PHILOSOPHICAL
TRANSACTIONS B

Phil. Trans. R. Soc. B.
doi:10.1098/not yet assigned

Paradox of Diversity in the Collective Brain

Supplementary Material

Robin Schimmelpfennig, Layla Razek, Eric Schnell, Michael Muthukrishna

We present a simple model of the paradox of diversity. We focus on diversity within a population, such as the diversity created by skill specialisation and the division of information and labour. We define diversity as variation in cultural traits across a population. These could be practices such as food-processing techniques, access to technologies such as the Internet, or technical skills, but also broader traits such as language, family structure, and occupation. Previous models have examined the evolution of the division of labour, showing how specialisation can lead to more successful populations, as measured by, for example, economic value or total yield of resources (1,2). We build on this prior work to theorize how a population should divide information to optimize cultural complexity and coordination between specialities.

Specialisation and the division of information and labour allow for cultural complexity to exceed the abilities of a single brain. As a verbal summary of the formalization below, consider a population of individuals with a fixed brain of capacity B, maximized due to constraints on brain size, such as the size of the birth canal (3). For ease of explanation, let $B = 10$. In order to survive, these individuals need to have some minimum skill across a range of domains M. For example, the ability to find and process food, build clothing and shelter, evade predators, heal the sick, and know the norms and laws of the group. Again, for ease of explanation, let $M = 10$. If everyone in a society has to learn all 10 domains, each reaches skill level $B/M = 10/10 = 1$. If, however, everyone had to learn only half the domains (5), then each person—and the society as a whole—could reach $B/5 = 2$. If they only had to learn 1 domain, then they would reach skill unit 10, half a domain then skill level 20, and so on.

The degree to which one can safely specialize is the degree to which one can rely on others to reliably provide the outputs of the knowledge in the remaining domains. This is a function of sociality, interdependence, and cooperation (4). That is, there need to be enough specialists in each domain to ensure that the information is not lost, and the products of labour are sufficient and reliable. For example, in a small town, there may be a single general physician who needs to know many, if not all domains of common medicine. But in New York, a doctor may specialise in a small part of the renal system and get very good at treating that one part. Other specialists will cover other domains. However, ongoing specialisation can create a new challenge. Individuals become smarter at a few domains and stupider at everything else. Ask an average adolescent in a WEIRD society how to grow wheat and you might starve. To summarize, at an individual level, specialisation leads to a higher skill level within a domain, but also a siloing of skills. At a population level, specialisation leads to a higher average skill level across all domains, but creates a coordination challenge that must be overcome for everyone to survive.

We formalize this logic in a model. Our model is predicated on the Cultural Brain Hypothesis (5), modelling division of information as a strategy for coping with an ever growing cumulative cultural corpus. The Cultural Brain Hypothesis predicts the coevolution of brain size, group size, social learning, and life history. Two broad strategies for dealing with the growing information are:

1. **Grow larger brains:** here we assume that larger brains help humans process and store more information (6–8). But there is a fundamental limit to brain size. Larger

and more complex brains are more costly than less complex brains because they require more calories (9), take longer to develop (6,10,11), and are harder to birth (12,13).

2. **Increase transmission fidelity:** Increasing transmission fidelity will help humans to learn faster and more efficiently. See Section 2.2 in the main text for the variety of genetic and cultural innovations that support increased transmission fidelity, ranging from better social learning to longer learning periods to media communication technologies.

There are limits to both bigger brains and transmission fidelity. For example, sufficient calories and safe, secure childhoods are common in much of the WEIRD world (though substantial inequality exists between and within countries), but as Lipschuetz et al. (12) show, bigger brains still predict emergency birth interventions such as Caesareans and forceps. Caesareans remove this selection pressure, but with other health costs, such as those created by the lack of a microbiome transfer (14,15); although new approaches, such as "vaginal seeding" may help mitigate these costs (16). Similarly, extending the juvenile period into a cultural adolescence for longer learning runs into trade-offs on lifetime earnings and reproduction (delayed birth of first child and difficulties reproducing at an older age, particularly among females). In contrast, a division of information and labour strategy has no limits to increasing cultural complexity, as long as the sociality and coordination challenges are met.

1.1 Model 1: Specialisation

A population can increase its average skill level by specialising and dividing information and labour. To formalize this logic, we assume the following:

- There are N individuals in the population.
- There are M domains to be learned. The aggregation of all individuals in the society must cover the full M range to ensure survival. Excellence in tool manufacturing is of little help if you have lost the ability to make food.

