Received: 13 January 2023 | Revised: 19 July 2023 | Accepted: 18 August 2023

DOI: 10.1111/desc.13449

SHORT REPORT

Developmental Science WILEY

Walking and falling: Using robot simulations to model the role of errors in infant walking

Ori Ossmy[1,†] | Danyang Han[2,†] | Patrick MacAlpine[3] | Justine Hoch[4] | Peter Stone[3,5] | Karen E. Adolph[2]

[1]Centre for Brain and Cognitive Development & Department of Psychological Sciences, Birkbeck, University of London, London, UK

[2]Department of Psychology, New York University, New York, New York, USA

[3]Department of Computer Science, The University of Texas at Austin, Austin, Texas, USA

[4]Department of Psychology and Human Development, Vanderbilt University, Nashville, Tennessee, USA

[5]Sony AI, Austin, Texas, USA

Correspondence
Karen Adolph, Department of Psychology, New York University, 6 Washington Place, Room 410, New York, NY 10003, USA.
Email: karen.adolph@nyu.edu

[†]Joint first authorship.

Portions of this work were presented at the biannual conference of the Society for Research in Child Development in 2019.

Funding information
National Institute of Child Health and Human Development, Grant/Award Number: R01-HD033486; Defense Advanced Research Projects Agency, Grant/Award Number: N66001-19-2-4035

Abstract

What is the optimal penalty for errors in infant skill learning? Behavioral analyses indicate that errors are frequent but trivial as infants acquire foundational skills. In learning to walk, for example, falling is commonplace but appears to incur only a negligible penalty. Behavioral data, however, cannot reveal whether a low penalty for falling is beneficial for learning to walk. Here, we used a simulated bipedal robot as an embodied model to test the optimal penalty for errors in learning to walk. We trained the robot to walk using 12,500 independent simulations on walking paths produced by infants during free play and systematically varied the penalty for falling—a level of precision, control, and magnitude impossible with real infants. When trained with lower penalties for falling, the robot learned to walk farther and better on familiar, trained paths and better generalized its learning to novel, untrained paths. Indeed, zero penalty for errors led to the best performance for both learning and generalization. Moreover, the beneficial effects of a low penalty were stronger for generalization than for learning. Robot simulations corroborate prior behavioral data and suggest that a low penalty for errors helps infants learn foundational skills (e.g., walking, talking, and social interactions) that require immense flexibility, creativity, and adaptability.

KEYWORDS

error, falling, penalty, reinforcement learning, simulated robot, walking

Research Highlights

- During infant skill acquisition, errors are commonplace but appear to incur a low penalty; when learning to walk, for example, falls are frequent but trivial.
- To test the optimal penalty for errors, we trained a simulated robot to walk using real infant paths and systematically manipulated the penalty for falling.
- Lower penalties in training led to better performance on familiar, trained paths and on novel untrained paths, and zero penalty was most beneficial.

Ori Ossmy and Danyang Han contributed equally to the work and shared the first authorship.

- Benefits of a low penalty were stronger for untrained than for trained paths, suggesting that discounting errors facilitates acquiring skills that require immense flexibility and generalization.

1 INTRODUCTION

What is the role of errors in learning? Commonsense beliefs, educational dogma, a long history of learning research, and modern artificial intelligence (AI) assume that errors signal the need for skill improvement and therefore inform learning (e.g., Ferris & Roberts, 2001; Kaelbling et al., 1996; Rescorla & Wagner, 1972). Presumably, a salient penalty for errors (i.e., negative feedback, or in AI lingo, "negative reward") promotes learning by highlighting undesired behaviors and incentivizing correction of prior mistakes. This process is common in learning highly structured, formal skills. For example, in math learning, teachers and parents typically provide children with clear examples as training input and explicit negative feedback for incorrect solutions. Children then repeat the same operations with the goal of eliminating errors and minimizing negative feedback. Likewise, in reinforcement learning in AI and computational modeling, negative reward—negative feedback for errors—is critical for learning (Kaelbling et al., 1996).

In contrast to formal skills such as math, foundational skills—walking, talking, social interactions, and so on—require immense flexibility, creativity, and adaptability to cope with the flux of everyday situations. From prior work, it is clear that infants do not learn foundational skills through structured training regimens with clear exemplars as training input as in math learning. Instead, learning foundational skills entails massive amounts of variable, time-distributed practice (Herzberg, Fletcher et al., 2022; Smith et al., 2018; Herzberg, Shilling et al., 2022). For example, when learning to walk, infants take thousands of steps per hour punctuated by frequent starts and stops. Most bouts contain steps in every direction (forward, backward, and sideways) and babies' paths are curved and winding (Adolph et al., 2012; Lee et al., 2018). Formal modeling with simulated robots shows that the variable training input that results from everyday activities leads to better learning than structured training on uniform, straight paths or on repetitive circular or square paths (Ossmy et al., 2018). But in addition to variability, infants' everyday training input also generates many mistakes. What role do errors play in infants' learning of foundational skills?

Behavioral data suggest that infants' natural walking regimen entails a low penalty for errors. Falling—an error in walking—is frequent during infants' natural activity, but falls appear to incur a negligible penalty. Walking infants fall 17 times per hour (Adolph et al., 2012). However, after falling, infants rarely cry, and caregivers rarely show concern, and infants return to play within two seconds (Han & Adolph, 2021). Moreover, infants are not deterred from walking: They move just as much before and after a fall and they do not avoid the objects or locations that were implicated in their falls.

