

# Misinformed leaders lose influence over pigeon flocks

Isobel Watts<sup>a\*</sup>, Máté Nagy<sup>a,b,c</sup>, Theresa Burt de Perera<sup>a</sup> and Dora Biro<sup>a\*</sup>

<sup>a</sup>*Department of Zoology, University of Oxford, Oxford, United Kingdom.*

<sup>b</sup>*Department of Collective Behaviour, Max Planck Institute for Ornithology & Department of Biology, University of Konstanz, Konstanz, Germany.*

<sup>c</sup>*MTA-ELTE Statistical and Biological Physics Research Group, Hungarian Academy of Sciences, Budapest, Hungary.*

\*Corresponding authors' email: [isobel.watts@zoo.ox.ac.uk](mailto:isobel.watts@zoo.ox.ac.uk), [dora.biro@zoo.ox.ac.uk](mailto:dora.biro@zoo.ox.ac.uk)

## Abstract

In animal groups where certain individuals have disproportionate influence over collective decisions, the whole group's performance may suffer if these individuals possess inaccurate information. Whether in such situations leaders can be replaced in their roles by better informed group mates represents an important question in understanding the adaptive consequences of collective decision-making. Here, we use a clock-shifting procedure to predictably manipulate the directional error in navigational information possessed by established leaders within hierarchically structured flocks of homing pigeons (*Columba livia*). We demonstrate that in the majority of cases when leaders hold inaccurate information they lose their influence over the flock. In these cases, incorrect information is filtered out through the rearrangement of hierarchical positions, preventing errors by former leaders from propagating down the hierarchy. Our study demonstrates that flexible decision-making structures can be valuable in situations where 'bad' information is introduced by otherwise influential individuals.

**Keywords: homing pigeon, hierarchy, leadership, clock-shift**

28

29

## 30 **Introduction**

31

32 Animal groups faced with making joint decisions can exploit variation in the quality of  
33 members' personal information by sharing decision-making: when navigating, for  
34 example, they can pool their personal knowledge to reduce overall navigational error  
35 (e.g. [1]). However, in some groups, decisions are not shared entirely  
36 "democratically", meaning that individuals do not contribute equally to decisions [2,  
37 3]. The group's performance then becomes disproportionately dependent on leaders'  
38 information quality. Because leadership does not necessarily correlate with  
39 competence [2, 4] the question arises whether (i) followers and/or leaders are  
40 sensitive to the quality of leaders' input, and, if so, (ii) whether groups have any  
41 scope for "overruling" inaccurate leadership.

42

43 Homing pigeon flocks form transitive leadership hierarchies where some individuals  
44 consistently contribute more to directional decisions than others [3], although a  
45 degree of decision-sharing is also evident among members [5, 6]. These hierarchies  
46 are stable across time [7], and important factors structuring them include individual  
47 differences in, e.g., navigational experience [8, 9] and speed [4]. Recent modelling  
48 work has shown that multi-level hierarchies can compensate for an increase in  
49 navigational error better than random networks [10], however, this advantage  
50 disappears when the most influential individuals have the highest error, as errors  
51 then propagate down the hierarchy. In such situations, hierarchical structuring could  
52 be detrimental.

53

54 Here, we examined whether this model prediction holds true in real homing pigeon  
55 flocks, by experimentally increasing the navigational error of an identified leader. We

manipulated leaders' personal information through "clock-shifting": a procedure known to systematically interfere with pigeons' use of the sun-compass for directional guidance, where clock-shifted ("misinformed") birds fly a predictably deviated route home [11]. We asked whether stable leadership results in such errors propagating down the hierarchy (i.e., the whole flock flies the incorrect, shifted route) or if some compensatory mechanism allows flocks to maintain a correct (non-shifted) route.

## **Methods**

We assigned 40 homing pigeons (age=2-5 years) to eight flocks of five birds, ensuring a comparable age distribution across flocks. During all flights, positional data were logged using 5Hz GPS loggers (Qstarz BT-Q1300ST) attached to the birds' backs via Velcro strips.

We trained flocks from our chosen site (Bladon, 51°49'23.48"N 1°21'26.29"W; distance and direction from home: 5.27km, 149.5°) through eight consecutive releases ("Stage 0"; figure 1a). We calculated leadership ranks for each member in each flock based on spatial positioning (with leaders nearer the front and followers nearer the back) and confirmed these through directional correlation delay analysis (figure S1; see [3] for details of both methods). We designated as "Stage-0 leaders" birds with the highest average rank over the last four training releases.

After training we conducted six experimental stages. We performed four clock-shifts (Stages 1, 2, 4 and 6; figure 1a), during which selected birds were placed in light-tight chambers until their internal clocks readjusted to an artificially shifted day-night cycle. All shifts corresponded to either an anticlockwise (fast) or clockwise (slow) 70°

shift in the sun's azimuth on the dates of release. Birds were then released in their original flocks. Stages 3 and 5 were control flights, not involving clock-shift.

GPS tracks were analysed in Matlab (Mathworks 2012b) and R (0.98.1014). We explored the effect of misinformed (i.e. clock-shifted) leaders and flocks by assessing, post-clock-shift (i) changes to the leadership hierarchy and (ii) deviations in the route flown. See electronic supplementary material for detailed Methods. All data were made available on the Dryad Digital Repository [12].

## Results

Birds identified as Stage-0 leaders occupied hierarchical positions where they had high influence (rank 1 or 2) on flocks' directional decisions in the majority of Stage 0 flights (figure S2). Although ranks initially showed fluctuations, as Stage 0 progressed these leaders' positions exhibited increasing stability, particularly in the final two flights (figures S3, S4). Following clock-shifting of Stage-0 leaders (Stages 1&2), most of these birds' average time delay values (reflecting whether and how soon their directional changes were copied by others) decreased (figure 2a) meaning that they were located significantly less often, than during Stage 0, in positions of high influence both within (figure 1b,c) and averaged across flights (Generalized Linear Mixed Model (GLMM) family binomial with *flock* as random factor a) Rank 1:  $Z=-2.75$ , effect=-2.12,  $P=0.006$ , b) Rank 1 or 2:  $Z=-3.76$ , effect=-3.05,  $P<0.001$ ; figures S2,S3,S4). This was accompanied by a significant, 2-place decrease in Stage-0 leaders' median rank (with only three of the eight Stage-0 leaders remaining top-ranked during the first clock-shifted flight, and none during the second), and while their ranks increased again during Stage 3, they did not recover fully (figure 2b). Other birds in the flock also showed changes in rank, but in none of them were these significantly in a consistent direction (figure S5). During the two flock clock-shift releases (Stages 4&6), flocks frequently split up, leaving too small a sample size to

110 assess changes in leadership hierarchies statistically. Out of eight flights in which the  
111 Stage-0 leader did not split and where at least three birds remained in the flock, on  
112 three occasions the Stage-0 leader assumed leadership.

