Supplementary material:

1. Parameters and levels examined

Variable	Type	Number of levels	Levels examined
Content bias	explanatory	11	b = 0.0 to 1.0 in steps of 0.1
Coordination bias	explanatory	11	c = -1.0 to 1.0 in steps of 0.1
Connectivity dynamic	explanatory	3	cd = high, moderate and low isolation
Memory	explanatory	4	mem = 1, 3, 5, 7
Mutation	control	1	m = 0.02

Table 1S. Parameters and levels examined.

2. Cognitive biases

2.1 Content bias

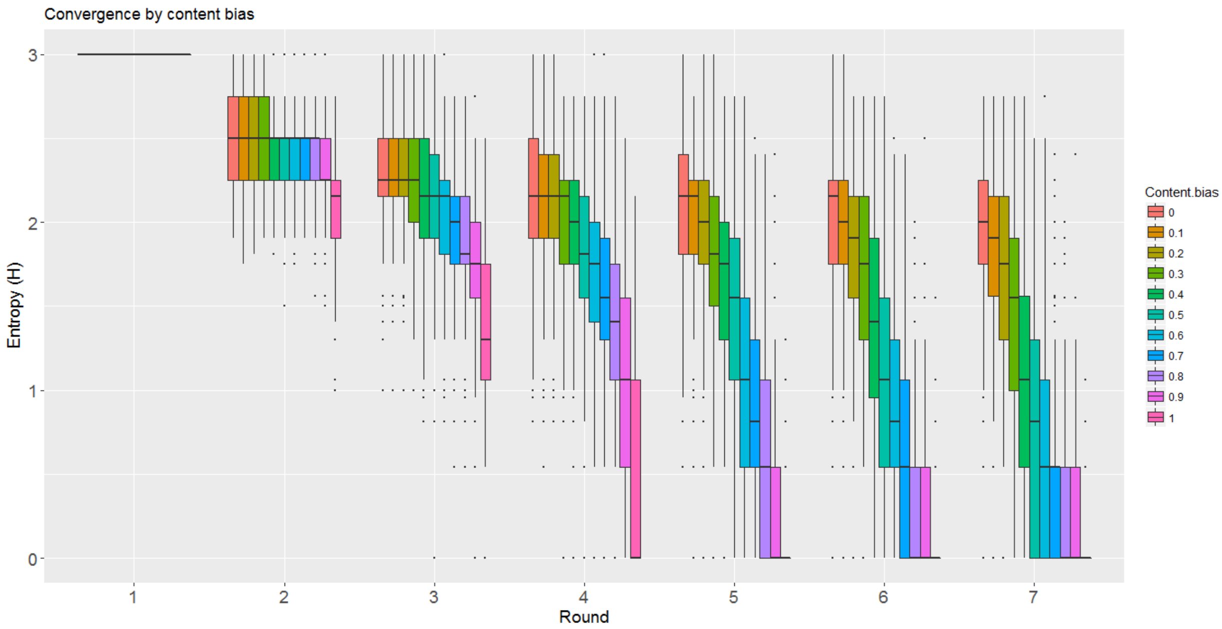

Figure 1S.: Entropy (H) by round by content bias. Data subset of neutral coordination bias. In this plot, a drift model corresponds to content bias = 0. In this and subsequent boxplots: middle line is median, 50% quantile; lower hinge, 25% quantile; upper hinge, 75% quantile; lower whisker is smallest observation greater than or equal to lower hinge - 1.5 * IQR; upper whisker is largest observation less than or equal to upper hinge + 1.5 * IQR.

One of the advantages of our model is that agents' responses are limited to a set of signals, so the number of permitted values of entropy is finite. It is therefore possible to accurately estimate conditional densities describing how the distribution of entropy changes over levels of content bias and over time. We compute conditional probability distributions of entropy by calculating the probability distribution of three entropy groups: high ($H \geq 2$ bits), medium (1 bit $\geq H < 2$ bits) and low ($H < 1$ bit). The probability of a non-convergent communication system (entropy higher than 2 bits) when content bias is 0 is approximately 99% by round 2: P(H high | Cont=.0, Gen=2) = 0.994, and it drops to 50% by round 7: P(H high | Cont=.0, Gen=7) = 0.504. When the content bias is 1, the probability of a non-convergent communication system is 68% by round 2, and almost 0% by round 7. On the contrary, the probability of a conventionalized communication system (entropy lower that 1 bit) when content bias is 0 goes from 0% to approximately 0.04%, whereas when content bias is 1 goes from 0% to 99%. More details in Fig. 2S.

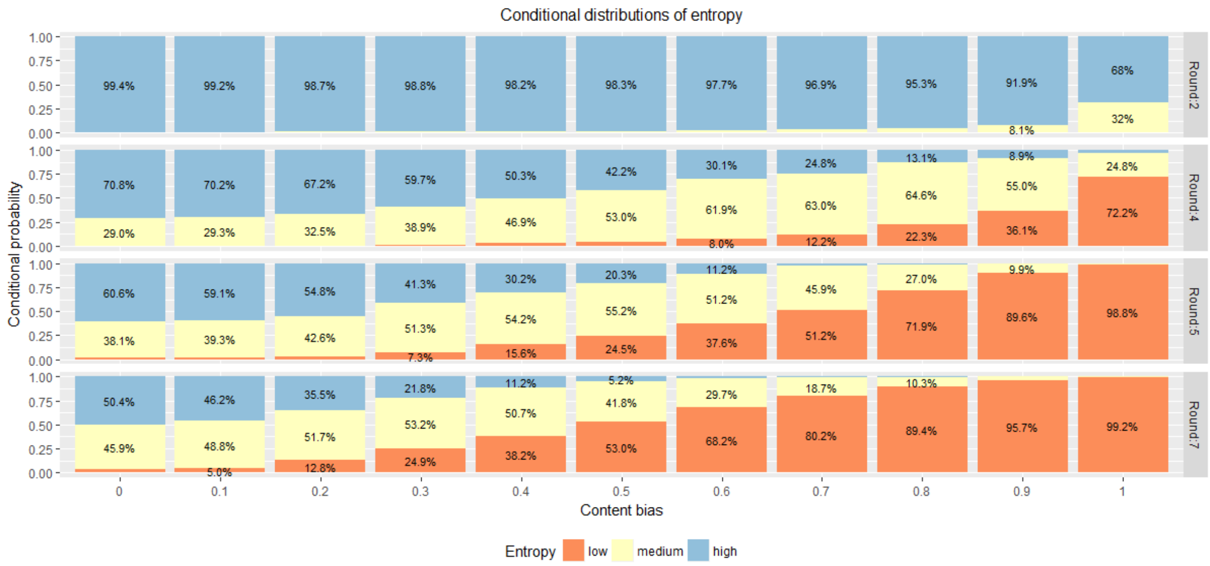

Figure 2S.: Conditional distributions of entropy over levels of content bias: *P(y | x)* against *x*. Bars represent conditional distributions of entropy (H): in blue high entropy (H ≥ 2 bits), in yellow medium entropy (1 ≥ H < 2), in red low entropy (H < 1). X-axis represents level of content bias, and y-axis represents conditional probability. For example, the probability of *high entropy (low convergence)* given that content bias is 0.5 (P(H high|Cont=.5)) goes from 0.9833 in round 2, to 0.0525 in round 7.

