

Document Version

Final published version

Licence

Dutch Copyright Act (Article 25fa)

Citation (APA)

Buijsrogge, A., Brethouwer, J. T., van de Rijt, A., & Lindelauf, R. (2026). Self-organized social distancing. *Journal of Mathematical Sociology*, 50(2), 113-149. <https://doi.org/10.1080/0022250X.2025.2513240>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.






Self-organized social distancing

Anne Buijsrogge, Jan-Tino Brethouwer, Arnout van de Rijt & Roy Lindelauf


To cite this article: Anne Buijsrogge, Jan-Tino Brethouwer, Arnout van de Rijt & Roy Lindelauf (09 Feb 2026): Self-organized social distancing, The Journal of Mathematical Sociology, DOI: [10.1080/0022250X.2025.2513240](https://doi.org/10.1080/0022250X.2025.2513240)



To link to this article: <https://doi.org/10.1080/0022250X.2025.2513240>

 View supplementary material 

 Published online: 09 Feb 2026.

 Submit your article to this journal 

 Article views: 34

 View related articles 

 View Crossmark data 



Self-organized social distancing

Anne Buijsrogge^a, Jan-Tino Brethouwer^a, Arnout van de Rijt^b, and Roy Lindelauf^{c,d}

^aDelft Institute of Applied Mathematics, TU Delft, Delft, The Netherlands; ^bDepartment of Political and Social Sciences, European University Institute, Fiesole, Italy; ^cFaculty of Military Sciences, Data Science in Military Operations, Netherlands Defence Academy, Breda, The Netherlands; ^dTilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, The Netherlands

ABSTRACT

During previous pandemics, social distancing was organized top-down, through the imposition of a minimum distance. An alternative approach toward social distancing asks individuals to try to maximize their distance to others. Here, we ask whether people can thus efficiently self-organize spatial arrangements. We studied 953 social distancing decisions made in 150 groups under controlled conditions. Results show that subject behavior approximates what optimal mathematical strategies achieve. At scale, the observed behavior produces greater distancing than the mere respecting of an externally imposed minimum distance. These findings suggest that the encouragement of maximal distancing may reduce the propagation of viruses that spread through close-range contact beyond what is achieved with minimum-distance policies alone.

ARTICLE HISTORY

Received 22 May 2023
Revised 26 February 2025
Accepted 24 April 2025

KEYWORDS


Experiment; health behavior; mathematical analysis; physical distancing; social distancing

1. Introduction

A central question in sociological scholarship is whether individuals naturally self-organize to generate optimal collective behavior or whether top-down intervention is needed to avoid organizational failure or inefficiency at the group level (Diekmann et al., 2014; Granovetter, 1985; Heckathorn, 1996; Horne & Mollborn, 2020; Macy, 1991; Schelling, 1971; Simpson & Willer, 2015). The present article identifies a new problem of self-organization. This problem became salient during the recent COVID-19 pandemic and comes up in pandemics more generally and other settings in which individuals must distance themselves from others within a confined space.

Social distancing, or physical distancing, refers to non-medical interventions that seek to reduce the propagation of a contagious disease that spreads through droplet contact or airborne transmission by limiting short-distance interaction (Courtemanche et al., 2020; Epton et al., 2022; McGrail et al., 2020). During the COVID-19 pandemic, a common social distancing measure employed by many countries as a national policy was the imposition of

CONTACT Roy Lindelauf  rha.lindelauf.01@mindef.nl  Faculty of Military Sciences, Data Science in Military Operations, Netherlands Defence Academy, Breda, The Netherlands

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/0022250X.2025.2513240>

© 2026 Netherlands Defence Academy

a minimum physical distance between individuals in public places considered safe. The exact distance prescribed by governments differed, from 2 m in Canada to 6 ft. in the U.S. to 1.5 m in Germany, 1.4 m in South Korea and 1 m in China (Fazio et al., 2021; Setti et al., 2020; Sun & Zhai, 2020). The implementation varied from footstep markings to standing signs to seats covered in red-white tape in alternating fashion, and exceptionally a theater adopting a mathematically backed policy maximizing the number of guests while maintaining social distancing (Blom et al., 2021).

Apart from enforcing what is unavoidably a somewhat arbitrary threshold distance, public policies did little to encourage keeping more than the prescribed distance where that was possible. As the risk of interpersonal transmission monotonically declines with distance, realizing greater spatial intervals reduces spread and eventually deaths (Chu et al., 2020). For example, if attendance at an event is modest, everyone sitting in the front rows separated by taped empty chairs is suboptimal. Fewer people would get sick if they were instead spread more evenly across the theater. The present study probes what can be expected to be gained from individuals being encouraged to maximize social distance beyond a legal minimum.

The specific aggregation problem we study is the following. People may not succeed at positioning themselves optimally, because they do not know how many people come after them and where they will choose to position themselves. Imagine being the second person to choose a position at a roundtable. One strategy would be to sit at 180 degrees from the first person. This would be optimal if no one else joins. But if one more person arrives, or three, or five, sitting at 180 degrees makes an optimal arrangement impossible. Will people facing this uncertainty be able to self-organize a reasonably efficient seating arrangement?

This problem shares some similarities with self-organization problems sociologists have previously studied but is distinct. A first major category is sociological work on *collective action*, which studies under what conditions individuals in the absence of formal intervention may overcome collective action problems through norms and social embeddedness (Diekmann et al., 2014; Granovetter, 1985; Heckathorn, 1996; Jiao et al., 2021; Macy, 1991; Oliver et al., 1985; Raub & Weesie, 1990). Collective action problems are characterized by a misalignment between individual incentives and collective incentives, leading individuals to deviate from what is good for the collective. In the social distancing problem we study here, by contrast, individual and group interests are aligned. Individual and group interests are also aligned in a second strand of research studies on *coordination problems*. This literature studies how self-organization through emergent behavioral conventions can overcome coordination problems in the absence of traffic rules or turn-taking regulations (Buskens et al., 2008; Centola & Baronchelli, 2015; Centola et al., 2018; Diekmann, 1985; Ellison, 1993; Guilbeault et al., 2021; Young, 1993).

What differentiates the social distancing problem is that the key issue in coordination problems is the uncertainty about what others will do (and about what others think others will do), while in the social distancing problem, the most important form of uncertainty is arguably about group size. However, uncertainty about what others will do still plays a role in the social distancing problem, as this can be expected to be informed by these others' expectations of group size. A third form of self-organization is seen in the context of collective intelligence, where the social challenge is to find an optimal solution to a complex problem. This field studies the conditions under which individuals manage to efficiently aggregate their insights to find that optimum (Becker et al., 2017; Frey & Van de Rijt, 2021; Friedkin & Bullo, 2017; Hong & Page, 2004; Keuschnigg & Ganser, 2017; Lorenz et al., 2011; Mason & Watts, 2012). A key feature of the social distancing problem is that it is not individuals' insights but rather the distributed knowledge of their physical presence that is aggregated.

We make both theoretical and empirical progress on the social distancing problem. Theoretically, we identify a number of individual seating strategies for which we prove upper bounds on performance. We also identify for each seating strategy upper bounds on average performance under the assumption of a uniform probability distribution of group size bounded at some n_{max} . Strikingly, we find that the common sense strategy of maximizing distance to those already present, ignoring who might come later, while not optimal, performs very well. We conjecture that this strategy will approximate empirical behavior. We conducted an empirical study in which we asked human subjects to sit down with the aim to maximize their distance to other group members. Analysis of the data confirms that humans by and large follow the common sense strategy and by doing so achieve a performance level that is much closer to optimal than to random behavior. In a large enough space, the common-sense strategy would lead a small enough number of people to achieve a minimal distance above a prescribed distance, e.g. 1.5 meters. This result suggests that in the absence of counteracting incentives, people may self-organize quite efficiently and form spatial arrangements that might reduce disease spread beyond what a legally enforced minimum achieves.

In practical settings, it may often already be difficult to motivate people to keep a minimal distance (see e.g. Epton et al., 2022; Lunn et al., 2020), let alone to distance beyond that minimum. For such settings where self-organized maximal distancing is impractical, the theoretical strategies we study here can be used to strategically mark locations with numbers corresponding to the order in which people enter a space. The strategies we study in this paper can be purposed to identify locations that are independent of the number of people ultimately entering the space, so that numbers need not be reassigned in between events. For very large settings, such as a stadium, one could alternatively enforce a top-down

policy in which people are divided into smaller subsections and let people self-organize within these smaller subsections. These strategic assignments can in principle reduce disease spread beyond what minimum distance enforcement can accomplish.

2. Theoretical seating strategies

The seating problem that our theoretical analysis addresses is as follows. Consider a roundtable with radius r and let N be the random number of persons that must be seated around the table. If $N = n$, it is clear that in the optimal seating the distance between all people must be $\frac{2\pi r}{n}$. However, if N is random and people cannot practically move once seated, can we find a strategy with a performance close to optimal for each realization of N ?

We start by defining our performance metrics: the maximum error, which we upper bound independently of the number N , and the average maximum error, which we also upper bound independently of N . To do so, we assume that N is uniformly distributed over a bounded range, with an upper limit n_{max} , as described in Subsection 2.1. Next, we describe all strategies and prove upper bounds on the performance metrics (Subsection 2.2). For readability, all proofs have been deferred to Appendix A.1.

2.1. Performance metrics

Our first performance metric is the maximum error $e_{max}(N)$: how “bad” do people get seated compared to optimal seating (a uniform seating over the circle). With a slight abuse of notation, we denote the minimum of all pairwise distances among n individuals – acknowledging that it also depends on their positions – as $d_{min}(n)$, and the minimum distance in the optimal seating as $d_{opt}(n)$. Accordingly, we define the maximum error as follows:

Definition 2.1. For any given number n , we define the *maximum error* as

$$e_{max}(n) = \frac{d_{opt}(n) - d_{min}(n)}{d_{opt}(n)} = 1 - \frac{d_{min}(n)}{d_{opt}(n)}.$$

This pair-wise distance minimum can be defined in two ways: (1) we can compute the Euclidean (direct) pair-wise distances $d_{Euc}(i, j)$ between all pairs i, j and then compute the minimum distance $d_{Euc}(n) = \min_{i, j} d_{Euc}(i, j)$, or, (2) we can compute the pair-wise distance along the circle $d_{circ}(i, j)$ between all pairs i, j and then compute the minimum distance $d_{circ}(n) = \min_{i, j} d_{circ}(i, j)$. Although our analysis focuses on arc lengths, it is important to note that optimizing arc lengths inherently optimizes Euclidean distances, which represent the path a virus would take.

Intuitively, it is clear that $d_{Euc}(n) \leq d_{circ}(n)$ and this can be easily shown. Therefore, any upper bound on the error using the distance along the circle is also an upper bound on the error using the Euclidean distance. In particular, the error along the circle is easier to compute. Thus, in the remainder, we will consider the minimum distance along the circle and refer to this as $d_{min}(n)$. Moreover, it is easy to show that the error is independent of the radius used and so we can assume w.l.o.g. a radius of $\frac{1}{2\pi}$ (implying a circumference of 1). Hence, we use

$$e_{max}(n) = 1 - nd_{min}(n). \tag{1}$$

We can now easily define our second performance metric: the average maximum error. Remark that the minimum distance, when $n = 1$, is by definition equal to 1.

Definition 2.2. Let $\mathbb{P}(N = n) = p(n)$ for $n \geq 1$ be the probability mass function of N . Then, we define the average maximum error as

$$\mathbb{E}[e_{max}(N)] = \sum_{n=2}^{\infty} p(n)(1 - nd_{min}(n)). \tag{2}$$

2.2. Strategies

In this section, we analyze several different types of seating strategies. In Section 2.2.1, we describe the split- k strategy, in which the largest line segment is always split into k smaller segments of equal size. In Section 2.2.2, we discuss the fixed error strategy, where the position of person j is such that the maximum error, after the arrival of person j , never exceeds some fixed value e_{fix} . Section 2.2.3 considers the fixed angle strategy, in which the position of person $j + 1$ has a fixed angle relative to the position of person j . Then, in Section 2.2.4, we conclude with a random strategy, where all persons are placed uniformly at random along the circle.

We summarize the worst-case performances of each of the strategies in Table 1.

Table 1. Performance measures for several strategies. For the expected maximum error, we assume that N has a uniform distribution in $[2, n_{max}]$ for some fixed value $n_{max} < \infty$ and φ denotes the golden ratio. Note that the bounds are independent of n_{max} and N .

Strategy	Upper bound on		
	$\mathbb{E}[e_{max}(N)]$		$e_{max}(N)$
Split-2	$1 - \frac{1}{\sqrt{2}}$	≈ 0.293	$\frac{1}{2}$
Fixed error	$1 - \frac{1}{2\log(2)}$	≈ 0.279	$1 - \frac{1}{2\log(2)}$ ≈ 0.279
Split-3	$1 - \frac{1}{\sqrt{3}}$	≈ 0.423	$\frac{2}{3}$
Fixed angle	$1 - \frac{\sqrt{\varphi}}{\sqrt{5}}$	≈ 0.431	$1 - \frac{1}{\sqrt{5}}$ ≈ 0.553
Random	1		1

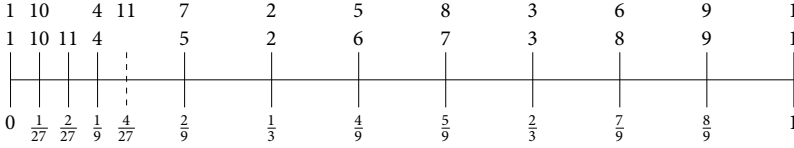


Figure 1. Two examples of a split-3 strategy for $N = 11$. Below the line are the distances from 0, above the line are two realizations of the person numbers that both give the same minimum distance.