In contrast to adults (especially older adults), infants may be "built" to fall (Han & Adolph, 2021), so the penalty for errors is naturally "discounted." Infants are short and low to the ground (short distance to impact), light weight (small mass at impact), and move slowly (low impact velocity). Therefore, an infant fall generates 18 times less potential energy than if infants were adult-sized and walked at adult speeds. Moreover, when infants lose balance, they quickly display a suite of a reactive behaviors that mitigate impact forces: They take quick reactive steps to maintain upright posture, grasp nearby furniture to slow the speed of falling, bend their knees during landing to reduce the impact, and outstretch their hands to break the fall. Infants mostly fall onto "padded" body parts such as hands, legs, and buttocks. Sensitive body parts such as the head and torso rarely impact the floor, and such impacts occur after "safer" body parts have already mitigated impact forces. The impact energy is further absorbed by infants' body fat and muscles such that the residual energy is rarely sufficient to break infants' malleable bones (Butte et al., 2000; Currey, 1979). Moreover, despite the high frequency of falls during infants' everyday activity, fall-related injuries are limited to a small proportion of infants (3%) primarily due to falling from a large height (e.g., off a balcony, down stairs; American Academy of Pediatrics, 2001; Borse et al., 2008; Lallier et al., 1999). Indeed, across toddlers' entire history of locomotion, most parents report that their infants never incurred a serious fall that left a mark, bruise, or cut or warranted a call to a doctor (Han & Adolph, 2021).

Although behavioral data suggest that infant falls are trivial, it cannot reveal whether a low penalty for error is optimal for learning foundational skills such as walking. One possibility is that infants' natural training regimen—with a low, negligible penalty for errors—is most beneficial for learning because infants are not discouraged from practicing their new skills. An alternative possibility is that a higher penalty for errors leads to better, faster learning by highlighting the causes of errors or by incentivizing faster reduction of errors. The ideal test of these possibilities would entail multiple training instances with a "standard" baby—as if that baby could learn to walk under one penalty condition and then start afresh under another penalty condition, with no carryover between conditions. Of course, such a test is impossible with real babies, but it is possible with robot simulations. A simulated robot can have a constant setting at the start of each training regimen such that each training instance can be independent and unrelated factors can be controlled.

Although both infancy and AI researchers are interested in developing systems that produce functional behavior, the two disciplines rarely capitalize on their complementary expertise, and relatively few studies have used the computational power of AI to test hypotheses in developmental science. The current study builds on previous research that used soccer-playing simulated robots to test the role of variable input in learning to walk (Ossmy et al., 2018). The prior work showed the feasibility and profitability of using simulated robots to test manipulations that are practically or ethically impossible with infants and children.

2 | CURRENT STUDY

We used robot simulations to test the effects of different penalties for errors on learning to walk. To retain maximal ecological validity in training, we conducted each simulation on one of five walking "paths" produced by actual infants during free play. Most critical, during training on the assigned path, each simulation was paired with one of five penalty conditions, ranging from zero penalty to 100 times the standard penalty. Then we evaluated the robot's walking performance on one of five tests—the trained path (to test learning) or on one of the four untrained paths (to test generalization). We repeated 100 independent training simulations for each of 125 path/penalty/test combinations (5 training paths × 5 penalty conditions × 5 tests) resulting in 12,500 independent simulated robot regimens in total—a level of precision, control, and magnitude that would be impossible with real infants.

Our outcome measures of walking performance were how far the robot walked and how well it walked in a 1-h test. Our primary hypothesis was that infants' natural learning input—where falling is trivial and has little impact on infants' subsequent activity—is a beneficial training regimen for flexible, functional walking. Thus, we predicted that the lower the penalty during training, the better the simulated robot would perform at test on both trained and untrained paths and that zero penalty would be most conducive for learning to walk. We also tested whether a lower penalty during training is more beneficial for learning to walk on trained versus untrained paths. If the effect of a lower penalty is stronger for trained paths, it would suggest that the benefits of low-penalty errors in training are specific to the initial training input. If the effect of a lower penalty is stronger for untrained paths, it would suggest that a low penalty for errors promotes generalization and flexibility of walking skill. Finally, based on prior robot simulations showing that varied training input is the most beneficial training regimen for learning to walk (Ossmy et al., 2018), we explored how changing the penalty for errors affects learning to walk in the context of more and less varied infant walking paths.

3 | METHODS

3.1 | Infant walking trajectories

We used the walking trajectories of 75 15-month-old infants (40 girls, 35 boys) during 20 min of free play in a large (6 m × 9 m) laboratory play room (Figure 1a) as in Ossmy et al. (2018). Families were recruited from the New York City area. All infants were born at term with no known disabilities. Parents reported infants' race and ethnicity as White (57.3%), Black (4.0%), Asian (6.7%), multiple races (22.7%), other (2.7%), or not reported (6.7%); 16% were Hispanic.

Walking trajectories were recorded from a fixed overhead view that captured the entire playroom. To define the trajectories, we first identified the timing of walking bouts with Datavyu (datavyu.org), a video coding software that time locks user-defined annotations to the relevant video frames. A primary coder scored the onset and offset of each walking bout and the number of steps per bout (as in Lee et al., 2018). A second coder independently scored 25% of each session to assess inter-observer agreement: $rs > 0.96$, $ps < 0.001$ for number of bouts, bout duration, and number of steps per bout. To define the shape of each trajectory and the angle between consecutive steps, the primary coder used Matlab software (DLTDataViewer digitizing tool) to manually digitize the location of each left-foot step using the overhead camera view that recorded the entire playroom. If an infant's foot was momentarily occluded, the coder estimated the location based on the preceding and following steps. We used the *xy* coordinates of these points to map the paths into the configuration of the playroom (adjusting for lens and perspective distortion). Using known distances, we verified that the digitizing method returned <1% error per bout. See an exemplar video clip of an infant free play session and the digitized walking trajectory at databrary.org/volume/1552/slot/63238.

To test the effects of penalty for falling on more and less varied infant walking trajectories, we created 5 groups of 15 infants using a k-means clustering algorithm with $k = 5$ (Spath, 1985). Clusters were based on variation in four aspects of infant walking as in Ossmy et al. (2018)—*bout shape* (curved or straight), *step direction* (angle of each step in 360°), *bout length* (steps per bout), and *number of bouts* (starts and stops). We calculated *variation in bout shape* as the standard error of bout curvature for bouts of ≥4 steps; we calculated bout curvature by averaging the overall bout curvature (the shortest distance between the start and end points of the bout divided by the total distance traveled) and step-to-step curvature (calculated the same way from each series of 3 points in the bout). We calculated *variation in step direction* for bouts of ≥2 steps based on the standard error of the change in degrees of the plane angle between each pair of steps. We calculated *variation in bout length* as the standard error of the number of infant steps per walking bout. Finally, we calculated *number of bouts* as the total number of stops.