- For each domain, learners must learn the same basic level of knowledge (*K*) before they can specialize further. This means that the first *K* amount an individual spends learning a domain overlaps with all other learners of that domain, as they must learn the basics first, and thus this *K* amount only contributes to the societies level of knowledge in the domain once. Above this amount *K*, learners can specialize in sub-domains, and so two learners can each contribute to increase the societal knowledge of the domain. This is equivalent to all engineering majors requiring some minimum mathematical training regardless of speciality.
    - For example, if $K = 0.4$ and two people invest 1 point worth of knowledge into the same domain, then the societal knowledge of that domain would be 1.6. This is because both people learn the same basic 0.4 worth of information, plus they both learn a unique 0.6 of information in that domain, resulting in $0.4 + 0.6 + 0.6 = 1.6$
- Learners have a fixed capacity *B* to learn new things, a function of brain size and time, which we assume is equal across the population and set at B=1.

Variable	Explanation	Type and Range
N	Population size	Integer, $N > 0$
M	Number of domains	Integer, $M > 0$
B	Capacity to learn domains	1
K	Basic knowledge threshold	$0 < K < 1$

Table 1: Summary of variables

In our model, the societal knowledge of a domain is the sum of knowledge spent above the threshold *K* by each learner of the domain, plus this threshold *K*. If no learner surpasses the threshold *K*, then it is the maximum of the knowledge spent by each learner of the domain. We make some simplifying assumptions in our model.

First, since we are not modelling comparative advantage between societies or cultural clusters, the total societal knowledge is the minimum of the societal knowledge of all the domains. This means a society's knowledge is only as great as it's least known domain.

Second, since we are modelling skill specialization at a societal level and not the individual level, we treat domains of knowledge as being discrete. In reality this is likely not the case, as the boundary between domains may be fuzzy and some skills may be interchangeable between domains.

Third, we assume an implicit optimization for comparative advantage. No domain is worth more than any other, but we try to equalize the number of individuals within any one domain. To achieve this, each individual will be assigned to $\lfloor \frac{2M}{N} \rfloor$ domains[1] and the remainder of this division, or M (mod N), will be assigned to an additional domain. It is then possible to ensure each individual spreads their efforts across all the domains they know such that the societal knowledge of each domain, and in turn the total knowledge level of the society, is:

$$\text{Societal Knowledge} = \frac{N - 2*M*K}{M} + K \quad (1)$$

1.1.1 Predictions

The results from the model illustrate a simple logic: ceteris paribus, with an increase in population size, specialisation can increase, or equivalently the number of domains learnt per person can decrease (see *Figure 1a*). Individuals can focus on learning fewer skills very well, as there are enough other people to learn the remaining skills. This is consistent with evidence that the benefits of diversity are more likely to emerge in larger groups (17).

Similarly, if the number of domains of knowledge increases (see *Figure 1b*), the population needs to learn more domains. However, there is an obvious trade-off: as the population and specialisation levels grow, coordination costs will emerge (18,19). In the next section, we model these coordination costs.

---

[1] If $\left\lfloor \frac{2M}{N} \right\rfloor = 0$, then we instead assign C individuals per domain, where C is the smallest number such that $\left\lfloor \frac{C*M}{N} \right\rfloor = 1$. This changes equation (1) to $\frac{N-C*M*K}{M} + K$.

