Science AAAS

Pure Reasoning in 12-Month-Old Infants as Probabilistic Inference
Erno Téglás, *et al.*
Science **332**, 1054 (2011);
DOI: 10.1126/science.1196404

Pure Reasoning in 12-Month-Old Infants as Probabilistic Inference

Ernő Téglás,[1,2]* Edward Vul,[3]* Vittorio Girotto,[4,5] Michel Gonzalez,[5]
Joshua B. Tenenbaum,[6]† Luca L. Bonatti[7]†

Many organisms can predict future events from the statistics of past experience, but humans also excel at making predictions by pure reasoning: integrating multiple sources of information, guided by abstract knowledge, to form rational expectations about novel situations, never directly experienced. Here, we show that this reasoning is surprisingly rich, powerful, and coherent even in preverbal infants. When 12-month-old infants view complex displays of multiple moving objects, they form time-varying expectations about future events that are a systematic and rational function of several stimulus variables. Infants' looking times are consistent with a Bayesian ideal observer embodying abstract principles of object motion. The model explains infants' statistical expectations and classic qualitative findings about object cognition in younger babies, not originally viewed as probabilistic inferences.

Exploiting statistical regularities of the environment predictively, to adapt behavior to future events, is a basic strategy in biology. Humans, even in infancy, use observed frequencies to learn words (*1*, *2*), spatiotemporal patterns (*3*, *4*), and visual object features (*5*). Adults can make rational statistical judgments on the basis of frequencies of previously experienced events or summaries of event frequencies (*6*, *7*). Even nonhuman mammals and many other organisms use experienced statistical regularities to modify their behavior (*8*, *9*).

However, humans also excel at reasoning about novel situations, flexibly combining abstract knowledge and perceptual information from disparate sources in "one-shot" intuitions to predict outcomes of events they have never before directly experienced. We call this ability "pure reasoning" to distinguish it from more data-driven means of forming expectations on the basis of statistical learning or finding patterns from repeated exposures. To see more clearly the difference between these two prediction modes and the one-shot nature of pure reasoning, consider several configurations of colored blocks arranged on a table (Fig. 1) and the following judgment: If the table is bumped so that one block falls off, is it more likely to be red or yellow? If the blocks are all close to the edges but there are twice as many red blocks (Fig. 1A), you will probably respond "red." However, if the yellow blocks are precariously stacked and located near the edge (Fig. 1B), you will likely respond "yellow." If the red blocks are located closer to the edge (Fig. 1C), you may again respond "red," although with less confidence, but you may revert to yellow if the precarious yellow stack in the center is doubled in height (Fig. 1D). People can make these judgments the first time they see these displays and make such common-sense predictions in a near-infinite variety of real-world situations, confidently and quickly. To do so by using only basic statistical learning mechanisms, recording over many such scenes all the ways any number of objects can be arranged and all the ways they tend to fall, would be almost impossible. Pure reasoning provides a more powerful and flexible approach (*10*). Considering the number of blocks of each color and their locations, together with intuitive knowledge of physical factors determining how they will move, a reasoner can construct a sample of ways that each configuration of blocks might fall if the table is bumped and then observe, purely within the imagined likely possibilities, which outcome appears most probable. This ability to flexibly combine multiple sources of information and knowledge to predict how a complex situation will unfold is at the core of human intelligence and is one of the biggest missing links in building artificial intelligence systems with humanlike "common sense."

Our goal is to probe the roots of this remarkable ability in human infants. Infants' reasoning abilities are typically studied by measuring their looking times to visually presented events as an index of surprise: Longer looking indicates greater violation of infants' expectations relative to their prior knowledge or greater novelty relative to their interpretation of habituation stimuli. Looking-time studies suggest that preverbal infants can reason about novel events depending on certain physical outcomes (*11*, *12*); object numerosities (*13*); other agents' beliefs, goals and behaviors (*14*–*16*); and the likely outcomes of simple random processes (*17*, *18*).

However, the richness, power, and coherence of infants' reasoning about future events remain unclear. Is it anything like the ability of adults to extrapolate likely future states for scenes such as those in Fig. 1? When presented with a complex dynamic environment with multiple objects that move and may be hidden from view as they move for several seconds, do infants form appropriate expectations about where different objects are likely to be observed at future times? Do their expectations about the future vary in systematic and rational ways over different initial configurations of objects in space and different temporal intervals for prediction?

We addressed these questions by using a combination of novel experiments and computational models (*19*). Our experiments independently varied several features of dynamic displays, such that forming appropriate expectations required infants to quantitatively integrate multiple sources of perceptual evidence with optimal weights that vary over time. Across these manipulations, we kept test events fixed and equal in salience so that infants' looking times had the potential to show variations in degrees of belief (or conversely, degrees of surprise) as their expectations changed.

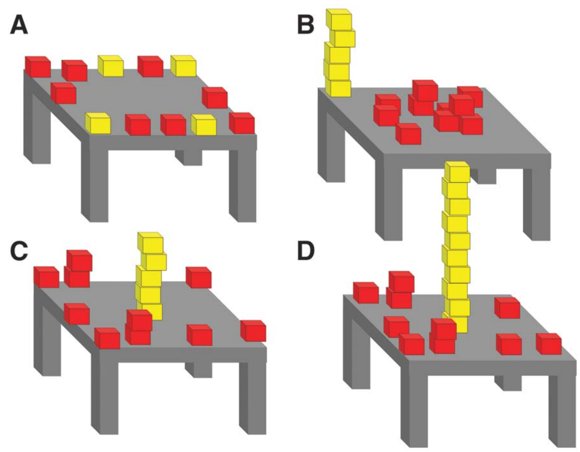

Fig. 1. Examples of common-sense predictions based on pure reasoning**.** If the table in this scene is bumped so that one block falls off the table onto the floor, is it more likely to be a red or a yellow block? Intuitions will vary according to the number of blocks of each type (**A**), their arrangement into more- or less-precarious stacks and their locations on the table (**B**), and interactions between all these factors (**C** and **D**).

