

# SCIENTIFIC REPORTS



OPEN

## The impact of variable commitment in the Naming Game on consensus formation

Xiang Niu<sup>1,2</sup>, Casey Doyle<sup>1,3</sup>, Gyorgy Korniss<sup>1,3</sup> & Boleslaw K. Szymanski<sup>1,2</sup>

Received: 22 September 2016

Accepted: 22 December 2016

Published: 02 February 2017

The Naming Game has proven to be an important model of opinion dynamics in complex networks. It is significantly enriched by the introduction of nodes committed to a single opinion. The resulting model is still simple but captures core concepts of opinion dynamics in networks. This model limitation is rigid commitment which never changes. Here we study the effect that making commitment variable has on the dynamics of the system. Committed nodes are assigned a *commitment strength*,  $w$ , defining their willingness to lose (in waning), gain (for increasing) or both (in variable) commitment to an opinion. Such model has committed nodes that can stick to a single opinion for some time without losing their flexibility to change it in the long run. The traditional Naming Game corresponds to setting  $w$  at infinity. A change in commitment strength impacts the critical fraction of population necessary for a minority consensus. Increasing  $w$  lowers critical fraction for waning commitment but increases this fraction for increasing commitment. Further, we show that if different nodes have different values of  $w$ , higher standard deviation of  $w$  increases the critical fraction for waning commitment and decrease this fraction for increasing commitment.

Attempts to understand social interactions on networks has led to many diverse models, each balancing the needs for capturing the essence of network dynamics while keeping the model simple and efficiently computable. One emergent model that effectively captures the large scale effects of idea spreading using only simple pairwise interactions between nodes is the binary agreement (BA) model, a two word version of the naming game model<sup>1–8</sup>. In this model, each node in a network is holding one or both of two possible opinions. Each time step speaker and listener are chosen randomly the first from the network, and the second from the nearest neighborhood of the speaker. Then, the speaker chooses an opinion from its list to share with the listener. If the listener already holds this opinion, then both the speaker and listener keep only that opinion. Otherwise, the listener simply adds the new opinion to its list. Allowed to evolve this way, the system will eventually reach a stable consensus state in which every node holds the same opinion.

One of the main benefits of the BA model is that it is restricted to just two opinions and, correspondingly, three possible states: A, B, and the mixed opinion state AB. A simulation using these rules can be used to observe the effects of two competing opinions in a society, eventually converging to a consensus on one opinion. The simplicity of this model also allows for studying with it how spreading is affected by such phenomena as the addition of committed agents (nodes that will never change opinion regardless of interaction)<sup>9</sup>. Prior work on committed agent by J. Xie *et al.* put a minority population  $p$  of committed agents with a single opinion into a network with the remaining nodes holding the opposite opinion. Doing so introduces a phase transition in the system at  $p_c \approx 0.1$ . Above  $p_c$  the system is in the absorbing regime and is pulled toward the consensus state which is reached in time on the order of  $T_c \sim \ln(N)$  where  $N$  is the network size. Below the critical point, however, the system will stay in an active steady state until some random extremely unlikely fluctuation pushes the system past a saddle point and into the absorbing regime, making the time to consensus to be on the order of  $T_c \sim \exp(N)$ .

An extension of this work allows for the existence of two competing committed groups<sup>10</sup>, creating three distinct regions in parameter space designated as generalizing a region of the absorbing regime into a beak with respect to the committed populations for each opinion. The first region, with a high ratio of  $p_A$  with respect to  $p_B$ , lies below the beak and represents regime absorbing towards the steady state for majority A. The second has

<sup>1</sup>Rensselaer Polytechnic Institute, Social Cognitive Networks Academic Research Center, Troy, NY, 12180, USA.

<sup>2</sup>Rensselaer Polytechnic Institute, Department of Computer Science, Troy, NY, 12180, USA. <sup>3</sup>Rensselaer Polytechnic Institute, Department of Physics, Applied Physics, and Astronomy, Troy, NY, 12180, USA. Correspondence and requests for materials should be addressed to X.N. (email: nx.niuxiang@gmail.com)

a high ratio of  $p_B$  with respect to  $p_A$ , lying above the beak and staying at the steady state for  $B$  majority. The beak region with a more even ratio of  $p_A$  to  $p_B$ , is associated with a saddle point of the system that allows it to switch between the two steady states.

Further extensions of the committed BA models have become rather popular<sup>11–14</sup>, but most of them retain the infinite commitment of individuals. Here, we introduce and systematically examine impact of extensions of notion of commitment, from making it finite, so it can be lost, but also gained and finally its strength varied among the population according to some distribution. We start with an extended analysis of Galehouse *et al.* *waning commitment* model which accounts for human tendency to lose the commitment over time. In this model, it happens after a committed node hears  $w$  consecutive opposing opinions. This property causes an increase in  $p_c$  which becomes a function of  $w$ :  $p_c(w) - p_c(\infty) \sim w^d$ , where  $p_c(\infty)$  represents a critical value with infinite commitment strength,  $d \approx -1.73$ . When waning commitment is applied to competing committed groups, the beak widens in parameter space.

Here we seek to extend the various committed agent models<sup>9,10,15,16</sup> by investigating further the effects of variable commitment as well as of heterogeneity of the commitment nodes on spread and stability of opinion in a network.

## Waning Commitment

The model of *waning commitment* first proposed in ref. 15 can be further examined via mean field analysis to investigate the impact of commitment loss and to analyze the model for large values of  $w$ . This model follows the rules of the binary agreement model with the addition of committed agents that are each assigned a commitment strength  $w$ ; a node receiving  $w$  consecutive opposite opinions becomes uncommitted. In the original binary agreement model there are only two opinions (A and B) and only three states ( $A, B$ , and  $AB$ , where  $AB$  represents a neutral state). The addition of commitment strength, however, creates multiple substates for committed nodes corresponding to how close they are to losing their commitment, ie  $A_0, \dots, A_w, B_0, \dots, B_w$  where the commitment strength for nodes in opinion A is  $w$ , and the commitment strength for nodes in opinion B is  $w'$ . For simplicity, unless otherwise stated it is assumed that  $w = w'$ . Supplementary Fig. S1 shows the transitions between the substates.

## Increasing and Variable Commitment

The other side of *waning commitment* is the scenario where individuals have the ability to gain commitment through consecutive interactions affirming their current opinion. Here, the system is no longer initialized with some set population of committed minorities, but instead with some individuals having the capacity to become committed during the simulation. For simplicity (and to make it analogous to prior work), each individual with the capacity to commit will only be able to do so for a single opinion, set beforehand. This means there will be three general types of nodes: those that have no capacity to commit, those that have the capacity to commit to A, and those that have the capacity to commit to B. The agents currently holding opinion A and able to commit to A ( $A_{0<a>}$ ) will become committed ( $A'_w$ ) after consecutively receiving the opinion A  $w$  times; likewise for agents holding opinion B. The substate transitions are shown in the Supplementary Fig. S3. Finally, here we also introduce a new symmetric model with *variable commitment*, where nodes can enter or exit states of commitment based on the opinions they hear. The substate transitions are shown in the Supplementary Fig. S4.