3.2 | Robot training paths

Following the cluster analyses, we created 5 robot training paths using the walking trajectories of the 15 infants in each group. For each infant session, we first excluded all the time between bouts when infants were not walking. From the remaining walking time, we randomly sampled a 4-min block of consecutive walking. Although infants often stopped for long periods between bouts, the robots were fully stabilized after 2 s, so longer pauses had no additional value for robot training. Thus, we inserted uniform stationary periods of 2 s between each infant bout.

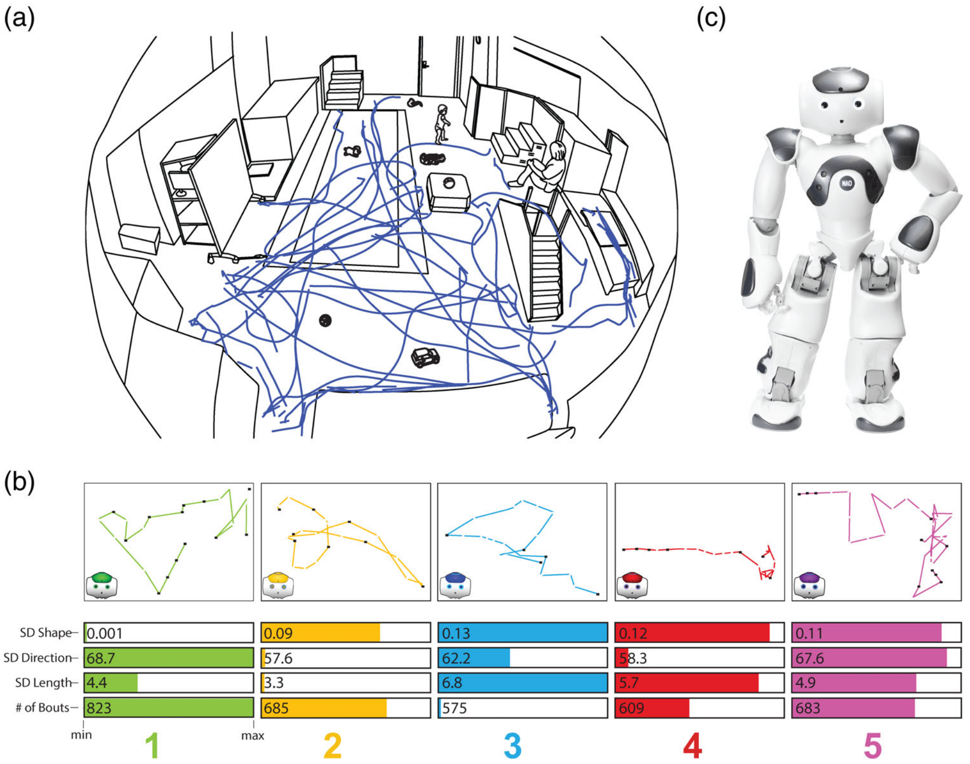

FIGURE 1 Study design. (a) Layout of the laboratory playroom. Blue line depicts an example path generated by a walking infant during 20 min of free play. (b) Five robot training paths based on concatenated walking trajectories generated by 75 infants during free play. Infants' walking trajectories were clustered into five groups based on trajectory features—variations in bout shape, step direction, bout length, and number of bouts. Top panel: exemplar segment from each robot training path (colored lines = path trajectory, dashes = stride length from left foot to left foot, black dots = stops). Bottom panel: relative combinations of features for each training path. Values are scaled from the minimum to the maximum across paths. (c) Simulated robot modeled after an Aldebaran Nao real robot.

(A few infants had less than 4 min of accumulated walking duration, so their paths were repeated until 4 min accumulated.) Finally, we concatenated the randomly sampled 4-min blocks from each of the 15 infants to create a 1-h training path (a realistic duration for training in terms of computational time complexity). Therefore, the concatenated training paths represented the combination of dimensions in each group of infant walking trajectories. Figure 1b (horizontal bars) shows the relative dimensions of the five concatenated training paths, distinguished by color. For example, the green training path was characterized by high variation in step direction, high number of bouts, and low variation in bout shape, and relatively low variation in bout length. The purple training path had relatively high variation along all dimensions. Figure 1b (path shapes) depicts exemplar portions of each training path.

We took the coordinates of each concatenated training path and mapped the start, stop, and shape of the walking bouts onto the robot walking field where each grid space is 1 $m^2 \times 1m^2$. During training, the simulated robot walked sequentially to each destination specified by the concatenated infant paths. When the infant stopped walking, the robot also stopped walking and stood in place.

3.3 | Robot simulations

The robot simulation was conducted in SimSpark, the simulator used in the RoboCup soccer 3D-simulation league competitions. The simulated robot was loosely modeled after the Aldebaran Nao real robot (www.aldebaran-robotics.com) with a height of 57 cm, a mass of 4.5 kg, and 22 degrees of freedom (six in each leg, four in each arm, and two in the neck); see Figure 1c. The simulated robot had proprioception of all joints, pressure sensors on its feet, two gyrometers, and an accelerometer. The joint preceptors and effectors enabled monitoring and control of the hinge joints. Joint effectors allowed the robot to specify the torque and direction in which to move.

We used a similar, open source, parameterized robot walk engine as in previous work (MacAlpine et al., 2012; MacAlpine & Stone, 2016; Ossmy et al., 2018) that first selects a path for the torso to follow, and then determines where the feet should be with respect to the torso location. The walk engine was parameterized using more than 40 parameters that specify the manner of robot walking, such as the maximum height of the foot from ground, maximum step size, and duration of a single step (Table 1). The parameters of the walk engine were initialized based on previous testing on the Nao real robot. See MacAlpine

TABLE 1 Parameters of the walk engine that are optimized during learning. For full details, see MacAlpine et al. (2012).