113  
114 To detect clock-shift deviation in tracks, we used “virtual vanishing bearings”,  
115 calculated as the subjects’ heading at given points with respect to the release point  
116 [13]. For each 500m concentric boundary centred around the release site, we  
117 calculated the flock medians of the difference in each individual’s heading between  
118 their last training release (non-shifted control) and test release for Stages 1, 2, 4 and  
119 6 (figure 2c,d). A deviation of zero indicates that the track is, at the given distance  
120 from release, identical to the training track, i.e., displays no effect of clock-shift.

121  
122 When all birds in the flock were clock-shifted (Stages 4&6), virtual vanishing bearings  
123 were significantly different from the null expectation (table 1), in the expected  
124 direction of deviation for compass control (i.e. anticlockwise for the Stage 4 fast-shift  
125 and clockwise for the Stage 6 slow-shift). However, when only the Stage-0 leader  
126 was clock-shifted (Stages 1&2) flock virtual vanishing bearing deviations were not  
127 significantly different from the null expectation. This was also true for the Stage 3 and  
128 5 controls (table 1).

## 129 130 **Discussion**

131  
132 Previous theoretical work predicted that in hierarchically structured decision-making  
133 errors by leaders propagate downwards, resulting in inaccurate collective decisions  
134 [10]. By introducing incorrect navigational information of a specific magnitude at the  
135 top of the hierarchy, we found this disadvantage could be overcome in pigeon flocks:  
136 our results showed that when only leaders were misinformed flocks retained their

existing routes, whereas when entire flocks were shifted they displayed deviated routes (albeit with a smaller-than-predicted deviation, as is common in birds familiar with the landscape, [14]; although see [15]). Thus we can infer that clock-shifting was successful, but that leaders alone were not able to “mislead” their flocks on erroneous routes.

Importantly, we also documented a drop in leaders’ hierarchical ranks (i.e. in the majority of flocks their input into the flock’s navigational decisions diminished) when they alone were clock-shifted. In considering this result, however, it is worth noting that Stage-0 leaders were not entirely stable in their leadership during training (figures S3,S4). Nonetheless, stability was greatest in releases 7-8, where all Stage-0 leaders ranked either in the top two (release 7) or the top (release 8) hierarchical positions. Thus the comparison with Stages 1&2 is clearest when considering the latter part of Stage 0. That leadership stability was gradually established is also supported by the observation that Stage 0 routes showed a clear learning curve, asymptoting also around releases 7-8 (figure S7). We therefore suggest that although there is noise in the system in the form of (i) flocks gradually settling on both leaders and routes and (ii) leaders varying in the extent to which they drop in rank, the combination of overall patterns within the leadership hierarchical data and analyses of route structures pre- and post-clock-shifting provide a sufficient (albeit noisy) signal for our conclusions.

We hypothesize that where a decrease was observed in leaders’ hierarchical rank, this could have been due to two non-mutually exclusive mechanisms. First, clock-shifting may have caused leaders to become uncertain in the quality of their own information. Clock-shifting places the sun-compass, an important navigational cue, in conflict with all other directional cues (e.g., visual, magnetic) in the bird’s environment. This conflict may have prompted leaders to place less weight on their

personal information and more on social information (i.e. the copying of flockmates).  
Uncertainty may also have reduced the flight speed of the clock-shifted leader, and  
since speed is associated with leadership in pigeons [4], slower flight may result in  
birds dropping down the hierarchy (although see figure S6, showing that loss of  
leadership cannot be explained purely by changes in speed). This mechanism  
requires no recognition by followers that their leader has incorrect information.  
Alternatively, flock members may have actively 'filtered out' the low quality  
information, by reducing their reliance on social information received from leaders.  
This could have been due to recognising the increased conflict between their and the  
leader's directional preference, or to detecting a cue (e.g. reduced speed) indicating  
uncertainty in the leader. Thus, the latter mechanism corresponds to followers  
'choosing' not to follow, and the former to leaders 'choosing' not to lead. At present  
we cannot distinguish between these alternatives.

Our study demonstrated that flexible decision-making structures can be valuable in  
situations where information with high error may be introduced by otherwise  
influential individuals. Our results have implications for both theoretical and empirical  
studies of collective motion and navigation, and highlight the importance of  
considering the effects of information quality and individual certainty in shaping inter-  
individual interactions during collective actions.

**Ethics:** Protocols used in this paper were approved by the Local Ethical Review Committee of Oxford University's Department of Zoology, reference number APA/1/5/ZOO/NASPA/Biro, 2013.

**Data Accessibility:** Data are available via Dryad (doi:10.5061/dryad.508j2).

**Competing interests:** We declare no competing interests.

**Authors' contributions:** IW, TBP and DB designed the study, IW collected the data, IW and MN analysed the data, and all authors contributed to writing the manuscript. All authors agree to be held accountable for the content therein and approve the final version of the manuscript.

**Acknowledgments:** We thank Oliver Padget and Benjamin Pettit for help with analyses, Tim Guilford for discussions, Lucy Larkman and Phil Smith for technical support, and four anonymous referees for helpful comments on the manuscript.

**Funding:** I.W. was supported by the BBSRC Doctoral Training Programme, DB by a Royal Society University Research Fellowship, MN by the Royal Society Newton Alumni scheme.

## References

- [1] Simons, A.M. 2004 Many wrongs: the advantage of group navigation. *Trends Ecol Evol* **19**, 453-455. (doi:10.1016/J.Tree.2004.07.001).
- [2] King, A.J., Douglas, C.M.S., Huchard, E., Isaac, N.J.B. & Cowlshaw, G. 2008 Dominance and Affiliation Mediate Despotism in a Social Primate. *Curr Biol* **18**, 1833-1838. (doi:10.1016/J.Cub.2008.10.048).



214 [3] Nagy, M., Akos, Z., Biro, D. & Vicsek, T. 2010 Hierarchical group dynamics in  
 215 pigeon flocks. *Nature* **464**, 890-U899. (doi:10.1038/Nature08891).