2.2 Coordination bias

Over rounds, the relationship between entropy and coordination bias remains characterized by a clear asymmetric distribution. When compared with neutral coordination, strong egocentric and allocentric behaviors reduce the entropy drop, and maintain variant diversity in the system. In the absence of content bias, entropy decreases only slightly towards its horizontal asymptote, which is highest when egocentric bias is strongest (Fig. 3S, supplementary material Fig. 4S).

Note: Our results agree with Walker et al. in recognizing a weak effect of coordination bias when content bias is strong, and a stronger effect of coordination bias when content bias is weak. Also, in both models, egocentric bias maintains diversity better than allocentric bias. However, when compared with a neutral model, allocentric bias reduced convergence in the current model, while increased it in Walker et al. This is likely due to different implementations of variant quality distributions. We will explore the effect of variant quality on convergence in future research.

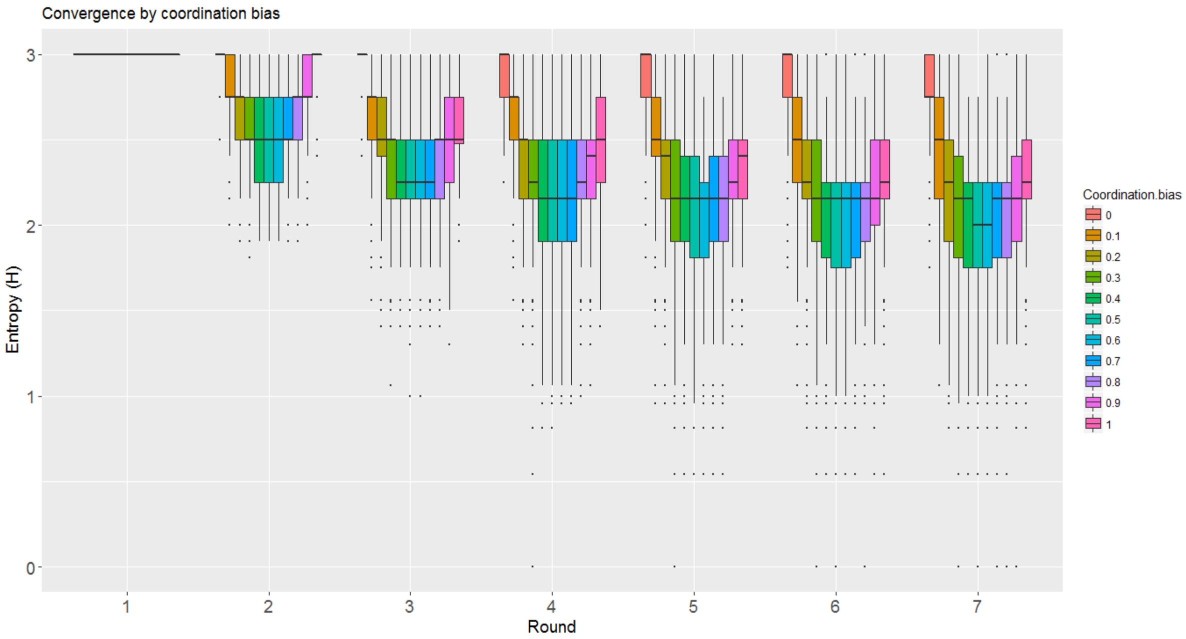

Figure 3S.: Entropy (H) by round by coordination bias. Data subset of neutral content bias. In this plot, a drift model corresponds to coordination bias = 0.5.

The probability of a non-convergent communication system (entropy lower than 1 bit) when coordination bias is 0 (fully egocentric) remains close to 100% over rounds. When coordination bias is 1 (fully allocentric) the probability of a non-convergent communication system is lower but still quite high (79.33% by round 7). The probability of a non-convergent communication system is lowest when coordination bias is neutral (Fig. 4S).

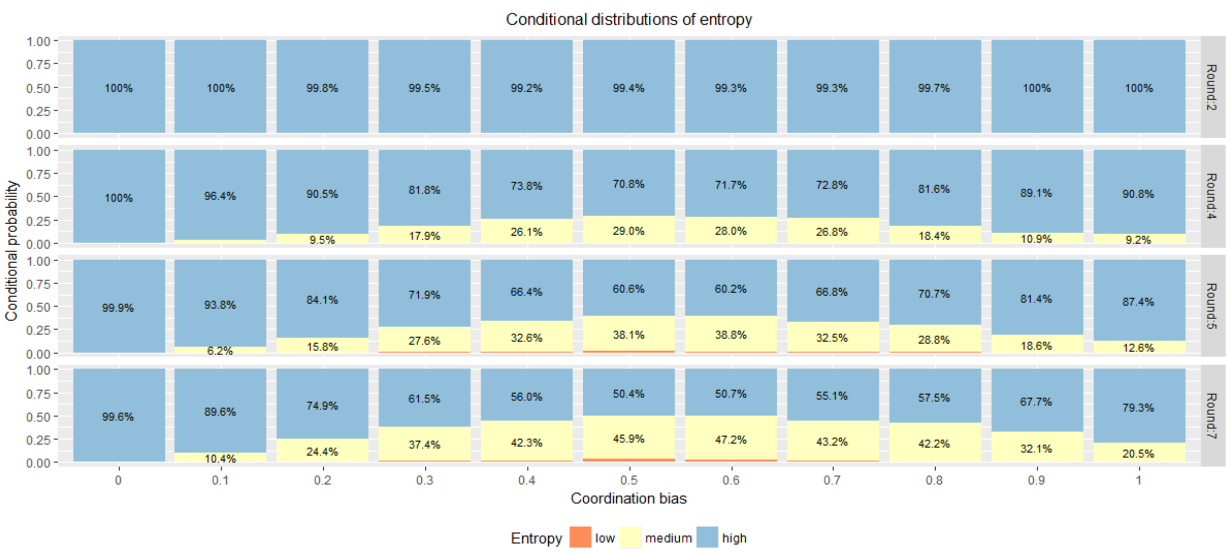

Figure 4S.: Conditional distributions of entropy over levels of coordination bias: *P(y | x)* against *x*. Bars represent conditional distributions of entropy: in blue high entropy (H ≥ 2 bits), in yellow medium entropy (1 ≥ H < 2), in red low entropy (H < 1). X-axis represents level of coordination bias, and right axis represents conditional probability.