Parameter
Maximum size of steps (radians)
Maximum size of steps for *x* coordinates (mm)
Maximum size of steps for *y* coordinates (mm)
How much center of mass is shifted from side to side (mm)
Height of the torso from ground (mm)
Maximum height of foot from ground during step (mm)
Fraction of a phase the swing foot remains still before moving
Fraction of a phase that the swing foot on the ground before lifting
Duration of single step (seconds)
Expected difference between commanded COM and sensed COM
Factor of how fast the step sizes change per time cycle
Maximum COM error before the steps are slowed (mm)
Maximum COM error before all velocity reach 0 (mm)
Constant offset between the torso and feet (mm)
Factor of the step size applied to the forwards position of the torso
Fraction of a phase that the swing foot spends in the air
Angle of foot when it hits the ground (radians)
Proportional controller values for the torso angles—tilt
Proportional controller values for the torso angles—roll
Proportional controller values for controlling COM (*x*)
Proportional controller values for controlling COM (*y*)
Proportional controller values for controlling COM (*z*)
Proportional controller values for moving arms (*x*)
Proportional controller values for moving arms (*y*)

et al. (2012) for a full description of the technical and mathematical details of the walk engine.

3.4 | Optimization procedure and reward structure

Before training, the robot could walk very slowly with the initial set of walking parameters. Based on the positive and negative rewards received during training, the robot learned to repeat or avoid repeating certain actions by optimizing the set of walking parameters as it walked toward a series of destinations (goToTarget optimization; MacAlpine et al., 2012; Ossmy et al., 2018) that followed the concatenated infant path. Because it was impractical to optimize more than 40 parameters, we selected a subset of 24 parameters, based on their impact on the speed and stability during previous testing with the Nao real robot (Table 1; see MacAlpine et al., 2012 and Ossmy et al., 2018). Over the course of the optimization, the robot learned to walk increasingly faster with fewer errors.

Over the training, the robot optimized the walking parameters to maximize the net rewards—the sum of the positive and negative rewards for walking to destinations minus the penalty (additional negative rewards) for falling. The robot received positive rewards based on the distance traveled toward the destination (according to the infant path). If the robot reached a destination ahead of time (faster than the time it took the infant to reach that destination—capped at 7 s), it received extra rewards based on the distance it could have traveled given the remaining time. If the robot reached a destination with a slower time than it took the infant to reach that destination, it received fewer positive rewards (note, falls interrupt walking and thus caused slower robot walk times). Training also included "stop destinations," so the robot received negative rewards for overshooting the destination. For full equations describing the robot reward structure, see MacAlpine et al. (2012).

3.5 | Penalty for falls

The optimization procedure and reward structure for the robot simulations were identical to MacAlpine et al. (2012); the only difference was the penalty value for falling during training. A critical part of our training procedure—that differed from MacAlpine et al (2012)—was that the robot simulations received different penalties (i.e., varying negative rewards) after falling during the optimization run. We manipulated the value of the penalty to test the role of errors in learning to walk. We used five penalty values based on the standard penalty value used in prior work (MacAlpine et al., 2012; Urieli et al., 2011): 0 (no penalty at all), 2.5 (half the standard penalty), 5 (standard penalty), 10 (double the standard penalty), and 500 (100 times the standard penalty). High penalty for error should lead the robot to avoid certain movements whereas low penalty for error should have minor or no influence on the learning procedure because it does not posit any constraint on the movement in subsequent training iterations.

3.6 | Learning outcome measures

To evaluate the training success for each penalty value, we tested 100 independent simulations on the robot's own training path (i.e., learning) and 100 independent simulations on each of the four untrained paths (i.e., generalization) for 1 h. We used two outcome measures: how *far* the robot walked during each 1-h test (total distance in meters); and how *well* the robot walked during the 1-h test—measured by the percent improvement in the robot's "fitness" score after training relative to the fitness score if the robot had not been trained at all (that is, with no parameter optimization; see MacAlpine et al., 2012; Urieli et al., 2011). Fitness scores for evaluation were computed as the sum of all the positive and negative rewards for walking according to destinations minus the standard penalty for falling during the post-training test walk (see full details about the fitness score in MacAlpine et al., 2012). That is, the fitness score was the net reward with the standard penalty for falling.

