An arms race between producers and scroungers can drive the evolution of social cognition

Michal Arbilly\textsuperscript{1,\textemdash}\textsuperscript{a,b}, Daniel B. Weissman\textsuperscript{2}, Marcus W. Feldman\textsuperscript{1} and Uri Grodzinski\textsuperscript{3}

\textsuperscript{1} Dept. of Biology, Stanford University, Stanford, CA 94305, USA
\textsuperscript{2} Institute of Science and Technology Austria, A-3400 Klosterneuburg, Austria
\textsuperscript{3} Dept. of Psychology, University of Cambridge, Cambridge CB2 3EB, UK

\textsuperscript{a} Current address: School of Biology, University of St Andrews, Bute Medical Building, Queen’s Terrace, St Andrews, Fife KY16 9TS, UK

\textsuperscript{b} Corresponding author: ma72@st-andrews.ac.uk; +44 7429 480069

\textbf{Short title:} A cognitive arms race between social foragers
ACKNOWLEDGEMENTS

This work was supported by the National Institute of Health (grant number GM28016 for M.W.F. and M.A.); the European Research Council (grant number 250152 for D.B.W.) and a fellowship from the Human Frontiers Science Program Organization (for U.G.).

We thank N.J. Boogert for helpful comments on the manuscript.
An arms race between producers and scroungers can drive the evolution of social cognition

The ‘social intelligence hypothesis’ states that the need to cope with complexities of social life has driven the evolution of advanced cognitive abilities. It is usually invoked in the context of challenges arising from complex intragroup structures, hierarchies and alliances. However, a fundamental aspect of group living remains unexplored as a driving force in cognitive evolution: the competition between individuals searching for resources (producers) and conspecifics that parasitize their findings (scroungers). In populations of social foragers, abilities that enable scroungers to steal by outsmarting producers, and those allowing producers to prevent theft by outsmarting scroungers, are likely to be beneficial and may fuel a cognitive arms race. Using analytical theory and agent-based simulations we present a general model for such a race that is driven by the producer-scrounger game, showing that the race’s plausibility is dramatically affected by the nature of the evolving abilities. If scrouning and scrounging-avoidance rely on separate, strategy-specific cognitive abilities, arms races are short-lived and have a limited effect on cognition. However, general cognitive abilities that facilitate both scrouning and scrounging-avoidance undergo stable, long-lasting arms races. Thus, ubiquitous foraging interactions may lead to the evolution of general cognitive abilities in social animals, without the requirement of complex intragroup structures.

Keywords: game theory; social foraging; social intelligence hypothesis; intraspecific arms race
INTRODUCTION

Recent decades have seen great interest in social cognition and its evolution, due largely to the non-trivial nature of such abilities (e.g., putting one’s self in others' shoes), as well as the idea that coping with social challenges may underlie the evolution of general intelligence (Shettleworth 2010). Indeed the latter suggestion, known as the ‘social intelligence hypothesis’ (Jolly 1966; Humphrey 1976; Byrne and Whiten 1988), relies heavily on the finding that species exhibiting advanced cognitive abilities often maintain elaborate social structures. While this hypothesis initially referred to humans and other primates, it has also been related to advanced cognition in other species, including corvids (Emery and Clayton 2004), hyenas (Holekamp 2007), and cetaceans (Marino 2002). However, regardless of such elaborate social structures, group-living animals face a more fundamental challenge which is often ignored in this context: social foraging (Giraldeau and Caraco 2000).

Foraging together for resources is a ubiquitous feature of group living, observed across taxa from insects to humans; it is perhaps one of the most common forms of social interaction, as it spans fundamental aspects of life such as food and shelter. Social foraging interactions have been framed in terms of the producer-scrounger (PS) game, in which individuals have the option either to produce (i.e. independently search for) resources, or scrounge them from producers (Barnard and Sibly 1981; Barnard 1984; Giraldeau and Caraco 2000; Giraldeau and Dubois 2008).
While scrounging saves the time and energy that must be invested in order to produce resources, it requires a sufficiently high frequency of producers in the population to be beneficial. The negative, frequency-dependent selection operating on these two strategies results in a mixed evolutionarily stable strategy (ESS) (Barnard and Sibly 1981; Barnard 1984; Giraldeau and Caraco 2000; Giraldeau and Dubois 2008). In such populations, selection can be strong enough to give rise to a suite of scrounging-avoidance tactics by producers, and consequent counter-tactics by scroungers (Barnard 1984; Coussi-Korbel 1994; Emery and Clayton 2001; Flynn and Giraldeau 2001; Bugnyar and Kotrschal 2002; Held et al. 2002; Bugnyar and Heinrich 2006; Shaw and Clayton 2013). Thus, selection on social foragers to outsmart each other can lead to an intraspecific evolutionary arms race (Dawkins and Krebs 1979) that results in increased cognitive abilities (Barnard 1984; Bugnyar and Kotrschal 2002; Grodzinski and Clayton 2010).

In many species of social foragers, the PS game may have selected for cognitive adaptations that involve plastic responses to the presence of others (an ‘audience effect’) (Barnard 1984; Byrne and Whiten 1988; Coussi-Korbel 1994; Norris and Freeman 2000; Emery and Clayton 2001; Flynn and Giraldeau 2001; Bugnyar and Kotrschal 2002; Held et al. 2002; Bugnyar and Heinrich 2006; Shaw and Clayton 2013). For example, in spice finches, as well as pigs and gorillas, producers keep their distance from potential scroungers (Byrne and Whiten 1988; Flynn and Giraldeau 2001; Held et al. 2002); in mangabeys and chimpanzees, producers lead scroungers away from food (Byrne and Whiten 1988; Coussi-Korbel 1994); scrub jays return to re-cache, in private, food items they have been observed by conspecifics to have been hiding (Emery and Clayton 2001);
Eurasian jays attempt to prevent auditory information of their caching activities from reaching potential scroungers (Shaw and Clayton 2013); scrounging ravens watch caching from a distance and delay their approach until the cacher (producer) has left (Bugnyar and Heinrich 2006), and scrounging chimpanzees may hide to watch conspecifics recover food, and emerge from hiding to steal it (Byrne and Whiten 1988).