3. Interactions between Content bias and Coordination bias

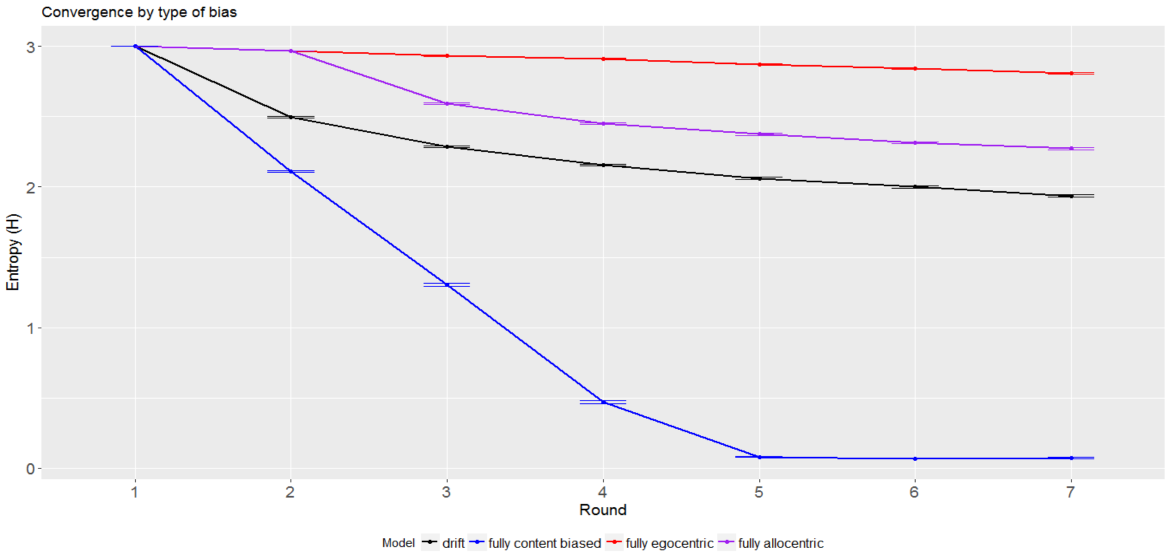

Figure 5S: Entropy by round by type of bias, averaged over the global data set, error bars indicate 95% CIs. In red: strongest egocentric bias when content bias is neutral. In purple: strongest allocentric bias when content bias is neutral. In blue: strongest content bias when coordination bias is neutral. In black: drift model (neutral content bias and neutral coordination bias). When compared to a drift model, content bias increases convergence while coordination bias, especially egocentric bias, decreases convergence.

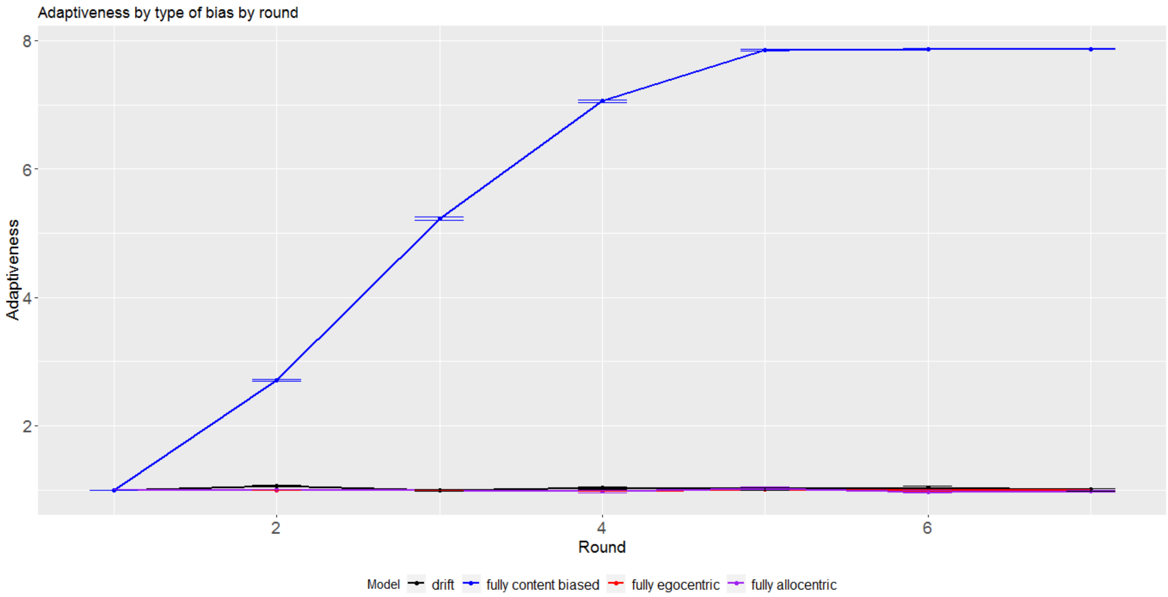

Figure 5SA: Adaptiveness by round by type of bias, averaged over the global data set, error bars indicate 95% CIs. In red: strongest egocentric bias when content bias is neutral. In purple: strongest allocentric bias when content bias is neutral. In blue: strongest content bias when coordination bias is neutral. In black: drift model (neutral content bias and neutral coordination bias). When compared to a drift model, only content bias increases adaptiveness.

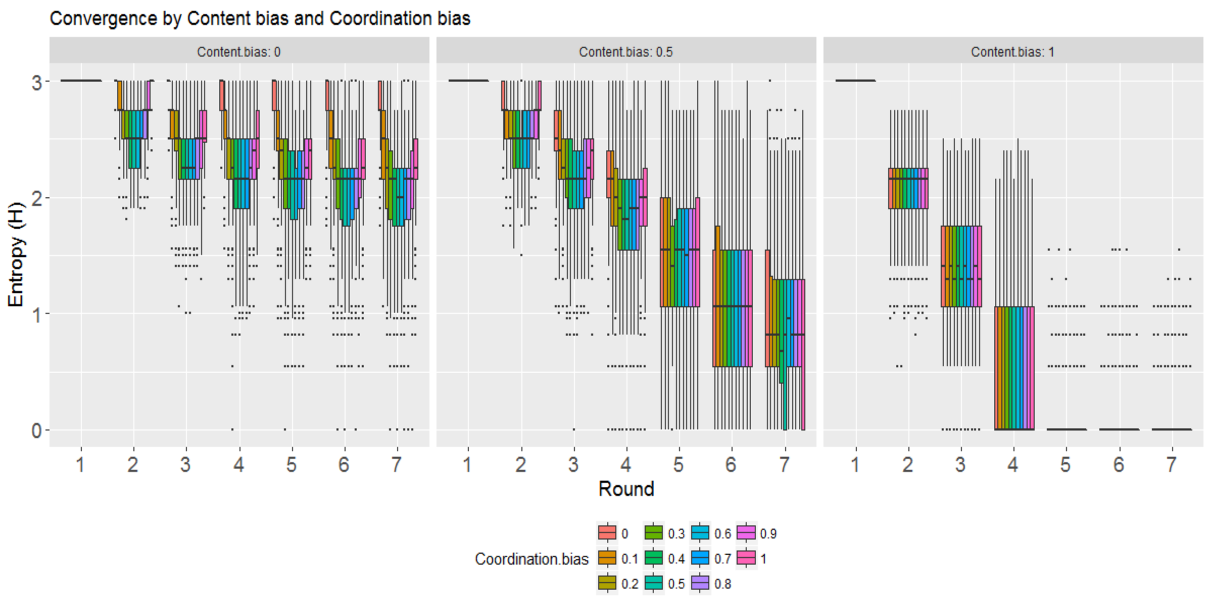

Figure 6S.: Entropy (H) by round by each combination of biases. Examples for content bias 0, 0.5 and 1. A drift model has a content bias of 0 and a coordination bias of 0.5. X-axis represents rounds from 1 to 7, Y-axis represents entropy in bits.

In all the rounds, the probability of a non-convergent communication system is higher when content bias is 0 (i.e., neutral) and coordination bias is 0 (i.e., strong egocentric bias). In contrast, the probability of a conventionalized communication system is higher when content bias is 1 (i.e., strong content bias) and coordination bias is 0.5 (i.e., neutral). Conditional distributions of entropy for all levels of coordination bias tend to equalize as content bias increases. A detailed multigraph of conditional distributions can be found in Fig. 7S.