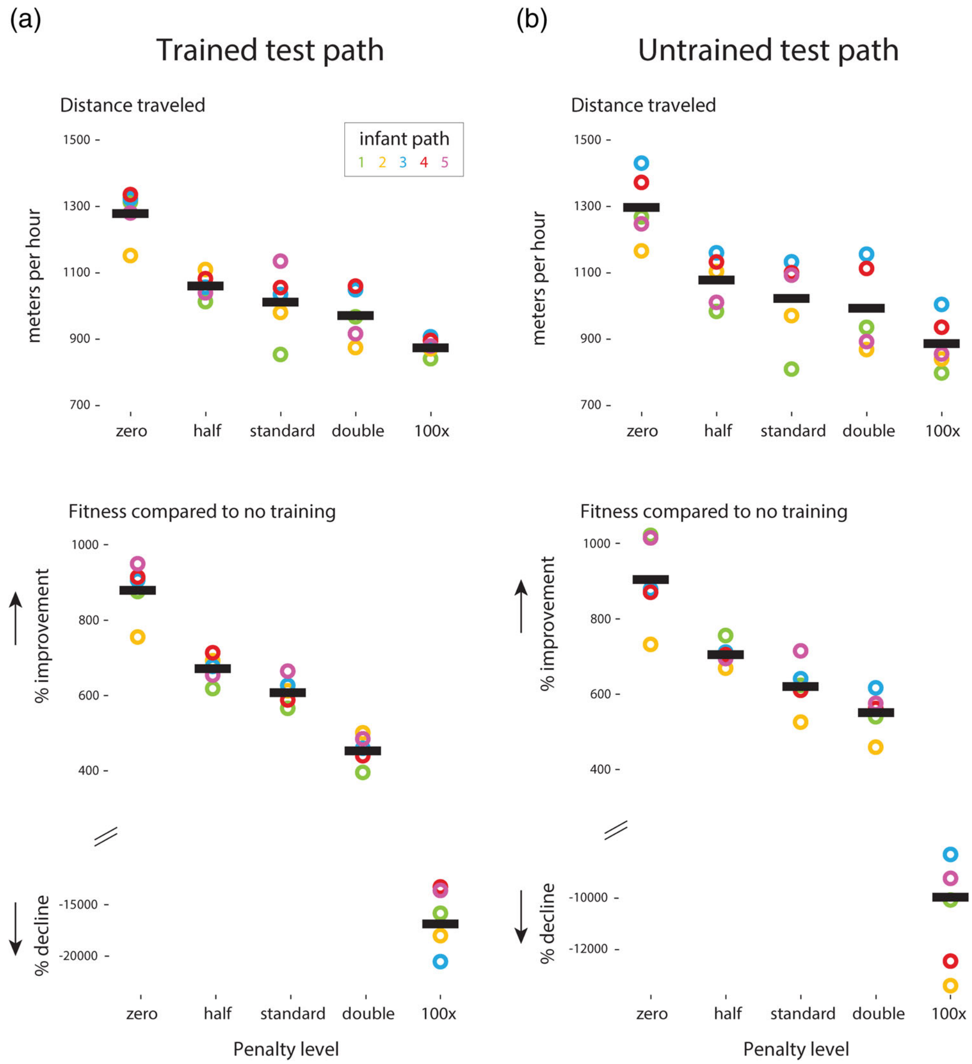

FIGURE 2 Effects of penalty on learning and generalization. (a) When tested on the trained path, robot simulations with a low penalty for falling travelled farther (top panel) and had greater improvement in fitness (bottom panel) relative to training with a high penalty for falling. (b) When tested on a novel, untrained path, robot simulations with a low penalty for falling traveled farther (top panel) and had greater improvement in fitness (bottom panel) relative to training with a high penalty for falling. Each symbol in (a) represents the average walking performance across 100 simulations of robot trained on the same infant path. Each symbol in (b) represents the average walking performance across 400 robot simulations trained on the same infant path (100 simulations tested on each of the four untrained paths).For example, 400 robot simulations trained on infant path 1 were tested on infant paths 2, 3, 4, and 5 with 100 simulations for each path. Symbol color denotes training path. Black lines denote means across paths.

4 | RESULTS

4.1 | Lower penalties led to better learning

Overall, lower penalties led to better learning—that is, better performance on the trained test path (identical to the robot's training path). Regardless of the training path, the robot walked farthest when it was trained with zero penalty for error compared to any non-zero penalty for error (Figure 2a, top panel). A 5 (penalty values) × 5 (training paths) ANOVA confirmed a main effect of penalty on distance walked, $F(4, 2475) = 8.63$, $p < 0.01$, a main effect of training path, $F(4, 2475) = 3.44$, $p < 0.01$, but no interaction between penalty and training path, $F(16, 2475) = 1.02$, $p = 0.42$. Sidak-corrected post-hoc tests on penalty showed that the robot walked farther when trained with zero penalty compared to all other penalties, $ps < 0.05$, and it walked for shorter distances when trained with 100 times the standard penalties compared to all other penalties, $ps < 0.01$. Sidak-corrected post-hoc tests on training path showed that the robot walked for shorter distances when trained on the least varied green path than the more varied red and blue paths, $ps < 0.05$.

Similarly, lower penalties led to better quality of walking. Regardless of training paths, the robot had the highest percent improvement in fitness scores when trained with zero penalty for error compared to any non-zero penalty for error (Figure 2a, bottom panel). A 5 (penalty values) × 5 (training paths) ANOVA confirmed a main effect of penalty on fitness scores, $F(4, 2475) = 4.21$, $p < 0.01$, a main effect of training path, $F(4, 2475) = 2.87$, $p < 0.03$, but no interaction between penalty and training path, $F(16, 2475) = .84$, $p = 0.62$. Sidak-corrected post-hoc tests on penalty showed that the robot had higher improvement in fitness scores when trained with zero penalty compared to all other penalties, $ps < 0.01$, and it had lower improvement in fitness scores when trained with 100 times the standard penalty compared to all other penalties, $ps < 0.001$. Sidak-corrected post-hoc tests on training paths showed that the robot had lower improvement in fitness scores when trained on the least varied green path compared to the most varied purple path, $p < 0.04$.

4.2 | Lower penalties led to greater generalization

As in the trained test path, a lower penalty for error improved robot walking in untrained test paths (different from the robot's training path). Figure 2b shows that decreasing the penalty improved walking distance (top panel) and fitness score (bottom panel). When the robot was trained with zero penalty, it walked farthest and displayed greatest improvement in fitness scores. ANOVAs with 5 (penalty values) × 5 (training paths) on distance and fitness confirmed a significant effect of penalty, $Fs(4, 9975) > 6.29$, $ps < 0.01$, no effect of infant path, $Fs(4, 9975) < 1.78$, $ps > .13$, and no interactions, $Fs(16, 9975) < 1.53$, $ps > .14$. Sidak-corrected post-hoc tests on distance and fitness showed that training with zero penalty led to better performance compared to training with all other penalties, and training with 100 times the standard penalty led to worse performance compared to training with all other penalties, $ps < 0.01$. In addition, the half penalty and the standard penalty conditions led to better performance compared to training with the double penalty, $ps < 0.04$.

