Cognitive Science*, *15*(1), Article e1663. https://doi.org/10.1002/wcs.1663

Gautam, S., Suddendorf, T., & Redshaw, J. (2022). Counterfactual thinking elicits emotional change in young children. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1866), Article 20210346. https://doi.org/10.1098/rstb.2021.0346

Goodman, J. (2015). Knowledge, counterfactuals, and determinism. *Philosophical Studies*, *172*(9), 2275–2278. https://doi.org/10.1007/s11098-014-0409-6

Guajardo, N. R., Parker, J., & Turley-Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. *British Journal of Developmental Psychology*, *27*(3), 681–702. https://doi.org/10.1348/026151008X357886

Harris, P. L. (1997). On realizing what might have happened instead. *Polish Quarterly of Developmental Psychology, 3*, 161–176.

Højsgaard, S., & Halekoh, U. (2006). The R package geepack for generalized estimated equations. *Journal of Statistical Software*, *15*(2), 1–11. https://doi.org/https://doi.org/10.18637/jss.v015.i02

Jones, A. K., Gautam, S., & Redshaw, J. (2024). Young children experience both regret and relief in a gain-or-loss context. *Cognition and Emotion*, *38*(1), 163–170. https://doi.org/10.1080/02699931.2023.2258586

Jones, A. K., Nelson, N. L., Gautam, S., & Redshaw, J. (2025). Children infer counterfactual information from others' facial expressions. *Journal of Experimental Child Psychology*, *259*, Article 106306. https://doi.org/10.1016/j.jecp.2025.106306

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136–153. https://doi.org/10.1037/0033-295X.93.2.136

Kahneman, D., & Varey, C. A. (1990). Propensities and counterfactuals: The loser that almost won. *Journal of Personality and Social Psychology*, *59*(6), 1101–1110. https://doi.org/10.1037/0022-3514.59.6.1101

Kominsky, J. F., Gerstenberg, T., Pelz, M., Sheskin, M., Singmann, H., Schulz, L., & Keil, F. C. (2021). The trajectory of counterfactual simulation in development. *Developmental Psychology*, *57*(2), 255–268. https://doi.org/https://doi.org/10.1037/dev0001140

Lenth, R. V. (2025). emmeans: Estimated marginal means, aka least-squares means. *Journal of Statistical Software*, 15(2). 1-11. https://doi.org/doi:10.32614/CRAN.package.emmeans

Leonard, J. A., Duckworth, A. L., Schulz, L. E., & Mackey, A. P. (2021). Leveraging cognitive science to foster children's persistence. *Trends in Cognitive Sciences*, *25*(8), 642–644. https://doi.org/10.1016/j.tics.2021.05.005

Levy, J. S. (2015). Counterfactuals, causal inference, and historical analysis. *Security Studies*, *24*(3), 378–402. https://doi.org/10.1080/09636412.2015.1070602

Lu, J., & Ganea, P. A. (2025). What Almost Happened? Using close-counterfactuals to prime a simulation mindset in children. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*. https://escholarship.org/uc/item/10m5b57v

Mandel, D. R., & Lehman, D. R. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality and Social Psychology*, *71*(3), 450–463.

McCormack, T., Ho, M., Gribben, C., O'Connor, E., & Hoerl, C. (2018). The development of counterfactual reasoning about doubly-determined events. *Cognitive Development*, *45*, 1–9. https://doi.org/10.1016/j.cogdev.2017.10.001

McCormack, T., O'Connor, E., Beck, S., & Feeney, A. (2016). The development of regret and relief about the outcomes of risky decisions. *Journal of Experimental Child Psychology*, *148*, 1–19. https://doi.org/10.1016/j.jecp.2016.02.008

Muradoglu, M., Dweck, C. S., & Cimpian, A. (2025). Children's ability concepts: Their development, content, and consequences. In *Advances in Child Development and Behavior* (Vol. 69, pp. 161–191). Elsevier. https://doi.org/https://doi.org/10.1016/bs.acdb.2025.07.001

