Chimpanzees spontaneously prepare for mutually-exclusive possibilities and collective context strengthens this behavior

Elizabeth Warren[1,*] , Charlie Pilgrim[2] , Mélisande Aellen[3] , Joe Morford[3] , Esther Herrmann[4] ,
Richard P. Mann[2] , Dora Biro[3] , Christopher Krupenye[1]

1. Department of Psychological & Brain Sciences, Johns Hopkins University, Baltimore, MD, 21218, USA
2. Department of Statistics, School of Mathematics, University of Leeds, Leeds, LS2 9JT, UK
3. Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA
4. School of Psychology, Sports and Health Sciences, University of Portsmouth, Portsmouth, PO1 2UP, UK

ORCIDs: EW (0000-0003-1183-4521), CP (0000-0002-3800-677X), MA (0000-0002-8408-0268), JM (0000-0003-3687-1783), EH (0000-0002-6791-0805), RPM (0000-0003-0701-1274), DB (0000-0002-3408-6274), CK (0000-0003-2029-1872)

*Correspondence: ewarren8@jh.edu

Abstract

In both humans and non-human animals, collectives can sometimes overcome individual cognitive biases or shortcomings to execute more rational behavior than individuals. To investigate differences in strategy and outcome between individuals and collectives in a logical reasoning task, we presented an inverted U-shaped tube to individuals and pairs of chimpanzees (*Pan troglodytes*) and examined their preparatory actions toward rewards which could fall from either end of the tube. Given that individual chimpanzees have typically produced a suboptimal one-handed strategy in past variants of this task, we predicted that pairs would outperform individuals primarily through subjects sharing the apparatus, each placing one hand under one tube end such that they collectively account for both possible outcomes. Unexpectedly, over half of our chimpanzees spontaneously produced the optimal two-handed behavior (covering both ends) on their own, providing evidence that *individuals* may be able to reason about mutually-exclusive future possibilities. This reduced the capacity for pairs to improve upon individual performance. Notably, however, we observed an increase in individual usage of the two-handed strategy in the collective setting. This individual improvement may have arisen from an effect of collective facilitation such as competition, suggesting an alternative mechanism through which collectives may outperform individuals.

Keywords: primates, cognition, rationality, logic, reasoning, collective action, cooperation, social behavior, decision-making, animal behavior

Introduction

For millennia, scholars have mused about the origins of rational behavior and whether it exists outside of humans (1). Individual organisms are faced with a plethora of decisions every day, spanning the mundane to the life-changing. These decisions can be rational, in the senses adopted by biology and economics, if they give rise to observable actions which achieve optimal fitness outcomes (B-rationality) or maximize utility and consistency (E-rationality)(2). Rationality – conceptualized through this behavioral, outcome-based lens – is widespread in the animal kingdom, where individuals and groups optimize their behavior toward improved fitness, resource acquisition, or optimal energy intake and expenditure. On the other hand, decisions can be rational, in the sense adopted by philosophy and psychology, if they recruit decision-making processes that involve reasoning, rather than only those involving heuristics (2). Heuristic processing, a system of decisional "shortcuts" based on rules and patterns (3), is fast and less cognitively effortful, but may be prone to pitfalls such as cognitive biases and illusory thinking. Reasoning, conversely, involves representing and reflecting on reasons for one's actions, which often entails considering multiple possible consequences of different choices and selecting the most optimal decision given the decision-maker's specific needs, ethics, and beliefs. This process, known as PP-rationality, is slower and more cognitively demanding, but less prone to reflexive errors. A diversity of reasoning abilities exist in humans, from logic to theory of mind, and it remains debated whether any of these capacities are shared with other species (2,4–6).

Fascinatingly, evidence has started to emerge that collectives can sometimes exhibit rational decision-making, at least in the B- or E- sense of rationality, even when their constituent members do not. For example, in a landmark discovery, Sasaki and Pratt (7) found that individual ants were susceptible to the 'decoy effect' – where the addition of an extra choice alternative changes an existing preference among options, producing an 'economically irrational' decision – yet colonies of ants resisted the decoy and reverted to rationality (in this case, consistency in preference). This is due not only to the simple aggregation of group members' sensory, processing and memory powers, but also to properties that emerge from the interaction of decision rules individuals use for combining their own information with that of others (8–13). This work shows that in some decision-making contexts, collectives can achieve biological forms of rationality, in this case in optimizing preferences and outcomes (collective B- or E-rationality), that individuals cannot. While this study has become a seminal proof-of-concept demonstration of emergent collective B/E-rationality in animal groups, a further open question, which we begin to tackle here, is whether collective decision-making can also produce 'reasoning-like' intelligence (i.e., collective rationality in the PP-sense) even in systems composed of agents that purportedly do not have the capacity to reason individually. That is, can we find evidence of emergent forms of collective reasoning in non-human animals?

Broadly, these questions fall within the study of "collective behavior", which describes scenarios where two or more individuals - each influenced by the choices and actions of others in the group - produce a joint behavioral outcome mediated by local-scale interactions among group-mates (14–16). We define "emergent collective reasoning" as a subset of collective behaviors in which forms of reasoning that are not present at the individual level emerge at the collective level, specifically when the process by which the collective pools inputs from individual minds resembles the process of reasoning that typically occurs within a single mind. In a paradigmatic case of emergent collective reasoning, individual members of a collective only possess heuristic mechanisms for decision-making, but, out of the system of interactions through which they combine their information and make collective decisions, a reasoning-like process emerges. For example, reasoning often involves mentally simulating multiple available options out of which the individual chooses the one it deems the most optimal. A similar process could emerge at the collective level even if each individual is only capable of simulating a single option, so long as their interactions provide a mechanism for making decisions over the many simulations that are collectively represented across individuals. While the emphasis here is on characterizing the process by which information from the minds of multiple individuals is aggregated into an emergent reasoning-like phenomenon, this kind of 'emergent collective reasoning' may also secure more optimal outcomes for collectives than those available to individuals. Notably, while emergent collective reasoning requires the development of a reasoning-like process across individuals, it nonetheless in principle could arise under a broad range of circumstances that vary in terms of the degree to which the collective's actions must be physically, sequentially, or temporally coordinated, the degree to which the outcome is equally distributed or identical across all participating individuals, and the degree to which the participating members individually mentally represent the behavior as collective.

We distinguish emergent collective reasoning from two other mechanisms of collective improvement related to reasoning. First, if the presence or absence of individual reasoning abilities varies across members of a species, larger collectives may simply be more likely to contain one or more individuals in possession of those reasoning capacities. In this case, which we term collective scaling of individual reasoning, the presence of more competent individuals may expand with group size and increase the likelihood that the decision of the collective is informed by these individuals' proficient reasoning, and therefore achieve more optimal outcomes. Alternatively, in the case of collective facilitation of individual reasoning, it is possible that collective contexts could elevate the motivations of individuals to deploy reasoning abilities that they, again, already possess, and thus produce a greater prevalence of *individual* decisions being shaped by reasoning. Here, although individuals may more reliably deploy their bonafide reasoning abilities in collective contexts, and collectives may secure more optimal outcomes as a result, the fundamental process resembling reasoning is taking place in the minds of individuals, rather than emerging across individuals.

As noted earlier, an open question in the comparative literature is whether even individual reasoning is unique to humans remains an open question in the comparative literature. Chimpanzees are an intriguing model to explore this capacity outside of humans, given their broadly sophisticated cognitive abilities and their position as one of humans' two closest living relatives (17–21). Moreover, their many collectively performed behaviors (hunting, coalitionary aggression, territory border patrols, etc.) (22,23) and their potential demonstration of collective reasoning in naturalistic hunting contexts (see below) make them ideal candidates for examining the possibility of individual or collective reasoning beyond humans. Although chimpanzees can and do make independent decisions driven by their own egocentric needs and desires, they live in cohesive and broadly cooperative groups that nest, travel, feed and breed together. As such, they exhibit numerous collective behaviors that result in jointly-experienced outcomes, although the benefits of these outcomes are not always shared equally among members of the collective. For example, a group of chimpanzees may cooperatively hunt a colobus monkey, and although not all participating hunters may receive equal or even any part in the prey, the ultimate capture of the monkey was the result of the combination of the individuals' efforts, even if they do not all receive equal benefit from the division of resources.

Although primates engage in a variety of collective behaviors, cooperative hunting of colobus monkeys by chimpanzees is an activity that provides a particularly useful lens through which to investigate reasoning processes, since it may involve chimpanzees simulating the potential escape paths of their prey. At Ngogo, in Kibale National Park, Uganda, larger hunting parties are more likely to successfully capture a monkey (24), and, at Taï Forest, in Taï National Park, Côte d'Ivoire, hunters are more likely than bystanders to receive a share of the meat from the captor, suggesting that these acts are cooperative (25,26). In at least some chimpanzee communities (e.g., Kasekela, in Gombe National Park, Tanzania, and Kanyawara, in Kibale National Park, Uganda), hunts are often spurred by specific individuals, termed 'impact' hunters, who initiate hunts at disproportionate rates (27). When a hunt begins, one or more chimpanzees ascend the trees toward a target. According to Boesch (28), who extensively studied this phenomenon, others initially "remain on the ground *in anticipation of the possible escape routes* of the colobus and ready themselves to join the pursuit" (28, italicized emphasis added). Boesch (28) describes these other hunters as either placing themselves in other trees "to block the progression of the prey," directly chasing the prey and attempting to catch it, or ambushing the prey by quietly positioning themselves in trees along the monkey's anticipated escape route. These blocking and ambushing behaviors raise the possibility that chimpanzees could be exercising reasoning capacities for modal logic (i.e., reasoning about possibilities) at the individual level: individuals might simulate multiple mutually exclusive possible future events (i.e., future escape attempts of the monkey) and coordinate their coverage of these possible escape routes. However, coverage of all escape routes could in principle also be achieved, without individuals possessing the logical reasoning capacity for simulating mutually exclusive possibilities, by a form of emergent collective reasoning: if each individual observes the locations of its fellow hunters and anticipates and blocks one remaining possible escape route, the group may be able to collectively represent and block all possible escape routes.