## Distributed Commitment

In *variable commitment* model, commitment is not innate property of a node, instead it can be acquired or lost as a result of consecutive interactions with the same opinion. This allows for a more fluid commitment model. It can be made even more sophisticated by creating varying degrees of commitment strength for individual nodes to account for the tendencies for some people to be more stubborn or flexible with their ideas. This is novel because all prior analysis assumed that every committed node has the same level of commitment. *Distributed commitment* accounts for the diversity of people by allowing each committed node to have a personal commitment value assigned according to a distribution assumed for the given system.

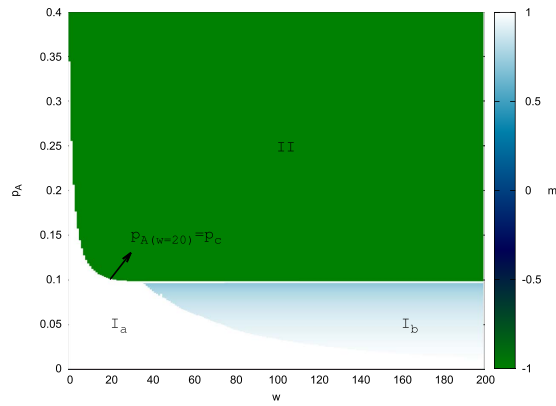
## Agent Selection

The final extension of a commitment model presented here investigates selection of committed agents based on their network connections rather than randomly. The effects of such selection on the critical system parameters can advance our understanding of how committed agents may work in real systems, as well as our ability to design efficient strategies for creating consensus in the presence of committed agents.

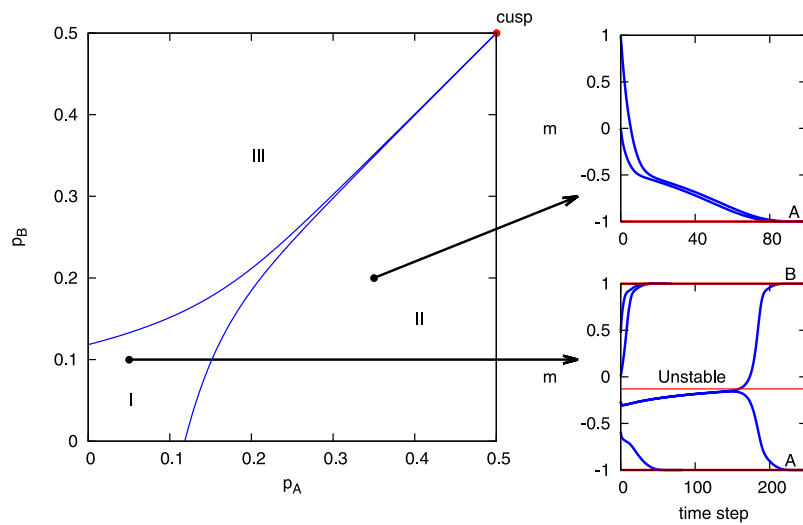
## Results

**Waning Commitment.** *Mean Field Analysis.* By deriving the probabilities of each state transition the system can undergo, *mean-field* equations are created leading to the critical values for the fraction of committed agents necessary to cause a phase transition in the system. As in the *constant commitment naming game*, each critical value represents a boundary separating two regions with different fixed points. Each can be found by using the bisection method on the *mean-field* equations as shown in Supplementary Table S2, yielding  $p_A$  in the single committed group model (shown in Fig. 1) and  $(p_A, p_B)$  in the two competing committed group model (shown in Fig. 2). The precision of this computation is  $10^{-10}$  due to using the standard 32 bit arithmetic and limited time for computation.

With a single committed group if their fraction is higher than the critical fraction, there is only one fixed point corresponding to an absorbing state for the committed opinion. In the opposite case, when fraction of committed agents is below the critical fraction, there are three fixed points, a steady state for the uncommitted opinion, a



**Figure 1.** Mean-field diagram of the waning commitment naming game with single committed group. The quantity  $m$  is calculated when the steady state condition of a system is satisfied. The system starts with the initial values of  $a_0 = c = 0, b_0 = 1 - p_A$ . Three regions of fixed points,  $I_a, I_b$  and  $II$  are shown and discusses in the text.

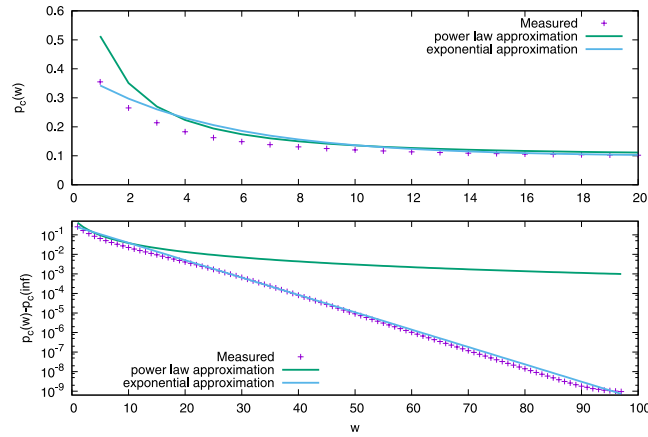


**Figure 2.** Mean-field diagram of the waning commitment naming game with two competing committed groups. The plots are calculated by the mean-field equations with commitment strength  $w = 10$ . (1) The left plot is a phase diagram containing three regions, in which the system has different fixed points. (2) The right plots show trajectories of points in regions I and II. The upper right plot is for the case where  $p_A = 0.35$  and  $p_B = 0.2$ . The lower right plot is for the case where  $p_A = 0.05$  and  $p_B = 0.1$ .

saddle state, and an absorbing state for the committed opinion. Importantly, though, at the critical point  $p_c$ , the steady state and the saddle state merge if the system is initialized with all non-committed nodes initially carrying the non-committed opinion. Figure 1 shows these two distinct regions separated by critical values of  $p_A$ . The quantity  $m = (b_0 - a_0) / (1 - \sum_{i=1}^w a_i - \sum_{j=1}^w b_j)$  is an order parameter quantifying the degree of dominance of one opinion over the other once the system has reached a steady state.

As shown in Fig. 1, in the first region (regions  $I_a$  and  $I_b$ ), the system has three fixed points: two steady states and one unstable state. The active steady state is defined by inequalities  $da \leq \varepsilon, db \leq \varepsilon, dc \leq \varepsilon, da_0 \leq \varepsilon, \dots, da_w \leq \varepsilon$  with  $\varepsilon = 10^{-5}$ . Under this definition, the active steady state appears in region  $I_b$ , which only occurs for the systems with high  $w$ . For lower values of  $w$  or  $p_A$ , the system is in the  $I_a$  region where it reaches a consensus for state  $B$ . Lastly, when the value of  $p_A$  is high enough the system enters region  $II$ , where there is only a single fixed point for the consensus state  $A$ .