While success in the PS game may be influenced by a number of traits, from body size to dominance ranking (Giraldeau and Beauchamp 1999), these observations suggest that potential targets for adaptation are likely to include data-processing and decision-making abilities. However, the cognition underlying such abilities is likely to entail a cost, which may be developmental, physiological and/or derived from prolonged data processing (e.g. Burger et al. 2008).

Here, we examine the conditions under which mutations in the cognitive apparatus that increase performance in the PS game provide sufficient benefit to outweigh such costs, and analyze the consequences of evolving general versus strategy-specific cognitive abilities. As the nature of the cognitive abilities involved in the aforementioned examples is far from clear, it is impossible to model them in any detail without restricting the generality of the model. To avoid this, we model these cognitive abilities simply as traits affecting the performance of producers and scroungers competing against each other (as detailed below). Consequently, our model is in fact much more general, and concerns any such traits.
We model a population of social foragers playing the PS game. We consider both the case of individuals playing pure social foraging strategies, and the more realistic case of mixed strategies. For simplicity we describe the pure strategy model first, and then extend it to include mixed strategies.

Basic model and the scrounging success probability function

Each generation consists of multiple rounds of foraging, and in each round some fraction of producers finds food and receives a payoff. A producer that finds food experiences scrounging attempts with probability $f$, which, if scrounging is successful, causes the payoff to be evenly split between the scrounger and the producer. The probability that a scrouning attempt is successful, $\sigma$, is determined by the difference $d$ in cognitive abilities between the scrounger and the producer: scroungers with relatively advanced cognitive abilities are more often successful at obtaining food, whereas producers with relatively advanced abilities are more often successful at avoiding loss of food to scroungers. It is then reasonable that $\sigma$ should increase monotonically with the cognitive difference $d$. We model this effect by assuming that $\sigma$ is a logistic function of $d$:

$$\sigma(d) = a + (1 - a) / (1 + e^{-ad})$$

The parameter $0 < a < 1$ determines the size of the effect that cognition has on scrounging success: it represents the lowest possible scrounging success rate, which occurs when a producer has an infinitely higher cognitive level than a scrounger ($i.e.$ $d \rightarrow -\infty$). In other words, if $a$ is large the influence of cognition should be weak, as scrounging is then likely to succeed regardless of the exact difference in cognitive abilities. We assume that the probability of successful scrounging is not
determined solely by cognitive abilities, but that these abilities are still significant in determining scrouning success \( i.e. \ 0 < a < 1 \). The parameter \( s \) determines the size of the effect that a single cognitive mutation has on the probability of successful scrouning:

each mutation changes \( \sigma \) by \( \approx s \) until it saturates at some maximum or minimum value for \( |sd| >> 1 \). The effects of these assumptions can be seen in Figure 1, which shows scrouning success probability as a function of \( d \) for different values of \( a \) and \( s \).

**Effect of cognitive mutations on fitness**

Given the probability of scrouning success \( \sigma \), the relative fitness advantage \( \alpha_p \) of a mutation that changes producers’ cognitive ability by \( \delta \) is given by

\[
\alpha_p(\delta) = e^{-\delta \gamma} \left[ 1 - f \sigma(d - \delta)/2 \right]/[1 - f \sigma(d)/2] - 1, \text{ where } \gamma \text{ scales the cost of increasing}
\]

cognitive ability. The advantage \( \alpha_s \) of a mutation that changes scroungers’ cognitive ability by \( \delta \) is given by

\[
\alpha_s(\delta) = e^{-\delta \gamma} \sigma(d + \delta)/\sigma(d) - 1.
\]

To be advantageous, the potential benefits of cognition-increasing mutations in relation to the PS game must outweigh their cost (corresponding roughly to \( s >> \gamma \)), which we will assume in what follows. However, this does not guarantee that increasing cognition is always favored, because the advantage of a mutation that increases cognitive ability in a producer (scrounger) depends on its current cognitive level relative to scroungers (producers) (Fig.2). When scroungers are slightly smarter than producers (i.e. \( d \) is small and positive), producers are selected to increase their cognitive level. When producers are slightly smarter than scroungers (\( d \) is small and negative) scroungers are selected to increase their cognitive level. Thus, small differences in cognitive level support an
evolutionary arms race between social foraging strategies (Fig. 2). On the other hand, cognitive differences that are too large have remarkably different consequences. If the producers are substantially smarter than the scroungers ($d$ is large and negative), or vice versa ($d$ is large and positive), the probability of successful scrounging, $\sigma$, is only slightly affected by further mutations, because it is close to either of its asymptotic values ($a$ or 1, respectively). Since the benefit of an increased cognitive level is low in such cases, it is outweighed by the cost, and selection will favor decreased cognitive levels (Fig. 2).

Effect of specialized vs. generalized cognition on the race

We consider two possibilities for the influence of cognition on the PS game. In the generalized cognition model (GCM), a single cognitive trait, ‘$C$’, determines both the ability to successfully scrounge and the ability to avoid being scrounged when producing. In the specialized cognition model (SCM), one trait, ‘$Cs$’, determines scrounging ability, while another, ‘$Cp$’, determines the ability of producers to avoid being scrounged. Comparing both types of models analytically and using agent-based simulations, we show that the SCM is unable to support a consistent arms race and results either in scroungers’ extinction or in a race to decrease cognitive level (a ‘backwards race’). The GCM, in contrast, exhibits persistent arms races, a result independent of the specific conditions.

