Title
Many Hands Don't Always Make Light Work: Explaining Social Loafing via Multiprocessing Efficiency

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Mieczkowski, Elizabeth
Turner, Cameron Rouse
Vélez, Natalia A
et al.

Publication Date
2024

Peer reviewed

Many Hands Don't Always Make Light Work: Explaining Social Loafing via Multiprocessing Efficiency

Elizabeth Mieczkowski[1], **Cameron Turner**[1,2], **Natalia Vélez**[2], **Thomas L. Griffiths**[1,2]
[1]Department of Computer Science, Princeton University
[2]Department of Psychology, Princeton University
{emiecz, c.rouse.turner, nvelez, tomg}@princeton.edu

Abstract

Humans collaborate to improve productivity and collective outcomes, but people do not always exert maximal effort towards accomplishing collaborative goals. Instead, individuals often expend less effort in groups, a phenomenon known as social loafing that is traditionally viewed as detrimental to productivity. However, theories from distributed computer systems suggest that social loafing might be a rational response to the diminishing returns expected from division of labor when group size increases. Here, we examine how considerations of task efficiency affect the perceived acceptability of withholding effort during a collaborative task. We conducted experiments varying workload and group size across scenarios in which all group members except for one are actively contributing to a common goal. We then compare participant judgments to a model inspired by latency speed-up in distributed systems. We find that people are systematically influenced by task efficiency, in addition to social norms, when judging social loafing.

Keywords: collaboration; social loafing; distributed computation; multiprocessing; collective intelligence

Introduction

Humans frequently work together to tackle challenges and accomplish goals beyond the capabilities of a single individual. In an effort to enhance productivity, we coordinate our actions with others in a way that allows us to leverage collective effort and overcome individual constraints on time and resources (Griffiths, 2020; Vélez et al., 2023). Collaboration is thus motivated by the underlying belief that "many hands make light work"; as we pool shared resources, skills, and effort, we should benefit from an increase in our collective output. However, figuring out how to combine effort effectively across group members is not a trivial endeavor. Decades of psychological research reveal a tendency for individuals to expend less effort when working as a group (Ringelmann, 1913; Steiner, 1972; Ingham et al., 1974; Petty et al., 1977). This withholding of labor, also known as social loafing, enables group members to reap the benefits of a collaborative task without contributing their fair share of effort, and is often considered to result in a loss of group-level productivity (Latané et al., 1979).

Numerous theories have aimed to address the causes of social loafing, mainly focusing on social and psychological phenomena such as arousal reduction, evaluation potential, visibility, and dispensability and matching of effort (Karau & Williams, 1993). However, these explanations often overlook a crucial point: tasks vary in the degree to which they should theoretically benefit from collaboration. One factor that can impact collaborative efficiency is workload: the amount of work to be completed in a given time span. Adding group members may not improve performance in situations where the workload is too small. Secondly, tasks may contain serial dependencies that prevent effective division of labor. These bottlenecks, whether due to spatial constraints, limited physical resources like tools and workstations, or cognitive barriers hindering communication and information flow, can restrict the contributions of additional group members, even when there is more work to be done.

In order to properly evaluate the complexities that emerge when relating individual effort to group productivity, we need a normative theory that formalizes how people should distribute resources and labor during collaborative tasks. There are parallels between these challenges and ones that have already been solved in the field of distributed computer systems. Classic findings in distributed systems provide a formal way of analyzing system-level efficiency in terms of the number of machines that can be added to a task before expected speed-up improvements plateau (Amdahl, 1967; Gustafson, 1988; Hill & Marty, 2008; Cassidy & Andreou, 2011). We propose that this branch of computer science can provide a theoretical framework for analyzing analogous problems faced by groups of people as they transition from working individually to collaborating (Vélez et al., 2023). In particular, these theories suggest that what might seem like a lapse in human group dynamics could, in fact, be a rational collaborative strategy. Individuals might refrain from exerting effort when they perceive that their involvement will not significantly enhance overall task efficiency. While the idea that individuals consider effort when deciding to contribute to a task is intuitive, a formal model can provide a principled way of identifying optimal labor allocations and whether people are sensitive to varying levels of efficiency.