In the model with two competing committed minorities, three regions arise: one with no dominance for either opinion, one with dominance for opinion  $A$ , and one with dominance for opinion  $B$ . Figure 2 shows these three different regions separated by a “beak” of critical values in parameter space. The first region (region  $I$ ) is within the beak, where the system has three fixed points corresponding to a consensus state for each of the two opinions as well as an unstable saddle state between them. Here the system can reach either consensus state depending largely on the initial conditions for the non-committed groups in the system ( $a_0, b_0$  and  $c$ ), or in the case of a relatively balanced beginning state can stay in the third unstable state until a random fluctuation pushes it into a consensus state. In the second and third regions (regions  $II$  and  $III$ ), there is only a single fixed point for a consensus on the dominant committed opinion. Here the system will reach the consensus for the dominant opinion



**Figure 3. The critical value function of the waning commitment naming game.** The top plot shows critical values of the fraction of committed agents with different commitment strengths. The bottom plot shows the differences between  $p_c(w)$  and  $p_c(\infty)$ . The critical value function is an exponential one especially when the commitment strength is large.

regardless of the initial conditions of the non-committed groups. The last area of interest in the case of competing committed agents is on the boundaries of the three regions, on which one of the consensus states and the unstable state merge together. This can only occur when all of the non-committed agents initially hold an opinion contrary to the more dominant committed side; i.e., at the points on the boundary between regions I and II, the initial values of  $a_0=0, c=0, b_0=1-p_A-p_B$  are needed; while for regions I and III, the needed initial values are  $b_0=0, c=0, a_0=1-p_A-p_B$ . These boundaries meet at the cusp of the beak, where  $p_A=p_B=0.5$ . At the cusp, all three states merge together.

*Single Committed Group Critical Value Function.* When there is only a single committed group in the system, the critical value of committed agents depends heavily on the commitment strength,  $w$ . The approximate value of  $p_c$  for specific values of  $w$  can be described as a critical value function  $p_c(w)$  obtained via a simplified version of the mean field equations and approximated values for the infinite  $w$  system denoted  $p_c(\infty)$  and the time  $t_w$  needed to reach the merged state:

$$p_c(w) \approx p_c(\infty) + p_c(\infty)r^w t_w \approx p_c(\infty) + 0.3 * 0.815^w \tag{1}$$

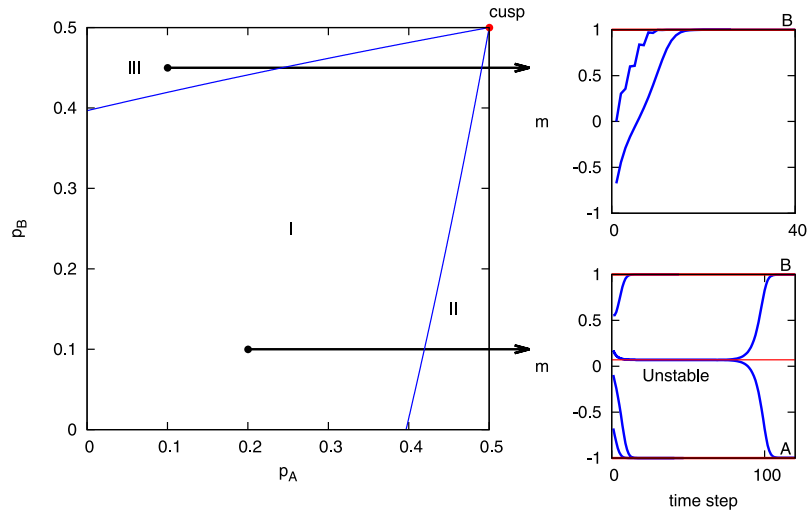
where  $r$  is a constant derived from Eq. 4 and simulation results with large  $w$ , as explained in the discussion of Eq. 4.

To ascertain validity of the approximate  $p_c(w)$  given by Eq. 1, in Fig. 3, we compare it to the value measured by the bisection method with equations in Supplementary Table S2 as well as previously proposed function by Galehouse *et al.* in ref. 15. The latter is a power law, and thus  $p_c(w) = p_c(\infty) + a(b + w^d)$ , where  $a \approx 2.792, b \approx 2.011, d \approx -1.729$ . This function fits well for low values of  $w$ , but deviates strongly for high values. In contrast, the exponential function  $p_c(w) = p_c(\infty) + k_q^w$  fits very well for all values of  $w$ , when  $k \approx 0.3, q \approx 0.815$ .

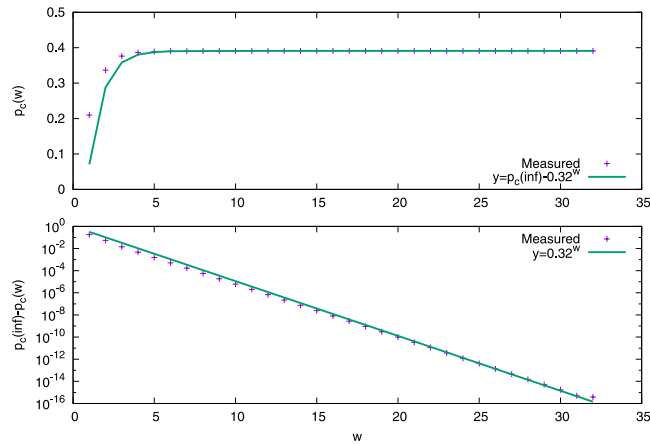
**Increasing Commitment. Mean Field Analysis.** Many of the qualitative aspects of the critical values in the *increasing commitment* model mirror those in the *waning commitment* case. As before, when there is only a single committed group the system undergoes a phase transition at some critical value  $p_c(w)$ , defined by a critical value equation similar to Eq. 1. Also as before, in model of the two competing committed minorities, three regions arise separated by a beak. The fixed points of the system still change depending on the region as can be seen in the phase diagram in Fig. 4. Here, the value of  $m$  is adjusted slightly to fit the new model:  $m = (b_{0<n>} - a_{0<n>}) / (1 - p_A - p_B)$ .

The first region is inside the beak and has three fixed points: a steady state for the dominant opinion, a saddle point, and a steady state for the unstable opinion. Now, however, the addition of *increasing commitment* allows these steady state points to be either consensus (absorbing) or active steady states, depending on whether any of the nodes achieved full commitment. The trajectories still show the system able to reach either both absorbing states or one stable and one unstable state. In the second and third regions, the system has only the single steady state fixed point for the dominant opinion. As before, the systems trajectory will force it towards this state regardless of other initial parameters. Finally, the last point of interest is the shape of the boundaries, which have lost all curvature and are linear in parameter space.