[1]Cognitive Development Centre, Central European University, H-1015 Budapest, Hungary. [2]Hungarian Academy of Sciences, H-1015 Budapest, Hungary. [3]University of California, San Diego, La Jolla, CA 92093, USA. [4]University IUAV of Venice, 30135 Venice, Italy. [5]CNRS and Aix-Marseille University, 13331 Marseille, France. [6]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [7]ICREA (Institució Catalana de Recerca i Estudis Avançats) and Universitat Pompeu Fabra, 08018 Barcelona, Spain.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: lucabonatti@mac.com (L.L.B.); jbt@mit.edu (J.B.T.)

We describe a Bayesian ideal observer model that predicts infants' looking times in our studies and extends to other aspects of infants' reasoning about the physical world, giving a unifying explanation of several classic results in infant cognition. This model shows how powerful pure reasoning capacities could derive from the operation of probabilistic inference mechanisms constrained by abstract principles of how objects act and interact over time.

Pure reasoning at 12 months. We probed preverbal infants' expectations about unknown future events when they witness dynamic scenes that contain multiple potentially relevant–but also potentially conflicting–sources of information, similar to (but simpler than) the examples in Fig. 1. Infants viewed movies in which four objects of two types, identified by different shapes and colors, bounced randomly inside a container with an opening on its lower side (movies S1 to S5). After several seconds of observed motion, an occluder covered the container's contents from view for some duration between 0 and 2 s. Finally, one object visibly exited through the bottom opening, and the occluder faded out. Monitoring infants' looking time to this final outcome allowed us to assess how surprised infants were to see an object of either type exit first.

Twelve kinds of movies were generated by manipulating three factors relevant to predicting these outcomes: the number of objects of each type in the scene (three instances of one type and one of the other type), their physical arrangement (objects of one type were always closer to the exit before occlusion than objects of the other type), and the duration of occlusion (0, 1, or 2 s). Forming correct expectations here requires the ability to integrate these three information sources, guided by abstract knowledge about how objects move: at a minimum, qualitative knowledge about solidity (objects are unlikely to pass through walls) and spatiotemporal continuity (objects tend to move short distances over brief time intervals). Infants appear to be sensitive to each of these information sources and knowledge systems individually (*11*, *20*). We asked whether they can also integrate them rationally to predict single future events.

A rational prediction of which object type will exit first should depend on both the number and the physical arrangement of the object types, but the relative importance of these factors should vary with occlusion duration. After a very brief occlusion, the objects' locations before occlusion are most predictive of which object type will exit first; however, when the occlusion is prolonged, proximity to the exit matters less because the objects continue moving in the container. Eventually, after a sufficiently long occlusion, only the number of each object type should be predictive.

In each of three experiments, infants saw four displays varying in whether the object that exited first belonged to the type with one or three instances and whether that type was near or far from the exit before occlusion. Occlusion duration was varied across experiments (Fig. 2A). Mean looking times (M) across all 12 displays showed exactly the rational pattern of predictions described above (Fig. 2, B to D). In experiment 1, with longest occlusion times (2 s), infants looked longer when the single unique object exited the container first [($M_{3\text{-instances}}$) = 11.9 s, $M_{1\text{-instance}}$ = 15.6 s; $F(1, 19)$ = 5.66, P = 0.028 under a repeated measures analysis of variance (ANOVA)], but distance from the exit had no effect [M_{Near} = 13.5 s, M_{Far} = 14.2 s; $F(1, 19)$ = 0.69, P = 0.42]. In experiment 2, with intermediate occlusion times (1 s), infants considered both factors, looking longer at the unique object outcome [$M_{3\text{-instances}}$ = 11.8 s, $M_{1\text{-instance}}$ = 15.0 s; $F(1, 19)$ = 4.65, P = 0.04] and also when an object located far from the opening before occlusion exited first [M_{Near} = 11.6 s, M_{Far} = 15.1 s; $F(1, 19)$ = 5.22, P = 0.03]. In experiment 3 with occlusion time of 0.04 s, looking times were insensitive to type numerosity [$M_{3\text{-instances}}$ = 14.0 s, $M_{1\text{-instance}}$ = 12.4 s; $F(1, 19)$ = 0.65, P = 0.43] but were significantly longer when an object far from the exit left the container first [M_{Near} = 10.2 s, M_{Far} = 15.7 s; $F(1, 19)$ = 16.5, P = 0.0007]. Numerosity and distance did not interact in any experiment [$F(1, 15)$ = 0.007, P = 0.93; $F(1, 17)$ = 2.09, P = 0.17; $F(1, 13)$ = 1.2, P = 0.29, respectively], suggesting that infants tended to consider both cues additively.

A Bayesian model of infants' pure reasoning. These experiments show that infants possess surprisingly sophisticated abilities to integrate multiple information sources and abstract knowledge in reasoning about future outcomes. We now analyze infants' expectations more quantitatively by comparing them with those of a Bayesian ideal observer equipped with only minimal computational resources and the minimal abstract knowledge about physical objects that, according to classic research, young infants possess.