Principles from computer systems motivate the prediction that considerations of group-level efficiency may contribute to expectations about social loafing in addition to social norms. For example, suppose all of the guests at a dinner party have just finished eating, and it is now time to clean the dishes. If there is only one sink, and several other guests have already crowded around it, should an additional person try to contribute, or is it better that they remain seated at the table? Previous work on social norms suggests that the guest at the table may stand up and help with the dishes once they see other guests doing so (Hackman & Morris, 1975; Schmidt & Tomasello, 2012; Turner et al., 2014; Roberts et al., 2017). As the number of people contributing to the clean-up increases,

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

there will be heightened normative pressure on those remaining seated to participate in order to avoid resentment and being perceived as uncooperative in the eyes of the productive group members (Mas & Moretti, 2009). However, what is presently unclear is how normative pressure interacts with group members' assessments of efficiency. The guest at the table may choose to remain seated, even though other guests are working, due to an implicit awareness of environmental constraints that hinder their ability to improve task performance. For example, if there are a limited number of plates to be cleaned, only so many guests can contribute to the task before there is nothing left to do. Similarly, if there are many dirty plates but only one sink, adding another helper will not reduce overall washing time due to a physical resource bottleneck.

Here, we assess whether people consider efficiency when deciding if an additional group member should contribute to a collaborative task, and we apply a novel modeling framework inspired by distributed computer systems to the study of human collaboration. We examine how assessments of task efficiency influence social loafing judgments in an experimental paradigm in which all group members except for one are actively engaged in completing a task. We find that both workload and group size significantly affect people's judgments of the acceptability of social loafing. People consider social loafing to be less acceptable for high workloads and small groups, and more acceptable for low workloads and larger groups. In addition, using a model inspired by Amdahl's Law in distributed systems, we find qualitative similarities between participants' judgments and the predicted speed-up that would have been gained if the social loafing agent contributed to the task. Broadly, these results demonstrate that people take into account distributed task efficiency when judging the efficacy and scope of collaboration.

Background

Collaboration and Group Performance

Humans are motivated to work with one another as a way to leverage collective resources, yet collaboration itself poses problems such as how to divide labor, coordinate actions, and share collective output amongst group members (Vélez et al., 2023). Previous work shows that individuals tend to collaborate more with one another in larger groups, and as group size grows, the benefits of coordination outweigh productivity losses due to team members reducing their individual effort levels (Mao et al., 2016). Another factor that may impact group performance is task complexity. Previous work has shown that the number of components and inter-dependencies in a task moderate the benefits of collaboration. In particular, collaboration improves the efficiency and quality of solutions for complex tasks, but not simpler ones (Almaatouq et al., 2021). Related work from cognitive psychology examines how people may hold social loafers responsible for the consequences of their actions. These responsibility judgments are influenced by factors such as the causal impact of the target's actions, how easily their contribution could have been replaced, and whether the outcome would have changed if the target had acted differently (Gerstenberg et al., 2018; Xiang et al., 2023; Wu & Gerstenberg, 2024).

Social Loafing and Norms

Previous work on social loafing and norms motivate the prediction that group size impacts collaborative effort. Research in social psychology indicates that individuals frequently exert less effort when working together. This phenomenon was first demonstrated by Ringelmann (1913), who found that when a group of people collectively pulled on a rope, the result was lower than each individual group member's output (Kravitz & Martin, 1986). This finding has been replicated across a wide range of populations in both physical and cognitive tasks (Ingham et al., 1974; Petty et al., 1977). Many theoretical accounts aim to explain these productivity losses in terms of social mechanisms such as social impact, contribution visibility, and dispensability of effort (Karau & Williams, 1993; Simms & Nichols, 2014). Latané (1981) theorized that when an experimenter makes a suggestion to a group, the request will be divided amongst each member, resulting in a reduction of individual effort. Kerr and Bruun (1983) found that group members exert less effort because they perceive their contribution to be more dispensable to the overall outcome. Jackson and Harkins (1985) proposed that individuals diminish their contributions to avoid an unequal division of labor when they expect social loafing from the other group members. Williams (1981) found that reduced identifiability in group settings led individuals to feel less motivated to contribute to a common goal. Group size additionally impacts collaborative behavior by placing greater normative pressure on each individual to behave similarly to the collective. When judging whether or not certain behavior is acceptable, people tend to take into account both descriptive norms (e.g., the average, or what everyone else is doing), as well as prescriptive norms (e.g., the ideal, or what everyone should be doing) (Bear & Knobe, 2017).