*Single Committed Group Critical Value Function.* The analysis of a critical value function for the case with a single committed group is also similar to that done for the *waning commitment* model. By applying bisection to the mean field equations, the real critical values can be obtained. Specifically, the general form is  $p_c(w) \approx p_c(\infty) + kq^w$  where  $k \approx -1, q \approx 0.32$ . Figure 5 shows the critical value function with this added commitment strength parameter. With the limitation of computing precision around  $10^{-16}$ , the computed critical



**Figure 4.** Mean-field diagram of the increasing commitment naming game with two competing committed groups. The plots are calculated by the mean-field equations with commitment strength  $w = 10$ . (1) The left plot is a phase diagram containing three regions. (2) The right plots show trajectories of points in region I and III. The top right is the case where  $(p_A = 0.1, p_B = 0.45)$  and the bottom right is the case where  $(p_A = 0.2, p_B = 0.1)$ .

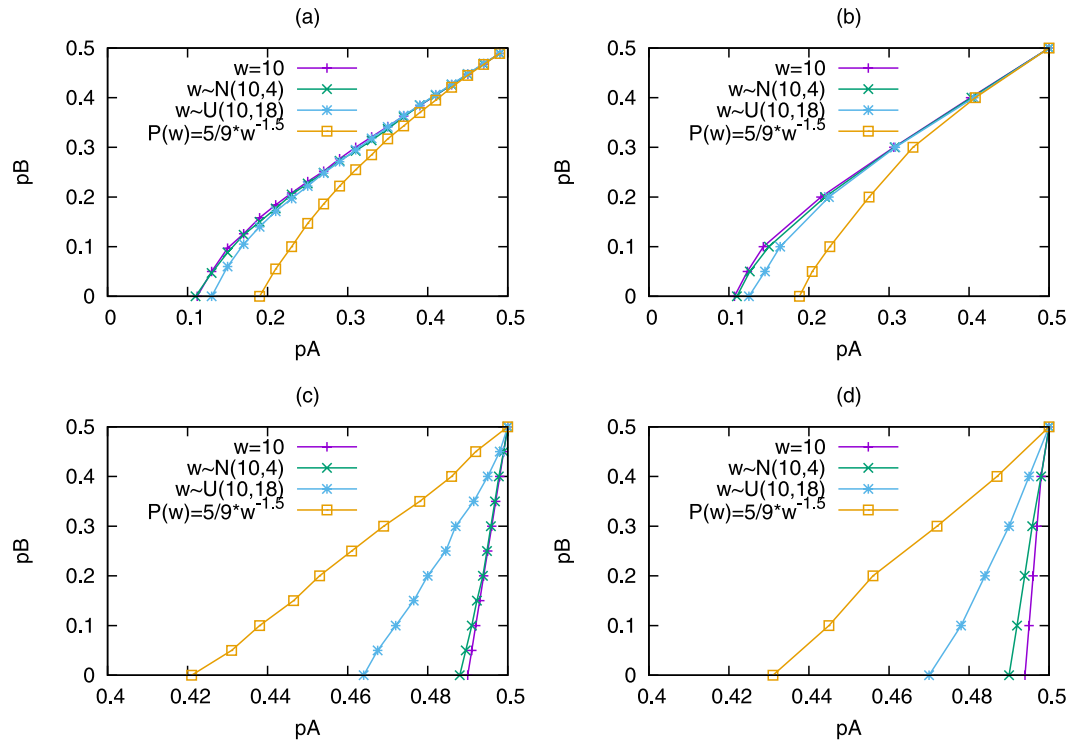


**Figure 5.** The critical value function of the increasing commitment naming game. The top plot shows critical values of the fraction of committed agents with different commitment strengths. The bottom plot shows the differences between  $p_c(w)$  and  $p_c(\infty)$ . Both plots shows our function fits data well.

value of  $p_c(33) = p_c(34) = \dots \approx p_c(\infty)$ . Since we cannot get critical values for  $w > 33$  with higher precision, in the experiment, we consider  $w = 33$  as computed infinite commitment strength.

**Distributed Commitment.** *Critical Value for a Single Committed Group.* As seen above, using a uniformly distributed variable commitment alters the value of  $p_c$  as  $p_c(w) \approx p_c(\infty) + kq^w$  where  $k > 0$  and  $q \approx 0.815$  for the waning commitment model, while  $k < 0$  and  $q \approx 0.32$  in the increasing commitment model. The most straightforward way to incorporate distributed commitment into these models is to simply use the average commitment strengths in these formulas. Indeed, there is a good deal of legitimacy to this method, as it can be verified by estimating the critical value to be  $p_c^{est} \approx p_c^{avg} = \sum \lambda_w p_c(w)$  where  $\lambda_w$  is the fraction of committed agents with commitment strength  $w$ , then comparing to the real critical values for systems with uniformly distributed commitment strengths. The main quantities of interest in this comparison are the relative errors for both the waning commitment model and the increasing commitment model. These quantities can be seen in Supplementary Fig. S5, where the relative errors are around 1% for small  $\bar{w}$ , and negligible for large  $\bar{w}$ . It also should be noted that since a group of distributed values can be represented by multiple subgroups of uniformly distributed values, then if the estimation is accurate for the uniform distribution it is very likely to be accurate for other distributions.

*Effects of Standard Deviation on the Critical Value Function.* Using the mean value estimation above, the effect of increasing standard deviation on the critical value function can be ascertained by studying the effect of a unitary increase in standard deviation on a system with an arbitrary distribution of commitment strengths. Doing so



**Figure 6. The critical value curves comparison.** All the experiments are based on networks generated by Barabasi-Albert (BA)<sup>17</sup> and Erdos-Renyi (ER) models with parameters  $N = 1000$ ,  $m_0 = m = 5$ ,  $p \approx 0.01$ , where the average degree of each network is 10, the mean of commitment strengths  $\bar{w}$  of each distribution is 10. The standard deviation in each case is 0, 2, 5, 196 and 16.3 for constant value, and then distributed normally, uniformly and according to power-law commitment strengths. Each critical value is averaged over 100 runs, half runs reaches consensus state A, while the other half reaches consensus state B. For simplicity, we just plot the critical values when  $p_A > p_B$ . Subfigures (a) and (b) show the critical value curves in the *waning commitment naming game*.  $p_A$ ,  $p_B$  are the initial fractions of committed agents at state A, B. Subfigures (c) and (d) show the critical value curves in the *variable commitment naming game*. The *variable commitment* yields very similar results to *increasing commitment*, but runs much faster.  $p_A$ ,  $p_B$  are the initial fractions of agents able to commit at state A, B, which are initialized as uncommitted, but are likely to become committed during interactions. Subfigures (a) and (c) are simulations in BA networks, while subfigures (b) and (d) are simulations in ER networks.