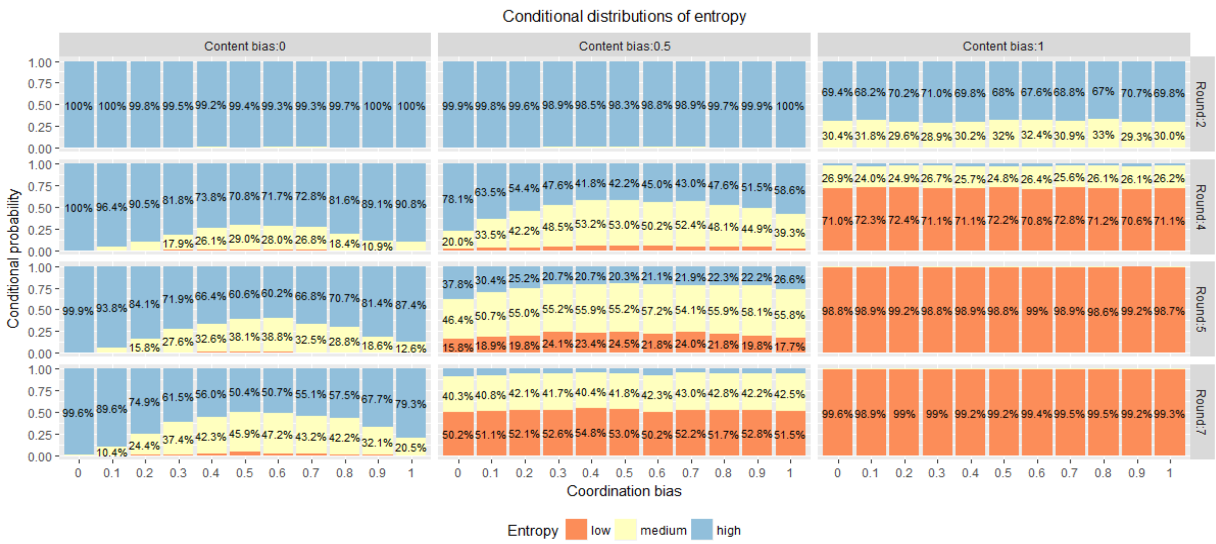

Figure 7S.: Conditional distributions of entropy over levels of content and coordination bias: *P(y | x, z)*. Bars represent conditional distributions of entropy: in blue high entropy (H ≥ 2 bits), in yellow medium entropy (1 ≥ H < 2), in red low entropy (H < 1). Y-axis represents level of content bias, X-axis represents level of coordination bias, and right axis represents conditional probability.

4. Memory

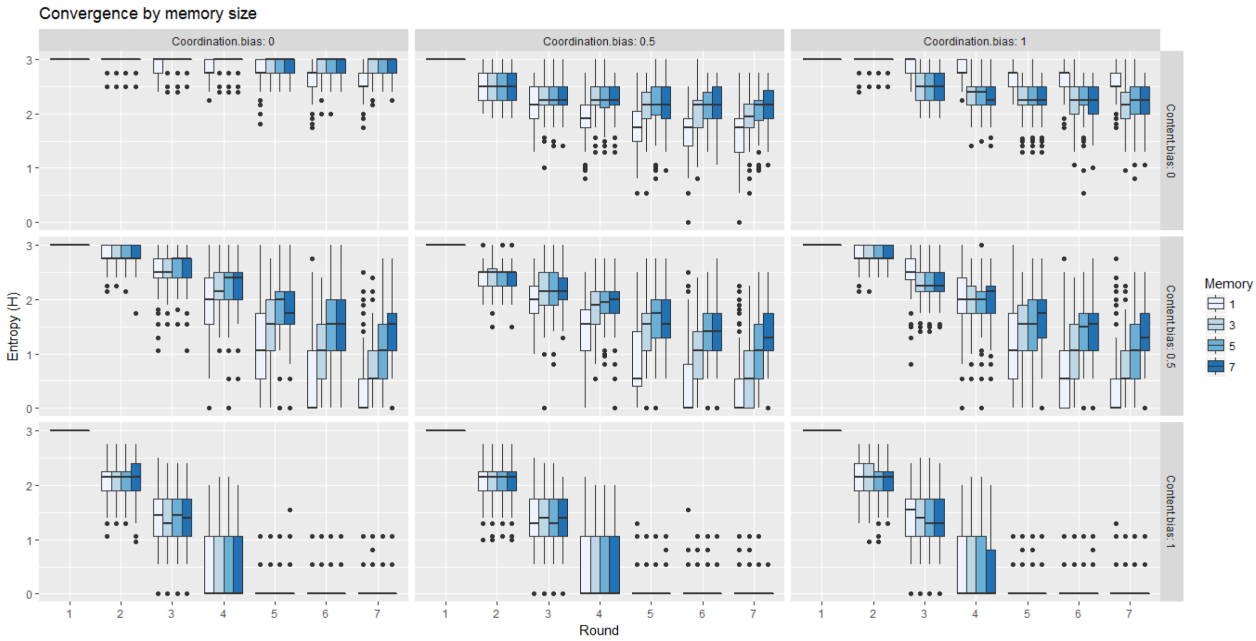

Figure 8S. Entropy (H) by round by each level of memory. X-axis represents rounds from 1 to 7, Y-axis represents entropy in bits. Coordination bias = 0 is fully egocentric and Coordination bias = 1 is fully allocentric. A drift model is represented by Coordination bias = .5 and Content bias = 0.

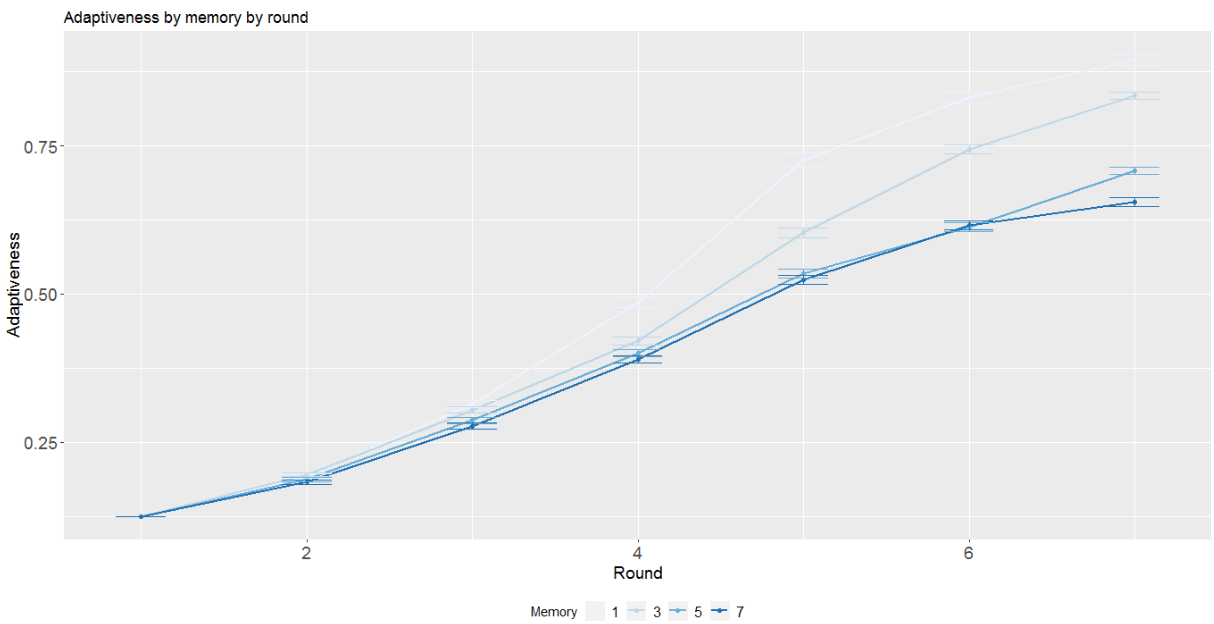

Figure 8SA. Adaptiveness by round by each level of memory, averaged over the global data set, error bars indicate 95% Cis. Memory size decreases adaptiveness of the cultural system when compared with memory = 1.

