Title

Euclid's Random Walk: Developmental changes in the use of simulation for geometric reasoning

Authors

Yuval Hart[1,2], L. Mahadevan[2,3,4,5], and Moira R. Dillon[6*]

Affiliations

[1]Department of Psychology, The Hebrew University of Jerusalem
[2]Paulson School of Engineering and Applied Sciences, Harvard University
[3]Department of Physics, Harvard University
[4]Center for Brain Science, Harvard University
[5]Department of Organismic and Evolutionary Biology, Harvard University
[6]Department of Psychology, New York University

*Correspondence should be addressed to:
Moira R. Dillon (moira.dillon@nyu.edu)

Acknowledgments

This work was supported by a National Science Foundation CAREER Award (DRL-1845924; to M. R. D.), a Jacobs Foundation Early Career Fellowship (to M. R. D.), a SEED Award from The Institute of Human Development and Social Change at NYU (to M. R. D), and the Mind, Brain, and Behavior Interfaculty Initiative at Harvard University (to Y. H. and L. M.). We thank O. Garcia, H. Huey, E. Schwartz, A. Kessler, N. Loncar, B. Reilly, the National Museum of Mathematics, and the museum's participating visitors. The authors declare no conflicts of interest.

Open Practices Statement

As specified in the main text, the design, protocol, and analysis plan for this study were preregistered prior to data collection on the Open Science Framework. Unplanned analyses are also specified in the text. The data and analysis code are publicly available, and the stimuli are available upon request. All of the materials are accessible at: https://osf.io/cxvz7

Abstract

Euclidean geometry has formed the foundation of architecture, science, and technology for millennia, yet the development of human's intuitive reasoning about Euclidean geometry is not well understood. The present study explores the cognitive processes and representations that support the development of intuitive reasoning about Euclidean geometry. One-hundred-twenty-five 7-12-year-old children and 30 adults completed a localization task in which they visually extrapolated missing parts of fragmented planar triangles and a reasoning task in which they answered verbal questions about the general properties of planar triangles. While basic Euclidean principles guided even young children's visual extrapolations, only older children and adults reasoned about triangles in ways that were consistent with Euclidean geometry. Moreover, a relation between visual extrapolation and reasoning appeared only in older children and adults. Reasoning consistent with Euclidean geometry may thus emerge when children abandon incorrect, axiomatic-based reasoning strategies and come to reason using mental simulations of visual extrapolations.

Keywords

spatial cognition; mathematical cognition; Euclidean geometry; reasoning; simulation; computation

1. Introduction

Our reasoning about everyday physical events, like how forces affect object trajectories, may be most successful when we consider how such events unfold over time (e.g., Battaglia, Hamrick, & Tenenbaum, 2013; Sanborn, Mansinghka, & Griffiths, 2013; Smith & Vul, 2013). For example, when asked what would happen if a ball attached to a string whirling around in a circle were suddenly released, about ⅓ of adult participants in one classic study incorrectly thought that the ball would continue on a curved, rather than straight, path (McCloskey, Caramazza, & Green, 1980, see also Caramazza, McCloskey, & Green, 1981; McCloskey, 1983; Proffitt & Gilden, 1989). But when given animated displays of the whirling ball versus static displays or linguistic descriptions, participants were more likely to choose the correct, linear trajectory than the incorrect, curved one (Hegarty, 2004; Smith, Battaglia, & Vul, 2018; Kaiser, Proffitt, Whelan, & Hecht, 1992).

While successful reasoning about the spatial and geometric properties of such dynamic physical events may naturally lend itself to mental simulations, what of successful reasoning about geometry itself, a mathematical cornerstone for physics and much of human achievement? Do such dynamic simulations play any role in our reasoning about the properties of static, immutable geometric objects, like planar triangles? Problems in geometry instead seem best answered by immediate inference (like Bhāskara’s seeing-is-knowing “Behold” proof of the Pythagorean theorem) or by step-by-step proof rooted in axiomatic deduction (like Euclid’s *Elements* 1.47 for the same theorem). But without Bhāskara’s brilliance or Euclid’s elements, what describes our intuitive reasoning about triangles?

Much prior work has addressed the role of visual imagery and visual routines for judgments about physical spatial entities (e.g., Mitrani & Yakimoff, 1983; Shepard & Metzler, 1971; Ullman, 1984; Weintraub & Virsu, 1972). Nevertheless, it remains unknown whether such visual and mental processes might also support our more general reasoning about abstract spatial entities, like those that underlie formal geometry. Evaluating this link is important not only for our understanding of geometry as a central cognitive achievement of the human mind but also for our development of effective geometry pedagogies, which traditionally communicate geometric abstractions through language, proofs, or static diagrams (Calero, Shalom, Spelke, & Sigman, 2019; Carraher, Schliemann, & Carraher, 1988; Duval, 2006; González & Herbst, 2013; Herbst & Brach, 2006; Zaslavsky, 2010; Zodik & Zaslavsky, 2007).

Recent work by Hart et al. (Hart, Dillon, Marantan, Cardenas, Spelke, & Mahadevan, 2018) has begun to address the possibility that dynamic mental simulations described by particular spatial properties might indeed support mature geometric intuitions used during reasoning about Euclidean objects, like planar triangles. In this study, adult participants tested in the laboratory and on Amazon Mechanical Turk were presented with a series of fragmented planar triangles varying greatly in size and were asked to use a mouse to drag a dot to the missing vertex of the triangles. Participants produced third corner locations that both underestimated the true vertex location and also were strikingly more accurate than those that would be produced if they had attempted one instantaneous, straight-line extrapolation from each of the given two corners with a noisy representation of the angle sizes (Mitrani & Yakimoff, 1983). Hart et al. (2018) thus modeled participants' localizations using a correlated random walk composed of two competing processes: one that maintained local, smooth motion; and another that globally corrected this motion's direction by the given angle sizes. Participants' localization accuracy was overall scale-dependent (error grew as triangles grew) because of the local noise associated with the random walk. Nevertheless, the global correction process inherently persevered the basic Euclidean principle of scale-invariant angle representations because extrapolations were corrected at a constant timescale as they unfolded. This model was able to account both for participants' underestimation of a triangle's missing vertex and also for the striking accuracy of their localizations.

