UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Herding cats: children’s intuitive theories of persuasion predict slower collective decisions in larger and more diverse groups, but disregard factional power

Permalink
https://escholarship.org/uc/item/7sj3b4bv

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Richardson, Emory
Hok, Hannah
Shaw, Alex
et al.

Publication Date
2023

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Herding cats: children's intuitive theories of persuasion predict slower collective decisions in larger and more diverse groups, but disregard factional power

Emory Richardson (emory.richardson@yale.edu), **Hannah Hok** (hok@uchicago.edu),
Alex Shaw (ashaw1@uchicago.edu), & **Frank Keil** (frank.keil@yale.edu)

Abstract

Collaboration can make collective judgments more accurate than individual judgments, but it also comes with costs in time, effort, and social cohesion. But how do we estimate these costs? In two experiments, we introduce children and adults to two teams in which the teammates disagree about the optimal solution to a novel problem, and ask which team would need more time to reach a consensus decision. We find that all ages expect slower decisions from teams with more people or factions, and expect the number of factions to matter more than the number of people. But only adults expect decisions initially endorsed by a stronger faction to be faster than those endorsed by a weaker faction. Results are discussed in context of children's reasoning about dominance, and models of time-rational collective decision-making.

Keywords: collective behavior, decision speed, group processes, conceptual development

Introduction

Reaching consensus can feel akin to herding cats: time-consuming and sometimes hopeless. But the struggle's not unique to committees of colicky faculty or poorly managed advisory panels. Differences of opinion are inevitable in groups, and time spent debating those differences adds up. Since people may agree on *what* to do without agreeing on *why*, discussions can easily involve more opinions than people, even in groups debating a yes-no decision about a single option. While some of those debates are sure to be more substantive than others, the clock ticks just as quickly for groups quibbling over minutiae as groups deliberating about substantive issues. And since one person's molehill may be another's mountain, dissent could continue to undermine consensus indefinitely. But it doesn't. We're not cats, after all; humans excel at collaboration and coordination. By adulthood, it seems commonsensical that collaborators need to weigh the costs of debate against the benefits. In some cases, getting consensus on your side may simply be too unlikely or too time-consuming to make a difference of opinion worth debating.

Clearly, the social dynamics that drive collective decision making are complex. But reasoning about how they contribute to decision speed doesn't seem to require much effort. For instance, it seems commonsensical that large groups will need more time than small groups to make decisions, or that groups in which a strong initial consensus can pressure dissenters to concede will make decisions more quickly than groups facing multipolar negotiations with no initial consensus at all. We suggest that these inferences feel effortless because they are generated by an intuitive theory (or suite of them) which inputs our beliefs about the constraints on a group decision and outputs systematic inferences about the ways we can influence the group's opinion dynamics — including outcomes, but also costs in time, effort, and social cohesion. Intuitive theories may begin as little more than a few salient cues and some beliefs about their causal connections (Keil, 2011; Mahr & Csibra, 2022), and they don't need to be particularly accurate or precise. They simply need to allow us to navigate a conceptual domain in everyday life, and be flexible enough to accommodate conceptual change and development.

Here, we provide evidence of systematic inferences about group decision speed in children and adults. We suggest these inferences may emerge from a few causally-connected intuitions about how group decision-making works: (1) expressing an opinion takes time, (2) debating differences takes even more, and while (3) not every difference of opinion is worth debating, a team's size and structure can make the cost-benefit tradeoffs of debate different for different teammates.

To illustrate how these intuitions generate predictions about decision speed, consider a robotics team deliberating over seven kinds of propeller for a drone (Fig 1). Discussion takes time, but any teammate can *concede* whenever they want — either because they've been convinced or because they simply don't think it's worth arguing further. However, one person's unilateral concession is only guaranteed to

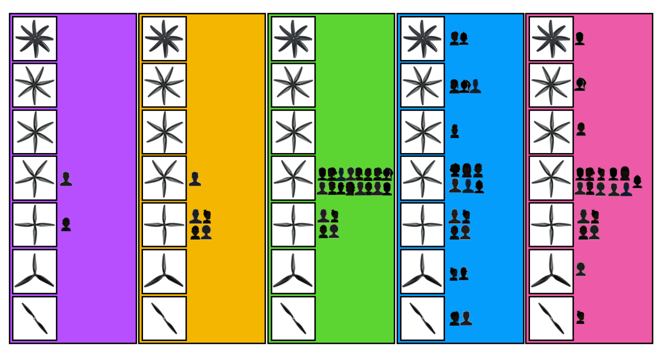

Figure 1: Each panel shows a robotics team divided into factions. Since the teammates disagree, they need to talk together to decide which propeller to use. Compare panels pairwise: which team took longer to decide, *assuming both teams decide to use the 4-blade propeller*? Team size is identical in Panels 3-5 and proportional distributions are identical in Panels 2-3, but intuitions about decision speed seem to vary semi-independently of the 4-blade faction's size or proportion.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*. ©2023 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

save time in Panel 1, where the debate will end as soon as either teammate concedes (assuming the other doesn't). By contrast, out of the five teammates in Panel 2, only one person can end the debate unilaterally by conceding: after all, even if one of the other four conceded, their former allies could continue to argue. And in every other Panel, no single person can unilaterally end the debate: the teammates *have to* spend time coordinating within and across factions in order to reach *any* consensus, regardless of whether they're arguing for their own propeller or simply trying to find an expedient option.