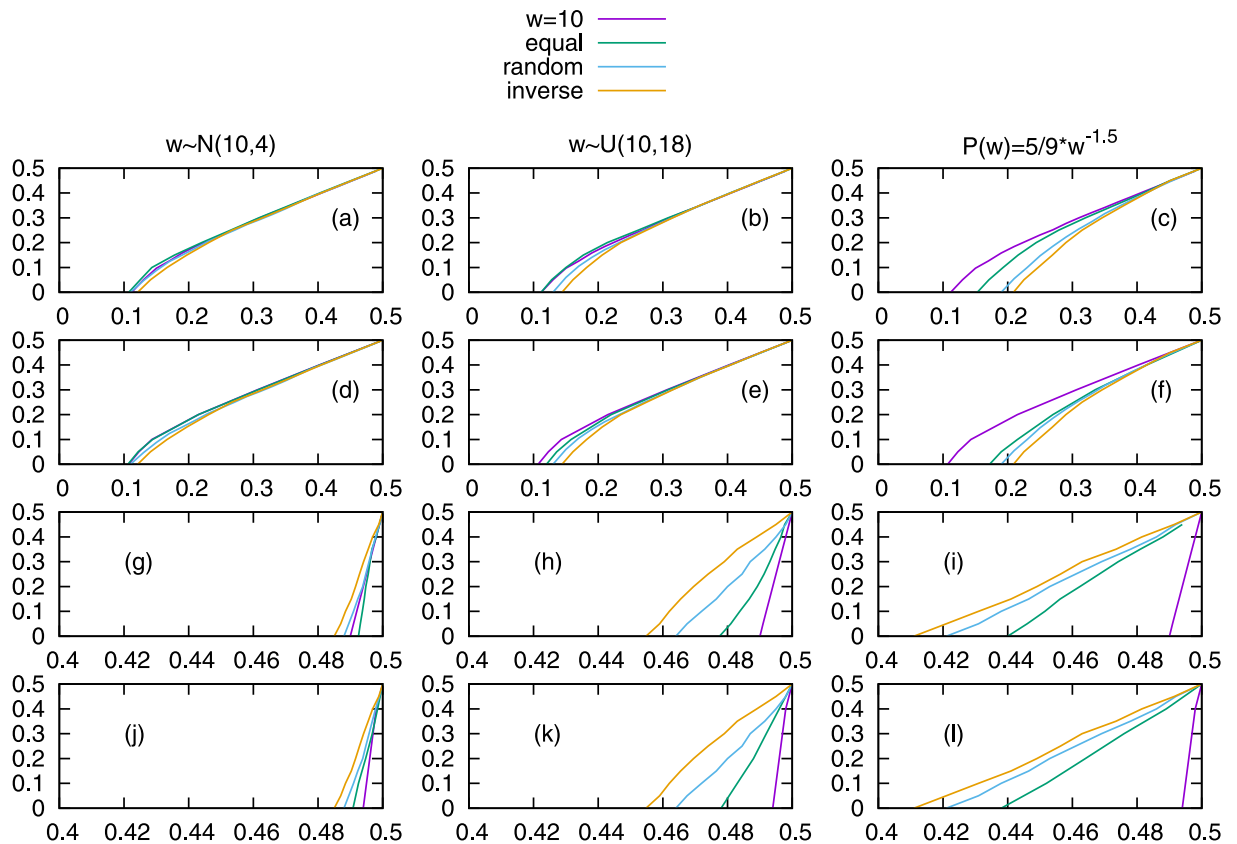
yields Eq. 9, which makes it clear that  $k$  is the controlling parameter of whether  $p_c$  is growing or shrinking with standard deviation. When  $k > 0$ ,  $p_c^{after} > p_c^{before}$  and thus the critical population is growing with increased standard deviation. When  $k < 0$ , however,  $p_c^{after} < p_c^{before}$  and thus the critical population shrinks with increased standard deviation. As mentioned before, the *waning commitment* model has  $k > 0$  while the *increasing commitment* model has  $k < 0$ . Therefore the two models react in opposite ways to an increase in standard deviation, with *waning commitment* growing its critical population while *increasing commitment* shrinking it.

**Competing Committed Minorities.** To test the validity of our analysis on systems with distributed commitment, we compare them to Monte Carlo simulations. Those can be done for any type of commitment distribution, so for completeness we compare four different types of systems: constant value, uniform distribution, normal distribution, and power-law distribution. For simplicity, for a single simulation, all committed agents follow the same distribution regardless of their opinion.

For the case where commitment strengths,  $w$  is set to 10, which also serves as the mean for all the other distributions. The uniform distribution has a width of 18 ( $\sigma = 3\sqrt{3}$ ). The normal distribution uses  $w \in [1, 19]$  and variance  $\sigma^2$ . Lastly, the power law follows  $p(w) = \frac{5}{9}w^{-1.5}$ ,  $w \in [1, 100]$ , which has  $\sigma = 16.3$ .

Based on simulations runs for each desired distribution of commitment strengths (constant commitment, uniform distribution, normal distribution, and power law distribution; all with  $\bar{w} = 10$ ) and then averaging the results yield the critical value curves for the committed populations shown in Fig. 6. For comparisons sake, these simulations are run on *waning commitment* (Fig. 6(a)) and *variable commitment* (Fig. 6(c)). *Variable commitment* is used instead of *increasing commitment* because the latter allows for infinitely committed nodes and often requires extremely large run times. Instead, using a *variable commitment* model where all the nodes with the ability to commit begin uncommitted (similar to the initial conditions of a pure *increasing commitment* simulation) allows for a close approximation because it forces the *increasing commitment* effect to occur before *waning commitment* is possible. This allows for the eventual loss of commitment to lower simulation time while





**Figure 7. The critical value curves comparisons of matching equal, inverse or random ranks of node degrees and commitments.** The committed agents are selected by three different strategies mentioned above. All the experiments are based on networks generated by Barabasi-Albert (BA) and Erdos-Renyi (ER) models with parameters  $N = 1000$ ,  $m_0 = m = 5$ ,  $p \approx 0.01$ , where the average degree of each network is 10, the mean of commitment strengths  $\bar{w}$  of each distribution is 10. Each critical value is averaged over 100 runs. For all subfigures, the x-axis is  $p_A$ , the y-axis is  $p_B$ . For simplicity, we just plot the critical values when  $p_A > p_B$ . Subfigures (a–f) show experiments of the *waning commitment naming game*, while subfigures (g–l) show experiments of the *variable commitment naming game*. Subfigures (a–c) and (g–i) are simulations in BA networks, while subfigures (d–f) and (j–l) are simulations in ER networks.

maintaining much of the behavior of the *increasing commitment model*. In each case, as the standard deviation increases the critical value curve moves to the right in the *waning commitment naming game*, and to the left in the *variable commitment naming game*. In other words, the wider distributed is the commitment strengths of a group of committed agents (holding the same opinion), the easier it is for them to change their opinions. The exact opposite is true for a group of agents able to commit to an opinion. Since the standard deviation are different for different distributions in Fig. 6, we run experiments with different standard deviations and fixed distribution type in Supplementary Fig. S7, which yield exactly the same conclusions.