Distributed Systems

Parallelism enables complex computations to be carried out simultaneously, leading to immense improvements in computational efficiency. However, coordinating multiple processors to tackle the same computations in parallel faces tremendous challenges (Baer, 1973; Almasi & Gottlieb, 1994; Kshemkalyani & Singhal, 2011). This literature focuses on the potential disadvantages that systems face when transitioning from one machine to distributed processing, as well as the best strategies for improving efficiency as tasks are allocated across processors. Amdahl's Law is a formula that predicts the theoretical speed-up expected when a computation is executed across multiple processors in terms of the proportion of the process that can be parallelized (Amdahl, 1967; Hill & Marty, 2008; Cassidy & Andreou, 2011; Hennessy & Patterson, 2011). In the simplest case, there are only two levels of parallelism – a completely parallelizable proportion of the task (f) that can be divided between all processors to achieve a speedup of s, versus a completely serial proportion that must be performed by one processor ($1-f$). We will consider a subtask to correspond to a specific proportion of an overall task. Amdahl's Law computes the predicted temporal speed-up for a fixed workload via

$$S_{time}(s) = \frac{1}{(1-f)+\frac{f}{s}} \tag{1}$$

As the number of processors increases, improvements in efficiency are limited by the proportion of the task that must be performed serially by a single processor due to time, resource, or memory dependencies. Amdahl's Law points to an inherent limit to which parallelism can enhance task efficiency; it is only beneficial to allocate available processors to a computation when there is unfinished work and additional processor involvement will reduce the time to completion.

Methods

Evaluating the Impact of Task Efficiency

We develop a paradigm for evaluating how people take into account efficiency when deciding if an additional group member should contribute to a collaborative task. In particular, we predict that if the decision to contribute to a collaborative task is influenced by task efficiency, then participants should consider both workload and group size when evaluating the acceptability of social loafing. Alternatively, if the decision to contribute to a collaborative task depends solely on social norms, then participants should only take group size into account when evaluating social loafing, regardless of the potential marginal impact of their contribution.

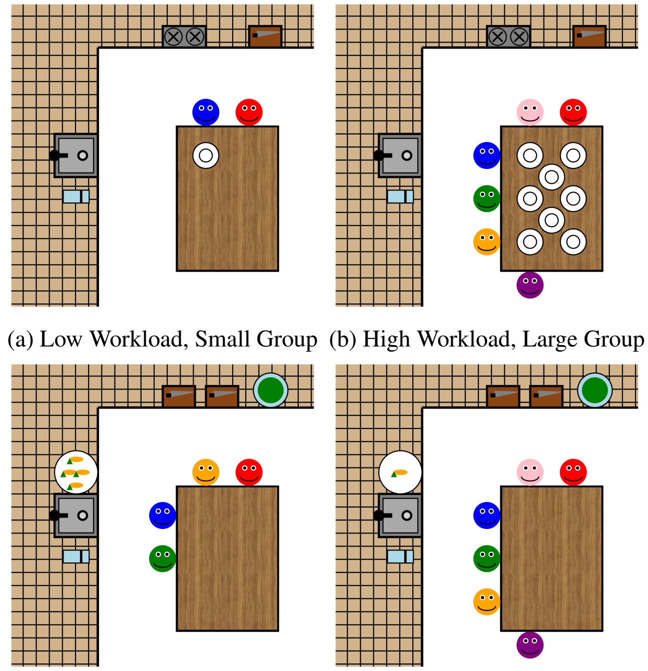
Participants were presented with a series of dinner party scenarios. In each scene, all agents except for one 'loafer' were actively engaged in a task. Participants were asked to evaluate the acceptability of this agent's social loafing behavior. All of these methods were preregistered prior to data collection at `https://aspredicted.org/8bj9q.pdf`.

Participants

We recruited 300 English-speaking adults from the Prolific platform in exchange for compensation ($2.40 for a 10-12 minute experiment). Fourteen participants who failed an attention check were excluded from subsequent analysis, resulting in a total of 286 participants (87 male, 196 female; mean age = 38.8, SD = 13.1). The experiment was performed with IRB approval (IRB #15959); all participants provided informed consent prior to the experiment.

Design and Stimuli

We designed stimuli that varied across two independent dimensions: (1) workload, or the amount of work to be completed, and (2) group size, or the number of agents present in the scene. To ensure robustness, each condition was presented in two distinct kitchen scenarios. In Dish Clean-Up, agents were seated at the table with a number of plates in front of them. Participants were told that the agents needed to carry the plates to the sink, then wash and dry them. In Dish Clean-Up, workload was defined as the number of plates that need to be cleaned. In Salad Preparation, agents were also seated at the table with a number of carrots next to the sink. Participants were told that the agents needed to clean and chop the carrots to add them to the salad bowl. In Salad Preparation, workload was defined as the number of carrots that need to be prepared.