We use agent-based computer simulations to demonstrate the population dynamics of this competition. Here, fitness is determined through explicitly-modeled repeated PS interactions rather than through the abovementioned fitness functions $\alpha_p$ and $\alpha_s$. A
population commences at some baseline cognitive level, labeled ‘level 0’ for
convenience. A generation consists of \( T \) repeated foraging interactions as described
above, and at each interaction scroungers are randomly assigned to successful producers
(one scrounger per producer); the probability of successful scrounging at each interaction
is determined by the relative cognition of the specific producer and scrounger involved,
using the aforementioned function \( \sigma(d) \). Mutations increasing (+1) or decreasing (-1)
cognitive level, as well as mutations in social foraging strategy (transforming a producer
into a scrounger or vice versa) occur randomly at a low rate, \( \mu \), at each generation. The
number of offspring per individual is set probabilistically, based on the relative payoff
accumulated through its lifetime (see SI section 2 for full details of the simulations).

In a population initially made up of individuals with equal cognitive abilities, the
scroungers’ relevant cognitive ability (\( C \) in the GCM, \( Cs \) in the SCM) typically increases
in either type of model (Fig. 3). This increases the probability that scrounging is
successful, intensifying the selective pressure on producers to avoid being scrounged and
leading producers’ relevant cognitive ability (\( C \) in the GCM, \( Cp \) in the SCM) to rise.
Improved producer cognitive ability, in turn, puts pressure on scroungers to re-adapt, and
the consequent positive feedback loop leads to a continuous evolution of increased
cognitive abilities in both producer and scrounger populations (Fig. 3). The rate of this
increase depends on the rate and magnitude of cognitive mutations: higher mutation rates
and higher values of \( s \) result in faster-evolving races (see SI). In the SCM, the escalation
in relevant cognitive abilities is accompanied by a decrease in the unused cognitive
abilities (\( Cs \) for producers, \( Cp \) for scroungers), due to their cost.
Such arms races occur in both the GCM and the SCM, and are temporarily stable as long as cognitive differences between foraging types are small, consistent with our analysis above showing that small $d$ values support an evolutionary arms race. However, the arms race is interrupted when either foraging type acquires a large cognitive advantage over the other ($|sd| >> 1$); such an advantage emerges stochastically due to the random processes in the simulation and finite population size. If producers have a sufficiently large advantage, the (unsuccessful) scroungers cannot obtain resources and are driven to extinction. Once this happens, they can only reappear through mutation that converts a producer into a scrounger. In the SCM, such mutants will have the high $C_p$ and low $C_s$ values typical of producers, but since this makes them unfit as scroungers, scroungers cannot recover from extinction (Fig. 3a). In the GCM, however, a mutant’s high cognitive ability $C$, inherited from its producer parent, will make it a good scrounger. This enables scroungers to re-invade the population, reestablishing the cognitive arms race from the current cognitive level of the population (Fig. 3b).

A large cognitive advantage for scroungers, in contrast, will not lead producers to extinction, due to the frequency-dependence of the PS game. Instead, mutations that decrease producers’ cognitive level will be favored because the benefits in reducing cognitive costs will outweigh their effect on scrouning-avoidance success (which is minimal under these conditions since scroungers are much smarter). Once producers’ cognitive levels are reduced, selection will act on the scroungers to follow suit for similar cost-saving reasons, resulting in a ‘backwards’ race. This ‘backwards’ race scenario is
likely to occur and escalate in SCM populations (Fig. 3c), but in GCM populations a
large cognitive advantage for scroungers will quickly be reduced by scroungers mutating
into producers while retaining their high C levels, thus reestablishing the race.

Mixed strategies

So far we have considered pure producers and pure scroungers. In nature, however, the
PS trait is usually manifested as a mixed strategy, and individuals have been observed to
employ both strategies to varying degrees based on their previous experience and on
social and environmental conditions (Mottley and Giraldeau 2000; Katsnelson et al.
2008; Morand-Ferron and Giraldeau 2010; Katsnelson et al. 2011). In simulations of both
of our models, inclusion of mixed strategies yields qualitatively similar results to those
described above: GCM races persist, while SCM races are bound to collapse. As in the
pure-strategy case of the SCM, gaps between $C_p$ and $C_s$ arise stochastically. If $C_s$
becomes much larger than $C_p$, the latter decreases (as in the pure case) to avoid cognitive
costs, resulting in a ‘backwards race’ (Fig. 3d). If $C_p$ becomes much larger than $C_s$,
selection favors foraging strategies that produce as often as possible, and scrouning
disappears from the population (Fig. 3e), as in the pure case. This disappearance of
scrouning from the population as its adaptive value decreases is plausible given that in
nature, social foraging strategies can be adjusted to provide better solutions for changing
environmental conditions (Mottley and Giraldeau 2000).
As shown above, arms races involving general cognitive abilities are not limited by the instability and short-life typical of those involving specialized cognition. However, it does not follow that these arms races will continue forever. For example, when cognition-costs become too high compared to their benefits the population may go extinct. As costs become too high the population may also become prone to invasion by migrants with baseline cognitive levels; such an invasion is possible because these migrants, despite their poor cognitive abilities, do relatively well altogether as they do not suffer such high cognitive costs. In this case, a cyclic pattern of escalation and collapse may emerge, as the population repeatedly regresses to the cognitive baseline and then restarts the arms race (Fig 3f). Alternatively, a general cognitive ability may co-evolve with other traits (such as foraging efficiency or diet, in our case), changing the very parameters considered here that govern the evolution of social cognition. Interestingly, increased general cognition resulting from the race may have pleiotropic benefits, such as enabling the exploitation of new food sources or habitats, which could themselves outweigh the costs of cognition. Conversely, if producers become better at exploiting food sources, producing may become much more profitable than scrounging. The consequent low frequency of scroungers will make the PS game less important, slowing down the cognitive arms race or drawing it to an end.

