nature human behaviour

ARTICLES
https://doi.org/10.1038/s41562-019-0814-0

Infants rationally decide when and how to deploy effort

Kelsey Lucca[1,2*], Rachel Horton[1] and Jessica A. Sommerville[1,3]

The ability to decide whether, when and how to try is central to human learning. We investigated whether infants can make rational inferences about when and how to try on a novel problem-solving task. After learning from an adult that the task was either easy, difficult or impossible to solve, infants varied in whether, when and how they tried based on the type of social evidence that they received and on their own ongoing experience with the task. Specifically, infants formed expectations about the task, their own ability to solve the task and the experimenter's ability to solve the task, in light of accumulating evidence across time that impacted their time spent trying, trying force, affect, and help-seeking behaviour on the task. Thus, infants flexibly integrate social input and first-hand experience in a dynamic fashion to engage in adaptive persistence.

Persistence in the face of obstacles is central to positive life outcomes, from heightened academic achievement to enhanced emotional well-being[1–3]. Nevertheless, persistence is metabolically expensive and rife with opportunity costs. What truly spurs learning is adaptive persistence: dynamically integrating information about one's own abilities and the competencies of social partners to decide whether, when and how to try. Here we report that infants vary the nature of their persistence adaptatively: they persist at solving an (impossible) task after learning that it is challenging for an adult to solve, they seek help from an expert after learning she can solve the task easily, and they stop trying when an adult cannot solve the task. Thus, infants decide whether, when and how to try based on a high-level integration of the context, their own experience and the competencies of others. Moreover, these experiences also have consequences for infants' affect and subsequent motivation on future tasks. These findings emphasize that infants are not erratic, impulsive and unpredictable in their pursuit of goals. Rather, infants selectively persist when they know it is worth the effort—that is, when their hard work is likely to pay off. This ability to make rational inferences regarding how and when to deploy effort is present within the earliest years of life, and serves as a foundation for the uniquely human capacity for extraordinary learning.

A rich variety of evidence suggests that infants are ubiquitous and smart social learners: they acquire a wide variety of knowledge by observing and interacting with others. Infants also implement this social learning to guide their own persistence: they are more likely to try hard to overcome a challenge if they have previously seen an adult model persist[4]. However, what remains unknown is how the dynamics of persistence unfold over time as a function of infants' expectations about task difficulty and their own experience, and whether the effects of these dynamic expectations extend beyond how much infants try to how hard they try, whether they elect to try themselves or recruit help from others, their affective reactions to the task and future motivation to solve similar tasks.

Two trenchant, interlinked issues exist regarding infants' social learning specifically, and their learning more broadly construed. The first concerns the degree to which infants' social learning or imitation is rational. Recent evidence suggests that infants do not naively copy the actions of others when presented with modelled behaviour. First, infants appear to make inferences about an actor's underlying intention, and seek to reproduce her intention, even after seeing a failed attempt at achieving a goal[5]. For example, Meltzoff[5] presented 18-month-olds with one of two models: one who successfully pulled apart a dumbbell-shaped toy and one who intended, but failed, to pull it apart. Infants reproduced the target action in both scenarios, suggesting they were able to infer the contents of the model's unobservable intention. Second, infants seek to utilize the most efficient means to achieve a goal, and imitate unusual means only when the context suggests that the means are an explicit part of an actor's intention[4]. For example, in a classic study by Gergely and colleagues[4], 14-month-olds were presented with a model who turned a light panel on with her head whilst having her arms restrained, or a model who performed this same action but had her hands clearly displayed on the table, and were then given the opportunity to act on it themselves. When the model's arms were restrained, infants turned the light on with their hands, suggesting they inferred that the model used her head because her hands were unavailable; whereas, when the model's arms were freely available, infants turned the light on with their head—suggesting that in the latter scenario infants inferred that the act of head-tapping, though unusual, was an intentional part of the demonstration and thus reproduced it.

The second issue our work addresses concerns the computational principles that underlie early learning. Bayesian models of learning posit that infants (and more mature learners) probabilistically entertain multiple hypotheses about the world (priors) based on observed evidence, and assign probabilities to these hypotheses that are subsequently adjusted in light of new information and experience in the world[6–9]. Relatedly, the naive utility calculus[10], nested within this Bayesian framework, views young learners as 'utility maximizers'. For example, infants estimate the anticipated utility of others' actions by integrating and weighing the costs and rewards of their actions, and then use this information to decide whether and how others will act, as well as form inferences about the beliefs, goals and desires of others. Though these frameworks have enriched and helped formalize theories early learning, there is only limited empirical evidence to support them[11].

The goal of the current work is to move these research endeavours forward. Our work adds to the growing literature on rational imitation by determining whether infants selectively reproduce effortful behaviour when they have reason to believe they will be

[1]Department of Psychology, University of Washington, Seattle, WA, USA. [2]Department of Psychology, Arizona State University, Tempe, AZ, United States. [3]Department of Psychology, University of Toronto, Toronto, ON, Canada. *e-mail: klucca@asu.edu

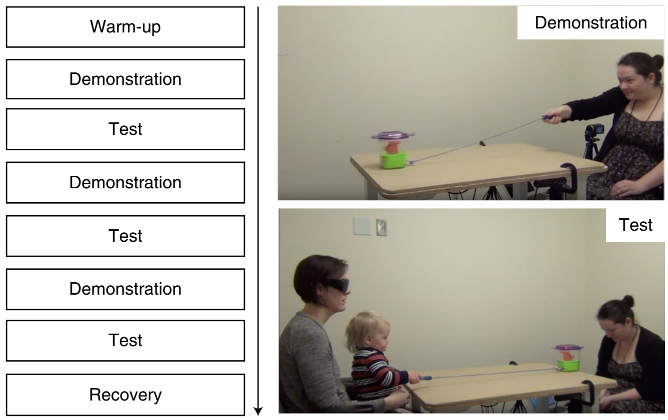

Fig. 1 | Experimental procedure. Schematic of experimental procedure (left) and screenshots taken from demonstration (top right) and test trials (bottom right).

successful. Moreover, we sought to assess whether infants can use social information and their ongoing experience not only to flexibly infer a social partner's intentions, but to adjust their inferences more broadly regarding the nature of the task, and their social partner's ability to solve the task. In pursuing these goals, we sought to move beyond asking whether and how infants would reproduce a single target behaviour, to examining the dynamic unfolding of all of infants' task-related actions, their help-seeking and affect, after watching a social demonstration. Our work directly contributes to an understanding of the computational principles underlying infants' inferences and behaviour by directly testing core tenets of these models. Specifically, by examining infants' responses across time as a function of interleaved experience, we can determine whether the probabilities infants assign to different hypotheses or expectations about the world change as a function of ongoing experience[12]. Moreover, evidence regarding the naive utility calculus has primarily focused on infants' application of this calculus to their implicit expectations and representations of others' behaviour (that is, infants infer others' goals based on the costs of the actions required to pursue those goals[11]); our work extends this framework by examining whether such a calculus may also guide infants' own, overt behaviour.

