Individual variations lead to universal and cross-species patterns of social behavior

Sang Hyun Choi[a,b], Vikyath D. Rao[a,b], Tim Gernat[b,c], Adam R. Hamilton[b], Gene E. Robinson[b,d,e], and Nigel Goldenfeld[a,b,1]

[a]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801; [b]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801; [c]Swarm Intelligence and Complex Systems Group, Department of Computer Science, Leipzig University, 04109 Leipzig, Germany; [d]Neuroscience Program, University of Illinois at Urbana-Champaign, Urbana, IL 61801; and [e]Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Contributed by Nigel Goldenfeld, October 16, 2020 (sent for review February 4, 2020; reviewed by Albert-László Barabási and Iain D. Couzin)

The duration of interaction events in a society is a fundamental measure of its collective nature and potentially reflects variability in individual behavior. Here we performed a high-throughput measurement of trophallaxis and face-to-face event durations experienced by a colony of honeybees over their entire lifetimes. The interaction time distribution is heavy-tailed, as previously reported for human face-to-face interactions. We developed a theory of pair interactions that takes into account individual variability and predicts the scaling behavior for both bee and extant human datasets. The individual variability of worker honeybees was nonzero but less than that of humans, possibly reflecting their greater genetic relatedness. Our work shows how individual differences can lead to universal patterns of behavior that transcend species and specific mechanisms for social interactions.

social network | individuality | heavy-tailed distribution

How individuals in a community interact with each other gives rise to collective emergent properties of the community (1–5). It reflects the individuals' personal preference, social roles, the external environment, and other numerous factors applicable to specific context. The distribution of interevent times, or waiting time between two consecutive events, for temporal social networks has been much studied because of its relation to information or disease spreading (3, 4). It has been shown that the heavy tail in the interevent time distribution is due to a decision-based queuing process, in which some tasks are more prioritized than others (6, 7). In contrast, the distribution of contact duration, instead of the interevent time, and its connection to the nature of social interactions have not been studied as much.

We have measured the duration of interactions among thousands of honeybees (*Apis mellifera*) in a hive, well-known eusocial insects that are easy to experimentally manipulate. Among many possible modes of honeybee social interaction, we focused on trophallaxis, which is mouth-to-mouth liquid food transfer. Trophallaxis occurs not only for feeding but also for communication (8, 9), making it a model system to study social interactions and collective effects (10, 11). To measure the interaction time, all of the honeybees in a colony were fitted with a barcode (12). A high-resolution machine vision camera imaged them at the rate of one frame per second. Then a customized algorithm detected each interaction event by analyzing the images and identified each bee, its position, and its orientation (12) (*Materials and Methods*).

Note that all of the honeybee data used in this work were originally generated by the authors for separate studies not for the purpose of testing our theory discussed in this paper or even to acquire the data needed for the theory. We used all of the data available to us, which were the trophallaxis social network data acquired in 2013 and analyzed in ref. 12 (1164_2013, 1140_2013, 1138_2013, 1174_2013, and 1170_2013, which are trials 1 to 5 in ref. 12) and the trophallaxis social network data acquired in 2016 from colonies with partial treatment with Juvenile Hormone analogue (13, 14) (789_JHA_2016 and 757_JHA_2016) (see *Materials and Methods* for more detail on the colony preparation). Our theory works for all of them, which indicates its robustness.

Fig. 1*A* shows that the distribution of honeybee interaction duration is heavy tailed. The exponents of the power law are –2.4, –2.3, –2.2, –2.2, –2.2, –2.7, –2.0, and –2.0 for each dataset listed in the legend from the top (1164_2013) to the bottom (757_JHA_2016). If every bee were the same and every interaction happened by chance, one might naively expect to see a peaked distribution such as a Gaussian. However, the observed heavy-tailed distribution suggests heterogeneity or variability among the population.

In order to improve the statistical power of our analysis, we also examined F2F events, where honeybees were close and oriented toward each other but not actually engaging in trophallaxis (15). F2F events occur about an order of magnitude more frequently than trophallaxis. F2F events include undetected trophallaxis and possible antennation, but the nature of the honeybee interaction during F2F is not as well defined as trophallaxis. Nevertheless, being F2F is a necessary but not sufficient condition for trophallaxis, so as long as the distance apart is not larger than the length of a honeybee, it would be expected that F2F events scale similarly to trophallaxis events. Indeed, coarse data for temporal networks retain some statistics of the actual interaction including heavy-tailedness of the contact duration distribution (16) (also confirmed by our results in Fig. 1*A*; see 1166_F2F_2013).

In our work, the detection of trophallaxis events is estimated using a two-step filtering process because we are not able to observe directly fluid transfer through the proboscis. Our

Significance

To probe the collective nature of social organisms, we measured the duration of food transfer (trophallaxis) and face-to-face events experienced by thousands of barcoded honeybees in a colony, over their entire lifetimes. The interaction time distribution is heavy-tailed, as previously reported for human face-to-face interactions. We develop a detailed theory for the universality of these findings and show that they reflect individual variations in the population. Using a measure of inequality from economics, we quantify these individual differences in both honeybees and humans. We show that bees have individual differences but exhibit less individuality than humans. This work demonstrates how individual variations can lead to universal patterns of behavior across different species and specific mechanisms for social interactions.

Author contributions: S.H.C., G.E.R., and N.G. designed research; S.H.C., V.D.R., T.G., A.R.H., and N.G. performed research; S.H.C., V.D.R., and N.G. analyzed data; and S.H.C. and N.G. wrote the paper.

Reviewers: A.-L.B., Northeastern University; and I.D.C., Max Planck Society.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: nigel@illinois.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2002013117/-/DCSupplemental.

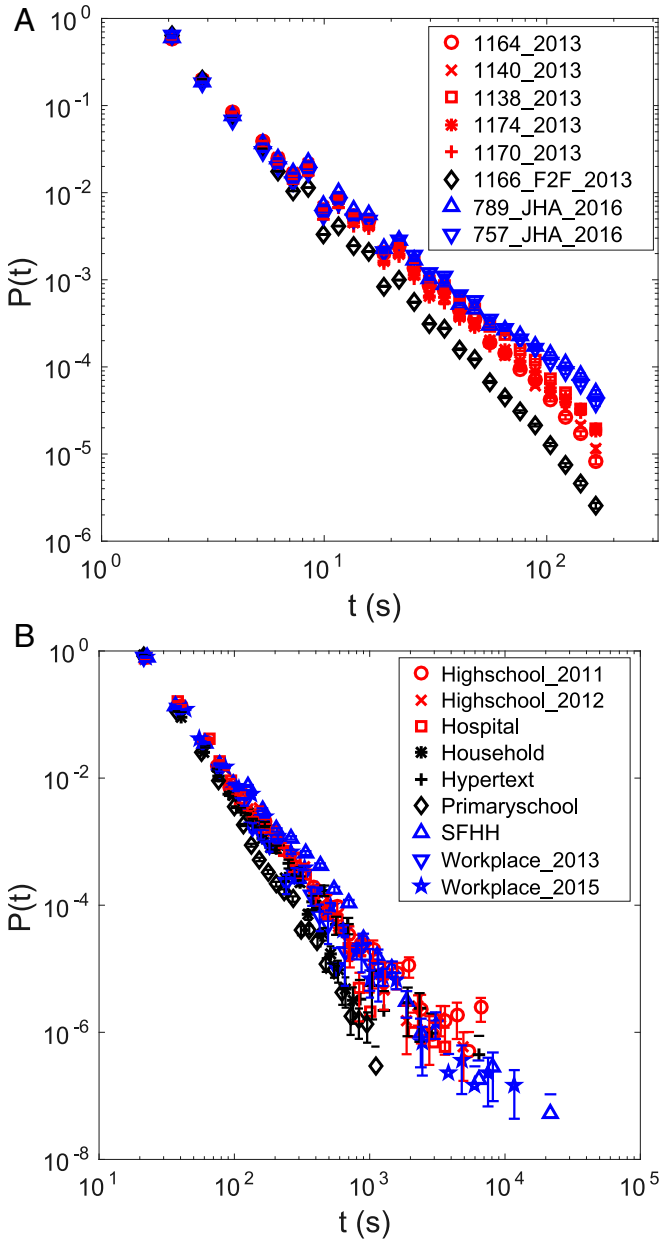

Fig. 1. (*A*) Duration of trophallaxis as well as F2F events that are not detected as trophallaxis follows a heavy-tailed distribution. The first number in the dataset name represents the number of observed honeybees in each colony, while the last number represents the year the experiment was performed. Note that the number of individuals in the last two datasets is actually smaller than what was originally observed (676 and 639 instead of 789 and 757, respectively) because we analyzed the data after applying an additional filter to exclude the time series during which the colonies were perturbed by the treatment of JHA, but we decided to keep the name of the datasets. The numbers of interactions used to make the plot are 302,221, 205,787, 191,795, 259,923, 329,170, 1,207,778, 136,529, and 115,965 in the order the datasets are listed in the legend. (*B*) Human face-to-face (F2F) interaction in various settings exhibits a heavy-tailed distribution. The numbers of interactions used are 10,677, 19,774, 14,037, 32,425, 9,865, 77,520, 26,039, 4,591, and 33,750 in the order the datasets are listed in the legend. In both *A* and *B*, error bars indicate the SE. In *B*, lower error bars for bins with count 1 could not be drawn in logarithmic scale because it extends to 0.

detection scheme is not subject to tracking error generated by bees that are misoriented or have other individual variations in visibility. Our method, described in detail in our earlier work (12), first selects bee pairs of close enough distance and orientation toward each other. Then we apply a second filter, selecting the bee pairs that are physically connected by proboscis through image processing. Through this second filter, we can filter out the pairs with inaccurate tracking of position or orientation because those pairs would not be connected by the proboscis. Even if there were innately poorly tracked bees, the multiple layers of filtering minimize the detection error on the trophallaxis events. There is quantitative evidence that there is minimal tracking error impact on our results, even without the second filter. The F2F dataset concerns bee pairs that satisfy only the geometric constraint regarding distance and orientation but are not detected as being connected by the proboscis. This dataset has not filtered out the pairs with inaccurate position and orientation through the second filter. However, as shown below, this F2F dataset gives the same power law and scaling laws as the trophallaxis datasets. In other words, the statistics of the data are retained regardless of whether the data have gone through the second filter, which is consistent with the notion that inaccuracy due to positional tracking error was already negligible after the first filter. Our earlier work (12) showed by subsampling of the trophallaxis interaction data at different sampling rates that the statistics of the trophallaxis interaction are robust against subsampling and false negatives, with the statistics of the times between trophallaxis events being robust to detection errors. This strongly suggests that the statistics of the duration of the trophallaxis events will also be robust.

