Coordinative Difficulty Drives Role-Taking in Group Tasks

Max Potter (potter.max2001@gmail.com)[1], Elizabeth Kao[1], Lalaine Emralino[1], Jack Terwilliger[1] & Federico Rossano[1]
[1]Department of Cognitive Science, University of California San Diego

Abstract

Human cooperation involves flexible role-taking. This form of cooperation has been attributed to collective intentionality – the ability to represent what "The Group" believes and desires. However, the factors determining when and why humans spontaneously assume roles remain understudied. We propose that a primary motivation for role-taking is to reduce coordinative complexity by making one's intentions transparent. We test this hypothesis in a naturalistic cooking study (N = 270, 90 trials) in which groups of two, three, and four collaborate, as additional members increase coordinative complexity by expanding possible task distributions and increasing inference misalignment. Our analysis shows that role-taking increases significantly with group size, supporting the hypothesis that roles are used to manage coordinative complexity. These findings provide insight into why humans choose certain cooperative strategies, suggesting that facilitating coordination and solidarity may be prioritized over optimizing productivity.

Keywords: Role-taking; groups; cooperation; coordination; collective intentionality; shared agency; multi-agent;

Introduction

Humans frequently cooperate to accomplish tasks impossible for an individual to achieve alone. A central feature of this cooperation is role-taking: the decomposition of complex tasks into individual specialized subtasks. Role-taking is pervasive in human cooperation from large-scale divisions of labor through career specialization (Durkheim, 1893) to consistent kinship and familial roles (Nadel, 1957), and even in moment-to-moment interactions, such as trading off between the roles of 'speaker' and 'listener' (Schegloff, 2007; Turner, 1962). While many species show forms of role-taking (Boesch, 2002; Dreyer et al., 2024; Goodall, 1986), humans appear to be uniquely distinct in their ability to spontaneously assume roles across a variety of contexts and in widely flexible ways (Tomasello, 2019). Understanding why humans take on and commit to roles is therefore critical for understanding the cognitive mechanisms that support flexible cooperation.

Despite roles constituting a critical component of human cooperation, the conditions that motivate humans to take on roles are understudied (Jara-Ettinger & Dunham, 2025). Studies rarely extend beyond dyads (Boyce et al., 2022) and often focus on existing or researcher-assigned roles rather than analyzing what causes role-formation.

Currently, a prominent explanation of role-taking is through the lens of organizational efficiency – divisions of labor can optimize parallelizable tasks (O'Brien, 1968) or introduce bottlenecks that hinder progress (Helmreich & Schaefer, 1994; Mieczkowski et al., 2025). This is supported by the finding that humans tend to assign task responsibility based on who would produce the most optimal outcome (Allen et al., 2015). However, humans also frequently persist in planned courses of action even when they are no longer optimal (Chu & Schulz, 2022; Zhai et al., 2024), suggesting that organizational efficiency alone is insufficient to explain why humans take on and commit to roles. This reveals a fundamental gap in knowledge: What causes humans to take on and maintain roles, even once they are no longer an efficient strategy?

To understand how role-taking might emerge, we turn to cognitive theories on collective intentionality that explain how humans reason about group-level goals and expectations. Shared intentionality has long been argued to be critical for cooperation (Bratman, 2014; Searle, 1990), with recent work arguing that "We"-minded reasoning may be directly responsible for the formation of roles (Tomasello, 2019). Group members cooperating can infer what the collective "we" needs to achieve a shared goal (Gilbert, 2013), naturally assuming roles to fulfill that need and adhere to a shared plan amongst group members (Chater et al., 2022; Tomasello, 2019). However, this inference is not always simple – especially when there are more group members and more possible distributions of responsibility, maintaining knowledge over who is doing what and aligning plans with partners can require complex inference. This complexity is even worse for more individualistic approaches, such as Bayesian Theory-Of-Mind, which demand that agents use a large number of inferences to determine each of their partners' beliefs, desires, and intentions in cooperation (Baker et al., 2017; Chater et al., 2022; Wu et al., 2021).

Roles may be an effective solution to circumvent the need for such costly inference. Roles come with a set of normative expectations about how the role-taker will act and what tasks they will fulfill (Eickers, 2024; Gilbert, 2016), and can be conceptualized independently of the agent that fulfills them (Rakoczy et al., 2014; Rakoczy, 2018). Since roles bind agents to a particular structure of possible actions, understanding a partner's role allows an agent to effectively predict that partner's actions without requiring a deeper inference about their mental states (Davis & Jara-Ettinger, 2022; Jara-Ettinger & Dunham, 2025). This, in turn, means that coordinating with established roles may require far less cognitive work (Ritchie, 2020).