The observer's knowledge of object dynamics is expressed in the form of a probabilistic model embodying the principles of solidity and spatiotemporal continuity described above. These principles can be formalized as a prior $P(S_t|S_{t-1})$ on how the state S_t of the world at time t depends

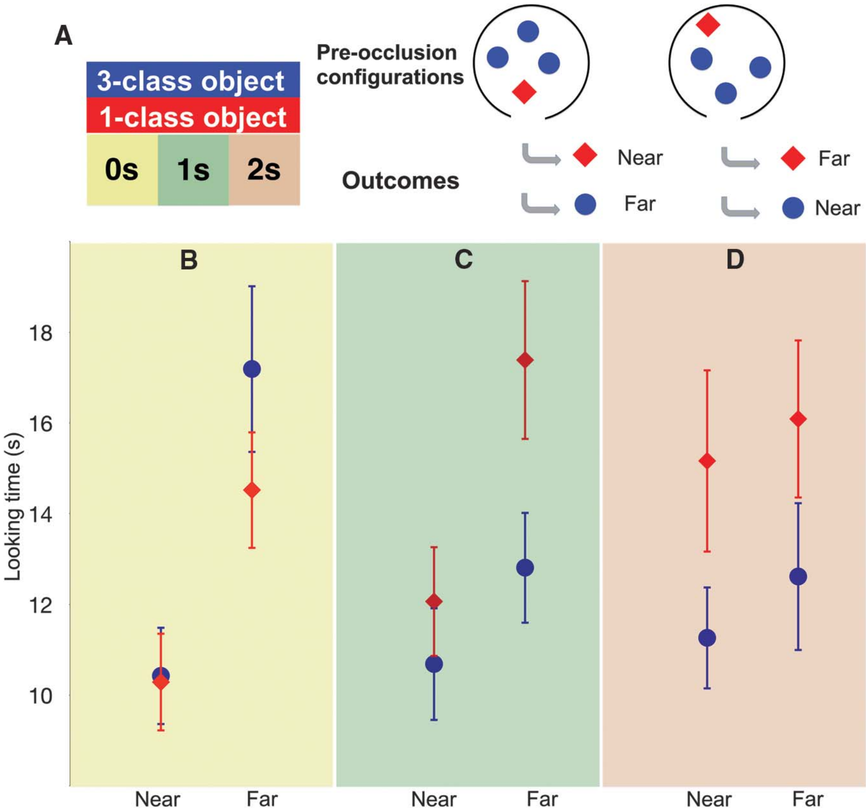

Fig. 2. Experiments probing infants' expectations in dynamic physical scenes. (**A**) Infants saw three objects of one type and one object of another type bouncing randomly inside a container. After some time, an occluder masked the objects, and one of four outcomes occurred: An object exited the container through the bottom opening that was either the common object kind or the unique object, with a position before occlusion that was either far from or near to the exit. The graph reports mean looking time (s, with SEM) of three experiments varying the duration of occlusion before the outcome. (**B**) After a short (0.04 s) occlusion, infants considered only the physical distance in forming their expectations, disregarding the number of objects of each type. (**C**) When occlusion duration was increased to 1 s, infants' looking times reflected both the number of objects of each type and their distance from the exit. (**D**) When the occlusion was longer still (2 s), infants' looking times reflected only the numerosities of each object type, regardless of their preocclusion distance from the exit.

probabilistically on the state at time $t-1$, which for simplicity we express as constrained Brownian motion: Objects move by accumulating small independent random spatial perturbations over time, subject to the constraint that they cannot pass through solid barriers (fig. S1).

The observer must also be equipped with some mechanism of inference and some notion of computational resources. Following state-of-the-art approaches in artificial intelligence and Bayesian models of adult cognition (*21–24*), we assume that predictions are computed approximately by Monte Carlo sampling. This process corresponds to a kind of hypothetical reasoning: Given a particular observed scenario, the observer has the capacity to consider possible future states of the world as they may unfold according to the observer's probabilistic model. A similar intuition for grounding probabilistic reasoning in representations of possible worlds was the basis for classic "mental models" accounts of adult cognition (*25*), although our treatment differs in explicitly formalizing probabilistic principles of knowledge representation and inference. Formally, the probability of a final outcome D_F given the observed data $D_{0,\ldots,F-1}$ is approximated as a sum of the scores of K hypothetical trajectories (sequences of states $S_{0,\ldots,F}$),

$$P(D_F|D_{0,\ldots,F-1}) \propto \sum_{k=1}^{K} P(D_F|S_F^k) \times \prod_{t=1}^{F} P(D_{t-1}|S_{t-1}^k)P(S_t^k|S_{t-1}^k) \qquad (1)$$

where the score is a product over time steps t of how well the kth hypothesis fits the observed data $P(D_t|S_t^k)$ and how probable it is under the prior on object dynamics $P(S_t^k|S_{t-1}^k)$. Intuitively, an observed outcome is expected insofar as many predicted future trajectories are consistent with it or unexpected if it is consistent with few predicted trajectories.

In this analysis, computational resources correspond to the number of hypothetical trajectories (the samples) that an observer can construct. In the limit of infinite samples, these Monte Carlo predictions correspond exactly to the posterior beliefs of the ideal Bayesian observer. This ideal observer forms expectations about which object will emerge first that are very similar to the pattern of looking times exhibited across our three experiments, trading off the influences of type numerosity and proximity, modulated by occlusion duration, just as infants do (Fig. 3). Note that because infants' looking times are typically inversely related to expectations, we compare looking times to $1 - P$(outcome) (*26*). Evaluated quantitatively, the modeled outcome probabilities explain 88% of the variance in infants' mean looking times across the 12 experimental conditions ($r = 0.94$, $df = 10$, $P < 0.0001$). By comparison, each of the three stimulus factors that we manipulated explains significantly less variance across these 12 conditions: occlusion duration, 1%; type numerosity, 12%; and proximity, 47%. Even their best linear combination explains only 61% of variance, with the added cost of two ad hoc free parameters.

In contrast to this analysis, infants—or, indeed, adults (*22–24*)—are unlikely to consider more than a small sample of possible trajectories. Accordingly, we have also analyzed the model under severe resource bounds, by using just one or two trajectories sampled from the Bayesian posterior to form expectations. Averaged over simulated participants and trials, this bounded model makes inferences almost identical to the Bayesian ideal (figs. S5 and S6) [$r(10) = 0.92$, $P < 0.05$; $r(10) = 0.93$, $P < 0.05$]. Thus even with very limited processing capacity, infants could make appropriate probabilistic predictions in our task.