Hart et al. (2018) also evaluated the relation between this model of participants' localizations and their reasoning about the general properties of triangles. A different group of adult participants on Amazon Mechanical Turk produced verbal judgments about the location and angle size of a triangle's missing corner after reading verbal descriptions of changes to the other two corners (e.g., "What happens to the angle size of the third corner of a triangle when the other two angles get smaller? Does the third corner angle size get bigger, get smaller, or stay the same size?"). Participants responded more accurately and more quickly when the described transformation resulted in a smaller versus larger triangle, suggesting that they were relying on a reasoning process that, like their localizations, was scale-dependent and tied to particular physical exemplars. Moreover, the model of the first group of participants' localizations predicted the categorical responses of the second group. Hart et al. (2018) speculated that adults might actively engage in mental simulation of these visual extrapolations to answer verbal reasoning questions about static geometric figures.

This work highlights, but does not directly address, several persistent questions about human geometric reasoning, including how formal education and individual development might affect the intuitive strategies humans adopt during geometric reasoning. Prior cross-cultural research testing children and adults from the United States, France, and a remote Amazonian village (Izard, Pica, Spelke, & Dehaene, 2011) and prior developmental research from a laboratory in the United States (Dillon & Spelke, 2018) had used tasks nearly identical to Hart et al. (2018) and found significant changes in geometric reasoning through development. Reasoning consistent with Euclidean geometry emerged universally across human cultures, regardless of formal schooling (Izard, Pica, Spelke, & Dehaene, 2011) at about 10-12-years of age (Dillon & Spelke, 2018; Izard, et al., 2011). While these cross-cultural and laboratory-based studies suggest universal developmental changes in geometric reasoning, they nevertheless provide no evidence of what cognitive processes, representations, or intuitive strategies might underlie those developmental changes. In particular, they do not reveal whether the spatial properties inherent in simple acts of visual triangle completion might be related to explicit judgments about the Euclidean properties of shapes. In the present work, we thus combine computational methods from statistical physics and developmental methods from basic research in cognitive science to examine the relations between visual triangle completion and verbal reasoning about the general properties of planar triangles across samples of children and adults. We speculate that reasoning consistent with Euclidean geometry may emerge in development when children abandon incorrect, axiomatic-based strategies and instead come to reason by an intuitive strategy rooted in mental simulations of visual extrapolations.

2. Methods

2.1. Child Participants

The use of human participants for this study was approved by the Institutional Review Board on the Use of Human Subjects at New York University. A sample size of 125 fluent English-speaking children between the ages of 7-12 was chosen in advance of data collection and was preregistered on the Open Science Framework (OSF). All participants were recruited from visitors to the National Museum of Mathematics in New York City. While the museum welcomes visitors of all ages, their target child age range is eight to eleven years. Most museum visitors reside in New York City or the surrounding suburbs. Most visitors are White, although household incomes varied widely. Museum visitors also likely have a strong interest in mathematics. Despite these specifications of our sample, the tasks in the present study have – rather uniquely – been used in previous studies with diverse populations, as reviewed above, and their results have been unaffected by education or culture. We thus consider the present sample's responses too as representative of the larger population at least in terms of the specific cognitive geometry probed here and in those prior studies.

Several unexpected outcomes related to the sample occurred during data collection. First, we had planned that each whole-year age group would include at least 20 children, but 125 participating children met the inclusion criteria before we could reach 20 children per age group (7 years: 19 children; 8 years: 17 children; 9 years: 28 children; 10 years: 30 children; 11 years: 19 children; 12 years: 12 children). Second, while our exclusion criteria were planned and preregistered, a greater number of children met these exclusion criteria than we had expected. We had planned to include an additional group pf 25 6-year-old children, moreover, apart from the main group of 125 older children. However, their exclusion rate was very high (12 out of the first 25 children tested, 2 for missing data and 10 for response properties in the Localization Task), and so we discontinued data collection with these younger children. In our main sample of 7- to 12-year-old children, an additional 61 children participated but were excluded for: missing data (6); technical failure (1); experimenter error (1); parental interference (1); and the properties of their responses in the Localization Task (52; see **SM; Fig. S1**). This last criterion, which by far led to the most exclusions, was specified in advance and based on Hart et al. (2018), who tested adults individually in the laboratory and presented three times the number of trials compared to the present task. This criterion turned out to have been too strict for the present study (see **SM**), not accounting for the age differences between studies, the more complex testing conditions in the museum compared to the laboratory, and the significantly reduced number of trials. To examine the robustness of our findings to this exclusion criterion, we thus repeated our main analyses as an unplanned analyses with the excluded sample (N = 52; 21 girls; 7 years: 17 children; 8 years: 9 children; 9 years: 10 children; 10 years: 6 children; 11 years: 6 children; 12 years: 4 children), and because those results are consistent with the analysis of the planned sample, we report them in the **SM**.

2.2. Adult Participants

Based on the findings with children presented below, we also tested an unplanned group of 30 adult participants (the maximum number of participants per age group in the child sample) between the ages of 21-36 years. This allowed us to examine whether the unexpected trends we observed in older children described below were also present in adults. An additional 7 adults also participated but were excluded because of the properties of their responses in the Localization Task (see **SM**); no adults met any of our other exclusion criteria. Adult participants were also recruited from visitors to the National Museum of Mathematics and completed the same tasks as children, presented exactly in the same way. None of the adults were participating children's parents or guardians.

2.3. Reasoning Task

The task materials and procedures were determined in advance and preregistered on the OSF. Participants first completed a geometric reasoning task (after Dillon & Spelke, 2018; Hart et al., 2018) that required them to produce verbal, categorical responses about the distance and angle properties of triangles given shape and size transformations to fragmented scalene triangles with only two visible corners (**Fig. 1A**). This task was presented on a large screen (65'' diagonal, 1920px x 1006px) and with the help of an adult experimenter. At the beginning of the task, participants saw a sample fragmented triangle (which never appeared during a test trial), displaying at first just the triangle's two base corners, then the complete triangle, then just the two base corners again. The experimenter then demonstrated what four different possible changes to those visible corners would look like, using a separate display with one button for each of the four possible changes: the visible angles growing in size; shrinking in size; moving apart; or moving together. The sample fragmented triangle had 30° base angles, and its base length was set to 0.7 of the full possible base length (**Table 1**). Although participants were tested only on static fragmented triangles, they could revisit the sample-changes display at any point during the task if they wanted to see those sample changes again. Participants were then told that for each fragmented triangle, they could be asked: whether the triangle's missing corner location would "move up," "move down," or "stay in the same place" after one of these changes; or whether its angle size would "get smaller," "get bigger," or "stay the same size" after one of these changes. To ensure that participants understood what each of these outcomes meant, the experimenter gestured as they described each one. For the location outcomes, the experimenter held one hand at chin height, then moved it up in space, then down in space (below chin height), and then back to chin height. For the angle-size outcomes, the experimenter formed an upside-down "V" shape with their hands, then made the "V" narrower, then wider (wider than its starting width), and then back to its starting width. In addition to providing these gestures during the task's introduction, the experimenter also displayed them during every test question. There were 8 possible questions (4 possible changes to the visible corners × 2 possible outcomes for the missing corner), and each question was presented twice, once per block of 8 questions with two total blocks for each participant. Those 8 questions were randomized within a block and paired with a random fragmented scalene triangle (**Table 1**). The second block presented the same questions but in a different order and with a different random triangle. Participants never saw the same question or triangle presented twice in a row. All images accompanying test questions were created by a custom Javascript code. Participants' responses were recorded by an experimenter's button press.