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

Within a group task, representing the roles of each group member therefore enables agents to effectively infer and maintain a shared plan to support cooperation (Banerjee & Gray, 2023). This also illuminates why roles may be beneficial even when inefficient – committing to a role provides structured information to scaffold planning and thinking (Chu et al., 2024), and abandoning that role incurs costs by forcing agents to re-plan and risk misalignment.

Under this view, role-taking is not just a strategy of efficient organization but also an inherently cooperative act designed to align group members' plans and intentions. As such, members who take on roles may also be viewed as being more cooperative partners. Humans tend to prefer more strongly cooperative partners in collaboration (Barclay & Willer, 2006), suggesting that taking on roles may also improve one's perceived value within a group. This lines up well with intuition from both Durkheim (1893) and Tomasello (2019), who argue that role-taking is highly connected to social standing and promotes solidarity between group members.

We use these insights to propose that humans may take on roles specifically in response to growing coordinative difficulty. It is well established that humans attempt to ease the difficulty of inference for their collaborative partners by making their actions and intentions more visible when cooperating (Vesper et al., 2011, 2016). Since role-taking has been argued to be especially powerful at constraining inferential difficulty, we expect that humans will take on roles significantly more often in response to growing complexity – using roles as a tool to support coordination amongst group members.

While there are many ways to increase the coordinative difficulty of a task, we focus specifically on the effect of increasing group size. Having more group members increases the space of possible task distributions and makes inferring a shared plan, as well as aligning those inferences between group members, more difficult (Skau, 2022). Relevant work has suggested that group tasks raise inference demands by making it more difficult to process information and structure interaction (Kolfschoten et al., 2014). Furthermore, more members working in parallel raises the likelihood of partners accidentally interfering with each other (Hutchins, 1996), making careful coordination much more important.

The jump between groups of size 2 to size 3 may present the most significant increase in difficulty due to the introduction of new problems not present in dyads, such as the addressee problem. Listeners in groups of size 3+ can no longer automatically assume relevance of communicative signals, needing to infer whether they were an intended recipient (Clark, 1992; Jovanovic, 2007) and whether the signal is ultimately relevant to them (Sperber & Wilson, 1986). Speakers also now often need to intentionally design their speech to designate which other members are 'addressees' and which are 'overhearers' (Clark & Carlson, 1982; Goffman, 1976), though some signals may be simply addressed to the entire group (Jankovic, 2018). As a result, cooperation in a group of size 3 or greater requires distinct forms of reasoning not present in dyads, suggesting that coordinative complexity increases significantly when moving beyond dyadic interaction.

To empirically test if coordinative complexity causes humans to take on roles, we assess role-taking behavior in a naturalistic collective cooking experiment. Participants collaborate in groups of size 2, 3, or 4 to cook omelets in a shared kitchen space. Modulating group size is especially relevant as, despite ample research examining coordinative inference in dyadic contexts, research examining role-taking in groups larger than two is significantly more rare (Boyce et al., 2022; Goldstone et al., 2024). We hypothesize that role-taking will be significantly more prevalent in larger groups, and that groups that take on roles will also indicate more positive sentiment towards other group members after the task is done. We expect these differences to be most significant between groups of size 2 and size 3, as this jump introduces additional sources of ambiguity not present in dyads. These findings would provide evidence that role-taking is a social tool for supporting cooperation, with humans intentionally leveraging roles to simplify group coordination even when they may not be the most efficient strategy for speed or productivity.

Methods

This study was preregistered at the Open Science Framework (OSF; https://osf.io/yvb6r). Note that two measures here (Individual Entropy and Task Entropy, described below) were not pre-registered and instead added post-hoc in response to suggestions for more individual-specific metrics to measure role-taking.

Task

A task's decomposition into multiple simultaneous subtasks can dictate the effectiveness of role-taking (Mieczkowski et al., 2025; O'Brien, 1968). For instance, consider the process of a group preparing a salad vs. preparing a crème brûlée. A salad is most optimally completed in parallel – each individual component can be chopped and prepared separately and then combined at the end, and thus the best possible group organization is a complete division of labor into roles. In contrast, a crème brûlée can only be completed sequentially – the cream must be heated into custard, then transferred to an oven water-bath, then cooled, then finally torched. Dividing into roles makes no sense for such a task, as most members will spend the majority of their time waiting.