Nyhout, A., & Ganea, P. A. (2019a). Mature counterfactual reasoning in 4- and 5-year-olds. *Cognition*, *183*, 57–66. https://doi.org/10.1016/j.cognition.2018.10.027

Nyhout, A., & Ganea, P. A. (2019b). The Development of the counterfactual imagination. *Child Development Perspectives*, *13*(4), 254–259. https://doi.org/10.1111/cdep.12348

Pawsey, H., Denison, S., & Friedman, O. (2024). Children use proximity and ability to infer distinct kinds of counterfactual closeness. *Developmental Psychology*, *61*(3), 572–578. https://doi.org/10.1037/dev0001774

Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, *114*(3), 389–404. https://doi.org/10.1016/j.jecp.2012.10.010

Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin*, *121*(1), 133–148. https://doi.org/10.1037/0033-2909.121.1.133

Stragà, M., Faiella, A., Santini, I., & Ferrante, D. (2023). "The game would have been better for me if…": Children's counterfactual thinking about their own performance in a game. *Thinking & Reasoning*, *29*(4), 663–697. https://doi.org/10.1080/13546783.2022.2130428

Teigen, K. H. (1995). How good is good luck? The role of counterfactual thinking in the perception of lucky and unlucky events. *European Journal of Social Psychology*, *25*(3), 281–302. https://doi.org/10.1002/ejsp.2420250304

Uller, C., Urquhart, C., Lewis, J., & Berntsen, M. (2013). Ten-month-old infants' reaching choices for "more": The relationship between inter-stimulus distance and number. *Frontiers in Psychology*, *4*, Article 84. https://doi.org/10.3389/fpsyg.2013.00084

Van Hoeck, N. (2015). Cognitive neuroscience of human counterfactual reasoning. *Frontiers in Human Neuroscience*, *9*, Article 420. https://doi.org/https://doi.org/10.3389/fnhum.2015.00420

Weisberg, D. P., & Beck, S. R. (2010). Children's thinking about their own and others' regret and relief. *Journal of Experimental Child Psychology*, *106*(2–3), 184–191. https://doi.org/10.1016/j.jecp.2010.02.005

Yang, F., & Frye, D. (2016). Early understanding of ability. *Cognitive Development*, *38*, 49–62. https://doi.org/10.1016/j.cogdev.2016.01.003

Yonas, A., Granrud, C. E., & Pettersen, L. (1985). Infants' sensitivity to relative size information for distance. *Developmental Psychology*, *21*(1), 161–167.

Appendix A

Check of Children's Understanding of Animals' Abilities

Here I aimed to determine whether 4 and 5-year-olds know which animal in my races was the fastest animal. That is, did the youngest children in my study know cheetahs are faster than pigs and sheep and monkeys are faster than cats and raccoons?

Methods

My final sample included 20 children aged 4-5 years (M_{age} = 5;0, range = 4;1 - 5;10, 13 girls and 7 boys). One additional child was seen but excluded due to interference.

Children saw an image of three animals lined up. The first set of three animals were lined up on a road and children were asked "which animal can run the fastest?". The next three animals were lined up at the bottom of a tree and children were asked "Which animal can climb the fastest?"

Discussion

Children responded correctly 95% of the time for the running trial, and 80% of the time for the climbing trial. Of the few errors children made, most of them were from 4-year-olds. Only one of the five errors was made by a 5-year-old. Furthermore, children who made an error still responded correctly to the other trial. Together, this suggests that 4 and 5-year-olds have the knowledge to identify the high-ability animals in my study.

Appendix B
Alternate Analyses

In Experiments 1 and 2, I initially ran GEE models for binomial data, but replaced them with models for poisson data. I was concerned about ceiling effects because nearly all children in the Almost conditions selected the high-ability animal. Here I report the original binomial models.

Experiment 1 Binomial GEE

A binomial GEE found a main effect of condition, F (1, 318) = 40.58, $p < .001$ and a main effect of age, $F(1, 318) = 15.21$, $p < .001$. There was no interaction between age and condition, $F(1, 318) = 0.16$, $p = .688$. To understand these results, I examined age by condition and found an effect of age for both Almost, $F(1, 318) = 8.91$, $p = .003$ and Easily conditions, $F(1, 318) = 7.04$, $p = .008$. The significant results in the Almost condition were indicative of ceiling effects given that nearly all children picked the second-place animal.