We compared the honeybee data with human data recorded by the SocioPatterns collaboration (16–24) to explore whether there are universal patterns of social interaction. Fig. 1*B* shows that human F2F interaction time in various settings also exhibits a heavy-tailed distribution. The exponents of the power law are –2.4, –2.7, –2.9, –3.4, –2.7, –3.6, –2.5, –2.9, and –2.6 for each dataset listed in the legend from the top (Highschool_2011) to the bottom (Workplace_2015). Such similarity across different systems indicates an unexpected universality governing the interaction in social systems and suggests that a minimal model (25) should be able to capture the salient features of the interactions.

To construct such a minimal model, we treat the social bond between bees as an effective particle. We focus on bees here, but the model is also applicable to humans. The bond is the edge in the usual network representation of social interactions. This effective particle has two states representing an interacting pair and a noninteracting pair. The particle jumps from one state to the other with a fixed rate ω in the reaction coordinate, as depicted in Fig. 2. Although jumping happens in both directions, we focus on the jumping from the interacting state to the noninteracting state because the interaction time is the waiting time for the first jump in that direction. The distribution of the first jump time $f(t,\omega)$ is obtained by multiplying the probability not to jump at each time step until time t with the probability density to jump at time t. The first part, the probability not to jump until t, is $(1-\omega dt)^{t/dt}$, where dt is the interval for a time step. Taking the limit of infinitesimal time steps yields $\lim_{dt\to 0}(1-\omega dt)^{t/dt}=e^{-\omega t}$. The distribution of the first jump time is then given by $f(t,\omega)=\omega e^{-\omega t}$.

Some pairs tend to have longer interactions than others, and this is reflected in their value of ω. To take into account this variability or heterogeneity within the community, we integrate the jump time distribution for a fixed ω over the distribution of rates $p(\omega)$. To determine $p(\omega)$, we use the Kramers theory for escape over a potential barrier (26): the distribution of ω is related to the distribution of energy barrier heights E through $\omega=\omega_0 e^{-E}$ where ω_0 is a constant (26). Here we use a dimensionless energy scale E, which is already normalized with possible fluctuations.

We propose that the barrier height distribution $p(E)$ follows the extreme value distribution for maxima. As illustrated in Fig. 2, a given bee has multiple candidate partners with which to interact. Each possible pair is associated with a certain barrier height. The pair with the highest barrier would spend more time together because it is more difficult for the particle to escape.

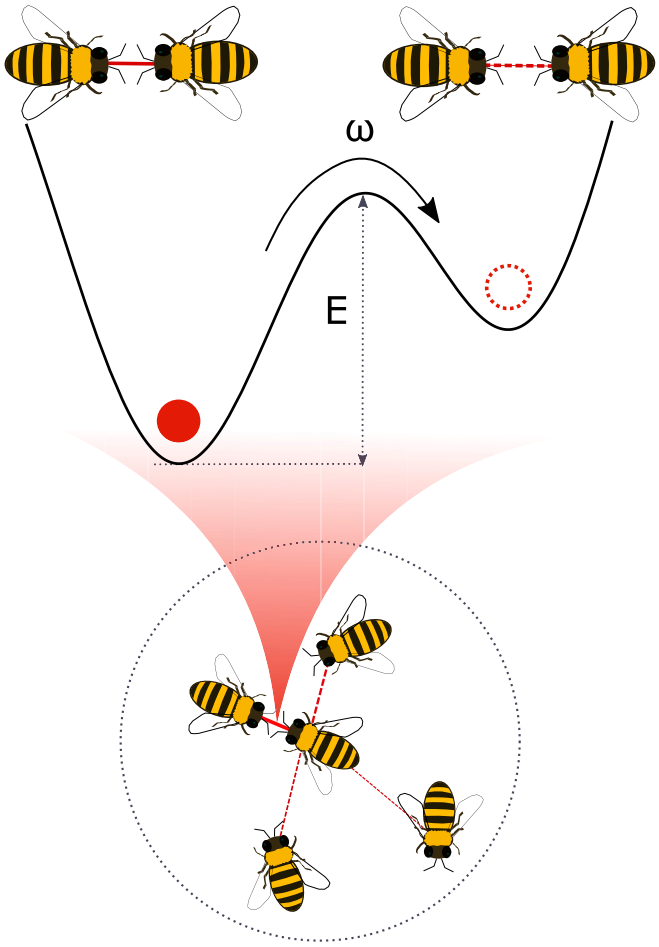

Fig. 2. Schematic picture for the theory of interaction durations. The dotted circular area indicates the neighborhood of the center bee. The neighbors, or the potential partners, inside the circle are connected with the center bee by a bond. The bond between two bees is represented by a particle in the reaction coordinate. This particle has two states representing an interacting pair and a noninteracting pair, and it changes its state by jumping over the energy barrier E with rate ω. Among the neighbors, the center bee is assumed to primarily interact with a bee with the highest barrier because they form the stablest pair.

This partner is thus interpreted as the most likeable and ends up forming a pair. Then the observed energy barrier E is the maximum among the neighbors. The distribution of the maximum is taken to be the Fisher–Tippett–Gumbel distribution for maxima (27, 28), which is appropriate if the parent distribution for barrier heights is localized, as seems reasonable. Then we express $p(E) = (\alpha+1)e^{-(\alpha+1)E}e^{-e^{-(\alpha+1)E}}$ where α is an undetermined parameter. For large E, $p(E) \sim e^{-(\alpha+1)E}$. We take the large E limit because the heavy-tailed behavior is observed at large t. Using $\omega = \omega_0 e^{-E}$, we find that $p(\omega) \sim \omega^{\alpha}$ for small ω, equivalent to large E. Combining this $p(\omega)$ with the exponential pair interaction time distribution, we get

$$f(t) = \int d\omega p(\omega) f(\omega, t) \sim t^{-(2+\alpha)} \qquad [1]$$

as the interaction time distribution for the population. The power law form suggests that the assumption about the parent distribution for barrier heights is valid. More details of this calculation are provided in *SI Appendix*. The quantity α in the exponent connects the community interaction time distribution $f(t)$ with the distribution of barrier heights $p(E)$.

We remark that a similar derivation for the heavy-tailed time distribution as shown in Eq. **1** arises in the theory for defect jumping (29) and a model of traps (30, 31) in glass. However, the interpretation of $p(E)$ in social interactions is different from the analogue in disordered materials. Atoms in a glass successively hop over multiple energy barriers in a rough potential landscape, so the integration over $p(E)$ is an average over the energy barriers experienced by one atom. On the other hand, our particles for the bond between a pair jump over one energy barrier to change their state. Thus, the integration over $p(E)$ is an ensemble average over the population.

Next we turn to verifications of the predictions of this theory. The simple theory predicts an exponential pair interaction time distribution. The quantile–quantile plots for pair interaction times (Fig. S1) suggest that the pair interaction time distributions for both honeybees and humans are better expressed by hyperexponential distributions, which are weighted sums of two exponential distributions. The theory is not affected by this additive modification, as discussed in *SI Appendix*. There are so many pairs in each colony that it is not practical to show the goodness of hyperexponential fit for each pair separately. Therefore, we devised a data collapse to show the fit of all pairs in one figure. Only the pairs that yield more than seven points of evaluation for the empirical cumulative distribution function (ECDF) were considered. The cumulative distribution function (CDF) for a hyperexponential function is $Y(t) = 1 - ge^{-\omega_1 t} - (1-g)e^{-\omega_2 t}$, where g is the weight and ω_1 and ω_2 are the rates for each exponential. We define a new variable $z \equiv \omega_1 t$ and rewrite the CDF as $F(z, g, \omega_1, \omega_2) \equiv (1/g)\left(1 - Y - (1-g)e^{-(\omega_2/\omega_1)z}\right) = e^{-z}$, where Y is ECDF. Then the x axis only depends on one variable z. If the data are well fitted by this functional form, plotting $F(z, g, \omega_1, \omega_2)$ against e^{-z} should produce a cloud of data points aligning with the $y = x$ reference line.

Fig. 3 *A* and *B* show that most honeybee and human pair interaction times are well fitted to hyperexponential distributions. Fig. 3 *A* and *B*, *Insets*, show the fit of a pair to provide some intuition of the fitting process. The fitted CDF tends to deviate more at small e^{-z}, or large t, because the CDF value of the fitting function approaches 1 for $t \to \infty$, while the ECDF value is 1 at the longest observed interaction time.

A second prediction from the model is the exponential barrier height distribution. Although E is not a directly measurable variable, the relation $\omega = \omega_0 e^{-E}$ enables us to indirectly measure $p(E)$ because the mean pair time associated with an energy barrier is $\tau = 1/\omega$. The relation $p(E) \sim e^{-(\alpha+1)E}$ implies that $p(\tau) \sim \tau^{-(2+\alpha)}$ for large τ, which has the same exponent as $f(t)$ in Eq. **1**. Therefore, comparing the exponent of the tail of $f(t)$ and $p(\tau)$ provides a test of the theory, in particular, the proposed functional form of $p(E)$.

Fig. 3 *C* and *D* demonstrate the same scaling between the tail of $f(t)$ and $p(\tau)$ for 1164_2013 and Primaryschool, respectively. The comparison of scaling for seven other honeybee datasets and eight other human datasets is shown in *SI Appendix*, Fig. S4. Here τ is obtained from fitting parameters ω, not from averaging of pair interaction times, because τ is the mean pair interaction time associated with a single energy barrier (*SI Appendix*). If we retain the full form for $p(E)$, i.e., including the superexponential term in the Fisher–Tippett–Gumbel distribution, $p(\tau)$ is expected to have a peak at small τ, which may explain the peak in Fig. 3*C*.

One might think that the identical scaling between $f(t)$ and $p(\tau)$ is a consequence of the so-called stable law (32) because τ is an average of t for pairs. Then, depending on the tail of $f(t)$, the distribution of τ would be given either by the central limit theorem, i.e., a Gaussian, or by a power law with the same exponent as $f(t)$. However, this is not correct, as explained in *SI Appendix*, because the parent distribution of $p(\tau)$ is not $f(t)$ but instead is the pair interaction time distribution $f(\omega, t)$.