216 [4] Pettit, B., Akos, Z., Vicsek, T. & Biro, D. 2015 Speed Determines Leadership and  
 217 Leadership Determines Learning during Pigeon Flocking. *Curr Biol* **25**, 3132-3137.  
 218 (doi:10.1016/j.cub.2015.10.044).

219 [5] Jorge, P.E. & Marques, P.A.M. 2012 Decision-making in pigeon flocks: a  
 220 democratic view of leadership. *J Exp Biol* **215**, 2414-2417. (doi:Doi  
 221 10.1242/Jeb.070375).

222 [6] Xu, X.K., Kattas, G.D. & Small, M. 2012 Reciprocal relationships in collective  
 223 flights of homing pigeons. *Phys Rev E* **85**. (doi:Artn 026120  
 224 Doi 10.1103/Physreve.85.026120).

225 [7] Santos, C.D., Neupert, S., Lipp, H.P., Wikelski, M. & Dechmann, D.K.N. 2014  
 226 Temporal and Contextual Consistency of Leadership in Homing Pigeon Flocks. *Plos*  
 227 *One* **9**. (doi:10.1371/journal.pone.0102771).

228 [8] Watts, I., Pettit, B., Nagy, M., Burt de Perera, T. & Biro, D. 2016 Lack of  
 229 experience-based stratification in homing pigeon leadership hierarchies. *Royal*  
 230 *Society Open Science*. (doi:10.1098/rsos.150518).

231 [9] Flack, A., Pettit, B., Freeman, R., Guilford, T. & Biro, D. 2012 What are leaders  
 232 made of? The role of individual experience in determining leader-follower relations in  
 233 homing pigeons. *Anim Behav* **83**, 703-709. (doi:10.1016/J.Anbehav.2011.12.018).

234 [10] Flack, A., Biro, D., Guilford, T. & Freeman, R. 2015 Modelling group navigation:  
 235 transitive social structures improve navigational performance. *J R Soc Interface* **12**.  
 236 (doi:10.1098/Rsif.2015.0213).

237 [11] Schmidt-Koenig, K. 1958 Experimentelle Einflußnahme auf die 24-Stunden-  
 238 Periodik bei Brieftauben und deren Auswirkungen unter besonderer  
 239 Berücksichtigung des Heimfindevermögens. *Z. Tierpsychol.* **15**, 301-331.

- 240 [12] Watts, I., Nagy, M., Burt de Perera, T. & Biro, D. 2016 Data from: Misinformed  
241 leaders lose influence over pigeon flocks. *Dryad Digital Repository*  
242 (doi:<http://dx.doi.org/10.5061/dryad.508j2>).
- 243 [13] Biro, D., Guilford, T., Dell'Omo, G. & Lipp, H.P. 2002 How the viewing of familiar  
244 landscapes prior to release allows pigeons to home faster: evidence from GPS  
245 tracking. *J Exp Biol* **205**, 3833-3844.
- 246 [14] Wallraff, H.G., Chappel, J. & Guilford, T. 1999 The roles of the sun and the  
247 landscape in pigeon homing. *J Exp Biol* **202**, 2121-2126.
- 248 [15] Wiltschko, R., Stapput, K. & Siegmund, B. 2005 Does familiarity with the release  
249 site reduce the deflection induced by clock-shifting? A comment to the paper by  
250 Gagliardo et al. (2005). *J Exp Biol* **208**, 4001-4002. (doi:10.1242/jeb.01850).

251

252

253

## Tables

Stage comparison	Estimate	z value	Bonferroni corrected p-value
Stage 1 - null	-3.40	-0.929	1
Stage 2 - null	-1.63	-0.446	1
Stage 4 - null	-13.5	-3.68	0.0024
Stage 6 - null	24.6	6.71	<0.001
Stage 3 release 1 - null	1.17	0.320	1
Stage 3 release 2 - null	-4.05	-1.10	1
Stage 5 - null	1.88	0.364	1
Stage 4- Stage 3 release 1	-16.8	-4.59	<0.0001
Stage 4- Stage 3 release 2	-17.3	-4.73	<0.0001
Stage 6- Stage 5	22.7	5.01	<0.0001

**Table 1. Comparisons between the median flock virtual vanishing bearings for each Stage and their corresponding null expectation.** Null expectation was calculated as a track with a mean deviation of 0, and a standard deviation exactly equal to that of each flock for the different Stages. We compared a Linear Mixed Model (LMM) with *stage*, and *distance* as fixed effects and *flock* as a random effect to a LMM without *stage* (Maximum Likelihood test:  $\text{Chisq}=130.8$ ,  $P<0.001$ ), then used post-hoc Tukey tests for pairwise comparisons between combinations of interest.

## Figures

**Figure 1. Experimental design and leadership analysis.** (a) Release protocol for the eight experimental flocks. (b-c) Momentary leader-follower interactions over 10s time windows for an example flock in releases 8 and 10, respectively. Coloured bands indicate who is following a particular leader at each time step. Dark grey indicates that a bird is not following other birds, either because it is top of the hierarchy or flying alone, or because interactions cannot be resolved using our time resolution. The Stage-0 leader (D) is highlighted with a light grey background; note

the disappearance of most of its leadership (coloured bands in (b)) when clock-shifted (c).

**Figure 2. Effect of clock-shift on time delays, leadership ranks and routes.** (a) Distribution of momentary time delay values ( $\tau$ ) for Stage-0 leaders for releases 8-10, averaged for each time step over the eight flocks. Positive values indicate being ahead of the mean of the flock. (b) Boxplots of median standardized ranks of Stage-0 leaders. \* indicates significance level between Stages (\* $<0.5$ , \*\*\* $<0.001$ ) and the numbers the effect sizes from post-hoc Tukey tests, only after a significant effect of *stage* in a LMM with *flock* as a random effect was found (see figure S4 for boxplots of the other ranked birds). Standardized rank 100 is equivalent to Rank 1 (top). (c) Median deviation in virtual vanishing bearings in the tracks of the four clock-shifted Stages compared with the same birds' final training flight. Bars indicate standard errors across all flocks. Dashed grey lines indicate the standard error of expected clock-shift across the four Stages. (d) Flight tracks of the Stage-0 leaders during their last training flight (black) and the four clock-shift releases. Colouring of lines matches figure 2c; dashed lines indicate the leader flew alone or the flock split (i.e. three or fewer birds remained). White circles show release site; black circles show home loft.