2.2.1. Split- k

In this strategy, we initially split all line segments on the roundtable in k segments, that is, we seat all persons according to the reciprocal of k . Then, after the placement of persons 1 till k , we split all the k segments again in k , so that when seating the next $k(k - 1)$ persons the minimum distance is always $\frac{1}{k^2}$. This process continues until all arriving persons are seated, an example of split-3 can be found in Figure 1, where we represent the circle as a line with person 1 on both ends of the line. This process does not give a unique location to each person, that is, we could also seat persons according to the second possibility in Figure 1. Note that both realizations of split-3 result in the same minimum distance for a given person and hence do give the same error. This strategy is related to so-called van der Corput sequences (see e.g. van der Corput, 1935).

We present the (expected) maximum error in case of the split- k strategy (for proofs see Appendix A.1).

Lemma 2.3 (Maximum error using the split- k strategy). *Consider the split- k strategy and let $k > 1$. For any realization of N , $e_{\max}(N) < 1 - \frac{1}{k}$.*

Theorem 2.4 (Expected maximum error using the split- k strategy). *Consider the split- k strategy and let $k > 1$. Furthermore, let N be uniformly distributed in $[2, n_{\max}]$ such that $k^l < n_{\max} \leq k^{l+1}$ for some $l \geq 1$. Then,*

$$\mathbb{E}[e_{\max}(N)] \leq 1 - \frac{1}{\sqrt{k}}.$$

In addition, when $n_{\max} = k^l$ for some $l \geq 1$, we also have

$$\mathbb{E}[e_{\max}(N)] < \frac{k - 1}{2k}.$$

Clearly, split-2 minimizes this error among all split- k strategies. Note that for all $n_{\max} < k$ Theorem 2.4 is not valid, and it is easy to verify that, using Equation (A1), in this case $\mathbb{E}[e_{\max}(N)] \leq 1 - \frac{2}{k}$ (i.e. the performance is as bad as the performance of person number 2 when $k \geq 4$). Thus, the split- k

strategy that is presented here is clearly not optimal for $k \geq 4$. In Appendix A.1, we present some improvements to this strategy. These improvements suggest that split-3 achieves the worst performance among all improved split- k strategies.

Since our experiment considers $n_{\max} \leq 11$, we also highlight a result that has a tighter bound for lower values of n_{\max} . It has a similar proof to Theorem 2.4 for $k = 2$ and can easily be computed for $k = 3$ and we therefore omit the proof.

Corollary 2.5. *Consider the split- k strategy and let $k > 1$. Furthermore, let N be uniformly distributed in $[2, n_{\max}]$ such that $k^l < n_{\max} \leq k^{l+1}$ for some $l \geq 1$. Then, for $k = 2$ and $l \leq 4$ and for $k = 3$ and $l = 1$,*

$$\mathbb{E}[e_{\max}(N)] \leq \frac{k - 1}{2k}.$$

2.2.2. Fixed error strategy

In this strategy, we bound the maximum error by imposing requirements on the minimum distance. That is, for any given n we have

$$e_{\max}(n) \leq e_{\text{fix}} \Leftrightarrow id_{\min}(i) \geq 1 - e_{\text{fix}} \quad \forall i \leq n, \tag{3}$$

where e_{fix} is the fixed error. This strategy leads to (De Bruijn & Erdős, 1949), where sequences of points on a circle are considered. Their main goal is to create a sequence of length n such that the largest and smallest distance between two points is as close as possible to $\frac{1}{n}$, regardless of n . The sequence they provide is shown to be the best possible sequence achieving their goal and therefore gives the best possible strategy for our problem when the number of persons is unknown and could be arbitrary large.

In the strategy from De Bruijn & Erdős (1949), person j is placed at position $\frac{\log(2j-1)}{\log(2)} \bmod 1$. This results in

$$d_{\min}(j) = \frac{\log\left(\frac{2j}{2j-1}\right)}{\log(2)}.$$

It is not difficult to see that when $j \rightarrow \infty$ we have $jd_{\min}(j) \rightarrow \frac{1}{2\log(2)}$ but also that $jd_{\min}(j)$ decreases towards this number. As a result, using this strategy we find that $e_{\max}(N) \leq 1 - \frac{1}{2\log(2)} \approx 0.279$ for any realization of N . Since the result from De Bruijn & Erdős (1949) is the best possible, the maximum error cannot be smaller than this value if we consider any possible realization of N . In particular, this upper bound also holds for the expected value. In Figure 2, we show the positions of the first 11 persons using the fixed error strategy.

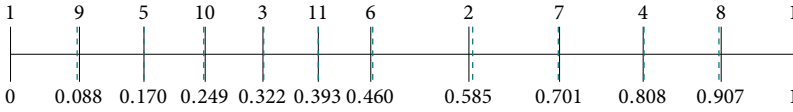


Figure 2. Example of the fixed error strategy when $e_{fix} = 1 - \frac{1}{2\log(2)}$ and $N = 11$ (solid lines). Above the line are the person numbers and below the line the positions (which are the distances from 0). The dashed lines are the positions when using the improved fixed error strategy for $N = n_{max} = 11$.

If, however, there is a known maximum of persons, this strategy can be slightly improved, see Appendix A.1 for the details. The dashed lines in Figure 2 show the small deviations from De Bruijn & Erdős (1949). The result of this improvement, when $n_{max} = 11$, is that $e_{max}(N) \leq 0.235$ for any realization of $N \leq n_{max}$ (compared to $e_{max}(N) \leq 0.262$ for the fixed error strategy given that $N \leq n_{max}$).

2.2.3. Fixed angle strategy

In the fixed angle strategy, we place persons along the roundtable with a distance based on some angle α , where $0 < \alpha < 2\pi$. The aim is to optimize the angle α , so that the maximum error is minimized. Note that for this strategy, the angle that minimizes the Euclidean (direct) distance is the same angle that minimizes the distance measured along the circle. As we consider a circle with a radius of $\frac{1}{2\pi}$, this implies a distance (along the circle) of $\frac{\alpha}{2\pi}$ between person j and $j + 1$, see also Figure 3. For notational convenience, we let $\beta = \frac{\alpha}{2\pi}$, that is, β represents the distance along the circle.

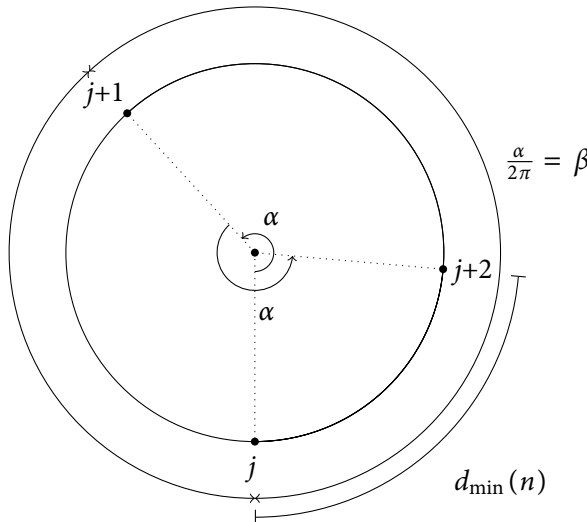


Figure 3. An example of the fixed angle strategy. Person j is w.l.o.g. located at the bottom of the table and the distance between person j and $j + 1$ is determined by the angle α .

The (circular) distance between any person i and j , defined as $d(i, j)$, is

$$d(i, j) = \min((j - i)\beta \bmod 1, 1 - (j - i)\beta \bmod 1).$$

From this equation, it is also clear that $d(i, j) = d(1, j + 1 - i)$, and so it suffices to consider the minimum distance between person 1 and all other possible persons $2, \dots, N$. As a result, the minimum distance $d_{\min}(n)$ for any number of n persons is

$$\begin{aligned} d_{\min}(n) &= \min_{1 \leq j \leq n-1} d(1, j + 1) \\ &= \min_{1 \leq j \leq n-1} \min(j\beta - \lfloor j\beta \rfloor, \lceil j\beta \rceil - j\beta), \end{aligned} \tag{4}$$

where the final equality follows by definition of the modulo function. Recall that we start with one person at the table and hence we only need to place $n - 1$ additional persons. It is known from (e.g. Niven, 1956) that the golden ratio is the angle that maximizes the minimum distance, that is, $\beta = \frac{1}{\varphi}$. The proof of the following theorem is hence omitted.

Theorem 2.6 (Maximum error using the golden ratio). *For any realization of N we find*

$$e_{\max}(N) < 1 - \frac{1}{\sqrt{5}},$$

when using the golden ratio.

Using Hurwitz’s theorem, we can provide a lower bound on the maximum error that gets very close to the upper bound of Theorem 2.6. Therefore, we find that the golden ratio minimizes the maximum error when considering a fixed angle strategy.

As for our other seating strategies, we also upper bound the expected maximum error using this strategy (see Appendix A.1 for the proof).

Theorem 2.7 (Expected maximum error using the golden ratio). *Consider the fixed angle strategy with the golden ratio. Furthermore, let N be uniformly distributed in $[2, n_{\max}]$ such that $F_{l+1} < n_{\max} \leq F_{l+2}$ for some $l \geq 1$, where F_k is the k -th Fibonacci number. Then,*

$$\mathbb{E}[e_{\max}(N)] < 1 - \frac{\sqrt{\varphi}}{\sqrt{5}}.$$

In addition, when $n_{\max} = F_{l+1}$ for some $l \geq 1$, we also have

$$\mathbb{E}[e_{\max}(N)] < 1 - \frac{\varphi^2}{2\sqrt{5}}.$$

2.2.4. Random strategy

In the random strategy, we place n persons uniformly at random along the circle. We find the following result (the proof can be found in Appendix A.1).

Theorem 2.8 (Expected error using the random strategy). *Suppose n persons are placed uniformly at random around a table. Then, for any fixed number n ,*
 $\mathbb{E}[E_{\max}(n)] = 1 - \frac{1}{n}$.

The randomness here is not yet in the number of persons who are seated around the table, but in their *position*. Hence, we also have the following corollary.

Corollary 2.9 (Maximum expected error using the random strategy). *Suppose N persons are placed uniformly at random around a table. Furthermore, let N be uniformly distributed in $[2, n_{\max}]$. Then,*

$$\mathbb{E}[E_{\max}(N)] = 1 - \frac{1}{n_{\max} - 1} \sum_{n=2}^{n_{\max}} \frac{1}{n} < 1.$$

While this upper bound may seem trivial, we note that it is not possible to improve the upper bound since for large enough n_{\max} the expected error gets arbitrary close to 1.

3. Experimental results

We conducted an experiment to evaluate the performance of actual human groups at social distancing and compare this to that of the theoretical strategies analyzed in the previous section. To this end, we recruited 31 military cadets of the Netherlands Defence Academy to perform a social distancing task in a large sports hall (details in Appendix A.2). Cadets were not offered any performance incentives. Each trial a specific number of cadets was asked to position themselves. We refer to this number as the group size, which varied between 2 and 11, each with probability 1/10. Cadets knew about group size probabilities but they learned what the group size was for a given trial only at the end of that trial. Cadets were first asked to stand in line, not being able to see the circle. Cadet numbers would be called one at a time. The cadet whose number was called had to position themselves at a location of their choice on the circle. They were instructed to stay put and not move once they had positioned themselves as the next cadet number was called. They were instructed to position themselves strategically with the aim of maximizing the ultimate distance between themselves and other participants at the end of the trial, not knowing how many cadets would be asked this trial. A total of 150 trials were collected with groups of at most 11 cadets. Figure 4 shows the

distribution of group size across the 150 trials. Two trials are colored orange because of irregularities that occurred in the writing of the results of these trials. According to the data, in these trials two cadets were seated on top of each other, which certainly did not occur during the experiment. Therefore, in the remainder of this paper, we only consider 148 trials. The distribution of group size is roughly uniform, with the exception of 11. The low numbers for 11 are due to a difference in the expected number of cadets and the actual numbers of cadets appearing at the experiment, which we explain in more detail in Appendix A.2. Therefore, in all results concerning the k -th person in a trial, $k = 11$ is left out. In Appendix A.2 we also explain why the distribution of group size slightly deviates from a uniform distribution.