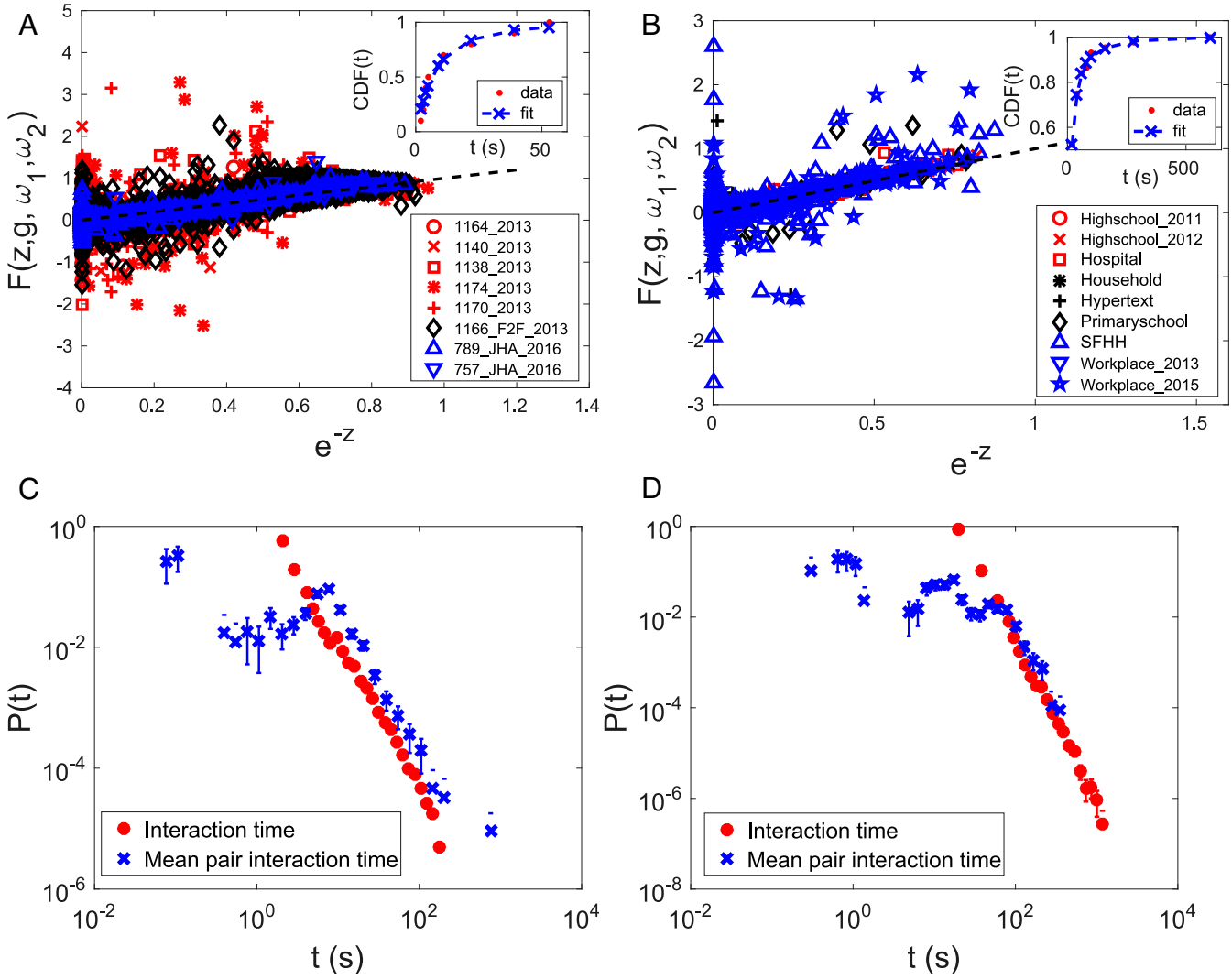

Fig. 3. (*A*) Pair interaction time distributions for honeybee pairs are fitted to hyperexponential functions and collapsed together. The numbers of pairs used to generate the plot are 197, 99, 328, 443, 561, 1,806, 46, and 20 in the order the datasets are listed in the legend. The pair interaction time distributions of these pairs were fitted. (*Inset*) Fitting ECDF of pair interaction times of a pair from 1164_2013. (*B*) The same plot as *A* but with human pairs. The numbers of pairs used are 37, 82, 59, 58, 39, 143, 98, 15, and 171 in the order the datasets are listed in the legend. (*Inset*) Fitting ECDF of pair interaction times of a pair from Highschool_2011. (*C*) Comparison between the scaling of interaction time distribution and mean pair interaction time distribution for 1164_2013. (*D*) The same plot as *C* but for Primaryschool. The number of mean pair interaction times used is the same as the number of pairs used for fitting, which is listed in *A* and *B*. The number of interaction times used is the same as what is listed in Fig. 1. Error bars indicate the SE, and lower error bars for bins with count 1 could not be drawn in logarithmic scale because it extends to 0.

The two energy barriers suggested by the hyperexponential pair interaction time distribution imply a multidimensional potential landscape of the reaction coordinate. Our model does not limit the number of barriers, allowing the pair interaction time distribution in principle to be an arbitrary sum $\sum_i g_i \omega_i e^{-\omega_i t}$, but the weight of further barriers seems to be too small to contribute to the dynamics. It is evident that different pairs are characterized by different barrier heights, but whether it is a specific pair or a specific individual that determines the barrier height cannot be determined by the analysis so far.

To explore the effect of individuality in social interactions, we calculated the Gini coefficient (33) for 1) the total interaction time spent by each individual, 2) the total number of interactions each individual had, and 3) the total number of partners with which each individual interacted. Widely used to express inequality in economics, the Gini coefficients have recently been used to quantify inequality in the activity level of eusocial insects (34, 35). Fig. 4 shows a graphical representation of the results, known as the Lorenz plot (36), for the total interaction time spent by bees and humans. The Lorenz plots for other variables are shown in *SI Appendix*, Fig. S5.

To read these results, note that in a Lorenz plot the greater the deviation from the $y = x$ reference line, the closer the Gini coefficient to unity, thus indicating a greater level of inequality. More inequality in our data means a greater contribution by individuals to the dynamics, signifying the effect of individuality. Fig. 4*A* shows that the Gini coefficients for honeybees are in the range 0.2 to 0.3, whereas for humans they are in the range 0.3 to 0.5. Thus, although individual bees are distinct, they are not as different from each other as humans are (Fig. 4*B*). The Lorenz plots and Gini coefficients for the total number of interactions and total number of partners provided in Fig. S5 show the same trend. The reduced individuality in honeybees compared to humans might be due to the average coefficient of relatedness being $r = 0.75$ among workers of the same colony as the queen was inseminated with a single male in these experiments, but further study is needed to verify this conjecture. Since the interaction time is a shared value between a pair, it is nontrivial to

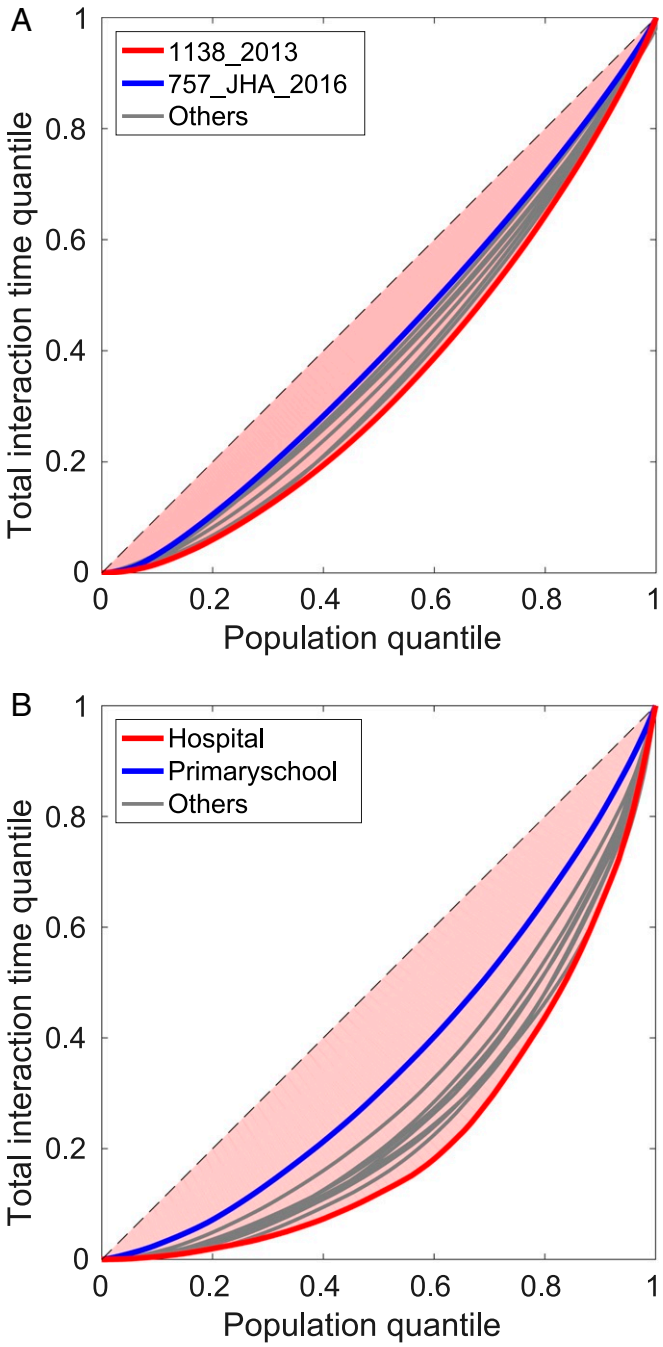

Fig. 4. Lorenz plots of the total time spent for interaction by honeybees and humans. (*A*) Gini coefficients of bees are as follows: 1164_2013, 0.2373; 1140_2013, 0.2111; 1138_2013, 0.3013; 1174_2013, 0.2760; 1170_2013, 0.2698; 1166_F2F_2013, 0.2089; 789_JHA_2016, 0.1941; and 757_JHA_2016, 0.1727. The numbers of data points used in the plot are the same as the numbers of individuals in each dataset, which are 1,164, 1,140, 1,138, 1,174, 1,170, 1,166, 676, and 639. (*B*) Gini coefficients of humans are as follows: Highschool_2011, 0.4333; Highschool_2012, 0.4879; Hospital, 0.5488; Household, 0.5012; Hypertext, 0.4576; Primaryschool, 0.2799; SFHH, 0.4937; Workplace_2013, 0.4493; and Workplace_2015, 0.3753. The numbers of individuals used are 126, 180, 75, 75, 113, 242, 403, 92, and 217.

completely separate the contribution of individuals. The effect of individuality in social interactions is therefore an open question but one that we have provided the tools to explore. Nevertheless, along with earlier studies of possible chemosensory mechanisms for individual identification (37), our results provide confirmation and quantification of the conjecture from recent work on the personality of honeybee workers as described in ref. 38 that some individuals are more likely to be interactive and engaged in food sharing, while others are less so.