In captive experiments, chimpanzees have not always shown robust individual reasoning abilities of this variety. In an elegant task developed by Redshaw and Suddendorf (29), a reward is dropped down an inverted Y-shaped tube, such that it may emerge from either of two ends. Subjects capable of mastering the logical operator “OR” (the reward may emerge from this end OR that end) and of anticipating multiple, mutually exclusive future events should plan accordingly by placing one hand beneath each end. Although three-year-old human children often fail this task, it is typically mastered by age four, a result that has been replicated across several task variations and multiple cultures (30,31). In the classic version of this task, chimpanzees perform comparably to three-year-olds, typically pursuing an apparent heuristic strategy of placing a single hand beneath just one end (29,30,32,but see 33 for a modified version that yields better performance). This suggests that, while chimpanzees do seem to anticipate the arrival of the reward, they may only be capable of simulating one possible future state at a time, potentially even treating it as a certain future event, as young children have been argued to do (34).

To investigate both individual and collective reasoning, the current study was designed to explore the performance and dynamics of collectives (here, pairs) versus individuals in this logical reasoning task. We tested individuals, then pairs, then individuals once more, with an inverted U-shaped tube apparatus. This task offered a stripped down opportunity to test the three mechanisms of collective improvement described above by providing the minimal conditions required for each mechanism, without requiring sophisticated coordination by subjects or lengthy training (as would have been necessary to equally distribute the outcome, e.g., to bait the apparatus with a token that chimpanzees could exchange to both receive a reward) and while still approximating collective hunting. Given that in past studies individual chimpanzees typically did not express the logical solution of covering both ends of an inverted Y-tube, this context provided scope for collectives to achieve more optimal strategies or outcomes in one of several ways. If collective reasoning can emerge in chimpanzees in the absence of individual logical reasoning capacities, we predicted that, in collective contexts, one individual may cover one end of the tube while a different individual covers the other. This would allow them to collectively achieve the optimal strategy (and outcome) without the burden of fully reasoning about each possible future event falling to any one individual. Similarly, if chimpanzees fail to express the optimal logical behavior as individuals not because they lack individual logical reasoning capacities but because the motor demands of the task are too great (32), pairs of apes may be able to overcome this problem by sharing access to the apparatus and collectively covering both ends. This pattern of results would functionally mimic the proposed mechanism of collective hunting in chimpanzees - each individual need only represent one possible future event for all possible events to be collectively represented and prepared for. Alternatively, it is possible that some chimpanzees do possess modal reasoning abilities and that the presence of such individuals in collectives is sufficient to allow collectives to outperform the majority of (non-reasoning) individuals, as may result from collective scaling of individual reasoning. Finally, if some chimpanzees do possess these individual reasoning capacities, we may see an effect of collective facilitation: that is, by inducing competition (or cooperation) (33,e.g., 35), collective contexts might heighten chimpanzees’ motivation and promote more rational strategies, even at the individual level. Beyond investigating these multiple mechanisms of collective improvement, we further predicted that higher intrapair tolerance and more similar dominance ranks would facilitate shared use of the apparatus (e.g., 36,37,38) and hence promote collective reasoning, whereas lower tolerance and greater dominance differences would facilitate individual monopolization of the apparatus (39), lowering the likelihood of successful collective reasoning.

Method

Subjects

24 adult chimpanzees (*Pan troglodytes* - 11 females, 13 males; ages 19-40 yrs), living at Ngamba Island Chimpanzee Sanctuary in Lake Victoria, Uganda participated in this study. The chimpanzees live in a large, multi-male, multi-female social group of 54 individuals, housed in a 95-acre outdoor forest enclosure adjacent to a 24 m x 10 m indoor sleeping area. Subjects had access to food, water, and enrichment *ad libitum* and participated in research daily, on a volunteer basis. All procedures were approved by the local ethics committee of the sanctuary and the Animal Care and Use Committee of Johns Hopkins University, and complied with international standards (the Weatherall report, The use of nonhuman primates in research), institutional guidelines, and the legal requirements set by Uganda. See **Table S1** in supplementary material for a full list of subjects.

Procedure

Apparatus and Testing

Testing for this experiment was conducted in the sleeping rooms of the indoor chimpanzee structure adjacent to their outdoor enclosure. The apparatus consisted of an inverted U-shaped tube constructed of opaque PVC pipe, which was mounted against a 90 x 90 cm, 10mm-thick clear polycarbonate panel (**Figure 1**). The ends of the tube were 60 cm apart, such that two chimpanzees could sit next to each other at each end with a small space in between their bodies, but such that an individual chimpanzee could still reach both ends while sitting in the center. The polycarbonate panel was mounted to the bars of the chimpanzee sleeping structure, on the experimenter's side, such that the ends of the U-shaped tube were 16 cm from the ground. Chimpanzees sitting across from the experimenter could see, but not access the tube, apart from three 7x4 cm reach holes cut into the bars beneath the polycarbonate panel (**Figure 1**), which provided subjects with access to the two ends of the tube. The U-shaped tube had a small opening at the top of the U-bend, inside of which was affixed a rotating plate, which could be turned to obstruct one or the other side of the tube. Thus, the experimenter could drop pieces of food into an opening at the apex of the tube and covertly control the side of the tube from which the food would fall. Subjects' line of sight allowed them to observe when the experimenter held the food over the entrance to the tube in preparation to drop it, but not the mechanism which controlled the side from which the reward would drop.

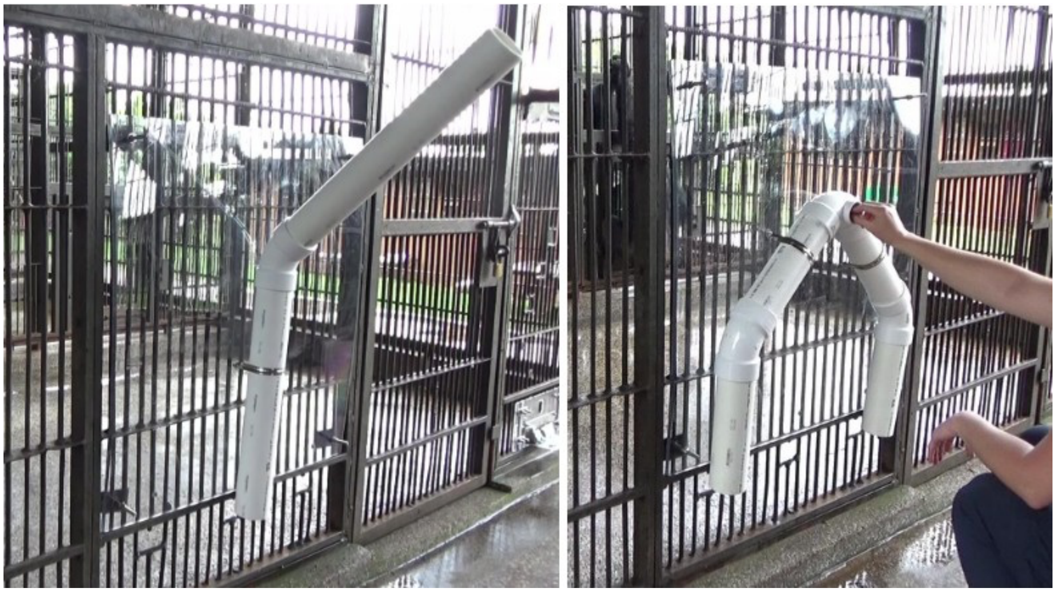


Figure 1: Image of "I" (left) and "U" (right) - shaped tubes used for the pre-training and experimental phases, respectively. Chimpanzees could hold their hands through holes in the mesh beneath tube ends in anticipation of falling food rewards. The right-hand image depicts the experimenter holding the food at the baiting location.

Pre-Training

Pre-Training I – Motor Demands

In order to confirm that subjects were capable of acquiring food from the apparatus, we first presented a straight "I"-shaped tube mounted in the same manner as the U-shaped tube (**Figure 1**). The purpose of this phase was to train subjects on the motor demands of the task – namely, placing their hand through the 'reach hole' beneath a tube end and covering the end of the tube sufficiently to catch the food. This phase also introduced subjects to the baiting cue: on each trial across all phases, the experimenter held the food above the tube and said "Ready?" in an excited tone of voice prior to dropping it. Critically, this phase did not train subjects on the U-tube apparatus, and provided no introduction to the possibility of multiple outcomes, nor did it train reaching into either of the reach holes occupied by the two ends of the U-tube in the main experiment.