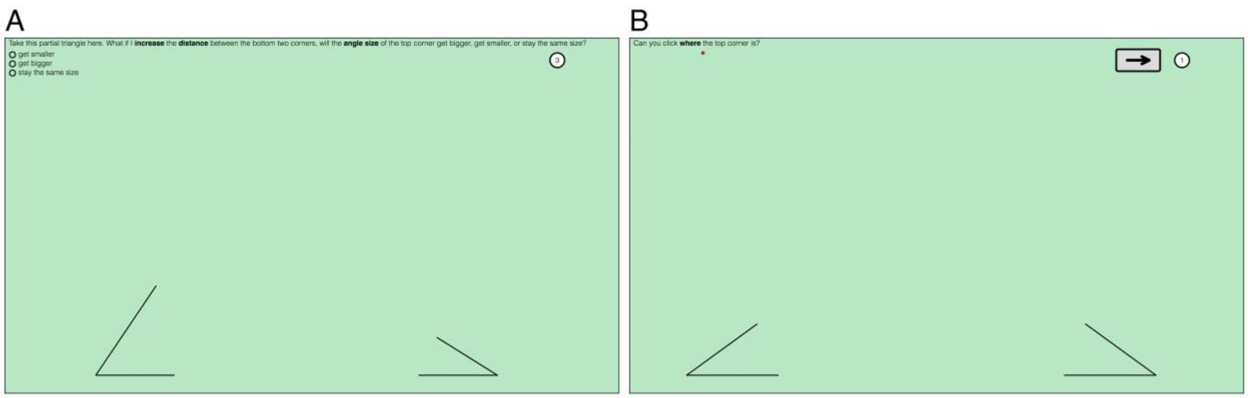

Fig. 1. A. Sample screen, Reasoning Task. The question at the top reads: "Take this partial triangle here. What if I **increase** the distance between the bottom two corners, will the **angle size** of the top corner get bigger, get smaller, or stay the same size?" Participants were provided with a set of scalene triangle corners and asked to make judgments about the third, missing corner after changes to the given corners. **B.** Sample screen, Localization Task. The question at the top reads: "Can you click **where** the top corner is?" Participants were provided with a set of isosceles triangle corners and asked to drag a dot to the vertex of the missing corner.

Table 1.

Properties of the triangle fragments presented in the Reasoning Task

Triangle	Base Length	Right base angle	Left base angle	Triangle size [area]
1	0.44	48°	32°	3.87
2	0.66	32°	40°	7.8
3	0.77	32°	56°	13.03
4	0.55	40°	32°	5.41
5	0.77	56°	48°	18.82
6	0.66	40°	48°	10.41
7	0.44	56°	40°	5.19
8	0.55	48°	56°	9.6

Note. (1 length unit = 1632 px [1920px x 0.85])

2.4. Localization Task

The task materials and procedures were determined in advance and preregistered on the OSF. Participants completed the Localization Task (after Hart et al., 2018; Izard et al., 2011) following the Reasoning Task. At the beginning of the task, participants again saw the sample fragmented triangle, displaying at first just the triangle's two base corners, then the complete triangle, then just the two base corners again. Participants were told that they would see more partial triangles and would be asked to use the mouse to click on the vertex location of the triangle's missing top corner. To ensure that participants understood the task, they completed one trial with this sample triangle. For the test trials, participants saw 49 fragmented triangles (**Fig. 1B**) and were asked to click on the location of a triangle's missing vertex. They received no feedback. Seven isosceles triangles were presented, which had 7 different side-length values combined with 2 angle sizes and 4 base lengths (**Table 2**). The presentation of these triangles was pseudo-random for each participant, not allowing the same triangle to be presented twice in a row. All participants used a single-button, child-sized mouse, and their responses were recorded based on where they clicked on the screen; reaction times were also recorded. All images accompanying test questions were created by a custom Javascript code.

Table 2.

Properties of the triangle fragments presented in the Localization Task

Triangle	Base Length	Base angles	Triangle size [area]
1	0.9	36°	14.7
2	0.4	36°	2.9
3	0.1	36°	0.18
4	0.04	36°	0.03
5	0.4	45°	4
6	0.1	45°	0.25
7	0.04	45°	0.04

Note. (1 length unit = 1632 px [1920px x 0.85])

3. Results

3.1. Child Results

3.1.1. Planned Analyses.

The following analyses were specified prior to data collection and preregistered on the OSF.

Reasoning Task.

First, a binomial mixed-model logistic regression revealed a significant effect of gender on children's overall accuracy, with boys performing better than girls ($P = 0.579$, 95% CI = [0.501, 0.653], $p = 0.048$). As planned, all analyses were thus repeated with gender as an additional predictor variable, but because those results were consistent with our primary planned analyses, they are reported in the **SM**.

A binomial mixed-model logistic regression evaluated the role on children's accuracy of: question type (about the position versus angle size of the missing corner); transformation (to the distance between the two given corners or their angle sizes); size of the transformation (whether the two given corners were described as getting farther/bigger versus closer/smaller); the two-way interactions between these variables; the implied area of the fragmented triangle; and age. As predicted, this regression revealed results consistent with prior studies (Dillon & Spelke, 2018; **Fig. 2**). In particular, children were more accurate on questions about the position versus angle size of the fragmented triangle's missing corner (*P* = 0.746, 95% CI = [0.672, 0.808], *p* < *.001*) and when there was a transformation to the angle sizes versus the distance between the two given corners (*P* = 0.716, 95% CI = [0.637, 0.783], *p* <*.001*). Children were also more accurate when they were asked about the position versus angle size of the missing corner after a distance transformation to the two given corners (*P* = 0.692, 95% CI = [0.597, 0.772], *p* < *.001*). Neither the size of the transformation (bigger or smaller) nor the implied area of the fragmented triangle presented with each question (continuous, in area units, see **Table 1**) affected children's accuracy (*p*s > *.490*). Finally, older children were more accurate on this task than younger children (age, in days, was treated as a continuous variable in this analysis; *P* = 0.538, 95% CI = [0.507, 0.568], *p* = *.016*).