The recently discovered heterogeneous food distribution in the *Camponotus sanctus* ant colony (39) may suggest individual variations in workers of this other well-known eusocial insect. Although the ratio of transferred food volume to maximal transferable volume during trophallaxis when the donor is a forager is measured to follow the same exponential distribution with the same parameter as the case when the donor is a nonforager (39), it does not necessarily mean the lack of individuality in ants because the individual variations may have been averaged out as the data of many pairs were analyzed collectively. If the data were analyzed for each pair, individual variations may have been observed. It is not the scope of our work, but it would be possible to study the effect of individuality on the food mixing due to trophallaxis of eusocial insects.

We have shown that high-resolution tracking can yield detailed multiscale information about the interactions and behavior of individuals within a community. Our results suggest that individual differences can lead to patterns of behavior that are universal and transcend species, context, and specific mechanisms for social interactions.

Materials and Methods

Animals. This All experiments took place at the University of Illinois Bee Research Facility, Urbana, Illinois. Experiments for 1164_2013, 1140_2013, 1138_2013, 1174_2013, and 1170_2013 were performed from July to October 2013. Dataset 1166_F2F_2013 was obtained from the same colony as 1164_2013. Each experiment started with a colony of 1,200 one-day-old adult worker honeybees (*A. mellifera*), but the activities of 1,164, 1,140, 1,138, 1,174, 1,170, and 1,166 bees were detected. Datasets 789_JHA_2016 and 757_JHA_2016 were collected in September and October 2016. Each experiment in 2016 started with a colony of 800 one-day-old adult worker honeybees, but the activities of 789 and 757 bees were detected. After filtering out the period of perturbation caused by the treatment of JHA, we detected the activities of 676 and 639 bees. The details of the experiments in 2013 are described in ref. 12, and here we portray the experiments in 2016 (13). The identification of each honeybee and the detection of trophallaxis were done in the same way as the 2013 experiments (12). The datasets from the 2016 experiments were collected as a part of a different study (13), but no difference with regard to the scaling law exposed in the paper was detected.

To minimize the effects of genetic variation on behavioral analyses, all experiments used adult worker bees from colonies headed by queens that had each been instrumentally inseminated with semen from a single drone. Due to haplodiploidy, this results in an average coefficient of relatedness of 0.75. Honeycomb frames containing pupae were removed from colonies 1 to 2 d prior to the beginning of each experiment and maintained in a dark incubator at 34°C and 50% relative humidity (13).

Observation Hive Construction and Monitoring. Eight hundred 0- to 24-h-old adult workers were cold anaesthetized and tagged (12). The bees were then placed with a barcoded, naturally mated queen into a glass-walled observation hive containing a single one-sided honeycomb frame. Colonies were provided with 300 g of honey and 30 g of artificial bee bread (nine parts honey, nine parts ground pollen, and two parts water). To prevent bees from spreading the honey onto the glass observation window, cells containing honey were covered in wax foundation scented with 5 mL of strawberry or orange food extract (McCormick & Company Inc.) to establish hive identity, thereby facilitating homing during orientation flights and foraging and preventing intracolony conflict during pharmacological treatments (*Juvenile Hormone Analogue Treatments*). Observation hives were housed in a dark room heated to 32°C and maintained at 50% humidity. Images were captured at a rate of one frame per second over the course of the entire experiment; infrared lighting illuminated the colony during periods of image capture from both the front and rear of the hive (via a blacklight) (12). Starting on the second night, the colony entrance was opened to allow the bees access to the exterior environment (12). The glass window of the observation hive was changed twice daily (ca. 9 AM and 8 PM) to prevent debris from obscuring the colony, but the hives were otherwise left undisturbed except during pharmacological treatments (*Juvenile Hormone Analogue Treatments*).

The barcode and trophallaxis identification was performed as detailed in ref. 12. During the trophallaxis detection, potential interaction partners were first selected according to the geometric constraint. First, the distance between the most anterior point of each bee's body on the anteroposterior axis was measured, and a bee pair was selected if the distance was between 3.4 and 6.9 mm. Then the selected pair was considered as the potential interaction partners if the angle between each bee's barcode orientation vector was between 90° and 180°. Among the potential partners, the pairs physically connected by the proboscis were detected as doing trophallaxis. On the other hand, the pairs satisfying the geometric constraint but not being connected by the proboscis were classified as F2F. The analyzed F2F dataset used close parameters of the geometric constraint. Identical parameters were used in trophallaxis detection for the 2013 experiment (12), and the more detailed description of the detection procedure is in ref. 12.

Juvenile Hormone Analogue Treatments. After 5 d of observation, colonies were dismantled, and the bees were placed in Plexiglas cages. The glass observation window was replaced with a transfer glass with removable portholes that could be centered over each sector, and the colony was placed in an area lit by a far-red LED light (Smart Vision Lights). Since far-red wavelengths are barely perceptible to bees, this allowed them to be transferred individually to the Plexiglas cage via forceps without risk of escape. Bees were then cold anaesthetized in groups of three and randomly assigned to one of three treatments: a Juvenile Hormone analogue (JHA) treatment, an acetone vehicle treatment, and a cold-anaesthetized sham treatment (13). Bees in the JHA group were treated topically with the JHA methoprene [200 ng per bee, a dose that is known to induce precocious foraging in bees (40)] dissolved in 1 μL of acetone and administered to the abdomen. The vehicle group was treated with 1 μL of acetone administered to the abdomen, and the sham treatment group was cold anaesthetized but otherwise untreated. The bees were also painted with a color code (Testor's Corporation) designating their treatment and photographed, allowing their barcode identification to be linked to their treatment (13). After all of the bees in a colony recovered from anaesthetization, they were returned to their observation hive. Because the colony was physically disturbed, trophallaxis observations from the day of the experiment were not used in any downstream analyses. Monitoring of the colony continued for 7 d after the JHA treatments, after which the colony was disassembled. More details on the JHA treatment and related experiments are in ref. 13.

Code Availability

The custom software to produce printable barcode images, detect barcodes in digital images, and detect trophallaxis is available at https://github.com/gernat/btools. The MATLAB codes used to analyze the data are available at https://github.com/schoi8/heavytail_scaling. The MATLAB code used to calculate the Gini coefficients and to generate the Lorenz plots is available at https://www.mathworks.com/matlabcentral/fileexchange/28080-gini-coefficient-and-the-lorentz-curve.

Data Availability. The honeybee interaction datasets from the 2013 experiments (12) analyzed in this study are available at http://www.beemonitoring.igb.illinois.edu. The F2F dataset is available at the Illinois Data Bank with the identifier https://doi.org/10.13012/B2IDB-4021786_V1. The honeybee interaction datasets from the 2016 experiments (13) analyzed in this study are available at the Illinois Data Bank with the identifier https://doi.org/10.13012/B2IDB-2712449_V1. The human F2F interaction datasets from the SocioPatterns Collaboration are available at http://www.sociopatterns.org/datasets/.

ACKNOWLEDGMENTS. This work was partially supported by National Institutes of Health Grant R01GM117467.

1. D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
2. A. L. Barabási, R. Albert, Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
3. A. L. Barabási, *Network Science* (Cambridge University Press, 2016).
4. P. Holme, J. Saramäki, Temporal networks. *Phys. Rep.* **519**, 97–125 (2012).
5. S. B. Rosenthal, C. R. Twomey, A. T. Hartnett, H. S. Wu, I. D. Couzin, Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4690–4695 (2015).
6. A. L. Barabási, The origin of bursts and heavy tails in human dynamics. *Nature* **435**, 207–211 (2005).
7. A. Vázquez *et al.*, Modeling bursts and heavy tails in human dynamics. *Phys. Rev. E* **73**, 036127 (2006).
8. M. L. Winston, *The Biology of the Honey Bee* (Harvard University Press, 1991).
9. S. Jarau, M. Hrncir, *Food Exploitation by Social Insects: Ecological, Behavioral, and Theoretical Approaches* (CRC Press, 2009).
10. J. Gräwer, H. Ronellenfitsch, M. G. Mazza, E. Katifori, Trophallaxis-inspired model for distributed transport between randomly interacting agents. *Phys. Rev. E* **96**, 022111 (2017).
11. E. Greenwald, E. Segre, O. Feinerman, Ant trophallactic networks: Simultaneous measurement of interaction patterns and food dissemination. *Sci. Rep.* **5**, 12496 (2015).
12. T. Gernat *et al.*, Automated monitoring of behavior reveals bursty interaction patterns and rapid spreading dynamics in honeybee social networks. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 1433–1438 (2018).
13. A. R. Hamilton, "Endocrine, transcriptomic and social regulation of division of labor in honey bees," PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois (2018).
14. S. H. Choi *et al.*, Honeybee trophallaxis event data for the origin of heavy tails in honeybee and human interaction times. Illinois Data Bank. https://doi.org/10.13012/B2IDB-2712449_V1. Deposited 14 October 2019.
15. S. H. Choi *et al.*, Honeybee F2F event data for the origin of heavy tails in honeybee and human interaction times. Illinois Data Bank. https://doi.org/10.13012/B2IDB-4021786_V1. Deposited 3 October 2019.
16. M. Génois, A. Barrat, Can co-location be used as a proxy for face-to-face contacts? *EPJ Data Sci.* **7**, 11 (2018).
17. SocioPatterns Collaboration (2018). http://www.sociopatterns.org/datasets/. Accessed 29 May 2019.
18. J. Fournet, A. Barrat, Contact patterns among high school students. *PLoS One* **9**, e107878 (2014).
19. P. Vanhems *et al.*, Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLoS One* **8**, e73970 (2013).
20. M. C. Kiti *et al.*, Quantifying social contacts in a household setting of rural Kenya using wearable proximity sensors. *EPJ Data Sci.* **5**, 1–21 (2016).
21. L. Isella *et al.*, What's in a crowd? Analysis of face-to-face behavioral networks. *J. Theor. Biol.* **271**, 166–180 (2011).
22. J. Stehlé *et al.*, High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS One* **6**, e23176 (2011).
23. V. Gemmetto, A. Barrat, C. Cattuto, Mitigation of infectious disease at school: Targeted class closure vs school closure. *BMC Infect. Dis.* **14**, 695 (2014).
24. M. Génois *et al.*, Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers. *Network Sci.* **3**, 326–347 (2015).
25. N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (CRC Press, 2018).
26. H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–360 (1940).
27. R. A. Fisher, L. H. C. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cambridge Phil. Soc.* **24**, 180–190 (1928).
28. E. J. Gumbel, Les valeurs extrêmes des distributions statistiques. *Ann. Inst. Henri Poincaré* **5**, 115–158 (1935).
29. J. T. Bendler, M. F. Shlesinger, Generalized Vogel law for glass-forming liquids. *J. Stat. Phys.* **53**, 531–541 (1988).
30. T. Odagaki, Glass transition singularities. *Phys. Rev. Lett.* **75**, 3701–3704 (1995).
31. C. Monthus, J. P. Bouchaud, Models of traps and glass phenomenology. *J. Phys. A: Math. Gen.* **29**, 3847–3869 (1996).
32. B. V. Gnedenko, A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, 1954).
33. C. Gini, "Variabilità e mutabilità (variability and mutability)" in *Memorie di Metodologica Statistica*, E. Pizetti, T. Salvemini, Eds. (Libreria Eredi Virgilio Veschi, Rome, 1955).
34. P. Tenczar, C. C. Lutz, V. D. Rao, N. Goldenfeld, G. E. Robinson, Automated monitoring reveals extreme interindividual variation and plasticity in honeybee foraging activity levels. *Anim. Behav.* **95**, 41–48 (2014).
35. J. Aguilar *et al.*, Collective clog control: Optimizing traffic flow in confined biological and robophysical excavation. *Science* **361**, 672–677 (2018).
36. M. O. Lorenz, Methods of measuring the concentration of wealth. *J. Am. Stat. Assoc.* **9**, 209–219 (1905).
37. W. M. Getz, R. E. Page, Jr, Chemosensory kin-communication systems and kin recognition in honey bees. *Ethology* **87**, 298–315 (1991).
38. A. Walton, A. L. Toth, Variation in individual worker honey bee behavior shows hallmarks of personality. *Behav. Ecol. Sociobiol.* **70**, 999–1010 (2016).
39. E. Greenwald, J. P. Eckmann, O. Feinerman, Colony entropy—Allocation of goods in ant colonies. *PLoS Comput. Biol.* **15**, e1006925 (2019).
40. G. E. Robinson, Regulation of honey bee age polyethism by juvenile hormone. *Behav. Ecol. Sociobiol.* **20**, 329–338 (1987).