Initially, the experimenter encouraged subjects to place their hand through the central reach hole to touch the end of the I-tube by indicating the location with her own hand, luring with additional pieces of food, using verbal encouragement, gesturing toward the end of the I-tube, and waiting to drop pieces of food until the subject had placed their hand in the correct location. This initial training ended once subjects were reliably willing to put their hands through the reach hole without luring or visual encouragement toward the correct location. Then, 18 evaluation trials were administered as follows: the experimenter held a piece of food above the opening to the I-tube, said "Ready?" and then dropped the food after five seconds, regardless of whether the subject had placed their hand through the reach hole in order to catch it. Initial training and evaluation trials occurred during a single session on each subject's first day of testing. Those subjects who successfully held their hand beneath the tube on 15/18 evaluation trials proceeded to the main experiment. Of the 24 subjects initially included in the sample, three subjects (all males) were omitted at this stage because they could not comfortably fit their hands into the reach hole to catch the food from the I-tube. Thus, the final sample for the main experiment consisted of 21 chimpanzees.

Pre-Training II – Observation of Outcomes

In order to provide subjects an opportunity to observe the possible outcomes of the U-tube, without directly interacting with it, subjects were presented with the U-tube, held in the experimenter's hand at a distance from the bars of the testing room, rather than mounted (as in 29). The experimenter recruited the subject's attention by calling their name and waving a piece of food, then dropped the food into the U-tube and allowed it to come out of one of the two ends to fall to the floor at her feet, out of reach from the subject. This process was repeated five more times, for a total of six trials, three times per side, in a counterbalanced order. The subject then received a large, unrelated reward to mitigate any frustration at not receiving the observed pieces of food, and the session was concluded. This session occurred on a separate day from both pre-training I and from the first individual phase of the main experiment.

Individual Pre-Pairs Phase

The purpose of the pre-pairs individual phase was to measure subjects' individual rates of deploying the rational strategy (covering both sides of the tube), and of successfully obtaining the food reward, prior to any intervention by social dynamics. This phase was a functional replication of the foundational study by Redshaw and Suddendorf (29), albeit with the larger, bars-mounted apparatus intended for subsequent pairs use. Subjects were admitted into the testing room, where the U-tube was mounted to the bars (**Figure 1**), and were tested in isolation. Subjects were administered 18 trials in a single session as follows: after subjects initially approached within one arm's length of the apparatus, the experimenter held a piece of food above the opening to the U-tube and said "Ready?". After five seconds, regardless of whether the subject had placed their hand or hands through the reach holes to cover either or both ends of the U-tube, the experimenter dropped the piece of food. If the subject caught the piece of food, she was able to eat it. If the subject did not catch the food (e.g., it fell through an uncovered side of the tube), the experimenter retrieved it and placed it back in the pail of food rewards. The experimenter then immediately began another trial by holding the next piece of food over the tube and repeating the baiting cue. On a separate day, all subjects then proceeded to the pairs phase of the experiment, regardless of their performance in this phase.

Pairs Phase

The purpose of the pairs phase was to evaluate whether pairs of chimpanzees, treated as a collective, were more or less likely to collectively cover both open ends of the U-tube, or more or less likely to successfully obtain the food reward, than individuals.

Pairs Selection

To assist in the selection process, the head caregiver at Ngamba Sanctuary completed a survey in which he rated the tolerance (0-3) and dominance differential (0-3) between each of the possible 210 pairings of the 21 chimpanzees participating in the main experiment. Pairs were then selected according to three criteria:

1. The head caregiver agreed that the pair could be enclosed into a room together and compete for food rewards without concern for safety or welfare.
2. Except one pair of subjects who could only be paired with each other (because they were housed separately from the rest of the group), each subject participated in at least two and no more than seven pairs. The number of pairs was balanced as much as possible between individuals, but ranged in order to prioritize a balance of demographic factors (below). The number of subjects participating in each number of pairs ($N_{pairs}(\#)$) was as follows:

$N_{pairs}(1)$= 2 individuals; $N_{pairs}(2)$= 2; $N_{pairs}(3)$= 4; $N_{pairs}(4)$= 6; $N_{pairs}(5)$= 5; $N_{pairs}(6)$= 1; $N_{pairs}(7)$= 1

3. Pairs were balanced across distribution of tolerance rating, dominance differential, and sex (M/M, M/F, and F/F), as much as possible within the limits of the sample.

The 21 subjects were initially mixed into 40 pairs. Two pairs were omitted from the experiment as a matter of welfare; they engaged in aggressive behavior upon being admitted into the testing room together and were thus excused before their paired trials commenced. Therefore, the final sample for the experiment included 38 pairs. See **Table S2** in supplementary material for a breakdown of the distribution of tolerance ratings, dominance differential ratings, and sex distribution for these pairings.

Pairs Procedure

Subjects were admitted together into the testing room, where the U-tube was mounted to the bars. At the start of the paired U-tube task, two caregivers were stationed along the right-hand side of the room such that subjects could approach them to intermittently receive juice rewards through the bars. Initially, caregivers called to the subjects and offered juice rewards to entice each of them to come to separate locations on the right-hand side of the room, equidistant from the apparatus. The experimenter then cued the subjects and the caregivers by saying “Ready?” and holding a reward over the opening to the apparatus, at which point the caregivers removed the juice, turned their backs toward the subjects and ceased interacting, and subjects had the opportunity to approach the tube to prepare to catch the food. The experimenter waited fifteen seconds from the time of the baiting cue to drop the food (to allow subjects time to walk across the room to approach the apparatus), and then dropped the food (and retrieved it if it had not been caught by one of the two subjects), as in the individual phase. The experimenter then cued the caregivers by saying “Okay!” and the caregivers turned around to re-engage with the subjects and offer juice. Once both subjects returned to the right-hand side of the room, equidistant from the apparatus, the experimenter initiated a new trial. This process was repeated for a total of 18 trials spread across two sessions occurring on separate days (1-10 days elapsed between these two sessions for any given pair); pairs received 9 trials per session. Subjects could participate in up to two different pairs per day.

Individual Post-Pairs Phase

The purpose of the individual post-pairs phase was to evaluate whether multiple exposures to the U-tube, including exposures in a social setting, had any impact on individuals' rate of covering both sides of the apparatus or rate of successfully obtaining the food. The procedure was identical to the procedure of the individual pre-pairs phase, again involving 18 trials administered in a single session. Subjects did not participate in the post-pairs phase until they completed all of their assigned pairings in the pairs phase.

Caregiver Survey

Finally, to independently characterize each pairs' social relationship, we collected information about tolerance and dominance differential within pairs using a survey administered to the caregivers at Ngamba Sanctuary. The survey consisted of twelve multiple-choice questions: ten hypothetical situations relating to intrapair tolerance, spanning situations such as food sharing and food stealing, comfort-seeking, co-sleeping, and grooming; one summary tolerance question asking caregivers to report the overall nature of the relationship between the pair (e.g., friendly, neutral, tense); and one question asking caregivers to select the dominance differential and direction between the pair (e.g., A is highly/ moderately/ a little/ equally dominant to B, or vice versa). See supplementary material for a reproduction of the full survey. Each pair was independently rated by two caregivers, neither of whom had observed that specific pair engage in the experiment. Caregivers at Ngamba Sanctuary had an average of 11 years (range: 1-21 years) of experience with these chimpanzees, and are in frequent contact and observation with the entire chimpanzee group on a daily basis. They have substantial opportunity to observe their group and dyadic dynamics in various situations, and were instructed to answer these questions based on their impressions of the chimpanzees' interactions as a whole, rather than their expectations about the experiment in particular.

Statistics and Analysis

We live-recorded the identity, order of approach, and covering behaviors of each subject across the three experimental phases. *Approach* was coded as any instance in which a subject moved within one arm's reach of the apparatus during the period in between the experimenter's baiting cue and the time at which she dropped the food reward. Covering behavior was recorded categorically as *left*, *right*, *both*, or *neither* with respect to the tube ends from the perspective of the experimenter, and *cover* was defined as any instance in which the subject extended their hand through the reach hole or the bars beneath the tube such that their fingertips could touch either the far side or the inside of the tube end. *Cover* was recorded at the time at which the food reward was dropped, regardless of any covering behavior prior to that point. An additional coder, blind to the purpose and predictions of this study, recorded the approach and cover behaviors for 20% of trials in each of the three experimental phases. We used Cohen's kappa to assess interobserver reliability and found excellent agreement with the live coding (approach: $k = 0.93$, cover: $k = 0.95$).

Regarding the caregiver surveys of intrapair tolerance and dominance, each of the first ten questions in the survey returned a tolerance score between 0-3 (see supplementary material for a full reproduction of the survey, including the tolerance scores for each response option for each question). These tolerance scores were averaged across the ten questions and then between caregivers to return a total mean tolerance score for each pair. The dominance differential question returned a dominant and subordinate designation and a dominance differential between 0 and 3, which was likewise averaged between caregivers for an overall dominance differential and direction. In cases where the caregivers disagreed as to the direction of the dominance within a pair, the dominance differential was set to 0 and the dominant and subordinate designations were set to "NA." Interrater reliability between caregivers was assessed using intraclass correlation coefficients with two-way, absolute-agreement, random-effects models with 95% confidence intervals. Agreement between caregivers was moderate for tolerance ($ICC = 0.65$) and good for dominance differential ($ICC = 0.71$). The summary tolerance question was used as a measure of internal validity: each caregiver's mean tolerance score across the ten hypothetical questions was compared to the tolerance score returned by the summary question, to ensure that the hypothetical situations posed by the survey accurately captured the tolerant or intolerant nature of the relationship in the minds of the caregivers. The correlation between the survey mean tolerance scores and the survey summary scores was moderate ($r = 0.66$).