**Agent Selection.** The previous sections show that the critical value is influenced by the variance of the commitment strength distribution. Thus, given a certain commitment strength distribution, there should be methods to enhance the ability of the committed agents to spread their opinion by placing high values of commitment on nodes with specific network properties.

The most direct way to go about this is to select the value of commitment strength based on the degree of an agent. This can be done in many different ways, but three of the most straight forward are matching by equal ranks, inverse ranks or random ranks of commitment strength and node degrees. Ranks are assigned by sorting agents in decreasing order of their degrees and commitments in decreasing order of their strength. All our previous experiments are based on the random matching.

Figure 7 shows the critical value curve comparisons among three matching methods. For each distribution in the *waning commitment naming game*, the agents committed to opinion A benefit from the equal rank matching (the critical values of  $p_A$  of equal rank is smaller than that of random rank), but suffer from the inverse rank matching. However, in the *variable commitment naming game*, the agents that are able to commit to A benefit from the inverse rank matching, but suffer from the equal rank matching.

## Discussion

**Waning Commitment.** Analysis of the mean field equations for the *waning commitment* model culminates in the discovery of three fixed points: the A consensus state, the B consensus state, and an unstable state. This is analogous to the results previously discussed for the *constant commitment* model, where inside the “beak” in parameter space, there are three fixed points: a consensus state in which all uncommitted agents have the same opinion, a saddle point, and an active steady state without agreement<sup>9,10</sup>. The obvious differences are that in the *waning commitment* model, the consensus states become more complete (as the loss of commitment allows even previously committed nodes to change their opinion) and the active steady state becomes a consensus state for the same reasons. The similarities make sense, since the *constant commitment* model is in fact just a special case of the *waning commitment* model in which the commitment strength is infinite. In that sense, the *waning commitment* is a more general model for human behavior.

In the simpler case with only a single committed group, the analysis of the mean field equations leads to a new form for the dependence of  $p_c$  on  $w$ . This new function outperforms one proposed in previous work<sup>15</sup> which is an average over multiple direct simulations for limited values of commitment strength and subject to the large experimental error inherent in direct simulation. Indeed, tests on the power law function proposed there for large values of  $w$  show great deviation from the values obtained via the bisection method. In contrast, the new exponential function proposed here is able to capture the behavior at large  $w$  while still remaining valid in the small  $w$  regime.

A final point of interest is the degree of validity of the approximations made in the bisection method. Due to computational precision of the mean field equations,  $w = 98$  is approximated to be equivalent to infinite commitment strength. This approximation holds surprisingly well considering that there are around 200 equations to simulate. In fact, the error between  $p_c(98)$  and  $p_c(\infty)$  is only  $\varepsilon = 9.0 \times 10^{-10}$ , when the precision is set to  $10^{-10}$ . This precision shows quite clearly that the analytic results obtained hold up to more exact values despite the difficulties of solving the intermediate equations corresponding to each of the substates of the system.

**Increasing Commitment.** *Increasing Commitment* is largely proposed to mirror *waning commitment* and round out the subject to set up *variable commitment*. As expected, in many ways the increasing commitment model is directly opposite of the *waning commitment* model. In fact, in the single committed group model, the critical value function is formulated in the same way with only a different set of parameters:  $p_c(w) = p_c(\infty) + kq^w$ . In the *waning commitment*,  $k \approx 0.3$ ,  $q \approx 0.815$  while in the *increasing commitment*,  $k \approx -1$ ,  $q \approx 0.32$ .  $k$  represents the direction of the change of the committed opinion, while  $q^w$  represents the displacement. With the increase of  $w$ , the displacement decreases and finally shrinks to 0. The two results actually display a very nice symmetry, and the specifics of the parameters speak volumes about the general behavior of the system. Clearly, due to the way the models have been constructed, they help/hinder the critical population via the same mechanism leading to equations of the same form. Yet the differences in the parameters also show the degree to which these changes manifest themselves. Namely, since the *waning commitment* model has a  $q$  much closer to 1, it will likely be less sensitive to changes in  $w$ , while the fact that it is much higher than the  $q$  in the *increasing commitment* model implies that it will be more likely to deviate from  $p_c(\infty)$ . This makes intuitive sense given the initial conditions of the system having a much higher density of nodes in opinion B than opinion A. When allowed to propagate, it is much more likely that a committed node will hear  $w$  consecutive opposing opinions than  $w$  consecutive confirming opinions. Thus, the special mechanism inherent to *waning commitment* is much more likely to occur than that in *increasing commitment*. Different initial conditions, however, can have very strong effects on the dynamics of the system. For instance, in the *variable commitment* naming game, it is shown that the *increasing commitment* effect can be allowed to dominate simply by initializing all of the nodes with the ability to commit into an uncommitted state so that they are forced to undergo the *increasing commitment* effect before there is even a possibility for *waning commitment* to take place. It is clear that alterations on the initial conditions can change the system dramatically, and defining them clearly is extremely important. In this case, initial conditions were chosen in a natural way to fit the goals of each experiment.

**Distributed Commitment.** The idea behind distributed commitment is to create a system with a more human-like heterogeneity in people’s willingness to change. In doing so, it becomes apparent that for the most part systems of differently distributed commitment values will behave similarly to our systems above, only with the constant  $w$  replaced by the average value of  $w$  across the system. Systems with the same mean do not behave identically, though, as the critical values also have some dependence on the standard deviation of  $w$  across the system. Like before, the *increasing commitment* and *waning commitment* cases behave in opposite ways: when the standard deviation increases, the critical values in the *waning commitment* model also increase while in the *increasing commitment* model the critical values decrease. These results can be confirmed by using Jensen’s Inequality<sup>18</sup> ( $\frac{f(y)-f(x)}{y-x} < \frac{f(z)-f(y)}{z-y}$ ,  $x < y < z$ ) given the strictly convex critical value function of *waning commitment* and the strictly concave critical value function of *increasing commitment* (for details see Supplementary materials). The observed behavior can be explained by considering high and low standard deviation on a node by node basis: high standard deviation will lead to some deeply committed nodes and some easily swayed nodes. For the *waning commitment model*, this means that many of the nodes are practically not committed at all; they will quickly lose their commitment and become normal nodes, leaving only the deeply committed nodes to try and dominate the system. This will obviously increase the value of  $p_c$ , as the committed population will, in effect, be only a fraction of those initially designated. On the other hand in the *increasing commitment* model, having these easily swayed nodes is a very good thing. Considering the relative rarity of the switching events in this model discussed earlier, having some group easily switched can create a strong foothold for the committed opinion early on in the simulations, decreasing the initial number of committed agents necessary to create consensus.