(a) Low Workload, Small Group (b) High Workload, Large Group

(c) Medium Workload and Group (d) Low Workload, Large Group

Figure 1: Example stimuli. Images (a) and (b) are sampled from Dish Clean-Up, and (c) and (d) from Salad Preparation.

In each stimulus, every agent was depicted with a unique color. The red-colored agent was present in every scene in a fixed position, and this agent was always the loafer, or the guest that remained seated at the table without contributing to the task. Stimuli for this study were generated using Python and the Matplotlib library. We represented a kitchen using a 7×7 grid, and programmatically populated all of the various layouts, objects, and agents in the scene. This approach allowed for a high degree of consistency across scenes.

We manipulated each scenario to reflect low (1 plate/carrot), medium (4 plates/carrots), or high (8 plates/carrots) workload, as well as small (2 agents), medium (4 agents), or large (6 agents) group size. This approach resulted in a total of 18 unique stimuli, or 3 workload levels × 3 group sizes × 2 kitchen scenarios, in a fully within-participants design. Participants were presented the 18 stimuli in random order. Representative examples of our stimuli are shown in Figure 1.

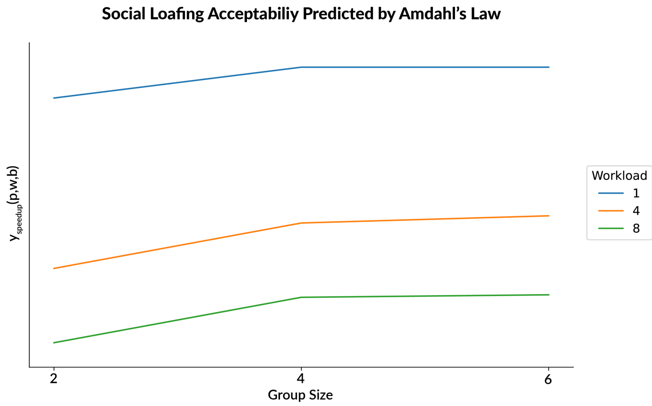


Figure 2: Predicted social loafing acceptability for varying workloads w and group sizes p based on $y_{speedup}(p,w,b)$, averaged over Dish Clean-Up and Salad Preparation. Results are presented on a log scale.

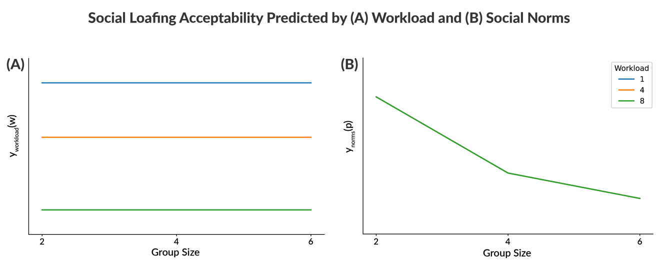


Figure 3: Predicted social loafing acceptability based on (A) workload $y_{workload}(w)$ and (B) social norms, $y_{norms}(p)$.

Procedure

To ensure data quality, we included an attention check at the beginning of the experiment, requiring participants to accurately identify the number of agents and plates depicted in a simple example stimulus. Participants were informed that every trial stimulus represented a different dinner party with new guests. In each trial, participants were told that the red agent remained seated at the table for the duration of the task. They were then asked to rate how acceptable they considered this guest's behavior. Participants reported their judgments on a Likert scale from 1 (completely unacceptable) to 10 (completely acceptable).

Distributed Systems Model

We formalized a predictive model for social loafing acceptability judgments based on a variant of expected temporal speed-up proposed by Amdahl's Law (Amdahl, 1967; Hill & Marty, 2008; Cassidy & Andreou, 2011; Hennessy & Patterson, 2011). Our model depends on a combination of (1) workload w, (2) group size p, (3) the number of subtasks n, and (4) the presence of environmental bottlenecks b_i that limit the number of agents who can successfully share work within a subtask i. Our model considers not only the difference in dividing up the overall workload between one fewer agent, but also the distinct levels of parallelizability across different subtasks due to different bottlenecks. The predicted speed-up that could be achieved if one more agent contributed to the task is given by

$$S_{time}(p,b) = \frac{1}{(1-\sum_{i=1}^{n} f(i)) + \sum_{i=1}^{n} \frac{f(i)}{s(p,b_i)}} \tag{2}$$