DISCUSSION

Our results suggest that a cognitive arms race improving performance of players in the producer-scrounger game can persist and escalate, but only if it involves a general
cognitive ability competing against itself. Arms races between two separate abilities may escalate temporarily, but are bound to collapse. In the present formulation, the increased stability of arms races, when they involve generalized rather than specialized cognitive abilities, is independent of the specific details of our model. Indeed, arms races involving a single trait should generally be more stable than those between two (or more) traits that mutate and evolve separately, since destabilizing asymmetries will arise less frequently in the former. That intra-species arms races should tend to persist for longer than inter-species ones is one possible implication.

Intraspecific evolutionary arms races are often mentioned in the context of sexual selection (Dawkins and Krebs 1979), sexual conflicts (Chapman et al. 2003), brood parasitism (Petrie and Møller 1991) and parent-offspring conflict (Kilner and Hinde 2008). Social foraging adds a further, rather general framework within which multiple, unrelated traits may each evolve by racing “against itself”. While the model we present here was designed with cognition in mind, it is, as stated above, certainly not limited to cognitive abilities. It appears that the producer-scrounger game can facilitate the evolution of many traits that improve scrounging and scrounging-avoidance: body-size, aggressiveness, motivation, and more.

That our model applies to a range of traits affecting interactions among foragers may indeed suggest that improved cognition is not the only possible consequence of social living. However, we believe that cognition might be especially affected, for two reasons. First, as mentioned in the introduction, there is strong evidence that cognitive abilities
such as information processing, learning, and decision-making have strong effects on
scrounging and scrounging-avoidance. Second, many other relevant traits, such as body
size, are likely to be under strong stabilizing selection as the cost of increasing them
becomes too high. For example, developing and maintaining a large body size requires
high energy intake and may entail a higher risk of predation (e.g. Blanckenhorn 2000;
Quinn, Hendry, and Buck 2001; Rotella, Clark, and Afton 2003; Bonduriansky and
Brassil 2005; Herczeg, Gonda, and Merilä 2009); an escalation of aggressive behavior is
likely to result in high rates of injury and death, and is evolutionarily unstable (Maynard
Smith and Price 1973). Improved cognitive abilities are likely to involve a fitness cost as
well, as demonstrated in some species of insects (Burger et al. 2008; Snell-Rood,
Davidowitz, and Papaj 2011). Seemingly complex abilities can be achieved through
surprisingly simple neural structures (Chittka et al. 2012), but it is quite possible that
brain size evolution is constrained by the energetic costs of maintenance (Isler and van
Schaik 2006). However, maintaining a large body to an extent that will make a difference
in success in the game most probably requires greater energy than maintaining a
cognitive tweak that will achieve the same difference. In other words, it seems likely that
cognitive mutations that improve foragers performance in the PS game will tend to cost
less than an increase in body size or aggressive behavior that could provide the same
improvement. Thus, while our model describes a scenario applicable to many traits,
cognition may be one of the few for which the benefit and cost parameters fall in the
region supporting an arms race.
We have seen that the arms race in a trait is only stable if the trait contributes to both scrounging and scrounging-avoidance; what general cognitive abilities, then, might serve both of these tactics? The requirement that the ability must be useful for such distinct behaviors strongly suggests some form of social cognition. An example consistent with our model is the strategies used by some corvid species to protect food-caches from being scrounged and to successfully pilfer others’ caches. It has been suggested that these strategies involve a general cognitive ability, and perhaps even some form of Theory of Mind (Bugnyar and Kotrschal 2002; Dally, Clayton, and Emery 2006; Grodzinski and Clayton 2010). Indeed, the finding that some cache-protection strategies require previous experience in pilfering (Emery and Clayton 2001) lends some support to this notion (Bugnyar and Kotrschal 2002; Dally, Clayton, and Emery 2006; Grodzinski and Clayton 2010). Our analysis shows that from an evolutionary perspective, evidence for advanced cognitive abilities makes it more likely that they are general rather than due to cognitive mechanisms that serve caching and pilfering separately, or else they would probably not have evolved.

Decades ago, an arms race of cognitive abilities (‘runaway intellect’) was proposed within the context of the social intelligence hypothesis (Humphrey 1976). It was also suggested that Theory of Mind itself is likely to involve increasing degrees of complexity (Premack 1988). We show that the fundamental and ubiquitous interactions between social foragers can give rise to an arms race of general cognitive abilities. This raises the question of whether traits such as attribution of knowledge and intentions to others have
evolved in a much wider range of taxa than currently suggested and, if not, what has inhibited them.

REFERENCES


FIGURE LEGENDS

Fig. 1. Successful scrounging probability, $\sigma$, for different values of cognition effect size $a$, and cognitive mutation effect size $s$. Dashed black line: $a = 0.7, s = 1.5$; solid gray line: $a = 0.5, s = 0.5$; solid black line: $a = 0.5, s = 1.5$; Dashed gray line: $a = 0, s = 1.5$.

Fig. 2. The selective advantage $\alpha_p$ to producers (solid line) and $\alpha_s$ to scroungers (dashed line) accorded by a (+1) cognitive mutation, as a function of $d$, the cognitive difference in favour of scroungers. The proportion of each foraging strategy is fixed at the proportion found to evolve in computer simulations (0.7 producing, 0.3 scrouning).

Parameters values used: $s = 1.5, a = 0.5, \gamma = 0.05$.