In our experiment, we gave 18-month-olds a task in which a rope that was attached to an out-of-reach transparent box could be pulled to reach a toy in the box (Fig. 1). Before the task, infants learned that an adult experimenter could either easily solve the problem (easy condition, $n=32$), found the problem challenging but solvable (hard condition, $n=32$) or was unable to solve the problem, even after working hard at it (impossible condition, $n=32$). These demonstrations thus provided infants with information about the potential difficulty of the task: in the easy condition, infants had evidence that the task should be easy; in the hard condition, infants had evidence that the task should be hard; and in the impossible condition, infants had evidence that the task should be impossible. Critically, these demonstrations also provided infants with information about the competency of their social partner: in the easy condition, infants learned that their social partner is highly capable; in the hard condition, infants learned that their social partner is competent, though not highly skilled; and in the impossible condition, infants learned that that their social partner was not competent.

We next examined infants' persistence on the same task that, unbeknownst to them, was actually impossible to solve (as the container was glued to the table); the only difference across conditions was the type of social input received before being presented with the task. This demonstration–test cycle was repeated three times with three different sets of toys. Finally, to investigate infants' subsequent motivation to solve the task when the task circumstances had changed, we gave infants a 'recovery trial' in which we presented them with a new, out-of-reach toy in a container, with the difference being that they could now freely pull the rope and obtain the toy.

Varying both the level of effort deployed by the adult model (high versus low) and whether the adult model was successful in achieving her goal (yes versus no) enabled us to determine whether infants' reproduction of effortful behaviour was indeed rational—that is, occurred selectively in the face of success but not failure—and additionally provided a test of whether infants apply a naive utility calculus to their own actions (by selectively undertaking costs when the likelihood of rewards was high). In advance of previous work[13], examining infants' persistence in a single moment of time, we measured infants' persistence across a cycle of three demonstration–action trials to gauge their changing expectations as a function of both what they saw and what they did. This temporal unfolding allowed us to test infants' ability to dynamically update their expectations as a function of ongoing experience in a manner that is highly ecologically valid, given that observational and first-hand experience are often interleaved in the real world. We implemented four different convergent dependent measures to shed light on the nature of infants' persistence and its consequences. Because persistence consists of both a temporal component (deciding how long to try) and an intensity component (deciding how hard to try), we measured both the time that infants spent trying and their pulling force. This combination of measures provides insight into the expectations held by infants about the solvability and difficulty of the task. We also assessed infants' positive and negative affect while trying to probe their expectations about how hard/easy the task should be. Negative affect (that is, frustration) signals that there is a mismatch between infants' perceived expectations and reality, whereas positive affect signals that their expectations and reality are aligned. Furthermore, we investigated infants' help-seeking from the experimenter to gauge their inferences about the experimenter's ability to solve the task; infants should seek help more frequently from experimenters that they perceive to be skilled. Finally, to examine the downstream motivational consequences of infants' experiences, we examined how much support infants required during the recovery trial when the task changed and was now possible to solve.

Our central aim—to understand how infants dynamically integrate observed and first-hand information in deciding when and how to act—was reflected in our statistical approach. We examined the interaction between condition and trial number across each of our dependent measures. When a significant interaction between condition and trial number emerged, we conducted planned post hoc comparisons to examine the effect of trial number on infants' behaviour within each condition. We hypothesized that infants would deploy effort rationally and that how long they tried for, how hard they pulled the rope and their attempts to recruit others would be guided by their evolving inferences about the difficulty of the task, as informed by the social evidence they received and their own experiences.

Results

Infants' time spent trying across trials varied as a function of condition ($t(189)=2.44$, $P=0.02$, unstandardized regression coefficient (b)$=0.19$, 95% confidence interval (CI)$=0.04$–0.33; Fig. 2). In the impossible and easy conditions, trying time decreased dramatically across trials ($t(63)=-4.56$, $P<0.001$, $b=-0.23$, 95% CI$=-0.33$ to -0.13 and $t(63)=-5.94$, $P<0.001$, $b=-0.30$, 95% CI$=-0.40$ to -0.20, respectively; see Fig. 2). In contrast, in the hard condition, infants' trying time remained relatively constant across time ($t(63)=-1.96$, $P=0.05$, $b=-0.12$, 95% CI$=-0.23$ to 0.00; effect sizes: easy condition: Cohen's $d=0.93$; impossible

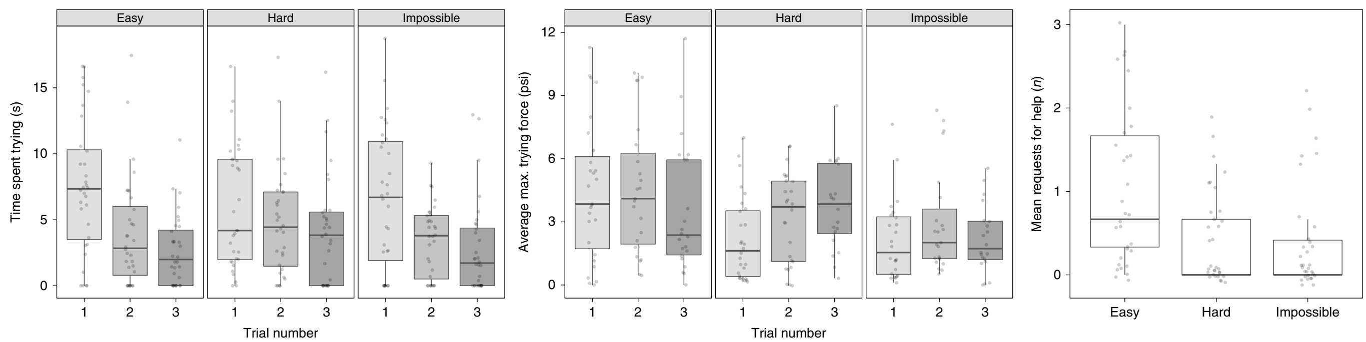

Fig. 2 | Infants' trying behaviour across all three indices of trying, conditions and test trials. Time spent engaging with task (left, $n = 288$ trials, condition by trial number interaction: $t(189) = 2.44$, $P = 0.02$, $b = 0.19$, 95% CI = 0.04–0.33); maximum pulling force on the rope in pounds per square inch (psi) (centre, $n = 213$ trials, condition by trial number interaction: $t(143.59) = 3.22$, $P = 0.002$, $b = 0.18$, 95% CI = 0.07–0.28); frequency of help-seeking behaviour (right, $n = 96$ infants, hard versus easy condition: $t(92) = -3.08$, $P = 0.003$, $b = -0.57$, 95% CI = −0.94 to −0.20; impossible versus easy: $t(92) = -3.19$, $P = 0.002$, $b = -0.59$, 95% CI = −0.96 to −0.22). The horizontal line dividing each box represents the median, the box ends show the upper and lower quartiles, and points represent individual data points. Data are jittered slightly on the horizontal axis to avoid overplotting.