In short, the harder it is for teammates to predict each others' behavior, the more coordination will be needed to make a decision; and the more coordination required, the slower the team will be. Our story is simply that if our intuitive theories make reasoning about these dynamics relatively effortless, we may be able to make more rational use of our time and effort in collective action. But the constraints on group decisions our intuitive theories are most sensitive to may change from early development to adulthood. We'll return to this point in a moment.

We predict that adults will expect slower decisions from teams with more people or factions. But we also predict that they'll also expect quicker decisions from teams in which consensus is already strong at the outset than from teams in which power is initially more equally distributed between factions. Why? Because consensus is not just an outcome; it's also an epistemic and normative influence on people's responses to disagreement (Morgan & Laland, 2012; Kameda, Toyokawa, & Tindale, 2022). For instance, adults defer more to polls showing a 16v4 majority than either a proportionally weaker 11v9 majority or numerically weaker 4v1 majority (Mannes, 2009). Being mutually aware that a team is converging on a consensus decision may put dissenters under increasing pressure to concede even if they disagree, and make the strongest faction increasingly difficult to fracture. Importantly, factional power is more than just degree of consensus. In teams with multiple factions, one faction's power over another may depend as much on their dynamics with a third as their own size or proportion. In other words, factional power is a matter of "party discipline" (i.e., how strictly individuals subordinate their idiosyncrasies to the interests of their factions) as well as conformist tendencies (i.e., deference to consensus). However, conceptual development in early childhood often produces qualitative changes in our intuitive theories (Carey, 2009). Here, we examine 6-to-9-year-olds inferences about how the number of people, factions, and consensus strength affect decision speed.

We predict that young children, like adults, will expect slower decisions from teams with more people or factions. Why? First, talk takes time, and reasoning about the relationship between time, effort, and task difficulty emerges early. Even four- and five- year-olds expect agents to take longer to complete more difficult physical tasks (Leonard, Bennet-Pierre, & Gweon, 2019); and by age six, children also begin to expect agents to take longer to solve more complex reasoning problems, unless the agent has seen the solution before (Richardson & Keil, 2022). So, children may infer that having more factions or people on a team leads to slower decisions because it makes coordination into a more complex or effortful task. Second, children may also be able to infer how much talk goes into resolving disagreements by drawing on their own experience of collaborative reasoning. Three-year-olds explicitly dispute statements they believe to be false (Köymen & Tomasello, 2018), but how much of their reasoning they verbalize depends on what they expect their collaborators to know already (Köymen, Mammen, & Tomasello, 2016). And while preschoolers do evaluate each others' reasoning, they only begin to engage in meta-talk comparing higher-order evidence such as their relative confidence or their informants' reliability between the ages of five and seven (Köymen & Tomasello, 2018). Along with believing that difficult tasks take longer and that coordination gets harder in larger groups, being inefficient collaborators could make children especially sensitive to how increasing the number of people or factions on a teams can slow down collective decisions.

However, reasoning about how consensus strength impacts decision speed may be more challenging for children. Why? First, at least one mechanism that allows adults to speed up group decisions seems to be less reliable in children: while preschoolers conform to majority opinion in both informational and normative contexts, *stronger* deference to proportionally *larger* majorities only emerges around age six or seven, even with only two factions to consider (Morgan, Laland, & Harris, 2015). That is, preschoolers are no more deferential to a 9v1 majority than a 6v4 majority — and they are selective about when they defer to majorities to begin with (Burdett et al., 2016; Haun, van Leeuwen, & Edelson, 2013, Pham & Buchsbaum, 2020). And children don't simply become more deferential; they also become more selective about deferring. Though seven year olds are more likely to defer when uncertainty is high, they are also more likely to point out when they think the emperor is clearly naked (Morgan et al, 2015). Second, strategic deference in group contexts is rarely just a matter of votes; it often depends on how we evaluate each others' approximate explanations of matters we only partially understand to begin with (Keil, 2006). Children are much less skilled than adults in adjudicating conflicting explanations, and often strikingly overconfident in their own knowledge (Kloo, Rohwer, & Perner, 2017; Mills & Keil, 2004). Taken together, these findings suggest that (1) disputes over idiosyncratic and fundamental differences may not be as strictly triaged or efficiently resolved in groups of children as in groups of adults, and that (2) at least one mechanism that speeds up decisions in adults — stronger epistemic deference to stronger consensus, particularly without argument — may be less reliable in children. Thus, while children may expect slower decisions from teams with more factions or more people, they may not expect consensus strength to increase decision speed.