Modeling infants' probabilistic and physical intuitions. If infants' expectations in our experiments truly reflect the origins of a broad "common sense" physical reasoning capacity and if this capacity is captured by our Bayesian model, then the same model should be able to account for expectations about a wider range of developmental situations.

Recent studies have suggested that infants and young children understand simple random processes. Observing the random drawing of some balls from a box containing differently colored balls, infants expect colors in the sample to be representative of proportions in the larger population, and vice versa (*18*). Probabilistic expectations may also be induced by the structure of environmental constraints, not only the distribution of object properties. For instance, when 3- and 5-year-olds (*17*) and 12-month-olds see a single ball bouncing within a bounded box containing three exits on one side and one on the opposite side, they anticipate that the ball will exit from the three-exit side; however, if the three-exit side is obstructed, such anticipation is absent. Our model explains all these results with no further assumptions (fig. S2 and Fig. 4). Spatiotemporal continuity as captured by the Brownian motion

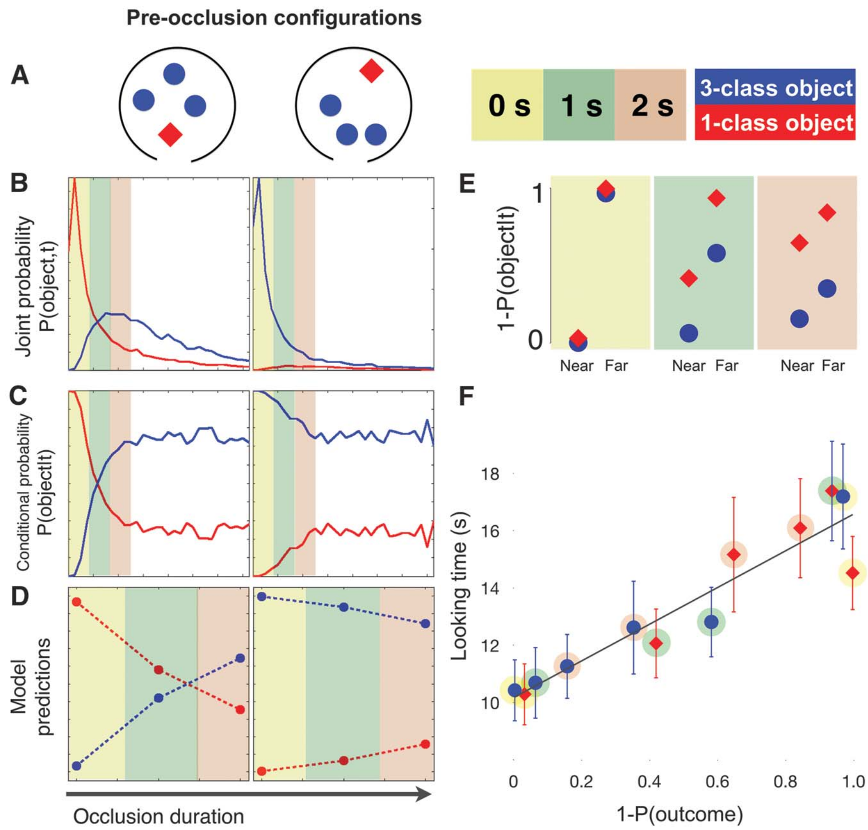

Fig. 3. The ideal Bayesian observer model. Starting with an unambiguous parse of the world into the two types of objects and their preocclusion positions (**A**), the model predicts the probability for each object type to be the first to exit as a function of occlusion duration and preocclusion distance from the exit. (**B**) The joint probability that a particular type of object exits at a particular point in time can be computed from a large number of Monte Carlo samples for each of the two starting scenarios. (**C**) Given the observation that an object first emerges at a particular time, we compute the conditional probability that it is of one type or another. (**D**) The predictions for our experiment consider only three points from the continuous distributions over time, corresponding to short (0 s; yellow), medium (1 s; green), and long (2 s; red) occlusion delays. (**E**) We combine these conditional probabilities from both starting scenarios to predict the joint effects of distance, object type numerosity, and occlusion duration on infants' expectations about which object type will emerge first, as found in experiments 1 to 3 (compare with looking-time data shown in Fig. 2, B to D). (**F**) Correlation between the model predictions (*x* axis) and infant looking times (*y* axis, s with SEM) in our three experiments. Each data point corresponds to one experimental condition.

prior drives the basic expectations about randomness, whereas the solidity constraint on Brownian motion incorporates the physical restrictions on possible or likely outcomes for each display.

Under the same physical principles, our model also explains classic findings on how young infants use visual motion to parse the world into a determinate number of objects. Infants' expectations were not originally viewed as rational probabilistic inferences nor analyzed quantitatively, but our model shows how they can be understood in these terms. In one well-known class of ambiguous displays (*27*), a foreground object occludes what could be two disconnected shorter objects or a single longer object with two parts extending above and below the occluder (Fig. 5A). When the display is static, infants show no preference for the one-object or two-object interpretation, but when the two parts move synchronously behind the occluder, infants expect they form a single object and are surprised if shown that they are two separate objects (Fig. 5B). Our model predicts this result by virtue of its stochastic prior on object motion and a natural version of Occam's razor that results from Bayesian inference under such a prior. It is certainly possible for the two parts to move synchronously left and right if they are two independent objects, but it is a coincidence: just one of many possible ways that two objects moving independently and randomly could move and thus relatively unlikely. However, synchronous motion must always occur if the two parts are two sides of a single rigid object, and thus a one-object interpretation receives much higher posterior probability for a Bayesian observer (Fig. 5C). When the objects are stationary, in contrast, there is essentially no evidence either way, and the observer is indifferent.