To avoid biasing participants towards or away from role-taking, we identify omelets as a viable middle-ground. Cooking an omelet involves 2 main sequentially-ordered steps: (1) chopping or whisking the ingredients before cooking, and (2) cooking the ingredients together to produce the omelet. Within these sequentially ordered steps, individual subtasks are able to be completed in parallel (eg: members may each chop different vegetables in step 1, or saute the eggs and vegetables separately in step 2). We therefore task participants with cooking omelets together to best identify when role-taking might emerge independently of task parallelizability.

Participants

A total of 270 participants were recruited across 90 experimental trials. Participants were undergraduate students from a large public university and received class credit for their participation. Participants were excluded if they had any significant structured cooking experience that could bias actions in a group cooking environment.

Design

The study used a between-subjects design with 1 factor (number of people in group) and 3 levels (two, three, or four people). 30 experimental trials were run for each condition (90 total). This sample size was determined through a basic power analysis (Faul et al., 2007). All experimental trials were conducted within the same kitchen (Fig 1) on a university campus.

Figure 1: Top: Kitchen used for all trials. Bottom: Close-up of materials provided on the left table

Procedure

Equipment and ingredients were set up prior to the participants' arrival, and always arranged as seen in Fig 1. To avoid biasing participants towards a particular organization, materials were stacked on a side table, requiring participants to choose how to organize the space. Additionally, materials were provided in multiples of 4 to ensure participants could overlap on the same task if desired, allowing for freedom in choosing what tasks to do and whether to specialize. Only two burners were available due to safety regulations; However, a post-hoc analysis found no significant difference in the degree to which burner-related tasks were distributed among participants relative to other tasks, suggesting this limitation had minimal impact on task allocation behavior.

Participants were informed that they would be cooking omelets alongside each other and that the trial would only end once they had collectively produced either two, three, or four omelets (depending on group size). Each participant was provided with a recipe to guide them and informed that they were required to use all specified ingredients. The experimenter would then note that the participants were free to organize however they saw fit, but they would ultimately be scored as a group at the end of the experiment based on the overall quality and time taken to complete all omelets. Video/Audio recordings began as soon as participants entered the kitchen, but the official timer would not begin until participants interacted with the materials. This allowed participants to introduce themselves and discuss the task before beginning the timer. Once participants completed all omelets, they were directed to fill out an online exit survey, which assessed sentiment towards other group members' contributions.

Data Analysis

Role-Taking

To analyze participant behavior, trial videos were manually annotated by 3 separate coders to identify the subtasks each participant engaged in throughout the entire duration of the trial. Start and end times were annotated to establish the overall duration of engagement with the task. Annotated tasks include the basic subtasks required to complete the omelet, as well as cleaning the environment and washing dishes before the omelets were done. Video coders had a reliability rating (Cohen's Kappa) of 0.9132, indicating high agreement between raters (Hallgren, 2012).

Analysis We use 3 separate measures that address different elements of role-taking in a group task – (1) Network Modularity, (2) Individual Entropy, and (3) Group Task Entropy. Network modularity assesses the degree to which a group's overall structure can be divided into distinct communities, reflecting a high-level division of labor into separate roles. Individual entropy assesses how an individual participant distributes their effort across tasks, reflecting whether participants chose to specialize in a small subset of tasks or generalize across many tasks. Finally, group task entropy assesses how each task was distributed amongst the group, reflecting whether a given task was mostly monopolized by a single agent or evenly distributed across all group members. These together offer a high-level structural view of role-taking as well as an individual-level perspective of specialization. We describe how we calculate each below. First, we compute role weights for each participant by taking the time the participant (p_i) spent on a given task and dividing it by the total time the group spent on that task.

$$\text{Role Weights}_{p_i} = \left[\frac{(\text{Time spent on task t})_{p_i}}{\Sigma_{j=1}^{n} (\text{Time spent on task t})_{p_j}} \; \forall \, \text{t} \in \text{T} \right]$$

Measure (1) (modularity) uses a strategy adopted from network theory known as modularity optimization, which identifies community structures within networks (Calderer & Kuijjer, 2021; Liu & Murata, 2010) and has been effective in finding relationships within biological networks (Monteiro et al., 2025). For each trial, we construct a bipartite network with two sets of nodes – Participants and Tasks. Connections between nodes are weighted by the participant's role weight for that task. We then use the modularity optimization algorithm presented by Beckett (2016), which is made specifically for weighted bipartite networks, to find the optimized modularity score of each group. Scores closer to 1 indicate that participants are easily partitioned into distinct communities of tasks. This rewards both complete specialization (all participants do distinct tasks) as well as modular specialization (eg: two participants take on the same role while a third does a distinct role). Scores closer to 0 indicate no clear partitioning and thus no division of labor.