The average maximum error, see Equation (2), achieved by humans is 0.294 when irregular trials are excluded (and 0.304 when they are included). Table 2 compares average human error with the average error achieved by each of our six theoretical strategies. Human error is not just much smaller than what can be achieved through random behavior but is also smaller than the fixed (golden) angle strategy and the split-3 strategy. As such it has an error that is much closer to optimal behavior than to random behavior. Average human error is, however, strictly greater than what is achieved by the split-2 and fixed error strategies. The variance of human error is comparable with what is achieved by a split-2 strategy. Table 2, columns 3 and 4, reports the results of t-tests of the null hypothesis that human error is smaller than

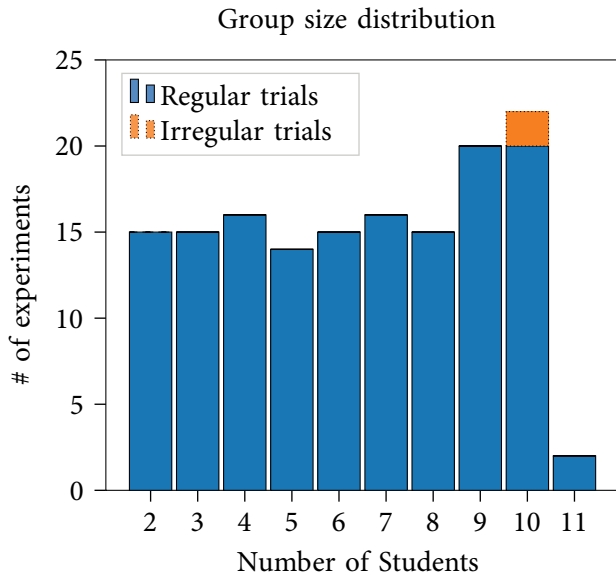


Figure 4. The distribution of the group size across all trials. The orange trials contain irregularities and are left out in the remainder of the work.

Table 2. The average error achieved by each of the strategies, when using the distribution of the regular trials in [Figure 4](#). The average error at experiment is 0.294 and the error variance is 0.026. The p-value and power with superscript 1 are results from a t-test in which we define our null hypothesis H_0 as “Humans perform better than the strategy mentioned” and hence our alternative hypothesis H_1 is “The strategy performs better than human”. Vice versa, the p-value and power with superscript 2 are results from a t-test with H_0 “The strategy performs better than human” and H_1 “Human perform better than the strategy mentioned”. All values are rounded up to 3 decimals and when smaller than computer epsilon they are denoted as 0. The fixed error finite strategy can be found in [Appendix A.1](#).

Strategy	Average error	Error variance	p-value ¹	Power ¹	p-value ²	Power ²
Split-2	0.214	0.028	1.546e-08	1.000	1.000	1.438e-09
Fixed error finite	0.225	0.035e-02	1.267e-07	1.000	1.000	1.571e-13
Fixed error	0.238	0.074e-02	4.837e-06	1.000	1.000	1.365e-09
Split-3	0.298	0.051	0.570	0.029	0.430	0.081
Fixed angle	0.371	0.008	1.000	3.604e-13	6.396e-09	1.000
Random	0.795	0.015	1.000	0.000	0.000	1.000

each of the theoretical strategies. The null hypothesis is rejected for the split-2 strategy and the fixed error strategy. Vice versa, columns 5 and 6 of [Table 2](#), report the results of t-tests of the null hypothesis that human error is greater. This null hypothesis can be rejected for both the random and the fixed angle strategy. In [Appendix A.3](#), we also show the probability that humans outperform the theoretical strategies and results from comparable Wilcoxon tests, see [Table A1](#). We conclude that self-organized group performance is only second to split-2 and fixed error, is comparable to what can be achieved with split-3, and is significantly better than the fixed angle and random strategies.

How do the cadets achieve these very good results? The answer is that the basic intuitive behavioral rules that they tended to follow happen to concord with optimal strategies. [Figure 5](#) shows the relative deviation from the middle of the gap when the k -th person is entering. In particular, this figure shows that the second person, $k = 2$, often chooses an angle of approximately 180 degrees. This is consistent with a split-2 strategy, which is among our best performing theoretical strategies. [Figure 5](#) shows a second mode, which is an angle of about 120 degrees that some cadets opt for. This angle is consistent with split-3 as well as with the fixed angle and fixed error strategies.

Another related feature of seating behavior that may have helped achieve high performance is that cadets – after choosing their gap – tend to choose the halfway point in between the two nearest neighbors, not only for the second person. It can be seen in [Figure 5](#) that the middle position, or a position close to the middle, is often chosen.

A final feature of cadet seating behavior that contributes to high group performance is that most cadets positioned themselves in the largest available gap. [Figure 6](#) shows the fraction of cadets choosing to stand in the largest gap by the order k in which they enter the circle. For most k

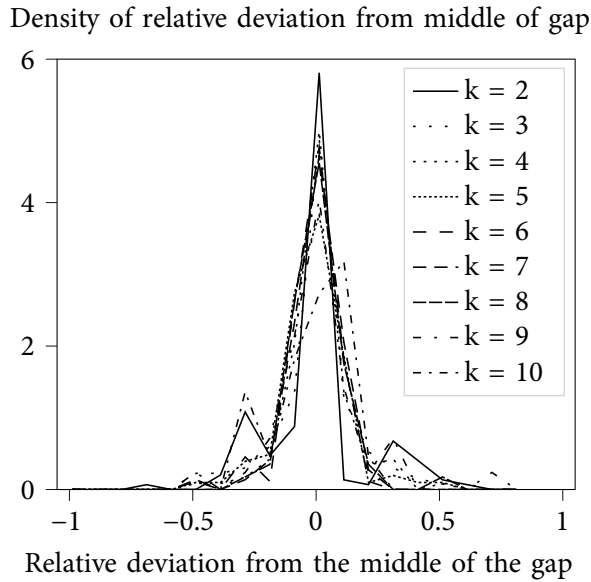


Figure 5. The density of the relative deviation from the middle of the gap, when the k -th person is entering.

a clear majority shows the expected behavior. We would expect that this proportion declines with k given that at higher k there are more gaps to choose from and their relative sizes are more difficult to discern. This is indeed what [Figure 6](#) shows, except for $k = 5$ and $k = 9$. A possible explanation for these exceptions is that at these points there often were multiple gaps available of approximately equal size. This is so because if a group approximately follows split-2, then at $k = 5$ and $k = 9$ the previous cadets will be positioned at four respectively eight positions that are roughly equally spaced. If a cadet chose a gap with a size of 80 degrees instead of an available gap of 90 degrees, the cadet chose what was almost but not quite the *largest* gap. In [Figure A3](#) in [Appendix A.3](#), we show that indeed most of the cadets chose a gap that differs at most 5 degrees (or 5 percent) from the largest gap. In particular when it was more difficult to see which gap was the largest, cadets may have decided based on convenience, as around half of the circle required less walking for cadets than the other half, see [Figure A1](#). [Figure 7](#) shows the fraction of decisions that contained less walking, denoted by “close”, and we see that in particular for $k = 5$ and $k = 9$ there is a much higher fraction of cadets that chose a position that was “close” to the area where they were waiting. These two arguments – hard-to-discern gap size differences and nearby gaps being more convenient – together explain the two dips in [Figure 6](#) in the fraction of decisions in the largest gap.

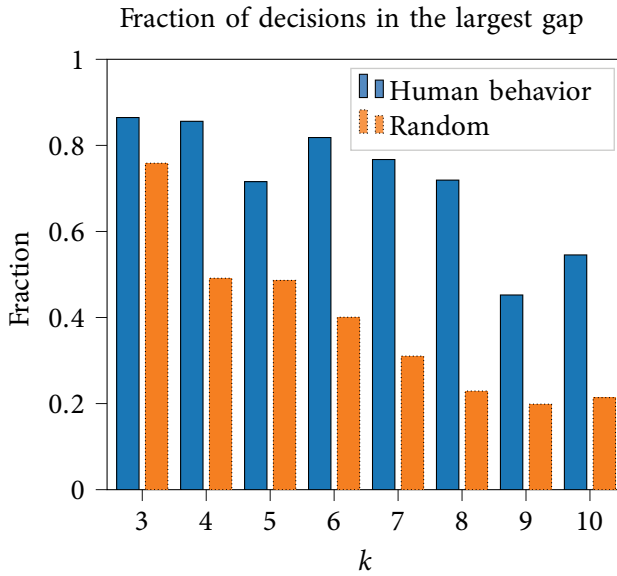


Figure 6. The fraction of decisions that are in the largest gap, when a person is seated as k -th in line, as chosen by experimental participants. Random denotes the fraction of the circle that, on average, represents the largest gap(s) – multiple when it is not unique. By definition, the second cadet *always* chooses the largest gap. In cases where the previous participants equally divided the circle, which happened frequently for $k=3$ and $k=5$, decisions must necessarily fall in the largest gap, because then all gaps are large.

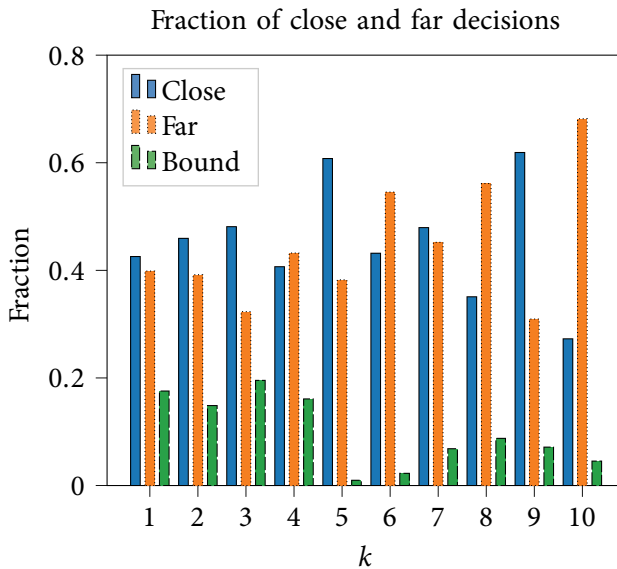


Figure 7. The fraction of decisions that contain certain degrees, see [Figure A1](#) where all degrees were located. Close refers to all degrees between 90 and 270, excluding the bounds and far refers to all degrees 270 till 360 and 0 till 90, again excluding the bounds. Bounds are all people that are seated exactly at 90 or 270 degrees.

4. Discussion

With regard to the problem of self-organization we posed in this paper, the results together provide a tentative positive answer to the question whether people can from the bottom up produce efficient spatial arrangements that reduce disease spread. Our findings suggest that one can in principle reduce the propagation of viruses that spread through close-range contact beyond what is achieved with familiar minimum-distance policies alone. When people are compelled to maximize distance, they can achieve a superior distancing performance. This is so because, as it turned out in our analysis of the roundtable setting we studied, intuitive behavior tended to concord with optimal behavior.

However, a number of important caveats and qualifications condition this tentative conclusion. First, the paper does not study the role of incentives. The question of how to effectively motivate people to try and maximize distance in times of a pandemic is beyond its scope (see e.g. Epton et al., 2022; Lunn et al., 2020). We assumed throughout that individuals were equally interested in social distancing. This ignores that less vulnerable individuals tended to violate social distancing rules during the COVID-19 pandemic (Bavel et al., 2020), which some have found to have generated distrust among the more vulnerable (Iacono et al., 2021). That said, the results of this study have relevance also in settings where people cannot reasonably be expected to effectively maximize distance because they do not care to or because it is too hard. For such situations, our mathematical strategies provide position numbers that a host could assign to people that achieve a better performance than with mere minimum-distance policy (such as the split-2 strategy and the fixed error strategy). Seats could be marked with numbers such that the k -th person entering takes the seat with number k .

Second, while the split-2 strategy was the best performing theoretical strategy for our experimental setting, the mathematics are clear that this does not always hold true. In particular when the bound on the number of persons that may need to be seated around the table is unknown, the upper bound on the error of the split-2 strategy is worse than that of the fixed error strategy, see Table 1. In such cases, the variance of the split-2 strategy is higher than that of the fixed error strategy. In addition, the average error of the split-2 strategy is more sensitive to the distribution of the number of persons that arrive at the roundtable compared to the fixed error strategy. We can hence conclude that depending on the setting and the required performance, a different strategy may be recommendable. The intuitive behavior displayed in this present empirical study may then be suboptimal.

Third, the empirical study was necessarily limited in scale to small groups of at most 11 subjects, raising the question how the results might hold up in larger populations. Subject behavior was found most consistent with a split-2

strategy, whereby individuals maximize the distance to those already present, ignoring who might come later. This theoretical strategy's good performance scales to larger populations. At the same time, we have presented evidence that the more subjects had already positioned themselves, the less accurate subsequent subjects were in identifying the largest open gap (Figure 6) and the middle positions in those gaps (Figure 5). This suggests that when holding space constant and increasing population size, performance will decrease. Alternatively, if instead we hold density constant but increase population size and space proportionately, then it will become increasingly difficult for individuals to identify optimal locations because of limited visibility and increased informational complexity. In large structures, e.g. a stadium, a two-stage approach may work whereby individuals are equally distributed to smaller spatial units, e.g. stadium sections, within which they can feasibly self-organize. Another limitation of the empirical study is that it relied exclusively on cadets as subjects, who may differ from the overall population in compliance and whose military training may contribute to an ability to assess distance.