For models at the collective level, we fitted generalized linear mixed models with a binomial error structure and a logit link function (GLMM, 40) using the *glmer* function from the lmerMultiMember package (41), in order to include the identities of both members of pairs as random effects. For models at the individual level, we used the *glmer* function of the lme4 package (42). P-values for individual fixed effects and interaction terms were calculated using likelihood ratio tests of the fitted model with and without the relevant term (R function *drop1*, 43). All analysis was conducted in RStudio (version 4.2.1), and all data visualizations were created using the R package ggplot2 (44).

For access to the raw data and the R code pertaining to this analysis, and for a video of each of the above coded behaviors, please see https://osf.io/dzv8c/?view_only=ca96b4d75cd447d887bee8013a46ccd5

Results

Overall Rate of Success across Phases

Our initial research prediction – relating to the reward-maximizing outcome of individual versus collective action – was that the overall rate of success would vary between individual and paired phases. We defined success as cases when one of the two subjects, or the only subject, covered the side from which the food ultimately fell, regardless of what other tube ends they did or did not cover. To compare success across all three phases, we conducted a binomial GLMM, with pair ID and subject ID(s) as random effects. This model revealed a nearly significant effect of phase on overall success ($\chi2_{(2)} = 5.64$, p=0.06), but the only significant difference between phases was an increase between pre- and post-pairs ($Z_{(1)}$=-2.34, p=0.041). The remaining differences between phases were not significant: individual-pre vs pairs: ($Z_{(1)}$=-0.59, p=0.809); individual-post vs pairs: ($Z_{(1)}$=0.21, p=0.973). See **Figure S1** in supplementary material for more details. We therefore focused the remainder of our analyses on the strategies subjects used to solve the task.

Sharing

We hypothesized that the most likely mechanism of emergent collective reasoning, if it occurred, would be shared use of the apparatus. If, as in previous experiments, subjects did not systematically use two hands to cover the apparatus independently, a pair could produce the same solution by having each subject cover one end of the tube such that both ends were still covered. This would constitute emergent collective reasoning. Each subject need only simulate one possible future event, and coordinate minimally such that her simulation is distinct from that of her partner, for the collective to represent and prepare for all possible events. Interestingly, we observed this apparatus-sharing behavior on only 12 trials (out of 684 total pairs trials), spread across 6 pairs (n_{trials}=1,1,1,2,2, and 5 trials involving sharing, respectively).

Unexpected Two-Hand Use at the Individual Level

Strikingly, in contrast with past findings, we found that many of our subjects were able to solve the task on their own, by using the apparently logical strategy of placing one hand under each end of the U-shaped tube (hereafter: two-hand use). 11 of the 21 subjects did so on at least one trial in the individual pre-pairs phase of the experiment, and 4 individuals did so on at least 50% of their trials in that phase. One of these subjects, remarkably, exhibited two-hand use on all 18 trials in the pre-pairs phase. Two-handed performance persisted in the pairs phase of the experiment, where, in addition to sustained or improved two-handed performance from those who had done so in the pre-pairs phase, 3 subjects began to use two hands in the pairs phase who had not done so in the pre-pairs individual phase. In the individual post-pairs phase, 8 subjects used two hands on at least 50% of trials. Subjects' use of two hands across the three phases is summarized in **Figure S2** (see also **Table S3**).

Collective Rate of Two-Hand Use Across Phases

In the pairs phase, although 'sharing' rates were surprisingly low, individual deployment of the two-hand strategy remained high. Indeed, out of our full sample of 684 pairs trials, the individual who interacted with the apparatus exhibited two-hand use in 299 trials. That many individuals demonstrated this apparently logical two-hand use at the individual level allowed us to investigate the two other mechanisms of collective improvement. The first of these mechanisms is collective scaling of individual reasoning, which predicts that, if reasoning is present at the individual level but varies across individuals, the likelihood that the collective will possess reasoning individuals and deploy the reasoning-guided solution will scale up linearly with group size. This strategy thus predicts greater two-hand use in collectives than in individuals. In order to examine this possible mechanism, we ran a binomial GLMM with the collective rate of two-hand use (i.e. two-hand use occurred as long as either (or the only) individual used two hands, regardless of the other individual's engagement) compared across the three phases, with subject ID(s) and pair ID included as random effects, and a random slope of phase nested within subject ID(s). The mean (± SE) rates of two-hand use were 0.21 ± 0.02 in the individual pre-pairs phase, 0.44 ± 0.02 in the pairs phase, and 0.34 ± 0.02 in the individual post-pairs phase. There was a nearly significant effect of phase on collective rate of two hand use ($\chi2_{(2)} = 0.8$, p=0.669), , but the differences between the pairs and individual pre- and individual post-pairs phases, respectively, were not significant (*individual-pre vs pairs*: $Z_{(1)}$=-0.92, p=0.599; *pairs vs individual-post*: $Z_{(1)}$=0.54, p=0.839). Rather, we observed a significant increase from the individual pre- to individual post-pairs phase ($Z_{(1)}$=0.08, p=0.996). This pattern is not consistent with collective scaling of individual reasoning, which predicted greater two-hand use in the pairs phase than in solo phases including in the post-pairs phase, where the number of individuals present is reduced back to one.

Tolerance and Dominance Effects

We did not find evidence of either emergent collective reasoning, as in sharing, or of collective scaling of individual reasoning. However, these mechanisms might have been constrained by intrapair factors like tolerance and dominance differential. To investigate such effects, we ran a binomial GLMM on data from the pairs phase, with two-hand use as the dependent measure, an interaction term of survey mean tolerance and survey dominance differential as fixed effects, and subject IDs and pair ID as random effects. This model identified a significant interaction between pairs' tolerance and their dominance differential on collective rate of two-hand use in the pairs phase ($\chi2_{(1)} = 6.83$, p=0.009). More tolerant pairs and pairs with a lower dominance differential tended to exhibit higher rates of individual deployment of the rational strategy. The positive effect of tolerance had a greater impact on the rate of two-hand use when the dominance differential was high, suggesting that in more tolerant dyads, even less dominant 'reasoning' individuals may still have had access to the apparatus (see **Figure 2**). This suggests that an otherwise simple mechanism of scaling up may be moderated by the social dynamic of the group, i.e. individuals may not be able to fully contribute their capacity to reason about multiple future events if they are inhibited by the presence of their partner.

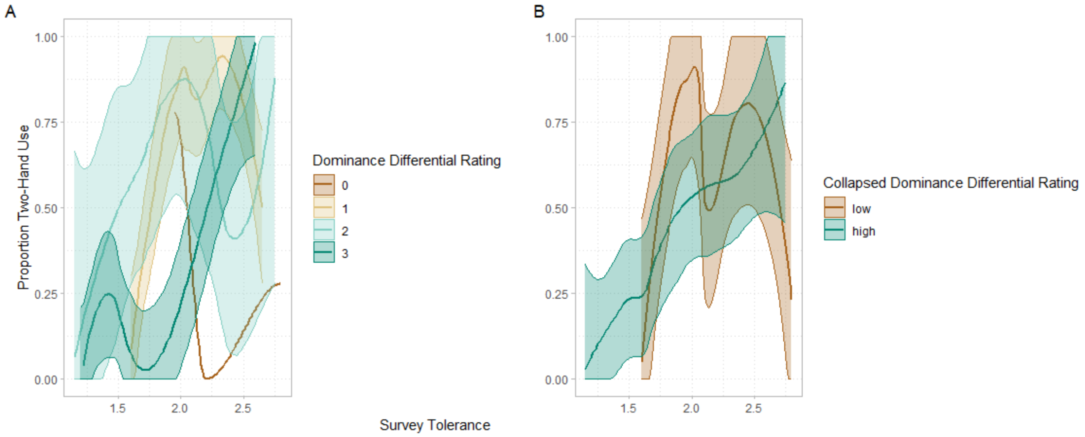


Figure 2: The effect of the interaction between pairs' tolerance and dominance differential scores on mean collective rate of two-hand use. ***A****: all four possible levels of dominance differential are depicted* ***B****: dominance differential is collapsed into 'low' (0-1), and 'high' (2-3), for the purpose of visualization.*

Collective Facilitation of Individual Reasoning

Finally, we investigated collective facilitation of individual reasoning, by examining the effect of individual versus collective context on individuals' deployment of the putatively logical strategy. Here, each trial in the pairs phase was evaluated independently, one with each member of the pair designated as the focal subject or partner subject, and then vice versa. To accurately assess changes in performance for the pairs phase, we filtered the data to include only trials in which the focal chimpanzee had completely unfettered access to the apparatus (i.e., the subject approached within one arm's length of the apparatus, and their partner either did not approach the apparatus at all, or did approach but did not cover either end of the tube, to account for the fact that dominant partners could otherwise prevent access to the apparatus). This filter also eliminated any instance of non-independent replication of the same trial; if one subject had unfettered access to the apparatus, then the other subject, by definition, did not, and that second subject's trial was therefore dropped when the first individual was viewed as the focal subject. A binomial GLMM found a significant effect of phase on rate of individual use of two hands within full-access trials ($\chi 2_{(2)} = 30.83$, $p < .001$; see **Figure 3**). Chimpanzees showed significantly greater rates of two-hand use in the pairs phase relative to the individual pre-pairs phase ($Z_{(1)}$=-2.7, p=0.017), and also in the post-pairs phase relative to the pre-pairs phase ($Z_{(1)}$=-5.15, $p < .001$). The difference between the pairs and post-pairs phases was not significant ($Z_{(1)}$=0.55, p=0.836). Thus, the collective context seems to have spurred an increase in individuals' deployment of the logical strategy, which persisted into the individual post-pairs phase. To assess the robustness and emergence of this effect, we ran a further model including, for each subject, all 18 pre-pairs and 18 post-pairs trials and only the first 18 pairs trials during which each subject had unfettered access to the apparatus. This analysis fully replicated the results of the full-dataset model, confirming a saltatory effect of the collective context on individual two-hand use, which persisted into the post-pairs phase (see supplementary materials for a breakdown of these results). This novel effect of collective facilitation likely resulted from competition, which has been shown to heighten chimpanzees' motivation and performance in cognitive tasks (e.g., 45,46,47).