4.3 | Lower penalties were more beneficial for generalization than for learning

Moreover, we found that lower penalties for error had a more pronounced effect on generalization (i.e., when robots were tested on a different path than their training path) than on learning (i.e., when robots were tested on the same path as their training path). To test whether lowering the penalty for errors was more beneficial for learning or generalization, we compared the difference in performance between robots trained with each pair of penalty values. Thus, we subtracted the robot's walking performance (total distance or improvement in fitness) with higher penalty from that of lower penalty for trained and untrained paths. As shown in Figure 3, the average differences in walking performance (lower penalty minus higher penalty) were positive in all trained and untrained tests (all above zero in figure), indicating that a lower penalty resulted in better walking performance than a higher penalty for both trained and untrained test paths. Critically, the differences in walking distance were significantly higher for untrained paths than trained paths for 8 of 10 test pairs (gray regions in Figure 3a), $ts(23) > 2.31$, $p < 0.03$. Likewise, the differences in fitness score improvement were higher for untrained paths than trained paths for 8 of 10 test pairs (gray regions in Figure 3b), $ts(23) > 1.93$, $ps < 0.05$.

5 | DISCUSSION

We used a simulated robot as an embodied model to test the optimal penalty for errors in learning to walk. We systematically manipulated the penalty for falling and assessed the influence on learning and generalization. Lower penalties improved robot walking performance for both familiar, trained walking paths and novel, untrained walking paths. And zero penalty was most conducive for learning and generalization. Moreover, a low penalty during training was especially beneficial when the robot had to generalize its learning to novel paths. Taken together, our findings suggest that low (or zero) penalty for errors is beneficial for learning foundational, flexible skills like walking.

5.1 | Low penalty for errors facilitates learning foundational skills

Why might a foundational skill like walking benefit from a low penalty for error? In contrast to more formal, structured skills such as math computations, foundational skills like walking require tremendous flexibility, creativity, and adaptability to cope with the everchanging flux of everyday situations (Adolph & Hoch, 2019). Because there is no fixed solution or formula for learning to walk—walkers must adjust their gait patterns on the fly when wearing different footwear, carrying varied objects, or walking over different ground surfaces—it requires an immense amount of varied practice to be flexible and generative. Accordingly, we propose that a low penalty for errors ensures that infants maintain high motivation to practice such that errors won't shut down the learning system. Indeed, after falling in free play, infants return to the same level of activity they exhibited prior to the fall—they continue to walk as much after a fall as before the fall and they return to play with the objects and locations that were implicated in the fall moments earlier (Han & Adolph, 2021).

Moreover, a low penalty for error may motivate infants to engage in new, challenging, creative activities. Because the penalty for falling is low, infants can continuously push the boundaries of their current skill repertoire (Han et al., 2021). As a result, infants spontaneously try new things. For example, as soon as infants can walk on the floor, they try to run, spin, and walk up and down elevations. Of course, these new skills incur more falls, but the benefits far outweigh the costs. In contrast, in academic settings, students' perceived "costs" are associated with lower academic performance and higher intention to avoid classroom engagement (Jiang et al., 2020). Likewise, students' anticipation of negative social consequences predicts less creative behaviors (Ivcevic & Hoffmann, 2021).

Nonetheless, a low penalty for errors does not preclude the possibility that infants learn from errors. Repeated errors in a uniform environment can inform infants' future actions. For example, after falling repeatedly into a squishy, visually marked foam pit, infants

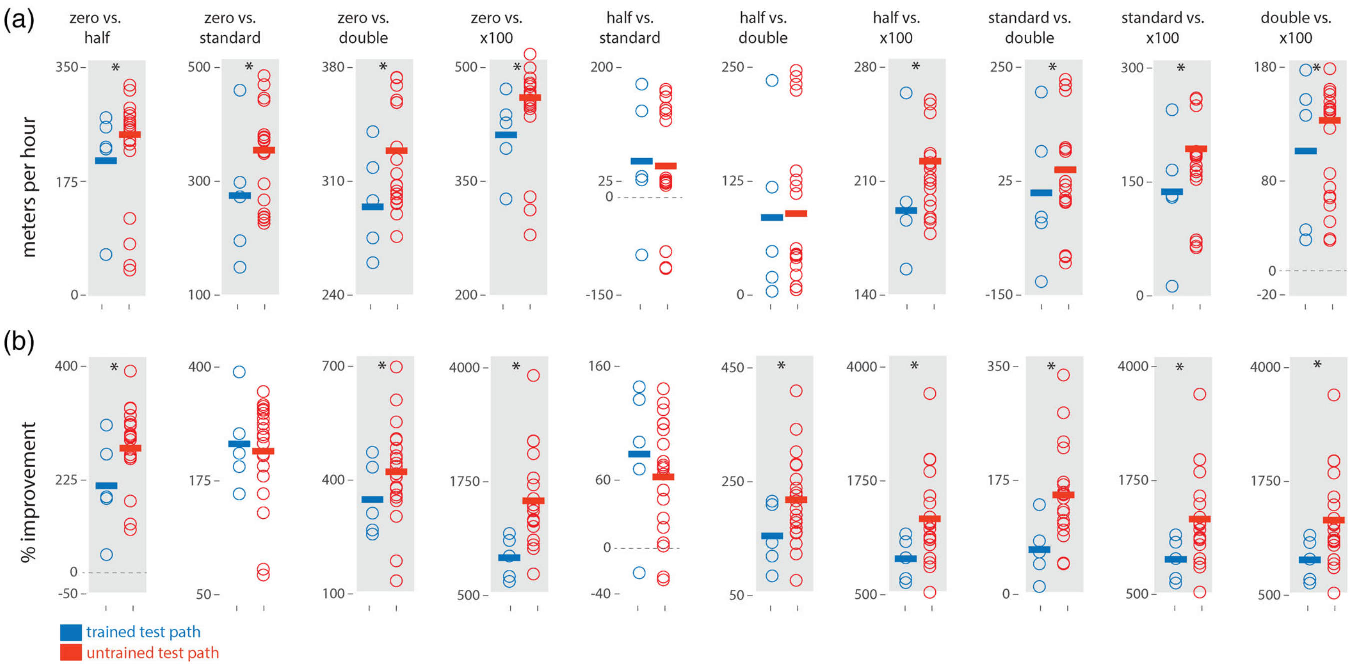

FIGURE 3 Stronger effects of penalty on generalization than learning. Graphs show all possible comparisons between penalty conditions (lower penalty minus higher penalty) for (a) walking distance and (b) improvement in fitness scores. Horizontal dashed lines represent 0 difference. Positive values indicate the lower penalty was more beneficial than the higher penalty. For each comparison, left plot (blue) shows trained path and right plot (red) shows untrained path. Each symbol denotes one combination of training and testing for 100 robot training simulations. Gray regions and asterisks denote significant differences between trained and untrained test paths.