PNAS
www.pnas.org

Supplementary Information for

Individual variations lead to universal and cross-species patterns of social behavior

Sang Hyun Choi, Vikyath D. Rao, Tim Gernat, Adam Hamilton, Gene E. Robinson, Nigel Goldenfeld

Corresponding Author: Nigel Goldenfeld
E-mail: nigel@uiuc.edu

This PDF file includes:

Supporting Information Text

1. Quantile-quantile plots of the pair interaction time distribution

Our model predicts the pair interaction time distribution to be exponential. We have generated the quantile-quantile (QQ) plots of pair interaction times of honeybees and humans to verify the prediction. If the data followed the proposed probability density, a QQ plot would show a linear line. QQ plot was meant to be a preliminary test to examine the probability density of the pair interaction times, so only a few pairs from each dataset have been tested. The pair with the most number of interactions from each dataset has been chosen to be demonstrated here for clearer examination of the distribution.

As shown in Fig. S1, although some pair interaction times appear to be exponential, there are data that deviate for long times. The deviation is always upward, which implies that it is not due to statistical fluctuations but is instead a systematic effect. The upward deviation indicates that the actual probability density decays more slowly than the proposed one. However, even the deviating data align well with the reference line for short times. The deviating region at long times also looks linear although it does not lie on the reference line. The linearity suggests that the deviating region is exponential as well, but with a slower decaying rate than the short time region. A weighted sum of two exponential distributions, sometimes known as a hyperexponential distribution, exhibits the same behavior. Its limiting behavior at small values of the argument is the faster decaying of the two summed exponentials whilst its limiting behavior at large values is the other one. Therefore, we deduce that the pair interaction time distribution is hyperexponential, and is well approximated by just two terms.

2. Calculation of the interaction time distribution.

As discussed in the main text, the probability density of the rates is $p(\omega) \sim \omega^{\alpha}$. Integrating the pair interaction time distribution over the rates gives

$$f(t) = \int_0^{\epsilon} d\omega p(\omega) f(\omega, t) = \int_0^{\epsilon} d\omega \omega^{\alpha+1} e^{-\omega t} = t^{-(2+\alpha)} \gamma(2+\alpha, \epsilon t) \qquad [S1]$$

where $\gamma(a, z) \equiv \int_0^z t^{a-1} e^{-t} dt$ is the incomplete gamma function. Since $p(\omega) \sim \omega^{\alpha}$ is for small ω limit, the integration is performed up to a finite value ϵ. $\gamma(a, z)$ approaches a constant $\Gamma(a) \equiv \int_0^{\infty} t^{a-1} e^{-t} dt$ as $z \to \infty$ (a is fixed in our case), which means that equation S1 would become $f(t) \sim t^{-(2+\alpha)}$ for large ϵt. $\gamma(a, z)$ starts to saturate to a constant value at z of order 1. Therefore, the large ϵt condition translates into $t > 1/\epsilon$. We assume ϵ to be small, but it is the largest value of ω in the population (equation S1). So the condition $t > 1/\epsilon$ is satisfied for most t, and we observe the heavy-tailed distribution $f(t) \sim t^{-(2+\alpha)}$.

The pair interaction time distribution turns out to be better represented by a hyperexponential distribution than an exponential distribution. However, this modification does not affect the final result $f(t) \sim t^{-(2+\alpha)}$. Now the pair interaction time distribution is $f_{pair}(t) = g\omega_1 e^{-\omega_1 t} + (1-g)\omega_2 e^{-\omega_2 t}$ where g is a weight on one of the two exponential terms. Since we now have two rates ω_1 and ω_2 with relation $\omega_1 > \omega_2$, the integration becomes

$$f(t) = \int_0^{\epsilon} d\omega_1 p(\omega_1) \int_0^{\omega_1} d\omega_2 p(\omega_2) \left\{ g\omega_1 e^{-\omega_1 t} + (1-g)\omega_2 e^{-\omega_2 t} \right\}. \qquad [S2]$$

The first term is calculated as

$$\int_0^{\epsilon} d\omega_1 \omega_1^{\alpha} \frac{g}{\alpha+1} \omega_1^{\alpha+2} e^{-\omega_1 t} = \frac{g}{\alpha+1} t^{-(3+2\alpha)} \gamma(3+2\alpha, \epsilon t) \sim t^{-(3+2\alpha)}. \qquad [S3]$$

The last asymptotic relation is for large t. Here "large t" means $t > 1/\epsilon$ as stated in the previous paragraph. The first integral of the second term gives us $(1-g)t^{-(2+\alpha)}\gamma(2+\alpha, \omega_1 t)$. Then the second term is calculated as

$$(1-g)t^{-(2+\alpha)} \int_0^{\epsilon} d\omega_1 \omega_1^{\alpha} \gamma(2+\alpha, \omega_1 t) \sim t^{-(2+\alpha)} \qquad [S4]$$

for large t because $\gamma(2+\alpha, \omega_1 t)$ is essentially a constant for that limit. Therefore, the interaction time distribution for the whole population becomes

$$f(t) \sim At^{-(3+2\alpha)} + Bt^{-(2+\alpha)} \sim t^{-(2+\alpha)} \qquad [S5]$$

where A and B are constants. As shown in equation S5, the scaling of $f(t)$ does not change even after the modification of the pair interaction time distribution.

Although we observe only two energy barriers, hence a hyperexponential pair interaction time distribution, the theory in principle could allow the pair interaction time distribution to be an arbitrary sum of exponential distributions $\sum_i g_i \omega_i e^{-\omega_i t}$. The same calculation as above gives $f(t) \sim t^{-(2+\alpha)}$ in this case as well.

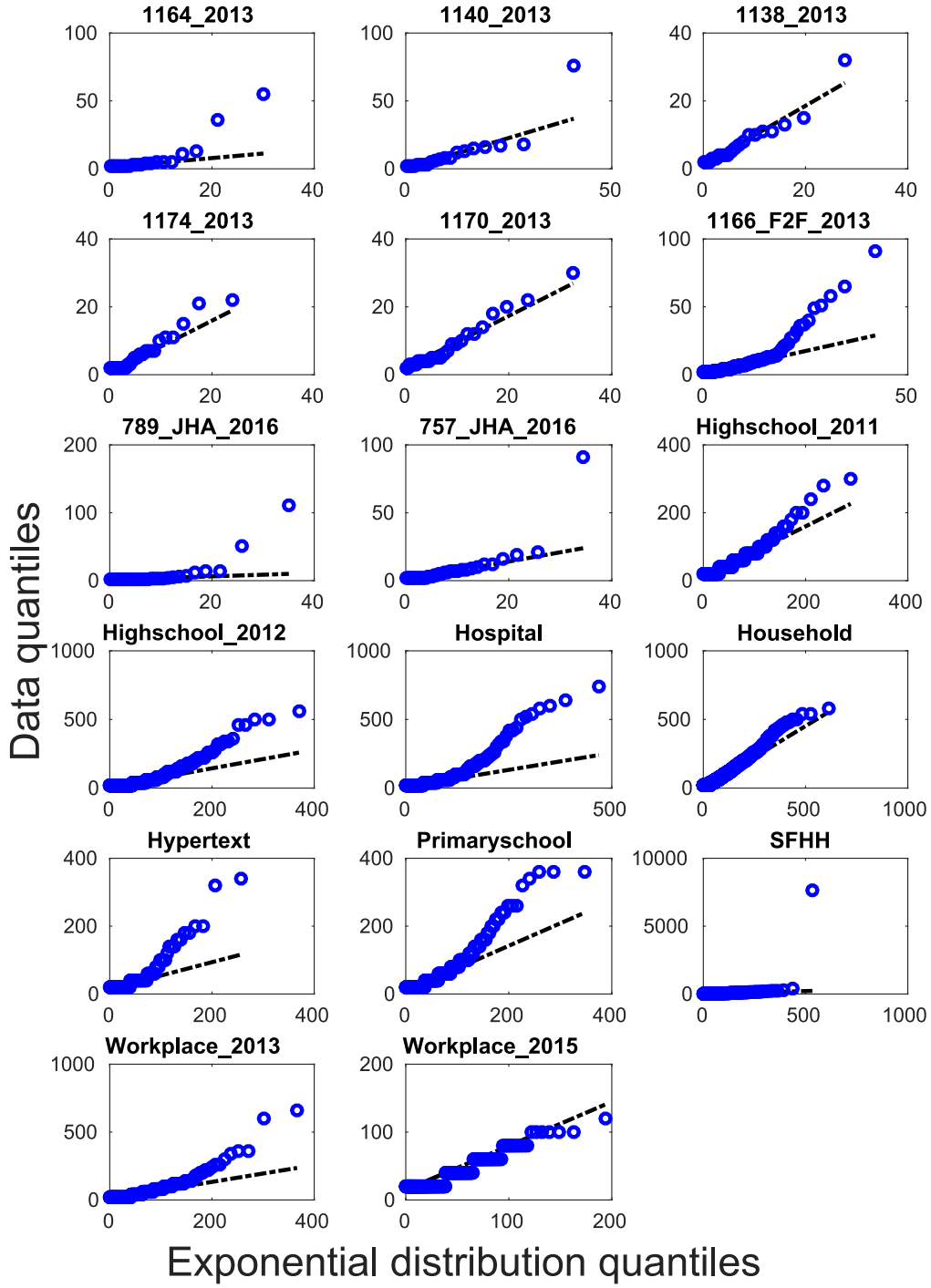

Fig. S1. Quantile-quantile plots of the most frequently-interacting pair from each honeybee and human dataset. If the data followed the proposed distribution, the data points would lie on the dashed reference line.