condition: Cohen's $d = 0.66$; hard condition: Cohen's $d = 0.32$; see Fig. 2). Infants' maximum pulling force across trials also interacted with condition ($t(143.59) = 3.22$, $P = 0.002$, $b = 0.18$, 95% CI = 0.07–0.28). Infants exerted consistently low levels of force in the impossible condition ($t(46.9) = 0.11$, $P = 0.91$, $b = 0.004$, 95% CI = −0.07 to 0.08; Fig. 2); consistently high levels of force in the easy condition ($t(46.35) = -1.57$, $P = 0.12$, $b = -0.06$, 95% CI = −0.14 to 0.02; Fig. 2); and increasingly high levels of force in the hard condition ($t(50.67) = 2.99$, $P = 0.004$, $b = 0.11$, 95% CI = 0.04–0.19; Fig. 2). Thus, infants use social input to form expectations of task difficulty, which are then revised in light of their own experience and subsequently guide when and how hard they try.

What do these findings reveal about infants' persistence as a function of social evidence and their own experience? First, they indicate that infants do not merely reproduce the actions of the experimenter: although infants in both the hard and impossible conditions saw effortful actions from the experimenter, those in the impossible condition showed decreasing time spent trying and low force across trials, whereas infants in the hard condition demonstrated consistent trying time and, increasing force across trials. Thus, infants' persistence is rational and may be guided by a utility calculus, resulting in their trying more and/or harder when given evidence that effort results in success. Second, these results demonstrate that infants integrate their own ongoing experience with information gleaned from observation in a dynamic, ongoing fashion. Infants exerted the most force in the easy condition, when they should be confident about the probability of success. Coupled with the fact that, in this condition, infants' overall time spent trying became increasingly short-lived across the experiment, these results suggest that infants believed that if an initial, intense attempt did not work then they were unlikely to solve the task. In the impossible condition, when failure seemed likely, infants withheld their effort. Infants in this condition engaged with the task only briefly and, when they did, exerted low levels of force, probably because they were producing more tentative or exploratory pulls on the rope. In contrast, in the hard condition, when infants have reason to believe that the task should be hard but ultimately solvable, they sustained their trying and ramped up their force over time, suggesting that they thought the source of their difficulty was insufficient quality of effort. The fact that infants initially exerted relatively low levels of force, then steadily increased the amount of force deployed over time, suggests that they may adopt a conservative approach in the face of uncertainty (that is, in terms of exactly how hard the task will be) by putting forth the minimum amount of effort required to achieve success and incrementally increasing their effort.

We next investigated infants' help-seeking behaviour in the form of manual gestures toward the experimenter; specifically, we investigated whether infants selectively sought help in the easy condition when provided with evidence that the experimenter was skilled at the task. On test trials, infants in the easy condition demonstrated high levels of help-seeking, whereas those in the hard and impossible conditions engaged in low levels of help-seeking (hard versus easy conditions: $t(92) = -3.08$, $P = 0.003$, $b = -0.57$, 95% CI = −0.94 to −0.20; impossible versus easy: $t(92) = -3.19$, $P = 0.002$, $b = -0.59$, 95% CI = −0.96 to −0.22; Fig. 2). These findings demonstrate that infants uniquely sought help when their social partner had expertise at solving the task (that is, in the easy condition), and also that infants in this condition were probably decreasing their own trying in favour of trying to solicit help from the experimenter. We next investigated whether infants consider the interaction between the experimenter's competence and their own ability to solve the task. We compared infants' help-seeking in situations when they actually required help (that is, during test trials when the task was impossible) to situations in which no help was needed (that is, during recovery trials when infants could freely solve the task on their own). Infants in the easy condition showed large decreases in help-seeking behaviour ($t(29) = -3.85$, $P < 0.001$, $b = -0.03$, 95% CI = −0.05 to − 0.01, Cohen's $d = 0.70$), and those in the hard condition showed moderate decreases in help-seeking-behaviour ($t(30) = -2.69$, $P = 0.01$, $b = -0.01$, 95% CI = −0.02 to −0.00, Cohen's $d = 0.48$). In contrast, infants in the impossible condition showed uniformly low rates of help-seeking across the test and recovery trials ($t(29) = -2.03$, $P = 0.05$, $b = -0.01$, 95% CI = −0.02 to −0.00, Cohen's $d = 0.37$). These findings demonstrate that infants monitor their own abilities and those of others and more frequently seek help when (1) they actually need it and (2) their social partner's competence at the task exceeds their own. Though it is possible that infants' help-seeking behaviours were simply requests to engage the experimenter in the activity, the finding that they exclusively sought out help when help was actually needed—and when the experimenter was highly capable of providing help—suggests that these gestures functioned to request help from the experimenter.

To provide more precise insight into infants' expectations about how difficult the task should be, we also measured the degree to which their trying behaviour was accompanied by positive or negative affect: we hypothesized that infants would exhibit greater negative affect when they expected the task to be easy to solve but it was not. Infants' negative and positive affect systematically varied as a function of the nature of the social input and their own experience with the task (condition by trial number interaction for