To be clear, however, the claim is not that children fail to recognize differences in consensus strength per se. Even preschoolers can accurately represent and compare small differences in numerical sets (Halberda & Feigensen, 2008;). Moreover, we think it's clear that children can make some inferences about power from relative group size (Pun, Birch, & Baron, 2016; Heck, Bas, & Kinzler 2021). But as noted above, reasoning about how consensus strength affects decision speed may involve capacities still developing between the ages of 6-9.

In two pre-registered experiments[1], we tested our predictions by presenting children and adults with pairs of robotics teams deciding which of seven kinds of propeller would make a drone fly the best. In each trial, the two teams vary in the number of people, factions, or both. Participants are told that the teammates on each team will have to talk together to decide which propeller to use. They then rate how sure they are that one team or the other would take longer to decide on a seven-point scale (with the midpoint indicating no difference), and briefly explain their reasoning.

Experiment 1

In Experiment 1, we asked children and adults to infer which of two teams would take longer to make a decision. Across three trials, we manipulated the number of people (*Size*), factions (*Diversity*), or both (*Contrast*). In the *Diversity* trial, two teams with the same number of people (10) were split into a different number of factions (2v7). In the *Size* trial, two teams with the same number of factions (2) differed in the number of people (10v20). In the *Contrast* trial, the team with more people (20v10) was split into fewer factions (3v7). We predict that both children and adults will expect slower decisions from teams with more factions or more people, and that they will treat the number of factions as more important than the number of people (i.e., in *Contrast*). However, we expect these inferences to be specific to decisions: in a second task following the experiment (*Build*), we ask which of two teams (20v10) would take longer to *build* their drone, *after* a consensus decision had been agreed upon. In the *Build* trial, we predict that participants will expect a *smaller* team to take longer than a larger team: whereas the task of reaching consensus divides a team against itself, many hands may make light work once consensus is reached. The *Build* and *Contrast* trials also help rule out a simple "more is more" heuristic. If participants are simply mapping the "more time" response to the team with more people or more factions, they will expect no difference in decision speed when one team has more people and the other has more factions, and they will infer that that the *larger* team will take more time to build a drone.

Participants. Based on power analysis, we recruited 80 children in two age groups (40 age 6-7, M=6.95, SD=.50, and 40 age 8-9, M=8.98, SD=.58; 34 girls, no non-binary genders reported), as well as 41 adults through MTurk. One additional child fussed out before completing the experiment and was replaced.

Procedure. After practicing with the response scale, children were told that they would see two teams each making a remote control drone, but that the teammates disagreed about which of seven kinds of propeller (differentiated by the number of blades, from two to eight) would make the drone fly the best. The experimenter told the child that they would see "which kind of propeller *each* person on *each* team *thinks* is best", and that the teammates would need to talk together to decide which kind of propeller to use. The child's job was "to say *which* team you think will take *longer to decide* which kind of propeller to use". They were then shown three trials in one of four counterbalanced orders. In each trial, participants first saw a group of students, represented as silhouettes, divided into two teams (allowing for easy visual comparison of the total number of people on each team), and then were shown each teammate "standing next to" the propeller they thought was best. The experimenter then told the participant "*So now, all the people on the blue team have to talk together to decide which propeller to use. And all the people on the green team have to talk together to decide which propeller to use. But, which team will take longer to decide: the blue team, the green team, or will they take the same amount of time?*". Children were then asked whether they were "*just a little sure, pretty sure, or very sure?*"; adults responded directly on a sliding scale. Participants were then asked to explain why they thought that team would take longer to decide. Finally, at the end of experiment, participants completed one trial of a second task: they were told that the next two teams had *already decided* which kind of propeller to use, and all agreed — but now, they needed to build their drone. One team was shown to have 10 people while the other had 20 people; participants were told that each team would start building at the same time, and asked which team would take longer to finish *building* their drone.

Results and Discussion. We conducted separate linear regressions on the child sample alone for each contrast *Type*, with responses centered on the midpoint of the 7-point scale and age in years centered on the midpoint of the children's age range (7.5 years), according to our pre-registered analysis plan. This makes the intercept equivalent to a one-sample t.test versus the scale midpoint while allowing us to simultaneously account for potential age effects. There was no effect of counterbalance for any measure or age group, so we reduced the model to just Ct_Values ~ Ct_AgeYears for each contrast *Type*. Adults and the child sample as a whole expected reaching consensus *decisions* about how to build a drone to take *longer* in *larger* teams than smaller teams, but physically *building* one after deciding to take *less* time (*Size* trial: $\beta_{Intercept}$ = -1.05, SE = .21, p<.0001, *95CI*: -1.46–-0.64; *Build* trial: $\beta_{Intercept}$ = 1.48, SE = .22, p<.0001, *95CI*: 1.04– 1.91). However, they also inferred that reaching consensus would take longer in teams divided into more factions, regardless of whether the more factious team was the *same size* (*Diversity* trial: $\beta_{Intercept}$ = 2.35, SE = .13, p<.001, *95CI*: 2.10–2.60) or *smaller* than the less factious