learned to associate the appearance of the foam pit with the consequences for locomotion. They increased exploratory touching of the foam pit, showed longer latency before crossing, and avoided the pit or used alternative strategies to cross (Han, et al., in press). However, learning fixed associations between an environmental stimulus and falling is not conducive for learning to walk in the real world (Adolph & Joh, 2009). Outside the confines of a structured laboratory task, most infant falls are due to "internal" causes—weak legs or poor upright balance control (legs collapse or infants lose balance when turning their head or lifting an arm while standing) or motor execution (trip over their own feet while walking or running)—all in different environmental contexts (Han, et al., 2021). Thus, if infants associated the causes of falls with a moderately high penalty, they would never learn to stand, turn their head, lift their arm, or walk or run. If they associated such internally generated falls with the features of the environment where the fall occurred, they would never want to walk at all.

5.2 | A low penalty for errors reinforces the benefits of a varied training regimen

Previous work shows that a more varied training regimen facilitates skill generalization (Ossmy et al., 2018). Simulated robots trained with real, varied infant paths outperformed robots trained on uniform, geometrically shaped paths (straight lines, circles, and squares) in robot soccer—a testing scenario that requires high flexibility in walking skills to cope with a continually changing environment. And robots trained on more varied infant paths outperformed robots trained on less varied infant paths. In the current study, we found consistent evidence that more varied infant paths (e.g., purple and red paths) are more beneficial for training than less varied infant paths (e.g., green path), regardless of the penalty values and testing environments. Critically, the current study took these findings one step further: We found that a varied training regimen accompanied by a low penalty for errors facilitates learning and generalization, whereas a high penalty for falling undermines the benefits of a varied training regimen.

Of course, infants' real-world learning environments encompass a more diverse array of penalties for errors compared to the uniform penalties used in this study. When infants learn to walk, the penalties for errors—here, infant falling—can be conceptualized as a "pyramid of pain" (Han & Adolph, 2021). Most infant falls are at the base of the pyramid. More than 90% of infants' falls are unremarkable and unimpactful, less than 10% of falls instigate infant fussiness and caregiver concern in the moment. In the middle portion of the pyramid, only 30% of infants experience memorable falls that are remembered by caregivers months later. Falls at the top of the pyramid are even more rare: 3% of infants have falls that require medical attention, and the dire statistics of fatal infant falls is limited to 0.0003% of infants (Borse et al., 2008).

This study focused on understanding the impact of the overwhelmingly low-penalty errors on learning and generalization. Future studies should consider the variability of error penalties and examine how changes in penalties during the learning process (e.g., increases and decreases in penalty value) affect skill acquisition and generalization. Potentially, such a manipulation would offer a more nuanced understanding of the effects of penalty on learning and generalization.

6 | CONCLUSIONS

What is the role of errors in learning? When learning foundational skills such as walking, robot simulations indicate that a low penalty for errors—perhaps no penalty at all—is most conducive for learning and generalization. Accordingly, we propose that infancy is an ideal period for learning flexible, creative, adaptive skills. Infancy naturally confers varied input for learning. Infants acquire foundational skills in a changing body with changing skills in a changing environment (Adolph & Hoch, 2019; Adolph et al., 2018; Adolph & Robinson, 2015). Thus, the natural variety in infant development ensures that infants will not learn "static facts" but rather generalize their skills to varied body-environment relations (Adolph & Joh, 2009). Moreover, infants' everyday errors typically incur a low penalty—falls are unimpactful, social gaffes are trivial, and grammatical errors and disfluencies in speech receive no explicit correction and do not impede communication (e.g., Bohannon & Stanowicz, 1988). Due to a confluence of factors—infants' unique body characteristics, protected environments, and social scaffolding from their caregivers—these features of the learning environment are exclusively available during infancy. At later stages, higher physical and social penalties can impede skill acquisition such as learning to ski or to speak a second language.

ACKNOWLEDGMENTS

This work was supported by National Institute of Child Health and Human Development grant R01-HD033486 and Defense Advanced Research Projects Agency grant N66001-19-2-4035 to Karen Adolph.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest. Peter Stone serves as the Executive Director of Sony AI America and receives financial compensation for this work. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.

DATA AVAILABILITY STATEMENT

We publicly shared research materials to illustrate the method, including an exemplar video clip of an infant free play session and the corresponding digitized walking trajectory (databrary.org/volume/1552).

ETHICS APPROVAL STATEMENT

All study procedures were approved by New York University Committee on Activities Involving Human Subjects and adhere to the ethical standards of the Federal Policy for the Protection of Human Subjects. Caregivers of all participants gave their informed consent prior to inclusion in the study.

PERMISSION TO REPRODUCE MATERIALS FROM OTHER SOURCES

No materials are reproduced from other sources.

ORCID

Danyang Han https://orcid.org/0000-0001-7113-8857
Karen E. Adolph https://orcid.org/0000-0003-2819-134X

REFERENCES

Adolph, K. E., Cole, W. G., Komati, M., Garciaguirre, J. S., Badaly, D., Lingeman, J. M., Chan, G. L., & Sotsky, R. B. (2012). How do you learn to walk? Thousands of steps and dozens of falls per day. *PsychologicaL Science*, *23*, 1387–1394. https://doi.org/10.1177/0956797612446346

Adolph, K. E., & Hoch, J. E. (2019). Motor development: Embodied, embedded, enculturated, and enabling. *Annual Review of Psychology*, *70*, 141–164. https://doi.org/10.1146/annurev-psych-010418-102836

Adolph, K. E., Hoch, J. E., & Cole, W. G. (2018). Development (of walking): 15 suggestions. *Trends in Cognitive Sciences*, *22*, 699–711. https://doi.org/10.1016/j.tics.2018.05.010

Adolph, K. E., & Joh, A. S. (2009). Multiple learning mechanisms in the development of action. In A. Woodward & A. Needham (Eds.), *Learning and the infant mind* (pp. 172–207). Oxford University Press.