The probability of a non-convergent communication system given a memory of 1 round drops from 96.04% in round 2 to 10.53% in round 7. Given memories of 3, 5 and 7 rounds, it drops from about 96%

in round 2 to 16.46%, 24.40% and 28.62% in round 7, respectively. Differences between levels of memory are higher at round 7. The probability of a convergent communication system given a memory span of one round is 69.10%. Given a memory span of 3 rounds, it is 55.26%, 5 rounds 42.58% and 7 rounds 37.54% (see Fig. 9S).

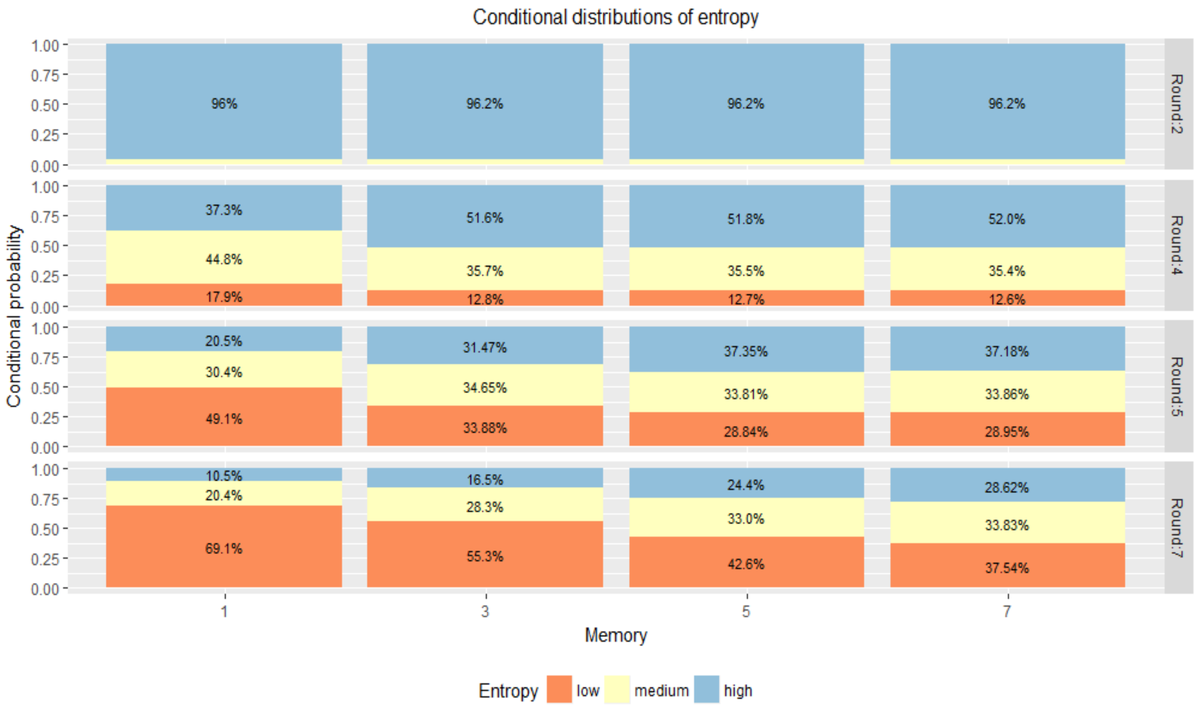

Figure 9S.: Conditional distributions of entropy over levels of memory: *P(y | x)* against *x*. Bars represent conditional distributions of entropy: in red high entropy (H ≥ 2 bits), in yellow medium entropy (1 ≥ H < 2), in blue low entropy (H < 1). X-axis represents memory limitation, and right axis represents conditional probability.

5. Connectivity dynamic

5.1 Connectivity dynamic (entropy)

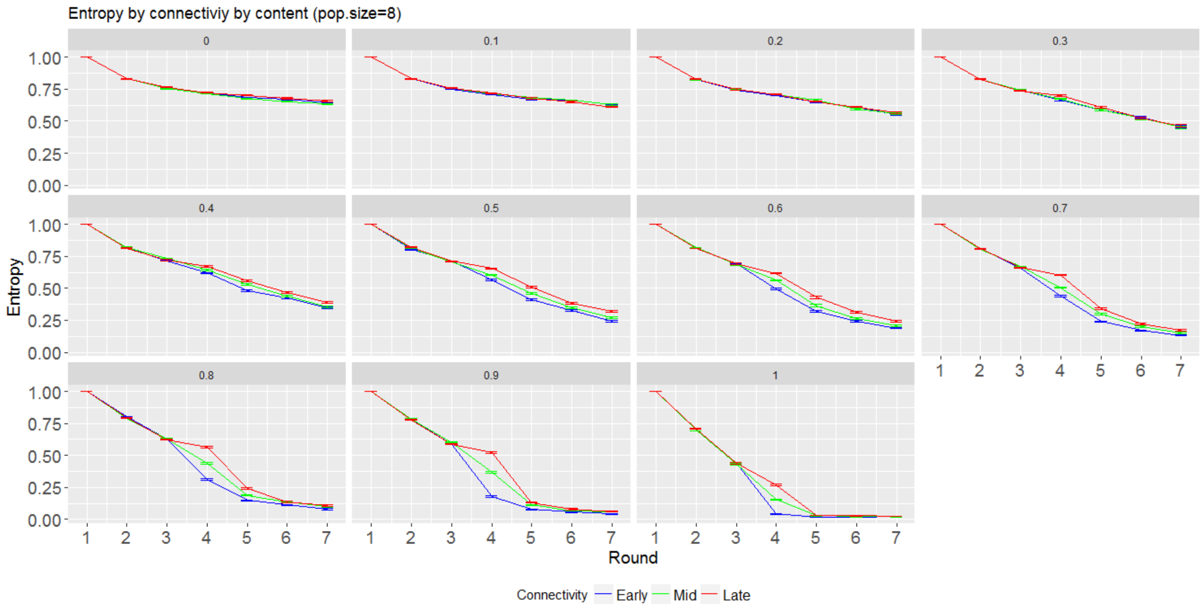

Figure 9S.: Entropy (H) by connectivity by content bias (when coordination = 0.5). X-axis represents rounds from 1 to 7, Y-axis represents mean entropy in bits (normalized values to the range of min. 0, max. 1). Drift (top-left). We run 1000 simulations for each parameter combination using 8-agent micro-societies. Error bars indicate 95% CIs.