Figure 1

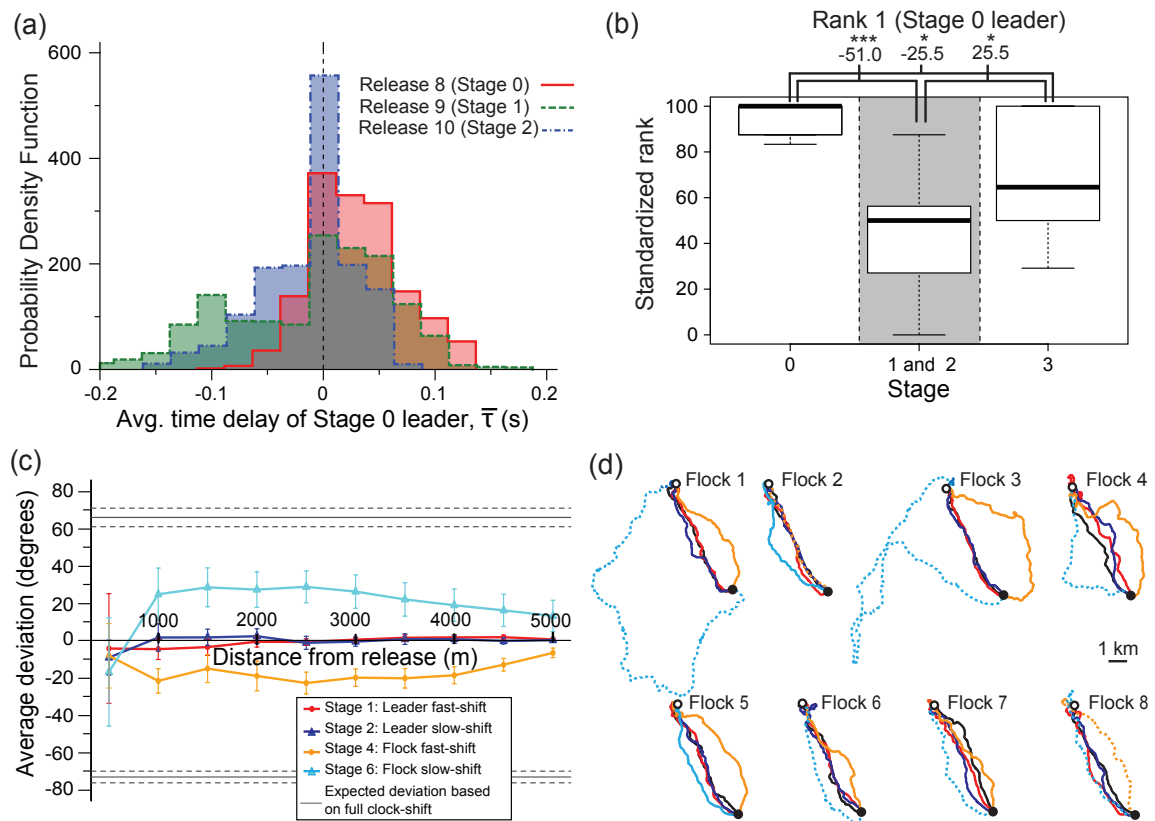
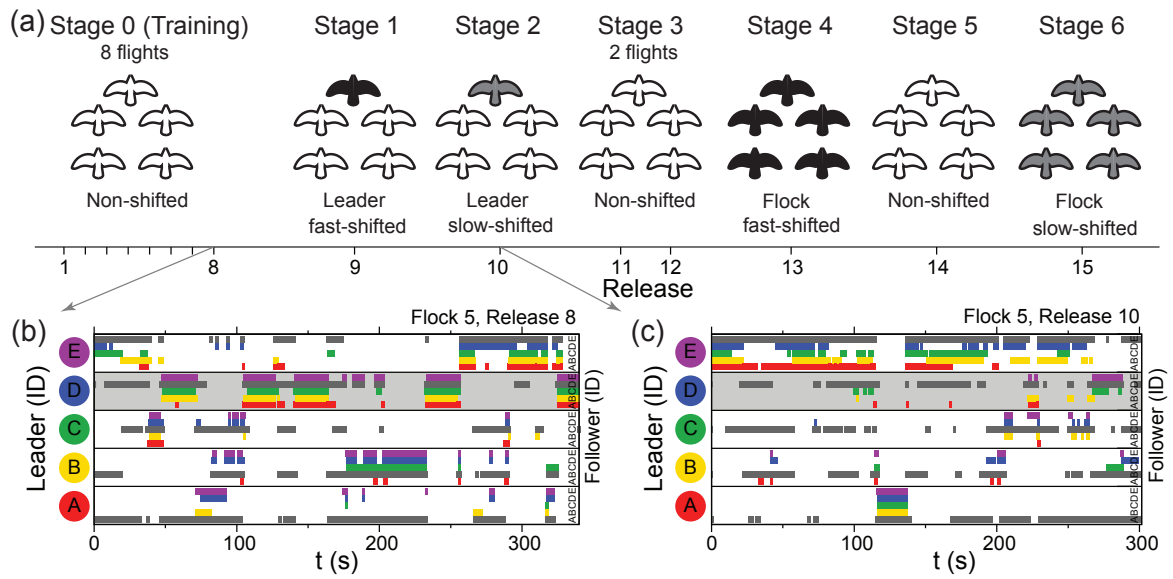


Figure 2



# Misinformed leaders lose influence over pigeon flocks

Isobel Watts<sup>a\*</sup>, Máté Nagy<sup>a,b,c</sup>, Theresa Burt de Perera<sup>a</sup> and Dora Biro<sup>a\*</sup>

<sup>a</sup>*Department of Zoology, University of Oxford, Oxford, United Kingdom.*

<sup>b</sup>*Department of Collective Behaviour, Max Planck Institute for Ornithology &  
Department of Biology, University of Konstanz, Konstanz, Germany*

<sup>c</sup>*MTA-ELTE Statistical and Biological Physics Research Group of Hungarian  
Academy of Sciences, Budapest, Hungary.*

## Supplementary Material

### Detailed Experimental Methods

#### Subjects and materials

40 homing pigeons, aged 2 to 5 years and bred at the Oxford University Field Station in Wytham, Oxfordshire, UK, served as subjects. They were chosen from among the ca. 120 subjects living together at the facility, by randomly selecting eight birds that had hatched in each of 2009, 2011 and 2012, and 16 that had hatched in 2014. These 40 birds were then randomly assigned to eight flocks of five birds, with the condition that each flock had the same age structure: one bird hatched in 2009, one in 2011, one in 2012, and two in 2014. During all homing flights, positional data were logged using miniature GPS loggers (Qstarz BT-Q1300ST) attached to the birds' backs via Velcro strips glued to trimmed feathers. The devices weighed 15.5g and logged time-stamped longitude and latitude coordinates at 5Hz (absolute spatial error: mean = 1.69m, 95th percentile = 4.33m [1]).