[1] Link to pre-regs, materials, power analyses, data: https://osf.io/9xtyu/?view_only=037914869b2c43b2bfeff1a3e4134bb7

team (*Contrast* trial: $\beta_{Intercept} = 1.71$, SE = .20, $p<.0001$, *95CI*: 1.31– 2.11). No age effects were observed in the *Diversity* or *Contrast* trials; however, age was significant for both *Build* and *Size*, with older children more likely than younger children to infer that larger groups would take more time to decide, and less time to build (*Size*: $\beta_{Ct\_AgeYears} = -0.50$, SE = .19, $p<.01$, *95CIs*: -0.87– -0.13; *Build*: $\beta_{Ct\_AgeYears} = -0.41$, SE = .20, $p=.039$, *95CI*: -0.80– -0.02). Following our preregistered analysis plan, we also conducted one-sample t.tests comparing each age group (6-7s, 8-9s, and adults) to chance separately for each measure. The expectation of slower decisions from the larger team was not significant for the youngest children in the *Size* trial (M = -0.60, $t(39)= -1.71$, $p<.095$, *95CI* -1.31—0.11); all other t.tests supported our primary analysis.

What do these results tell us about participants' reasoning process? If participants were simply mapping a "more time" response to the team with more people or more factions, they wouldn't have expected team size to have opposite effects on cognitive decision speed (*Size* trial) and physical build speed (*Build* trial). A more-is-more heuristic also fails to explain why participants would expect decision speed to depend more on the number of factions than the number of people. Instead, participants' inferences appear to reflect their beliefs about how teams resolve (potential) disagreements between individuals. But Experiment 1 alone doesn't tell us what those beliefs are. For instance, one might simply assume *an* outcome (either majority rule, or whichever propeller seemed best to the participant themselves), and infer decision speed from the number of opponents remaining to be convinced. This is akin to the kind of reasoning predicted by our account, but because it's blind to differences in power that make some outcomes more likely than others, it will often generate counterintuitive predictions. For instance, one might expect convincing four people to always require the same amount of time, regardless of the number of factions and people in them (e.g., 16v4, 16v1v1v1v1, 2v4, 1v4, etc). In Experiment 2, we control for the numerical and proportional size of the winning and losing factions as well as the total number of people and factions by asking participants to infer which team would take longer given that *both* teams chose the *same* propeller. We chose to describe this propeller as *optimal* in order to minimize variability in participants' responses stemming from their different assumptions about whether the team had chosen accurately. Thus, it's possible that stipulating an *inaccurate* winning faction would lead participants to reason differently about decision speeds.

Experiment 2

Experiment 2 probes participants' reasoning about how consensus strength affects decision speed. We predict that all ages will infer slower decisions from teams with more factions or people. But if the winning faction has less power relative to its opponents on a small team than the winning faction has on a large team, we predict that while adults will expect consensus strength to matter more than size, children will infer the opposite. For instance, adults might expect a minority rule outcome on a team of six to take longer than a majority rule outcome on team of twelve, children will infer the opposite. However, because Experiment 1 and the pilot data for Experiment 2 suggested that younger children's (ages 6-7) size inference may not differ from chance despite recognizing the impact of the number of factions, our preregistration treats older children as the primary developmental contrast for the trials in which size and factional power are contrasted. Younger may show the same pattern as older children; but if they do not differ from chance, further work would be needed to understand why.

Participants. Based on power analysis, we recruited 100 children in two age groups (50 age 6-7, M=6.88, SD=.67, and 50 age 8-9, M=8.98, SD=.67; 60 girls, no non-binary genders reported), as well as 50 adults through MTurk. Two children fussed out before completing the experiment and were replaced; six adults were screened out and replaced before completing the experiment for failing an attention check.

Materials. We created four trials intended to contrast different dimensions of the distribution of opinions on each