Fourth, the problem as currently formulated ignores interactions individuals might have outside the focal setting. If the space studied brings together people from different network clusters, e.g. in meetings or classrooms (Manzo, 2020; Weeden & Cornwell, 2020), then optimal arrangements may put together individuals sharing offices or living quarters, so that they can be jointly placed at a greater distance from others (Block et al., 2020; Centola, 2020).

Fifth, our analysis made the assumption that group size is uniformly distributed and bounded. An important open task is to thoroughly and systematically investigate how strategies, particularly the split-2 strategy, perform under different distributions that more closely match empirical distributions in relevant contexts. One attractive alternative distribution of group size that an anonymous reader suggested is the exponential distribution, which has the features of being unbounded and memoryless.

A natural extension of the present research is the development of strategies for other classes of seating arrangements. Our mathematical results may inspire but do not generalize to analytical solutions for other settings such as theater or classroom settings (Weeden & Cornwell, 2020) with more than one spatial dimension. Whether humans can also effectively self-organize social distancing in such alternative seating arrangements is not studied here and is an open question.

The question of how feasible self-organization is besides a matter of efficiency in disease prevention also a matter of public support for public policy. Also during future pandemics, the public will likely grow increasingly impatient with restrictions on movement imposed top-down. Sociologists have found such restrictions to have had significant negative impact on societal cohesion (Bavel et al., 2020; Bol, 2020; Ward, 2020). Exploring to what degree

populations can be allowed to self-organize with milder movement restrictions may help reduce such negative impacts.

Acknowledgements

We would like to thank all cadets that participated in the experiment and the NLDA for the possibility to perform our experiment. We would like to thank Robbert Fokkink, Bart van Ginkel and Trineke Palm for their assistance in preparing and performing the experiment, and Antoni Brentjes, Robbert Fokkink and Marijn ten Thij for the fruitful discussions on different strategies.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data and code availability

All data and source code used for data analysis in this study are freely available in de-identified format through a dedicated and open-access GitHub repository: <https://www.github.com/Roy-Tilburg/SelfOrgSD>

References

- Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Molly J. Crockett, Alia J. Crum, Karen M. Douglas, James N. Druckman, John Drury, Oeindrila Dube, Naomi Ellemers, Eli J. Finkel, James H. Fowler, Michele Gelfand, Shihui Han, S. Alexander Haslam, Jolanda Jetten, Shinobu Kitayama, . . . , Robb Willer. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, 4(5), 460–471.
- Becker, J., Brackbill, D., & Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, 114(26), E5070–E5076.
- Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, 4(6), 588–596.
- Blom, D., Pendavingh, R., & Spieksma, F. (2021). Filling a theater during the COVID-19 pandemic. *INFORMS Journal on Applied Analytics* 52(6): 473–484. <https://doi.org/10.1287/inte.2021.1104>
- Bol, T. (2020). Inequality in homeschooling during the corona crisis in the Netherlands. first results from the liss panel.
- Buskens, V., Corten, R., & Weesie, J. (2008). Consent or conflict: Coevolution of coordination and networks. *Journal of Peace Research*, 45(2), 205–222.
- Centola, D. (2020). Considering network interventions. *Proceedings of the National Academy of Sciences*, 117(52), 32833–32835.
- Centola, D., & Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, 112(7), 1989–1994.
- Centola, D., Becker, J., Brackbill, D., & Baronchelli, A. (2018). Experimental evidence for tipping points in socialconvention. *Science*, 360(6393), 1116–1119.

- Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H. J., on behalf of the COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. *Lancet*, 395 (10242), 1973–1987.
- Courtemanche, C., Garuccio, J., Le, A., Pinkston, J., & Yelowitz, A. (2020). Strong social distancing measures in the United States reduced the COVID-19 growth rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed COVID-19 cases across the United States. *Health Affairs*, 39(7), 1237–1246.
- De Bruijn, N., & Erdős, P. (1949). Sequences of points on a circle. *Proceedings of the Section of Sciences of the Koninklijke Nederlandse Akademie van Wetenschappen te Amsterdam*, 52(1), 14–17.
- Diekmann, A. (1985). Volunteer's dilemma. *Journal of Conflict Resolution*, 29(4), 605–610.
- Diekmann, A., Jann, B., Przepiorka, W., & Wehrli, S. (2014). Reputation formation and the evolution of cooperation in anonymous online markets. *American Sociological Review*, 79 (1), 65–85.
- Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society* 61(5), 1047–1071. <https://doi.org/10.2307/2951493>.
- Epton, T., Ghio, D., Ballard, L. M., Allen, S. F., Kassianos, A. P., Hewitt, R. Katherine Swainston, Wendy Irene Fynn, Vickie Rowland, Juliette Westbrook, Elizabeth Jenkinson, Alison Morrow, Grant J. McGeechan, Sabina Stanesco, Aysha A. Yousuf, Nisha Sharma, Suhana Begum, Eleni Karasouli, Daniel Scanlan, Gillian W. Shorter, . . . John Drury. (2022). Interventions to promote physical distancing behaviour during infectious disease pandemics or epidemics: A systematic review. *Social Science and Medicine* 303(June), 114946. <https://doi.org/10.1016/j.socscimed.2022.114946>.
- Fazio, R. H., Ruisch, B. C., Moore, C. A., Granados Samayoa, J. A., Boggs, S. T., & Ladanyi, J. T. (2021). Social distancing decreases an individual's likelihood of contracting COVID-19. *Proceedings of the National Academy of Sciences*, 118(8), e2023131118.
- Frey, V., & Van de Rijt, A. (2021). Social influence undermines the wisdom of the crowd in sequential decision making. *Management Science*, 67(7), 4273–4286.
- Friedkin, N. E., & Bullo, F. (2017). How truth wins in opinion dynamics along issue sequences. *Proceedings of the National Academy of Sciences*, 114(43), 11380–11385.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481–510.
- Guilbeault, D., Baronchelli, A., & Centola, D. (2021). Experimental evidence for scale-induced category convergence across populations. *Nature Communications*, 12(1), 327.
- Heckathorn, D. D. (1996). The dynamics and dilemmas of collective action. *American Sociological Review* 61(2), 250–277. <https://doi.org/10.2307/2096334>.
- Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101(46), 16385–16389.
- Horne, C., & Mollborn, S. (2020). Norms: An integrated framework. *Annual Review of Sociology*, 46(1), 467–487.
- Iacono, S. L., Przepiorka, W., Buskens, V., Corten, R., & van de Rijt, A. (2021). COVID-19 vulnerability and perceived norm violations predict loss of social trust: A pre-post study. *Social Science & Medicine*, 291, 114513.
- Jiao, R., Przepiorka, W., & Buskens, V. (2021). Reputation effects in peer-to-peer online markets: A meta-analysis. *Social Science Research*, 95(March), 102522 <https://doi.org/10.1016/j.ssresearch.2020.102522>.

- Keuschnigg, M., & Ganser, C. (2017). Crowd wisdom relies on agents' ability in small groups with a voting aggregation rule. *Management Science*, 63(3), 818–828.
- Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, 108(22), 9020–9025.
- Lunn, P. D., Timmons, S., Belton, C. A., Barjaková, M., Julienne, H., & Lavin, C. (2020). Motivating social distancing during the COVID-19 pandemic: An online experiment. *Social Science & Medicine*, 265 (November), 113478. <https://doi.org/10.1016/j.socscimed.2020.113478>.
- Macy, M. W. (1991). Chains of cooperation: Threshold effects in collective action. *American Sociological Review* 56(6), 730–747. <https://doi.org/10.2307/2096252>.
- Manzo, G. (2020). Complex social networks are missing in the dominant COVID-19 epidemic models. *Sociologica*, 14(1), 31–49.
- Mason, W., & Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, 109(3), 764–769.
- McGrail, D. J., Dai, J., McAndrews, K. M., & Kalluri, R. (2020). Enacting national social distancing policies corresponds with dramatic reduction in COVID-19 infection rates. *PLOS ONE*, 15(7), e0236619.
- Niven, I. (1956). *Irrational numbers* (No. 11). The Mathematical Association of America.
- Oliver, P., Marwell, G., & Teixeira, R. (1985). A theory of the critical mass. i. interdependence, group heterogeneity, and the production of collective action. *American Journal of Sociology*, 91(3), 522–556.
- Raub, W., & Weesie, J. (1990). Reputation and efficiency in social interactions: An example of network effects. *American Journal of Sociology*, 96(3), 626–654.
- Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1(2), 143–186.
- Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Borelli, M., & Miani, A. (2020). *Airborne transmission route of COVID-19: Why 2 meters/6 feet of inter-personal distance could not be enough* (Vol. 17) (No. 8). MDPI.
- Simpson, B., & Willer, R. (2015). Beyond altruism: Sociological foundations of cooperation and prosocial behavior. *Annual Review of Sociology*, 41(1), 43–63.
- Sun, C., & Zhai, Z. (2020). The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable Cities and Society*, 62(November), 102390. <https://doi.org/10.1016/j.scs.2020.102390>.
- van der Corput, J. G. (1935). Verteilungsfunktionen I & II. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen* Vol. 38. pp. 1058–1066.
- Ward, P. R. (2020). A sociology of the COVID-19 pandemic: A commentary and research agenda for sociologists. *Journal of Sociology*, 56(4), 726–735.
- Weeden, K. A., & Cornwell, B. (2020). The small-world network of college classes: Implications for epidemic spread on a university campus. *Sociological Science*, 7, 222–241. doi:10.15195/v7.a9.
- Young, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society* 61(1), 57–84. <https://doi.org/10.2307/2951778>.

Appendix A.

A.1. Mathematics

In this section, all proofs of mathematical theorems are presented, as well as some (heuristic) improvements of the methods that are presented.

A.1.1. Split- k

Proof of Lemma 2.3. For any realization of N , we can find an $x \geq 0$ such that $k^x < N \leq k^{x+1}$. For this value of x we have $d_{\min}(N) = \frac{1}{k^{x+1}}$ and so, by using (1),

$$e_{\max}(N) = 1 - Nd_{\min}(N) < 1 - \frac{k^x}{k^{x+1}} = 1 - \frac{1}{k},$$

for any realization of N . This concludes the proof. \square

Proof of Theorem 2.4. By definition in Equation (2) we have

$$\begin{aligned} \mathbb{E}[e_{\max}(N)] &= \sum_{n=2}^{n_{\max}} \frac{1}{n_{\max} - 1} (1 - nd_{\min}(n)) \\ &= 1 - \frac{1}{n_{\max} - 1} \left(\sum_{x=1}^l \sum_{n=k^{x-1}+1}^{k^x} \frac{n}{k^x} + \sum_{n=k^l+1}^{n_{\max}} \frac{n}{k^{l+1}} \right) \\ &= 1 - \frac{1}{n_{\max} - 1} \left(\sum_{x=1}^l \frac{(k^x - k^{x-1})(k^x + k^{x-1} + 1)}{2k^x} \right. \\ &\quad \left. + \frac{(n_{\max} - k^l)(n_{\max} + k^l + 1)}{2k^{l+1}} \right), \end{aligned}$$

where the final step follows by computing the average of all terms in the summation. Then, we can rewrite the second term of the above, by cancelling out k^{x-1} and using a telescoping sum, and rewrite the third term, so that we find

$$\begin{aligned} \mathbb{E}[e_{\max}(N)] &= 1 - \frac{1}{n_{\max} - 1} \left(\frac{(k^l - 1)(k + 1) + (k - 1)l}{2k} \right. \\ &\quad \left. + \frac{n_{\max}(n_{\max} + 1) - k^l(k^l + 1)}{2k^{l+1}} \right) \\ &= 1 - \frac{k^{2l+1} + [(k - 1)l - (k + 2)]k^l + n_{\max}(n_{\max} + 1)}{2k^{l+1}(n_{\max} - 1)}. \end{aligned} \tag{A1}$$

Since we are interested in an upper bound on the expected value, we take derivatives with respect to n_{\max} and find that the maximum of $\mathbb{E}[e_{\max}(N)]$ is attained whenever

$$(n_{\max}^* - 1)^2 = 2 + k^{2l+1} + [(k - 1)l - (k + 2)]k^l. \tag{A2}$$

While this optimum value n_{\max}^* might be non-integer, as n_{\max} should be, using this optimum value does provide an upper bound on the expected maximum error, even though it may be slightly worse compared to when n_{\max}^* would have been integer. It is possible to verify that

indeed $n_{\max}^* \in (k^l, k^{l+1}]$. Consider all cases where $l \geq \frac{k+2}{k-1}$. Then, we can upper bound the expected value as follows

$$\begin{aligned} \mathbb{E}[e_{\max}(N)] &\leq 1 - \frac{(n_{\max}^* - 1)^2 - 2 + n_{\max}^*(n_{\max}^* + 1)}{2k^{l+1}(n_{\max}^* - 1)} \\ &= 1 - \frac{2(n_{\max}^* - 1) + 3}{2k^{l+1}} \\ &< 1 - \frac{n_{\max}^* - 1}{k^{l+1}} \\ &< 1 - \frac{\sqrt{k}k^l}{k^{l+1}} = 1 - \frac{1}{\sqrt{k}}, \end{aligned} \tag{A3}$$

where the first inequality follows by using (A2), the first equality follows by cancelling out $n_{\max}^* - 1$, the second inequality follows trivially and the final inequality follows by using that $(n_{\max}^* - 1)^2 > k^{2l+1}$ if $l \geq \frac{k+2}{k-1}$.