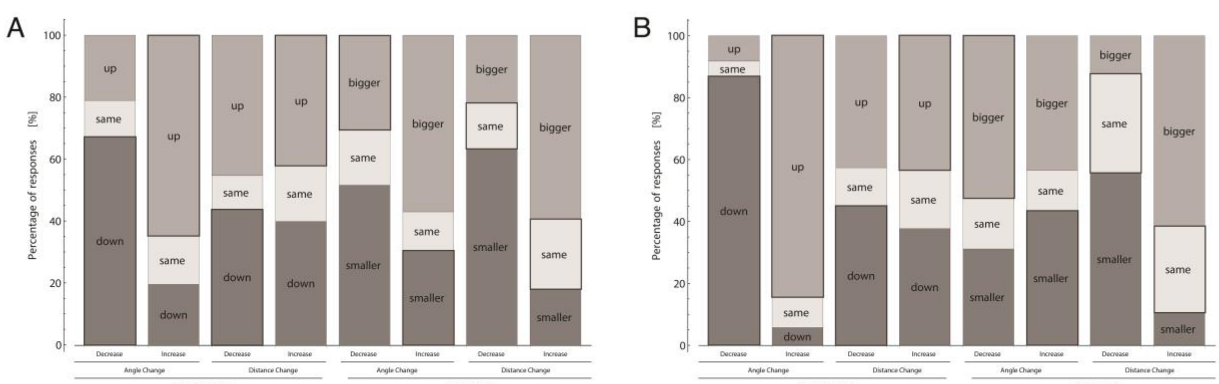

Fig. 2. The percentage of **A**. younger (< 10 years) and **B**. older (≥ 10 years) children's responding in the Reasoning Task about the general properties of triangles. Children were asked

to reason about changes to the position and angle size of a missing corner of incomplete triangles after changes to the angle sizes or distances between the two given corners (see **Fig. 1**).

Localization Task.

For each child and for each of the 7 triangle side lengths, we calculated the localization error in the y direction (the true vertex location – the mean of the child's estimates) and the standard deviation in the y direction of the child's estimates. Using a linear regression, we first evaluated the growth in each child's error with growing triangle side lengths. We then evaluated, across the sample of children, the relation between error growth by side length and age using a linear regression. As predicted, across the sample of children, error grew significantly as triangle side-length grew ($p < .001$), suggesting an overall scale dependence in children's visual extrapolations of the triangles' missing parts. Moreover, as predicted, we found that the error grew less in older versus younger children ($p = .039$).

After Hart et al. (2018), we then evaluated the slope of the log of the standard deviation of each child's localization estimates as a function of the log of triangle side length. This slope, or *scaling exponent*, is equivalent to the power law by which the standard deviation of the estimates scales with triangle side-length. The scaling exponent represents one of the two competing processes in the correlated-random-walk model described above, which characterizes the extrapolation process. It represents the global correction of the local noise associated with maintaining smooth motion in the direction of the given angle sizes.

(1) $$d^2\theta/dt^2 = 1/\tau(1/\xi(\theta - \theta_0) - d\theta/dt) + \eta(t)$$

(2) $$dx/dt = v_p\, cos(\theta)$$

(3) $$dy/dt = v_p\, sin(\theta)$$

The model parameters include τ, an inertial relaxation timescale for local smoothness, v_p, a characteristic speed of extrapolation progress, ξ, a timescale for the global error correction, and η(t), a noise term. The more correction events that occur, the closer the scaling exponent is to 0.5 versus 1. Scaling exponents less than 1 suggest that correction events are occurring, and scaling exponents closer to 0.5 suggest that correction events are occurring at a more frequent timescale. Extrapolations with scaling exponents close to 0.5 thus better preserve the angle sizes of the triangle's given corners, allowing greater consistency with Euclidean geometry.

We predicted that our data would be well described by this model, yielding localization errors that underestimated the true vertex location and scaling exponents that were less than 1. We also predicted that since older children more consistently *reason* in line with Euclidean geometry (as revealed by prior work), their *localizations* would also better reflect Euclidean geometry, resulting in smaller scaling exponents.

Consistent with the model from Hart et al. (2018), children tended to underestimate the location of a triangle's vertex (**Fig. 3**; **Fig. S2**) and most of their scaling exponents were less than 1: Children produced a median scaling exponent of 0.83 (95% CI = [0.80, 0.86], range = [0.56, 1.14]). Contrary to our prediction, however, the relation between scaling exponent and age was not significant ($p = .666$): We did not find evidence that older children corrected their visual extrapolations more than younger children.

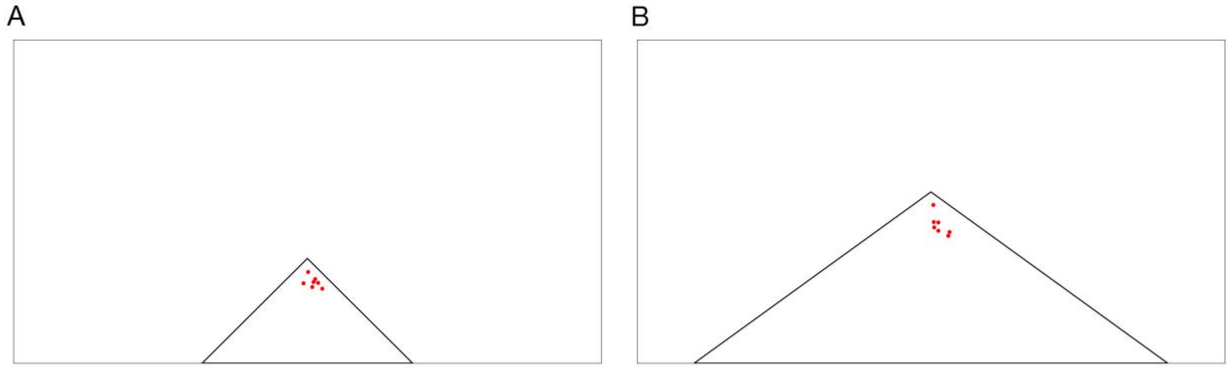

Fig. 3. Example responses from an 8-year-old child on the Localization Task on **A.** a smaller triangle with 0.4 times the base-length metric and 45° angles and **B.** a larger triangle with 0.9 times the longest base length and 36° angles.

Relation between reasoning and simulation.