Adolph, K. E., & Robinson, S. R. (2015). Motor development. In R. M. Lerner, L. Liben, & U. Muller (Eds.), *Handbook of child psychology and developmental science* (7th ed., Vol. 2 Cognitive processes, pp. 113–157). Wiley.

American Academy of Pediatrics. (2001). Falls from heights: Windows, roofs, and balconies. *Pediatrics*, *107*, 1188–1191.

Bohannon, J. N., & Stanowicz, L. (1988). The issue of negative evidence: Adult responses to children's language errors. *Developmental Psychology*, *24*, 684–689. https://doi.org/10.1037/0012-1649.24.5.684

Borse, N. N., Gilchrist, J., Dellinger, A. M., Rudd, R. A., Ballesteros, M. F., & Sleet, D. A. (2008). *CDC childhood injury report: Patterns of unintentional injuries among 0–19 year olds in the United States, 2000–2006*. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Butte, N. F., Hopkinson, J. M., Wong, W. W., Smith, E. O., & Ellis, K. J. (2000). Body composition during the first 2 years of life: An updated reference. *Pediatric Research*, *47*, 578–585. https://doi.org/10.1203/00006450-200005000-00004

Currey, J. D. (1979). Changes in the impact energy absorption of bone with age. *Journal of Biomechanics*, *12*, 459–469. https://doi.org/10.1016/0021-9290(79)90031-9

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of second language writing*, *10*, 161–184. https://doi.org/10.1016/S1060-3743(01)00039-X

Han, D., & Adolph, K. E. (2021). The impact of errors in infant development: Falling like a baby. *Developmental Science*, *24*, e13069. https://doi.org/10.1111/desc.13069

Han, D., Cole, W. G., Joh, A. S., Liu, Y., Robinson, S., & Adolph, K. E. (in press). Pitfall or pratfall? Behavioral differences in infant learning from falling. *Journal of Experimental Psychology: General*.

Han, D., Rachwani, J., Liu, Y., & Adolph, K. E. (2021, April 7–9). *Why infants fall: Development changes the causes of errors as infants acquire new skills* [Conference presentation]. Society for Research in Child Development (virtual).

Herzberg, O., Fletcher, K. K., Schatz, J. L., Adolph, K. E., & Tamis-LeMonda, C. (2022). Infant exuberant object play at home: Immense amounts of time-distributed, variable practice. *Child Development*, *93*, 150–164. https://doi.org/10.1111/cdev.13669

Herzberg, O., Shilling, M. S., Tamis-LeMonda, C. S., & Adolph, K. E. (2022, July 1-10). *Infants' spontaneous locomotor activity at home.* [Conference presentation]. International Congress of Infant Studies, Ottawa, Canada.

Ivcevic, Z., & Hoffmann, J. D. (2021). The creativity dare: Attitudes toward creativity and prediction of creative behavior in school. *Journal of Creative Behavior*, *56*, 239–257. https://doi.org/10.1002/jocb.527

Jiang, Y., Kim, S. I., & Bong, M. (2020). The role of cost in adolescent students' maladaptive academic outcomes. *Journal of School Psychology*, *83*, 1–24. https://doi.org/10.1016/j.jsp.2020.08.004

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, *4*, 237–285. https://doi.org/10.1613/jair.301

Lallier, M., Bouchard, S., St-Vil, D., Dupont, J., & Tucci, M. (1999). Falls from heights among children: a retrospective review. *Journal of Pediatric Surgery*, *34*, 1060–1063. https://doi.org/10.1016/S0022-3468(99)90564-X

Lee, D. K., Cole, W. G., Golenia, L., & Adolph, K. E. (2018). The cost of simplifying complex developmental phenomena: A new perspective on learning to walk. *Developmental Science*, *21*, e12615. https://doi.org/10.1111/desc.12615

MacAlpine, P., Barrett, S., Urieli, D., Vu, V., & Stone, P. (2012). *Design and optimization of an omnidirectional humanoid walk: A winning approach at the RoboCup 2011 3D simulation competition* (Vol. 1). Association for the Advancement of Artificial Intelligence.

MacAlpine, P., & Stone, P. (2016). UT Austin Villa RoboCup 3D simulation base code release. *RoboCup 2016: Robot Soccer World Cup XX*, 135–143.

Ossmy, O., Hoch, J. E., MacAlpine, P., Hasan, S., Stone, P., & Adolph, K. E. (2018). Variety wins: Soccer-playing robots and infant walking. *Frontiers in Neurorobotics*, *12*, 19. https://doi.org/10.3389/fnbot.2018.00019

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99), Appleton-Century-Crofts.

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, *22*, 325–336. https://doi.org/10.1016/j.tics.2018.02.004

Spath, H. (1985). *The cluster dissection and analysis theory FORTRAN programs examples*: Prentice-Hall, Inc.

Urieli, D., MacAlpine, P., Kalyanakrishnan, S., Bentor, Y., & Stone, P. (2011). On optimizing interdependent skills: A case study in simulated 3D humanoid robot soccer. In K. Tumer, P. Yolum, L. Sonenberg, & P. Stone (Eds.), *Proeedings of the 10th International Conference on Autonomous Agents and Multiagent Systems* (pp. 769–776). IFAAMAS.

How to cite this article: Ossmy, O., Han, D., MacAlpine, P., Hoch, J., Stone, P., & Adolph, K. E. (2023). Walking and falling: Using robot simulations to model the role of errors in infant walking. *Developmental Science*, e13449. https://doi.org/10.1111/desc.13449