**Independent and normally-distributed response thresholds are biologically feasible for social insects**

Lynch C.M.1, Pavlic, T.P.1, 2, 3, 4

1School of Life Science, Arizona State University, Tempe, AZ, USA 85287

2School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ, USA 85287

3School of Complex Adaptive Systems, Arizona State University, Tempe, AZ, USA 85287

4School of Sustainability, Arizona State University, Tempe, AZ, USA 85287

**Abstract**

Measuring heterogeneous traits among members of a group can be onerous when the trait of interest is not visible. Heterogeneous patterns of behavior can be partially summarized by descriptions of division of labor (DOL) in social insects, and the most widely-cited causal mechanism behind DOL is the response threshold. A comprehensive understanding of the form and function of heterogeneity requires an understanding of how response thresholds are distributed among workers and tasks. However, the task-associated stimuli for many tasks are not known, and it can be difficult to get repeated measures of individual insects. It is therefore challenging to measure these thresholds directly. Instead, we can make inferences about true response threshold distributions by measuring colony-level dynamics of different proposed threshold distributions and comparing these dynamics to those of real ants. We find that the most common distribution type for response threshold models - a normal distribution where every threshold is independent of both the individual and task - produces the most realistic dynamics of the distributions tested here. This distribution yields levels of division of labor that are consistent with that of real social insects, it successfully predicts positive scaling between colony size and division of labor, activity levels are right-tailed, and it tends to minimize activity while it also keeps the stimulus at a manageable level.

**Introduction**

Grouping behaviors are nearly ubiquitous in the animal kingdom, and can take on a number of different forms. On one end of the social spectrum, cockroaches aggregate as a response to environmental stressors (Jeanson et al., 2005) while on the other end of the spectrum social insect colonies are considered superorganisms (Holldobler & Wilson, 2009), as the group can be considered a coherent vehicle for gene propagation (Reeve & Holldobler, 2007). Group living can confer a number of energetic and safety benefits (Krause et al., 2002; Ward & Webster, 2016), which are realized through the integrated actions of individuals within the group. An understanding of the heterogeneity of individuals within the group is therefore critical to the study of collective behavior (Farine et al., 2015). Despite the extensive empirical studies connecting individual variation with collective behavior (Metcalfe et al., 2016; Seebacher & Krause, 2017; Killen et al., 2017), relatively few theoretical studies account for heterogeneity (Jolles et al., 2020).

Social insect colonies are excellent and tractable examples of complex adaptive systems, self-organizing without the need of a centralized control structure (Gordon, 1996). As such, they have been used as a model system for many empirical and theoretical studies of collective behavior (Seeley et al., 2012; Khuong et al., 2016; Huang & Robinson, 1996; Greenwald et al., 2018). Social insects also display division of labor (DOL), which is among the more extreme versions of individual heterogeneity within a group, playing a key role in the evolutionary transitions in individuality (Michod, 2007). Among social insects, division of labor is the degree to which individuals within a group specialize in one or more tasks (Gorlick et al., 2004). It can also be described as a stable pattern of variation among workers with respect to the tasks they perform (Beshers & Fewell, 2001). While phenotypic variation can result from differences in many individual traits such as age (Naug & Gadagkar, 1998), morphology (Kwapich et al., 2018) and experience (Ravary et al., 2007) as well as several different proximate mechanisms (Beshers & Fewell, 2001), the majority of studies on the impact of variability on colony level focus on the distribution of response thresholds (Oldroyd & Fewell, 2007; Jeanson & Weidenmuller, 2014).

Response thresholds are the propensity of individuals to respond to task-specific stimuli, indicating a group-level demand for a task (Page & Mitchell, 1998). According to the response threshold hypothesis, individual workers in a colony intrinsically differ in the stimulus level at which they start working (their response threshold), and by performing the task associated with that stimulus, they reduce the stimulus and therefore also reduce the probability that this task will be performed by others (Jeanson & Weidenmüller, 2014). While response thresholds can be made to vary with experience (Theraulaz et al., 1998), models employing a fixed threshold have been successful in recreating some behaviors associated with division of labor (Brahma et al., 2018; Pankiw & Page, 2000; Ulrich et al., 2018; but see Weidenmuller et al., 2019; Ulrich et al., 2021).

Most papers which study the relationships between variation in thresholds and variation in behavior usually assume a normal distribution of thresholds (Myerscough & Oldroyd, 2004; Jeanson et al., 2007; Richardson et al., 2011; Ulrich et al., 2020; Ulrich et al., 2021; Lin, 2021; Lynch et al., *in prep;* Smith et al., *in prep*). This seems like a reasonable assumption, as many social insect traits are normally distributed (Tay & Crozier, 2001; Schultz et al., 2002; DeHeer & Vargo, 2006; Parr et al., 2007; Couvillon et al., 2010, Frank et al., 2018; ). However, it is difficult to validate this assumption, as no one has yet to measure response threshold distributions, either because the task-associated stimulus for a task is not known (Beshers & Fewell, 2001) or because it is challenging to get repeated measurements of individuals (Detrain et al., 2019; Leitner et al., 2019). Still, assumptions about individuals can be tested by studying group dynamics (Fewell & Bertram, 1999), and since the choice of threshold distribution can drastically affect performance (Wu et al., 2020), normal, idiosyncratic distributions can be compared to other proposals for how work can be distributed among individuals. For instance, workers can share tasks (Helms & Fewell, 2004), a few elite workers can perform most of the work (Pinter-Wollman et al., 2012), and workers can be divided into discrete castes which specialize in a single or a subset of tasks (Wheeler, 1991). Each of these scenarios can be modeled with a different distribution of response thresholds.