Here, n is the number of parallelizable subtasks that can be divided amongst agents in each scene. We informed participants that the overall scenarios were composed of different subtasks. We assume that each subtask contributes equally to the overall task, i.e. that the proportion $f(i)$ is a uniform function of the number of subtasks. For every scene, $0 \leq \sum_{i=1}^{n} f(i) \leq 1$. $\sum_{i=1}^{n} f(i) = 1$ if every subtask is parallelizable to some extent. Therefore, $1 - \sum_{i=1}^{n} f(i)$ gives us the fraction of the task that cannot be parallelized, or split amongst group members. We assume that Dish Clean-Up and Salad Preparation are composed of two distinct subtasks: Dish Clean-Up involves bringing the plates to the sink and washing them, and Salad Preparation involves washing the carrots and chopping them on a cutting board. These assumptions are considered further in the Discussion section.

For each of the parallelizable subtasks i, $s(p,b_i)$ estimates the proportional increase in speed that can be achieved by adding the pth agent given the current number of agents $p-1$ and serial bottlenecks b_i, i.e.

$$s(p,b_i) = \frac{\min(p,b_i)}{\min(p-1,b_i)} \tag{3}$$

To unpack this function, consider a scenario in which there are five agents but only two sinks, thereby $p=5$ and $b_i=2$. In this case, adding an additional agent to the task will have no impact on task efficiency due to the limited number of workstations that can be used to wash the dishes. Alternatively, if there are five agents and six sinks, i.e. $p=5$ and $b_i=6$, then as long as there is enough work to be done, recruiting another agent to wash the dishes will improve task completion time due to the availability of workstations. In this experiment, b_i was fixed across the different conditions for each scenario. Since $s(p,b_i)=1$ when adding the pth agent will lead to no speed-up for a given subtask i, we can equivalently re-write our model in terms of the number of total (parallelizable and serial) subtasks m, or

$$S_{time}(p,b) = \frac{1}{\sum_{i=1}^{m} \frac{f(i)}{s(p,b_i)}} \tag{4}$$

Amdahl's Law is calculated given a fixed workload w. To compute the impact of a given speed-up on overall task efficiency given different workload levels, we scale speedup $S_{time}(p,b)$ by w. We consider social loafing acceptability to be inversely proportional to speed-up, or the "slow-down" that occurs when the social loafing agent does not contribute. Therefore, we predict loafing acceptability given speed-up efficiency, or $y_{speedup}(p,w,b)$ (depicted in Figure 2), to be

$$y_{speedup}(p,w,b) \propto \frac{1}{w \cdot S_{time}(p,b)} \tag{5}$$

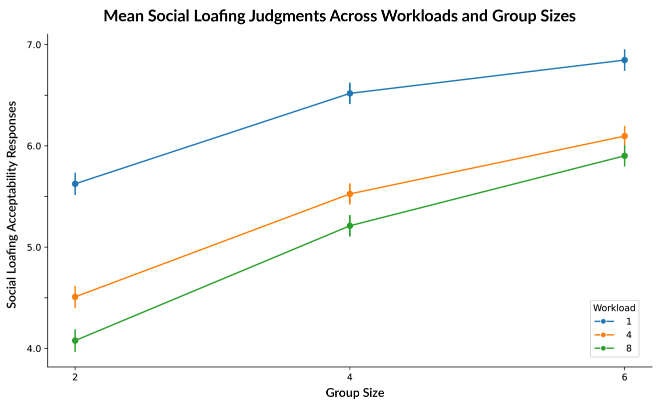


Figure 4: Mean participant responses to social loafing judgments for varying workloads and group sizes. Error bars represent standard errors.

In order to motivate the need for this predictive model, we also consider several baseline models. To begin, if perceived acceptability of withholding effort depends only on the amount of work to be completed w, then social loafing judgments should be predicted by $y_{workload}(w) \propto \frac{1}{w}$. Alternatively, if perceived acceptability of withholding effort depends solely on social norms, then a simple normative model should take into account group size but not workload. In this case, social loafing judgments given p group members will be predicted by $y_{norms}(p) \propto \frac{1}{p}$. Model predictions based on $y_{workload}(w)$ and $y_{norms}(p)$ are depicted in Figure 3.