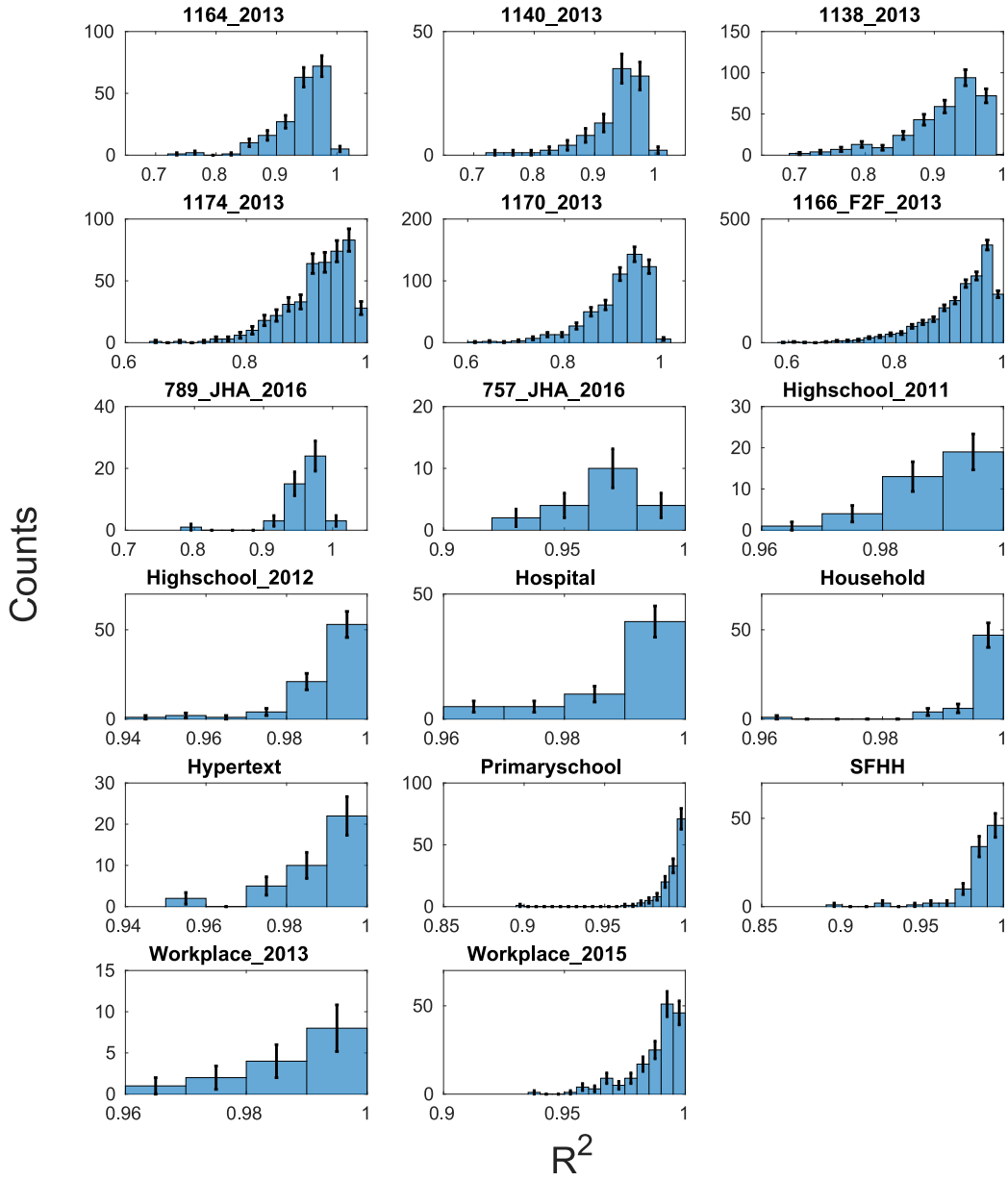

Fig. S2. Histogram of R^2 of fitting of the pair interaction time distribution to hyperexponential distribution for all honeybee and human datasets. For all figures, error bars indicate the standard error.

3. Fitting of the pair interaction time distribution to the hyperexponential distribution.

We have fitted the empirical cumulative distribution function (ECDF) of the pair interaction times to the cumulative distribution function (CDF) of a hyperexponential distribution $F_{pair}(t) = 1 - ge^{-\omega_1 t} - (1-g)e^{-\omega_2 t}$. Fig. S2 shows the histogram of R^2 values of the fitting. It is not a single value but a distribution of values because fitting has been done to each pair and R^2 has been collected from each pair.

As shown in Fig. S2, most values of R^2 are close to 1, which means that the pair interaction time distributions are well represented by a hyperexponential distribution.

Fig. S3 shows the histogram of weight g on the first exponential term in the hyperexponential distribution. Again it shows not a single value but a distribution of values because the fitting has been done for each pair.

The peak near 1 in Fig. S3a-f shows that one of the two exponential terms has a much higher weight than the other in these datasets. On the other hand, other datasets show more uniform distributions, signifying that both exponential terms, or both energy barriers, have non-negligible effect on many pairs. The broad distributions in Fig. S3i-q suggest that the multi-dimensionality of the potential landscape of social interaction is more prominent in humans than in honeybees. Yet, as shown in Fig. S3g and h, honeybee colonies with partial JHA treatment exhibit similar distributions of weights as humans. Despite the robust heavy tail in the interaction time distribution, Fig. S3 shows that the presence of honeybees with JHA treatment does influence the social interaction of honeybees. The precise way that it affects the honeybee social interactions needs further study and is beyond the scope of this work.

4. Exponential barrier height distribution.

Our model predicts that the barrier height distribution is the extreme value distribution for maxima. However, since the heavy tail at large time is dominated by high energy barriers, we take the large barrier height limit and obtain the exponential distribution $p(E) \sim e^{-(\alpha+1)E}$ where E is the energy barrier height in units of an effective energy scale that depends on details of social interactions that are outside the level of description of the present theory and α is some parameter. Using Kramers' rate formula $\omega = \omega_0 e^{-E}$ where ω_0 is some constant, we express the mean pair interaction time distribution as $p(\tau) \sim \tau^{-(2+\alpha)}$. It has the same scaling as the interaction time distribution $f(t) \sim t^{-(2+\alpha)}$. So by comparing the exponent of $f(t)$ and $p(\tau)$, we verify the exponential energy barrier height distribution.

Fig. S4 shows the scaling of the probability densities of interaction time $f(t)$, mean pair interaction time from fitting $p(\tau_{ij})$ and mean pair interaction time from averaging $p(\tau_i)$ for all honeybee and human datasets. Here the index i labels a pair, and the index j labels an energy barrier. τ_{ij} is the mean escape time for each energy barrier and is obtained from fitting of the pair interaction time distribution. With this index notation, CDF of a hyperexponential distribution to which we have fitted ECDF of the pair interaction time distribution is written as $F_i(t) = 1 - g_{i1}e^{-\omega_{i1}t} - g_{i2}e^{-\omega_{i2}t}$ where g_{ij} is the weight with $g_{i2} = 1 - g_{i1}$ and ω_{ij} is the rate of each exponential term. These g_{i1}, ω_{i1} and ω_{i2} are fitting parameters. We obtain two mean pair interaction times τ_{i1} and τ_{i2} for each pair i by taking the reciprocal of ω_{i1} and ω_{i2} respectively. On the other hand, τ_i is average of all interaction times for a pair. So $\tau_i = \frac{1}{n_{i1}+n_{i2}}\left(\sum_{k=1}^{n_{i1}} t_{i1k} + \sum_{k=1}^{n_{i2}} t_{i2k}\right)$ where t_{ijk} is the interaction time, the index k labels each interaction event, and n_{ij} is the number of interactions associated with each energy barrier j. In this notation, the mean pair interaction time from fitting is $\tau_{ij} = \frac{1}{n_{ij}}\sum_{k=1}^{n_{ij}} t_{ijk}$. So the relation between τ_{ij} and τ_i is written as $\tau_i = \sum_j g_{ij}\tau_{ij}$ where the weight is $g_{ij} \equiv n_{ij}/\sum_j n_{ij}$.

In Fig. S4, data points that extend to very large values (such as of order of 10^{14} seconds) were omitted because they are an artifact from fitting. These outliers correspond to one or two pairs out of all analyzed pairs in the colony. The artifact arises because the pair interaction time distribution for some pairs is better described by an exponential distribution than a hyperexponential distribution. In such a case, ω_{i2} takes a very small value close to 0 because $1 - g_{i1}e^{-\omega_{i1}t} - (1-g_{i1})e^{-\omega_{i2}t} \to g_{i1}(1 - e^{-\omega_{i1}t})$ for $\omega_{i2} \to 0$. Then taking the value of g_{i1} close to 1 gives us $1 - e^{-\omega_{i1}t}$ which is CDF of an exponential distribution. $\omega_{i2} \to 0$ corresponds to $\tau_{i2} \to \infty$, which explains the large τ_{ij}'s in some datasets. The pair interaction time distributions that yield such large τ_{ij} indeed have g_{ij} values close to 1 and are fitted well to exponential distributions.

For all datasets, $p(\tau_{ij})$ exhibits the same scaling as $f(t)$, verifying the exponential distribution of energy barriers. On the other hand, as shown in Fig. S4l and n, $p(\tau_i)$ doesn't show the same scaling for some datasets, and this feature is discussed in the next section.

5. The mean pair interaction time distribution and the stable law.

In this section, we discuss the difference between the two kinds of mean pair interaction time distributions $p(\tau_{ij})$ and $p(\tau_i)$ to emphasize that $p(\tau_{ij})$ is the right distribution to use. τ_{ij} represents the mean time needed to jump over an energy barrier j, while τ_i is the average of the whole pair interaction times. So the quantity associated with the energy barrier E through $\omega = \omega_0 e^{-E}$ is τ_{ij}. It would be more accurate to write the Kramers rate formula as $\omega_{ij} = \omega_0 e^{-E_{ij}}$ according to this notation. τ_i is the normalized linear sum of τ_{ij}'s.