Here, we compare the abilities of different biologically-relevant response threshold distributions to reproduce real patterns of behavior. We test whether each distribution can successfully predict scaling relationships between colony size and division of labor (Jeanson et al., 2007; Holbrook et al., 2011), produce right-tailed activity levels (Dornhaus et al., 2009, Richardson et al., 2011; Pinter-Wollman et al., 2012; Yamanaka et al., 2019), and produce realistic levels of division of labor (Gorelick et al., 2004; Dornhaus et al., 2009; Jandt et al., 2009; Holbrook et al., 2011). We also determine whether or not these distributions are capable of minimizing overall activity while completing the same amount of work as an algorithm designed to minimize the work signal. This could give researchers an indication as to which definition of inter-individual variability best represents the behaviors associated with division of labor.

**Methods**

*Generating response threshold distributions*

Let **𝛉** be the matrix that contains the K response thresholds for N ants:

\boldsymbol \theta  = \begin{bmatrix} 
    \theta_{11}  & \dots & \theta_{1N}\\
    \vdots & \ddots & \vdots \\
    \theta__{K1} &    \dots    & \theta__{KN} 
    \end{bmatrix}

where 𝛉ij is ant j’s response threshold for task i. Note that for softmax functions, each ant will also need a response threshold for the rest state, which is typically set to 0 (Lynch et al., *in prep*). This would represent a row vector of all 0’s for 𝛉0, j = 1, 2, …, N. This row vector is used for all distribution types and is added after the construction of **𝛉**.

For the task sharing distribution (Fig. 1A), all elements of the matrix are set equal to a constant intercept IT, the single free parameter of this distribution:

\theta_{ij} = I_T

There are two types of workers in the elite distribution (Fig. 1B): elites and non-elites. There are NE elites in the colony (and therefore there are N-NE non-elites) whose response thresholds differ from non-elites by ΔE. All workers also have a base threshold of IE. Elites have the same threshold for all tasks, while non-elites have the same threshold for all of their tasks. This can be summarized by the piecewise function:

\theta_{ij} = \begin{cases} 
      I_E-\Delta_E/2 & j \leq N_C \\
      I_E+\Delta_E/2 & j >N_C
   \end{cases}

Note that when we perform the sensitivity analysis, we freely vary N and then choose NE by multiplying N by ρE, which is a random number drawn from a uniform distribution between 0 and 1.

When a colony has castes (Fig. 1C), we assume that every task has its own caste. We also assume that all tasks are equally important, and therefore have the same number of workers in each caste, NC. To calculate NC, we take the floor function of N/K, or N_C = \lfloor N/K \rfloor . Members of the same caste will have the same threshold for their particular task. They also have a different, higher threshold for all other tasks. In cases where N/K does not yield a whole number, the remaining ants are all given the same, high threshold. The difference between the high and low thresholds is ΔC. This process can be summarized by:

\theta_{ij} = \begin{cases} 
      I_C-\Delta_C/2 & \text{for} \ i \in \{ N_C (j-1) + 1, 2, ..., N_C\} \\
      I_C+\Delta_C/2 & \text{o.w.}
   \end{cases}

Finally, we assume that colonies with normally distributed thresholds have no internal, correlational structures (Fig. 1D). Therefore, every draw from the normal distribution is purely independent of both the individual and the task:

\theta_{ij} \sim \mathcal{N}(\mu,\,\sigma^{2})

where μ is the mean and σ2 is the standard deviation of the distribution.

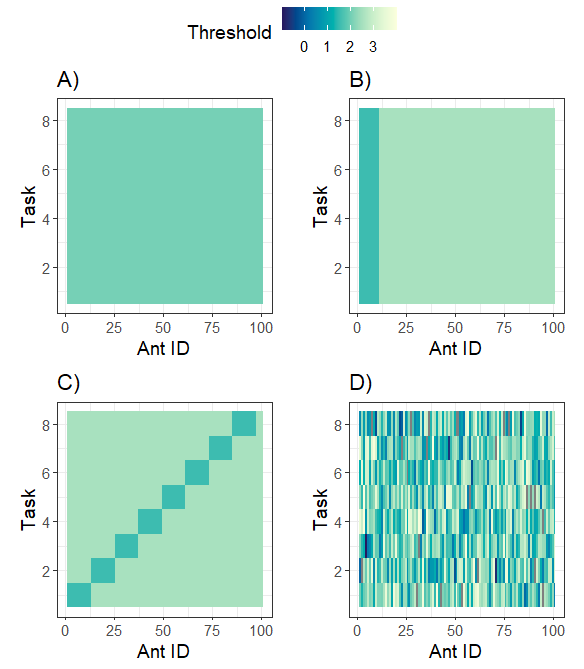


Figure 1: Examples of the response threshold matrix **𝛉** for different distributions. A) task sharing, B) elite, C) caste, and D) normal distribution. IT = IE = μ = 2, ΔE = ΔC = 𝜎2 = 1, NC = 12, NE = 10.