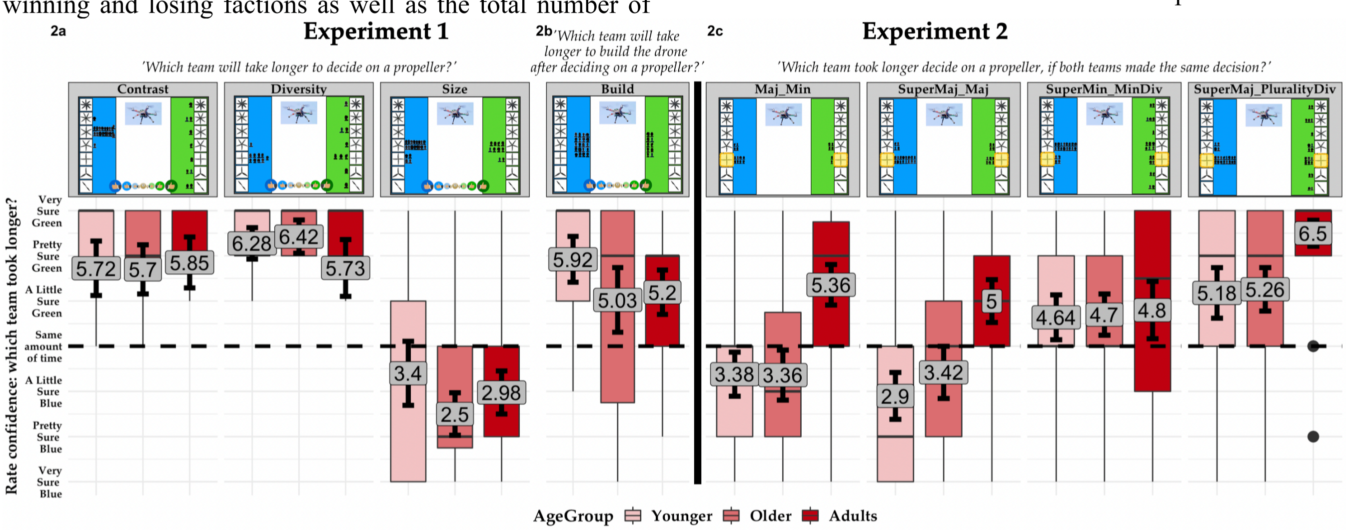

Figure 2: **Results**. In Exp. 1, participants each rated which team would take longer to decide (Fig. 2a: *Contrast, Diversity, Size*), or which team would take longer to build a drone in the final trial (Fig. 2b: *Build*). In Exp. 2 (Fig. 2c), participants rated which team had taken longer to decide, supposing that both had chosen the propeller highlighted in yellow.

team: the size of each team, the number of options initially endorsed, and the proportion and number of teammates who had initially disputed the group's final decision. In two trials (*Maj_Min*, *SuperMaj_vs_Maj*), one team was twice the size of the other, but each team was split between two options, and choosing the correct propeller would require the team to convince 4 people to change their answer (*Maj_Min*: 8v4 or 2v4; *SuperMaj_vs_Maj*: 16v4 or 6v4). In the other two trials (*SuperMin_MinDiv*, *SuperMaj_PluralityDiv*), each team was the same size, but one team was split between all six options while the other team was split between only two options , with a either a plurality or majority initially endorsing or opposing the correct propeller (*SuperMin_MinDiv*: 4v16 or 4v6v3v2v2v1; *SuperMaj_PluralityDiv:* 16v4 or 6v4v3v2v2v1).

Procedure. The procedure was similar to Experiment 1, with the following changes. (1) First, during the introduction, participants were additionally told that "*the kind of propeller that's actually the best for the kind of drone these teams are both building is the one* 4-blades", after which the 4-blade propeller was highlighted in yellow and remained highlighted for the remainder of the experiment. (2) Second, after seeing during each trial what each teammate on each team thought was best, participants were prompted to remember which propeller was actually best. (3) Third, the experimenter told participants to pretend that both teams had ultimately chosen the correct propeller, saying: "*Now the teammates on each team have to talk together to decide which propeller to use. And each team might decide to use the 4-blade propeller, or they might not. And we don't know which propeller they'll choose after they talk. But, let's pretend we do know. Let's pretend that after they talk, both the blue team and the green team do decide to use the 4-blade propeller. So, which team do you think had to talk for longer, if both teams decided to use the 4-blade propeller: did the blue team take longer, did the green team take longer, or did they both take the same amount of time?*". (4) Finally, after rating how sure they were that one team or the other would take longer and explaining why, the experimenter told the participants "*Now we're done pretending for a minute. Remember, we don't actually know which propeller each team will decide to use — but, I want to know which propeller you think each team will use*", and for each team, asked the participant to predict whether the team would decide to use the 4-blade propeller after talking.

Results and Discussion. Experiment 2 provides direct evidence against a number of heuristics simpler than the kind of reasoning about disagreement we have in mind. Across trials, children and adults made systematic inferences even when we controlled (1) the total number of people, (2) the total number of factions, (3) the number of "losers" (4) the proportion of "losers", and (5) the number and proportion of "winners".

On the *SuperMaj_PluralityDiv* and *SuperMin_MinDiv* trials, each team had 20 teammates; as predicted, children and adults expected slower decisions when they were divided into 7 factions than when they were divided into only 2 factions — not only when the team with more factions was contrasted with a team with a stronger winning faction (*SuperMaj_PluralityDiv*: 16-winners-vs-4-losers-in-1-faction and 6-winners-vs-14-losers-in-6-factions: $M_{younger}$= 5.18, $t(49) = 4.24$, $p<.001$; M_{older}=5.26, $t(49) = 5.28$, $p=.003$,, M_{adult} =6.50, $t(49) = 17.43$, $p<.001$), but also when contrasted with a team with the same number and proportion of both winners and losers (*SuperMin_MinDiv*: 4-winners-vs-16-losers-in-1-faction and 4-winners-vs-16-losers-in-6-factions: $M_{younger}$= 4.64, $t(49) = 2.59$, $p=.013$; M_{older}=4.70, $t(49) = 3.08$, $p=.003$,, M_{adult} =4.80, $t(49) = 2.54$, $p=.014$). One-way ANOVAs revealed that younger children were significantly less confident than adults on the *SuperMaj_PluralityDiv* trial; but older children's responses were not significantly different from either younger children's or adults' for either trial (*SuperMaj_PluralityDiv*: F(2, 147)=54.77, $p<.001$, $\eta_p^2 = .13$; *Younger—Adult*: $t(147)=-4.11$, $p < .001$; *Younger—Older*: all p's ns; *SuperMin_MinDiv*: F(2, 147)=0.65, *p= ns; Older—Adult*: $t(147)=-0.27$, $p <$ ns; *Younger—Adult*: $t(147)=-0.43$, $p <$ ns).