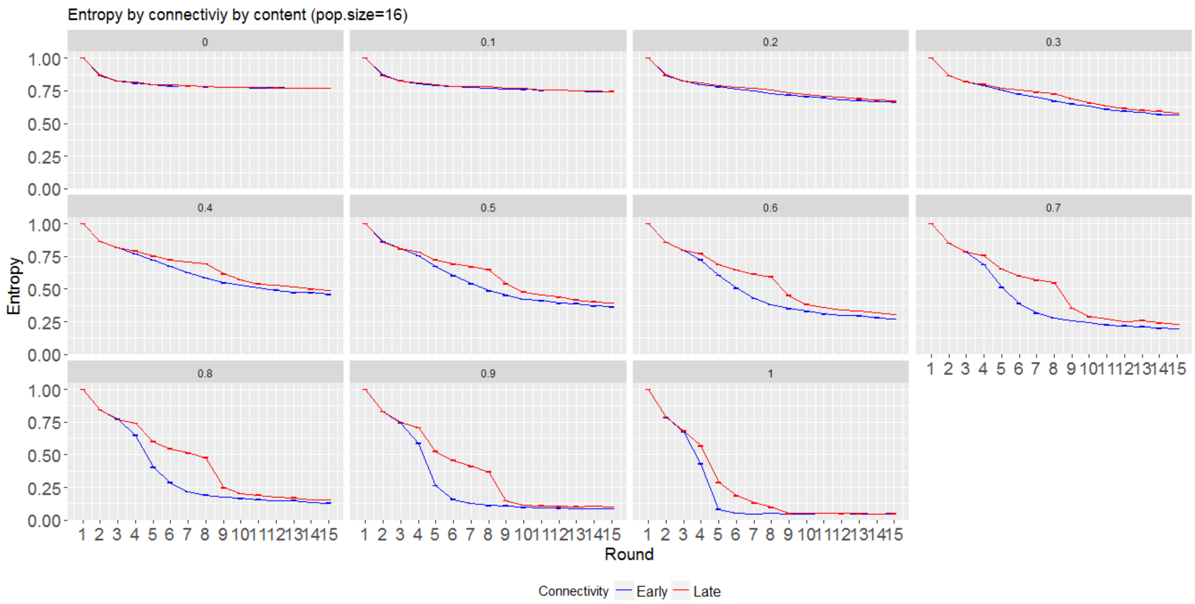

Figure 10S.: Entropy (H) by content bias by connectivity (complete memory, coordination = 0.5, pop.size=16). X-axis represents rounds from 1 to 7, Y-axis represents mean entropy in bits (normalized values to the range of min. 0, max. 1). Drift model (top-left). We run 1000 simulations for each parameter combination using 16-agent micro-societies. Error bars indicate 95% CIs.

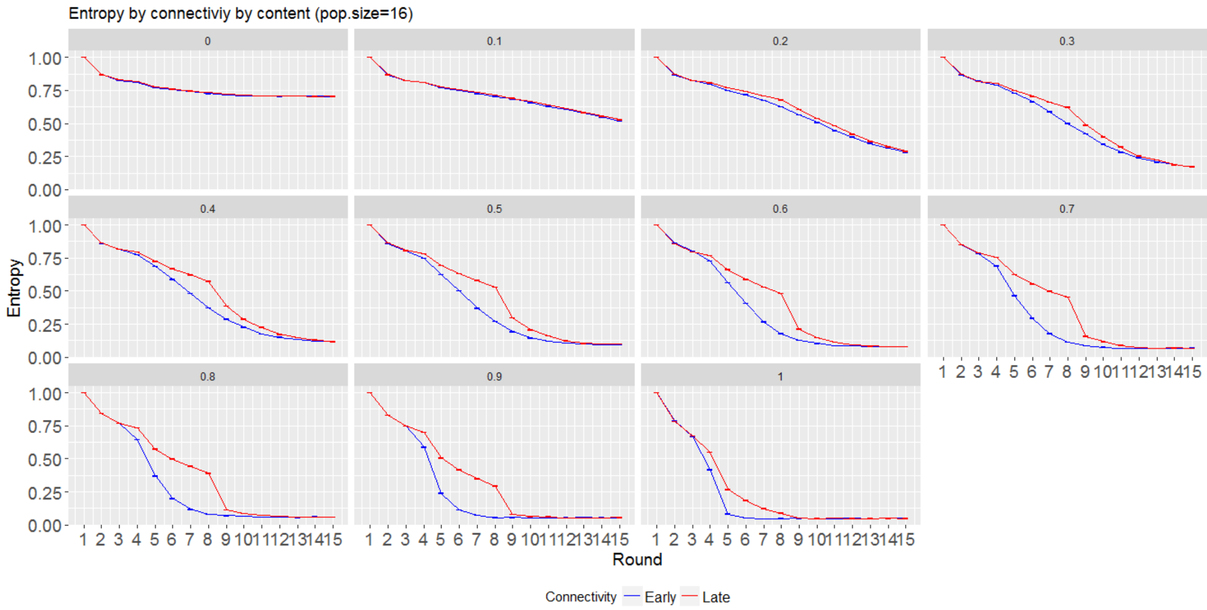

Figure 11S.: Entropy (H) by content bias by connectivity (memory = 3, coordination = 0.5, pop.size=16). X-axis represents rounds from 1 to 7, Y-axis represents mean entropy in bits (normalized values to the range of min. 0, max. 1). Drift model (top-left). We run 1000 simulations for each parameter combination using 16-agent micro-societies. Error bars indicate 95% CIs.

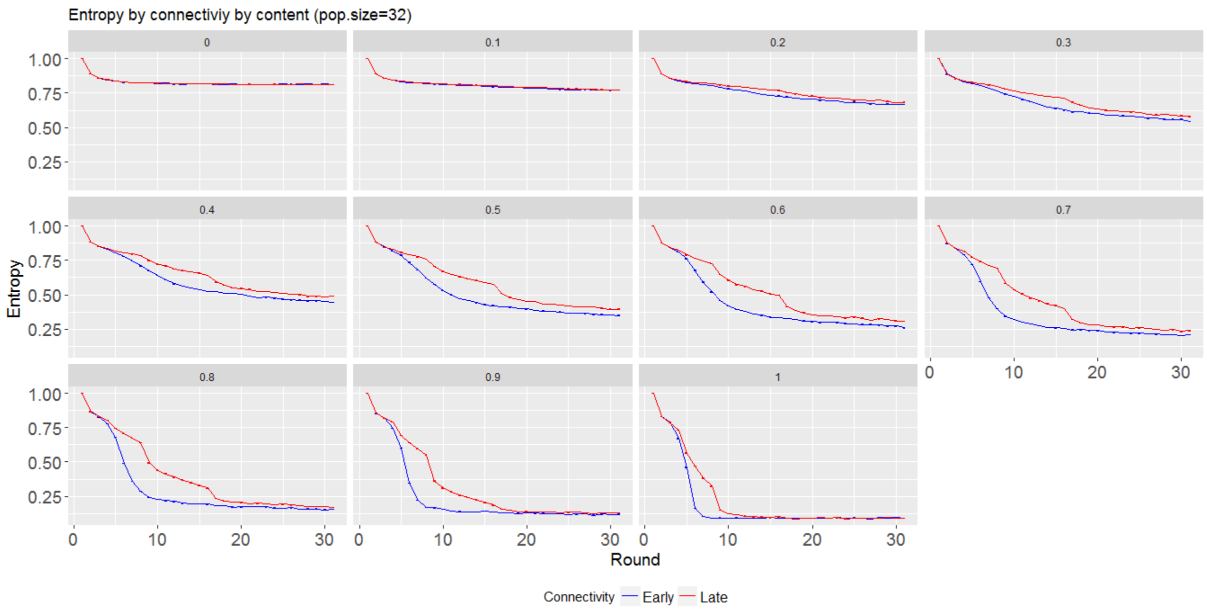

Figure 12S.: Entropy (H) by content bias by connectivity (complete memory, coordination = 0.5, pop.size=32).. X-axis represents rounds from 1 to 7, Y-axis represents entropy in bits. Drift (top-left). We run 1000 simulations for each parameter combination using 32-agent micro-societies. Error bars indicate 95% CIs.