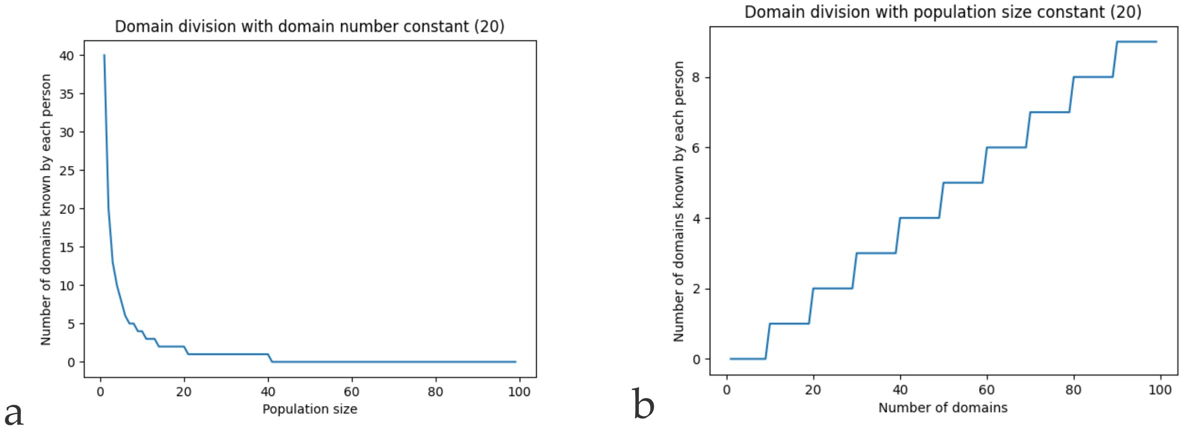

Figure 1. Domain division. For both figures we show the number of domains learnt per person, as we vary either population size (a) or total number of domains (b). We show the minimum number of domains each person must specialise in, recalling that some people will also have to specialise in one extra domain. In *Figure 1a*, given a fixed number of domains ($M = 20$), when the population size increases the number of domains each person specialises in decreases. In *Figure 1b*, for a fixed population size ($N = 20$), when the number of domains increases the number of each person specialises in also increases.

1.2 Model 2: Coordination

To model this coordination problem, we introduce network structure. We construct a network such that difficulty in coordinating is proportional to network path distance. If Person A and Person B share a skill set, then they should be able to coordinate with no issues and so have a direct link in our network. If Person A and Person B do not share a skill set, they have to go through mutual connections based on overlapping skill sets.

To capture the coordination cost for the population as a whole, we can measure the efficiency of the network using Average Dyadic Efficiency ($ADE = \frac{1}{N(N-1)}\sum_{i\neq j}^{N}\frac{1}{d_{i,j}}$) (20), where $d_{i,j}$ is path distance. A larger ADE represents a more efficient network.

For our purposes we will be randomly creating networks with different sets of parameters. We do so by randomly assigning A individuals to each domain, where A is an integer greater than 2. By assigning at least 2 individuals per domain, we can ensure that our network can be connected. We do this assignment while ensuring that every individual is assigned to at least $\lfloor\frac{A*M}{N}\rfloor$ domains, AxM (mod N) individuals learns one additional domain, and no individual learns the same domain twice. This algorithm can be made more efficient by ensuring that individuals who already connected in one domain will not share any additional domains, but this algorithm is computationally expensive, hard to scale up to large values of A, and not necessarily realistic to how real-world skill specialization occurs, so we ignore this possibility. We try different values for A to test which ones best improve network efficiency without harming the performance of the society. Societal knowledge is measured as:

$$\text{Societal Knowledge} \approx \frac{N - A*M*K}{M} + K \qquad (2)$$ [2]

In the case where $A = 2$ equation (2) is the same as equation (1) above as they both represent having 2 individuals learning each domain.

[2] Equation (2) may not be exact as we do not ensure the same constraints as for equation (1), such as separating individuals learning the same number of domains as much as possible. However, it provides a close enough estimate to demonstrate the relationship between societal knowledge and efficiency.