On the *Maj_Min* and *SuperMaj_Maj* trials, each team was divided into 2 factions that left each team with the same number of "losers" to convince, but also made one team on each trial twice the size of the other (*Maj_Min*: 8v4-and-2v4; *SuperMaj_Maj*: 16v4-and-6v4). As predicted, adults inferred on both trials that the decision would have been slower when the winning faction was proportionally weaker, but children inferred that decisions would have been slower in the numerically larger team, even though the winning faction was proportionally stronger (*Maj_Min*: $M_{younger}$= 3.38, $t(49) = -2.56$, $p=.014$; M_{older}= 3.36, $t(49) = -2.33$, $p=.024$; M_{adult} = 5.36, $t(49) = 6.11$, $p<.001$; *Supermajority*: $M_{younger}$= 2.90, $t(49) = -4.27$, $p=.001$; M_{older}= 3.42, $t(49) = -2.02$, $p=.049$; M_{adult} = 5.00, $t(49) = 4.24$, $p<.001$). As predicted, these age differences were significant for older children (*Maj_Min*: *Older-Adult*: $t(147)=-5.71$, $p < .001$; *SuperMaj_Maj*: *Older-Adult*: $t(147)=-4.28$, $p < .001$). The pattern for younger children also differed from adults, but was indistinguishable from older children (*Maj_Min*: *Younger-Adult*: $t(147)=-5.65$, $p < .001$; *SuperMaj_Maj*: *Younger-Adult*: $t(147)=-5.69$, $p < .001$.

Since *SuperMaj_Maj* and *SuperMin_MinDiv* each contrasted two teams in which the winning faction was the initial majority, these results also speak against the possibility that inferences about decision speed are simply an artifact of assuming that only one of the teams (that without a majority) would need any time at all to make a decision. But when asked to predict each team's final decision, all ages expected majority rule — and to a lesser extent, plurality rule (i.e., in teams with many factions) — regardless of whether the propeller endorsed by the initial majority (or plurality) faction was the optimal decision or not. In other words, both children and adults *predicted* majority rule (and to a lesser extent, plurality rule), but their inferences about decision speed were not simply an artifact of assuming majority rule as a *fait accompli*.

General Discussion

Knowing how to pick your battles can make group decisions more cost-effective. One cost is time. Taken together, our experiments show that children and adults make systematic inferences about the time needed to reach consensus by reasoning about how opinion dynamics are constrained by groups' size and structure. Adults and children as young as six expected more people or factions to slow down decisions. And these inferences were specifically about reaching consensus decisions: all ages expected that once consensus had been reached, teams with more people would be *faster* at *building* drones. But while adults expected stronger initial consensus to speed up consensus-congruent decisions (and slow down consensus-incongruent decisions), children expected slower decisions from larger teams even when consensus was stronger than on the smaller team. We doubt children's size-over-strength inferences are due to a failure to realize that consensus was stronger on one team than the other. Even preschoolers can easily distinguish the vote ratios (2:1, 3:2, 4:1) we used in the two-faction trials (Halberda & Feigenson, 2008). And children inferred that the larger team's majority-rule decision had been slower than the smaller team's despite knowing the smaller team had followed minority rule.

Developmental changes in commonsense reasoning about how power asymmetries constrain decision-making may shed light on our capacity for collective action. Though children in our experiments didn't expect power asymmetries to constrain decision speeds, our results don't suggest that they lack any concept of factional power. For instance, we suspect that if we *equated* team size instead of pitting it *against* proportions in two-faction teams (e.g., 18v2-or-12v8, or 8v2-or-2v8), children's inferences might match adults. But doing so would also differentiate teams on at least two other dimensions children could use to infer decision speed (e.g., the number of "losers" and majority vs minority rule for an 8v2-or-2v8 contrast). And Experiment 2 already shows that children's inferences are similar to adults' when team size is equated (in the two trials where we contrasted a two-faction team with a many-faction team). One potential explanation for these results is that children and adults infer a faction's power over its opponent(s) from their relative sizes, but children expect less within-faction cohesiveness than adults. Why?