Experiment 2 Binomial GEE

A binomial GEE found marginally significant main effects for condition, $F(1, 242) = 3.87$, $p = .050$ and age, $F(1, 242) = 3.79$, $p = .053$ and no interaction between age and condition $F(1, 242) = 0.02$, $p = .883$. I also examined age by condition and found a significant effect of age in the Should condition, $F(1,242) = 8.05$, $p = .005$, but not the Almost condition, $F(1, 242) = 0.94$, $p = .334$.

Appendix C

Survey of Adults' Beliefs about Animals' Abilities

For each story in my study, I needed an animal people believed was especially fast and two others they believed were slower and similarly matched. I considered telling stories featuring running, climbing, swimming, and jumping competitions. I showed adults the animals I considered featuring in my competitions and asked them to rate each animal based on its ability.

Methods

I sent 18 adults (mostly undergraduates) an excel file that contained prompts and provided space for their responses. Participants rated animals based on how quickly they could complete a running race, climbing race, swimming race, flying race, or jumping contest. For example, for the climbing race participants were asked "How fast can each animal finish a climbing race?" Emojis of the animals[5] they were meant to rate were in the leftmost column. Participants rated each animal on a scale of 0 to 7 (1 = extremely slow, 7 = extremely fast, 0 = cannot complete the race) in the cell next to it, which was highlighted green. I asked about 22 animals for running, 11 animals for climbing, 13 animals for swimming, 13 animals for flying, and 14 animals for jumping. I asked participants to rate some animals on multiple abilities if I was considering that animal as a candidate for more than one competition. For example, I asked participants to rate the tiger on how fast it can run and how high it can jump.

Discussion

To understand how fast participants thought each animal was and the level of agreement between participants, I examined the mean and standard deviation of the

[5] All emojis showed the animal's whole body. To maintain a consistent style, I mostly used Apple emojis

responses for each animal; see Table C1 for running and climbing ratings. I also created histograms to visually represent the variance of responses; see Figures C1 for running ability and C2 for climbing ability.

Even though I asked adults about five abilities, I only showed children a running race and a climbing race. In the first two experiments, I chose the cheetah (M = 6.94, SD = 0.24), pig (M = 2.56, SD = 0.86), and sheep (M = 2.94, SD = 0.80) to compete in a running race and the monkey (M = 6.17, SD = 0.86), raccoon (M = 4.28, SD = 0.96), and cat (M = 4.56, SD = 1.20) to compete in a climbing race. In the final study I added a fourth animal to each race that received a rating similar to the two slower animals featured in the previous races. For the running race I added a cow (M = 3.50, SD = 1.15) and for the climbing race a mouse (M = 3.39, SD = 1.72).