It remains to show that the upper bound also holds when $1 \leq l \leq \frac{k+2}{k-1}$. In these cases, we need to show that optimum value of n_{\max}^* in (A2) satisfies

$$(n_{\max}^* - 1) \geq \frac{1}{\sqrt{k}}k^{l+1} - \frac{3}{2},$$

so that it immediately follows from (A3) that the expected error is indeed bounded above by $1 - \frac{1}{\sqrt{k}}$.

Thus, we need to show that

$$2 + k^{2l+1} + [(k-1)l - (k+2)]k^l \geq \left(\frac{1}{\sqrt{k}}k^{l+1} - \frac{3}{2}\right)^2,$$

which is equivalent to

$$(k-1)l - (k+2) + 3\sqrt{k} \geq \frac{1}{4k^l}.$$

This inequality is clearly satisfied whenever $l \geq 1$, but it is also satisfied for all k for which $3\sqrt{k} \geq k+2$ (since $(k-1)l \geq 1$ for all $k \geq 2$ and $l \geq 1$). The latter is true for all $k \in [1, 4]$, which includes those k where $\frac{k+2}{k-1} \geq 2$. Thus, the first statement follows.

To show the second statement, we note that for $n_{\max} = k^l$ Equation (A1) equals

$$\mathbb{E}[e_{\max}(N)] = 1 - \frac{k+1}{2k} - \frac{(k-1)l}{2k(k^l-1)} < \frac{k-1}{2k},$$

for any l . This concludes the proof. \square

A.1.1.1. Improvements for split- k For all $k > 3$, we can improve heuristically as follows. Assume that for all $k^l < n \leq k^{l+1}$, we can only use the positions $\frac{j}{k^{l+1}}$ $j = 0, \dots, k^{l+1}$ (some of which are already occupied by $1 < n \leq k^l$). Then, we first pick the position that is “almost” in the middle of the (largest) interval. After that we can divide each interval such that the minimum distance is (roughly) halved, until all positions are filled. Then, we repeat this procedure for all newly created intervals, where we first pick the same “almost” middle position from each interval, and so on. We claim that, without stating any proof, this procedure results

in $e_{\max}(N) \leq \frac{\lfloor \frac{k}{2} \rfloor + 1}{k}$ for all N . This results from the fact that person number $(\lfloor \frac{k}{2} \rfloor - 1)k^x + 1$ is always the person number that ensures that the minimum distance is $\frac{1}{k^{x+1}}$ and hence creates the largest error. That means that, when k is large enough, this procedure approximates split-2. Note that for all k that equal some power of 2, split- k has the same performance as split-2 when always dividing each interval in half.

A.1.2. Improvements for the fixed error strategy

In this section, we present an improvement of the fixed error strategy from (De Bruijn & Erdős, 1949). Section 4 of that paper provides a maximum error for an infinite amount of persons that need to be seated (in their paper referred to as smallest length of intervals based on sequences of points on a circle), as discussed in Section 2.2.2 of this paper. But if we take a closer look at the proof method in Section 4 of their paper, we can determine better bounds for a finite amount of persons to be seated. To this end, we present a finite seating sequence with the positions of all persons that possibly need to be seated.

We first show the main result, then we present the finite seating sequence. After that, we prove that for any optimal finite seating sequence, the lengths of all the intervals are uniquely determined. That is, the distances between all neighbors on the circle are uniquely determined. After proving our main result, it follows that the finite seating sequence we present is optimal. That is, it is not possible to improve upon this result.

For notational convenience, we let $\hat{n} = \lceil \frac{n_{\max}}{2} \rceil$ (i.e. $2\hat{n} = n_{\max}$ if n_{\max} is even, and $2\hat{n} = n_{\max} + 1$ if n_{\max} is odd). The main result of this section is as follows.

Theorem A.1 (Maximum error using the improved fixed error strategy). *For any realization of $N \leq n_{\max}$ we find*

$$e_{\max}(N) \leq 1 - \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})},$$

where H_x is the x -th harmonic number, when using the improved fixed error strategy.

We define our seating sequence as follows.

Definition A.2 (Seating sequence s). Person number j is seated at position s_j , with

$$s_j = \begin{cases} \frac{2H_{b(1)} - 2H_{(2j-1)b(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})} \bmod 1 & \text{if } 1 \leq j \leq \hat{n} \\ \frac{2H_{b(1)} - H_j - H_{j-1}}{2(H_{2\hat{n}} - H_{\hat{n}})} \bmod 1 & \text{if } \hat{n} < j \leq 2\hat{n} \end{cases},$$

and

$$b(j) = 2 \lfloor \log_2 \left(\frac{2\hat{n}}{j-1} \right) \rfloor.$$

Furthermore, we let $\lambda(w)$ denote how “close” we are to the optimal seating for a given seating sequence w .

Definition A.3 For all seating sequences w of maximum length n_{\max} , we define $\lambda(w)$ as

$$\lambda(w) = \min_{2 \leq n \leq n_{\max}} n d_{\min}^{(w)}(n),$$

cf. (1), where $d_{\min}^{(w)}(n)$ is the minimum distance with n persons for a given seating sequence w , and we define λ so that

$$\lambda = \sup_w \lambda(w).$$

We first show that the lengths of the intervals for any finite optimal seating sequence are uniquely determined. By uniquely determined we mean that if the interval lengths are sorted on length, the sorted sequence is unique. If two seating sequences end up with the same sorted interval lengths at every step, we regard those as equivalent.

Lemma A.4. *For any optimal finite seating sequences, the lengths of the intervals are uniquely determined, after every person has taken their seat.*

Proof. Let $w := (w_1, \dots, w_{2n})$ be a seating sequence of even length. After all $2n$ persons have taken their seat, the circle is split up into $2n$ intervals. Let the length of these intervals be a_1, a_2, \dots, a_{2n} , sorted in descending order.

Since a_{2n} is the smallest interval, we know that $a_{2n} \geq \frac{\lambda(w)}{2n}$, by definition of $\lambda(w)$. We want to find similar bounds for all intervals. We do this by looking at the size of the intervals right before the last person takes their seat, so when the other $2n - 1$ persons have already taken their seat. The seating of a person splits an interval into two and therefore creates exactly 2 new intervals. This means that before the last person took their seat, at least one of the three intervals $a_{2n-2}, a_{2n-1}, a_{2n}$ must have already existed. The interval or intervals that already existed must be at least of size $\frac{\lambda(w)}{2n-1}$ by using the same argument. Since a_{2n-2} is the largest interval of the three, it must in particular hold that $a_{2n-2} \geq \frac{\lambda(w)}{2n-1}$.

We can repeat this step, by looking at the intervals right before the last two persons take their seat, or right before the last three persons take their seat, etcetera. This gives the following bounds for all intervals with an even index:

$$a_{2n} \geq \frac{\lambda(w)}{2n}, \quad \dots, \quad a_{2n-2k} \geq \frac{\lambda(w)}{2n-k}, \quad \dots, \quad a_2 \geq \frac{\lambda(w)}{n+1}. \tag{A4}$$

For any a_j with j uneven, $a_j \geq a_{j+1}$ is a lower bound and so we now have a lower bound for every interval a_i , for all $1 \leq i \leq 2n$.

Since all intervals together form the circle with circumference 1, we know that

$$a_{2n} + a_{2n-1} + \dots + a_1 = 1.$$

Combining this equation with all the lower bounds in (and below) (A4), we get

$$\lambda(w) \left(\frac{2}{2n} + \frac{2}{2n-1} + \dots + \frac{2}{n+1} \right) \leq 1,$$

or equivalently,

$$\lambda(w) \leq \frac{1}{2(H_{2n} - H_n)}. \tag{A5}$$

Letting n go to ∞ here provides the upper bound of $\frac{1}{2 \log 2}$ as found in (De Bruijn & Erdős, 1949). However, Equation (A5) also provides an upper bound for finite n . We will show in the proof of Theorem A.1 that the seating sequence as in Definition A.2 is tight on this bound, and thus optimal.

For any optimal seating sequence, it must therefore hold that $\lambda(w)$ is equal to the bound in Equation (A5). The upper bound on $\lambda(w)$ in Equation (A5) is determined by summing over the lower bounds of the length of all intervals in Equation (A4). This means that for any optimal seating sequence, since Equation (A5) holds with equality, every bound in Equation (A4) must hold with equality as well. Since every interval present after every person has taken their seat

has a lower bound in Equation (A4), this means that this uniquely determines the lengths of intervals after every person has taken their seat. \square

Lemma A.5. *Let w be an optimal finite seating sequence of length $2n$. After $2j$ persons in the sequence have been seated, let the lengths of the intervals be a_1, a_2, \dots, a_{2j} , sorted in descending order. It then holds that*

$$\frac{\lambda}{j+i-1} > a_{2i-1} \geq a_{2i} \geq \frac{\lambda}{j+i}, \quad (\text{A6})$$

for $i = 1, \dots, j$. Furthermore, every time a person is seated, it is seated in the interval with largest length, and the two newly created intervals are smaller than all preexisting intervals.

Proof. We prove (A6) by backward induction, with the second statement as an immediate result. From Lemma A.4, we immediately find that (A6) holds for $j = n$.

We now need to show that if (A6) holds when $2j$ persons have been seated, it also holds when $2j - 2$ persons have been seated. Furthermore, we show that both person $2j$ and $2j - 1$ are seated in what was the largest interval at the time, and the two newly created intervals are smaller than all preexisting intervals.

Suppose that indeed (A6) holds. From this induction hypothesis we can determine that for the smallest two intervals it holds that $\frac{\lambda}{2j-1} > a_{2j-1} \geq a_{2j} \geq \frac{\lambda}{2j}$. By definition of λ , in any optimal seating sequence, any interval smaller than $\frac{\lambda}{x}$ cannot exist after x persons have been seated. Therefore, the two intervals a_{2j-1} and a_{2j} were created when the $2j$ -th person took the seat. So, after $2j - 1$ persons were seated, a_{2j-1} and a_{2j} still formed one interval together, of length $a_{2j-1} + a_{2j} =: I_y$. By our induction hypothesis we now have

$$\frac{\lambda}{j-\frac{1}{2}} > I_y \geq \frac{\lambda}{j}. \quad (\text{A7})$$

This means interval I_y becomes the interval with the largest length. Thus person $2j$ was placed in the interval with largest length, and the two intervals created by the seating of person $2j$ are smaller than all preexisting intervals.

Now we look at the intervals with lengths a_{2j-3} and a_{2j-2} . These are the smallest intervals after $2j - 1$ persons have been seated, since the intervals with lengths a_{2j-1} and a_{2j} no longer exist (i.e. these have been created by person $2j$). From the induction hypothesis we find that $\frac{\lambda}{2j-2} > a_{2j-3} \geq a_{2j-2} \geq \frac{\lambda}{2j-1}$. Since these intervals are therefore too small to be created by the seating of one of the first $2j - 2$ persons, it must mean that these two intervals were created when the $(2j - 1)$ -th person was seated. So, after only $2j - 2$ persons were seated, a_{2j-3} and a_{2j-2} formed one interval together. This interval has a length of $I_z := a_{2j-3} + a_{2j-2}$, and by induction hypothesis it holds that

$$\frac{\lambda}{j-1} > I_z \geq \frac{\lambda}{j-\frac{1}{2}}. \quad (\text{A8})$$

From (A7) and (A8) it follows that

$$\frac{\lambda}{j-1} > I_z \geq \frac{\lambda}{j-\frac{1}{2}} > I_y \geq \frac{\lambda}{j}. \quad (\text{A9})$$

Thus, I_z is the interval with the largest length and hence person $2j - 1$ was also seated in the interval with largest length, and the two newly created intervals by the seating of person $2j - 1$ are smaller than all preexisting intervals.

Therefore, after $2j - 2$ persons have taken their seats, there exist $2j - 2$ intervals with lengths $I_z > I_y \geq a_1 \geq a_2 \geq \dots \geq a_{2j-4}$. Using (A9) and (A6) for $i = 1, \dots, (j - 2)$, which holds by assumption, we have the bounds on interval lengths required by the induction step. \square

Theorem A.6. *For any optimal finite seating sequences, the lengths of the intervals are uniquely determined after any number of persons are seated.*

Proof. In Lemma A.5 we proved that, in any optimal finite seating sequence, any person must always be seated in the largest interval to split it into the two smallest intervals. If the length of these two smaller intervals is unique, then the length of the original largest interval must be unique too. Since we have shown in Lemma A.4 that the lengths of the intervals are unique when all persons are seated, it follows from backward induction that all lengths of the intervals are unique after any number of persons are seated. \square

As a final lemma before proving our main result, we show some properties of the function $(2j - 1)b(j)$. This function plays an important role in the seating sequence, see Definition A.2.

Lemma A.7 (Properties of $(2j - 1)b(j)$).

The function $(2j - 1)b(j)$ has the following properties.