One reason may be that individual inefficiencies in communication add up quickly in large groups. And since children are much less skilled than adults at resolving conflict through meta-talk and reason-giving (Köymen & Tomasello, 2018), it wouldn't be unreasonable for them to expect slower decisions from larger groups in general; after all, when consensus is controlled, so do adults. But while in adults, strength-dependent deference to consensus can short-circuit endless dissent and redundant commentary, this deference is only beginning to emerge around ages 6-7 (Morgan, 2015; Schmidt et al., 2016). Moreover, while children can explicitly justify the use of different decision-making procedures in different contexts (Helwig & Kim, 1999; Hok, et al, 2023), a more rigid sense of procedural justice (e.g., "everyone should have their say") could make majority rule as time-consuming as unanimous consensus. And since social dynamics in collective decisions are to some extent consequences of our beliefs about how to manage them (cf. Pietraszewski, 2022), it's possible that children and adults were simply accurately reporting their experience of group decision-making. But lacking direct tests of children's collective decision times, whether or not children's inferences accurately reflect their experience — and whether developmental changes in reasoning about factional power could improve speed-accuracy tradeoffs in collective decision-making — are questions for future work.

Expecting such reasoning from children might seem odd. But managing speed-accuracy tradeoffs is no less critical to groups than individuals, and group size and structure are endogenous constraints on group decisions. Lower decision thresholds (e.g., plurality or majority instead of supermajority or unanimity) can speed up decisions, but economic models of committee deliberations show that they can also give stubborn minorities leverage over moderate majorities if swing voters are impatient (Chan, et al, 2018). Similar dynamics are observed in other species. When *temnothorax* ants urgently need to find a better nest, they lower their quorum threshold — enabling the "votes" of a smaller number of scouts to trigger a migration (Pratt & Sumpter, 2006). And when schooling fish choose a foraging patch, increasing the number of no-preference voters makes it harder for strong-preference minorities to overrule weak-preference majorities (Couzin et al., 2011). Though constraints on group decisions aren't unique to humans, a capacity for commonsense reasoning about complex social dynamics may make us especially efficient at guiding collective action (Heyes, 2016).

One capacity that may be critical is the ability to shift between different forms of "government" as constraints change. Our scenario emphasized deliberative consensus, and implied egalitarianism. But in other cases, social status may influence decisions more than egalitarian deliberation. Even infants infer dominance relations between two individuals from patterns of deference in zero-sum conflicts (Mascaro & Csibra, 2012) and expect a smaller individual to defer to a larger one (Thomsen, et al, 2011). But more work is needed to understand how we balance individual dominance or prestige against power based on alliances. For instance, infants may expect an agent with one physically large ally to make way on a narrow bridge for an agent with two smaller allies (Pun, et al., 2016; but see Yousif & Keil, 2021). And preschoolers infer that the smaller of two groups is more likely to be "in charge" while the larger group is more likely to "get the stuff" (Heck, et al., 2021). But physical and factional size not only confer different kinds of power; as children and adults in our studies infer, they may also incur different kinds of costs. Further research into children's reasoning about factional power may shed light on how we learn to manage group dynamics. Good intuitions about when power asymmetries make some battles not worth the time may help ensure that reaching consensus isn't always akin to herding cats.

References

Burdett, E. R. R., Lucas, A. J., Buchsbaum, D., McGuigan, N., Wood, L. A., & Whiten, A. (2016). Do Children Copy an Expert or a Majority? Examining Selective Learning in Instrumental and Normative Contexts. *PLOS ONE*, *11*(10), e0164698. https://doi.org/10.1371/journal.pone.0164698

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Chan, J., Lizzeri, A., Suen, W., & Yariv, L. (2018). Deliberating Collective Decisions. *The Review of Economic Studies*, *85*(2), 929–963. https://doi.org/10.1093/restud/rdx028

Couzin, I. D., Ioannou, C. C., Demirel, G., Gross, T., Torney, C. J., Hartnett, A., Conradt, L., Levin, S. A., & Leonard, N. E. (2011). Democratic Consensus in Animal Groups. *Science*, *334*(6062), 4. https://doi.org/10.1126/science.1210280

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–1465. https://doi.org/10.1037/a0012682

Haun, D. B. M., van Leeuwen, E. J. C., & Edelson, M. G. (2013). Majority influence in children and other animals. *Developmental Cognitive Neuroscience*, *3*, 61–71. https://doi.org/10.1016/j.dcn.2012.09.003

Heck, I. A., Bas, J., & Kinzler, K. D. (2021). Small groups lead, big groups control: Perceptions of numerical group size, power, and status across development. *Child Development*, *93*(1), 194-208. https://doi.org/10.1111/cdev.13670

Helwig, C. C., & Kim, S. (1999). Children's Evaluations of Decision-Making Procedures in Peer, Family, and School Contexts. *Child Development*, *70*(2), 502–512. https://doi.org/10.1111/1467-8624.00036

Heyes, C. (2016). Who Knows? Metacognitive Social Learning Strategies. *Trends in Cognitive Sciences*, *20*(3), 204–213. https://doi.org/10.1016/j.tics.2015.12.007