*Simulations of ant behavior*

In these discrete time Markovian models, the probability that ant j will perform task i is given by a softmax function when she is in the inactive state:

\phi(x_{ti}, \theta_{ij})_{ij} = \frac{e^{k(x_{ti}-\theta_{ij})}}{\sum_{m=1}^K e^{k(x_{tm}-\theta_{mj})}}

where xti is the task-associated stimulus for task i at time t,, K is the total number of states (so there are K-1 tasks) and k is the steepness parameter of the probability curve (the equivalent of the inverse temperature). Note that the rest state is where i = 0, and xt0 = **𝛉**i0 = 0. Once an ant is engaged in a task, she must at some point return to the rest state before starting another task. On average, an ant will spend 𝝉 = 5 timesteps on the task, or have a ⅕ probability of stopping the task.

Next, the task demand stimulus for each task is updated with the following function (Jeanson et al., 2007):

x_{ti} = x_{t-1i} + \frac{\frac{\alpha N 
\delta}{K-1}-\alpha n_{ti}}{N}

where N is the colony size, nti is the number of ants in state i at time t, 𝛂 is the amount of stimulus an ant can remove from the environment in a single timestep, and 𝛅 is the amount that each task increases every time step. xt0 does not change over time, and all other xti is initialized to 0. xti can take on both positive and negative values, so we consider a value of 0 to be an ideal intermediate value of some task-associated stimulus. For instance, 0 could represent a desirable ambient temperature of a bumble bee hive around 25° C (O'Donnell & Foster, 2001). Individual bees will fan at higher temperatures (positive values of x) and will incubate brood at lower temperatures (negative values of x).

When comparing the performances of these models, we want to ensure that the stimulus does not increase at a rate faster than the rate at which the colony can remove this stimulus. Thus we want to find the maximum 𝛅 such that the stimulus stays constant throughout time (Lynch et al., *in prep*). This occurs when:

\delta = \frac{\tau}{\tau+1}

We do not vary 𝝉, so the upper limit for 𝛅 when 𝝉 = 5 is 5/(5+1) = ⅚.

For every simulation run, we drew random values for each free parameter from a uniform distribution. We chose to vary the free parameters that manipulated **𝛉** so we could directly compare the performance of each distribution type on one another. For all models, K = 9, 𝛂 = 3, k = 2, and 𝛅= 0.6. N can also be any integer between 10 and 1000. For the task sharing distribution, we only varied IT, which could vary between -5 and 5. For the elite distribution, both IE and ΔE vary between -5 and 5, while ρE varies between 0 and 1. For the caste distribution, both IC and ΔC vary between -5 and 5. For the normal distribution, μ and 𝜎2 also vary between -5 and 5.

For every simulation we measure division of labor as well as the kurtosis of activity for a single task. Division of labor is derived from Shannon’s mutual entropy and allows us to determine the degree to which individuals specialize on one or a few tasks (Gorelick et al. 2004; Dornhaus et al., 2009). Activity within tasks (as well as other attributes such as velocity) tend to be right tailed so that most individuals spend only a little bit of time on a task whereas a few individuals spend a lot of time on it (Dornhaus et al., 2009, Waters et al., 2010; Richardson et al., 2011; Pinter-Wollman et al., 2012). In other words, these activity distributions will have a high kurtosis level. As all tasks are treated equally in this model, we measure kurtosis with respect to only task 1. That is, we count the number of timesteps each ant in a colony is performing task 1 and measure the kurtosis of the resulting activity distribution.

We ran 2,500 simulations for each type of distribution, drawing random values for each free parameter each time. Each simulation run was repeated 5 times, and the performance measures we each averaged across these 5 repetitions. Each simulation was run for 1,000 timesteps. In total there were 2,500 × 5 × 4 = 50,000 simulations. Simulations were run in MATLAB (version 2021b).

*Model Performance of Distributions Relative to Optimization Algorithm*

Here we compare the performance of each distribution type to a colony whose thresholds are set by a genetic algorithm. This colony represents a benchmark, a measurement of how well these natural distributions compare to an artificially derived one. This model is designed to minimize the distance between the average stimulus levels across tasks and time (\hat{x}_{..} ) and some arbitrarily defined target which lies between the maximum and minimum threshold values set for the distributions (x^*). More concretely, this algorithm minimizes the quadratic loss function L = ( x^* - \hat{x}_{..} )^2.

To make these comparisons, we first run the genetic algorithm (Appendix 1) to find an optimal 𝛉 for 25 ants and 4 tasks (Fig. 2) and then run 100 simulations with this distribution, each time measuring 5 different performance measurements. We then run 1,000 simulations for each other type of distribution where the free parameters of each are randomly selected from a uniform distribution. We then find the set of free parameters that minimizes L, and then run 100 more replicates of this optimal set of parameters and again measure 5 different performance metrics for each replicate. We then compare the levels of each performance metric across all distribution types with a pairwise Wilcoxon test (with a Holm-Bonferroni correction for multiple comparisons) to establish ranks. A distribution is given a higher rank for minimizing each performance metric. If two distributions are part of the same significance group though, then they are given an equal rank. Ranks were finally normalized so that their sum was equal to 1+2+3+4+5 = 15.