It is very tempting to use τ_i instead of τ_{ij} because simply averaging the pair interaction times is much easier than fitting the pair interaction time distributions. Fig. S4 shows whether it matters which mean pair interaction time we use; while $p(\tau_i)$ has the same scaling as $p(\tau_{ij})$ and thus as the population interaction time distribution $f(t)$ for most datasets, it is not the case for Household (Fig. S4l) and Primaryschool (Fig. S4n). The deviation of $p(\tau_i)$ is clearer in Primaryschool (Fig. S4n) because of better statistics associated with the larger dataset. These two particular datasets differ from the rest by having $f(t)$ that

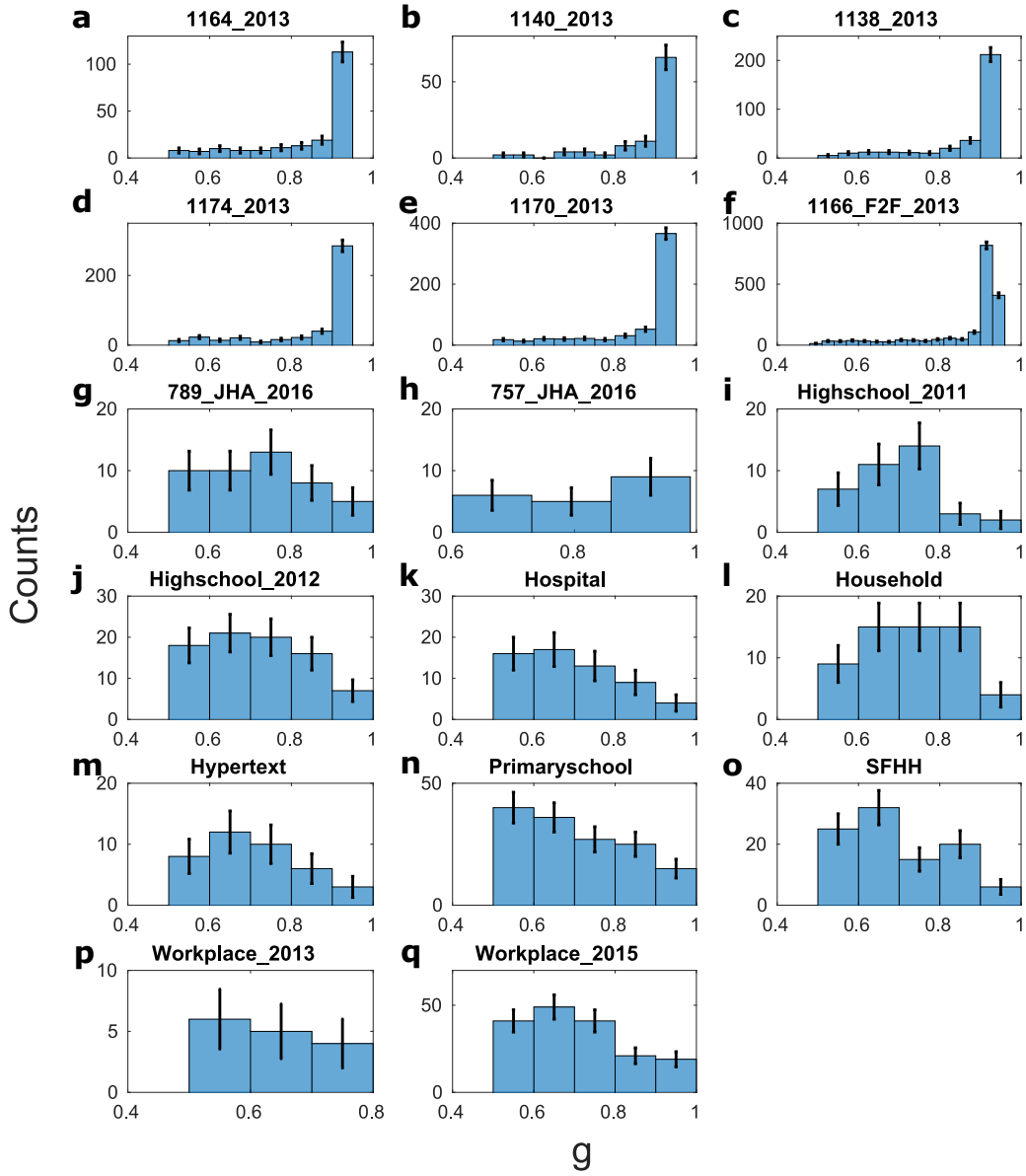

Fig. S3. Histogram of weight g in hyperexponential fitting of the pair interaction time distribution for all honeybee and human datasets. g can only be between 0.5 and 1. For all figures, error bars indicate the standard error.

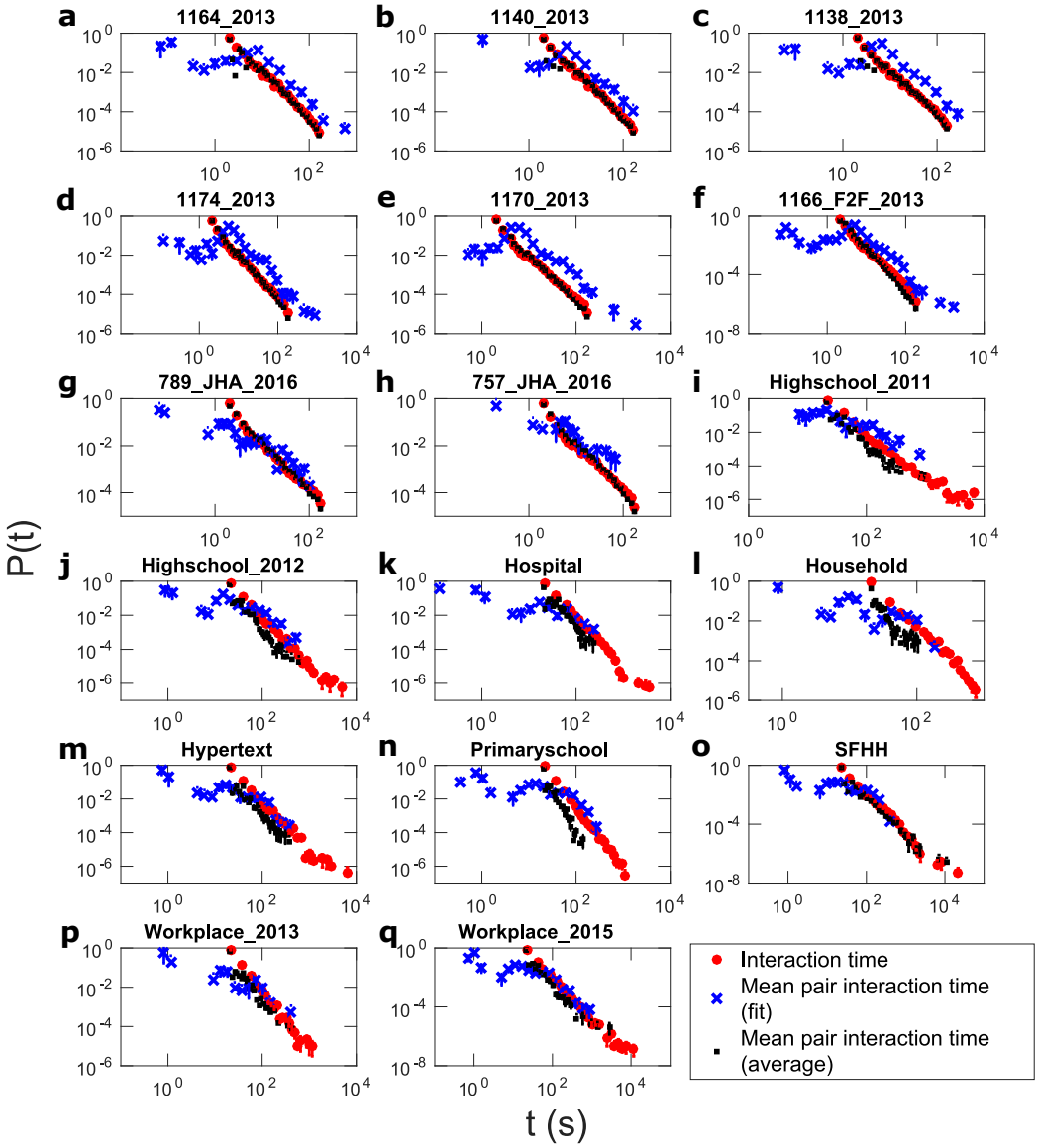

Fig. S4. Comparison of scaling between the interaction time distribution and the mean pair interaction time distributions for all honeybee and human datasets. The probability densities of mean pair interaction times obtained by different methods are separately plotted. From **a** to **q**, the number of mean pair interaction times from averaging used to generate the plot is 200723, 143571, 129653, 174317, 212685, 472914, 88441, 76810, 1710, 2220, 1139, 891, 2196, 8316, 9889, 754, 4273, which is the same as the number of detected pairs. The number of mean pair interaction times from fitting used is 197, 99, 328, 443, 561, 1806, 46, 20, 37, 82, 59, 58, 39, 143, 98, 15, 171, which is the same as the number of pairs used for fitting. For all figures, error bars indicate the standard error. Some error bars are not visible because they are smaller than the marker size. Lower error bars for bins of count 1 could not be drawn on a logarithmic scale.

decays faster than t^{-3}; the exponent of $f(t)$ of Household is -3.4 and that of Primaryschool is -3.6. The other datasets have $f(t)$ with exponents from -2 to -3, with specific values are provided in the main text.

Now we explain why the value of the exponent of $f(t)$ is relevant to whether or not $p(\tau_i)$ has the same scaling as $p(\tau_{ij})$ by using the so-called stable law (S1), sometimes referred to as the generalized central limit theorem. A probability density is said to be stable if a linear combination of independent and identically distributed (i.i.d.) random variables drawn from this probability density is also a random variable from the same probability density up to location and scale parameters. The normal distribution is a well-known example of a stable distribution. The stable law states that the probability density of normalized sums of i.i.d. random variables converges to a stable distribution (S1), and specifies the conditions that determine the appropriate stable distribution. In our case, the random variables are simply the pair interaction times.

We are interested in a set of random variables, drawn from an empirical probability distribution that we will call the parent or source distribution. We wish to know what will be the probability distribution for a linear combination of these variables. If the parent distribution is stable, then the stable law will imply that the linear combination is also distributed according to the same distribution. If not then it will be distributed according to a different distribution.