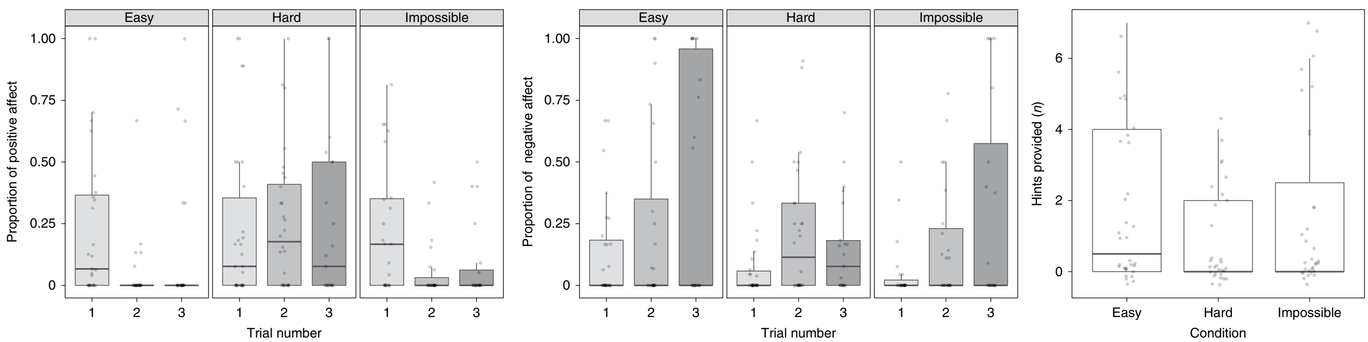

Fig. 3 | Infants' affect and recovery trial support across conditions. Frequency of positive affect across test trials (left, $n = 218$ trials, condition by trial number interaction: $Z = -5.46$, $P < 0.001$, $b = -1.22$, 95% CI = −1.67 to −0.79); negative affect across test trials (centre, $n = 218$ trials, condition by trial number interaction: $Z = 2.57$, $P = 0.01$, $b = 0.63$, 95% CI = 0.16–1.12); and number of hints provided during the recovery trial (right, $n = 94$ infants: easy versus hard condition, $Z = 2.71$, $P = 0.007$, $b = 0.65$; 95% CI = 0.19–1.13; hard versus impossible condition, $Z = 2.62$, $P = 0.009$, $b = 0.63$, 95% CI = 0.17–1.11). The horizontal line dividing each box represents the median, the box ends show the upper and lower quartiles, and points represent individual data points. Data are jittered slightly on the horizontal axis to avoid overplotting.

negative affect: $Z = 2.57$, $P = 0.01$, $b = 0.63$, 95% CI = 0.16–1.12; positive affect: $Z = -5.46$, $P < 0.001$, $b = -1.22$, 95% CI = −1.67 to −0.79; Fig. 3). Infants in the easy condition exhibited the most negative affect, growing dramatically more frustrated across the duration of the experiment ($Z = 7.34$, $P < 0.001$, $b = 1.00$, 95% CI = 0.74–1.27) and reducing their positive affect ($Z = -4.45$, $P < 0.001$, $b = -0.95$, 95% CI = −1.38 to −0.55). In the impossible condition, infants initially demonstrated low levels of negative affect—they knew the task should be impossible and therefore were not frustrated or surprised when it was, in fact, impossible. However, by the end of the experiment, after continually being given a problem they knew they could not solve, they grew moderately frustrated ($Z = 6.51$, $P < 0.001$, $b = 1.35$, 95% CI = 0.96–1.77) and reduced their positive affect ($Z = -5.71$, $P < 0.001$, $b = -1.09$, 95% CI = −1.48 to −0.73). In contrast, infants in the hard condition showed only a modest increase in negative affect ($Z = 4.79$, $P < 0.001$, $b = 0.69$, 95% CI = 0.41–0.98) and maintained positive affect across the duration of the experiment ($P = 0.30$; Fig. 3). These affective reactions also had motivational consequences: during the recovery trial, infants in the easy and impossible conditions required significantly more support from the experimenter to solve the task than those in the hard condition (easy versus hard condition: $Z = 2.71$, $P = 0.007$, $b = 0.65$; 95% CI = 0.19–1.13; hard versus impossible condition: $Z = 2.62$, $P = 0.009$, $b = 0.63$, 95% CI = 0.17–1.11; Fig. 3). Thus, infants' affect is driven by their expectations about task difficulty, not merely by whether or not they can complete the task, and these affective consequences have a lasting impact on their motivation to complete similar tasks.

Discussion

Together, these results demonstrate that infants engage in adaptive persistence: they strategically decide when, whether and how to try based on the social input they receive and their own first-hand experience. These findings underscore the importance of social learning as a complex process in which infants are active agents who dynamically act on and witness events in the world, and integrate this information to rationally decide when and how to learn from others, as opposed to agents who indiscriminately adhere to and follow the behaviours of others. Critically, infants in the current experiment adopted qualitatively different approaches based on their integration of evidence from observed and direct experiences. This research builds on work examining infants' rational imitation, which has established that infants consider contextual cues in deciding who and how to imitate, and moves this field forward by demonstrating that infants simultaneously weigh and integrate information observed from others and that experienced directly—a task most akin to their everyday experiences—to form inferences not only about the a social partner's intentions, but also about social partner competency and task difficulty.

Most critically, this research demonstrates that infants are able to update their expectations about the world around them in real time in light of new and changing evidence, and act in line with principles of computational models of learning. When infants expect tasks to be challenging but solvable they sustain effort, primarily conserve affect and remain motivated to solve future tasks on their own. In contrast, when tasks are unexpectedly difficult but social partners are skilled, infants trade off trying for help-seeking and become increasingly frustrated and demotivated. Indeed, there was a negative relation between the amount of time infants spent trying and the number of help requests they produced ($t(281) = -2.07$, $P = 0.04$ $r = -0.12$, 95% CI = −0.24 to −0.00; Fig. 4) and, critically, this effect appears to be driven by infants in the easy condition, who have reason to believe that seeking help may be a more adaptive strategy than trying directly (correlation between time spent trying and number of help requests, easy condition: $t(89) = -2.57$, $P = 0.01$, $r = -0.26$, 95% CI = −0.44 to −0.06; hard condition: $P = 0.85$; impossible condition: $P = 0.51$; Fig. 4). And, when infants expect tasks to be impossible, they simply reduce their trying and show moderate changes in affect and motivational decreases. In sum, this work highlights that cognition, even in the earliest stages of development, is a complex process that dynamically updates and integrates new evidence.

Though computational models of early learning have been influential in guiding research in developmental science, little empirical work has tested whether these models can accurately predict infants' overt behaviour in the world. Our findings provide support for Bayesian models of learning that argue that infants probabilistically entertain multiple hypotheses about the world and assign probabilities to these hypotheses that are then adjusted via experience[6–9]. Infants' behaviour in the current experiment was structured around these exact principles: they formed hypotheses about the nature of the task based on the social evidence at hand—that is, whether the task should be easy, hard or difficult to solve. Infants then modified these hypotheses in an ongoing manner in light of new experiences and accumulating evidence, as demonstrated by the finding that they dynamically changed, modified and updated their strategies across the duration of the experiment. These results also provide empirical support for a naive utility calculus model of cognition that proposes that social cognition is based on a set of computational principles that maximizes the utility of actions.