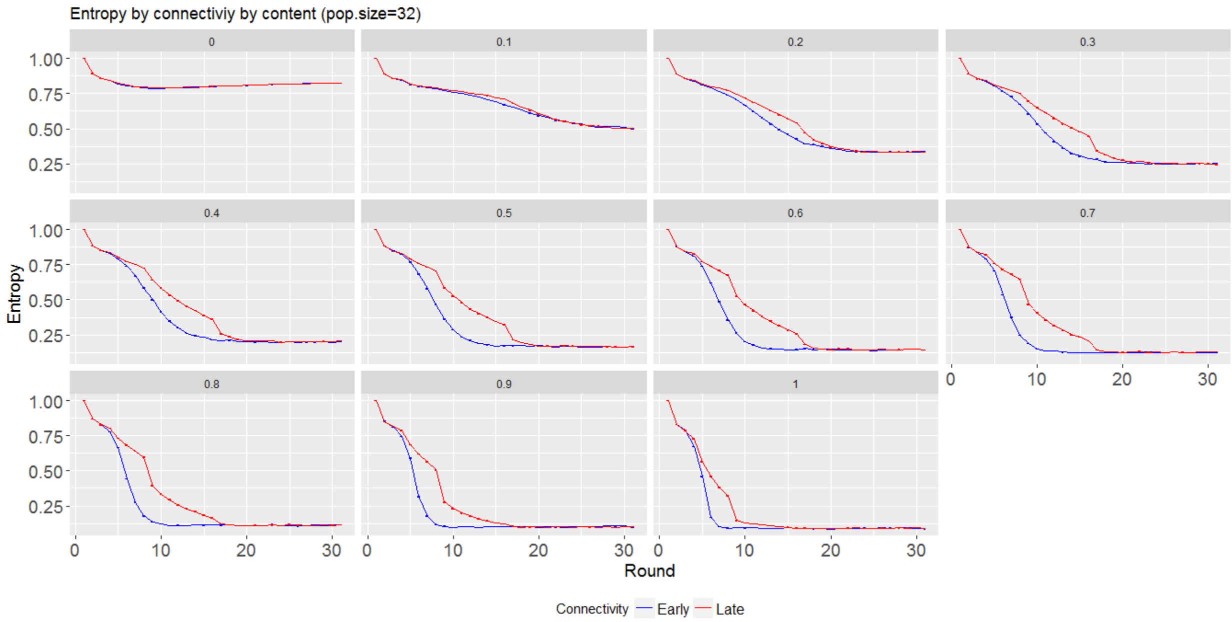

Figure 13S.: Entropy (H) by content bias by connectivity (memory = 3, coordination = 0.5, pop.size=32).. X-axis represents rounds from 1 to 7, Y-axis represents entropy in bits. Drift (top-left). We run 1000 simulations for each parameter combination using 32-agent micro-societies. Error bars indicate 95% CIs.

5.1 Connectivity dynamic (time to convergence)

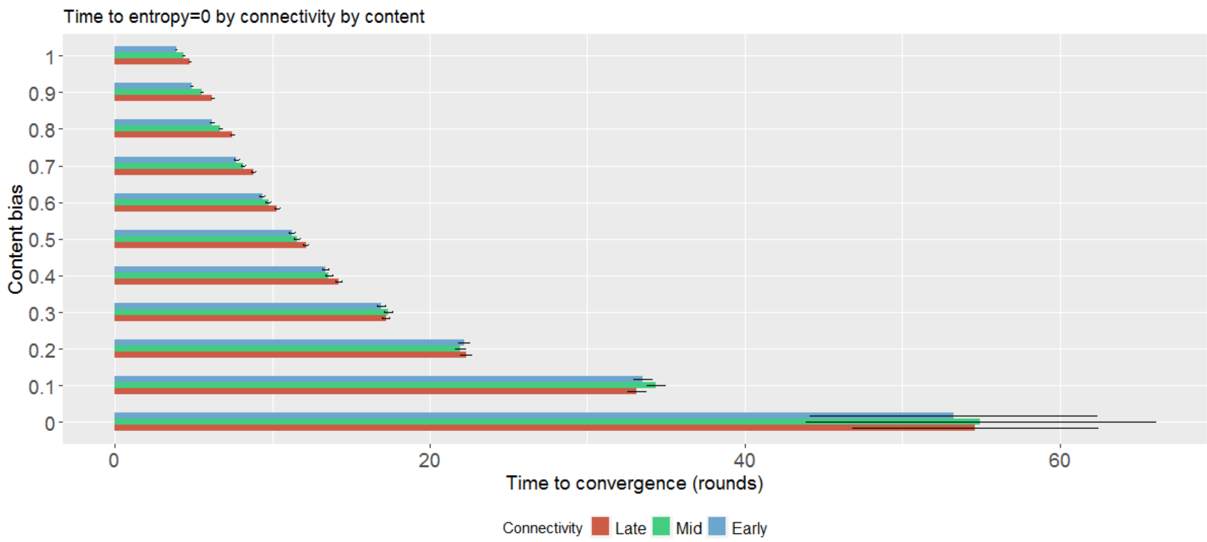

Figure 14S: Time to convergence (number of rounds until entropy 0 is reached for the first time) averaged over each level of connectivity and content bias after 350 rounds, when coordination bias is neutral. We run 1000 simulations for each parameter combination using 8-agent micro-societies. Late connectivity delays time to convergence when compared to mid and early models of connectivity. Populations with early connectivity reach convergence faster. Content bias amplifies the effect of the connectivity dynamic. Error bars indicate 95% CIs.

Note that error bars are larger under low content bias due higher variability: in these conditions, proportionally, more micro-societies reach entropy 0 by chance.