The distribution of the linear combination of random variables depends on the asymptotic behavior of the parent distribution. If the parent distribution $f(x)$ has the asymptotic behavior $f(x) \sim x^{-(\alpha+1)}$ with $\alpha \geq 2$ for large x, the distribution of the sums converges to a normal distribution. This result is the central limit theorem. On the other hand, if $0 < \alpha < 2$, the parent distribution itself is stable, so the limit distribution of the sums converges to the parent distribution $x^{-(\alpha+1)}$.

In our case, the parent distribution is $p(\tau_{ij})$, and the distribution of the normalized sums is $p(\tau_i)$. The datasets with the same scaling of $p(\tau_{ij})$, $p(\tau_i)$ and $f(t)$ all have $f(t)$ with exponents from -2 to -3. According to the theory in the main text, $f(t)$ and $p(\tau_{ij})$ always have the same exponents, which is verified by all the datasets in Fig. S4; thus $p(\tau_{ij})$ has exponents in the range from -2 to -3. It can be re-written as $p(\tau_{ij}) \sim \tau_{ij}^{-(\alpha+1)}$ with $1 < \alpha < 2$. Indeed, we see that α is in the range where $p(\tau_{ij})$ is stable, and so we expect the same scaling between $p(\tau_i)$ and $p(\tau_{ij})$, assuming that the variables are indeed i.i.d. In conclusion, $f(t)$ has the same scaling as $p(\tau_{ij})$ due to our theory, while $p(\tau_{ij})$ and $p(\tau_i)$ have the same scaling because $p(\tau_{ij})$ is the parent distribution of $p(\tau_i)$. Thus all three variables have the same scaling when $\alpha + 1$ is in the range from -2 to -3.

On the other hand, Household (Fig. S4l) and Primaryschool (Fig. S4n) have $f(t)$, and thus $p(\tau_{ij})$, with exponents -3.4 and -3.6 respectively. These exponents correspond to $\alpha + 1$. For these datasets, $\alpha > 2$, showing that $p(\tau_{ij})$ is not a stable distribution. Therefore, the stable law predicts that $p(\tau_i)$ does not have the same scaling as $p(\tau_{ij})$ and $f(t)$. The deviation of $p(\tau_i)$ from the other two distributions in Fig. S4l and n is consistent with this prediction.

Although the stable law predicts that the probability density of the normalized sum of i.i.d. random variables that are not stable converges to a Gaussian, $p(\tau_i)$ is not Gaussian as shown in Fig. S4l and n. It is not Gaussian because the number of summed τ_{ij} is only two, as there are two energy barriers for each pair. The convergence to a Gaussian only occurs if the number of summands is large (e.g. more than 10). In addition, even if the number of summands were large, burstiness of temporal social networks of honeybees and humans suggests that τ_{ij} could be not i.i.d. Then $p(\tau_i)$ would be expected to exhibit an exponential tail rather than a Gaussian tail, because the distribution of the normalized sum of correlated random variables converges to the Fisher-Tippett-Gumbel distribution, which has an exponential tail (S2).

In summary, $f(t)$ and $p(\tau_{ij})$ always have the same scaling, as predicted by our theory. However, as demonstrated in Fig. S4l and n, $p(\tau_i)$ may or may not have the same scaling as $f(t)$ depending on the exponent that characterizes the asymptotic behavior of $f(t)$, in accord with the stable law. This observation shows that we cannot ignore the fact that there are two energy barriers per pair, and we indeed need to obtain the mean pair interaction times by fitting the pair interaction time distribution to the hyperexponential function.

Note that the stable law does not hold between $f(t)$ and $p(\tau_{ij})$ although they have the same exponent. τ_{ij} is a normalized sum of pair interaction times associated with one energy barrier whose probability density is exponential. So the parent distribution for τ_{ij} is not $f(t)$. Furthermore, each τ_{ij} has its own parent distribution labeled by indices i and j, which is an exponential distribution with different parameter values. The stable law concerns the normalized sum of random variables from the same probability density. Therefore, one cannot predict the form of $p(\tau_{ij})$ from $f(t)$.

Conversely, our theory is only able to predict the form of $f(t)$ from $p(\tau_{ij})$ precisely because $f(t)$ is not the parent distribution for τ_{ij} and so the stable law does not hold between them. As a result, the scaling of $p(\tau_{ij})$ is purely determined by $p(E_{ij})$.

6. Lorenz plots.

Heterogeneity or variability in the population is represented by the barrier height distribution in our model. However, the model does not dictate whether the barrier height is determined by a specific pair or a particular individual, which poses the question about the individuality in social interaction. To explore the effect of individuality in a way that is independent of our theory, we have calculated the Gini coefficients (S3) for the total number of interactions and the total number of partners in addition to the total interaction time shown in the main text. We have chosen the Gini coefficient to quantify the different degree of dominance in social interactions by each individual because we are analyzing interaction times, which are shared between a pair, and thus it is nontrivial to decouple individual contribution.

Fig. S5 shows the Lorenz plots (S4) for the total number of interactions and the total number of partners of honeybees and humans. The larger deviation from the reference line in Fig. S5b and d than in Fig. S5a and c respectively indicates that individual differences are larger in human communities than honeybee communities. However, Fig. S5a and c still show a deviation from the straight line, signifying that honeybee individuals are different. This result still does not tell us how much individuality contributes to social interaction. Given that the energy barrier is a pairwise property, we can only tell that the

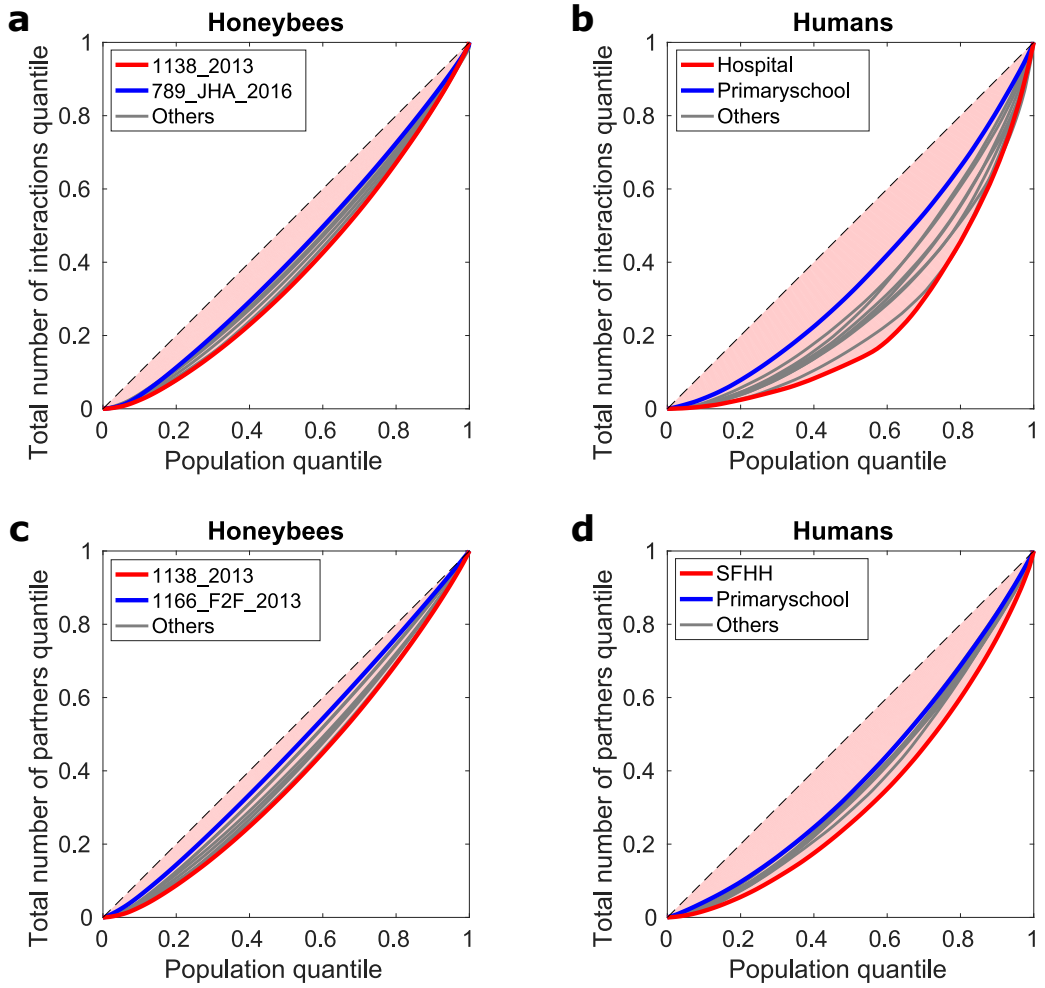

Fig. S5. Lorenz plots for total number of interactions and total number of partners for honeybees and humans. **a**, Gini coefficients for total number of interactions of honeybees are following. 1164_2013: 0.2125, 1140_2013: 0.1823, 1138_2013: 0.2526, 1174_2013: 0.2462, 1170_2013: 0.2362, 1166_F2F_2013: 0.1949, 789_JHA_2016: 0.1609, 757_JHA_2016: 0.1667. **b**, Gini coefficients for total number of interactions of humans are following. Highschool_2011: 0.3509, Highschool_2012: 0.4361, Hospital: 0.5293, Household: 0.4992, Hypertext: 0.4089, Primaryschool: 0.2595, SFHH: 0.4410, Workplace_2013: 0.4011, Workplace_2015: 0.3426. **c**, Gini coefficients for total number of partners of honeybees are following. 1164_2013: 0.1758, 1140_2013: 0.1619, 1138_2013: 0.2222, 1174_2013: 0.2081, 1170_2013: 0.1897, 1166_F2F_2016: 0.0916, 789_JHA_2016: 0.1284, 757_JHA_2016: 0.1301. **d**, Gini coefficients for total number of partners of humans are following. Highschool_2011: 0.2612, Highschool_2012: 0.2538, Hospital: 0.2828, Household: 0.2498, Hypertext: 0.2643, Primaryschool: 0.2215, SFHH: 0.3474, Workplace_2013: 0.2537, Workplace_2015: 0.2243.

pairwise heterogeneity is essential for the heavy tail to appear in the interaction time distribution, as demonstrated in this report, but whether the pairwise heterogeneity is purely a property of a pair independent of individuals forming the pair is still open for further study.

References

[S1] Gnedenko BV, Kolmogorov AN (1954) *Limit Distributions for Sums of Independent Random Variables.* (Addison-Wesley).

[S2] Bertin E, Clusel M (2006) Generalized extreme value statistics and sum of correlated variables. *J. Phys. A: Math. Gen.* 39(24):7607–7619.

[S3] Gini C (1912) Variabilità e mutabilità (variability and mutability). *Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi (1955) ed. Bologna.*

[S4] Lorenz MO (1905) Methods of measuring the concentration of wealth. *Publications of the American statistical association* 9(70):209–219.