Hok, H., Gerdin, E., Zhao, X., & Shaw, A. (2023). When Should the Majority Rule?: Children's Developing Intuitions About Majority Rules Voting [Preprint]. SSRN. https://doi.org/10.2139/ssrn.4409263

Kameda, T., Toyokawa, W., & Tindale, R. S. (2022). Information aggregation and collective intelligence beyond the wisdom of crowds. *Nature Reviews Psychology*. https://doi.org/10.1038/s44159-022-00054-y

Keil, F. (2011). The Hidden Strengths of Weak Theories. *Anthropology and Philosophy*, *10*(1–2), 61. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4190847/

Keil, F. C. (2006). Explanation and Understanding. *Annual Review of Psychology*, *57*(1), 227–254. https://doi.org/10.1146/annurev.psych.57.102904.190100

Kloo, D., Rohwer, M., & Perner, J. (2017). Direct and indirect admission of ignorance by children. *Journal of Experimental Child Psychology*, *159*, 279–295. https://doi.org/10.1016/j.jecp.2017.02.014

Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology*, *52*(3), 423–429. https://doi.org/10.1037/dev0000089

Köymen, B., & Tomasello, M. (2018). Children's meta-talk in their collaborative decision making with peers. *Journal of Experimental Child Psychology*, *166*, 549–566. https://doi.org/10.1016/j.jecp.2017.09.018

Köymen, B., & Tomasello, M. (2020). The Early Ontogeny of Reason Giving. *Child Development Perspectives*, *14*(4), 215–220. https://doi.org/10.1111/cdep.12384

Leonard, J. A., Bennett-Pierre, G., & Gweon, H. (2019). Who is better? Preschoolers infer relative competence based on efficiency of process and quality of outcome. *Proceedings of the Cognitive Science Society*, 7.

Levine, M. E., & Plott, C. R. (1977). Agenda Influence and Its Implications. *Virginia Law Review*, *63*(4), 561. https://doi.org/10.2307/1072445

Mahr, J. B., & Csibra, G. (2022). A Short History of Theories of Intuitive Theories. In J. Gervain, G. Csibra, & K. Kovács (Eds.), *A Life in Cognition* (Vol. 11, pp. 219–232). Springer International Publishing. https://doi.org/10.1007/978-3-030-66175-5_16

Mannes, A. E. (2009). Are We Wise About the Wisdom of Crowds? The Use of Group Judgments in Belief Revision. *Management Science*, *55*(8), 1267–1279. https://doi.org/10.1287/mnsc.1090.1031

Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences*, *109*(18), 6862–6867. https://doi.org/10.1073/pnas.1113194109

Mills, C. M., & Keil, F. C. (2004). Knowing the limits of one's understanding: The development of an awareness of an illusion of explanatory depth. *Journal of Experimental Child Psychology*, *87*(1), 1–32. https://doi.org/10.1016/j.jecp.2003.09.003

Morgan, T. J. H., & Laland, K. N. (2012). The Biological Bases of Conformity. *Frontiers in Neuroscience*, *6*. https://doi.org/10.3389/fnins.2012.00087

Morgan, T. J. H., Laland, K. N., & Harris, P. L. (2015). The development of adaptive conformity in young children: Effects of uncertainty and consensus. *Developmental Science*, *18*(4), 511–524. https://doi.org/10.1111/desc.12231

Pham, T., & Buchsbaum, D. (2020). Children's use of majority information is influenced by pragmatic inferences and task domain. *Developmental Psychology*, *56*(2), 312–323. https://doi.org/10.1037/dev0000857

Pietraszewski, D. (2022). Toward a computational theory of social groups: A finite set of cognitive primitives for representing any and all social groups in the context of conflict. *Behavioral and Brain Sciences*, *45*, e97. https://doi.org/10.1017/S0140525X21000583

Pratt, S. C., & Sumpter, D. J. T. (2006). A tunable algorithm for collective decision-making. *Proceedings of the National Academy of Sciences*, *103*(43), 15906–15910. https://doi.org/10.1073/pnas.0604801103

Richardson, E., & Keil, F. C. (2022). Thinking takes time: Children use agents' response times to infer the source, quality, and complexity of their knowledge. *Cognition*, *224*, 105073. https://doi.org/10.1016/j.cognition.2022.105073

Schmidt, M. F. H., Rakoczy, H., Mietzsch, T., & Tomasello, M. (2016). Young Children Understand the Role of Agreement in Establishing Arbitrary Norms-But Unanimity Is Key. *Child Development*, *87*(2), 612–626. https://doi.org/10.1111/cdev.12510

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and Mighty: Preverbal Infants Mentally Represent Social Dominance. *Science*, *331*(6016), 477–480. https://doi.org/10.1126/science.1199198

Yousif, S. R., & Keil, F. C. (2021). How We See Area and Why It Matters. *Trends in Cognitive Sciences*, *25*(7), 554–557. https://doi.org/10.1016/j.tics.2021.03.017