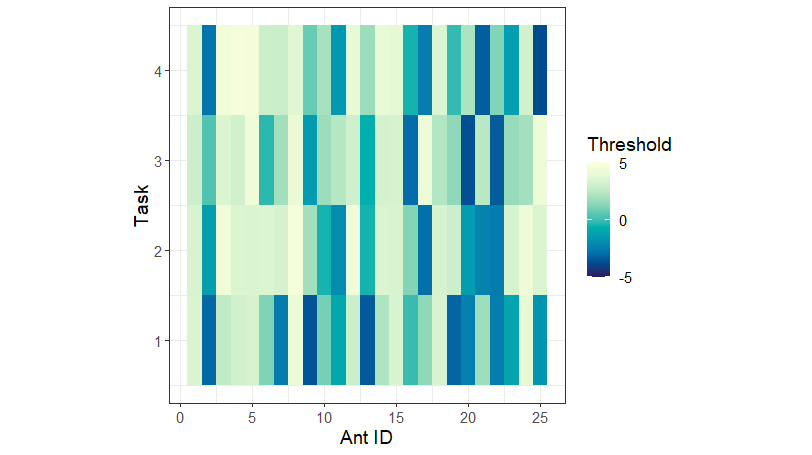


Figure 2: Optimal set of thresholds produced by the genetic algorithm. The x-axis represents individual ants, the y-axis represents tasks, and the color gives the strength of the threshold.

The 5 performance measures are: proportion of active workers, worker variation, number of state switches, the number of timesteps to equilibrium, and the loss L. The proportion of active workers sums the number of workers that are active each timestep and then divides this sum by N. It is then averaged across timesteps. The variation of active workers is similar, except we take the standard deviation at the end rather than the mean. To get the number of task switches, we count how often each ant switches tasks throughout the duration of the simulation, then average across ants. To find the number of timesteps across equilibrium, we first average x across all tasks and then find the difference between sequential values of this average. We then smooth the resulting vector with a moving average filter, and find the first timestep where this smooth vector is greater than -0.00025 and less than 0.00025. The first element where this happens gives the number of timesteps to equilibrium. If this value is never reached (as in cases where one or more of the stimuli consistently increase or decrease), then the function returns an NA. Most simulations that did reach equilibrium did so before 500 timesteps, so to get L we averaged x across all tasks and across the final 500 timesteps of the simulation.

*Data analytics*

To test whether different threshold distributions produced different levels of division of labor across simulations, we performed a pairwise Wilcox test, using a Holm-Bonferonni correction for multiple comparisons. We then determined whether or not the division of labor indices from the simulations overlapped those of real social insect colonies (Dornhaus et al., 2009; Jandt et al., 2009; Holbrook et al., 2011). We also performed a pairwise Wilcox test on the kurtosis values of each distribution.

To determine whether or not division of labor naturally scales with colony size for each distribution type, we performed a beta regression between the division of labor index, each of the free parameters of the threshold distributions as well as the log of colony mass. Colony mass in this case ignores the mass of brood and queens, only capturing the mass of workers. The mass of a harvester ant is approximately 0.003 grams, so we would multiply this number by N and then log transform the result. We perform the log transformation as this is typical for scaling studies (Harrison et al., 2022). We use a beta regression as the DOL index varies between 0 and 1. To test for the presence of interactions between these parameters, each beta regression was built with all possible interaction structures and then we selected the model which minimized AIC and BIC. As division of labor tends to increase with colony size (Holbrook et al., 2011), we determine whether or not the regression coefficient for N (as well as any possible interactions with N) is statistically significant and positive.

**Results**

*Range of division of labor indices for each distribution type*

Division of labor in real social insects tends to range between 0.15 and 0.6 (Dornhaus et al., 2009; Jandt et al., 2009; Holbrook et al., 2011). However, not all distribution types are capable of generating these values (Fig. 3). While each distribution type is significantly different from each other (all pairwise p-values < 0.001), only the caste and normal distributions produce division of labor indices in the biologically-determined range. The elite distribution is able to capture the very bottom of this range, but the task sharking distribution misses it altogether.

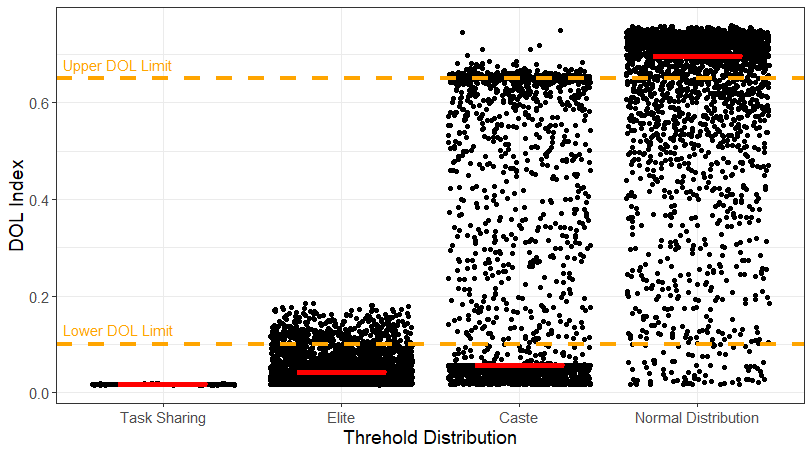


Figure 3: The division of labor index (y-axis) across 1000 simulations for each distribution type (x-axis). Each point is a simulation, and the red bar represents the median of each distribution type. The upper orange line represents the upper limit of DOL measurements in real social insects while the other line represents the lower limit.