The second and third measures both use Shannon entropy (Anand & Bianconi, 2009), which assesses the degree to which a distribution is highly concentrated or uniform. To calculate individual entropy (2), we compute the entropy of each participant's role weight vector described above. This is similar to the Role Stability measure in Banerjee and Gray (2023), using entropy rather than standard deviation. A lower entropy indicates that a participant specialized on a subset of tasks rather than spreading effort across many tasks.

To calculate group task entropy (3), we make a distribution (per trial) for each task, where a task's distribution consists of each participant's role weight for that task. We then compute the entropy of each distribution and average across all tasks in the trial. Here, a lower entropy indicates that a task was completed mostly by one participant rather than evenly spread across the whole group.

To make these measures comparable across group sizes, we normalize each metric. For group task entropy (3), we adopt a standard normalization strategy of dividing by log(N), representing the entropy as a proportion of the maximum possible entropy for that system. The other 2 measures can also structurally differ across group sizes due to joint constraints on participants; however, there are not similarly standardized methods of normalization for these. Instead, we simulate 1000 random trial runs (role weight distributions) for each group size. We then compute the expected distributions of modularity and entropy for trials of each size, and use those distributions to compute Z-scores of our observed measures. These measures therefore identify to what degree a group or individual specializes relative to what would be expected from a random organization of the same group size, making results directly comparable across conditions.

Solidarity

Analysis To analyze solidarity, we generate a rating of 'other-value' from participant responses to the exit survey questions. Other-value was computed by taking the group's average responses to the following questions: "How much did you rely on the other participants in the kitchen to complete the cooking task?", "How much do you feel other participants hindered or helped you while completing the cooking task?", and "How important do you feel the other participants were in helping complete the cooking task?".

Responses were on a scale of -2 to 2, with negative values indicating a negative valence and vice versa. To analyze the results, we fit a linear mixed effects model to predict 'other-value' from group size and modularity. To control for non-independence between participants in the same group, trial group is included as a random effect.

Results

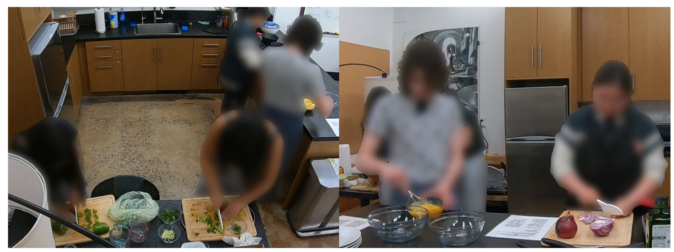

Figure 2: Example of a clear division of labor in a group of 4. Each participant completed separate subtasks in parallel.

Role-Taking

To test our hypothesis that groups of size 3+ should differ significantly from dyads, and less so from each other, we use planned contrasts with a Holm-Bonferroni correction applied (Holm, 1979). Contrasts are defined prior to data collection, specifying which condition/group comparisons to test based on our hypotheses. This controls for Type I and Type II errors more effectively than standard post-hoc comparisons such as ANOVAs or suites of t-tests (Garofalo et al., 2022), and provides greater statistical power in situations where multiple comparisons must be made (Thompson, 1990). To test our hypothesis, we run three contrasts: Groups of size 2 vs. 3 and 4, groups of size 2 vs. 3, and groups of size 3 vs. 4.

Groups of size 3 and 4 (together) had significantly higher modularity z-scores than groups of size 2: $c_0 = 7.14$, 95% CI = $[6.40, 7.88]$, $p < 0.001$ (Cohen's d = 1.96). This indicates a significant increase in role-taking when moving beyond dyads, supporting our hypothesis that groups of size 3+ exert a greater pressure to take roles than groups of size 2.