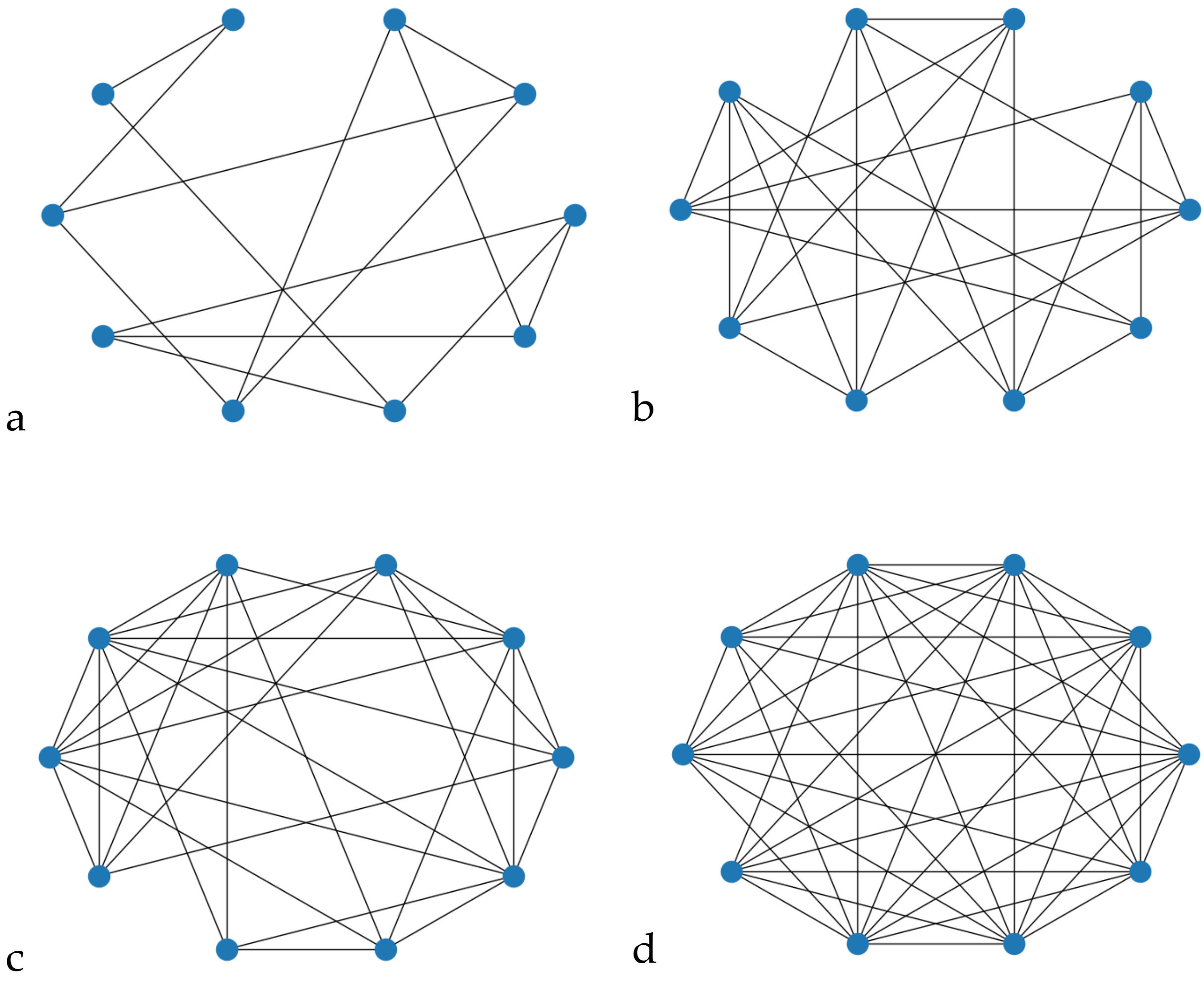

Figure 2. Network Efficiency. We draw 4 networks randomly generated following the instructions explained above using different sets of parameters. All networks have 10 individuals ($N = 10$). The remaining network parameters are: (a) $M = 15$, $A = 2$; (b) $M = 25$, $A = 2$; (c) $M = 15$, $A = 3$; (d) $M = 25$, $A = 3$. These networks have the following Average Dyadic Efficiency (ADE): (a) ADE = 0.533; (b) ADE = 0.719; (c) ADE = 0.867; (d) ADE = 0.922. The difference between (a) vs (b) and (c) vs (d) is the introduction of more domains of knowledge to be learnt ($M = 15$ to $M = 25$). When there are more domains of knowledge then everyone is responsible for learning more domains (see *Figure* 1). This in turn creates more connections between people and also increases the network efficiency. The difference between (a) vs (c) and (b) vs (d) is the requirement for more people to learn each domain ($A = 2$ to $A = 3$). This creates more links as individuals are connected to two others in every domain, rather than only one other. It also requires each person to learn more domains, which further increases the number of connections in the network. Both these reactions increase network efficiency.

1.2.1 Predictions

We simulate how different population sizes, number of domains, and values for *A* affect the network efficiency for 10,000 randomly generated networks constructed with the constraints described above[3]. The following predictions hold:

1. When population size increases for a constant number of domains, then the efficiency of our network decreases (see *Figure 3a*).
2. When the number of domains increases for a constant population size, then the efficiency of the network increases (see *Figure 3b*).
3. When the number of people learning each domain (*A*) increases, then network efficiency increases and societal knowledge decreases.

Overall, this creates a trade-off between the knowledge level of a population and the coordination cost within the population. As population size increases the knowledge level increases, but network efficiency decreases (or coordination costs increase). To keep this coordination cost in check, the number of domains (*M*) should increase, but this will lower the knowledge level of our population. There may be many strategies to optimize this trade-off in the diversity of domains, such as those verbally described in the main text. Our formalization offers a framework for making more specific predictions about these in future work.