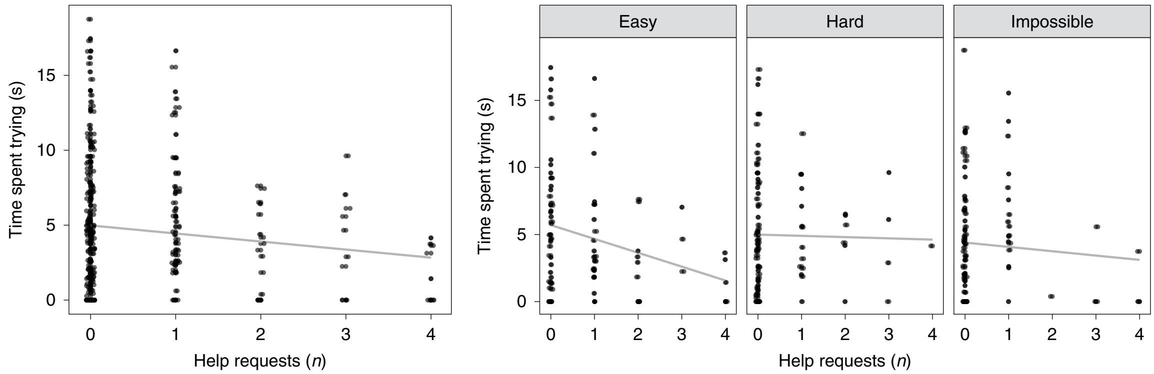

Fig. 4 | Infants' help-seeking behaviour plotted against time spent trying. Data are presented collapsed across all conditions (left, $t(281) = -2.07$, $P = 0.04$, $r = -0.12$, 95% CI = −0.24 to −0.00) and broken down by condition (right, easy condition: $t(89) = -2.57$, $P = 0.01$, $r = -0.26$, 95% CI = −0.44 to −0.06; hard condition: $P = 0.85$; impossible condition: $P = 0.51$). $n = 284$ trials. Data are jittered slightly on the horizontal axis to avoid overplotting.

In the current experiment, infants selectively performed metabolically costly acts to overcome a challenge only when they had evidence that there was a high probability of an achievable reward.

Infants are not only strategic in when and how to try themselves—they are also acutely aware of when it is most adaptive to stop trying altogether and seek help from a social partner. Research has established that, in early childhood, children are highly sensitive to the reliability and competency of social partners[14] and, as early as the end of the second year of life, infants preferentially seek help from a social partner who will consistently and accurately provide them with information[15–18]. Our results demonstrate that this ability is present by 18 months: infants preferentially request help from those who are experts on a task (that is, the experimenter in the easy condition) compared to those who are capable, though not highly skilled (that is, the experimenter in the hard condition). By examining infants' help-seeking in contexts in which help was actually needed versus not needed, the current work also elucidated an advanced level of sophistication in infants' metacognitive abilities. Infants in the current experiment were keenly aware of when help was actually needed, and selectively requested it when they were unable to complete the task themselves; they decreased their help-seeking during the recovery trial when they could freely solve the task themselves. These findings suggest that infants not only monitor their own abilities[19], but also undertake a relative comparison of their own abilities to those of another person, to decide when to seek help from others versus when to act on their own.

The current results also shed light on the intimate relationship between affect and cognition. Specifically, we show how examination of infants' affective responses to challenging situations provides information about the nature of their task representations, as well as their motivation to pursue related tasks in the future. Infants in all conditions received an identical, frustrating and impossible task. Nevertheless, not all infants responded negatively. It was only when there was a mismatch between infants' expectations and their reality (in the easy condition) that there were dramatic affective consequences. Infants in the easy condition, who expected the task to be easy, grew increasingly frustrated across the duration of the experiment whereas infants in the hard and impossible conditions, who expected the task to be difficult, retained positive affect. Critically, these findings provide evidence that it is not merely presenting infants with a difficult task that is frustrating—rather, frustration accrues when infants' expectations about task difficulty are inaccurate. These results not only demonstrate the tight intertwining of cognition and affect, but also highlight the pedagogical importance of setting appropriate expectations, even within infancy.

Although this work builds on previous work in that it assessed infants' changing behaviour over time using multiple converging measures, open questions remain. For example, it will be important to explore whether these effects hold over longer time scales and in infants' home environments. Recent research by Lucca, Horton and Sommerville[20] suggests these results may be stable across time and settings, because the findings demonstrated that infants' persistence is highly correlated across contexts. Thus, although the current work measured infants' persistence in a single, tightly controlled setting, it is reasonable to assume that these behaviours would transfer to more ecologically valid settings, such as a child's home environment. Future research is needed to test this claim and explore the range of factors that influence early persistence.

Overall, our results demonstrate that infants are rational decision-makers[21]: they dynamically integrate social evidence and their own experience to decide when, whether and how to deploy effort. This ability may underlie infants' capacity for rapid and extraordinary learning.

Methods

We have complied with all relevant ethical regulations to conduct this work, which was approved by the Institutional Review Board at the University of Washington: Protocol Title: 'Infants' Persistence: Behaviours Throughout Early Childhood', ID: STUDY00003312. Parents signed a consent form before participation, and consent to publish images of research participants was obtained. Participants received a small toy for participating. Data collection and analysis was not performed blind to the conditions of the experiment; all coding of data was performed blind to the conditions of the experiment.

Participants. Ninety-six full-term, typically developing 18-month-olds (38 females, mean age = 18.50 months, range = 17.67–19.30 months) participated. Participants were recruited from a university-maintained database and identified by their parents as either White ($n = 69$), Asian ($n = 3$), Hispanic ($n = 2$), declined to report ($n = 1$) or mixed race ($n = 21$). Data from 16 additional infants were excluded due to equipment failure ($n = 2$), refusal to participate ($n = 13$: fuss out, $n = 8$, did not interact with warm-up toys, $n = 5$) or parental interference ($n = 1$). No statistical methods were used to predetermine sample size, but our sample size is similar to those reported in previous publications on early persistence[13].

Procedure. In the testing room ($3.2 \times 2.7\,m^2$), infants sat on their parents' lap at a table 40 inches away from the experimenter. The experimental session was divided into four components: (1) warm-up, (2) demonstration, (3) test and (4) recovery. Infants were randomly assigned to one of three between-subject conditions: easy ($n = 32$), hard ($n = 32$) or impossible ($n = 32$), which varied in how much effort the experimenter exerted while solving the problem and whether they were ultimately successfully in solving it. Throughout the experiment, parents wore occluding glasses and were instructed to remain completely neutral and not to interfere with their infants' behaviour in any way during the task. Between each of the four phases, parents faced away from the table towards a wall with four large pictures. Parents were instructed to keep their child from looking as the experimenter set up, and were told they could interact with their infant in any way they chose. This procedure and central hypothesis were pre-registered on OSF (https://osf.io/f87ks) on 8 January 2018. We also collected data from 12-month-olds as part of the same project, but these are not reported in this manuscript. We explain this deviation along with others in a registration on OSF (https://osf.io/y9qaj), as well as in the Supplementary Methods.

Warm-up trials. Infants participated in three warm-up trials. In each warm-up trial, the experimenter presented infants with a single familiar toy. After presenting the toy, the experimenter prompted them to play, touch and interact with the toy for approximately 30 s. After infants had interacted with the toy, the experimenter introduced the next toy. The goal of the warm-up phase was to rule out the possibility that infants' failure to interact with the task during the test trials was not due to shyness or lack of interest in interacting with the experimenter. It also allowed us to acclimate infants to the testing room and experimenter. Infants who failed to interact with a toy in two out of the three warm-up trials were excluded from the final sample ($n = 5$).