Children's accuracy in the Reasoning Task and Localization Task may nevertheless rely on properties inherent to Euclidean geometry. We thus hypothesized that individual children's scaling exponents would be related to their individual reasoning success such that the more frequently a child corrected their visual extrapolations in the Localization Task, the greater their accuracy in the Reasoning Task. This relation would be especially evident in older children, moreover, who may more often adopt a strategy of mentally simulating visual extrapolations during reasoning.

First, a binomial mixed-model logistic regression across the entire sample of children probing the relation between scaling exponent and reasoning accuracy was not significant (P = 0.271, 95% CI = [0.097, 0.562], $p = .117$). Nevertheless, this first analysis did not take into account the difference in reasoning accuracy for older versus young children. An additional binomial mixed-model logistic regression predicting accuracy by scaling exponent, age, and their interaction did not provide evidence that age moderated the relation between scaling exponent and reasoning (Scaling Exponent: $P = 0.995$, 95% CI=[0.038, 1], $p = .226$; Age: $P = 0.660$, 95% CI=[0.486,0.800], $p = .071$; Scaling Exponent *Age: $P = 0.344$, 95% CI=[0.182, 0.554], $p = .142$).

3.1.2. Unplanned Analyses.

Relation between reasoning and simulation.

To better understand the relation between reasoning and simulation and the differences between younger versus older children beyond what we could infer from the two planned analyses, we conducted two additional unplanned analyses. First, we repeated the same regressions as in the planned analysis, but this time treated children below 10 years of age (N = 64) and above 10 years of age (N = 61) as different groups. This decision was motivated by prior results from the literature on children's and adults' geometric reasoning across cultures: Prior studies had indicated 10 years of age as approximately the age at which reasoning becomes conformal with Euclidean geometry (Dillon & Spelke, 2018; Izard et al., 2011). This age split, as opposed to the continuous treatment of age in our moderation analysis, may better capture the developmental changes in children's reasoning, especially if there is not much change in reasoning before age 10 years and not much change in reasoning after age 10 years. In addition to splitting the sample based on the findings and conclusions of prior work, we also conducted a change-point analysis on children's accuracy on our Reasoning Task, with age binned by month and using a binary segmentation method (Scott & Knott, 1974) with a Bayesian Information Criterion (BIC) penalty type. We found one change point at 10 years 3 months (**Fig. S3**). As a test of robustness, we thus repeated our analysis using this age split, and because it revealed results consistent with the split at 10 years, we report those results in the **SM**.

First, a binomial mixed-model logistic regression predicting reasoning accuracy by scaling exponent, age (≥ 10 years versus <10 years), and their interaction found no significant effect of scaling exponent ($P = 0.492$, 95% CI = [0.176, 0.814], $p = .966$) but a significant effect of age ($P = 0.920$, 95% CI = [0.613, 0.988], $p = .016$). This analysis was further characterized by a scaling exponent by age interaction ($P = 0.093$, 95% CI = [0.010, 0.522], $p = .059$). Individual contrasts revealed no relation between scaling exponent and reasoning for younger children ($P = 0.492$, 95% CI = [0.176, 0.815], $p = .966$), but a significant relation between scaling exponent and reasoning for older children ($P = 0.090$, 95% CI = [0.016, 0.381], $p = .013$).

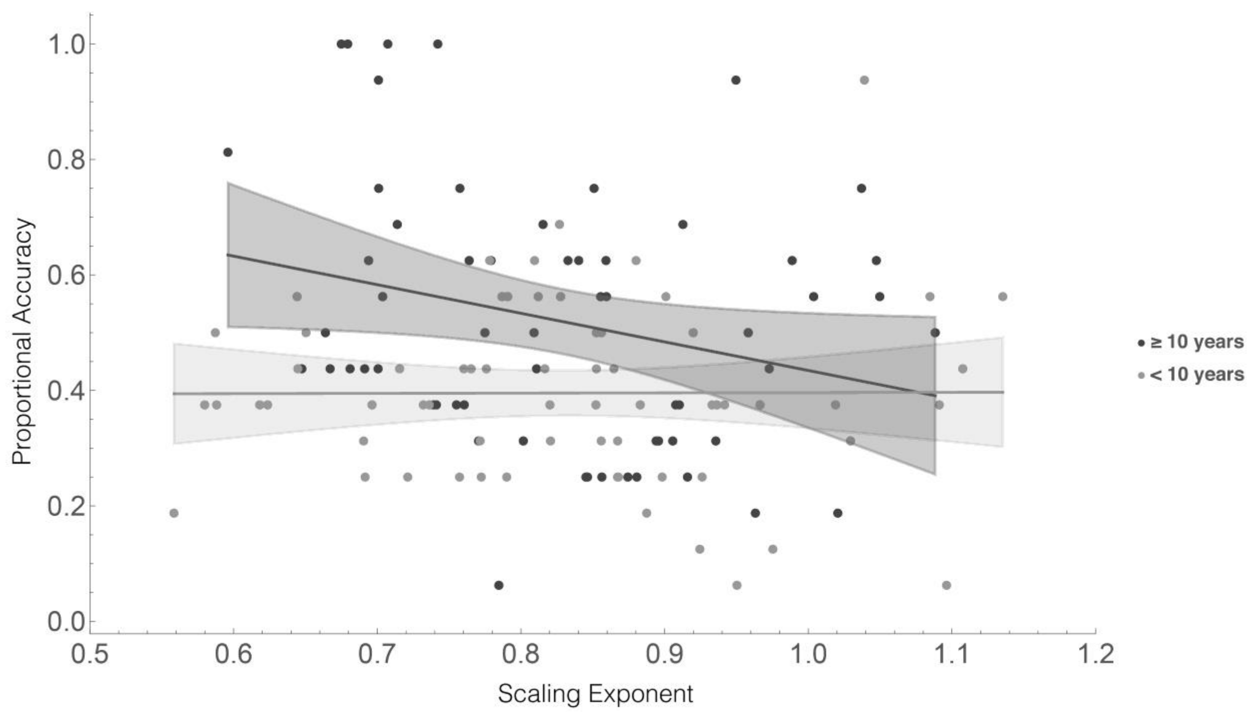

Fig. 4. The relation between the scaling exponent from the Localization Task and accuracy in the Reasoning Task across younger (< 10 years, light grey) and older (≥ 10 years, dark grey) children, 95% CIs are depicted for each regression line.