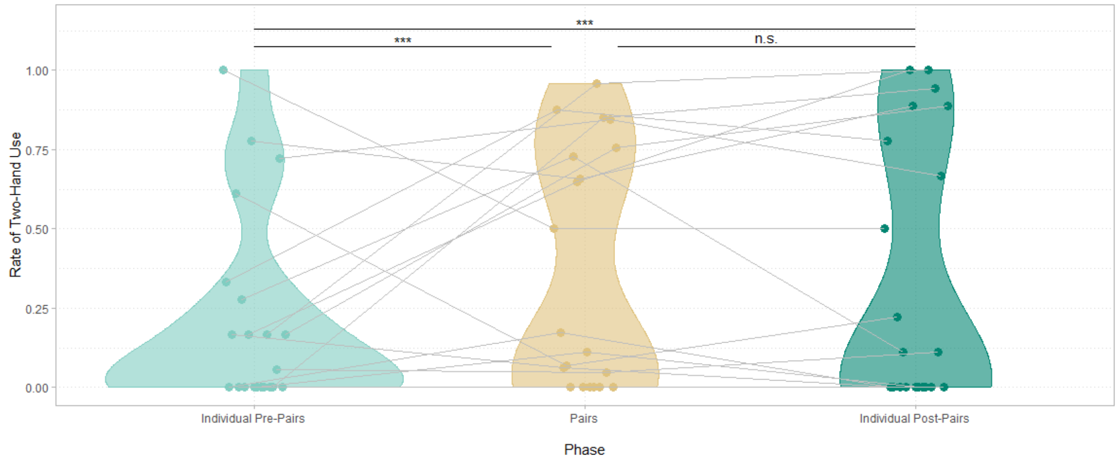


*Figure 3: Mean individual rate of use of two hands across phases during trials in which the subject had full access to the apparatus. Points represent individual rates, lines represent individual change between phases. *** - $p<0.001$, n.s. - $p>0.05$*

Learning

Effect of Repeated Exposures to Apparatus on Use of Two Hands

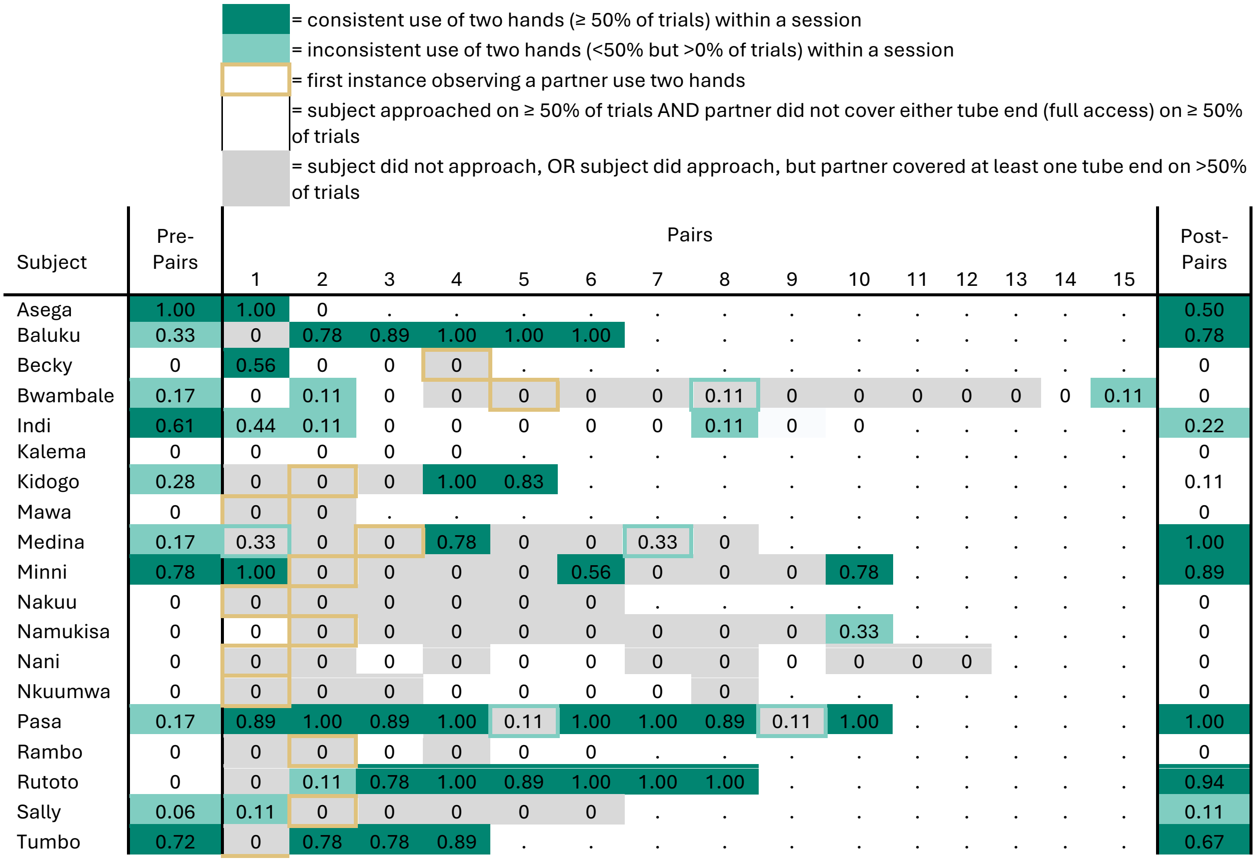
One alternative is that the increase in individual use of two hands was not due to the collective context at all; rather, subjects simply learned the strategy over time, with experience. In order to investigate any effects of experiential learning on the two-hand behavior, we fitted a binomial GLMM with overall trial number (continuous across all phases for each individual) and session number (each session of 18 trials in the individual phases or 9 trials in the pairs phase) as fixed effects and the identity of the focal subject and the pair ID (as above) as random effects. Critically, this model found no significant effect of overall trial number across all phases on individual use of two hands ($\chi 2_{(1)} = 0.75$, p=0.385). The effect of session number was also not significant ($\chi 2_{(1)} = 0.5$, p=0.481). Of note, neither of the two subjects who experienced the most trials across the pairs phase (Nani and Bwambale, 108 and 126 trials, respectively), showed proficient use of two hands, which further supports the suggestion that the overall increase in two-hand use from the pre-pairs phase to the post-pairs phase was not a function of simple exposure. We therefore find no evidence for a general effect of learning. The significant effect of phase in the above full-access model is therefore most parsimoniously explained by differences between the circumstances of the phases, and not by general learning.

Social Learning

Another possibility is that the collective context mattered but only insofar as it provided opportunities for subjects to socially learn, by observing their partner's two-handed use.

However, we find no evidence that individuals systematically learned to employ the two-handed strategy from first observing others doing so. The majority of individuals (11/14) who used the two-handed strategy in any trial produced their first instance of two-hand use in their first individual phase, before any exposure to paired trials. Of those who did not first produce the behavior in their individual pre-pairs phase (3/14 - see Becky, Namukisa, and Rutoto, below), only one (Namukisa) had the opportunity to observe the behavior from another individual before producing it herself; the other two innovated the behavior without having observed it from a partner. Similarly, of those who produced the behavior consistently (>50% of trials within a single session), only 2 of the 11 (see Kidogo and Medina, below) observed a partner use the two-hand strategy prior to their own first consistent sessions. Notably in the case of both of these subjects (Kidogo and Medina), their first consistent use of two hands emerged in the first pairs session during which they had full access to the apparatus on at least 50% of the trials, and it is therefore not possible to parse whether their consistency was a consequence of having seen a partner use the two-handed strategy, or simply a matter of access related to dominance within their pairs. See **Table 1** for more details.

Table 1. Progression of two-hand use per subject per session, depicting the percentage of trials on which the subject used two hands in each session. Note that pre- and post-pairs sessions consisted of 18 trials each, and the subject participated alone, while paired sessions consisted of 9 trials each, and each pair completed two sessions. Paired sessions with the same partner were not necessarily consecutive for either member of any given pair. Individuals participated in between two and fifteen paired sessions (depicted as cells with numerical values).