Demonstration trials. The purpose of the demonstration trials was for infants to learn how easy or difficult it was to complete the target task (that is, to obtain an out-of-reach toy), as well as the experimenter's ability to successfully solve the task. There was a total of three demonstration trials, each of which was directly followed by a test trial. In the demonstration trial, the experimenter attempted to obtain an out-of-reach toy (located inside a closed transparent container). The container was 1 m outside of the experimenter's reach, and accessible only by pulling a rope that was attached to the container. A different toy was placed inside the container on each trial in a fixed order: a dinosaur bath toy, a yellow knobby ball and a red plastic car, respectively. Without the infant seeing, the experimenter placed the container on the edge of the infant's side of the table and stretched the rope attached to the container across the table with the handle of the rope placed at the edge of the table on the experimenter's side. The experimenter then instructed parents to turn around and to stay away from the table so the infants could not reach the toy. Once parents put on their occluding glasses and infants were visually attentive, the experimenter pointed to the toy and said "Hi <child's name>, I want to get that toy". The experimenter then demonstrated that she could not reach the toy. The experimenter then tried to get the toy by using the rope attached to the container.

In the easy condition, immediately after showing that she could not reach the toy, the experimenter pulled the rope and quickly moved the container across the table, successfully obtaining the toy. In the hard condition, the experimenter attempted to get the toy by pulling on the rope a total of five times before successfully obtaining the toy on her fifth and final try. Each try consisted of the experimenter first pointing to the toy and either saying "I want to get that toy" or "Look <child's name>, I want to get that toy", and then pulling on the rope. The experimenter appeared to exert a great deal of force in attempting to pull the rope. After each unsuccessful try, the experimenter said "hmm", while looking at the container. When the experimenter successfully obtained the toy in the easy and hard conditions, she picked up the toy and said "I got it!" The impossible condition was identical to the hard condition, except that the experimenter did not successfully retrieve the toy on the fifth try, ending the trial by saying "hmm". At the end of each trial, parents were instructed to turn towards the wall, away from the table and experimenter.

Test trials. The purpose of the test trials was to see how long and in what ways infants attempted to get the toy that the experimenter had just demonstrated trying to get. Without the infant looking, the experimenter replaced the board on top of the table with one that had a container, identical to that used in demonstration but glued to the table. Thus, the container was affixed to the table and, unbeknownst to the infant, it was not possible to move the container by pulling the rope (or via any other means). The board was positioned with the container at the edge of the experimenter's side of the table, and the rope stretched across the table with the rope handle placed on the edge of the infant's side of the table. Once the table was secured in place, the experimenter instructed parents to turn towards the table and to get as close to the table as they could. Once the infant was at the table the experimenter started a timer. Throughout the test trial the experimenter did not look in the infant's direction. The experimenter watched the infant from a monitor below the table, and ended the trial either after the infant stopped interacting with the rope for 15 s or 120 s had passed. If the rope fell off the table or became out of reach for the infant, the experimenter placed the rope on the table within the infant's reach.

Recovery trial. After completing the third and final test trial, the infant participated in a recovery trial in which he/she could easily retrieve the toy by pulling on the rope (that is, the container was not affixed to the table top). When the infant was facing away from the table and experimenter, the experimenter set up the table with a container placed on the edge of the experimenter's side of the table and the rope stretched across the table, with the handle at the edge of the table on the infant's side. Once the table was secured in place, the experimenter instructed the parent to turn towards the table and to get as close to the table as they could. Once the infant was at the table, the experimenter started a timer. If the infant did not get the toy or interact with the rope after 15 s, the experimenter prompted the infant to get the toy by saying "<Child's name> it's your turn, can you get ...?" while pointing at the toy. If, after an additional 15 s, the infant did not get the toy or interact with the rope, the experimenter prompted him/her to get the toy just as before and provided a hint by pulling slightly at the rope, showing that the container could be moved forward. The experimenter continued this prompting at 15-s intervals until the infant finally obtained the toy. A maximum of nine hints was provided.

Coding. All coding was done offline by coders who were naive to condition, using Datavyu software (datavyu.org). During the experiment, infants were given varying amounts of time to interact with the task during the test trial as a function of whether they were engaged with the task or not (the experimenter stopped the trial 15 s after the infant stopped engaging with the task). We designed the task in this way so as to not disrupt infants while they were trying to complete the task, which may inadvertently cue them that they shouldn't try in subsequent trials. The shortest trial length of any trial was 20 s. Moreover, after the first 20 s of the trial, infants generally diminished their trying behaviours and began playing with the rope, making the coding of trying actions versus play actions extremely difficult. For these reasons, we limited all coding and analyses to the first 20 s of each test trial. Coders were naive to the experimental condition the infants were in, whereas data collection and analysis were not performed naive to the conditions of the experiments.

Measures of persistence. *Duration of trying behaviour.* A primary coder watched videos of participants and coded, in seconds, the amount of time infants spent trying to pull the rope to obtain the toy. If infants pulled the rope and looked at the target object immediately before, during or after pulling the rope, the coder considered the behaviour to be 'trying'. If infants threw the rope and/or swung it side-to-side without looking toward the target object, the coder coded the behaviour as 'play/not trying'. A secondary coder independently coded 100% of videos to establish an intraclass correlation coefficient of 0.95 ($P < 0.001$).

Strength of trying behaviour. A 5-kg S-type load cell was connected to the toy's rope and hidden inside the target toy to measure infants' pulling strength in pounds per square inch (PSI). The load cell was connected via a discrete wire to a laptop, which stored infants' PSI data. In any given trial in which infants' behaviour was coded as intentionally trying (by a human observer), we extracted infants' maximum PSI. Maximum PSI was extracted during the 20-s period following the experimenter's initiation of the trial.

Additional measures. *Help-seeking behaviour.* A primary coder watched videos of participants and coded the number of communicative bids the infant produced. Based on previous research demonstrating that infants' gestures function as requests for help[22], any reaching or pointing gestures that were directed towards the target object were considered help-seeking communicative bids. We coded only those help-seeking behaviours directed toward the experimenter, because the caregiver had their vision occluded and therefore any behaviour directed toward the caregiver was unlikely to be a signal for help. Help-seeking behaviour was coded during test trials and during the recovery trial. A secondary coder independently coded 100% of the videos to establish an intraclass correlation coefficient of 0.93 ($P < 0.001$).

Affect. Because we were interested in infants' emotional reactivity while engaged in trying behaviours, we coded infants' affect only during moments of trying. During instances when infants' behaviour was coded as trying, coders watched close-up recordings of their faces and coded facial expressions using a coding scheme adapted from Repacholi and colleagues[23]. Facial expressions were rated on a freeze-frame every 500 ms during periods of trying. Though this was rare, it is possible that infants simultaneously displayed features of positive and negative affect. As in previous research[24,25], we limited our coding of affect to infants' facial expressions (as opposed to bodily gestures). For a review on the power of facial expressions in determining affect, see ref. [24].