#### Training and testing procedures

Prior to starting the experiment, all birds received basic training, which consisted of multiple flock and solo flights from sites 2-3km from the loft in the four cardinal directions, while carrying plasticine dummy weights of the same shape and weight as the GPS devices. Subjects were transported to release sites by car and released at 10-minute intervals to prevent them merging en route.

We then trained flocks to home from our chosen site (Bladon, 51°49'23.48"N 1°21'26.29"W; distance and direction from home: 5.27km, 149.5°) through eight consecutive releases, over a period of one week ("Stage 0"; see figure 1a of main manuscript for experimental design). We staggered the start of training, with four of the flocks only starting training once the other four had completed Stage 0 and had begun the clock-shifting procedure. This was due to space limitations in the clock-shift lofts.

We calculated leadership ranks for each member in each flock using two different methods. First, leadership ranks were assigned based on spatial positioning, using each bird's average distance from the centre of the flock along the front-back axis projected onto the flock's direction of motion. Birds nearer the front of the flock were given higher ranks, while those nearer the back were given lower ranks. Second, we calculated ranks through directional correlation delay analysis, which evaluates the temporal relationship between directional changes that each possible pairing of birds within the flock performs. In these pairwise comparisons, the bird that performs the same sequence of movement changes as another bird, but, on average, delayed in time, is designated the follower, and the other the leader. We resolved leader-follower relations between all pairs of birds, giving each pair a tau value (the average time delay between the two given birds). These pairwise leader-follower relations were then used to construct leadership hierarchies for each flock. A given bird's leadership score was calculated as the average of its tau values, with the most



positive value identifying the bird at the top of the hierarchy. Ref. [2] provides further detail on both methods.

We found a significant correlation between the leadership values obtained through the two methods (Linear Mixed Model (LMM) with *flock* as a random factor: comparison to model without time delay using Maximum Likelihood,  $\chi^2 = 9.90$ ,  $P = 0.002$ ; figure S1). In other words, birds at the front of the flock were also those whose directional changes were consistently copied by birds further behind. In our subsequent analyses we used the leadership ranks obtained through analysis of spatial positioning, and identified as “Stage-0 leaders” the birds with the highest average rank over the last four training releases. The Stage-0 leader was positioned at the top of the hierarchy in 21 of the total of 32 (8 flocks x 4 flights) Stage 0 flights, and positioned in the top two in 29 of the 32 flights. Therefore, in 91% of Stage 0 flights the bird we identified as the Stage-0 leader was positioned in one of the two positions with the greatest influence on the flocks’ directional decisions. Figure S2a shows the standardized ranks of the Stage-0 leaders over Stage 0.

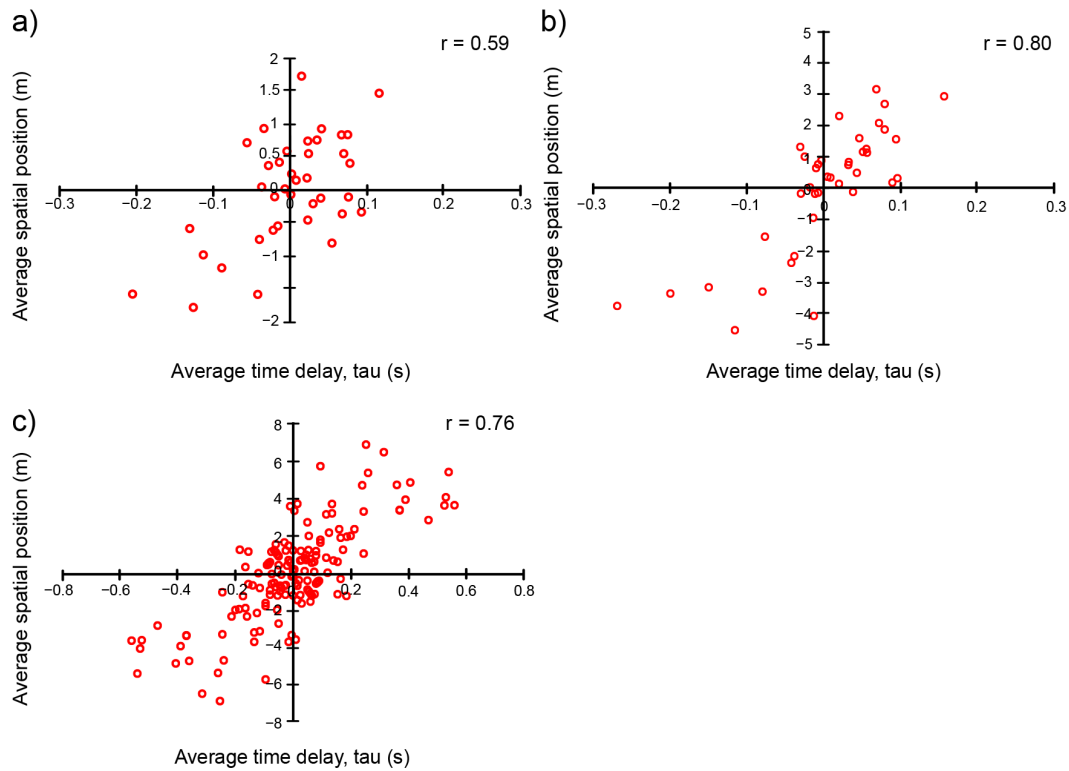
After training we conducted six experimental stages. We performed four clock-shifts (Stages 1, 2, 4, and 6; see figure 1a), during which selected birds were placed in light-tight chambers until their internal clocks had readjusted to an artificially shifted day-night cycle. All shifts corresponded to either an anticlockwise (fast) or clockwise (slow) 70° shift in the sun’s azimuth on the dates of release. The experiment was run over a three-month period, with a six-week break before Stage 3 for half the flocks and before Stage 5 for the other half, due to the staggered starts. Therefore, birds were clock-shifted either two or three hours depending on the time of the year, in order to maintain a roughly 70° shift.