In another class of ambiguous displays, an object emerges from alternate sides of a single large occluding screen or a split screen with a visible gap in the middle (*28*) (Fig. 5D). Only one object is ever visible at a time, but the motion could be produced by either a single object traveling behind the occluder(s) or else two objects successively emerging from opposite sides. Infants are surprised to see this scene with the visible gap if they have previously seen only a single object placed behind the screen (*28*) (Fig. 5, D and E) but not if they have previously seen two objects. Our ideal observer forms the same expectations: Two objects can easily produce this motion without appearing in the visible gap, but a single object can do so only if it takes a physical jump across the gap in a single time step, which is possible but highly unlikely under the Brownian motion prior (Fig. 5F). Moreover, both infants and our model make the inverse inference: Seeing an object emerge from both sides of the split screen without appearing in the visible gap, they expect there to be two objects behind the screen rather than one (*29*, *30*) (fig. S4). Lastly, both infants and our model can use the spatiotemporal relations between the speed of an object, the size of the occluder, and the duration of the delay between an object's disappearance behind an occluder and reappearance on the other side to infer the likely existence of one or two objects (*31*) (fig. S3).

Across all the studies described above, our model is able to capture the main ordinal trends in how infants' looking times vary by condition, but it does not provide the stronger quantitative fits that we found for our experiments (Fig. 3F). Differences in the model's subjective probabilities are often too extreme relative to the observed differences in infants' looking times (e.g., Figs. 4, E and F, and 5, E and F). There are several possible reasons for this, which suggest ways that future modeling and experimental studies of infant reasoning could be improved. Previous experimental work adopted a variety of designs intended only to uncover qualitative effects of binary stimulus manipulations on looking times; typical studies report data from few conditions, often using qualitatively different test events across experiments or conditions. In contrast, our experiments parametrically varied multiple dynamical aspects of scenes while keeping test events fixed, allowing for a more sensitive test of model predictions. More quantitative predictions for the classic phenomena described above could be tested with novel parametric designs such as those used in our experiments.

There are also several ways that our modeling approach can be improved and refined. Assuming a linear relationship between outcome probabilities and looking times is too simple, and future work should explore more complex, nonlinear dependencies such as log-likelihoods or information-theoretic measures of surprise (*32*). The rich literature on infants' object perception suggests a need for more sophisticated ideal-observer models, with a more detailed specification of how infants represent the physical properties of objects (*33*, *34*). Lastly, an ideal-observer model, even with resource and processing constraints, provides only a coarse approximation to the psychological mechanisms of infant cognition. A more fine-grained processing model might make stronger quantitative predictions for a broader range of experimental designs. Still, it is intriguing that, with only minimal assumptions about infants' computational resources and their knowledge of object motion, we can explain some of the most basic, early developing abilities to parse the world into a discrete set of objects as rational probabilistic inferences.

Conclusion. Preverbal infants' ability to reason about complex unseen events is surprisingly sophisticated: 12-month-olds can represent the crucial spatial, temporal, and logical aspects of dynamic scenes with multiple objects in motion and integrate these cues with optimal context-sensitive weights to form rational expectations consistent with a Bayesian observer model. Although classic work in judgment and decision-making has suggested that people often fail to follow Bayesian principles in deliberate, explicit reasoning, Bayesian models have recently provided compelling accounts of more intuitive, implicit inference and prediction abilities in adults and older children (*35*, *36*). The present studies carry this approach further back to the roots of cognition by demonstrating a systematic relation between infants' looking times and rational probabilistic expectations in a complex task.

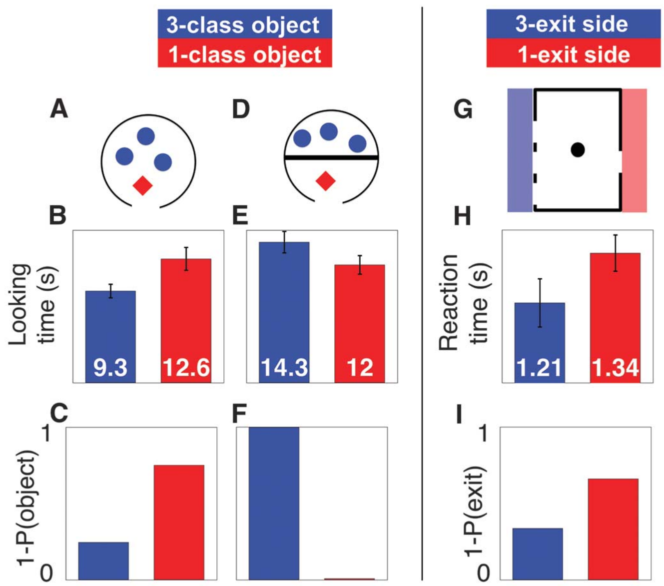

Fig. 4. Model predictions for infants' expectations about random events. (**A**) Schematic representation of experiments from (*17*). Infants saw scenes similar to those of the current experiments, with long preexit occlusions (2 s). (**B**) Infants' mean looking times were longer when an object of the less-numerous type exited (means in s, with SEM). (**C**) The model's predictions for the relative probabilities of which object type is more likely to exit. (**D**) When a barrier separated the objects of the more numerous kind from the exit, (**E**) infants looked longer when one of these objects exited, indicating that they took the physical constraints of the scene into account. (**F**) The solidity constraint of our model yields matching predictions: The common objects are unlikely to "jump" over the barrier and thus unlikely to exit. In another paradigm (**G**), older children see a ball bouncing inside a box with three exits on one side and one exit on the opposite side. After the display is occluded, the ball exits the box. (**H**) Children are slower to react when the ball emerges from the one-exit side. (**I**) The model's constrained Brownian motion predicts that the ball is more likely to exit from the three-exit side.