Testing groups individually confirms these results. Groups of size 3 had significantly higher modularity z-scores than groups of size 2: $c_0 = 9.34$, 95% CI = $[8.30, 10.38]$, $p < 0.001$ (Cohen's d = 1.59). Groups of size 4 also had significantly higher modularity z-scores than groups of size 3, albeit with a weaker effect size: $c_0 = 4.41$, 95% CI = $[2.93, 5.88]$, $p < 0.001$ (Cohen's d = 0.80). These

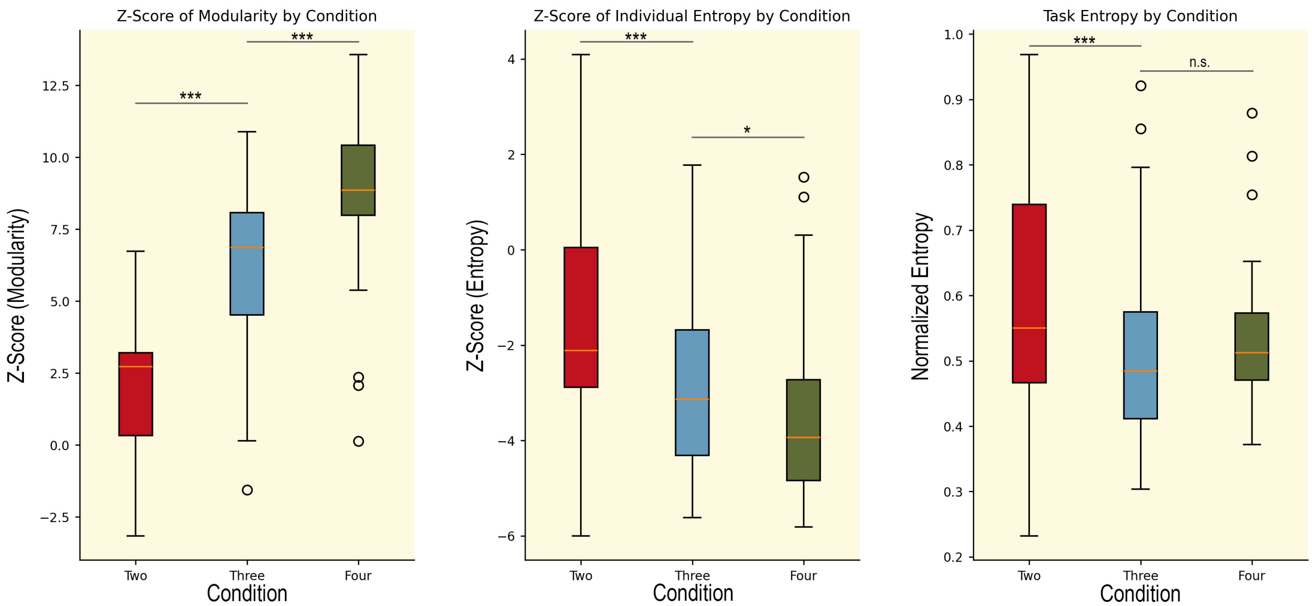

Figure 3: Left: Z-Scored Modularity, measuring overall group division of labor. Middle: Z-Scored Individual Entropy, measuring individual specialization in a subset of tasks. Right: Normalized Task Entropy, measuring how much each task was monopolized or spread out among individuals.

together indicate that role-taking increased substantially between groups of size 2 and 3, with a weaker but significant increase between groups of size 3 and 4.

Entropy Analysis Normalized Shannon entropy scores for each individual participant's role weight distribution strongly support these results. Groups of size 3 and 4 had significantly lower individual entropy z-scores than groups of size 2: $c_0 = -3.03$, 95% CI = $[-3.57, -2.50]$, $p < 0.001$ (Cohen's d = −0.85). This indicates a significant decrease in individual entropy (indicating more specialization in participants) in groups of size 3 and 4 compared with groups of size 2, aligning with our hypotheses.

Testing groups individually again confirms these results. Groups of size 3 had significantly lower individual entropy z-scores than groups of size 2: $c_0 = -3.76$, 95% CI = $[-4.51, -3.00]$, $p < 0.001$ (Cohen's d = −0.68). Groups of size 4 also had significantly lower individual entropy z-scores than groups of size 3, again with a much weaker effect: $c_0 = -1.45$, 95% CI = $[-2.52, -0.38]$, $p \approx 0.026$ (Cohen's d = −0.34). These together indicate that individuals specialized significantly more often in larger groups, with the largest difference occurring between groups of size 2 and 3.

Normalized Shannon entropy scores for overall task distributions help clarify the results from modularity and individual-level specialization. Groups of size 3 and 4 had significantly lower overall task entropy scores than groups of size 2: $c_0 = -0.54$, 95% CI = $[-0.59, -0.50]$, $p < 0.001$ (Cohen's d = −0.44). This indicates a significant decrease in task entropy (indicating tasks were more often completed by a single individual) in groups of size 3 and 4 compared with groups of size 2, aligning with our hypotheses.