Fig. 3. Examples of GCM and SCM population dynamics in agent-based simulations, under various conditions. Black and white panels show producer frequency over time; color panels show mean cognitive level over time: GCM (b, f) - two lines representing mean $C$ levels for producers (red) and scroungers (teal); SCM (a, c-e) – four lines representing mean level of specialized cognitive ability for producing, $C_p$, in producers (red) and scroungers (blue) and mean level of specialized cognitive ability for scrounging, $C_s$, for producing, $C_p$, in producers (orange) and scroungers (teal). In mixed-strategy simulations (d-e), 0-50% producing is included under ‘scrooungers’, 60-100% producing is included under ‘producers’. Where red line is not visible it is hidden by the teal or blue lines. In all simulations population size $n = 100$; cognitive cost is a fractional deduction of size $\gamma = C / 100$ in GCM, $\gamma = (C_p + C_s) / 100$ in SCM; scrounging
success baseline probability $a = 0.5$; mutation rate $\mu = 1/n$ for all genes; mutations in $C/Cp/Cs$ increase or decrease cognitive ability by 1. Note that the $y$ axis scales in colored panels vary. (a) SCM, pure producing/scrounging (PS), $s = 1.5$. (b) GCM, pure PS, $s = 1.5$. (c) SCM, pure PS, $s = 0.5$. (d) SCM, mixed PS, $s = 1.5$. (e) SCM, mixed PS, $s = 1.5$. (f) GCM, pure PS, random inwards migration of individuals with baseline cognitive level ($C=0$).
Figure 1

Probability of successful scrounging, $\sigma$

Advantage of scroungers, $d$
Figure 2

Selective advantage of cognitive mutation of size +1

Scroungers' cognitive advantage, d

Scroungers' cognitive advantage, d

Scroungers' cognitive advantage, d
Supplementary Information

We present below a mathematical analysis of the model for small parameter values, complemented with computer simulations of evolving populations for a larger parameter range.

Symbols used throughout the paper and supplementary material are summarized in table S1.

1. Analytical model

We focus on a population of $n$ individuals, a variable fraction $\varphi$ of which are pure scroungers, and $1 - \varphi$ of which are pure producers. Each individual has a cognitive level $C$ that determines how good it is at scrouning (for scroungers) or avoiding being scrounged (for producers). Each generation comprises many foraging steps. In one such step, a fraction $\rho$ of the producers each finds one unit of food. Each scrounger is then randomly assigned to one of the producers that found food and attempts to scrounge. If multiple scroungers are assigned to the same producer, only one (chosen randomly) is allowed to attempt to scrounge. With probability $\sigma$ (given above), the scrouning attempt is successful and the producer and scrounger split the food evenly. Once all the foraging steps are completed, the next generation is produced according to a Wright-Fisher process, with each individual’s fitness proportional to the total amount of food it acquired, multiplied by a factor $e^{-\alpha y}$ to account for the cost of additional cognition. In this analytical section, we will focus on parameter values $\rho,s << 1$ and $\alpha = 0$; we examine
larger values ($\rho = 0.25, s \sim 1, a = 0.5$) in the simulation section below.

1.1. Frequency of scroungers

In the following analysis, we will assume that mutations are sufficiently rare ($n\mu << 1$) that there is a separation of timescales. Each foraging type is usually fixed for a single cognitive level, and producers and scroungers are present in the population at the equilibrium frequency determined by the difference in cognitive levels. Occasionally, a mutation occurs that changes either an individual’s cognitive ability (by one level) or its foraging type. If the mutation initially confers a selective advantage $a$, it has a chance to escape drift and increase in frequency in the population. We assume that if it does so, it fixes and the frequency of scroungers relaxes to its new equilibrium value before the next mutation occurs. We now find the equilibrium frequency of scroungers (note that this is always complemented by the frequency of producers).

We assume that each scrounger independently chooses a producer with food from which it attempts to scrounge, but that if multiple scroungers choose the same producer, only one can actually make a scrounging attempt. (This is the same model as in the simulation section 2.1.2 below.) Since there are $n\varphi$ scroungers and $n\rho(1-\varphi)$ producers with food, the probability $f$ that a producer that finds food will face a scrouning attempt is

$$f = 1 - \left[1 - \frac{1}{n\rho(1-\varphi)}\right]^{n\varphi},$$

and the probability that, in a given foraging step, a scrounger will find a producer with food that is also available to be scrounged from is $f\rho(1-\varphi)/\varphi$.

Assuming that $n$ is large enough that many producers find food every foraging step
\( n \rho (1 - \varphi) \gg 1 \), \( f \) is approximately

\[
f = 1 - \exp \left( - \frac{\varphi}{\rho(1 - \varphi)} \right) \quad \text{(2)}
\]

The two foraging types then have relative fitnesses (up to an overall normalizing constant) \( w_p \) and \( w_s \) given by their expected payoffs, adjusted for the cost of cognition:

\[
w_p = \rho \left[ 1 - \frac{\sigma(d)}{2} f \right] \quad \text{(3)}
\]

\[
w_s = e^{-\varphi d} \frac{\sigma(d)}{2} f \frac{\rho(1 - \varphi)}{\varphi} \quad \text{(4)}
\]

The equilibrium frequency of scroungers, \( \varphi^* \), is the value of \( \varphi \) at which \( w_p = w_s \); scroungers go extinct if \( w_p(\varphi = 0) > w_s(\varphi = 0) \).