**Agent Selection.** The final tests of variably committed individuals establishes how the locations of the committed agents within the structure of the network impact the critical values of the systems. By designating nodes of higher/lower degree as committed agents, large changes in the critical values can be induced. This is an idea often tested on prior models of committed nodes<sup>5</sup>, leading to the conclusion that “selecting a small number of the nodes with the highest degree works best” for a committed minority to maximize their influence. Indeed, this can still be seen in the case of variable commitment. However, since the seed selection problem in the Naming Game has not been fully analyzed, a more in depth study is left for future work.

## Methods

**Waning Commitment.** *Mean Field Analysis.* Supplementary Table S1 shows the change in the number of agents at each substate after one system interaction. In the *mean-field* analysis, each interaction is considered a system time step, differing from simulations on BA and ER networks where  $N$  interactions are considered a system time step (where  $N$  is the network size).

To simplify the equations, we define the quantity  $r$  as  $r \equiv b + c/2$ , following the approach used in ref. 19, where  $a, b, c$  represent the fraction of agents at states A, B, AB. In order to analyze the steady states of this naming game, we set all derivatives to zero, which yields the simple formula (for  $w \geq 1$ )

$$da_w/dt = -a_w r^w = 0 \quad (2)$$

where  $a_0, \dots, a_w, b_0, \dots, b_w, r$  represent the fraction of agents at substates  $A_0, \dots, A_w, B_0, \dots, B_w$ . This equation implies that for finite values of  $w$ , either  $a_w = 0$  or  $r = 0$ , but applying these equations recursively reveals  $a_{w-1} = a_w r = 0$ , and thus  $a_1 = \dots = a_{w-1} = 0$  (and similarly  $b_1 = \dots = b_{w-1} = 0$ ). This leads to a final set of equations to be solved:

$$\begin{aligned} a_0 + a_w &= a \\ b_0 + b_w &= b \\ a_w r &= 0 \\ b_w(1 - r) &= 0 \end{aligned} \quad (3)$$

Solving these equations under the assumption that  $w \geq 1, w' \geq 1$  gives three possible fixed points:

- If  $r = 0$ ; then  $a = 1, b = c = 0$ , yielding a consensus state on A.
- If  $r = 1$ ; then  $b = 1, a = c = 0$ , yielding a consensus state on B.
- If  $0 < r < 1$ ; then  $a_w = b_w = 0, a_0 = a, b_0 = b$ , due to the rate equations  $da_0/dt = ac + cc - ab = 0$  and  $db_0/dt = bc + cc - ba = 0$  under the restriction  $a + b + c = 1$ . Solving simply yields  $a = b = (3 - \sqrt{5})/2 \approx 0.38, c = \sqrt{5} - 2$ , an unstable fixed point.

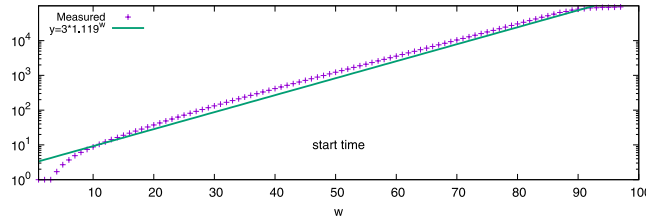
**Single Committed Group Critical Value Function.** By using the mean field equations in Supplementary Table S1 in conjunction with the previously stated fact that committed agents are lost at a rate of  $da_w/dt = -a_w r^w$ , it can be deduced that after  $t$  interactions the total number of committed agents will be approximately  $a_w - a_w r^w t$ . When  $p_A = p_c$ , the saddle state and steady state merge. Hence, we can assume that after times  $t_w, t_{w+1}$  for large values of  $w$ , systems with commitment strengths  $w$  and  $w + 1$  will be close to the merged state (with the similar number of remaining committed agents). Therefore

$$\begin{aligned} p_c(w) - p_c(w) r^w t_w &\approx p_c(w + 1) - p_c(w + 1) r^{w+1} t_{w+1} \\ &\approx \dots \approx p_c(\infty) - p_c(\infty) r^\infty t_\infty \\ &\approx p_c(\infty) \\ p_c(w) &\approx p_c(\infty) + p_c(w) r^w t_w \end{aligned} \quad (4)$$

With computing precision at  $10^{-10}$ , the computed critical value for  $p_c(98) = p_c(99) = \dots = p_c(\infty) + \varepsilon$ , where  $p_c(\infty) = 0.09788910377$  is the theoretical critical value of infinite commitment strength,  $\varepsilon = 9.0 \times 10^{-10}$ . Thus, with this precision  $p_c(98) \approx p_c(\infty)$ , so  $w = 98$  will be considered as the value of computed “infinite” commitment strength. The critical values are, as before, calculated using bisection on the mean field equations. When there is only a single committed minority and in the system initial state only the committed agents hold opinion A, the critical value is defined by Eq. 4:  $p_c(w) \approx p_c(\infty) + p_c(w) r^w t_w$ . This can be made more exact by looking at the known values for the “infinite” commitment case. Comparing Eq. 4 with simulation results for  $p_A = p_c(98)$ , we find that the merged state occurs for  $r \approx 0.728$ .

The value of  $t_w$  is much harder to ascertain, but can be approximated as the system start time, which represents the time from the initial state to the merged state. Three running time samples for  $w = 10, 30, 90$  can be found as Supplementary Fig. S2, in which we distinguished three periods of which the first call start time, and denoted  $t_w$  is used in Eq. 4. Figure 8 shows the start times for systems with different  $w$ 's. For this analysis, it will be approximated as  $t_w \approx 3 * 1.119^w$ , thus  $p_c(w) \approx p_c(\infty) + p_c(\infty) r^w t_w \approx p_c(\infty) + 0.3 * 0.815^w$ .

**Increasing Commitment.** Here  $A_0, AB$ , and  $B_0$  represent that agents currently holding opinion A, AB, B;  $A_{0 < a >}$  represents agents currently holding opinion A and able to commit to A ( $A_{0 < b >}$  for those in A able to



**Figure 8.** The start times of system with each commitment strength  $w$  at its critical point.

commit to B); and  $A_{0<n>}$ ,  $AB_{<n>}$ ,  $B_{0<n>}$  represent normal agents currently holding opinion A, AB, B respectively. The values  $a'_1, \dots, a'_w, b'_1, \dots, b'_w, a_{0<k>}, c_{<k>}, b_{0<k>}$  represent the fraction of agents at substates  $A'_1, \dots, A'_w, B'_1, \dots, B'_w, A_{0<k>}, AB_{<k>}, B_{0<k>}$ .