- (1) *The function $f : [1, \dots, \widehat{n}] \rightarrow [\widehat{n} + 1, \dots, 2\widehat{n}]$, $f(j) = (2j - 1)b(j)$ is bijective.*
- (2) *If $(2j - 1)b(j) = \max_{i \leq j} (2i - 1)b(i)$, then $2jb(j) > 2\widehat{n}$ and $b(j) > 1$.*
- (3) *If $(2j - 1)b(j) = \min_{i \leq j} (2i - 1)b(i)$, then $(j - 2)b(j) \leq \widehat{n}$.*

Proof. We first prove the first property. It is easy to show that $(2j - 1)b(j)$ is always an integer in $[\widehat{n} + 1, \dots, 2\widehat{n}]$, since by definition of $b(j)$ we have

$$\frac{\widehat{n}}{2j - 1} = 2^{\log_2 \left(\frac{\widehat{n}}{2j-1}\right) - 1} < b(j) \leq 2^{\log_2 \left(\frac{2\widehat{n}}{2j-1}\right)} = \frac{2\widehat{n}}{2j - 1}.$$

In particular, $(2j - 1)b(j)$ is a *unique* integer in this interval. Suppose the reverse holds and that it is not unique. Then there must exist a number x so that $x = (2j - 1)b(j) = (2k - 1)b(k)$ for $j \neq k$. Since $b(j)$ and $b(k)$ are multiples of two, and $2j - 1$ and $2k - 1$ are odd numbers, this would imply that either x has two unique prime factorizations (which is impossible) or that $b(j) = b(k)$ and so $j = k$. Thus, $(2j - 1)b(j)$ is a unique integer in $[\widehat{n} + 1, \dots, 2\widehat{n}]$.

To prove the second property, we note that for each i , we can consider the binary representation of $(2i - 1)b(i)$ and write

$$(2i - 1)b(i) = \sum_{k=\log_2(b(i))}^{\log_2(b(1))} 2^k c_k, \tag{A10}$$

where c_k is the k^{th} binary number. In fact, the right hand side is the addition of several $b(j)$'s, $j \geq i$. Note that the smallest value in the sum, for $k = \log_2(b(i))$, equals $b(i) = 2^k c_k$, with $c_k = 1$ since $(2i - 1)b(i)$ is clearly an odd multiple of $b(i)$.

Let $1 = j^{(1)} < \dots < j^{(\ell)} < \dots < j^{(m)}$ denote the argument of the ℓ -th local maximum, that is,

$$(2j^{(\ell)} - 1)b(j^{(\ell)}) = \max_{i \leq j^{(\ell)}} (2i - 1)b(i) = \sum_{k=\log_2(b(j^{(\ell)}))}^{\log_2(b(1))} 2^k d_k^{(\ell)}.$$

By using the first property of this lemma, we always have that for $j^{(m)}$ the maximum value equals $2\widehat{n}$. Therefore, if $j^{(m)} = 1$, the maximum value (and minimum value) equals $b(1) = 2\widehat{n}$, and the statement is trivial.

Suppose that $j^{(m)} > 1$. Starting from $j^{(1)} = 1$, where we have the binary representation of $b(1)$, to $j^{(m)}$, with binary representation of $2\widehat{n}$, we exactly know which powers of 2 need to be added to $b(1)$ in order for us to obtain $2\widehat{n}$. Since $b(\cdot)$ is a non-increasing function, each local maximum

with argument $j^{(\ell)}$ occurs with the addition of the highest power of 2 that is in $2\widehat{n}$, but not yet in $(2j^{(\ell-1)} - 1)b(j^{(\ell-1)})$. In fact, this is the addition of $b(j^{(\ell)})$, see below (A10). In particular we have $2\widehat{n} - (2j^{(\ell)} - 1)b(j^{(\ell)}) < b(j^{(\ell)})$, since after the addition of $b(j^{(\ell)})$ we can add at most $b(j^{(\ell)}) - 1$ to the sum before the sum reaches $2\widehat{n}$ due to $b(\cdot)$ being non-increasing. Moreover, $b(j) > 1$ for all $j \leq \frac{\widehat{n}+1}{2}$ and therefore, all even numbers in $[\widehat{n} + 1, \dots, 2\widehat{n}]$ occur before all odd numbers. In particular, $2\widehat{n}$ is even *and* the final local maximum, and therefore $b(j^{(m)}) > 1$.

To prove the final property, we consider $\max_{i \leq j} (2b(1) - (2i - 1)b(i))$. As $2b(1)$ is strictly larger than $2\widehat{n}$, it allows us to proceed the proof in the exact same way as the second property. This concludes the proof. \square

Proof of Theorem A.1. We prove this theorem by showing that the seating sequence as in Definition A.2, has a maximum error as stated. Recall that s_j is the j -th person in the seating sequence s and let I_l, I_r be the length of the two intervals created by placing s_j to the left and to the right of person j respectively. For this theorem to hold, it suffices to show that

$$d_{\min}^{(s)}(j) = I_l \geq \frac{1}{2j(H_{2\widehat{n}} - H_{\widehat{n}})}, \quad (\text{A11})$$

for any value of j and then use the definition of e_{\max} , see (1).

We prove Equation (A11) in two steps. First, we define a new sequence, from which we can obtain the seating sequence s by shifting all points, and show that for this new sequence (A11) holds. Trivially, this does not alter the distances between all persons.

The sequence ν we consider in the proof is

$$\nu_j = \begin{cases} \frac{2H_{2\widehat{n}} - 2H_{(2j-1)b(j)}}{2(H_{2\widehat{n}} - H_{\widehat{n}})} & \text{if } 1 \leq j \leq \widehat{n} \\ \frac{2H_{2\widehat{n}} - H_j - H_{j-1}}{2(H_{2\widehat{n}} - H_{\widehat{n}})} & \text{if } \widehat{n} < j \leq 2\widehat{n} \end{cases}.$$

The sequence s is then obtained by shifting all points such that person 1 is now seated at position 0, which we formally show at the end of the proof. To show that (A11) holds for ν , we consider two cases. First, we show that this holds in the case of $1 < j \leq \widehat{n}$, and afterwards we show this holds for the case $\widehat{n} < j \leq 2\widehat{n}$.

Case 1: $1 < j \leq \widehat{n}$

In the case of $1 < j \leq \widehat{n}$, ν_j is placed at

$$\frac{2H_{2\widehat{n}} - 2H_{(2j-1)b(j)}}{2(H_{2\widehat{n}} - H_{\widehat{n}})}. \quad (\text{A12})$$

To determine I_l and I_r , we need to know which of the seating locations ν_1, \dots, ν_{j-1} are closest to ν_j . Specifically, we are looking for the largest ν_l and smallest ν_r for which holds that $\nu_l \leq \nu_j \leq \nu_r$. These persons l and r are the closest two persons to the left and right of person j respectively.

To determine the location of ν_l and ν_r , we look at the definition of the sequence. The location of each point ν_i is completely determined by the value of $(2i - 1)b(i)$. Moreover, the higher this value, the smaller ν_i . If the value of $(2i - 1)b(i)$ is close to $(2j - 1)b(j)$, then certainly ν_i is close to ν_j . Since we consider a circle, we have to be careful when placing the largest or smallest value of $(2i - 1)b(i)$ currently present, as this person has a neighbor that has the smallest or largest value respectively of $(2i - 1)b(i)$ currently present.

Specifically, for any j such that $(2j - 1)b(j) \neq \max_{i \leq j} (2i - 1)b(i)$ or $(2j - 1)b(j) \neq \min_{i \leq j} (2i - 1)b(i)$, ν_l and ν_r with $(2l - 1)b(l) > (2j - 1)b(j) > (2r - 1)b(r)$ are the closest in value. Now we determine how close this value can get. The cases where

$(2j - 1)b(j)$ equals the maximum or minimum value currently present are the boundary cases and are considered afterwards.

Note that $b(i)$ is a monotone decreasing function, which is always a power of 2. Therefore, for all v_i placed before v_j , $(2i - 1)b(i)$ is always divisible by $b(j)$. Clearly, v_l and v_r are placed before v_j . The closest v_l and v_r could possibly get to v_j , is if l and r are such that $(2l - 1)b(l) = 2jb(j)$ and $(2r - 1)b(r) = (2j - 2)b(j)$. This is always possible.

Thus, the two closest points possible are

$$v_l = \frac{2H_{2\hat{n}} - 2H_{2jb(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})}$$

and

$$v_r = \frac{2H_{2\hat{n}} - 2H_{(2j-2)b(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})}.$$

The size of the interval I_l therefore is

$$\begin{aligned} I_l = v_j - v_l &= \frac{2H_{2jb(j)} - 2H_{(2j-1)b(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})} \\ &= \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \sum_{i=(2j-1)b(j)+1}^{2jb(j)} \frac{2}{i} \\ &\geq \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} b(j) \frac{2}{2jb(j)} = \frac{1}{2j(H_{2\hat{n}} - H_{\hat{n}})}, \end{aligned}$$

while for I_r we find

$$\begin{aligned} I_r = v_r - v_j &= \frac{2H_{(2j-1)b(j)} - 2H_{(2j-2)b(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})} \\ &= \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \sum_{i=(2j-2)b(j)+1}^{(2j-1)b(j)} \frac{2}{i} \geq \frac{1}{(2j-1)(H_{2\hat{n}} - H_{\hat{n}})}. \end{aligned}$$

Similarly, we can upper bound I_l and find that $I_l < \frac{1}{(2j-1)(H_{2\hat{n}} - H_{\hat{n}})}$ and hence $I_r > I_l$ for each j . Moreover, this also shows that after $j - 1$ persons, for $j \geq 2$, the smallest interval is at least of size $\frac{1}{2(j-1)(H_{2\hat{n}} - H_{\hat{n}})}$. This is larger than I_l can possibly get with j persons present. Thus, it follows by an inductive argument that I_l is always the smallest interval that exists.

It remains to be shown that similar bounds hold on I_l and I_r for both boundary cases. Suppose that $(2j - 1)b(j) = \max_{i \leq j} (2i - 1)b(i)$, which only changes v_l . Then it must hold that $2jb(j) > 2\hat{n}$ by the second property of Lemma A.7. Thus, there must then be an integer $l < j$ such that $(2l - 1)b(l) = jb(j)$, which corresponds to the largest position that is currently taken. Note that if $b(l) = b(j)$, it immediately follows that $l < j$ for $l > 1$, and that if $b(l) > b(j)$ this also follows by definition of the function $b(\cdot)$.

Therefore, in this case the closest point to the left is

$$v_l = \frac{2H_{2\hat{n}} - 2H_{jb(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})},$$

and so the interval length becomes

$$I_l = v_j + 1 - v_l.$$

Using (A12) and the just derived closest left point, we find that

$$\begin{aligned} I_l &= \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \sum_{i=(2j-1)b(j)+1}^{2\hat{n}} \frac{2}{i} + \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \sum_{i=\hat{n}+1}^{jb(j)} \frac{2}{i} \\ &> \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \left(\sum_{i=(2j-1)b(j)+1}^{2\hat{n}} \frac{2}{i} + \sum_{i=2\hat{n}+2}^{2jb(j)+1} \frac{2}{i} \right), \end{aligned} \quad (\text{A13})$$

where the inequality follows by lower bounding each term in the second summation using $\frac{1}{x} > \frac{1}{2x} + \frac{1}{2x+1}$. Note that we indeed have $2jb(j) > 2\hat{n} + 1$ in this boundary case since $2jb(j)$ is even. Using that $\frac{1}{x} \geq \frac{2}{x-1} - \frac{1}{x-2}$ we immediately find

$$\begin{aligned} I_l &> \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \left(\sum_{i=(2j-1)b(j)+1}^{2\hat{n}} \frac{2}{i} + \sum_{i=2\hat{n}+2}^{2jb(j)} \frac{2}{i} + 2 \left(\frac{2}{2jb(j)} - \frac{1}{2jb(j)-1} \right) \right) \\ &\geq \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \left(\sum_{i=(2j-1)b(j)+2}^{2\hat{n}} \frac{2}{i} + \sum_{i=2\hat{n}+2}^{2jb(j)} \frac{2}{i} + 2 \frac{2}{2jb(j)} \right) \\ &\geq \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} b(j) \frac{2}{2jb(j)}, \end{aligned} \quad (\text{A14})$$

where the second inequality follows using $2jb(j) - 1 \geq (2j - 1)b(j) + 1$ (from the second property of Lemma A.7 we know that in this boundary case we have $b(j) > 1$) and the final inequality follows by using $i \leq 2jb(j)$ for each i in both summations and then counting all the terms.

To show that I_l is the minimum distance after j persons are seated, we also need to upper bound I_l , as before. Using (A13) and $\frac{1}{x} < \frac{1}{2x} + \frac{1}{2x-1}$ we immediately find that

$$I_l < \frac{1}{2(H_{2\hat{n}} - H_{\hat{n}})} \left(\sum_{i=(2j-1)b(j)+1}^{2\hat{n}} \frac{2}{i} + \sum_{i=2\hat{n}+1}^{2jb(j)} \frac{2}{i} \right) < \frac{1}{(2j-1)(H_{2\hat{n}} - H_{\hat{n}})}.$$

This concludes the first boundary case.