In order to find a simple approximate expression for \( \varphi^* \), note that for scroungers to be maintained in the population, it must not be too easy for producers to find food, i.e., \( \rho \) must be small. Assuming that \( \rho \ll \varphi \), we have \( f \approx 1 \), and therefore from (3) and (4)

\[
\varphi^* \approx \frac{\sigma e^{-\varphi d}}{2 - \sigma(1 - e^{-\varphi d})} \quad \text{(5)}
\]

Further assuming that the cognitive gap is not likely to grow to levels such that it imposes an enormous selective disadvantage, we have \( |\gamma d| \ll 1 \), and we can further approximate

\[
\varphi^* \approx \frac{\sigma}{2} \left[ 1 - \gamma d \left( 1 - \frac{\sigma}{2} \right) \right] \quad \text{(6)}
\]

Figure S1 shows the accuracy of this approximation.
1.2. Selective coefficients of mutations

As stated in the main text, the selective advantage of a mutation increasing cognitive abilities by $\delta$ is $\alpha_p(\delta)$ for producers and $\alpha_s(\delta)$ for scroungers, where $\alpha_p$ and $\alpha_s$ are given by:

\[
\alpha_p(\delta) = e^{-\gamma} \frac{1 - f \sigma(d - \delta)/2}{1 - f \sigma(d)/2} - 1 \quad (7)
\]

\[
\alpha_s(\delta) = e^{-\gamma} \frac{\sigma(d + \delta)}{\sigma(d)} - 1. \quad (8)
\]

We assume that scroungers are at a frequency $\hat{\varphi} >> \rho$, so that $f = 1$ in the above equations. Further assuming that each mutation has only a small effect on scrouning probability or cognitive cost ($|s\delta| << 1$ and $|\gamma\delta| << 1$), the right-hand sides of (7) and (8) are approximately given by the first-order Taylor expansions in $\delta$:

\[
\alpha_p(\delta) = \delta \left[ \frac{s\sigma(d)(1 - \sigma(d))}{2 - \sigma(d)} - \gamma \right] + O(\delta^2) \quad (9)
\]

\[
\alpha_s(\delta) = \delta \left[ s(1 - \sigma(d)) - \gamma \right] + O(\delta^2). \quad (10)
\]

(Recall that we assume that $a=0$.) The behaviors of $\alpha_p(1)$ and $\alpha_s(1)$ as functions of $d$ are shown in figure S2.

1.3. Speed of the cognitive arms race

In populations experiencing a stable cognitive arms race, the gap between the cognitive levels of the two foraging types settles down to a roughly steady value $\hat{d}$ at which both types increase in cognitive level at the same rate. In the rare-mutation regime we are
considering, this rate is given by the mutation supply times the probability that mutants
with a cognitive ability increased by one unit ($\delta = +1$) will fix. Assuming that $\alpha_p(1)$ and
$\alpha_s(1)$ are small compared to one, the probability of fixation is only $2\alpha$ (Ewens 2004).
The two foraging types therefore evolve higher cognition at the same rate when

\[ n(1 - \hat{\varphi})\mu(2\alpha_p(1)) = n\hat{\varphi} \mu(2\alpha_s(1)), \text{ i.e., when} \]

\[ (1 - \hat{\varphi})\alpha_p(1) = \hat{\varphi} \alpha_s(1). \quad (11) \]

Inserting the above expression (6) for $\hat{\varphi}$, and (9) and (10) for $\alpha_p$ and $\alpha_s$, and assuming
that food is hard to find ($\rho << 1$), this reduces to

\[ (1 - \sigma)\left(1 - \frac{sd\sigma}{2}\right) = \gamma d\sigma \left(1 - \frac{\sigma}{2}\right). \quad (12) \]

When the cost of cognition is low ($\gamma << s$), the expression simplifies further to $d \approx 2 / s\sigma$,
and we find that the rates of advance balance at a cognitive gap of

\[ \hat{d} \approx \frac{2}{s}, \quad (13) \]

corresponding to a probability of scrounging success $\hat{\sigma} \approx 0.9$. This is illustrated in Figure
S3, where the blue and purple curves intersect at $d \approx 2/s$. At this value of $d$, both
producers and scroungers accumulate cognitive mutations at a rate of approximately

\[ n\left(1 - \hat{\varphi}\right)\mu(2\alpha_p(1)) = n\hat{\varphi} \mu(2\alpha_s(1)) = \frac{n\mu s}{10} \left(1 - \frac{10\gamma}{s}\right). \quad (14) \]
Note that if the cost of cognition is too large \((\gamma > \frac{s}{10})\), a stable race is impossible; either cognition is too costly to ever evolve, or else it quickly becomes too costly for producers to keep up with scroungers, leading to the collapse of the race.

1.4. Mixed strategies and sexual reproduction

Above, we have focused on asexual populations of individuals following pure producing or scrounging strategies. We now consider the opposite limit, in which nearly all individuals follow a single mixed foraging strategy, scrounging with probability \(\varphi\) and producing with probability \(1 - \varphi\). Equivalently, we can consider a sexual population of individuals following pure strategies with frequent recombination between the foraging locus and the cognition locus or loci. In both cases, producer-scrounger interactions will typically occur between individuals with the same cognitive genotype, and cognitive mutations will, on average, be present in an individual acting as a scrounger with frequency \(\varphi\), and in an individual acting as a producer with frequency \(1 - \varphi\). We will continue to assume that the dynamics of the foraging strategy locus are fast, so that the scrouning frequency can be approximated by its equilibrium value given the current cognitive genotype of the population, \(
\varphi = \hat{\varphi}\) (for the mixed strategy case, this means that we assume that the population is at the evolutionarily stable foraging strategy).

In this case, the selective coefficients on cognitive mutations of size \(\delta\) are
\[ \alpha_p(\delta) = e^{-\gamma} \left[ \frac{\hat{\nu}_p + (1 - \hat{\nu}_p)\left[1 - f\sigma(d - \delta)\right]/2}{(1 - f\sigma(d))/2} \right] - 1 \]  

(14)

\[ \alpha_s(\delta) = e^{-\gamma} \left[ 1 - \hat{\nu}_s + \hat{\nu}_s - \frac{\sigma(d + \delta)}{\sigma(d)} \right] - 1. \]  

(15)

When food is hard to find \( (\rho < 1) \) and mutations are small \( (|s\delta|, |y\delta| < 1) \), these are both approximately

\[ \alpha(\delta) \approx \delta \left[ \frac{s\sigma(1 - \sigma)}{2} - \gamma \right]. \]  

(16)

Since mutations increasing the producing and scrounging cognitive levels both appear in the population at rate \( n\mu \), the rate of increase of cognition (assuming \( \delta = 1 \) for all mutations) is

\[ n\mu \alpha(1) = n\mu \left( \frac{s}{8} - \gamma \right) \]  

(17)

as long as cognitive levels remain roughly balanced \( (|s\delta| << 1, \text{ so that } \sigma \approx \frac{1}{2}) \).