Table C1

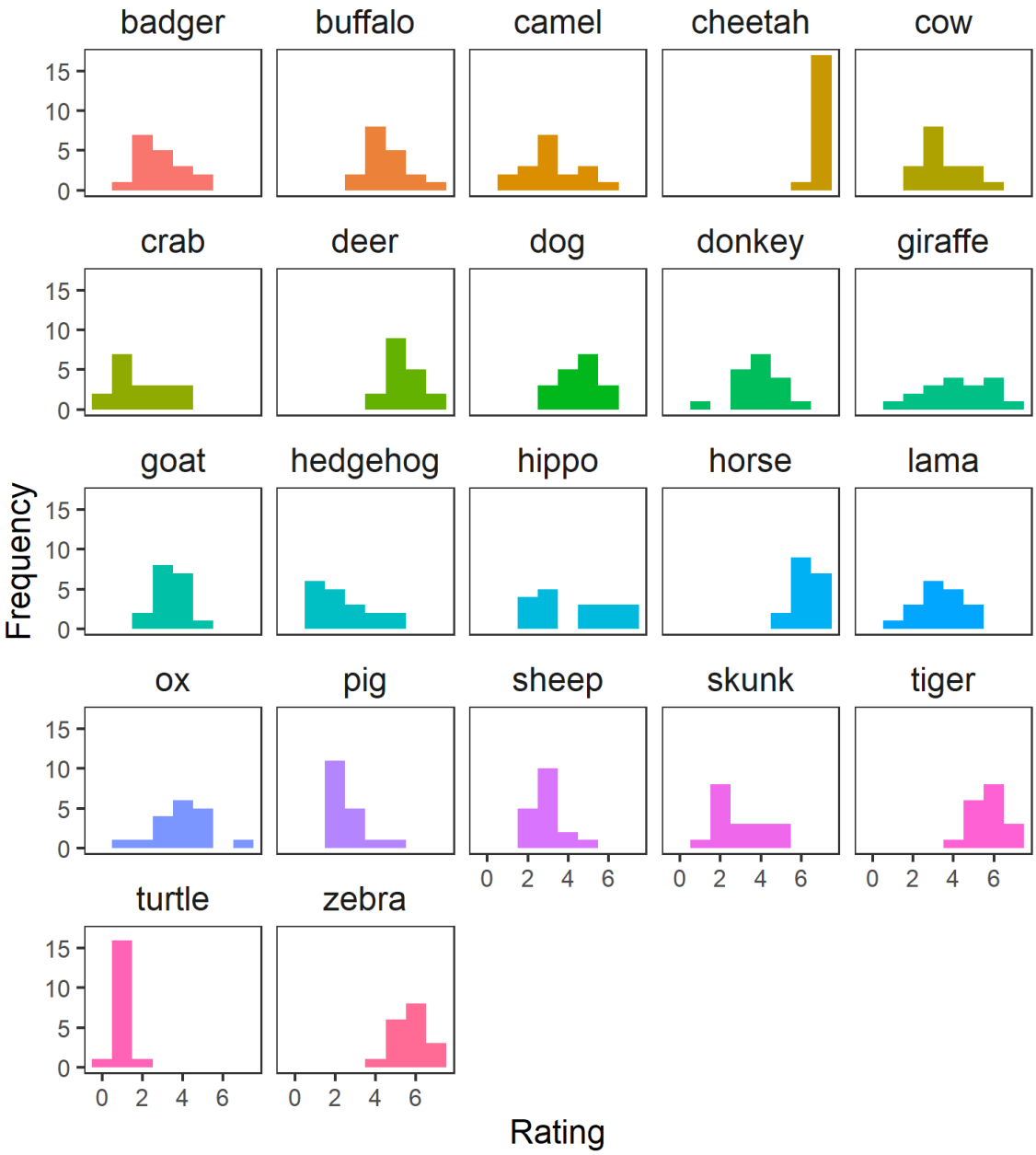
Mean and Standard Deviation of the Animals in Running and Climbing Races

Running			Climbing		
Animal	M	SD	Animal	M	SD
Horse	6.28	0.69	Chipmunk	5.94	1.26
Turtle	1.00	0.34	Cat	4.56	1.20
Dog	4.56	0.98	Raccoon	4.28	0.96
Tiger	5.72	0.83	Monkey	6.17	0.86
Cheetah	6.94	0.24	Gorilla	5.39	1.42
Zebra	5.72	0.83	Mouse	3.39	1.72
Donkey	3.89	1.13	Gecko	4.72	1.67
Deer	5.39	0.85	Orangutan	5.56	1.20
Bison	4.56	1.04	Hedgehog	1.17	1.58
Bull	3.94	1.35	Skunk	2.06	1.59
Cow	3.50	1.15	Badger	1.83	1.47
Pig	2.56	0.86			
Sheep	2.94	0.80			
Goat	3.39	0.78			
Camel	3.22	1.40			
Llama	3.33	1.14			
Giraffe	4.22	1.66			
Hedgehog	2.39	1.38			
Skunk	2.94	1.26			
Badger	2.89	1.13			
Crab	1.89	1.32			
Hippo	4.28	1.90			

Note: Results are rounded to two decimal places.