We next explored whether this result was due to differences in effort or motivation in younger versus older children. In particular, if the hardest working or most motivated children were older, corrected their localizations more, and thought more deeply during reasoning, this might lead to both better scaling exponents and more accurate reasoning. If we correct for the time older children took to complete the Localization Task (as a proxy for their effort; reaction time, in seconds, was log-transformed to better align the scales of the variables, allowing for model convergence) and evaluate the relation between scaling exponent and reasoning, we find that the relation persists ($P = 0.080$, 95% CI = [0.009, 0.448], $p = .032$) and that time does not independently predict reasoning ($P = 0.428$, 95% CI = [0.258, 0.617], $p = .495$). The relation between scaling exponent and reasoning in older children is thus not likely due to overall effort or motivation.

Finally, a close investigation of children's responses lent further support to the suggestion that common Euclidean principles drive both visual extrapolation and geometric reasoning in older but not younger children. First, older children tended to produce reasoning responses that, like the extrapolation process, showed some scale dependence, for example, responding more accurately when the transformed triangle was smaller versus larger than the original (**Fig. 2**). Younger children, in contrast, tended to produce reasoning responses that directly conflicted with properties of extrapolation. The majority of younger, but not older, children reasoned, for example, that the missing third angle of a triangle would change *in the same direction as* (as opposed to *inversely to*) the change to the other two angles (**Fig. 2**). Even a very noisy extrapolation of such an angle transformation would be unlikely to yield this response in a majority of children. Thus, older children's reasoning errors were—and younger children's errors were not—consistent with the properties of visual extrapolation.

3.2. Adult Results

3.2.1. Unplanned Analyses.

After seeing these results with children, we collected an additional unplanned, small sample of adult participants to further evaluate two surprising findings, namely that children's scaling exponents, which inherently reflect the Euclidean principle of scale-invariant angle measures: (1) do not improve with age; and (2) are associated with reasoning only at older ages, i.e., when reasoning is conformal with Euclidean geometry.

First, consistent with the findings from the child sample, a linear regression revealed no evidence of an effect of age on the scaling exponent across the entire child and adult sample (P = 0.500, 95% CI = [0.500, 0.501], $p = .303$). To further evaluate this null effect, we conducted a Bayesian regression, which calculated the posterior distribution of slopes characterizing the relation with a region of practical equivalence of -0.005 to 0.005. This analysis suggested that there was no effect of age on the scaling exponent (slope = 0.0015, 95% CI = [-0.0014, 0.0044], posterior probability of the null effect of age = 99.14%).

Second, consistent with the findings with older children, adults' performance on the Reasoning Task was conformal with Euclidean geometry (**Fig. 5**). For adults, as for older children, moreover, individuals' scaling exponents were related to their reasoning success (P = 0.017, 95% CI = [0.0002, 0.62], $p = .080$).

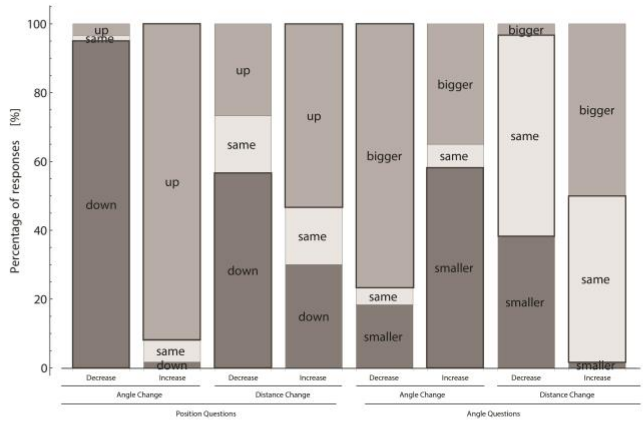

Fig. 5. The percentage of adults' responding in the Reasoning Task about the general properties of triangles.

4. Discussion

Two tasks required children and adults to make judgments about the properties of visually fragmented triangles. The patterns of performance on these tasks suggested both continuity and change in geometric cognition through development. First, a correlated-random-walk model from statistical physics characterized children's localizations of the missing third corners of triangles of different sizes, as it had in prior studies examining adults' localizations. The model revealed that while the random noise associated with triangle-side extrapolation decreased as children got older, the timescale with which they corrected that noise in line with the basic Euclidean principle of scale-independent angle-size information did not change. And so, children may require no explicit knowledge of this Euclidean principle (or its relevance to a visual shape completion task) when extrapolating the missing parts of planar shapes. Instead, basic Euclidean principles guiding visual extrapolation may be present from early in human development, perhaps due to experiences with the continuous edges and surfaces in scenes and objects or to the very structure of our brain systems dedicated to everyday spatial tasks

(Ayzenberg & Lourenco, 2019; Elder & Goldberg, 2002; Feldman, 2001; Field, Hayes, & Hess, 1993; Lee & Yuille, 2006; Walther, Chai, Caddigan, Beck, & Fei-Fei, 2011). Moreover, sensitivities to straight and oriented trajectories for moving through spaces and recognizing objects are observable in infancy and young childhood, even in the absence of typical visual experience (Kellman & Spelke, 1983; Landau, Gleitman, & Spelke, 1981; Slater, Mattock, Brown, & Bremner, 1991), and the tradeoff between maintaining a straight line at a certain angle and maintaining a smooth line with no sharp corrections is even inherent in the navigational abilities of a variety of animal species (Cheung, Zhang, Stricker, & Srinivasan, 2007), including dung beetles (Peleg & Mahadevan, 2016), birds (Wiltschko & Wiltschko, 2005), sharks (Papastamatiou, Cartamil, Lowe, Meyer, Wetherbee, & Holland, 2011), and insects (Wehner, Michel, Antonsen, 1996). Future research exploring whether other animal species incorporate basic Euclidean principles into their visual extrapolations, moreover, could evaluate whether such principles are reflective of our uniquely human capacity to learn geometry, our experiences in the spatial world shared by other animals (e.g., Hubel & Wiesel, 1962, 1965; Rubin, Nakayama, & Shapley, 1996; von der Heydt, Peterhans, & Baumgartner, 1984), or any evolutionarily inherited Euclidean biases in perception and cognition.

Second, the present study found that children's verbal reasoning about the general properties of triangles changed markedly as children got older, consistent with prior studies with diverse populations (Dillon & Spelke, 2018; Izard et al., 2011). In particular, younger children seemed to respond to reasoning questions by simple, though erroneous size-based heuristics that conflicted with Euclidean principles. For example, younger children responded that the missing angle of a fragmented triangle changed in the same direction as (as opposed to inversely to) the change to the other two angles. In contrast, older children and adults tended to respond to questions about the side and angle properties of planar triangles in general accord with formal, Euclidean geometry. Nevertheless, neither older children nor adults were perfectly Euclidean: Both groups showed some scale dependence in their reasoning, for example, by responding more accurately when the described transformations to the triangles made triangles smaller versus bigger. This was true even though the participants in the present study may have more interest and practice in math compared to others who have been tested in prior studies and others in the general population. Their similar performance to other populations thus further supports the suggestion that some intuitive reasoning about geometry is largely unaffected by culture, education, or even expertise (see, e.g., Amalric & Dehaene, 2016; 2018; Butterworth, 2006).