Suppose that $(2j-1)b(j) = \min_{i \leq j} (2i-1)b(i)$, which only changes v_r . Then $(2j-2)b(j) \leq \hat{n}$ by the third property of Lemma A.7 and so there must be an integer $r < j$ such that $(2r-1)b(r) = 2(2j-2)b(j)$. Hence, $b(r) \geq 4b(j)$.

Therefore, the closest point to the right is

$$v_r = \frac{H_{2\hat{n}} - 2H_{2(2j-2)b(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})},$$

and using similar steps that led to (A14) we find

$$\begin{aligned} I_r &= 1 - v_j + v_r \\ &= 1 - \frac{2H_{2\hat{n}} - 2H_{(2j-1)b(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})} + \frac{H_{2\hat{n}} - 2H_{2(2j-2)b(j)}}{2(H_{2\hat{n}} - H_{\hat{n}})} \\ &> \frac{1}{(2j-1)(H_{2\hat{n}} - H_{\hat{n}})}. \end{aligned}$$

This concludes the first case.

Case 2: $\hat{n} < j \leq 2\hat{n}$

In the case of $\hat{n} < j \leq 2\hat{n}$, v_j is placed at

$$\frac{2H_{2\hat{n}} - H_j - H_{j-1}}{2(H_{2\hat{n}} - H_{\hat{n}})}.$$

It immediately follows that the two closest points possible are

$$v_l = \frac{2H_{2\hat{n}} - 2H_j}{2(H_{2\hat{n}} - H_{\hat{n}})}$$

and

$$v_r = \frac{2H_{2\hat{n}} - 2H_{(j-1)}}{2(H_{2\hat{n}} - H_{\hat{n}})},$$

except when $j = \hat{n} + 1$. In that case, $v_r = 1$, which is a position that is not been occupied by a person according to (A12). However, position 0 is occupied, which is the same position given that we consider a circle. It immediately follows that

$$I_r = I_l = \frac{H_j - H_{j-1}}{2(H_{2\hat{n}} - H_{\hat{n}})} = \frac{1}{2j(H_{2\hat{n}} - H_{\hat{n}})},$$

and therefore it again follows from an inductive argument that I_l is always the smallest interval that exists (even though it is not unique).

To conclude the proof, we need to show that the seating sequence s is obtained from the sequence v , by shifting the sequence. To this end, let

$$s_j = \left(v_j + \frac{2H_{b(1)} - 2H_{2\hat{n}}}{2(H_{2\hat{n}} - H_{\hat{n}})} \right) \text{ mod } 1.$$

From Lemma A.7 we have $\hat{n} < b(1) \leq 2\hat{n}$ and therefore we know that this is a shift that puts all persons “to the left”, by less than 1. Therefore, the modulo makes sure all values of the sequence are again between 0 and 1. This concludes the proof. □

A.1.3. Fixed angle strategy

Proof of Theorem 2.7. The proof goes along similar lines as the proof of Theorem 2.4. To determine $d_{\min}(n)$, we note that by construction of the continued fraction expansion of φ , for the k -th term we either have $\left\lfloor \frac{F_{k+1}}{\varphi} \right\rfloor = F_k$ or $\left\lceil \frac{F_{k+1}}{\varphi} \right\rceil = F_k$, where F_k is the k -th Fibonacci number. Additionally, the continued fraction expansion is known to be the best local approximation to an irrational number. That is, given any q_k , there exists a unique number p_k so that $\frac{p_k}{q_k}$ approximates φ . It thus follows from (4) that, for any $F_{k+1} < n \leq F_{k+2}$,

$$\begin{aligned} d_{\min}(n) &= \min_{1 \leq j \leq n-1} d(1, j+1) \\ &= d(1, F_{k+1} + 1) = \left| \frac{F_{k+1}}{\varphi} - F_k \right|. \end{aligned}$$

For the k -th Fibonacci number it is known that

$$(1 - \varphi)^k = F_{k+1} - \varphi F_k \quad \text{and} \quad (1 - \varphi)^k = \left(\frac{-1}{\varphi} \right)^k,$$

so that we find that $d_{\min}(n) = \frac{1}{\varphi^{k+1}}$ for $F_{k+1} < n \leq F_{k+2}$.

Recall that we required $F_{l+1} < n_{\max} \leq F_{l+2}$. By definition in Equation (2) we have

$$\begin{aligned} \mathbb{E}[e_{\max}(N)] &= 1 - \frac{1}{n_{\max} - 1} \left(\sum_{k=1}^{l-1} \frac{1}{\varphi^{k+1}} \sum_{n=F_{k+1}+1}^{F_{k+2}} n + \frac{1}{\varphi^{l+1}} \sum_{n=F_{l+1}+1}^{n_{\max}} n \right) \\ &= 1 - \frac{1}{n_{\max} - 1} \left(\sum_{k=1}^{l-1} \frac{1}{\varphi^{k+1}} \frac{F_{k+3} + 1}{2} (F_{k+2} - F_{k+1}) \right. \\ &\quad \left. + \frac{1}{\varphi^{l+1}} \frac{n_{\max} + F_{l+1} + 1}{2} (n_{\max} - F_{l+1}) \right) \\ &< 1 - \frac{1}{n_{\max} - 1} \left(\frac{\varphi^2}{2\sqrt{5}} \sum_{k=1}^{l-1} (F_{k+2} - F_{k+1}) \right. \\ &\quad \left. + \frac{1}{\varphi^{l+1}} \frac{n_{\max}(n_{\max} + 1) - F_{l+1}(F_{l+1} + 1)}{2} \right) \\ &= 1 - \frac{1}{n_{\max} - 1} \left(\frac{\varphi^2}{2\sqrt{5}} (F_{l+1} - 1) \right. \\ &\quad \left. + \frac{1}{\varphi^{l+1}} \frac{n_{\max}(n_{\max} + 1) - F_{l+1}(F_{l+1} + 1)}{2} \right), \end{aligned} \tag{A15}$$

where the first equality follows by definition of $d_{\min}(n)$, the inequality by noting that $\frac{F_{k+1}}{\varphi^{k+1}} > \frac{1}{\sqrt{5}}$ and the final equality follows due to cancelling terms. We find that the maximum of the upper bound on $\mathbb{E}[e_{\max}(N)]$ is attained whenever

$$(n_{\max}^* - 1)^2 = 2 + \frac{\varphi^2}{\sqrt{5}} (F_{l+1} - 1) \varphi^{l+1} - F_{l+1}(F_{l+1} + 1),$$

where we again note that this value is not necessarily integer (which does not matter for the upper bound to hold) and that it is possible to verify that indeed $n_{\max}^* \in (F_{l+1}, F_{l+2}]$. Thus, we find

$$\begin{aligned} \mathbb{E}[e_{\max}(N)] &< 1 - \frac{1}{n_{\max}^* - 1} \left[\frac{(n_{\max}^* - 1)^2 - 2}{2\varphi^{l+1}} + \frac{n_{\max}(n_{\max} + 1)}{2\varphi^{l+1}} \right] \\ &= 1 - \frac{2(n_{\max}^* - 1) + 3}{2\varphi^{l+1}}. \end{aligned}$$

If we show that the optimum value of n_{\max}^* satisfies

$$(n_{\max}^* - 1) \geq \sqrt{\varphi} \frac{\varphi^{l+1}}{\sqrt{5}} - c,$$

for some $0 \leq c \leq \frac{3}{2}$, we have shown the first part of the statement. We claim that this is true for

$$c^2 = 2 + \frac{1}{5\varphi}. \tag{A16}$$

We indeed have $c^2 \leq \frac{9}{4}$, as required.

We need to show that

$$2 + \frac{\varphi^2}{\sqrt{5}} (F_{l+1} - 1)\varphi^{l+1} - F_{l+1}(F_{l+1} + 1) \geq \left(\sqrt{\varphi} \frac{\varphi^{l+1}}{\sqrt{5}} - c \right)^2,$$

for c as in (A16). Using the definition of the Fibonacci number, $F_k = \frac{\varphi^k - (1-\varphi)^k}{\sqrt{5}}$, we find the left-hand side to be the same as

$$\begin{aligned} &2 + \frac{\varphi^2}{\sqrt{5}} \left(\frac{\varphi^{l+1} - (1-\varphi)^{l+1}}{\sqrt{5}} - 1 \right) \varphi^{l+1} - \left(\frac{\varphi^{l+1} - (1-\varphi)^{l+1}}{\sqrt{5}} \right) \left(\frac{\varphi^{l+1} - (1-\varphi)^{l+1}}{\sqrt{5}} + 1 \right) \\ &= (\varphi^2 - 1) \frac{\varphi^{2(l+1)}}{5} - (\varphi^2 + 1) \frac{\varphi^{l+1}}{\sqrt{5}} + \frac{(-1)^{l+2}}{5\varphi} - \frac{(1-\varphi)^{2(l+1)}}{5} + \frac{(1-\varphi)^{l+1}}{\sqrt{5}} + 2. \end{aligned}$$

Thus, since by definition $\varphi^2 = \varphi + 1$, it remains to show that, by combining the two displays above,

$$(2c\sqrt{\varphi} - (\varphi + 2)) \frac{\varphi^{l+1}}{\sqrt{5}} + \frac{(-1)^{l+2}}{5\varphi} + \frac{(1-\varphi)^{l+1}}{\sqrt{5}} \left(1 - \frac{(1-\varphi)^{l+1}}{\sqrt{5}} \right) + 2 - c^2 \geq 0, \tag{A17}$$

for c as in (A16). Now, we note that we only have to consider $l \geq 3$, since for all $n_{\max} \leq F_4$ the first case of the statement does not apply (as $F_4 = 3$). For all even l , if the equation is satisfied for $l = 2$ in the third term, it is certainly satisfied for all larger l in that term. Thus, we only need to show that

$$(2c\sqrt{\varphi} - (\varphi + 2)) \frac{\varphi^{l+1}}{\sqrt{5}} \geq -\frac{(1-\varphi)^3}{\sqrt{5}} \left(1 - \frac{(1-\varphi)^3}{\sqrt{5}} \right), \tag{A18}$$

by definition of c . We do not yet prove this condition, as the condition for odd l turns out to be stricter than this one.

Hence, consider l odd (implying it is larger than or equal to 3). Then, using (A16), this condition is equivalent to showing that

$$(2c\sqrt{\varphi} - (\varphi + 2)) \frac{\varphi^4}{\sqrt{5}} \geq \frac{2}{5\varphi}, \quad (\text{A19})$$

since for all odd l the other terms are positive. In particular, we use the value $l = 3$, since any larger value of l only increases the left-hand side (if it is indeed positive).

If we rewrite this equation and then square both sides, we find that we need to show whether

$$\begin{aligned} 2 + \frac{1}{5\varphi} &\geq \left(\frac{\varphi + 2}{2\sqrt{\varphi}} + \frac{1}{\sqrt{5\varphi^5}} \right)^2 \\ &= \frac{5\varphi + 5}{4\varphi} + \frac{(\varphi + 2)}{\sqrt{5}\varphi^6} + \frac{1}{5\varphi^{11}}. \end{aligned}$$

Rewriting the first term of the right-hand side as $2 + \frac{5-3\varphi}{4\varphi}$, we need to determine if

$$\varphi^{10} \geq (5 - 3\varphi) \frac{5}{4} \varphi^{10} + \sqrt{5}\varphi^6 + 2\sqrt{5}\varphi^5 + 1,$$

and by using that $\varphi^j = F_j\varphi + F_{j-1}$, we again rewrite this to

$$\varphi(45 - 18\sqrt{5}) \geq -\frac{111}{4} + 11\sqrt{5} + 1.$$

Since the left-hand side is positive and the right-hand side is negative, (A19) holds.

Now we still need to show that (A18) also holds. Since for even l we had $l \geq 4$, we find that the left-hand side of (A18) is larger than the left-hand side of (A19). For the right-hand side we now show the reverse: the right-hand side of (A19) is larger than the right-hand side of (A18). That is, we have the following equivalent statements:

$$\begin{aligned} \frac{2}{5\varphi} &\geq -\frac{(1-\varphi)^3}{\sqrt{5}} \left(1 - \frac{(1-\varphi)^3}{\sqrt{5}} \right), \\ 2 &\geq (5 - \sqrt{5})\varphi + 2\sqrt{5} - 8, \\ \frac{10}{4} &\geq \sqrt{5}, \end{aligned}$$

where the first rewriting uses $(1-\varphi)^j = F_j(1-\varphi) + F_{j-1}$. This proves the first part of the statement.

To show the second statement, we note that for $n_{\max} = F_{l+1}$ Equation (A15) immediately results in the second part of the statement. \square

A.1.4. Random

Proof of Theorem 2.8. Consider placing n people uniformly at random around the table. We can fix one person at an arbitrary point at the table, let this point be 0. This means we still have to place $n - 1$ others. Each of these will be placed at some point $X_i \sim U[0, 1]$, $i = 1, \dots, n - 1$. Suppose for some realization x_1, \dots, x_{n-1} of X_1, \dots, X_{n-1} we find $x_1 \leq \dots \leq x_{n-1}$. This means that the minimum distance between two persons is

- x_1 iff $x_1 \leq x_{j+1} - x_j$ for all $j = 1, \dots, n - 2$ and $x_1 \leq 1 - x_{n-1}$;
- $x_{j+1} - x_j$ for some $j = 1, \dots, n - 2$ iff $x_{j+1} - x_j \leq x_{k+1} - x_k$ for all $k \neq j$ and $x_{j+1} - x_j < x_1$ and $x_{j+1} - x_j \leq 1 - x_{n-1}$;
- $1 - x_{n-1}$ iff $1 - x_{n-1} \leq x_{j+1} - x_j$ for all $j = 1, \dots, n - 2$ and $1 - x_{n-1} \leq x_1$.