*Kurtosis of activity level*

The distributions of worker activity within tasks tends to be either exponential (Pinter-Wollman et al., 2012) or gamma-distributed where the shape parameter is close to 1 (Yamanaka et al., 2019). Therefore, the kurtosis values of these distributions is either exactly or approximately 9. Conversely, a normal distribution has a kurtosis level of 3. Once again, each each distribution type is significantly different from each other (all pairwise p-values < 0.001) and they each have different ranges. The median kurtosis levels of the caste, task sharing, and elite distributions are all close to that of a normal kurtosis level (Fig. 4), although the range for the caste distribution is higher than the other two. Conversely, the median kurtosis value produced by the normal distribution of thresholds is very close to that of the exponential/gamma distribution. The range of kurtosis values also spans a wide margin, indicating that a normal distribution can be made to form different activity patterns in an ant colony.

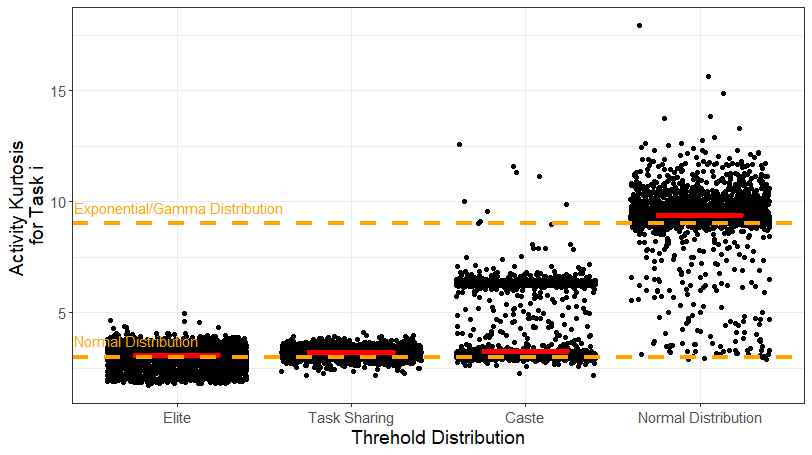
**

Figure 4 : The division of labor index (y-axis) across 1000 simulations for each distribution type (x-axis). Each point is a simulation, and the red bar represents the median of each distribution type. The upper orange line represents the kurtosis of an exponential distribution as well as a gamma distribution whose shape parameter is 1. This kurtosis value is 9. The lower orange line represents the kurtosis of a normal distribution, 3.

*Scaling relationships between colony size and division of labor*

The most parsimonious beta regression for most distribution types (lowest AIC/BIC) was the one which had only the main effects (Normal Distribution: phi = 15.288, p-value < 0.001, pseudo R2 = 0.561; Task Sharing: phi = 246228, p-value < 0.001, pseudo R2 = 0.19; Elite: phi = 43.136, p-value < 0.001, pseudo R2 = 0.005). The one exception was the caste distribution model, which included an interaction between mass and IC (Caste: phi = 14.013, p-value < 0.001, pseudo R2 = 0.6584). The main effect of mass was significant and positive for both the task sharing and normal distributions (Task Sharing: estimate = 0.0032, p-value < 0.001; Normal: estimate = 0.029, p-value < 0.001). On the other hand, this effect was not significant for the elite distribution (Elite: -0.007, p-value = 0.624). It also was not significant for either the main effect nor the interaction for the caste distribution (Caste: Mass main effect estimate = 0.012, p-value = 0.405; Main effect \* IC interaction estimate = -0.002, p-value = 0.654).

*Performance of natural distributions vs genetic algorithm*

Even though the underlying characteristics of virtual agents are different across distribution types, colony performance for each remained largely the same, each performing as well as the genetic algorithm (Fig. 5). The median loss was equivalent across all distribution types (although it should be noted that the task sharing loss was slightly lower than that of the genetic algorithm, but it was not different from the other distribution types, so they were all given the same rank), they all had the same number of task switches, and workers were equally active for all types. The only performance metric which saw significant differences between distribution types was the ability to minimize the variation of the task-associated stimuli. Here, the Caste distribution prevailed. The only metric in which the genetic algorithm performed better than the other models was that the stimulus came to equilibrium in fewer timesteps

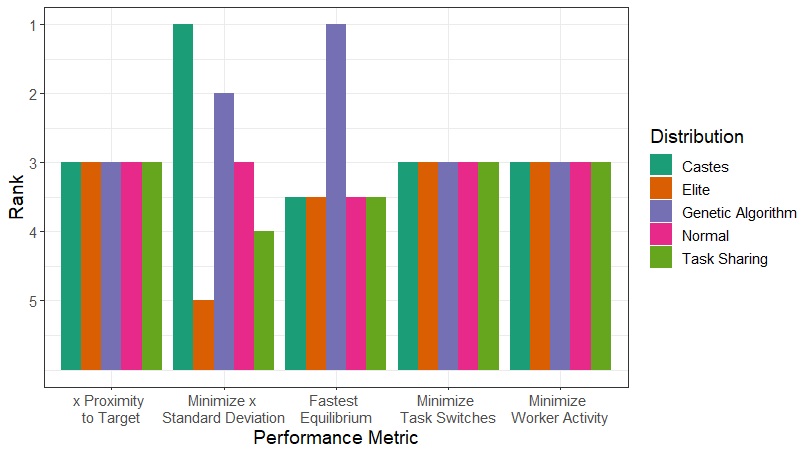


Figure 5: Ranks (y-axis) for each distribution type (color) within each performance metric (x-axis).

**Discussion**

Stochastic models which use response thresholds as a mechanism for division of labor in social insects usually assume a normal distribution. That is, the threshold for every individual and every task is independent of another and drawn from a normal distribution. We validate this assumption by showing that normal distributions are better than other reference distributions at reproducing a small set of biological behaviors. Normal distributions can reproduce realistic levels of division of labor, right-skewed activity levels, and successfully predict scaling relationships between colony size and division of labor. Normal distributions (as well as the other reference distributions) also perform as well as an artificial distribution produced by a genetic algorithm in terms of some colony performance metrics.