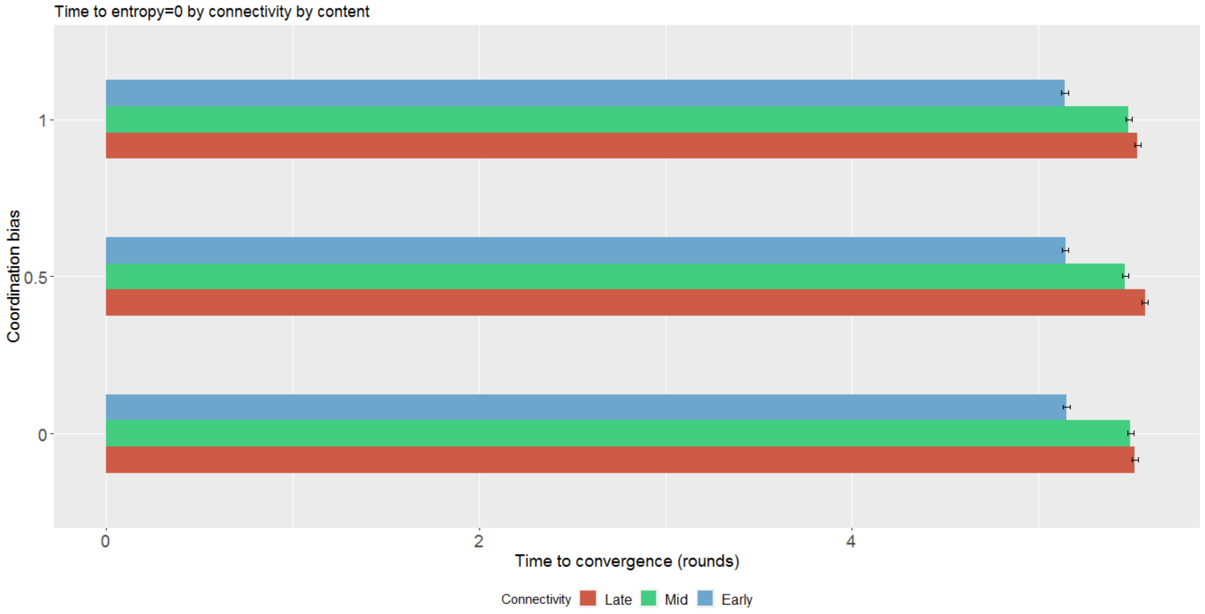

Figure 15S: Time to convergence (number of rounds until entropy is lower than 0) averaged over each level of connectivity and coordination bias after 7 rounds. Error bars indicate 95% CIs. All levels of content bias included.

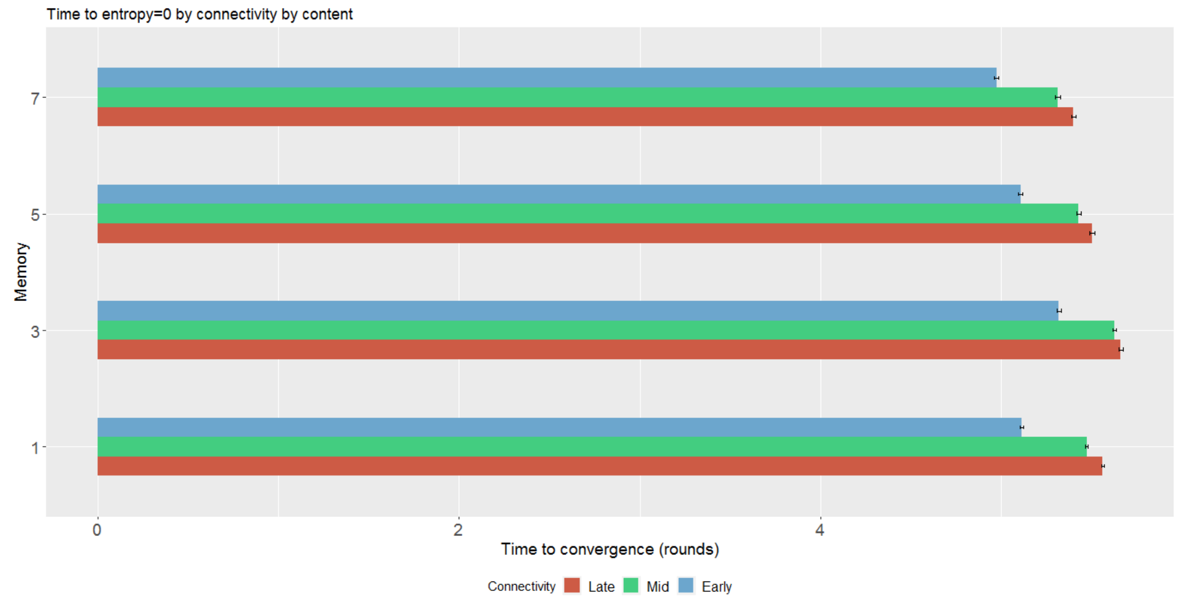

Figure 16S: Time to convergence (number of rounds until entropy is lower than 0) averaged over each level of connectivity and memory after 7 rounds. Error bars indicate 95% CIs. All levels of content bias included.

5.2 Connectivity dynamic (conditional entropy distributions)

The connectivity dynamic also plays an important role when it comes to explain entropy distributions at different times of the evolutionary process. Assuming a highly content-biased population (e.g. 0.8), the probability of a non-convergent communication system given a rapid connectivity dynamic drops from 100% in round 2 to 10% in round 4. Given a gradual connectivity dynamic, it drops from 100% by round 2 to 22% in round 4, and given a punctuated connectivity dynamic, it only drops from 100% in round 2 to 73% in round 4. Later on, as all the connectivity dynamics tend to complete a fully connected network, the probabilities of a convergent communication system tend to equalize (see Fig. 17S).

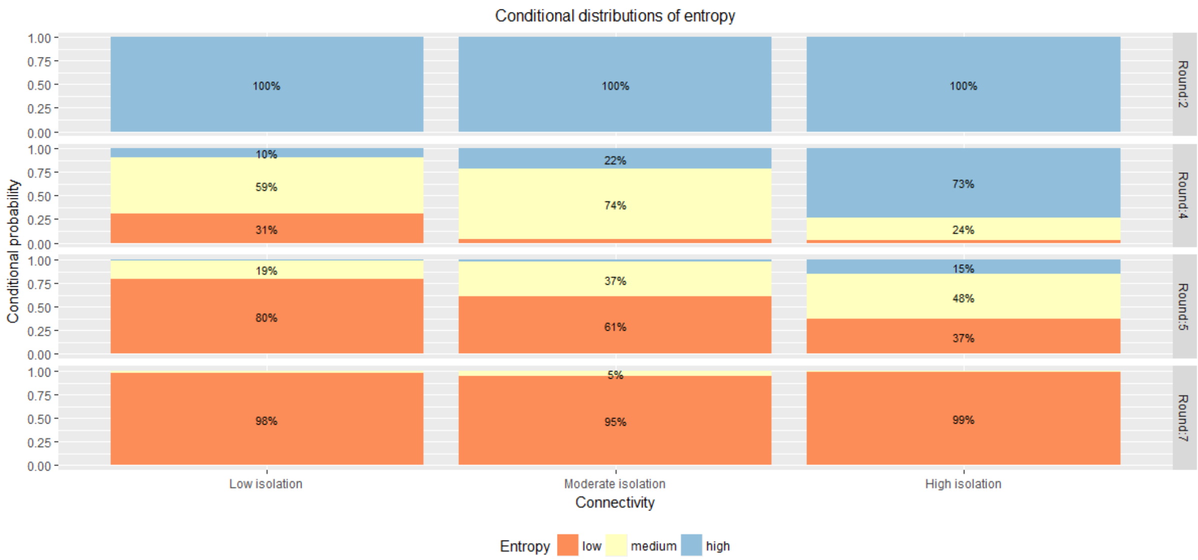

Figure 17S.: Conditional distributions of entropy over levels of connectivity dynamic: *P(y | x)* against *x*. Bars represent conditional distributions of entropy: in blue high entropy (H ≥ 2 bits), in yellow medium entropy (1 ≥ H < 2), in red low entropy (H < 1). X-axis represents level of connectivity, and right axis represents conditional probability. Results for content bias = .8.