Results

We analyzed our data using a linear mixed-effects model that predicted continuous judgments of social loafing acceptability, $y_i \in [1, 10]$, as a function of the discrete fixed effects of workload x_1 and group size x_2. To enable pairwise comparisons between factor levels, workload and group size were each coded using two binary indicator variables measured against the baseline middle level. In particular, $x_{1,1}$ denoted Low Workload, $x_{1,2}$ denoted High Workload, $x_{2,1}$ denoted Small Group Size, and $x_{2,2}$ denoted Large Group Size. We also included all interactions between workload and group size levels, a fixed intercept term $\overline{\alpha}$, random effects across participants α_i, and residual error ε. These analyses were performed using statsmodels, NumPy, and pandas in Python.

All regression coefficients for the main fixed effects of workload and group size were significant. First, examining the effects of workload, we found that low workload produced a significant increase in acceptability judgments compared to the baseline of medium workload ($\beta_{1,1} = 1.00$, SE = 0.11, 95% CI = [0.79, 1.21], $z = 9.27$, $p < 0.001$). High workload produced a significant decrease in acceptability judgments compared to medium workload ($\beta_{1,2} = -.31$, SE = 0.11, 95% CI = [-0.52, -0.10], $z = -2.89$, $p < 0.001$). Together, these results suggest that as workload increases, social loafing is judged to be less acceptable.

Next, examining the effects of group size, we found that small group size produced a significant decrease in how acceptable social loafing was considered, compared to the medium group size ($\beta_{2,1}$ = -1.01, SE = 0.11, 95% CI = [-1.22, -0.80], $z = -9.41$, $p = 0.004$). Large group size led to a significant increase in social loafing acceptability, compared to the medium group size ($\beta_{2,2}$ = 0.57, SE = 0.11, 95% CI = [0.36, 0.78], $z = 5.32$, $p < 0.001$). These results suggest that as group size increases, social loafing is also considered to be more acceptable. None of the estimates for the fixed effects of the interactions between workload and group size levels were significant: (1) low workload and small group ($\beta = .12$, SE = 0.15, 95% CI = [-0.18, 0.42], $z = 0.78$, $p = 0.439$), (2) low workload and large group (β, SE = 0.15, 95% CI = [0.11, -0.54], $z = -1.61$, $p = 0.106$), high workload and small group (β, SE = 0.15, 95% CI = [-0.42, 0.18], $z = 0.43$, $p = 0.425$), (4) high workload and large group (β, SE = 0.15, 95% CI = [-0.18, 0.41], $z = 0.76$, $p = 0.446$).

Discussion

Social loafing is typically considered to result from group members selfishly withholding effort during collaborations, thereby reducing group productivity. However, division of labor is not guaranteed to improve task completion. Research in distributed computing suggests that scaling the size of a system cannot always enhance performance in cases of insufficient workload and serial dependencies that prevent multiple processors from working on a computation simultaneously. Here, we assess whether people analogously take into account system-level efficiency when evaluating individual agents' contributions to a group task, and we present a modeling framework inspired by distributed systems to formalize factors that may influence collaborative performance.

Influence of Task Efficiency on Collaborative Effort

In order to test whether people are aware of the inherent limitations that a group member's contribution will have on a collaborative task, we measured the extent to which participants found a loafer's behavior acceptable across a highly controlled set of stimuli. We found that both workload and group size, but not their interaction, were significant predictors of social loafing acceptability. In particular, participants judged social loafing to be less acceptable for high workloads compared to lower workloads, and for small groups compared to larger groups. It is less acceptable to loaf when there is more work to be completed, and when there are fewer agents to do it. Put together, our results suggest that social loafing may not always be sub-optimal – instead, people consider an individual's impact on task efficiency when judging whether it is acceptable to withhold effort during collaboration.

Distributed Systems Model Evaluation

We also characterize a precise normative framework inspired by expected speed-up in Amdahl's Law to model how people should withhold effort from a task depending on workload, group size, and environmental constraints. Our findings reveal a correspondence between model predictions and observed behavior across different conditions. Particularly notable is the model's capacity to differentiate between low, medium, and high workloads, as well as predict the same general increase in social loafing acceptability as group size increases. Interestingly, our model also begins to capture the plateau that occurs with a medium sized group. This plateau emerges due to the limited number of cutting boards in the salad preparation scenario, as well as there only being one sink and one towel in the dish-washing scenario. As group size exceeds the number of bottlenecks that prevent parallelization in each scene, efficiency improvements reach a saturation point. That is, when there are more group members than workstations, adding an additional agent will have a diminishing impact on the time required to complete the task. This finding indicates that people are aware of the theoretical limit to collaborative gains that may exist when environmental bottlenecks prevent subtask parallelization.