A facial expression was coded as 'positive' if the infant smiled (for example, upturning of mouth). This was often (though not always) accompanied by cheek elevation and brow raising. If the infant did not display a clear positive affect, the coder rated the facial expression as 'neutral/non-positive'. Fifty per cent of participants were coded by a secondary coder to establish high inter-rater reliability (intraclass correlation coefficient: 0.97, $P < 0.001$).

A facial expression was coded as 'negative' if the infant frowned. This was often accompanied by an anger expression, wrinkled nose, furrowed brow and/or crying. If infants did not display clear negative affect, the coder rated their facial expressions as neutral/non-negative. Fifty per cent of participants were coded by a secondary coder to establish high inter-rater reliability (intraclass correlation coefficient: 0.96, $P < 0.001$).

Number of hints given during recovery. To assess the level of support required by infants to solve the task when it was possible to solve it, a coder watched videos of the recovery trial and coded the number of hints infants were given before solving the problem. Importantly, the number of hints that the experimenter provided was based strictly on the amount of time that passed since the start of the trial, and was not contingent on infants' behaviour or help-seeking. Fifty per cent of participants were coded by a secondary coder to establish high inter-rater reliability (intraclass correlation coefficient: 0.96, $P < 0.001$).

Statistical approach

In the current study, we were interested in testing how infants' problem-solving approaches are impacted by observational information suggesting that the task should be easy, hard or impossible, and by information indicating that the task was either very difficult or impossible to solve, generated from their own experience. To test these questions, we conducted a series of analyses in which we analysed the effect of condition (easy, hard or impossible) and trial number on each of our outcome measures of interest (that is, time trying, trying force, affect, help-seeking). When significant main effects emerged from these analyses, we conducted post hoc contrasts to determine the source of these differences. When a significant interaction between trial number and condition emerged, we conducted within-condition comparisons to test how infants' behaviour changed over time.

For each outcome variable of interest, we fit a mixed-effects model. For variables that followed a linear distribution (that is, time trying, trying force), linear mixed models were fit[26]. These variables were square-root transformed after adding a constant of ten to remove skewness and normalize distribution. For variables that followed a nonlinear distribution (that is, affect, help-seeking), generalized linear mixed models were fit[27]. This approach confers several advantages over traditional repeated-measures analysis of variance. First, it allows us to standardize our analytic approach across all variables of interest, despite the fact that our outcome variables followed different distributions (linear, binomial, Poisson). Second, since many of our outcome variables relied on infants' trying in a given trial (for example, we could not measure force in a given trial if infants did not try), this approach gave us more power to handle the unbalanced design that resulted from missing data within individuals[28].

In each mixed-effects model, we included condition (easy, hard or impossible), trial number (1, 2 or 3) and the interaction between the two terms as fixed-effect predictors. Pseudo-replication, which was caused by three repeated observations per infant, was taken into consideration by including the infant's ID in the model as a random effect. This approach controls for variation in the response driven by multiple consecutive and correlated test trials within an individual. To assess model fit, we compared the fit of the full model (including all predictors of interest) to the reduced model (including only the random effect) using likelihood ratio tests via the anova command. For these model comparisons, we present chi-square statistics, degrees of freedom, change in AIC (ΔAIC) and P values.

For each model that improved in fit relative to the null model, we present the Z- or t-statistic, degrees of freedom, P value, unstandardized regression coefficients (b) and 95% CIs. Cohen's d-values were computed by calculating the mean difference between the first and last trials, and then dividing this value by the pooled standard deviation. In cases in which a significant interaction between condition and trial number emerged, we tested for a simple main effect of trial number within each condition by running a follow-up linear (or generalized linear) mixed model, with trial number as a fixed effect and infant ID as a random effect. If a main effect of trial number was present, we conducted planned one-tailed post hoc contrasts within conditions between the first and third trials. If a main effect of condition was present, we conducted planned one-tailed contrasts across conditions in the first and third trials. We selected only the first and third trials to conserve the number of tests run, and also because we were primarily interested in assessing infants' initial approach to the task and where they ended up performing at the end of the experiment. Outliers detected at s.d. = 3 above/below the mean were removed from analyses. Removal of outliers did not impact the results. Except where noted, data from 32 infants (96 trials) in each condition were included in each analysis (with a total of 288 trials across 96 infants). Analyses were performed in R (v.1.1.463)[29] using the functions anova and the lmer and glmer of package lme4[30]. All reported tests were two-tailed, and data met the assumptions of the statistical tests used. Additional analyses are presented in Supplementary Results

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data are publicly available at https://github.com/klucca/Lucca_et_al_Effort_2019.

Code availability

All code used for the analyses in the manuscript can be found at https://github.com/klucca/Lucca_et_al_Effort_2019.

Received: 22 May 2019; Accepted: 12 December 2019;
Published online: 20 January 2020