**Mean Field Analysis.** Supplementary Table S3 shows, the change of numbers of agents able to commit to A at each substate after one system interaction. In order to analyze the steady states of this naming game, we set all derivatives to zero, and solve the equation set. If  $w \geq 1$ ,

$$da'_w/dt = a_{0<a>}(1 - r)^w = 0 \tag{5}$$

Applying this recursively for finite  $w$  leads to  $a'_1 = a_{0<a>}(1 - r) = 0, \dots, a'_1 = \dots = a'_{w-1} = 0$ . Similarly,  $b'_1 = \dots = b'_{w-1} = 0, a_{0<a>} = b_{0<b>} = 0, c_{<a>} = c_{<b>} = 0, c = c_{<n>}$ . This leads to a final set of equations to be solved:

$$\begin{aligned} (3a/2 + c)c - ra_{0<n>} &= 0 \\ (3b/2 + c)c - (1 - r)b_{0<n>} &= 0 \\ ra_{0<n>} + (1 - r)b_{0<n>} - (3a/2 + 3b/2 + 2c)c &= 0 \end{aligned} \tag{6}$$

Once solved, these equations give three possible solutions for the system

- If  $r = 0$ ; then  $a = 1, b = c = 0$ , and  $a_{0<b>} = p_B$ . The system reaches a consensus state on A.
- If  $r = 1$ ; then  $b = 1, a = c = 0$ , and  $b_{0<a>} = p_A$ . The system reaches a consensus state on B.
- If  $0 < r < 1$ ;  $a_{0<b>} = b_{0<a>} = 0, a'_w = p_A$ , and  $b'_w = p_B$ . The system may reach a consensus state, an active steady state or an unstable state.

where  $p_A, p_B$  represent the total fraction of agents with potential to commit to A, B.

**Distributed Commitment.** *Effects of Standard Deviation on the Critical Value Function.* The impact of changes of standard deviation on the critical values of variable commitment model can be ascertained quite generally using an arbitrary distribution of commitment strengths. Using the mean value estimation described in the Results section, the estimated critical value of the system can be written as

$$p_c = \frac{p_c(w_{a_1}) + \dots + p_c(w_{a_n})}{n} \tag{7}$$

where there are  $n$  committed agents and the commitment strength of each agent is  $w_{a_1}, w_{a_2}, \dots, w_{a_n}$ . Any increase in standard deviation preserving average of integer commitments can be decompose into a series of pairs of changes  $(w_{a_i} - 1, w_{a_j} + 1)$  where  $1 < w_{a_i} \leq w_{a_j}$  each pair making the same change. Indeed, after a pair of changes,  $w'_{a_i} = w_{a_i} - 1$  and  $w''_{a_j} = w_{a_j} + 1$ . To find the effect of this change, the value of  $p_c^{after} - p_c^{before}$  must be calculated using Eq. 7. Doing so yields

$$p_c^{after} - p_c^{before} = \frac{p_c(w_{a_i} - 1) + p_c(w_{a_j} + 1) - p_c(w_{a_i}) - p_c(w_{a_j})}{n} \tag{8}$$

which after plugging in the exponential form of  $p_c(w)$  becomes

$$p_c^{after} - p_c^{before} = \frac{k}{n} q^{w_{a_i}-1} (1 - q^{w_{a_j}-w_{a_i}+1}) (1 - q) \tag{9}$$

## References

1. Baronchelli, A., Felici, M., Loreto, V., Caglioti, E. & Steels, L. Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat. Mech.* **P06014** (2006).
2. Dall'Asta, L., Baronchelli, A., Barrat, A. & Loreto, V. Agreement dynamics on small-world networks. *Europhys. Lett.* **73**, 969 (2006).
3. Barrat, A., Baronchelli, A., Dall'Asta, L. & Loreto, V. Agreement dynamics on interaction networks with diverse topologies. *Chaos* **17**, 026111 (2007).

4. Lu, Q., Korniss, G. & Szymanski, B. K. Naming games in two-dimensional and small-world-connected random geometric networks. *Phys. Rev. E* **77**, 016111 (2008).
5. Lu, Q., Korniss, G. & Szymanski, B. K. The naming game in social networks: community formation and consensus engineering. *J. Econ. Interact. Coord.* **4**, 221–235 (2009).
6. Baronchelli, A. & Diaz-Guilera, A. Consensus in networks of mobile communicating agents. *Phys. Rev. E* **85**, 016113 (2012).
7. Gao, Y., Chen, G. & Chan, R. H. Naming game on networks: let everyone be both speaker and hearer. *Sci. Rep.* **4**, 6149 (2014).
8. Lou, Y. & Chen, G. Analysis of the “naming game” with learning errors in communications. *Sci. Rep.* **5**, 12191 (2015).
9. Xie, J. *et al.* Social consensus through the influence of committed minorities. *Phys. Rev. E* **84**, 011130 (2011).
10. Xie, J. *et al.* Evolution of opinions on social networks in the presence of competing committed groups. *PLoS One* **7**, e33215 (2012).
11. Zhang, W. *et al.* Social influencing and associated random walk models: asymptotic consensus times on the complete graph. *Chaos* **21**, 025115 (2011).
12. Singh, P., Sreenivasan, S., Szymanski, B. K. & Korniss, G. Accelerating consensus on coevolving networks: the effect of committed individuals. *Phys. Rev. E* **85**, 046104 (2012).
13. Zhang, W., Lim, C. & Szymanski, B. K. Analytic treatment of tipping points for social consensus in large random networks. *Phys. Rev. E* **86**, 061134 (2012).
14. Mistry, D., Zhang, Q., Perra, N. & Baronchelli, A. Committed activists and the reshaping of status-quo social consensus. *Phys. Rev. E* **92**, 042805 (2015).
15. Galehouse, D. *et al.* Impact of network connectivity and agent commitment on spread of opinions in social networks. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics*, 2318–2329 (2014).
16. Doyle, C., Sreenivasan, S., Szymanski, B. K. & Korniss, G. Social consensus and tipping points with opinion inertia. *Phys. A: Stat. Mech. App.* **443**, 316–323 (2016).
17. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Modern Phys.* **74**, 47–97 (2002).
18. Jensen, J. L. W. V. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.* **30**, 175–193 (1906).
19. Treitman, Y., Lim, C., Zhang, W. & Thompson, A. Naming game with greater stubbornness and unilateral zealots. In *IEEE 2nd Network Science Workshop (NSW)*, 126–130 (2013).

## Acknowledgements

This work was partially supported by the Army Research Laboratory under Cooperative Agreement Number W911NF-09-2-0053 (Network Science CTA), and by the ONR Contract N00014-15-1-2640. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies either expressed or implied of the Army Research Laboratory or the U.S. Government.

## Author Contributions

Designed research: X.N., B.K.S., G.K., C.D.; Performed research: X.N.; Analyzed data: X.N., B.K.S., G.K., C.D.; Wrote and edit the paper: X.N., C.D., G.K., B.K.S.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Niu, X. *et al.* The impact of variable commitment in the Naming Game on consensus formation. *Sci. Rep.* **7**, 41750; doi: 10.1038/srep41750 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017