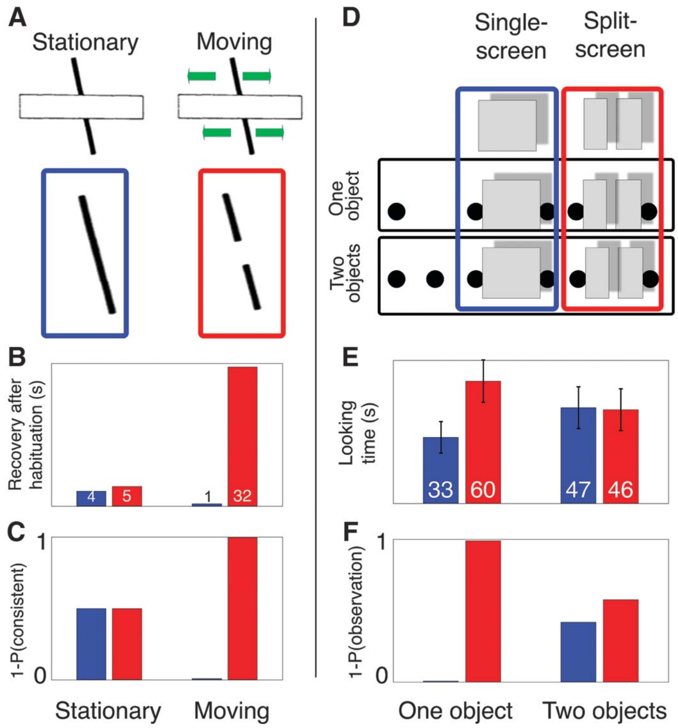

Fig. 5. Model predictions for object cognition in infants. (**A**) Infants as young as 3 months were habituated to two bars behind an occluder that either remained stationary or moved synchronously (*27*). (**B**) After the occluder was removed, infants recovered more quickly from habituation (s) when they saw two independent bars than when they saw a conjoined bar, but only when the bars moved synchronously during habituation (right), not when they were stationary (left). (**C**) The model considers two possible world parses (two independent objects or one conjoined one), but synchronous motion is much less likely for two independent objects. In the stationary condition, two- and one-object parses are equally likely, but in the motion condition the single-object parse becomes much more likely. (**D**) The 2.5-month-old infants were habituated to either one object or two objects, which were then occluded behind one large or two small screens (*28*). Infants saw an object emerge from the left side of the screen and then an object emerge from the right side. (**E**) Infants looked longer when an object emerged from both sides in the two-screen, one-object event, but not in the other events (means in s, with SEM). (**F**) Given an initial unambiguous parse of the display into objects and the one- or two-screen layout, the model samples possible trajectories in which an object emerges on the left followed by the right side of the display without emerging elsewhere. When one object and two screens are present, it is very unlikely for an object to emerge from the right side without being seen in the middle; when two objects are present, or if one large screen is present, this observation is much more likely.

It is unclear how exactly the workings of our model correspond to the mechanisms of infant cognition, but the strong model fits suggest at least a qualitative similarity between the two. We suggest that the commonality lies in the ability to generate physically plausible candidates for future world states, consistent with the observed present. More work is needed to discover the precise form of the representations infants use to effectively construct and weight these hypotheses.

How do these sophisticated inferential abilities arise in development? We have emphasized their one-shot nature: Just as with adults' expectations in Fig. 1, infants' sensitivity to graded outcome probabilities in our displays varies systematically and rationally with the numerosities of different object types, their spatial configuration, and occlusion duration but does not depend on seeing these displays many times as needed for traditional statistical learning. However, the statistics of an infant's experience could still play a role in how this capacity is constructed. Pure reasoning requires the ability to represent the space of possible future events (*37*), as well as some abstract knowledge of how physical objects move, and, although the relevant physical knowledge could be innate (*12*), it could also be acquired through a structured statistical learning approach (*38*) from patterns of object movement (*39*).

More broadly, our work opens the possibility that a cognitive architecture based on probabilistic generative models, complemented by abstract knowledge representations, can account for early common-sense reasoning beyond the limited physical domain we explored. Equipped with richer abstract knowledge, this approach can explain a wide range of common-sense predictions as the result of pure reasoning, such as the ability to make rational inferences about other agents on the basis of core psychological principles (*15*, *40*, *41*) or the ability to reason about more challenging physical situations (as in the displays of Fig. 1) on the basis of more sophisticated physical principles. It thus offers a way to explain how developing humans may come to make progressively richer inferences about their increasingly complex world.