Individual testing finds that groups of size 3 had significantly lower task entropy scores than groups of size 2: $c_0 = -0.50$, 95% CI = $[-0.56, -0.44]$, $p < 0.001$ (Cohen's d = −0.45). However, we find that groups of size 3 and size 4 do not differ significantly: $c_0 = -0.08$, 95% CI = $[-0.17, 0.00]$, $p \approx 0.162$. These results demonstrate that, while both group structure and individual-level specialization moved towards role-taking in larger groups, the proportion of members contributing to any specific task did not decrease between groups of size 3 and 4. This suggests that the resultant division of labor was not necessarily optimized for parallelizability.

This is supported by a post-hoc test examining how role-taking impacted overall task efficiency. Fitting a linear model to predict the overall duration (in seconds) of the trial from modularity and condition finds that while group size significantly increased overall duration (size 3: $\beta = 403.18$, $p \approx 0.022$, size 4: $\beta = 623.44$, $p \approx 0.003$), modularity is not a significant predictor of trial duration ($\beta = -25.57$, $p \approx 0.230$). Taking on roles resulted in an insignificant benefit to overall task efficiency, supporting the idea that roles were formed to ease cooperation but not necessarily to directly improve productivity.

Collectively, these tests verify our initial hypothesis – role-taking occurs significantly more often in larger groups, with the largest increase appearing between groups of size 2 and size 3. Furthermore, formed roles did not provide any benefit towards task efficiency. These results overall align with our prediction that role-taking would increase in response to greater coordinative complexity, supporting the view that role-taking functions as a solution to manage such complexity.

Solidarity

To assess whether role-taking predicted perceived value of other group members while accounting for group size and non-independence within groups, we fit a linear mixed-effects model. We predict individual other-value scores from group-level modularity and group size (treated as a numeric predictor), including random intercepts for each group.

Modularity (Z-scored) was a significant positive predictor of other-value ($\beta = 0.15$, $p < 0.001$), indicating that groups exhibiting greater role specialization also reported higher perceived value of collaborators. In contrast, group size was negatively associated with other-value ($\beta = -0.38$, $p < 0.001$), independent of modularity. These results suggest that role-taking is positively related to perceived value within the group, even as increasing group size may reduce solidarity.

Discussion

Our results provide important empirical evidence to support recent claims that humans can leverage roles to simplify coordinative inferences in cooperation (Banerjee & Gray, 2023; Jara-Ettinger & Dunham, 2025; Ritchie, 2020). We extend these claims to show that coordinative difficulty itself may be a direct cause of role-taking in humans, demonstrating that role-taking increases significantly as a result of increasing group size.

Larger groups consistently adopted a group-wide division of labor (measured via network modularity) and individually specialized to a far greater extent (measured via participant-level entropy). However, even in groups that split labor, tasks were still often shared between a subset of participants rather than totally monopolized. Combining this with the finding that groups that took on roles were no faster than groups that did not, the results provide substantial evidence that roles were adopted to support cooperation independently of any benefits to overall task efficiency.

Though role-taking increases significantly with each additional group member, we consistently identify the strongest increase occurring between groups of size 2 and 3. While increasing group sizes consistently increases the difficulty of coordination (Kolfschoten et al., 2014; Skau, 2022), groups of size 3 specifically introduce new inference problems not present in dyads, such as determining the intended recipients and relevance of coordinative signals (Clark, 1992; Jovanovic, 2007). Therefore, role-taking appears to increase proportionally to increases in coordinative difficulty, supporting the claim that humans use role-taking to manage complexity.

This claim is further supported by the finding that groups that take on roles also consistently rate their partners as being more important and valuable to the task. Since humans are known to prefer more cooperative partners (Barclay & Willer, 2006), this finding is consistent with the view that role-taking is a socially cooperative act. At the same time, the findings on group solidarity are not necessarily incompatible with organizational efficiency-based explanations either – humans would likely also appreciate partners that are more efficient collaborators, so further work should aim to distinguish between social and efficiency-based explanations. However, task efficiency is unlikely to explain the results found here, as we identified no benefit to overall speed from role-taking.