Figure C1

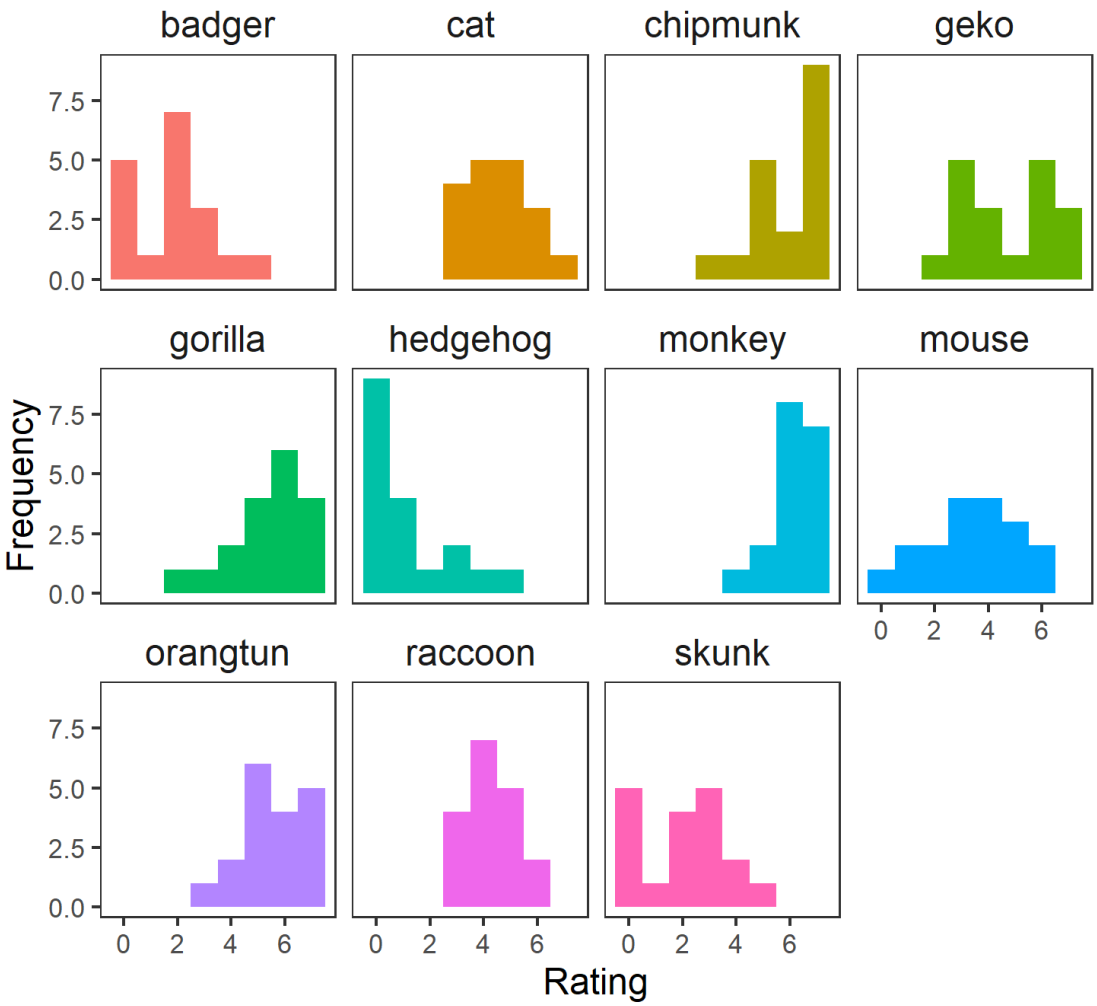
Histograms for Running Ability



Note: Animals were rated on a scale of 0 to 7 (1 = extremely slow, 7 = extremely fast, 0 = cannot compete in the race). Ratings are displayed on the x-axis and the number of times an animal received a given rating is displayed as frequency on the y-axis.

Figure C2

Histograms for Climbing Ability



Note: Animals were rated on a scale of 0 to 7 (1 = extremely slow, 7 = extremely fast, 0 = cannot compete in the race). Ratings are displayed on the x-axis and the number of times an animal received a given rating is displayed as frequency on the y-axis.