References and Notes

1. J. R. Saffran, E. L. Newport, R. N. Aslin, *J. Mem. Lang.* **35**, 606 (1996).
2. M. Peña, L. L. Bonatti, M. Nespor, J. Mehler, *Science* **298**, 604 (2002).
3. J. Fiser, R. N. Aslin, *J. Exp. Psychol. Learn. Mem. Cogn.* **28**, 458 (2002).
4. S. C. Creel, E. L. Newport, R. N. Aslin, *J. Exp. Psychol. Learn. Mem. Cogn.* **30**, 1119 (2004).
5. D. Kersten, P. Mamassian, A. Yuille, *Annu. Rev. Psychol.* **55**, 271 (2004).
6. L. Cosmides, J. Tooby, *Cognition* **58**, 1 (1996).
7. G. Gigerenzer, U. Hoffrage, *Psychol. Rev.* **102**, 684 (1995).
8. M. D. Hauser, E. L. Newport, R. N. Aslin, *Cognition* **78**, B53 (2001).
9. J. M. Toro, J. B. Trobalón, *Percept. Psychophys.* **67**, 867 (2005).
10. V. Girotto, M. Gonzalez, in *Norms in Human Development*, L. Smith, J. J. Vonèche, Eds. (Cambridge Univ. Press, Cambridge, 2006), pp. 220–236.
11. R. Baillargeon, E. S. Spelke, S. Wasserman, *Cognition* **20**, 191 (1985).
12. E. S. Spelke, K. Breinlinger, J. Macomber, K. Jacobson, *Psychol. Rev.* **99**, 605 (1992).
13. K. Wynn, *Nature* **358**, 749 (1992).
14. A. L. Woodward, *Cognition* **69**, 1 (1998).
15. K. H. Onishi, R. Baillargeon, *Science* **308**, 255 (2005).
16. A. M. Kovács, E. Téglás, A. D. Endress, *Science* **330**, 1830 (2010).
17. E. Téglás, V. Girotto, M. Gonzalez, L. L. Bonatti, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19156 (2007).
18. F. Xu, V. Garcia, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5012 (2008).
19. Materials and methods are available as supporting material on *Science* Online.
20. S. P. Johnson *et al.*, *Child Dev.* **74**, 94 (2003).
21. S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach* (Prentice Hall, Englewood Cliffs, NJ, ed. 3, 2009).
22. S. D. Brown, M. Steyvers, *Cognit. Psychol.* **58**, 49 (2009).
23. E. Vul, M. Frank, G. Alvarez, J. Tenenbaum, in *23rd Annual Conference on Neural Information Processing Systems 2009*, Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, A. Culotta, Eds. (Advances in Neural Information Processing Systems 22, Curran, Red Hook, NY, 2009), p. 1955.
24. R. Levy, F. Reali, T. L. Griffiths, in *22nd Annual Conference on Neural Information Processing Systems 2008*, D. Koller, D. Schuurmans, Y. Bengio, L. Bottou, Eds. (Advances in Neural Information Processing Systems 21, Curran, Red Hook, NY, 2008), p. 937.
25. P. N. Johnson-Laird, P. Legrenzi, V. Girotto, M. S. Legrenzi, J.-P. Caverni, *Psychol. Rev.* **106**, 62 (1999).
26. Assuming that relative differences in looking times reflect the relative strengths of infants' expectations given the initial conditions, we normalize these outcome probabilities to sum to 1 over the set of test outcomes that infants saw. Although an information theoretic (logarithmic) measure of surprise may be more general, in this case, where looking times to only one of two alternatives are compared, $1 - P$ is a natural metric to evaluate the relative expectations of these alternatives.
27. P. J. Kellman, E. S. Spelke, *Cognit. Psychol.* **15**, 483 (1983).
28. A. Aguiar, R. Baillargeon, *Cognit. Psychol.* **39**, 116 (1999).
29. F. Xu, S. Carey, *Cognit. Psychol.* **30**, 111 (1996).
30. E. S. Spelke, R. Kestenbaum, *Psychol. Fr.* **31**, 67 (1986).
31. T. Wilcox, A. Schweinle, *Infant Behav. Dev.* **26**, 253 (2003).
32. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, Hoboken, NJ, ed. 2, 2006), pp. xxiii, 748.
33. C. Kemp, A. Jern, F. Xu, in *23rd Annual Conference on Neural Information Processing Systems 2009*, Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, A. Culotta, Eds. (Advances in Neural Information Processing Systems 22, Curran, Red Hook, NY, 2009), p. 925.
34. R. Baillargeon, J. Li, W. Ng, S. Yuan, in *Learning and the Infant Mind*, A. Woodward, A. Needham, Eds. (Oxford Univ. Press, Oxford, 2009), pp. 66–116.
35. T. L. Griffiths, J. B. Tenenbaum, *Psychol. Sci.* **17**, 767 (2006).
36. J. B. Tenenbaum, T. L. Griffiths, C. Kemp, *Trends Cogn. Sci.* **10**, 309 (2006).
37. J. A. Fodor, *The Language of Thought* (Crowell, New York, 1975).

38. J. B. Tenenbaum, C. Kemp, T. L. Griffiths, N. D. Goodman, *Science* **331**, 1279 (2011).
39. P. Sinha, B. Balas, Y. Ostrovsky, J. Wulff, in *Object Categorization: Computer and Human Vision Perspectives*, S. J. Dickinson, A. Leonardis, B. Schiele, M. J. Tarr, Eds. (Cambridge Univ. Press, Cambridge, 2009), pp. 301–323.
40. G. Gergely, G. Csibra, *Trends Cogn. Sci.* **7**, 287 (2003).
41. C. L. Baker, R. Saxe, J. B. Tenenbaum, *Cognition* **113**, 329 (2009).

Acknowledgments: The research was supported by grants Friuli-Venezia-Giulia "PsyScope XL," Ministerio de Ciencia E Innovación PSI2009-08232PSIC, the James S. McDonnell Foundation Causal Learning Research Collaborative grant, Office of Naval Research grants N00014-09-0124 and N00014-07-1-0937, Army Research Office MURI W911NF-08-1-0242, the Swiss and Global-Fondazione Ca' Foscari, and Marie Curie 035975 Disorders and Coherence of the Embodied Self. E.T. and L.L.B. designed the experiments and analyzed the data. E.T. ran the experiments. J.B.T. and E.V. conceived the model and the simulations. E.T., E.V., J.B.T., and L.L.B. wrote the paper. All authors discussed the research and revised the text and the analyses. We thank D. Amati, L. Filippin, F. Gandolfo, N. Goodman, A. Isaja, and N. Sebastian for support and suggestions. J. Mehler and the Language, Cognition, and Development Laboratory in Trieste, Italy made the experimental work possible. We are deeply grateful to them.