These findings help resolve a standing discrepancy between organizational efficiency-based accounts of role-taking and human tendencies to commit to roles even when they become suboptimal (Chu & Schulz, 2022; Zhai et al., 2024), suggesting that humans use roles as a social tool to support their partners' ability to coordinate and plan around them. We argue that humans take on roles as a way to publicly constrain their permissible actions and make their intentions more easily inferable, and do so specifically when there is greater difficulty in aligning plans between members. Given that roles constitute a critical component of human cooperation (Jara-Ettinger & Dunham, 2025), the findings here provide important empirical evidence to help explain why and when humans take on roles in group tasks.

It is important to note some potential limitations of this study. For one, cooking tasks may bias participants towards role-taking due to associations with popular cooking shows or knowledge about commercial kitchens, and it is possible that other tasks may elicit different strategies. Second, we manipulate coordinative difficulty by varying group size, but there are a number of other factors that are relevant for coordination not examined here. Task difficulty, differences in task expertise, barriers to communication, and increased time pressure (to name a few examples) could all contribute towards coordinative complexity and may produce different strategies than those produced by changing group size. In addition, increasing group size may also introduce more salient social pressures that could bias participants away from individualistic strategies. More coordinators also limit physical space, which could favor station-based strategies due to difficulties moving around. These could contribute towards role-taking independently of an explicit increase to coordinative difficulty, warranting further investigation to understand how group size and other factors may interact to affect resultant group organization strategies.

More broadly, this work highlights a critical need for more multi-agent (3+) research in Cognitive Science. We have demonstrated that human cooperative strategy – namely, the decision to take on roles – can vary greatly with group size. Humans are highly social creatures, and a large amount of day-to-day interaction occurs in groups larger than 2. These social situations introduce new sources of ambiguity, new demands on social inference, and likely amplify social pressures to a greater extent than is possible to observe in simple dyadic interaction. Moving beyond the dyad will therefore be critical for cognitive scientists to develop a more complete and ecologically valid understanding of human cognition.

References

Allen, K., Jara-Ettinger, J., Gerstenberg, T., Kleiman-Weiner, M., & Tenenbaum, J. B. (2015). Go fishing! responsibility judgments when cooperation breaks down. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*.

Anand, K., & Bianconi, G. (2009). Entropy measures for networks: Toward an information theory of complex topologies. *Physical Review E*, *80*.

Baker, C., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. (2017). Rational quantitative attribution of beliefs, desires, and percepts in human mentalizing. *Nature Human Behavior*, *1*.

Banerjee, S., & Gray, W. D. (2023). Role stability and team performance in a 4-player cooperative cooking game. *Virtual MathPsych/ICCM*.

Barclay, P., & Willer, R. (2006). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*.

Beckett, S. J. (2016). Improved community detection in weighted bipartite networks. *Royal Society Open Science*. https://doi.org/https://doi.org/10.1098/rsos.140536

Boesch, C. (2002). Cooperative hunting roles among Taï Chimpanzees. *Human Nature*, *13*(1), 27–46.

Boyce, V., Hawkins, R., Goodman, N., & Frank, M. (2022). Two's company but six is a crowd: Emergence of conventions in multiparty communication games. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*.

Bratman, M. (2014). *Shared agency: A planning theory of acting together*. Oxford University Press.

Calderer, G., & Kuijjer, M. L. (2021). Community Detection in Large-Scale Bipartite Biological Networks. *Frontiers in Genetics*. https://doi.org/10.3389/fgene.2021.649440

Chater, N., Zeitoun, H., & Melkonyan, T. (2022). The paradox of social interaction: Shared intentionality, we-reasoning and virtual bargaining. *Psychological Review*, *129*(3), 415–437.

Chu, J., & Schulz, L. (2022). "because i want to": Valuing goals for their own sake. *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*.

Chu, J., Tenenbaum, J. B., & Schulz, L. E. (2024). In praise of folly: Flexible goals and human cognition. *Trends in Cognitive Science*.

Clark, H. H. (1992). *Arenas of language use*. University of Chicago Press.

Clark, H. H., & Carlson, T. B. (1982). Hearers and speech acts. *Language*.

Davis, I., & Jara-Ettinger, J. (2022). Hierarchical task knowledge constrains and simplifies action understanding. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Dreyer, T., Haluts, A., Korman, A., Gov, N., Fonio, E., & Feinerman, O. (2024). Comparing cooperative geometric puzzle solving in ants versus humans. *Proceedings of the National Academy of Sciences*, *122*(1).

Durkheim, E. (1893). *The division of labour in society*.

Eickers, G. (2024). Scripts and social cognition. *Ergo an Open Access Journal of Philosophy*.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods*, *39*(2), 175–191.