When only Stage-0 leaders (n=8) were being clock-shifted (Stages 1 and 2), we also placed the remaining members of the flocks (n=32) in a light-tight chamber, but their experimental sunrise and sunset times coincided with true sunrise and sunset. Birds remained in the clock-shifting chambers between four and seven days (depending on weather conditions) before being transported in light-tight containers to the release site and released in their original flocks. Light-tight containers were used to ensure clock-shifted birds were unable to begin re-adjusting their internal body clocks already prior to being tested, i.e., while being transported to the release sites.

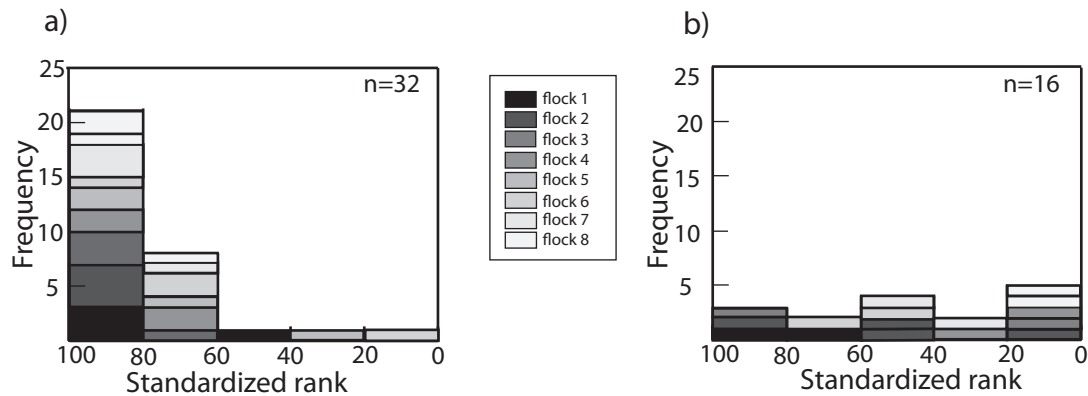
#### GPS data processing

Upon birds' return to the loft, we downloaded their data using QTravel software (Qstarz V.3.2). GPS tracks were analysed in Matlab (Mathworks 2012b) and R (0.98.1014). The geodetic latitude and longitude coordinates provided by the GPS were first converted to X and Y Universal Transverse Mercator (UTM) coordinates using UTM projection. We then processed tracks by discarding all points other than those where speed was continuously above  $5\text{ms}^{-1}$  for 10s before or after the given point (i.e. flight) and trimmed tracks to end once birds had reached to within 100m of the loft.

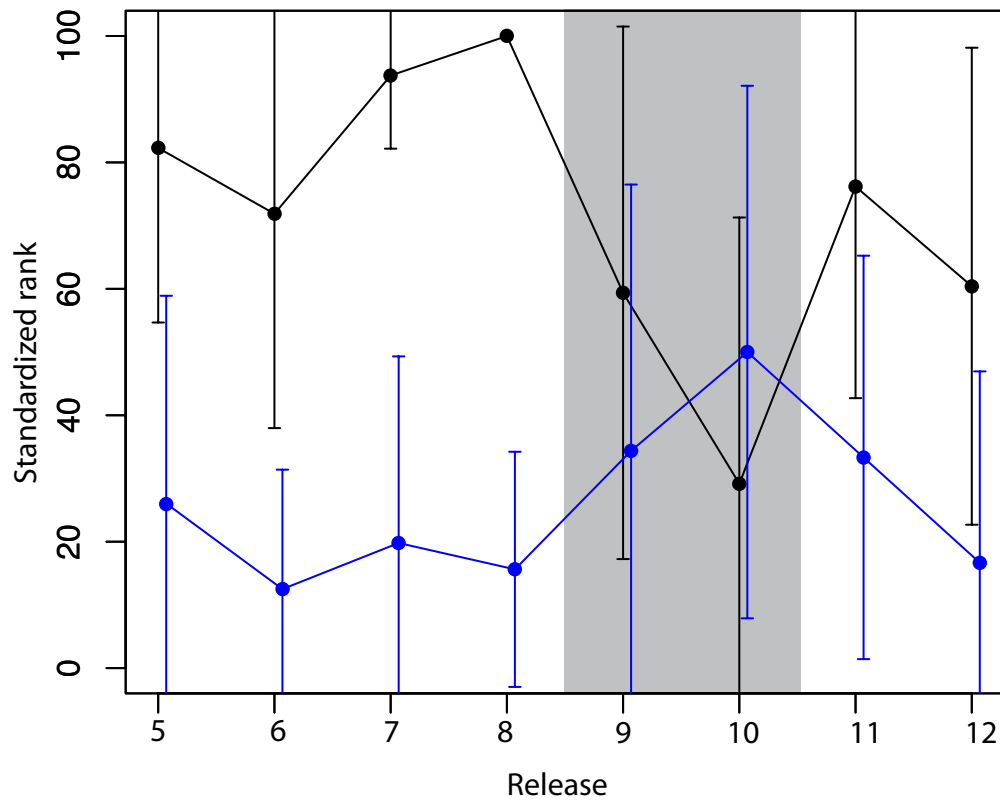
## Supplementary Figures



**Figure S1: Correlation between time delay ( $\tau$ ) values and those based on birds' spatial positioning within the flock.** Correlations between  $\tau$  and spatial positioning were calculated for each bird as an average over each pair in releases 5-8 with **a)** a distance filter of 10m, and **b)** a distance filter of 30m. Distance filters refer to a threshold value: for data fixes where the distance between two birds is greater than this threshold, those fixes are not used in the calculation of  $\tau$  or spatial positioning. **c)** Shows the pair-wise average  $\tau$  for each release, plotted against a pair's spatial positioning. Correlations were calculated using Pearson correlation tests between the two variables in question.



**Figure S2: Histograms of the standardized rank positions of the Stage 0-leaders during a) Stage 0 and b) Stages 1&2 (clock-shift).** During stage 0, leaders were posited at rank 1 in 21 of the 32 flights and positioned rank 1 or 2 in 29 of the 32 flights. Rank standardization accounts for the varying size of flocks in some releases due to birds splitting. Standardized rank 100 is equivalent to rank 1 (top of the hierarchy) and standardized rank 0 is equivalent to the bottom of the hierarchy. These two distributions are significantly different (Kolmogorov-Smirnov test, p-value calculated using bootstrapping  $n=1000$ ,  $D = 0.62$ ,  $P < 0.001$ ).