Supporting Online Material
www.sciencemag.org/cgi/content/full/332/6033/1054/DC1
Materials and Methods
Figs. S1 to S8
Movies S1 to S5

12 August 2010; accepted 21 April 2011
10.1126/science.1196404

Experimental Repetitive Quantum Error Correction

Philipp Schindler,[1] Julio T. Barreiro,[1] Thomas Monz,[1] Volckmar Nebendahl,[2] Daniel Nigg,[1] Michael Chwalla,[1,3] Markus Hennrich,[1]* Rainer Blatt[1,3]

The computational potential of a quantum processor can only be unleashed if errors during a quantum computation can be controlled and corrected for. Quantum error correction works if imperfections of quantum gate operations and measurements are below a certain threshold and corrections can be applied repeatedly. We implement multiple quantum error correction cycles for phase-flip errors on qubits encoded with trapped ions. Errors are corrected by a quantum-feedback algorithm using high-fidelity gate operations and a reset technique for the auxiliary qubits. Up to three consecutive correction cycles are realized, and the behavior of the algorithm for different noise environments is analyzed.

Information in a quantum computer is extremely vulnerable to noise induced by the environment and thus needs to be protected with quantum error correction (QEC) techniques. Pioneering theoretical work in this field has shown that all errors can be corrected for if imperfections of the quantum operations and measurements are below a certain (error) threshold and the correction can be applied repeatedly (*1*–*3*). Such error thresholds depend on details of the physical system, and quantifying them requires a careful analysis of the system-specific errors, the en- and decoding procedures, and their respective implementation (*4*). It is currently accepted that gate error probabilities ranging from 10^{-4} to 10^{-5} are tolerable (*5*), which seem to be in reach with technical improvements in conjunction with dynamical control techniques (*6*). In addition, fault-tolerant operation requires highly efficient, repeatable algorithms to minimize the computational overhead. So far, all experimental implementations (*7*–*12*) are limited to a single correction cycle, where the only experimental implementation in a scalable system (*10*) relies on projective measurements and classical feedback. Because high-fidelity measurements take time and potentially disturb the qubit system, it can be advantageous to use a measurement-free QEC algorithm based on implicit quantum feedback (*4*, *7*). Also, in contrast to previous expectations (*13*), these measurement-free protocols lead to error thresholds comparable to those of their measurement-based counterparts (*14*).

We demonstrate repeated QEC with a system of trapped $^{40}Ca^{+}$ ions as qubits, and multiple repetitions of the algorithm are enabled by a toolbox consisting of high-fidelity quantum operations (*15*, *16*), an optimized pulse sequence (*17*), and a qubit-reset technique that has a negligible effect on the system of qubits. The performance of the implementation is assessed with quantum process tomography in the presence of phase-flip errors, and its behavior is analyzed for different environments that show correlated and uncorrelated phase noise. Our approach is based on the three-qubit repetition code capable of detecting and correcting phase-flip errors on a single qubit (*1*, *4*). This algorithm protects against phase noise, which is the dominant error source in our ion-trap quantum computer, causing gate errors as well as decoherence.

As indicated in Fig. 1A, each QEC cycle consists of (i) encoding the system qubit $\{|0\rangle, |1\rangle\}$ and two auxiliary qubits (ancillas) into an entangled state, (ii) error incidence, (iii) detecting and correcting the error, and (iv) resetting the ancillas. Initially, the qubit to be protected is in the state $|\Psi\rangle = \alpha|+\rangle + \beta|-\rangle$, where $|\pm\rangle = 1/\sqrt{2}(|0\rangle \pm |1\rangle)$, and the two ancilla qubits are both prepared in the state $|1\rangle$. In the encoding stage, they are mapped into the entangled state $\alpha|+++\rangle + \beta|---\rangle$. Next, a single-qubit phase-flip error may change $|\pm\rangle$ to $|\mp\rangle$. In the decoding and correction stage, the error is identified by a simple majority vote, and the system qubit is corrected accordingly. It should be noted that this protocol maps the information in and out of the protected state between QEC cycles. Each cycle is concluded by resetting the ancilla qubits while preserving the information on the system qubit.

The textbook implementation of a single cycle of this QEC procedure would consist of a circuit using four controlled-NOT (CNOT) and one controlled controlled-NOT (Toffoli) gate operations (*4*) (Fig. 1B). Although the process fidelities of available CNOT (92%) (*18*) and Toffoli (80%) (*19*) implementations could possibly be improved, it seems more promising to pursue an approach based on global Mølmer-Sørensen entangling gate operations (fidelity of 99%) (*15*, *20*). These operations provide a universal set of gates in combination with individually addressed Stark-shift gates and collective single-qubit rotations (*17*, *21*). Moreover, the optimization procedure of (*17*) allows us to rigorously simplify the pulse sequence for a complete algorithm based on this set of gates. Two additional refinements lead to the algorithm used for the optimization (Fig. 1B). First, the space of optimized solutions is increased by adding an arbitrary unitary operation, U, acting only on the ancillas before resetting them. Second, the encoding stage can be simplified by adding an operation, D, and its inverse, D^{-1}, that commutes with any phase error. As a result, the encoding stage consists of a single entangling operation, and the decoding stage can be implemented with a total of eight pulses with only three entangling operations (Fig. 1C). Formally, this encoding implements a stabilizer code with the generators $G = \{\sigma_y^{(1)}\sigma_z^{(2)}\sigma_y^{(3)}, \sigma_y^{(1)}\sigma_y^{(2)}\sigma_z^{(3)}\}$, which are tensor products of the Pauli operators $\sigma_{x,y,z}^{(i)}$ acting on qubit i (*4*).

The QEC protocol is realized in an experimental system consisting of a string of three $^{40}Ca^{+}$ ions confined in a macroscopic linear Paul trap. Each

[1]Institut für Experimentalphysik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria. [2]Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria. [3]Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, Otto-Hittmair-Platz 1, A-6020 Innsbruck, Austria.

*To whom correspondence should be addressed. E-mail: markus.hennrich@uibk.ac.at