Garofalo, S., Giovagnoli, S., Orsoni, M., Starita, F., & Benassi, M. (2022). Interaction effect: Are you doing the right thing? *Plos One*.

Gilbert, M. (2013). *Joint Commitment: How we Make the Social World*. Oxford University Press.

Gilbert, M. (2016). Joint Commitment: What it is and why it matters. *Phenomenology and Mind*, *9*, 18–26.

Goffman, E. (1976). Replies and responses. *Language in Society*.

Goldstone, R. L., Andrade-Lotero, E. J., Hawkins, R. D., & Roberts, M. E. (2024). The Emergence of Specialized Roles Within Groups. *Topics in Cognitive Science*, *16*(2), 257–281.

Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: The Belknap Press of Harvard Univ. Press.

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in quantitative methods for psychology*, *8*(1), 23–24.

Helmreich, R. L., & Schaefer, H.-G. (1994). Team performance in the operating room. In *Human error in medicine* (pp. 225–253). Lawrence Erlbaum Associates, Inc.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*.

Hutchins, E. (1996). *Cognition in the Wild*. MIT Press.

Jankovic, M. (2018). Collective intentionality and language. In *The routledge handbook of collective intentionality*. Routledge.

Jara-Ettinger, J., & Dunham, Y. (2025). The institutional stance. *Behavioral and Brain Sciences*.

Jovanovic, N. (2007). *To whom it may concern – addressee identification in face-to-face meetings* [Doctoral Dissertation]. University of Twente.

Kolfschoten, G., French, S., & Brazier, F. (2014). A discussion of the cognitive load in collaborative problem-solving: The decision-making phase. *EURO Journal on Decision Processes*.

Liu, X., & Murata, T. (2010). An Efficient Algorithm for Optimizing Bipartite Modularity in Bipartite Networks. *JACIII*, *14*(4), 408–415.

Mieczkowski, E., Mon-Williams, R., Bramley, N. R., Lucas, C. G., Vélez, N. A., & Griffiths, T. (2025). A normative account of specialization: How task and environment shape role differentiation in collaboration. *Proceedings of the 47th Annual Meeting of the Cognitive Science Society*.

Monteiro, B. L., Souza, C. S., Maruyama, P. K., Camargo, M. G. G., & Morellato, L. P. C. (2025). Applying plant-pollinator network to identify priority species for conservation in a biodiversity hotspot. *Biological Conservation*,

302. https://doi.org/https://doi.org/10.1016/j.biocon.2025.110979
Nadel, S. F. (1957). *The Theory of Social Structure*. Cohen & West, Ltd.
O'Brien, G. (1968). The measurement of cooperation. *Organizational Behavior and Human Performance*.
Rakoczy, H., Gräfenhain, M., Clüver, A., Shulze Dalhoff, A., & Sternkopf, A. (2014). Young children's agent-neutral representations of action roles. *Journal of Experimental Child Psychology*.
Rakoczy, H. (2018). Development of collective intentionality. In *The routledge handbook of collective intentionality*. Routledge.
Ritchie, K. (2020). Minimal cooperation and shared agency. Springer.
Schegloff, E. (2007). *Sequence Organization in Interaction* (Vol. 1). Cambridge University Press.
Searle, J. R. (1990). Collective Intentions and Actions. In *Intentions in Communication*. MIT Press.
Skau, S. (2022). Analysis of the cognitive processes involved in creating and sustaining cooperative group activity. *Frontiers in Psychology*.
Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Basil Blackwell.
Thompson, B. (1990). Planned versus unplanned and orthogonal versus nonorthogonal contrasts: The neo-classical perspective.
Tomasello, M. (2019). The Role of Roles in Uniquely Human Cognition and Sociality. *Journal for The Theory of Social Behaviour*, *50*(1), 2–19.
Turner, R. (1962). Role-Taking: Process Versus Conformity. In *Human Behaviour and Social Processes*. Routledge; Kegan Paul.
Vesper, C., Schmitz, L., Safra, L., Sebanz, N., & Knoblich, G. (2016). The role of shared visual information for joint action coordination. *Cognition*.
Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*.
Wu, S., Wang, R., Evans, J., Tenenbaum, J., Parkes, D., & Kleiman-Weiner, M. (2021). Too Many Cooks: Bayesian Inference for Coordinating Multi-Agent Collaboration. *Topics in Cognitive Science*.
Zhai, S., Cheng, S., Moskowitz, N., Shen, M., & Gao, T. (2024). The development of commitment: Attention for intention. *Child Development*.