First and foremost, any model which purports to explain division of labor should reproduce values of division of labor found in nature. Unsurprisingly, this constraint necessitates differences among individuals in the colony. Colonies that possess identical individuals - task sharing colonies - produce extremely low levels of division of labor. The same applies to a slightly lesser extent to colonies which only have elite workers. Division of labor is higher here relative to task sharing colonies, however it seems impossible for these models to produce DOL indices higher than 0.2. Conversely, colonies which have castes or have normally distributed response thresholds are able to remake the entire range of division of labor. Altering the parameters of either of these models can allow the colony to have any index between 0 and 0.8.

Only normal distributions, though, seem capable of reproducing right-tailed activity levels of social insects. It has been shown that the amount of time that ants spend on individual tasks follows either an exponential distribution (Pinter-Wollman et al., 2012) or a gamma distribution (Yamanaka et al., 2019). These distributions have kurtosis levels of 9, which only the normal distribution of thresholds was able to reproduce. The normal distribution for activity level has only a kurtosis of 3, which was a good estimate for the kurtosis values of the task-sharing, elite, and caste threshold distributions.

Response thresholds have also been used to investigate the scaling relationship between size, activity, and metabolism (Jeanson et al., 2007; Lin 2021) as division of labor tends to increase with colony size (Jeanne, 1986; Thomas & Elgar, 2003; Holbrook et al., 2011; Amador-Vargas et al., 2015). Task-sharing and the normal distribution models both predict this relationship, whereas the correlation between colony size and division of labor is not present for the elite and caste distributions, indicating that these are poor models to use for scaling studies. Ultimately, only the normal distribution passes all of these tests simultaneously, indicating that it should continue to be used for future studies.

**Works cited**

Amador-Vargas, S., Gronenberg, W., Wcislo, W. T., & Mueller, U. (2015). Specialization and group size: brain and behavioural correlates of colony size in ants lacking morphological castes. Proceedings of the Royal Society B: Biological Sciences, 282(1801), 20142502.

Beshers, S. N., & Fewell, J. H. (2001). Models of division of labor in social insects. Annual review of entomology, 46(1), 413-440.

Bonabeau, E., Theraulaz, G., & Deneubourg, J. L. (1996). Quantitative study of the fixed threshold model for the regulation of division of labour in insect societies. Proceedings of the Royal Society of London. Series B: Biological Sciences, 263(1376), 1565-1569.

Brahma, A., Mandal, S., & Gadagkar, R. (2018). Emergence of cooperation and division of labor in the primitively eusocial wasp Ropalidia marginata. Proceedings of the National Academy of Sciences, 115(4), 756-761.

Couvillon, M. J., Jandt, J. M., Duong, N. H. I., & Dornhaus, A. (2010). Ontogeny of worker body size distribution in bumble bee (Bombus impatiens) colonies. Ecological Entomology, 35, 424-435.

DeHeer, C. J., & Vargo, E. L. (2006). An indirect test of inbreeding depression in the termites *Reticulitermes flavipes* and Reticulitermes virginicus. Behavioral Ecology and Sociobiology, 59, 753-761.

Detrain, C., Pereira, H., & Fourcassié, V. (2019). Differential responses to chemical cues correlate with task performance in ant foragers. Behavioral ecology and sociobiology, 73(8), 1-11.

Dornhaus A, Holley JA, Franks NR. Larger colonies do not have more specialized workers in the ant Temnothorax albipennis. Behavioral Ecology. 2009 May 19;20(5):922-9.

Farine, D. R., Montiglio, P. O., & Spiegel, O. (2015). From individuals to groups and back: the evolutionary implications of group phenotypic composition. Trends in ecology & evolution, 30(10), 609-621.

Fewell, J. H., & Bertram, S. M. (1999). Division of labor in a dynamic environment: response by honeybees (Apis mellifera) to graded changes in colony pollen stores. Behavioral ecology and sociobiology, 46(3), 171-179.

Frank, E. T., Wehrhahn, M., & Linsenmair, K. E. (2018). Wound treatment and selective help in a termite-hunting ant. Proceedings of the Royal Society B: Biological Sciences, 285, 20172457.

Gordon, D. M. (1996). The organization of work in social insect colonies. Nature, 380(6570), 121-124.

Gorelick, R., Bertram, S. M., Killeen, P. R., & Fewell, J. H. (2004). Normalized mutual entropy in

biology: quantifying division of labor. The American Naturalist, 164(5), 677-682.

Greenwald, E. E., Baltiansky, L., & Feinerman, O. (2018). Individual crop loads provide local control for collective food intake in ant colonies. Elife, 7, e31730.

Harrison, Jon F., Andrew Biewener, Joanna R. Bernhardt, Joseph R. Burger, James H. Brown, Zach N. Coto, Meghan E. Duell et al. "White paper: an integrated perspective on the causes of hypometric metabolic scaling in animals." Integrative and Comparative Biology 62, no. 5 (2022): 1395-1418.

Helms Cahan, S., & Fewell, J. H. (2004). Division of labor and the evolution of task sharing in queen associations of the harvester ant Pogonomyrmex californicus. Behavioral ecology and sociobiology, 56(1), 9-17.