[3] See code in Supplementary. Also available at: https://github.com/schnelleric/Knowledge-Diversity-Model

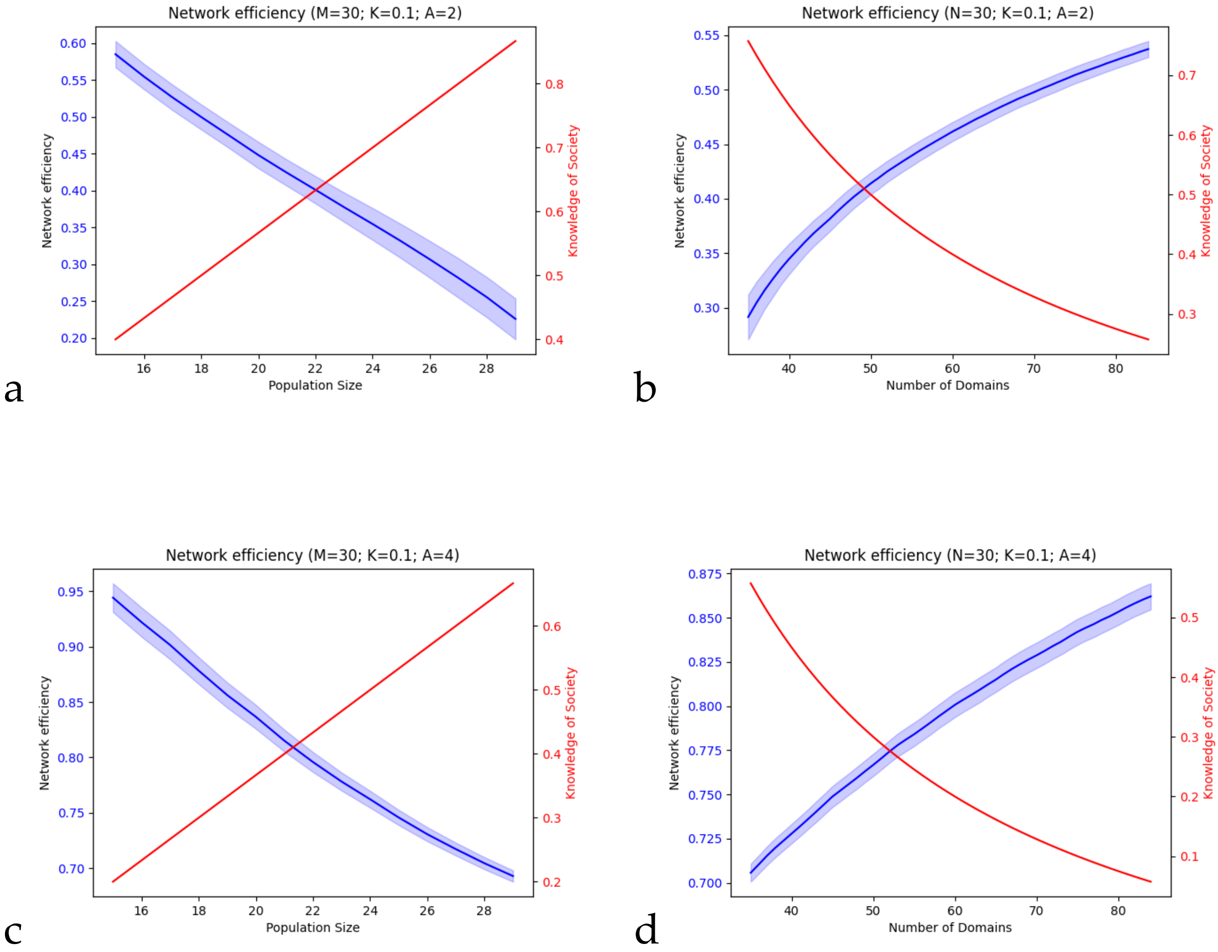

Figure 3. Network Efficiency. For all figures we show the network efficiency, as calculated by the Average Dyadic Efficiency (ADE) alongside societal knowledge as calculated in equation (2). We vary either population size (a and c) or total number of domains (b and d). For all figures $K = 0.1$ and for Figures 3a and 3b $A = 2$, whereas for figures 3c and 3d $A = 4$. The distribution of networks is randomly assigned per the rules listed above and the value in the graphs are the average ADE over 10,000 runs with error bars representing the standard deviation. In Figure 1a and 1c, given a fixed number of domains (M = 30), when population size increases the network efficiency decreases and societal knowledge increases. In Figure 1b and 1d, for a fixed population size (N = 30), when the number of domains increases the network efficiency also increases and societal decreases. When A is larger, Figures 3c and 3d as compared to 3a and 3b, Network efficiency is higher but societal knowledge is lower.