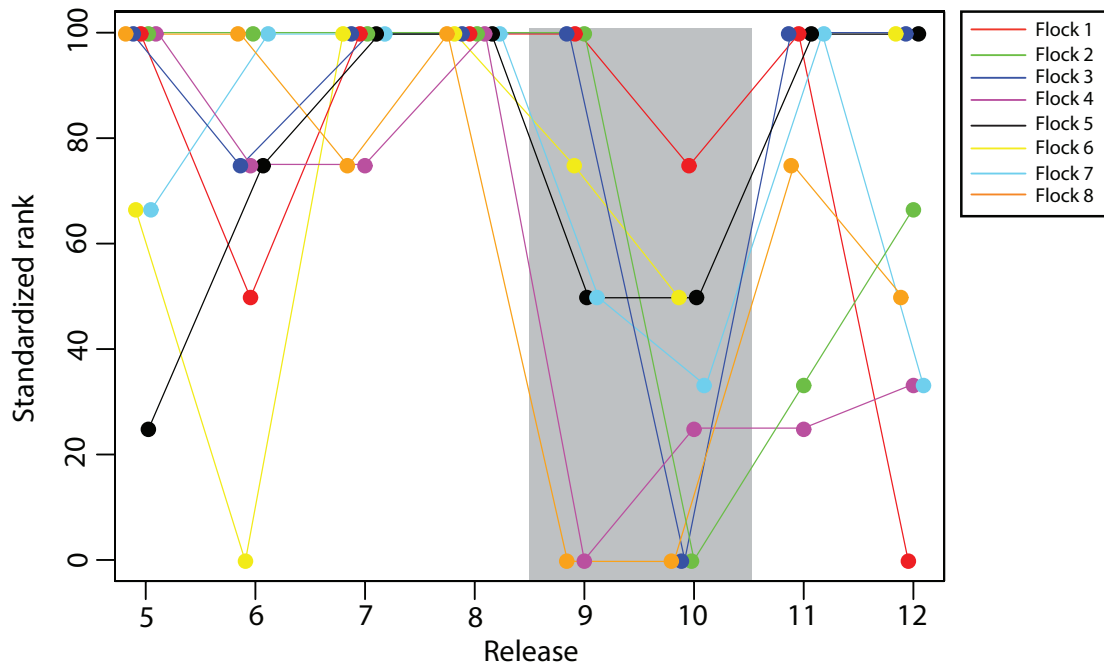


**Figure S3: Mean standardized ranks of birds identified as the highest and the lowest ranked after Stage 0 (i.e. in releases 5-8), across Stages 0, 1, 2 and 3.**

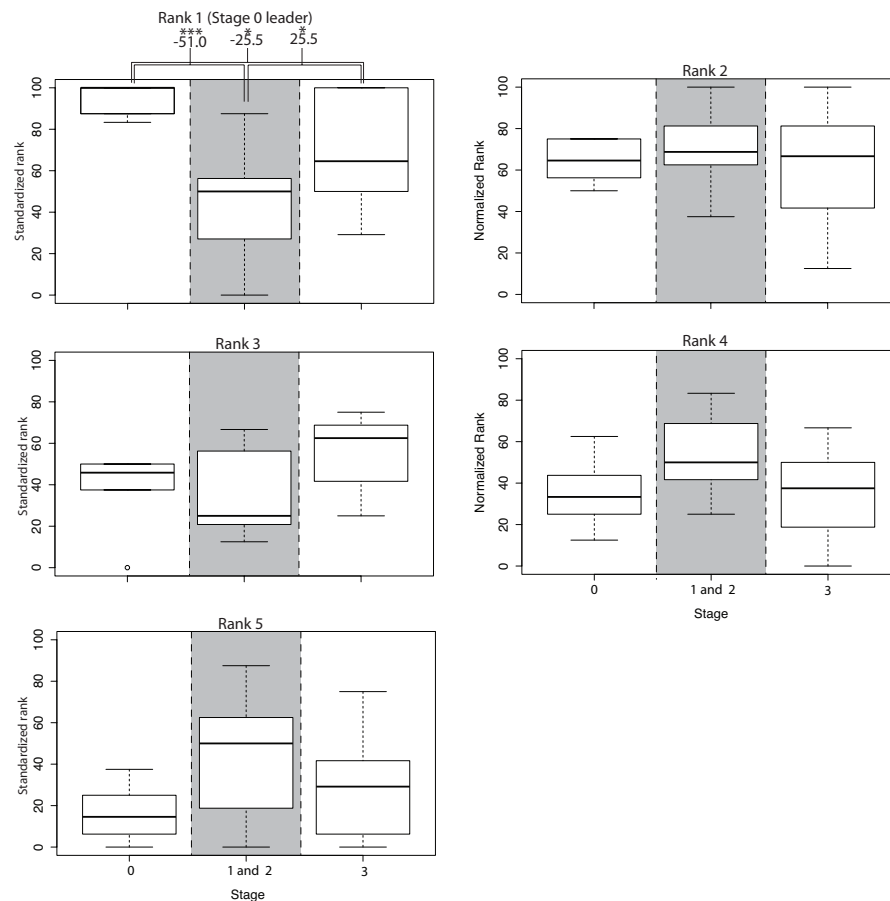
The black line indicates the bird with the highest average rank during Stage 0 (i.e. the Stage-0 leader), and the blue line the bird with the lowest average rank (i.e. 5).

Error bars indicate the standard deviation around the mean. Rank standardization accounts for the varying size of flocks in some releases due to birds splitting.