Holbrook, C. T., Barden, P. M., & Fewell, J. H. (2011). Division of labor increases with colony size in the harvester ant Pogonomyrmex californicus. Behavioral Ecology, 22(5), 960-966.

Hölldobler, B., & Wilson, E. O. (2009). The superorganism: the beauty, elegance, and strangeness of insect societies. WW Norton & Company.

Huang, Z. Y., & Robinson, G. E. (1996). Regulation of honey bee division of labor by colony age demography. Behavioral Ecology and Sociobiology, 39(3), 147-158.

Jandt, J. M., Huang, E., & Dornhaus, A. (2009). Weak specialization of workers inside a bumble bee (Bombus impatiens) nest. Behavioral Ecology and Sociobiology, 63(12), 1829-1836.

Jeanne, R. L. (1986). The evolution of the organization of work in social insects. Monitore

Zoologico Italiano-Italian Journal of Zoology, 20(2), 119-133.

Jeanson, R., Rivault, C., Deneubourg, J. L., Blanco, S., Fournier, R., Jost, C., & Theraulaz, G. (2005). Self-organized aggregation in cockroaches. Animal behaviour, 69(1), 169-180.

Jeanson, R., Fewell, J. H., Gorelick, R., & Bertram, S. M. (2007). Emergence of increased division of labor as a function of group size. Behavioral Ecology and Sociobiology, 62(2), 289-298.

Jeanson, R., & Weidenmüller, A. (2014). Interindividual variability in social insects–proximate causes and ultimate consequences. Biological Reviews, 89(3), 671-687.

Jolles, J. W., King, A. J., & Killen, S. S. (2020). The role of individual heterogeneity in collective animal behaviour. Trends in ecology & evolution, 35(3), 278-291.

Khuong, Anaïs, Jacques Gautrais, Andrea Perna, Chaker Sbaï, Maud Combe, Pascale Kuntz, Christian Jost, and Guy Theraulaz. "Stigmergic construction and topochemical information shape ant nest architecture." Proceedings of the National Academy of Sciences 113, no. 5 (2016): 1303-1308.

Killen, S. S., Marras, S., Nadler, L., & Domenici, P. (2017). The role of physiological traits in assortment among and within fish shoals. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1727), 20160233.

Krause, J., Ruxton, G. D., & Ruxton, G. (2002). Living in groups. Oxford University Press.

Kwapich, C. L., Valentini, G., & Hölldobler, B. (2018). The non-additive effects of body size on nest architecture in a polymorphic ant. Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1753), 20170235.

Leitner, N., Lynch, C., & Dornhaus, A. (2019). Ants in isolation: obstacles to testing worker responses to task stimuli outside of the colony context. Insectes Sociaux, 66(3), 343-354.

Lynch, C.M., Pavlic, P. Modeling response thresholds for multiple simultaneous stimuli using a Boltzmann-sampling approach. *In prep.*

Lynch, C.M., Wilson, R., Dornhaus, A. Stop and go: exploring alternative mechanisms for dynamic task allocation in social insects - response and satisfaction thresholds trade off time invested and work accomplished differently

Metcalfe, N. B., Van Leeuwen, T. E., & Killen, S. S. (2016). Does individual variation in metabolic phenotype predict fish behaviour and performance?. Journal of fish biology, 88(1), 298-321.

Michod, R. E. (2007). Evolution of individuality during the transition from unicellular to multicellular life. Proceedings of the National Academy of Sciences, 104(suppl\_1), 8613-8618.

Myerscough, M. R., & Oldroyd, B. P. (2004). Simulation models of the role of genetic variability in social insect task allocation. Insectes Sociaux, 51(2), 146-152.

O'Donnell, S., & Foster, R. L. (2001). Thresholds of response in nest thermoregulation by worker bumble bees, Bombus bifarius nearcticus (Hymenoptera: Apidae). Ethology, 107(5), 387-399.

Oldroyd, B. P., & Fewell, J. H. (2007). Genetic diversity promotes homeostasis in insect colonies. Trends in ecology & evolution, 22(8), 408-413.

Page Jr, R. E., & Mitchell, S. D. (1998). Self-organization and the evolution of division of labor. Apidologie, 29(1-2), 171-190.

Pankiw, T., & Page Jr, R. E. (2000). Response thresholds to sucrose predict foraging division of labor in honeybees. Behavioral ecology and sociobiology, 47(4), 265-267.

Parr, C. L., Andersen, A. N., Chastagnol, C., & Duffaud, C. (2007). Savanna fires increase rates and distances of seed dispersal by ants. Oecologia, 151(1), 33-41.

Pinter-Wollman, N., Hubler, J., Holley, J. A., Franks, N. R., & Dornhaus, A. (2012). How is activity distributed among and within tasks in Temnothorax ants?. Behavioral Ecology and Sociobiology, 66(10), 1407-1420.

Radcliff, N. J. 1991. Forma analysis and random respectful recombination. In Proc. 4th

Int. Conf. on Genetic Algorithms, San Mateo, CA: Morgan Kauffman.

Ravary, F., Lecoutey, E., Kaminski, G., Châline, N., & Jaisson, P. (2007). Individual experience alone can generate lasting division of labor in ants. Current Biology, 17(15), 1308-1312.

Reeve, H. K., & Hölldobler, B. (2007). The emergence of a superorganism through intergroup

competition. Proceedings of the National Academy of Sciences, 104(23), 9736-9740.