References

1. Banerjee, P. N. & Tamis-LeMonda, C. S. Infants' persistence and mothers' teaching as predictors of toddlers' cognitive development. *Infant Behav. Dev.* **30**, 479–491 (2007).
2. Duckworth, A. L., Peterson, C., Matthews, M. D. & Kelly, D. R. Grit: perseverance and passion for long-term goals. *J. Pers. Soc. Psychol.* **92**, 1087–1101 (2007).
3. Eskreis-Winkler, L., Shulman, E. P., Beal, S. A. & Duckworth, A. L. The grit effect: predicting retention in the military, the workplace, school and marriage. *Front. Psychol.* **5**, 1–12 (2014).
4. Gergely, G., Bekkering, H. & Kiraly, I. Rational imitation in preverbal infants. *Nature* **415**, 755 (2002).
5. Meltzoff, A. N. Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol.* **31**, 838–850 (1995).
6. Schulz, L. E. & Bonawitz, E. B. Serious fun: preschoolers engage in more exploratory play when evidence is confounded. *Dev. Psychol.* **43**, 1045–1050 (2007).
7. Gopnik, A. et al. A theory of causal learning in children: causal maps and Bayes nets. *Psychol. Rev.* **111**, 3–32 (2004).
8. Xu, F. & Tenenbaum, J. B. Sensitivity to sampling in Bayesian word learning. *Dev. Sci.* **10**, 288–297 (2007).
9. Gopnik, A. & Wellman, H. M. Reconstructing constructivism: causal models, Bayesian learning mechanisms, and the theory theory. *Psychol. Bull.* **138**, 1085–1108 (2012).
10. Jara-Ettinger, J., Gweon, H., Schulz, L. E. & Tenenbaum, J. B. The naïve utility calculus: computational principles underlying commonsense psychology. *Trends Cogn. Sci.* **20**, 785 (2016).
11. Liu, S., Ullman, T. D., Tenenbaum, J. B. & Spelke, E. S. Ten-month-old infants infer the value of goals from the costs of actions. *Science* **358**, 1038–1041 (2017).
12. Jara-Ettinger, J., Floyd, S., Tenenbaum, J. B. & Schulz, L. E. Children understand that agents maximize expected utilities. *J. Exp. Psychol. Gen.* **146**, 1574–1585 (2017).
13. Leonard, J. A., Lee, Y. & Schulz, L. E. Infants make more attempts to achieve a goal when they see adults persist. *Science* **1294**, 1290–1294 (2017).
14. Kidd, C., Palmeri, H. & Aslin, R. Rational snacking: young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition* **126**, 109–114 (2013).
15. Birch, S. A. J., Akmal, N. & Frampton, K. L. Two-year-olds are vigilant of others' non-verbal cues to credibility. *Dev. Sci.* **13**, 363–369 (2010).
16. Kachel, G., Moore, R. & Tomasello, M. Two-year-olds use adults' but not peers' points. *Dev. Sci.* **21**, e12660 (2018).
17. Zmyj, N., Daum, M. M., Prinz, W., Nielsen, M. & Aschersleben, G. Fourteen-month-olds' imitation of differently aged models. *Infant Child Dev.* **12**, 250–266 (2012).
18. Zmyj, N., Buttelmann, D., Carpenter, M. & Daum, M. M. The reliability of a model influences 14-month-olds' imitation. *J. Exp. Child Psychol.* **106**, 208–220 (2010).
19. Goupil, L., Romand-Monnier, M. & Kouider, S. Infants ask for help when they know they don't know. *Proc. Natl Acad. Sci. USA* **113**, 3492–3496 (2016).
20. Lucca, K., Horton, R. & Sommerville, J. Keep trying!: parental language predicts infants' persistence. *Cognition* **193**, 104025 (2019).
21. Xu, F. & Kushnir, T. Infants are rational constructivist learners. *Curr. Dir. Psychol. Sci.* **22**, 28–32 (2013).
22. Goubet, N., Rochat, P., Maire-Leblond, C. & Poss, S. Learning from others in 9–18-month-old infants. *Infant Child Dev.* **15**, 161–177 (2006).

23. Repacholi, B. M., Meltzoff, A. N., Hennings, T. M. & Ruba, A. L. Transfer of social learning across contexts: exploring infants' attribution of trait-like emotions to adults. *Infancy* **21**, 785–806 (2016).
24. Messinger, D. S. & Messinger, S. Positive and negative: infant facial expressions. *Psychol. Sci.* **11**, 1–6 (2002).
25. Ekman, P. & Friesen, W. V. *Facial Action Coding System: Investigator's Guide* (Consulting Psychologists Press, 1978).
26. West, B. T., Welch, K. B. & Galecki, A. T. *Linear Mixed Models: A Practical Guide using Statistical Software*. (Chapman and Hall/CRC, 2014).
27. McCullagh, P. & Nelder, J. A. *Generalized Linear Models* (Chapman & Hall, 2008).
28. Baayen, R. H., Davidson, D. J. & Bates, D. M. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* **59**, 390–412 (2008).
29. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2014).
30. Bates, D. & Maechler, M. lme4: Linear mixed-effects models using S4 classes. R version 0.999375-35 https://github.com/lme4/lme4/ (2010).

Acknowledgements

We thank the researchers of the Early Childhood Cognition Lab at the University of Washington for their help in participant recruitment, data collection and coding. In particular, we thank A. Sedlacek, Y. Xu, J. Lee, S. Cho, K. McManus, M. Rozaniti, K. Ventura and P. Carpentier. We also thank B. Kuykendall for his assistance with the force gauge installation and data extraction, as well as the families and infants who participated in this study. This work was supported by a grant from the Society of Research on Child Development awarded to K.L., and through a grant from NICHD (no. 1R01HD076949-01) awarded to J.A.S. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

K.L., R.H. and J.A.S. made substantial contributions to the conception and design of the work. R.H. collected the data. K.L. coordinated data coding and performed all data analyses. J.A.S. provided critical oversight and feedback of the work. K.L. and J.A.S. wrote the manuscript. R.H. provided critical feedback.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0814-0.

Correspondence and requests for materials should be addressed to K.L.

Corresponding author(s): Kelsey Lucca

Last updated by author(s): 11/11/2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All behavior was coded using Datavyu software. Infants' force data was collected using MM01 MultiDAQ software.
Data analysis	Analyses were performed in R (Version 1.1.463; R Development Core Team, 1999) using the functions anova and the lmer and glmer of the package lme4 (Bates, Maechler, & Dai 2010).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data are made publicly available here: https://github.com/klucca/Lucca_et_al_Effort_2019

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This was a quantitative experimental study.
Research sample	Ninety-six full-term, typically developing 18-month-olds (38 females, M = 18.50 months, range = 17.67 - 19.30) participated. Participants were recruited from a university-maintained database and identified by their parents as White (n = 69), Asian (n = 3), Hispanic (n = 2), declined to report (n = 1), and mixed race (n = 21).
Sampling strategy	Sample sizes were selected based on prior research. Infants were selected from the university-maintained database based on their age.
Data collection	Pen and paper were used to record participant metadata, SONY video cameras were used to record participant behavior, a force gauge was used to measure infants' trying behavior, and a computer stored infants' force data. The main experimenter was not blind to experimental condition and study hypotheses, but the researchers who coded the data were blind to experimental condition and study hypotheses. The only person present during the procedure aside from the participant and experimenter was the participant's parent, who were occluding glasses so they were naïve to the experimental procedure.
Timing	Data collection began on January 18, 2018 and ended on August 17, 2018.
Data exclusions	Data exclusion criteria were pre-established. Outliers more than 3 standard deviations above the mean were removed from analysis (the exact number for each analysis is reported in text).
Non-participation	Exclusion criteria were preestablished. Data from 16 additional infants were excluded due to equipment failure (n = 2), refusal to participate (n = 13: fuss out, n = 8, did not interact with warm up toys, n = 5), or parental interference (n = 1).
Randomization	Participants were randomly assigned to one of three experimental conditions.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☐	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology
☒	☐	Animals and other organisms
☐	☒	Human research participants
☒	☐	Clinical data

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☒	☐	MRI-based neuroimaging

Antibodies

Antibodies used	NA
Validation	NA

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above.
Recruitment	See above for recruitment information. Our participants were primarily white, which may impact the generalizability of the findings to other racial groups.
Ethics oversight	We have complied with all relevant ethical regulations to conduct this work, which was approved by the IRB at the University of Washington: Protocol Title: "Infants' Persistence: Behaviors Throughout Early Childhood", ID: STUDY00003312.

Note that full information on the approval of the study protocol must also be provided in the manuscript.