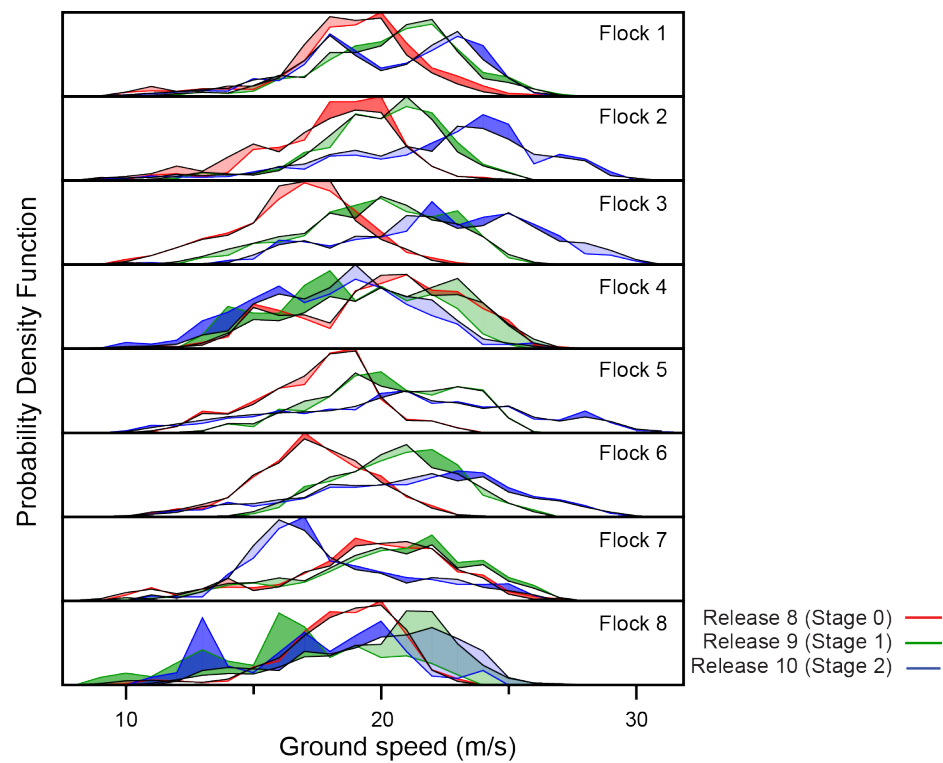
Standardized rank 100 is equivalent to rank 1 (top of the hierarchy) and standardized rank 0 is equivalent to the bottom of the hierarchy.



**Figure S4: Standardized ranks of birds identified as Stage-0 leaders (i.e. those with the highest rank in releases 5-8), across Stages 0, 1, 2 and 3.** Points are spread on the x-axis to enable identification of different flocks. Rank standardization accounts for the varying size of flocks in some releases due to birds splitting. Standardized rank 100 is equivalent to rank 1 (top of the hierarchy) and standardized rank 0 is equivalent to the bottom of the hierarchy.

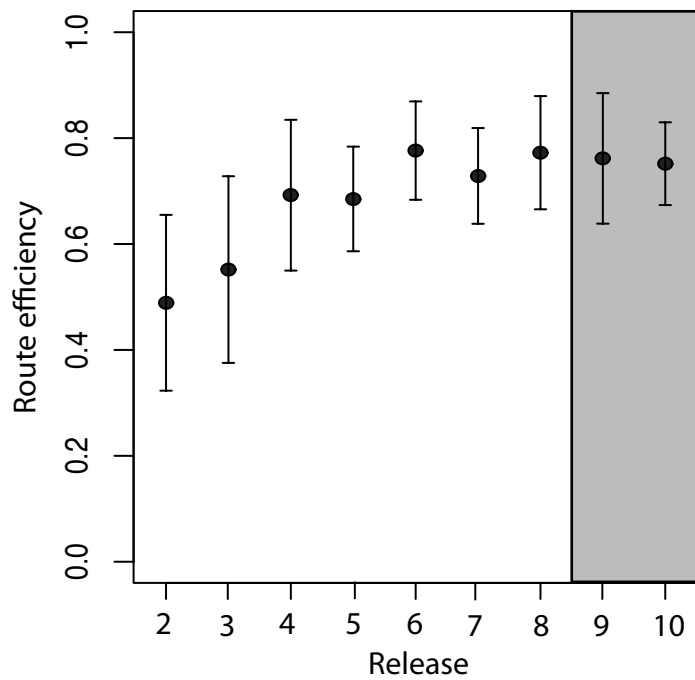


**Figure S5: Standardized ranks of each bird ranked after Stage 0, in Stages 0, 1&2 and 3.** Box plots show the change in standardized rank for all flocks over each stage. Grey shading represents the clock-shift stages. We tested for significant differences in the mean ranks between the three stages by comparing a LMM for each plot with *standardized rank* as the response variable, *stage* as a fixed variable and *flock* as a random factor, to a model without *stage* using a maximum likelihood ratio test. Significance was only found for Rank 1 ( $P < 0.001$ ), thus for all other birds the standardized ranks did not significantly change between stages. The lowest-ranked birds' position is useful for comparison, since they, like leaders, are subject to a ceiling/floor effect. We then used a post-hoc Tukey test to compare the means of the three stages within Rank 1: Stage 1-2,  $P < 0.001$ , effect size = -51.0; Stage 1-3,  $P = 0.02$ , effect size = -25.5; and Stage 2-3,  $P = 0.02$ , effect size = 25.5). These levels of significance are shown as asterisks and the effect sizes as numerals below them (in the top left panel).



**Figure S6: Ground speed distribution for the Stage-0 leaders compared to the rest of the birds for all 8 flocks.** Ground speed was calculated for each bird from the GPS trajectories. Ground velocities are the vectorial sums of two vectors: the flight velocity compared to the air and velocity of the air due to wind. Only those data points were used where a bird had at least two neighbours within a radius of 30m. Assuming that all birds of a flock were flying under similar conditions, the Stage-0 leader's Probability Density distribution (coloured line) was compared to that of the rest of the flock (shown in black). Dark shaded areas indicate where the distribution for the Stage-0 leaders is above the rest of the flock, and light shaded areas indicate where the distribution for the Stage-0 leaders is below the rest of flock. In Flocks 1, 2, and 3 the Stage-0 leaders flew faster compared to the flock in all non-shifted and the two clock-shifted releases. In Flocks 4 and 8 the Stage-0 leaders flew with similar speeds as the rest of the flock in the non-shifted release, but flew slower in the two clock-shifted releases. In flocks 5, 6 and 7 the Stage-0 leaders flew with similar speeds as the rest of the flock in all non-shifted and the two clock-shifted releases.





**Figure S7: Mean route efficiency of the eight flocks across Stages 0-2.** Route efficiency was calculated as the straight-line distance between the release site and the home loft, divided by the actual distance travelled by the bird. Individual birds' efficiencies were then averaged to obtain a single value for each flock (note that this was done since within the same flock birds' efficiencies were not independent). The grey shading around releases 9 and 10 represents the clock-shift Stages 1-2. Error bars indicate the standard deviation around the mean of the eight flocks.