References

1. Henrich J, Boyd R. Division of Labor, Economic Specialization, and the Evolution of Social Stratification. Curr Anthropol. August 2008;49(4):715–24.

2. Nakahashi W, Feldman MW. Evolution of division of labor: Emergence of different activities among group members. J Theor Biol. Mai 2014;348:65–79.

3. Muthukrishna M, Doebeli M, Chudek M, Henrich J. The Cultural Brain Hypothesis: How culture drives brain expansion, sociality, and life history. Tarnita CE, Herausgeber. PLOS Comput Biol. 8. November 2018;14(11):e1006504.

4. Henrich J, Muthukrishna M. The Origins and Psychology of Human Cooperation. Annu Rev Psychol. 2021;72(1):207–40.

5. Muthukrishna M, Doebeli M, Chudek M, Henrich J. The Cultural Brain Hypothesis: How culture drives brain expansion, sociality, and life history. PLOS Comput Biol. 11. August 2018;14(11):e1006504.

6. Herculano-Houzel S, Collins CE, Wong P, Kaas JH. Cellular scaling rules for primate brains. Proc Natl Acad Sci. 27. Februar 2007;104(9):3562–7.

7. MacLean EL, Hare B, Nunn CL, Addessi E, Amici F, Anderson RC, u. a. The evolution of self-control. Proc Natl Acad Sci. 20. Mai 2014;111(20):E2140–8.

8. Sol D, Bacher S, Reader SM, Lefebvre L. Brain Size Predicts the Success of Mammal Species Introduced into Novel Environments. Am Nat. 1. Juli 2008;172(S1):S63–71.

9. Isler K, van Schaik CP. Metabolic costs of brain size evolution. Biol Lett. Dezember 2006;2(4):557–60.

10. Charvet CJ, Finlay BL. Chapter 4 - Embracing covariation in brain evolution: Large brains, extended development, and flexible primate social systems. In: Hofman MA, Falk D, Herausgeber. Progress in Brain Research [Internet]. Elsevier; 2012 [zitiert 27. März 2021]. S. 71–87. (Evolution of the Primate Brain; Bd. 195). Verfügbar unter: https://www.sciencedirect.com/science/article/pii/B9780444538604000040

11. Herculano-Houzel S, Mota B, Lent R. Cellular scaling rules for rodent brains. Proc Natl Acad Sci. 8. August 2006;103(32):12138–43.

12. Lipschuetz M, Cohen SM, Ein-Mor E, Sapir H, Hochner-Celnikier D, Porat S, u. a. A large head circumference is more strongly associated with unplanned cesarean or instrumental delivery and neonatal complications than high birthweight. Am J Obstet Gynecol. 1. Dezember 2015;213(6):833.e1-833.e12.

13. Fischer B, Mitteroecker P. Covariation between human pelvis shape, stature, and head size alleviates the obstetric dilemma. Proc Natl Acad Sci. 5. Mai 2015;112(18):5655–60.

14. Sandall J, Tribe RM, Avery L, Mola G, Visser GH, Homer CS, u. a. Short-term and long-term effects of caesarean section on the health of women and children. The Lancet. Oktober 2018;392(10155):1349–57.

15. Shao Y, Forster SC, Tsaliki E, Vervier K, Strang A, Simpson N, u. a. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. Nature. Oktober 2019;574(7776):117–21.

16. Dominguez-Bello MG, De Jesus-Laboy KM, Shen N, Cox LM, Amir A, Gonzalez A, u. a. Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer. Nat Med. 2016;22(3):250–3.

17. Pescetelli N, Rutherford A, Rahwan I. Diversity promotes collective intelligence in large groups but harms small ones [Internet]. PsyArXiv; 2020 Jan [zitiert 9. Januar 2020]. Verfügbar unter: https://osf.io/b8q2c

18. Putnam RD. E pluribus unum: Diversity and community in the twenty-first century the 2006 johan skytte prize lecture. Scand Polit Stud. 2007;30(2):137–74.

19. Weber RA, Camerer CF. Cultural Conflict and Merger Failure: An Experimental Approach. Manag Sci. April 2003;49(4):400–15.

20. Pasquaretta C, Levé M, Claidière N, van de Waal E, Whiten A, MacIntosh AJJ, u. a. Social networks in primates: smart and tolerant species have more efficient networks. Sci Rep. Mai 2015;4(1):7600.