Alternatively, we considered several baseline models that take into account the amount of work to be completed or normative social pressures as group size increases. None of these models qualitatively compare to the trends found in human responses, suggesting that people base judgments about collaborative contributions on a more complex model of how much the task could be sped-up if another agent participated.

However, there are several limitations that arise from applying a distributed systems model to human collaboration. Basic multiprocessing models lack key aspects of human group dynamics, such as task allocation ambiguity, communication overhead, fairness, and social norms. For example, standard distributed systems are typically not evaluated when there is only one unit of workload, because it is illogical to parallelize a single computation across processors. Alternatively, in human responses, there appear to be differences in perceived acceptability of social loafing even at the smallest workload level across group sizes. This difference may occur due to factors like uncertain task allocation. In distributed systems, there is no ambiguity about which processor will be assigned to subtasks due to scheduling algorithms (Davis & Burns, 2011). However, in human groups it is often unclear who will work on each subtask unless communicated otherwise. In our data, people may be simulating what will happen if none of the guests stand up to complete the task, which becomes more probable when there are only two agents present (Kwon et al., 2023). Unlike distributed systems, participants' judgments may also be influenced by social concerns. It may be counter-normative to make no attempt to wash the dishes, even when doing so would have no impact on task efficiency. This explanation poses a potential challenge in comparing multiprocessing systems to human behavior. It is important to consider the influence of social norms, such as whether or not an invited guest should contribute to a task, as well as the differences that arise cross-culturally in social loafing acceptability (Earley, 1989).

Our model also makes assumptions about the given tasks and environments that impose limitations to its ecological validity. Because we designed stimuli that are static screenshots of time-dependent tasks, the model must explicitly assume the relative proportions of overall task time required to complete each subtask. Our stimuli are highly controlled "toy" examples of real-world tasks. In order to confirm the model's validity, it will be crucial to replicate these findings using multi-player paradigms in which participants are able to allocate roles, resources, and subtasks in real time (Shrout & Rodgers, 2018; Yarkoni, 2022).

Future Directions

Our work provides a starting point for examining collaborative behavior from a multiprocessing perspective. Moving forward, this work can be extended in at least three ways. First, distributed systems motivate several new predictions about how individual contributions affect team performance. These theories predict that the enhancements achieved via parallelization are limited by not only workload, but also serial dependencies in time, memory, or resources that prevent computations from being distributed between processors. More concretely, changing the number of bottlenecks in each scene – such as the number of sinks or cutting boards – should shift the saturation point at which adding more agents will no longer improve task completion. Future work should assess how the presence of physical and cognitive bottlenecks impact the efficiency of collaborative groups and judgments of social loafing behavior. Second, our framework can be applied not just to predict third-party judgments, but also to test how considerations of task efficiency affect participants' first-hand effort allocation in real-time collaborations. Third, our framework can also be extended to capture properties that are *unique* to human collaborations as compared to distributed systems, such as how considerations of task efficiency may interact with culturally-specific norms about whether a guest is expected to contribute to household tasks.

Conclusion

Collaboration enables humans to achieve goals that no one individual could do on their own – however, finding ways to efficiently combine our efforts is itself a challenge. Here, we propose that distributed systems provide a rich source of hypotheses about how the structure of collaborative groups impacts performance, and we find that theories inspired by these systems capture human judgments about when it is acceptable not to contribute to a collaborative task. Put together, our work provides the first steps towards understanding human collaborations as a natural distributed system.

Resources for Reproducibility

The full stimulus set, behavioral dataset, and codebase can be found at `https://github.com/emieczkowski/SocialLoafingEfficiency`.

Acknowledgments

This work was supported by a grant from the Templeton World Charity Foundation.

References

Almaatouq, A., Alsobay, M., Yin, M., & Watts, D. J. (2021). Task complexity moderates group synergy. *Proceedings of the National Academy of Sciences*, *118*(36), e2101062118.

Almasi, G. S., & Gottlieb, A. (1994). *Highly parallel computing (2nd ed.)*. Benjamin-Cummings Publishing Co., Inc.

Amdahl, G. M. (1967). Validity of the single processor approach to achieving large scale computing capabilities. In *Proceedings of the April 18-20, 1967, Spring Joint Computer Conference* (pp. 483–485).

Baer, J.-L. (1973). A survey of some theoretical aspects of multiprocessing. *ACM Computing Surveys (CSUR)*, *5*(1), 31–80.