2. Agent-based simulations

2.1. General properties

2.1.1. The population

We simulated a population of \( n = 100 \) haploid social foragers. Foragers’ PS behavior was determined by their genotype in the \( F \) gene, which was their probability to play the producer strategy. A pure producer carries the \( F \) genotype of 1, a pure scrounger carries 0; an agent with an \( F \) genotype of, for example, 0.7, played producer with probability 0.7 on any given foraging step, and scrounger with probability 0.3. In pure-strategy simulations, the only possible \( F \) alleles were 0 and 1, and in mixed-strategy simulations...
there were 11 possible alleles: 0, 0.1, 0.2 … 1. Alleles included in the simulation were assigned at equal probabilities in the population’s first generation.

Cognitive ability was determined by the $C$ gene in the General Cognition Model (GCM) and by the $Cp$ and $Cs$ genes in the Specialized Cognition Model (SCM). In both models, foragers’ cognitive level in the first generation was set to zero, *i.e.* in the GCM all foragers had the 0 allele in the $C$ gene, and in the SCM all foragers carried the 0 allele in both the $Cp$ and $Cs$ gene.

### 2.1.2. The producer-scrounger game

The lifetime of one generation included a series of 50 PS interactions, or steps, which were independent of each other and their order unimportant. This number of steps was chosen in order to allow foragers to interact with a large sample of the population. At the beginning of each step, all foragers draw a PS strategy according to their $F$ genotype, and those who play producer receive a set payoff (*e.g.* 4, although the value does not matter) at a probability of 0.25. This probability was set to introduce a cost to the producer strategy and to allow an effective PS game. Foragers who play scrounger are then assigned randomly and independently to producers who found food. We assume only one scrounger can join each successful producer, and therefore if two or more scroungers are assigned to the same producer, only one of them will be able to attempt scrouning. The difference $d$ between the (relevant) cognitive levels of the scrounger and producer involved was calculated as $d = [\text{scrounger’s } C] - [\text{producer’s } C]$ in the GCM and as $d = [\text{scrounger’s } Cs] - [\text{producer’s } Cp]$ in the SCM. In the case of successful scrouning, the
scrounger receives half of the producer’s food finding. It should be noted that when
setting scrounging success probability to 1, such populations evolve a PS ratio fluctuating
around 0.5:0.5. Scrounging success baseline probability, \( a \), used to calculate the
probability of successful scrounging \( \sigma(d) \), was set to 0.5 in all simulations, to allow for a
stable producer-scrounger game under simulation conditions on one hand, and for
cognition to play a significant role in determining the probability of successful
scrounging on the other. Increasing the value of \( a \) will result in lower collapse rates in the
SCM; however, this contributes little to how cognitive abilities affect scrounging success
and we therefore ignore it here.

2.1.3. Selection and reproduction

After completing 50 interactions, the foragers reproduce asexually, in proportion to their
relative lifetime accumulated payoff, and immediately die (population size remains
constant). Offspring are genetically identical to their parent, but for mutations, which
occur in all genes at a rate of \( \mu = n^{-10} \). Mutations in the \( F \) gene change it within the
simulation’s defined allele pool. In the cognition genes, a mutation changes the mutated
allele by one level, either increasing (+1) or decreasing it (-1). We allow the population to
evolve for 10,000 generations; under each parameter set, we repeated the simulation 100
times.
2.2. Factors affecting race speed and collapse rate

2.2.1. Cognitive mutation effect size (s)

We simulated populations under three possible values of $s$ (mutation effect size): 0.25, 0.5 and 1.5. In both GCM and SCM populations, a larger $s$ resulted in faster races; in SCM populations, it also resulted in a higher collapse rate (Figure S4). When $s$ is large, a single mutation increasing cognitive level of a scrounger (producer) when the cognitive difference between producers and scroungers is small or zero, entails a significant increase in the probability of successful scrouning (successful scrouning avoidance).

Such a mutant has a relatively large advantage over other individuals and the mutation is therefore likely to spread rapidly. This spread, in turn, provides a background on which a counter-mutation will have a large advantage, in the same manner. In GCM population with pure strategies, this will be a mutation increasing the producer’s or scrounger’s cognitive level $C$; in SCM populations, this will be a mutation increasing the cognitive level relevant to the foraging allele (i.e. $Cp$ for producers and $Cs$ for scroungers).

Stochastic events, resulting from the probabilistic nature of the payoff distribution, scrounger-to-producer assignment, selection, reproduction, and finite population size may lead one side to accumulate more mutations and increase its cognitive level before the other side responds, thus increasing the absolute value of $d$. When the cognitive advantage of either side is large enough (e.g. $d = -3$ when $s = 1.5$), a single cognitive mutation of $\pm 1$ has a negligible effect on the scrouning or scrouning avoidance success of its carrier, and its spread rate is very low. As we explain in the main text, if such a large cognitive difference is to the advantage of producers, scroungers are likely to go extinct before they can increase their cognitive level; if the difference is in favor of
scroungers, producers may begin decreasing their cognitive level because a lower
cognitive cost outweighs the slight reduction in scrounging-avoidance abilities. In GCM
populations a large cognitive difference is transient, as $F$ gene mutations transfer
cognitive abilities between strategies, and the race resumes at roughly the same point
where it stopped. In SCM populations the race collapses or escalates “backwards” under
these conditions, because mutations in $F$ uselessly transfer low $Cs$ levels from producers
(who had no use for high $Cs$ levels) to scroungers and low $Cp$ levels from scroungers to
producers. How large the cognitive difference $d$ must be to cause such effects depends on
$s$. When $s$, the effect size of cognitive mutations is large, relatively small values of $d$ will
be sufficient for the race to collapse, leading to higher collapse rates for large $s$, as does
the generally faster race associated with large $s$ values.