■ = consistent use of two hands (≥ 50% of trials) within a session
▨ = inconsistent use of two hands (<50% but >0% of trials) within a session
□ = first instance observing a partner use two hands
▢ = subject approached on ≥ 50% of trials AND partner did not cover either tube end (full access) on ≥ 50% of trials
▒ = subject did not approach, OR subject did approach, but partner covered at least one tube end on >50% of trials

Subject	Pre-Pairs	Pairs 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Post-Pairs
Asega	1.00	1.00	0	.	.	.	.	.	.	.	.	.	.	.	.	.	0.50
Baluku	0.33	0	0.78	0.89	1.00	1.00	1.00	.	.	.	.	.	.	.	.	.	0.78
Becky	0	0.56	0	0	0	.	.	.	.	.	.	.	.	.	.	.	0
Bwambale	0.17	0	0.11	0	0	0	0	0	0.11	0	0	0	0	0	0	0.11	0
Indi	0.61	0.44	0.11	0	0	0	0	0	0.11	0	0	.	.	.	.	.	0.22
Kalema	0	0	0	0	0	.	.	.	.	.	.	.	.	.	.	.	0
Kidogo	0.28	0	0	0	1.00	0.83	.	.	.	.	.	.	.	.	.	.	0.11
Mawa	0	0	0	.	.	.	.	.	.	.	.	.	.	.	.	.	0
Medina	0.17	0.33	0	0	0.78	0	0	0.33	0	.	.	.	.	.	.	.	1.00
Minni	0.78	1.00	0	0	0	0	0.56	0	0	0	0.78	.	.	.	.	.	0.89
Nakuu	0	0	0	0	0	0	0	.	.	.	.	.	.	.	.	.	0
Namukisa	0	0	0	0	0	0	0	0	0	0	0.33	.	.	.	.	.	0
Nani	0	0	0	0	0	0	0	0	0	0	0	0	0	.	.	.	0
Nkuumwa	0	0	0	0	0	0	0	0	0	.	.	.	.	.	.	.	0
Pasa	0.17	0.89	1.00	0.89	1.00	0.11	1.00	1.00	0.89	0.11	1.00	.	.	.	.	.	1.00
Rambo	0	0	0	0	0	0	0	.	.	.	.	.	.	.	.	.	0
Rutoto	0	0	0.11	0.78	1.00	0.89	1.00	1.00	1.00	.	.	.	.	.	.	.	0.94
Sally	0.06	0.11	0	0	0	0	0	.	.	.	.	.	.	.	.	.	0.11
Tumbo	0.72	0	0.78	0.78	0.89	.	.	.	.	.	.	.	.	.	.	.	0.67

Umugenzi	0	0	0	0	0	0	0	0	0	.	.	.	.	.	.	.	0
Yoyo	0.17	0.56	0.89	0.89	0.89	0	0	0	0	0.78	0.78	.	.	.	.	.	0.89

Discussion

In this study, we set out to investigate individual and collective performance in a logical reasoning task in which a food reward could emerge from either of two ends of an apparatus. Subjects could prepare for these two mutually exclusive possibilities through individual reasoning, with one subject accounting for both possibilities and placing their hands beneath both ends to ensure their acquisition of the reward. Alternatively, we hypothesized that - even in the absence of individual reasoning - collectives might be able to harness an emergent strategy we termed emergent collective reasoning: if two subjects each anticipate and prepare (with a single hand each) for a different possibility, the coordination and combination of these one-handed preparatory actions could allow the collective, even incidentally, to produce the optimal cover-both strategy. However, we found little evidence for collective reasoning in this task: pairs were not more successful than individuals at collectively covering the end from which the reward would emerge, and we observed only exceedingly rare instances (12/684 trials) of pairs exhibiting a shared strategy of each individual covering one end of the apparatus. We also found no evidence for a simple scaling up of reasoning in collectives relative to individuals, a sort of collective reasoning by proxy which could have occurred if some individuals reason and collectives simply have a greater likelihood of containing one or more of those individuals. Strikingly, however, we found that individual chimpanzees deployed the logical strategy of covering both ends on their own, and that the collective context facilitated more reliable deployment of this behavior.

In contrast to several past studies with apes (29,30,32) and one with monkeys (48), we present the first evidence that nonhuman primates readily and consistently produce the optimal rational behavior of reaching with two hands to catch food falling from an inverted Y (or U)-shaped tube. That is, they prepared for either of two possible mutually exclusive future events. More than half of our subjects spontaneously displayed this behavior from their first 18-trial session, some at extremely high rates and/or from the first trial. Notably, two-handed use was observed very occasionally in a very small number of individuals in past studies but never consistently, fueling the interpretation that the behavior was morphologically possible but that the individuals displaying it did not grasp a logical understanding of possibilities (49). In stark contrast to these findings, many of our subjects consistently deployed the logical two-handed strategy at high rates from the time they first used it onward.

Our findings corroborate but also meaningfully extend work by Engelmann and colleagues (33,47), which - in contrast to the studies referenced above - has provided evidence that apes may be capable of representing mutually-exclusive possibilities and logical “OR” relations (the reward might emerge from the left OR the right tube). In one study (33), chimpanzees could manually stabilize either of two platforms containing food. In the Y-shaped tube condition, an inverted Y-tube was located above the two platforms, such that when a rock was dropped down the tube, it could land on either platform and, if the platform was not stabilized, the food would be lost. In the single tube condition, an I-shaped tube was located over only one of the platforms. Chimpanzees stabilized both platforms more often in the Y-shaped tube condition, where the rock could emerge from either tube end, than the single tube condition. While this finding suggests that the chimpanzees may have anticipated and prepared for mutually-exclusive possibilities in the Y-shaped tube condition, the actions taken by the chimpanzees (stabilizing trays to protect food) are less precisely targeted to a very specific possible future state (stone or food emerging from one tube end) than are the actions displayed in our task (placing a hand beneath the tube to catch the falling item). These actions also required training in the task by Engelmann et al. (33) but occurred spontaneously in our own. Finally, critics have argued that chimpanzees in the Engelmann et al. (33) study, desiring both visible food rewards, could solve the task by representing an AND relation (I want this food AND that food) rather than an OR relation (49). The classic task used in our study is not open to this AND alternative and therefore our data substantially strengthen the argument that chimpanzees are capable of representing mutually-exclusive possibilities and logical OR relations. At least one alternative does exist, however. In both our study and that of Engelmann et al. (33) (as well as most existing studies with human children), if tube ends are simply an attractive stimulus because of their ability to produce food, subjects may ignore the baiting and potential path of the food entirely, and simply focus on reaching for the rewarding element of the apparatus. This is an issue for future research, explored in two follow-up studies presented in Warren et al. (*in prep*).

The performance of consistent two-handers in this experiment resembles the behavior produced by four-year-old children in the same task (29), which was taken as evidence that children can reason logically about mutually exclusive possibilities. By this token, we have grounds to question the supposition that process-reasoning (PP-rationality) is unique to humans. One outstanding question is why our task elicited such frequent and consistent use of the logical two-handed strategy, even in the initial individual pre-pairs phase, compared to other studies using this paradigm. Although 12 of our 21 subjects had some relevant experience participating in the study by Engelmann et al. (33), only about half of these subjects (seven individuals) ever produced the two-handed behavior in our study. Critically, all but one of the nine fully naive subjects innovated the two-handed strategy within our task, and the majority were consistent two-handers (e.g., Kidogo, Minni, Rutoto, Tumbo, Yoyo), demonstrating that experience in previous tasks cannot account for our chimpanzees' success. One possibility is that our adaptations to the apparatus to make it suitable for paired use helped subjects overcome some of the other motor and cognitive demands of this task. Indeed, in a control study, Lambert & Osvath (32) found that chimpanzees failed to produce a two-handed response even in a task that didn't require anticipating mutually-exclusive possibilities (e.g., when two rewards were concurrently dropped down separate tubes), suggesting that the configuration of the setup may be relevant to chimpanzees' production of a two-handed behavior. Our apparatus differed in several key ways. First, our apparatus was mounted to the bars in front of the subjects, permitting them to support their weight by grasping the ends of the tubes in front of them, as they might normally do with the bars, while still using their hands to catch the food rewards. Second, the ends of the apparatus were wider-spaced and lower to the ground than other versions of this task, to facilitate the possibility that two subjects might sit at the apparatus together, but this may have provided a more physically comfortable arm and shoulder position for the subjects to use two hands simultaneously. Finally, the U-shaped design of this tube may have helped subjects overcome any potential gravity bias (e.g., 50), given that there was no part of the tube which led straight downward from the baiting location, and the two possible paths for the food reward may therefore have been more salient to the subjects. It is also possible that this population of chimpanzees was primed to succeed at this task due to factors such as their history of participation in research or their highly enriching, free-ranging outdoor enclosure. Although we can demonstrate from our pattern of results that specific experience on a similar task (33) was not a requirement for chimpanzees to express the correct behavior, it is still possible that these subjects are generally more practiced at analytical tasks. Future research with additional populations and variants of the task will help to shed light on these open questions. Regardless of the driver of success in this population, though, the novel behavior resembling logical reasoning that we have documented here also points to exciting horizons for future investigations into PP-rationality and the mechanisms of modal logic in non-human primates.

Excitingly, we also found evidence that the collective context fueled an increase in logical two-handed use by individuals. Overall, the rate of two-hand use was highest in the pairs phase, and higher in the post-pairs than in the pre-pairs phase. This pattern suggests that there may be some social effect of pairs that induced more individuals to produce or improve the correct strategy, and that this strategy, once acquired or stabilized, persists in individual performance. Yet, we found little evidence that collectives succeeded by sharing the tube and covering different ends. Individual's performance in the pairs phase may have been shaped by the social dynamics of the individual pairs, in that their relationship with their partner may have facilitated or hindered their access to the apparatus; competent two-handers may not be able to freely apply the two-handed strategy if they were paired with a dominant individual who was not tolerant of their approach. The complexity of these dynamics is particularly evident in the significant interaction between tolerance and dominance, where the effect of tolerance is relatively flat in pairs with a low dominance differential, but more relevant to the rate of two-hand use when there is a high dominance differential. These effects of tolerance and dominance support the premise that, in chimpanzees, competition within a collective is a relevant factor in the optimality of the outcomes they are able to achieve, especially within a task where the benefit of the outcome is not shared collectively.