Without loss of generality we can assume that $1 - x_{n-1}$ takes the minimum value. Remark that there are $(n - 1)!$ possible orderings of the random variable and each of these has n possible minimum distances.

This means that we have

$$\mathbb{E}[D_{\min}(n)] = n! \int_{\frac{n-1}{n}}^1 \int_{(n-2)(1-x_{n-1})}^{2x_{n-1}-1} \int_{(n-3)(1-x_{n-1})}^{x_{n-2}+x_{n-1}-1} \dots \int_{1-x_{n-1}}^{x_2+x_{n-1}-1} (1-x_{n-1}) dx_1 \dots dx_{n-2} dx_{n-1},$$

where most upper and lower bounds of the integrals follow by using that $1 - x_{n-1} \leq x_{j+1} - x_j$ for all $j = 1, \dots, n - 2$. The lower bound for x_1 and both bounds for x_{n-1} follow by definition. As a result, the lower bound of each integral is indeed smaller than its upper bound.

The first integral is trivial, and the second integral until the $(n - 2)^{th}$ follow by using the change of variable $y_j = x_j - j(1 - x_{n-1})$ for $j = 2, \dots, n - 2$. We then find

$$\begin{aligned} \mathbb{E}[D_{\min}(n)] &= n! \int_{\frac{n-1}{n}}^1 (1-x_{n-1}) \frac{1}{(n-2)!} (nx_{n-1} - (n-1))^{n-2} dx_{n-1} \\ &= n! \int_{\frac{n-1}{n}}^1 \frac{1}{n!} (nx_{n-1} - (n-1))^{n-1} dx_{n-1} = \frac{1}{n^2}, \end{aligned}$$

where the second equality follows by partial integration. The result follows using (1). □

A.2. Experimental details

Our study complies with all relevant ethical regulations. The study was approved by the IRB at the Technical University of Delft, HREC, under ID numbers 1766, 1892, and 2465 and informed consent was obtained from all participants.

Participants. We recruited 31 cadets of the Netherlands Defence Academy (NLDA). Participation in the experiment was not for credit, though we offered some snacks as reward.

Procedures. The experiment was conducted on April 26, 2022 between 13:00h and 17:00h CEST in the NLDA sports hall. The hall has three rooms. Each room had a large circle drawn on the floor. The angle of the circle was indicated on the floor expressed in degrees, at every 10 degree-mark, so 0 degrees, 10 degrees, 20 degrees, and so on until 350 degrees. A week prior to the experiment, cadets received information on experiment location and times, as well as instructions on the experimental task. At the start of the experiment these instructions were given again. Cadets were then asked three questions that verified their understanding of the experiment. If they got these questions wrong, we then explained the questions until they understood. The experiment consisted of 10 rounds. Each round consisted of 5 trials. Each round the 31 cadets were split into three groups of at most 11 cadets. This was done in a way that minimized overlap between groups. Each of the three groups occupied one of the three rooms of the sports hall. This process yielded 50 trials per group times three groups is 150 trials and 803 individual location decisions. The 803 individual location decisions exclude the first decision of each trial, which is trivial.

The positioning task. Each trial a specific number of cadets was asked to position themselves. This was a number between 2 and 11, each with probability 1/10. Cadets knew about these probabilities but they learned what the number was for a given trial only at the end of that trial. Cadets were first asked to stand in line, not being able to see the circle, see [Figure A1](#) for a

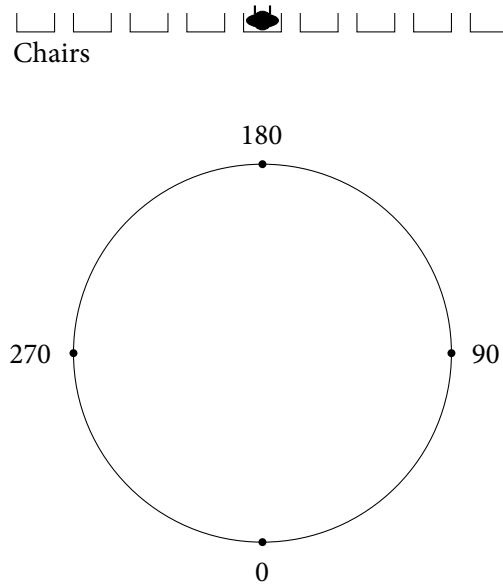


Figure A1. A graphical illustration of the experiment. The cadets were waiting until being called at the chairs, not seeing the experiment. The degrees on the circle show where the degrees were located compared to where the cadets waited.

graphical illustration. Cadet numbers would be called one at a time. The cadet whose number was called had to position themselves at a location of their choice on the circle. They were instructed to stay put and not move once they had positioned themselves and the next cadet number was called. They were instructed to position themselves strategically with the aim of maximizing the ultimate distance between themselves and other participants at the end of the trial, not knowing how many cadets would be asked this trial.

Low numbers for 11 and other data discrepancies. Upon the preparation of the experiment, we knew there could be at most 37 cadets and so we prepared for 33 to 37 cadets to join the experiment. At the day of the experiment, there were a total number of 31 cadets, from which one attended only early afternoon and one only in the late afternoon (with a small gap in between them). Therefore, at all times we had at least 29 cadets joining and at most 30 cadets.

Therefore, at the beginning of the experiment we decided to use the setup for 33 cadets. For each room, there would be a list of 11 cadet numbers, so whenever a cadet number was not there, we would simply call the next cadet number. We continued until there were as much cadets seated as required by the list, or, until there were no more cadets. This made the probability distribution of cadets in [Figure 4](#) not uniform. On the other hand, cadets were aware of their group size and hence knew the maximum number possible.

This explains 17 deviations from the original uniform distribution that was set up for 33 cadets. In 13 cases, the number of cadets entered should have been 11, and was 10 (9 times) and 9 (4 times) instead. In three cases this made the number of cadets entered going from 10 to 9, and once it went down from 9 to 8 (all “missing” cadet numbers were scheduled at the same group). The final two deviations were manual errors, but can be explained since the manual notes show that it was forgotten to call the next cadet number whenever a cadet number was not there. Due to these errors, it occurred once that the number of cadets entered went from 5 to 4 and from 8 to 7.

Determining the order of which cadets seat around the table. The order in which cadets are invited to seat around the table is random, though we added some conditions. In each round, consisting of 5 trials, a cadet would never be asked to be seated as k -th person twice. In addition, we limited the amount of times that the same cadet would be asked to be seated as k -th person as follows. Every 6 trials we increased the amount of times a cadet is allowed to be seated as k -th person by one. The number 6 is arbitrary, though the higher the number, the more difficult it becomes to find an order. We did not use these conditions on student numbers of which we knew they would not be invited to be seated. Since in the end we had fewer students than expected, these conditions may not always have been satisfied. Every number of cadets that would be invited to be seated was used once every 10 trials.

A.3. Experimental results

In this section, we show some additional experimental results.

Table A1. The probability that the strategy mentioned is strictly worse than what humans achieved. The p-value with superscript 1 is the result from a Wilcoxon test in which we define our null hypothesis H_0 as “Humans perform better than the strategy mentioned” and hence our alternative hypothesis H_1 is “The strategy performs better than human”. Vice versa, the p-value with superscript 2 is the result from a Wilcoxon test with H_0 “The strategy performs better than human” and H_1 “Human perform better than the strategy mentioned”. All values are rounded up to 3 decimals and when smaller than computer epsilon they are denoted as 0.

Strategy	Probability	p-value ¹	p-value ²
Split-2	0.385	2.104e-07	1.000
Fixed error finite	0.297	4.077e-07	1.000
Fixed error	0.345	1.539e-05	1.000
Split-3	0.581	0.628	0.372
Fixed angle	0.696	1.000	2.720e-09
Random	0.993	1.000	0.000

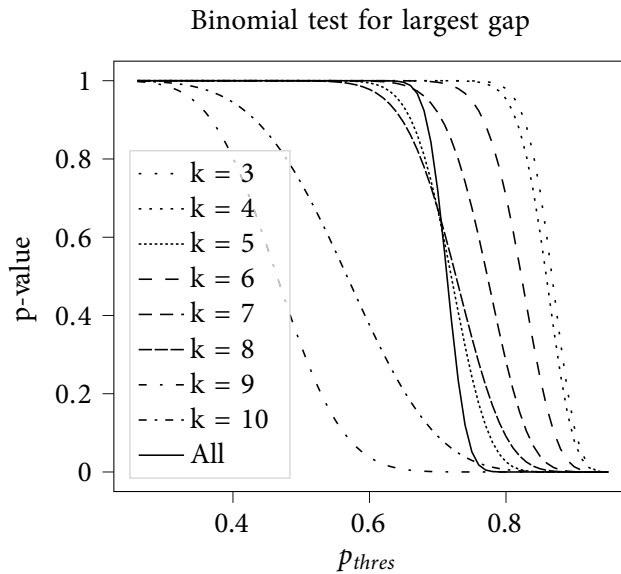


Figure A2. The fraction of persons, that are number k to enter the room, that are seated in (one of) the largest gap(s). We perform a binomial test where we define H_0 as “The probability that someone seats in the largest gap is at least p_{thres} ” and H_1 as the reverse. Note that the first and second person always sits in the largest gap, so those are left out. $k = 11$ is only included in “All”.

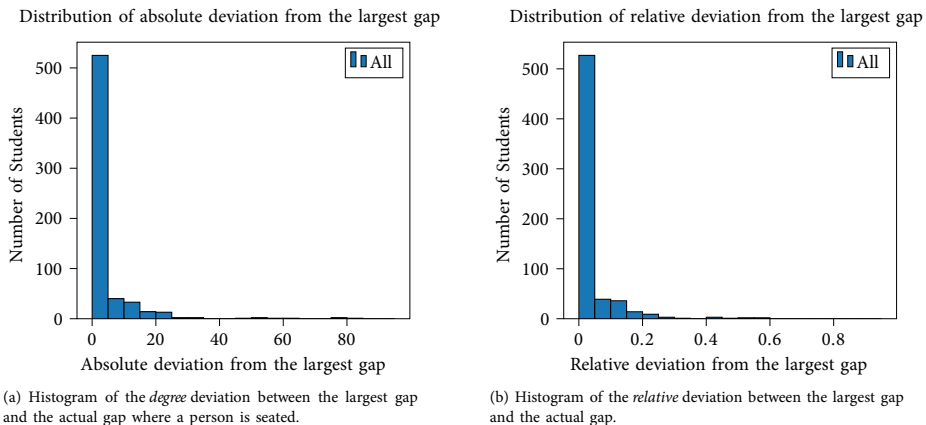


Figure A3. Two histograms showing the distribution of the deviation from the largest gap. A deviation of 0 implies that the cadets are seated in the largest gap.

A.4. Evidence of learning in subject behavior

We explored the possibility of learning in the dynamics of subject behavior. Figure A4 shows for each trial the gap between the first and second subject. There are three data points per trial number because there were three rooms in which trials were conducted simultaneously. In each room 50 trials were conducted. It can be observed in the figure that gaps far away from the two central tendencies of 180 and 120 degrees occur only during the first 13 trials. This is suggestive evidence that behavior became more routine-like and settled on two basic strategies. We thank an anonymous reviewer for suggesting this analysis.

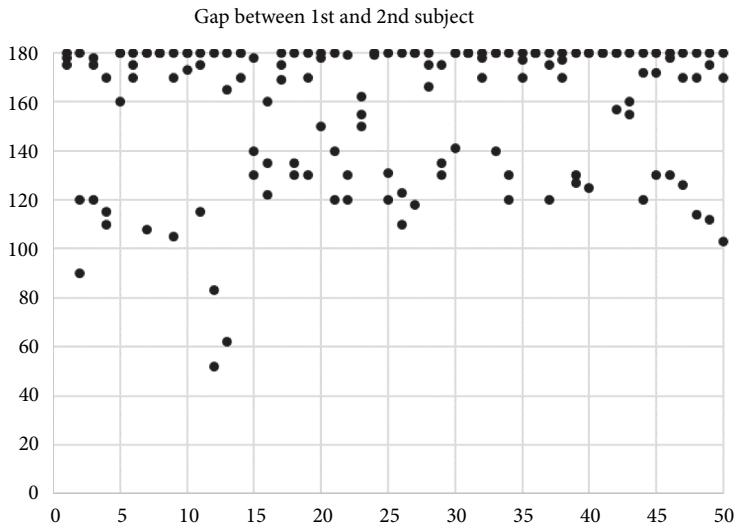


Figure A4. Gap between the first and second experimental subject in each of 50 trials conducted in 3 rooms.