Bear, A., & Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *Cognition*, *167*, 25–37.

Cassidy, A. S., & Andreou, A. G. (2011). Beyond Amdahl's law: An objective function that links multiprocessor performance gains to delay and energy. *IEEE Transactions on Computers*, *61*(8), 1110–1126.

Davis, R. I., & Burns, A. (2011). A survey of hard real-time scheduling for multiprocessor systems. *ACM Computing Surveys*, *43*(4), 1–44.

Earley, P. C. (1989). Social loafing and collectivism: A comparison of the United States and the People's Republic of China. *Administrative Science Quarterly*, 565–581.

Gerstenberg, T., Ullman, T. D., Nagel, J., Kleiman-Weiner, M., Lagnado, D. A., & Tenenbaum, J. B. (2018). Lucky or clever? From expectations to responsibility judgments. *Cognition*, *177*, 122–141.

Griffiths, T. L. (2020). Understanding human intelligence through human limitations. *Trends in Cognitive Sciences*, *24*(11), 873–883.

Gustafson, J. L. (1988). Reevaluating Amdahl's law. *Communications of the ACM*, *31*(5), 532–533.

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. *Advances in Experimental Social Psychology*, *8*, 45–99.

Hennessy, J. L., & Patterson, D. A. (2011). *Computer architecture: a quantitative approach*. Elsevier.

Hill, M. D., & Marty, M. R. (2008). Amdahl's law in the multicore era. *Computer*, *41*(7), 33–38.

Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology*, *10*(4), 371–384.

Jackson, J. M., & Harkins, S. G. (1985). Equity in effort: An explanation of the social loafing effect. *Journal of Personality and Social Psychology*, *49*(5), 1199–1206.

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, *65*(4), 681–706.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology*, *44*(1), 78–94.

Kravitz, D. A., & Martin, B. (1986). Ringelmann rediscovered: The original article. *Journal of Personality and Social Psychology*, *65*, 681–706.

Kshemkalyani, A. D., & Singhal, M. (2011). *Distributed computing: principles, algorithms, and systems*. Cambridge University Press.

Kwon, J., Zhi-Xuan, T., Tenenbaum, J., & Levine, S. (2023). When it is not out of line to get out of line: The role of universalization and outcome-based reasoning in rule-breaking judgments. *PsyArXiv*.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*(4), 343–356.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, *37*(6), 822–832.

Mao, A., Mason, W., Suri, S., & Watts, D. J. (2016). An experimental study of team size and performance on a complex task. *PloS One*, *11*(4), e0153048.

Mas, A., & Moretti, E. (2009). Peers at work. *American Economic Review*, *99*(1), 112–145.

Petty, R. E., Harkins, S. G., Williams, K. D., & Latane, B. (1977). The effects of group size on cognitive effort and evaluation. *Personality and Social Psychology Bulletin*, *3*(4), 579–582.

Ringelmann, M. (1913). Research on animate sources of power: The work of man. In *Annales de l'institut national agronomique* (Vol. 12, pp. 1–40).

Roberts, S. O., Gelman, S. A., & Ho, A. K. (2017). So it is, so it shall be: Group regularities license children's prescriptive judgments. *Cognitive Science*, *41*, 576–600.

Schmidt, M. F., & Tomasello, M. (2012). Young children enforce social norms. *Current Directions in Psychological Science*, *21*(4), 232–236.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*, 487–510.

Simms, A., & Nichols, T. (2014). Social loafing: A review of the literature. *Journal of Management Policy and Practice*, *15*(1), 58–67.

Steiner, I. D. (1972). *Group process and productivity*. Academic Press.

Turner, C. R., Nielsen, M., & Collier-Baker, E. (2014). Groups' actions trump injunctive reaction in an incidental observation by young children. *PloS one*, *9*(9), e107375.

Vélez, N., Christian, B., Hardy, M., Thompson, B. D., & Griffiths, T. L. (2023). How do humans overcome individual computational limitations by working together? *Cognitive Science*, *47*(1), e13232.

Wu, S. A., & Gerstenberg, T. (2024). If not me, then who? Responsibility and replacement. *Cognition*, *242*, 105646.

Xiang, Y., Landy, J., Cushman, F. A., Vélez, N., & Gershman, S. J. (2023). Actual and counterfactual effort contribute to responsibility attributions in collaborative tasks. *Cognition*, *241*, 105609.

Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, e1.