The individual improvement in use of two hands from pre-pairs into pairs and then largely sustained into the post-pairs phase suggests a form of collective facilitation inducing individuals to use the two-handed strategy (an effect of social pressure, social competition, or social learning). Notably, this saltatory effect of the collective context was present within the first 18 trials. We found no evidence to support the suggestion that subjects innovated or improved their use of two hands purely as a function of repeated exposures to the apparatus. There was no significant increase in the use of two hands across session or trial number, and the majority of subjects first innovated the behavior in their very first session, some from their first trial. The pattern by which subjects acquired and stabilized their use of two hands also does not suggest social learning; all but one subject had already produced the two-handed behavior at least once prior to witnessing a partner doing so, and only two saw a partner use the behavior prior to producing it consistently themselves. Additionally, several subjects who never produced or became consistent at using the two-handed behavior did see a partner do so on multiple occasions, suggesting that social learning was not the primary mechanism of acquisition or stabilization of this behavior. Participation in the collective phase of the task did, however, appear to drive an increase in individual use and consistency of the two-hand strategy, which suggests that social pressure or competition may have driven certain subjects to develop or improve their two-handed behavior. The presence of another subject could drive individual improvement in two (non-exclusive) ways: first, the presence of a conspecific may raise arousal or increase motivation to monopolize food resources, which could, in turn, sharpen focus and surmount passivity, resulting in more active and thoughtful engagement with the apparatus. This interpretation aligns with several past studies that have documented competition’s motivating influence on chimpanzee cognitive performance (e.g., 45,46,47). Second, it is possible that the presence of a second individual could induce the active subject to conceive of multiple roles or actions, where they might represent how the other could act on the apparatus alongside them and then realize that they could take that action themselves instead. Numerically, the mean rate of two-hand use does decrease slightly from pairs to post-pairs, reflecting the behavior of those individuals who used the two-hand strategy in pairs only, suggesting some possible effect of competition specifically, but they were outnumbered by those who improved and sustained their improvement, and this difference is not significant. It is difficult to say, therefore, whether the effect of the collective context simply drove a sustainable form of learning that persisted into the individual post-pairs phase, or whether the active presence of a conspecific still further influenced certain subjects’ strategy. Regardless, competition and collective facilitation are both valid mechanisms of collective outcome-based rationality. If participation in a collective, where the individual’s behavior is inherently influenced by the choices and actions of others, causes the individual to perform better, the collective as a whole may still achieve more optimal outcomes or strategies than an individual alone.

In this study, we investigated three potential mechanisms of collective improvement in a logical reasoning task. We did not find evidence for emergent collective reasoning or collective scaling of individual reasoning. Future work should continue to investigate these hypothesized mechanisms of collective improvement, either using similar paradigms but varying key parameters of the task, such as equalizing rewards (e.g., by dropping a token that apes can trade to both receive a reward), altering the degree of coordination that is required, increasing the size of the collective, or by deploying entirely disparate paradigms with apes or other social species. Excitingly, we did find evidence that individual chimpanzees may possess modal logical reasoning capacities. We also identified a novel mechanism of collective improvement that we term collective facilitation, in which the collective context drives more robust deployment of (apparent) individual reasoning capacities, an effect that persisted even beyond the collective context. It is worth noting that we examined this effect in dyads – the smallest possible collective – in order to tightly monitor the effects of single versus multiple actors. There is, therefore, substantial opportunity to investigate whether and how the effect of social context found in this investigation scales up to larger collectives. Finally, we found evidence for effects of tolerance and dominance differential on the expression and suppression of the logical reasoning strategy. Each of these factors, in that they may impact individual capacity to act rationally, has the potential to affect collective outcomes, where the actions of an individual affect the choices available to others and potentially influence others' chance to benefit from the collective behavior. Although much of this analysis was exploratory in nature, and therefore limited in its ability to confirm mechanisms about chimpanzee collective behavior, it reveals a rich ground for future study, including further exploration into the roles of learning and of social dynamics on individual and group-level outcomes. We also note that, although our current paradigm did not ultimately afford the opportunity to model emergent collective reasoning via the distribution of the reasoning effort of multiple individuals, it remains an area of great potential in the interpretation of collective behavior in primates, in natural or simulated versions of behaviors such as collective foraging, group navigation, and cooperative territorial defense.

Acknowledgements

We are grateful to the wonderful staff at Ngamba Island Chimpanzee Sanctuary who made this work possible, especially Dr. Joshua Rukundo, Head Caregiver Innocent Ampeire, and caregivers Phillip Sekulya, Amos Okello, Boris Waiga, Stanley Kyama, Hillary Aligumisiriza, and Lydia Namayanja, who assisted with data collection and completed our intrapair tolerance and dominance survey. We also thank the Uganda National Council for Science and Technology and the Uganda Wildlife Authority for supporting our research. We thank Mareesa Islam and Georgette Djenadoum for assistance with behavioral coding of video recordings and reliability coding, respectively. This research was funded by a grant from the Templeton World Charity Foundation Inc. (TWCF-2021-20647) to D.B., R.P.M., and C.K. C.K. was also supported by the CIFAR Azrieli Global Scholars program.

References

1. Aristotle. Parts of Animals. Movement of Animals. Progression of Animals [Internet]. Cambridge, MA: Harvard University Press; 1937 [cited 2025 May 30]. (Loeb Classical Library). Available from: https://www.loebclassics.com/view/LCL323/1937/volume.xml

2. Kacelnik A. Meanings of rationality. In: Nudds M, Hurley S, editors. Rational animals. Oxford University Press; 2006. p. 87–106.

3. Kahneman D. Thinking, Fast and Slow. Macmillan; 2011.

4. Penn DC, Holyoak KJ, Povinelli DJ. Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. Behavioral and Brain Sciences [Internet]. 2008 Apr [cited 2025 Sep 17];31(2):109–30. Available from: https://www.cambridge.org/core/product/identifier/S0140525X08003543/type/journal_article

5. Krupenye C, Call J. Theory of mind in animals: Current and future directions. WIREs Cognitive Science [Internet]. 2019 [cited 2025 Sep 17];10(6):e1503. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/wcs.1503

6. O'Madagain C, Tomasello M. Shared intentionality, reason-giving and the evolution of human culture. Philosophical Transactions of the Royal Society B: Biological Sciences [Internet]. 2022 Jan [cited 2025 Sep 17];377(1843):20200320. Available from: https://royalsocietypublishing.org/doi/10.1098/rstb.2020.0320

7. Sasaki T, Pratt SC. Emergence of group rationality from irrational individuals. Behavioral Ecology [Internet]. 2011 Mar [cited 2025 May 30];22(2):276–81. Available from: https://doi.org/10.1093/beheco/arq198

8. Biro D, Sumpter DJT, Meade J, Guilford T. From Compromise to Leadership in Pigeon Homing. Current Biology [Internet]. 2006 Nov [cited 2025 May 30];16(21):2123–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0960982206021555

9. Sumpter DJT. Collective Animal Behavior [Internet]. Princeton University Press; 2010 [cited 2025 May 30]. Available from: https://www.degruyterbrill.com/document/doi/10.1515/9781400837106/html

10. Biro D, Sasaki T, Portugal SJ. Bringing a Time–Depth Perspective to Collective Animal Behaviour. Trends in Ecology & Evolution [Internet]. 2016 Jul [cited 2025 May 30];31(7):550–62. Available from: https://www.sciencedirect.com/science/article/pii/S0169534716300027

11. Mann RP. Collective decision making by rational individuals. Proceedings of the National Academy of Sciences [Internet]. 2018 Oct [cited 2025 May 30];115(44). Available from: https://pnas.org/doi/full/10.1073/pnas.1811964115

12. Mann RP. Collective decision-making by rational agents with differing preferences. Proceedings of the National Academy of Sciences [Internet]. 2020 May [cited 2025 May 30];117(19):10388–96. Available from: https://pnas.org/doi/full/10.1073/pnas.2000840117

13. Mann RP. Optimal use of simplified social information in sequential decision-making. Journal of The Royal Society Interface [Internet]. 2021 Jun [cited 2025 May

30];18(179):20210082. Available from: https://royalsocietypublishing.org/doi/full/10.1098/rsif.2021.0082

14. Camazine S, Deneubourg JL, Franks NR, Sneyd J, Theraulaz G, Bonabeau E. Self-Organization in Biological Systems [Internet]. Princeton University Press; 2003 [cited 2025 Sep 17]. (Princeton Studies in Complexity). Available from: https://www.torrossa.com/en/resources/an/5622276

15. Couzin ID, Krause J. Self-Organization and Collective Behavior in Vertebrates. In: Advances in the Study of Behavior [Internet]. Elsevier; 2003 [cited 2025 Sep 17]. p. 1–75. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0065345403010015

16. Sumpter Djt. The principles of collective animal behaviour. Philosophical Transactions of the Royal Society B: Biological Sciences [Internet]. 2005 Nov [cited 2025 Sep 17];361(1465):5–22. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2005.1733