The present work also addresses two questions about the cognitive mechanisms underlying human geometric reasoning that prior work had not been able to address: What developmental change in cognitive representations and processes might underlie a change in reasoning from incorrect and axiomatic to nearly Euclidean? And what would it mean for our understanding of human intuitive cognitive geometry to qualify this reasoning as *nearly* Euclidean? While prior work had speculated that older children naturally become "little Euclids," reasoning by intuitive knowledge of geometric rules (e.g., Dillon & Spelke, 2018; Izard et al., 2011), the present work instead suggests that older children and adults fall short of reasoning that is perfectly consistent with formal, Euclidean geometry. Instead, older children and adults appear to engage only some Euclidean principles during simple tasks of visual triangle completion and during verbal tasks of explicit geometric reasoning. We suggest, therefore, that older children and adults may perform better on tasks of Euclidean reasoning not because they become "little Euclids," but because they adopt an intuitive reasoning strategy that relies on the mental simulations of their visual extrapolations, which include some Euclidean elements.

Developmental discontinuity in Euclidean reasoning may thus emerge when children abandon axiomatic strategies and begin to engage in dynamic simulations to solve novel geometric reasoning problems. For older children and adults, moreover, the strength of the Euclidean elements guiding these simulations may contribute to their individual success in reasoning in accord with Euclidean geometry.

Given the correlational design of the present study as well as some unplanned analyses, this suggestion is speculative. Nevertheless, the present work raises new questions for future exploration. For example, if simulation is a relatively effective intuitive strategy for geometric reasoning that older children and adults rely on, and younger children's extrapolations already incorporate basic Euclidean properties that are predictive of reasoning success, then why do younger children not engage in simulation during reasoning? One possibility is that younger children do not recognize the relevance of their simulations to the reasoning problem. Simply telling a younger child to dynamically imagine the missing parts of and the transformations to fragmented triangles during a reasoning task might thus make their performance look more like older children's. Instruction to imagine the dynamic unfolding of physical events has improved, for example, even young children's reasoning about the trajectories of balls moving through opaque tubes (e.g., Joh, Jaswal, & Keen, 2011; Palmquist, Keen, & Jaswal, 2017). Future studies using such explicit verbal instruction or implicit priming could begin to evaluate both whether mental simulation of visual extrapolations about geometry and its static planar figures is available to younger children as a reasoning strategy and whether such simulation is causally related to reasoning success.

Another possibility for why younger children may not engage in simulation for reasoning is that limits to younger children's memory and attention, in general, or other properties of their simulations, in particular, may affect their ability to engage in simulation as a reasoning strategy. For example, while there were many similarities between older and young children's visual extrapolations in the Localization Task, engaging in mental simulation of these visual extrapolations for reasoning requires both visualizing a transformation to a given triangle and also performing extrapolations on that imagined triangle. Our current tasks do not examine whether younger and older children might differ in such abilities. Moreover, younger children had more local noise in their simulations than older children. Future studies might begin to explore whether introducing noise into the displays accompanying reasoning questions for older children and adults might lead them to adopt language-based heuristics instead of simulation-based strategies for solving reasoning problems (see Perfecto, Donnelly, & Critcher, 2019). Such studies could lead to the investigation of how individuals decide, more generally, whether reasoning by language-based heuristics or simulation might be more or less effective when faced with novel problems in geometry, mathematics, or other domains. Moreover, such findings could ultimately inform pedagogies aimed at teaching and testing geometric formalisms, rules, and abstractions.

While problems in geometry may seem best answerable by immediate inference or deductive proof, intuitive geometric reasoning may instead rely on noisy, dynamic simulations. The achievements enabled by Euclidean geometry are manifest throughout human history, and Euclidean geometry has often been held up as *the* model of abstract thought. And yet our findings suggest that Euclid himself, like the rest of us, may have taken quick random walks in his mind before he plodded step by step on the printed page.

References

Amalric, M., & Dehaene, S. (2018). Cortical circuits for mathematical knowledge: evidence for a major subdivision within the brain's semantic networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*, 20160515.

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences*, *113*, 4909-4917.

Ayzenberg, V., & Lourenco, S. F. (2019). Skeletal descriptions of shape provide unique and privileged perceptual information for object recognition. *BioRxiv*, 518795. https://doi.org/10.1101/518795

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332. https://doi.org/10.1073/pnas.1306572110

Benavoli, A., Mangili, F., Ruggeri, F., & Zaffalon, M. (2015). Imprecise Dirichlet Process With Application to the Hypothesis Test on the Probability That X ≤ Y. *Journal of Statistical Theory and Practice*, *9*(3), 658–684. https://doi.org/10.1080/15598608.2014.985997

Butterworth, B. (2006). Mathematical expertise. In K.A. Ericsson, N. Charness, P.J. Feltovich, & R.R. Hoffmann (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 553–568). Cambridge: Cambridge University Press

Calero, C. I., Shalom, D. E., Spelke, E. S., & Sigman, M. (2019). Language, gesture, and judgment: Children's paths to abstract geometry. *Journal of Experimental Child Psychology*, *177*, 70–85. https://doi.org/10.1016/j.jecp.2018.07.015

Caramazza, A., McCloskey, M., & Green, B. (1981). Naive beliefs in "sophisticated" subjects: misconceptions about trajectories of objects. *Cognition*, *9*(2), 117–123. https://doi.org/10.1016/0010-0277(81)90007-X

Carraher, T. N., Schliemann, A. D., & Carraher, D. W. (1988). Mathematical concepts in everyday life. *New Directions for Child and Adolescent Development*, *1988*(41), 71–87. https://doi.org/10.1002/cd.23219884106

Cheung, A., Zhang, S., Stricker, C., & Srinivasan, M. V. (2007). Animal navigation: the difficulty of moving in a straight line. *Biological Cybernetics*, *97*(1), 47-61.