Richardson, T. O., Christensen, K., Franks, N. R., Jensen, H. J., & Sendova-Franks, A. B. (2011). Ants in a labyrinth: a statistical mechanics approach to the division of labour. PLoS One, 6(4), e18416.

Schulz, D. J., Sullivan, J. P., & Robinson, G. E. (2002). Juvenile hormone and octopamine in the regulation of division of labor in honey bee colonies. Hormones and behavior, 42, 222-231.

Seebacher, F., & Krause, J. (2017). Physiological mechanisms underlying animal social behaviour. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1727), 20160231.

Seeley, T. D., Visscher, P. K., Schlegel, T., Hogan, P. M., Franks, N. R., & Marshall, J. A. (2012). Stop signals provide cross inhibition in collective decision-making by honeybee swarms. Science, 335(6064), 108-111.

Tay, W. T., & Crozier, R. H. (2001). Mating behaviour of Rhytidoponera sp. 12 ants inferred from microsatellite analysis. Molecular Ecology, 10, 167-173.

Thomas M & Elgar M (2003) Colony size affects division of labour in the ponerine ant Rhytidoponera metallica. Naturwissenschaften 90(2):88-92.

Theraulaz, G., Bonabeau, E., & Denuebourg, J. N. (1998). Response threshold reinforcements and division of labour in insect societies. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265(1393), 327-332.

Ulrich, Y., Saragosti, J., Tokita, C. K., Tarnita, C. E., & Kronauer, D. J. (2018). Fitness benefits and emergent division of labour at the onset of group living. Nature, 560(7720), 635-638.

Ulrich, Y., Kawakatsu, M., Tokita, C. K., Saragosti, J., Chandra, V., Tarnita, C. E., & Kronauer, D. J. (2020). Emergent behavioral organization in heterogeneous groups of a social insect. BioRxiv.

Ulrich, Y., Kawakatsu, M., Tokita, C. K., Saragosti, J., Chandra, V., Tarnita, C. E., & Kronauer, D. J. (2021). Response thresholds alone cannot explain empirical patterns of division of labor in social insects. PLoS biology, 19(6), e3001269.

Ward, A., & Webster, M. (2016). Sociality: the behaviour of group-living animals.

Waters, James S., et al. "Allometric scaling of metabolism, growth, and activity in whole colonies of the seed-harvester ant Pogonomyrmex californicus." The American Naturalist 176.4 (2010): 501-510.

Waters, J. S., Toth, J., Harrison, J. F., & Fewell, J. H. (2017, March). Metabolic Allometry and the Scaling of Interaction Patterns with Ant Colony Size. In INTEGRATIVE AND COMPARATIVE BIOLOGY (Vol. 57, pp. E441-E441). JOURNALS DEPT, 2001 EVANS RD, CARY, NC 27513 USA: OXFORD UNIV PRESS INC.

Weidenmüller, A., Chen, R., & Meyer, B. (2019). Reconsidering response threshold models—short-term response patterns in thermoregulating bumblebees. Behavioral Ecology and Sociobiology, 73(8), 1-13.

Wheeler, D. E. (1991). The developmental basis of worker caste polymorphism in ants. The American Naturalist, 138(5), 1218-1238.

Wu, A. S., Mathias, D. H., Giordano, J. P., & Hevia, A. (2020, May). Effects of response threshold distribution on dynamic division of labor in decentralized swarms. In The Thirty-Third International Flairs Conference.

Yamanaka, O., Shiraishi, M., Awazu, A., & Nishimori, H. (2019). Verification of mathematical models of response threshold through statistical characterisation of the foraging activity in ant societies. Scientific reports, 9(1), 1-8.

**Appendix 1: Construction of genetic algorithm**

The optimization goal of this genetic algorithm is to produce a threshold distribution matrix 𝛉 which would allow a colony of N virtual ants to perform T tasks such that they keep the task associated stimulus for each task as close to an arbitrarily defined threshold as possible. If \hat{x}_{..}  is the average stimulus level across tasks and across time and x^* is the target, then the algorithm finds 𝛉 which minimizes the quadratic loss function L = ( x^* - \hat{x}_{..} )^2.

𝛉 is an N×T matrix which contains the response thresholds for each ant for each task. This matrix represents the behavioral chromosome of the algorithm, and while it functions as a matrix within the simulation, during the selection, mating, and mutation steps, this matrix is converted into a 1 dimensional array and then converted back into a matrix before the simulation is run. This array will be referred to as the chromosome. Response thresholds can vary continuously in the range [-5 5], so we use a continuous genetic algorithm.

The initial population of 1,000 chromosomes are created by drawing from a uniform distribution between -5 and 5. We then run simulations to estimate the loss of each chromosome. The 4 chromosomes that minimize L the most are then mated with one other randomly selected survivor to produce the next generation of 1,000 chromosomes. The parents are directly transferred to the next generation, but the other 996 chromosomes are produced by crossing over the chromosomes of the parents at 3 randomly chosen sites. Each new response threshold pnew is blended with parent a’s threshold, pa and parent b’s threshold, pb with the formula p_{new} = \beta p_a + (1-\beta)p_b where β is a randomly selected value between 0 and 1 (Radcliff, 1991). After crossover, each pnew in each chromosome outside of the parents has a 20% chance of mutating. If a response threshold mutates, then it is replaced with a new random value between -5 and 5. Loss is then estimated for each of these new chromosomes. This process is repeated for 2,500 generations. The best performing chromosomes and their associated loss values are stored across generations and is used to produce the final 𝛉.