17. Völter CJ, Call J. Younger apes and human children plan their moves in a maze task. Cognition [Internet]. 2014 Feb [cited 2025 May 30];130(2):186–203. Available from: https://www.sciencedirect.com/science/article/pii/S0010027713002096

18. Wittig RM, Crockford C, Langergraber KE, Zuberbühler K. Triadic social interactions operate across time: A field experiment with wild chimpanzees. Proceedings of the Royal Society B: Biological Sciences [Internet]. 2014 Mar [cited 2025 May 30];281(1779):20133155. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2013.3155

19. Krupenye C, Kano F, Hirata S, Call J, Tomasello M. Great apes anticipate that other individuals will act according to false beliefs. Science [Internet]. 2016 Oct [cited 2025 May 30];354(6308):110–4. Available from: https://www.science.org/doi/10.1126/science.aaf8110

20. Schmelz M, Grueneisen S, Kabalak A, Jost J, Tomasello M. Chimpanzees return favors at a personal cost. Proceedings of the National Academy of Sciences [Internet]. 2017 Jul [cited 2025 May 30];114(28):7462–7. Available from: https://pnas.org/doi/full/10.1073/pnas.1700351114

21. Lewis LS, Wessling EG, Kano F, Stevens JMG, Call J, Krupenye C. Bonobos and chimpanzees remember familiar conspecifics for decades. Proceedings of the National Academy of Sciences [Internet]. 2023 Dec [cited 2025 May 30];120(52):e2304903120. Available from: https://pnas.org/doi/10.1073/pnas.2304903120

22. Goodall J. The chimpanzees of Gombe: Patterns of behaviour. Cambridge, MA: Harvard University Press; 1986.

23. Muller MN, Mitani JC. Conflict and Cooperation in Wild Chimpanzees. In: Advances in the Study of Behavior [Internet]. Academic Press; 2005 [cited 2025 May 30]. p. 275–331. Available from: https://www.sciencedirect.com/science/article/pii/S0065345405350078

24. Mitani JC, Watts DP. Demographic influences on the hunting behavior of chimpanzees. American Journal of Physical Anthropology [Internet]. 1999 [cited 2025 May 30];109(4):439–54. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291096-8644%28199908%29109%3A4%3C439%3A%3AAID-AJPA2%3E3.0.CO%3B2-3

25. Boesch C. Cooperative hunting in wild chimpanzees. Animal Behaviour [Internet]. 1994 Sep [cited 2025 May 30];48(3):653–67. Available from: https://www.sciencedirect.com/science/article/pii/S0003347284712851

26. Samuni L, Preis A, Deschner T, Crockford C, Wittig RM. Reward of labor coordination and hunting success in wild chimpanzees. Communications Biology [Internet]. 2018 Sep [cited 2025 May 30];1(1):1–9. Available from: https://www.nature.com/articles/s42003-018-0142-3

27. Gilby IC, Machanda ZP, Mjungu DC, Rosen J, Muller MN, Pusey AE, et al. "Impact hunters" catalyse cooperative hunting in two wild chimpanzee communities. Philosophical Transactions of the Royal Society B: Biological Sciences [Internet]. 2015 Dec [cited 2025 May 30];370(1683):20150005. Available from: https://royalsocietypublishing.org/doi/10.1098/rstb.2015.0005

28. Boesch C. Cooperative hunting roles among taï chimpanzees. Human Nature [Internet]. 2002 Mar [cited 2025 May 30];13(1):27–46. Available from: https://doi.org/10.1007/s12110-002-1013-6

29. Redshaw J, Suddendorf T. Children's and Apes' Preparatory Responses to Two Mutually Exclusive Possibilities. Current Biology [Internet]. 2016 Jul [cited 2025 May 30];26(13):1758–62. Available from: https://linkinghub.elsevier.com/retrieve/pii/S096098221630416X

30. Suddendorf T, Crimston J, Redshaw J. Preparatory responses to socially determined, mutually exclusive possibilities in chimpanzees and children. Biology Letters [Internet]. 2017 Jun [cited 2025 May 30];13(6):20170170. Available from: https://royalsocietypublishing.org/doi/10.1098/rsbl.2017.0170

31. Redshaw J, Suddendorf T, Neldner K, Wilks M, Tomaselli K, Mushin I, et al. Young Children From Three Diverse Cultures Spontaneously and Consistently Prepare for Alternative Future Possibilities. Child Development [Internet]. 2019 [cited 2025 May 30];90(1):51–61. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/cdev.13084

32. Lambert ML, Osvath M. Comparing chimpanzees' preparatory responses to known and unknown future outcomes. Biology Letters [Internet]. 2018 Sep [cited 2025 May

30];14(9):20180499. Available from: https://royalsocietypublishing.org/doi/full/10.1098/rsbl.2018.0499

33. Engelmann JM, Völter CJ, Goddu MK, Call J, Rakoczy H, Herrmann E. Chimpanzees prepare for alternative possible outcomes. Biology Letters [Internet]. 2023 Jun [cited 2025 May 30];19(6):20230179. Available from: https://royalsocietypublishing.org/doi/10.1098/rsbl.2023.0179

34. Leahy B, Huemer M, Steele M, Alderete S, Carey S. Minimal representations of possibility at age 3. Proceedings of the National Academy of Sciences [Internet]. 2022 Dec [cited 2025 May 30];119(52):e2207499119. Available from: https://www.pnas.org/doi/10.1073/pnas.2207499119

35. Hare B, Call J, Tomasello M. Do chimpanzees know what conspecifics know? Animal Behaviour [Internet]. 2001 Jan [cited 2025 May 30];61(1):139–51. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0003347200915185

36. Schneider AC, Melis AP, Tomasello M. How chimpanzees solve collective action problems. Proceedings of the Royal Society B: Biological Sciences [Internet]. 2012 Oct [cited 2025 Sep 17];279(1749):4946–54. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2012.1948

37. Melis AP, Hare B, Tomasello M. Engineering cooperation in chimpanzees: Tolerance constraints on cooperation. Animal Behaviour [Internet]. 2006 Aug [cited 2025 May 30];72(2):275–86. Available from: https://www.sciencedirect.com/science/article/pii/S0003347206001461

38. Hare B, Melis AP, Woods V, Hastings S, Wrangham R. Tolerance Allows Bonobos to Outperform Chimpanzees on a Cooperative Task. Current Biology [Internet]. 2007 Apr [cited 2025 May 30];17(7):619–23. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0960982207010172

39. Koomen R, Herrmann E. Chimpanzees overcome the tragedy of the commons with dominance. Scientific Reports [Internet]. 2018 Jul [cited 2025 May 30];8(1):10389. Available from: https://www.nature.com/articles/s41598-018-28416-8

40. Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language [Internet]. 2008 Nov [cited 2025 May 30];59(4):390–412. Available from: https://www.sciencedirect.com/science/article/pii/S0749596X07001398

41. Paridon J van, Bolker B, Alday P. Multiple membership random effects [Internet]. 2023 [cited 2025 May 30]. Available from: https://jvparidon.github.io/lmerMultiMember/

42. Bates D, Maechler M, Bolker [aut B, cre, Walker S, Christensen RHB, et al. lme4: Linear Mixed-Effects Models using 'Eigen' and S4 [Internet]. 2024 [cited 2024 May 9]. Available from: https://cran.r-project.org/web/packages/lme4/index.html

43. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language [Internet]. 2013 Apr [cited 2025 May 30];68(3):255–78. Available from: https://www.sciencedirect.com/science/article/pii/S0749596X12001180

44. Wickham H. ggplot2: Elegant graphics for data analysis [Internet]. Springer-Verlag New York; 2016. Available from: https://ggplot2.tidyverse.org

45. Hare B, Call J, Agnetta B, Tomasello M. Chimpanzees know what conspecifics do and do not see. Animal Behaviour [Internet]. 2000 Apr [cited 2025 May 30];59(4):771–85. Available from: https://www.sciencedirect.com/science/article/pii/S0003347299913775

46. Hare B, Tomasello M. Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. Animal Behaviour [Internet]. 2004 Sep [cited 2025 May 30];68(3):571–81. Available from: https://www.sciencedirect.com/science/article/pii/S0003347204001678

47. Engelmann JM, Völter CJ, O'Madagain C, Proft M, Haun DBM, Rakoczy H, et al. Chimpanzees consider alternative possibilities. Current Biology [Internet]. 2021 Oct [cited 2025 May 30];31(20):R1377–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0960982221012525

48. Suddendorf T, Watson K, Bogaart M, Redshaw J. Preparation for certain and uncertain future outcomes in young children and three species of monkey. Developmental Psychobiology [Internet]. 2020 [cited 2025 May 30];62(2):191–201. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/dev.21898

49. Redshaw J, Suddendorf T. Can chimpanzees conceive of mutually exclusive future possibilities? A Comment on: "Chimpanzees prepare for alternative possible outcomes" (2023), by Engelmann *et al.* Biology Letters [Internet]. 2024 Jun [cited 2025 May 30];20(6):20230409. Available from: https://royalsocietypublishing.org/doi/10.1098/rsbl.2023.0409

50. Tomonaga M, Imura T, Mizuno Y, Tanaka M. Gravity bias in young and adult chimpanzees (Pan troglodytes): Tests with a modified opaque-tubes task. Developmental Science [Internet]. 2007 [cited 2025 May 30];10(3):411–21. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-7687.2007.00594.x