Dillon, M. R., & Spelke, E. S. (2018). From map reading to geometric intuitions. *Developmental Psychology*, *54*(7), 1304–1316. https://doi.org/10.1037/dev0000509

Duval, R. (2006). A Cognitive Analysis of Problems of Comprehension in a Learning of Mathematics. *Educational Studies in Mathematics*, *61*(1), 103–131. https://doi.org/10.1007/s10649-006-0400-z

Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, *2*(4), 5–5. https://doi.org/10.1167/2.4.5

Feldman, J. (2001). Bayesian contour integration. *Perception & Psychophysics*, *63*(7), 1171–1182.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, *33*(2), 173–193. https://doi.org/10.1016/0042-6989(93)90156-Q

González, G., & Herbst, P. (2013). An Oral Proof in a Geometry Class: How Linguistic Tools Can Help Map the Content of a Proof. *Cognition and Instruction*, *31*(3), 271–313. https://doi.org/10.1080/07370008.2013.799166

Hart, Y., Dillon, M. R., Marantan, A., Cardenas, A. L., Spelke, E., & Mahadevan, L. (2018). The statistical shape of geometric reasoning. *Scientific Reports*, *8*(1), 12906. https://doi.org/10.1038/s41598-018-30314-y

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*(6), 280–285. https://doi.org/10.1016/j.tics.2004.04.001

Herbst, P., & Brach, C. (2006). Proving and Doing Proofs in High School Geometry Classes: What Is It That Is Going On for Students? *Cognition and Instruction*, *24*(1), 73–122. https://doi.org/10.1207/s1532690xci2401_2

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106-154.2.

Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, *28*, 229–289.

Izard, V., Pica, P., Spelke, E. S., & Dehaene, S. (2011). Flexible intuitions of Euclidean geometry in an Amazonian indigene group. *Proceedings of the National Academy of Sciences*, *108*(24), 9782–9787. https://doi.org/10.1073/pnas.1016686108

Joh, A. S., Jaswal, V. K., & Keen, R. (2011). Imagining a way out of the gravity bias: Preschoolers can visualize the solution to a spatial problem. *Child Development*, *82*(3), 744-750.

Kaiser, M. K., Proffitt, D. R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology. Human Perception and Performance*, *18*(3), 669–689.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*(4), 483–524. https://doi.org/10.1016/0010-0285(83)90017-8

Landau, B., Gleitman, H., & Spelke, E. (1981). Spatial knowledge and geometric representation in a child blind from birth. *Science (New York, N.Y.)*, *213*(4513), 1275–1278.

Lee, T. S., & Yuille, A. L. (2006). Efficient coding of visual scenes by grouping and segmentation. In *Bayesian Brain: Probabilistic Approaches to Neural Coding* (pp. 141–185). The MIT Press.

McCloskey, M. (1983). Intuitive Physics. *Scientific American*, *248*(4), 122–130.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear Motion in the Absence of External Forces: Naïve Beliefs About the Motion of Objects. *Science*, *210*(4474), 1139–1141. https://doi.org/10.1126/science.210.4474.1139

Mitrani, L., & Yakimoff, N. (1983). A model of perceived intersection of two converging line segments. *Perception & Psychophysics*, *33*(2), 186–192. https://doi.org/10.3758/BF03202838

Palmquist, C. M., Keen, R., & Jaswal, V. K. (2018). Visualization instructions enhance preschoolers' spatial problem-solving. *British Journal of Developmental Psychology*, *36*(1), 37-46.

Papastamatiou, Y. P., Cartamil, D. P., Lowe, C. G., Meyer, C. G., Wetherbee, B. M., & Holland, K. N. (2011). Scales of orientation, directed walks and movement path structure in sharks. *Journal of Animal Ecology*, *80*(4), 864-874.

Peleg, O., & Mahadevan, L. (2016). Optimal switching between geocentric and egocentric strategies in navigation. *Royal Society Open Science*, *3*(7), 160128.

Perfecto, H., Donnelly, K., & Critcher, C. R. (2019). Volume Estimation Through Mental Simulation. *Psychological Science*, *30*(1), 80–91. https://doi.org/10.1177/0956797618813319

Proffitt, D. R., & Gilden, D. L. (1989). Understanding natural dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(2), 384–393. https://doi.org/10.1037/0096-1523.15.2.384

Rubin, N., Nakayama, K., & Shapley, R. (1996). Enhanced perception of illusory contours in the lower versus upper visual hemifields. *Science (New York, N.Y.)*, *271*(5249), 651–653.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review*, *120*(2), 411–437. https://doi.org/10.1037/a0031912

Scott, A. J. & Knott, M. (1974). A Cluster Analysis Method for Grouping Means in the Analysis of Variance, *Biometrics* 30(3), 507-512.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Slater, A., Mattock, A., Brown, E., & Bremner, J. G. (1991). Form perception at birth: Cohen and Younger (1984) revisited. *Journal of Experimental Child Psychology*, *51*(3), 395–406.

Smith, K. A., Battaglia, P. W., & Vul, E. (2018). Different Physical Intuitions Exist Between Tasks, Not Domains. *Computational Brain & Behavior*, *1*(2), 101–118. https://doi.org/10.1007/s42113-018-0007-3

Smith, K. A., & Vul, E. (2013). Sources of uncertainty in intuitive physics. *Topics in Cognitive Science*, *5*(1), 185–199. https://doi.org/10.1111/tops.12009

Ullman, S. (1984). Visual routines. *Cognition*, *18*(1), 97–159. https://doi.org/10.1016/0010-0277(84)90023-4

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science (New York, N.Y.)*, *224*(4654), 1260–1262.

Walther, D. B., Chai, B., Caddigan, E., Beck, D. M., & Fei-Fei, L. (2011). Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(23), 9661–9666. https://doi.org/10.1073/pnas.1015666108

Wehner, R., Michel, B., & Antonsen, P. (1996). Visual navigation in insects: coupling of egocentric and geocentric information. *The Journal of Experimental Biology*, *199*(1), 129-140.

Weintraub, D. J., & Virsu, V. (1972). Estimating the vertex of converging lines: Angle misperception? *Perception & Psychophysics*, *11*(4), 277–283. https://doi.org/10.3758/BF03210378

Wiltschko, W., & Wiltschko, R. (2005). Magnetic orientation and magnetoreception in birds and other animals. *Journal of Comparative Physiology A*, *191*(8), 675-693.

Zaslavsky, O. (2010). The Explanatory Power of Examples in Mathematics: Challenges for Teaching. In M. K. Stein & L. Kucan (Eds.), *Instructional Explanations in the*

Disciplines (pp. 107–128). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-0594-9_8

Zodik, I., & Zaslavsky, O. (2007). Is a visual example in geometry always helpful? *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education, v.4 (Seoul, Korea)*, 265–272.