2.2.2. The effect of cognitive cost
We assume that a higher cognitive level incurs a cost, which may be developmental,
physiological or derive from the possibly longer processing times associated with high
cognitive levels. We use a cost proportional to the agent’s cognitive level, which is a
fractional deduction from the final accumulated payoff. For example, maintaining
cognitive level 10 resulted in deductions of 10% of payoff, regardless of PS strategy.
Unsurprisingly, setting the cost to 0 (Figure S4a) resulted in faster races (compare to
Figure S4b). In SCM populations, it also caused the cognition genes which were
irrelevant to the foraging strategy ($Cs$ for producers and $Cp$ for scroungers) to drift rather
than decrease, as there was no selection acting on them in either direction. Additionally,
fewer race collapses occurred in such populations, but this was the case only for lower $s$
values (see table S2).

2.3. Evolution in the $F$ gene

In simulations where the $F$ gene was free to evolve, the frequency of producers and
scroungers fluctuated; the $F$ gene inevitably co-evolves with the genes determining
cognitive level, but the interaction is complex due to the negative frequency-dependence
defining the PS game. To examine the effect of these fluctuations on the arms race, and to
explore the nature of arms races in SCM populations where scroungers cannot become
extinct, we ran a set of simulations with no fluctuations by holding the frequencies of
producers and scroungers constant, at 0.7 and 0.3, respectively. This ratio was based on
the frequencies observed in populations where the cognitive level was held at 0 with no
cognitive mutations, while the $F$ gene was allowed to evolve (producer frequency for the
last 100 of 10,000 generations was $0.697 \pm 0.009$ mean±SD; Population size $n = 100$, $s =
1.5$; 100 simulation repeats). To still allow transfer of cognitive abilities between
producers and scroungers (a key feature of the GCM) while keeping PS frequencies fixed,
we allowed $F$ gene mutations (at a rate of $\mu = n^{-10}$) which changed one producer into a
scrounger and one scrounger into a producer (retaining their cognitive levels).

Incidentally, the effective mutation rate was thus doubled.

The arms race in GCM populations was faster under constant PS frequencies (Figure
S4d), which can be expected given the higher mutation rate. This result does, however,
indicate that the fluctuations in PS frequencies that are typical of the PS game are not the
driving force behind the arms race, as might be hypothesized. Interestingly, the fact that scroungers could not go extinct did not promote consistent arms races in SCM populations. Instead of extinction, once a large gap formed in cognitive abilities between producers and scroungers, scroungers decreased their cognitive level and the race did not progress (Figure S4d).

References

Figure S1. The equilibrium frequency of scroungers, \( \phi \), as a function of their cognitive advantage over producers, \( d \). The blue curve shows the exact value obtained by solving Eqs. (3) and (4) for \( w_p = w_s \), while the purple curve shows the approximate value given by Eq. (6). The other parameters are \( a = 0 \), \( s = 1 \), \( \gamma = 0.05 \), \( \rho = 0.05 \). As long as the probability of scrounging success, \( \sigma(d) \), is larger than \( 2\rho \), and the cost of cognition is not too high (\( \gamma d \ll 1 \)), the frequency of scroungers is approximately \( \phi \sim \sigma / 2 \).
Figure S2. The selective advantage $\alpha$ for a mutation increasing the relevant cognitive ability by one unit in producers (blue) and scroungers (pink), as a function of the cognitive advantage of scroungers, $d$. The solid curves show the exact equations (7) and (8), while the dashed curves show the approximations (9) and (10). Negative values indicate that mutations reducing cognitive abilities are favored. The parameters are $a = 0$, $s = 0.1$, $\gamma = 0.005$, $r = 0.05$. 
Figure S3. The rate at which mutations increasing cognitive ability are fixed, divided by $n\mu$, the number of mutations in the population per generation, as a function of the cognitive advantage of scroungers, $d$. The rate for producers is shown in blue, and for scroungers in pink. When the scroungers' rate is higher than that of the producers, the population will tend to move to higher values of $d$. Thus, the population will tend to move from $d = 0$ to $d \approx 20 = 2/s$ where the rates match. At very large values of $d$, producers tend to lose cognitive ability faster than scroungers, leading to a backwards race. At sufficiently negative values of $d$, scroungers are unable to get food and go extinct, removing the selection for cognitive abilities in producers. The parameters are $a = 0$, $s = 0.1$, $\gamma = 0.005$, $r = 0.1$. 
Figure S4. Mean and standard error of cognitive level among foraging strategies in agent-based-computer simulations of the GCM and SCM, under various assumptions. Each mean is calculated for generations 9,901-10,000, for 100 repeats of each simulation. Columns marked with (*) are means calculated for less than 90 repeats, i.e. at least 10 repeats did not have the marked genotype in at least one of the 100 generations considered (see table 2 for detailed account of valid data points). The three column groups in each subfigure corresponds to different values of \( s \), slope coefficient of the scrounging success probability function. All simulations are for population size \( n = 100 \), \( T = 50 \) time steps, \( G = 10,000 \) generations, mutation rate \( \mu = 0.01 \). (a) Pure social foraging strategies; cognitive level cost \( \gamma = 0 \). (b) Pure social foraging strategies; \( \gamma = C/100 \). (c) Mixed social foraging strategies (producing probability \( \leq 0.5 \) alleles are grouped under “scrounger”, producing probability \( > 0.5 \) alleles are grouped under “producer”); \( \gamma = C/100 \). (d) Pure social foraging strategies at fixed frequency of 0.3:0.7 scroungers to producers (i.e. no evolution in \( F \) gene); \( \gamma = C/100 \).