Testing measures of animal social association by computer simulation

David J. White[1,3] **& V. Anne Smith**[2]

([1] Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA 19104, USA; [2] School of Biology, University of St. Andrews, St. Andrews, UK)

(Accepted: 27 August 2007)

Summary

Techniques used to measure patterns of affiliation among social animals have rarely been tested for accuracy. One reason for this lack of validation is that it is often impossible to compare sample data to the true distribution of social assortment of a group of animals. Here we test some methods of assessing social assortment by using a computer simulation of organisms whose assortment patterns were under our control. We created male and female organisms that moved in a direction that was based on a social bias parameter. As the weight of this parameter increased, organisms were more likely to move in the direction of others of their sex. We then created virtual observers to sample assortment of the organisms under different social bias conditions. Observers used three different techniques of measuring assortment. These were (1) group membership: noting all organisms that were associated in the same 'group', (2) nearest neighbour: noting the nearest organism to a randomly selected individual and (3) neighbourhood: noting all organisms near a selected individual. Neighbourhood was taken either by all-occurrence sampling or by focal sampling the associations of randomly selected individuals. Some techniques emerged as more sensitive than others under different conditions and biases were revealed in some measures. For example, the group membership method was biased toward finding significant assortment differences between the sexes when no difference actually existed. Nearest neighbour was insensitive to finding a difference in assortment between sexes when one existed. Focal sampling was less sensitive to finding effects than all-occurrence sampling. The computer simulation revealed properties of each technique that would have been impossible to detect in the field.

Keywords: social assortment, social, animal, agent-based model, association index.

[3] Corresponding author's e-mail address: whitedj@psych.upenn.edu

Introduction

Documenting patterns and frequencies of social assortment among animals in groups can provide insights into levels of biological organization that are not possible to understand at the individual level (Hinde, 1976; West-Eberhard, 1983; Whitehead & Dufault, 1999). A large literature has amassed of studies on social assortment in animals of a variety of taxa from insects, birds, marine mammals, and primates (Whitehead & Dufault, 1999). Along with this abundance of studies, an abundance of techniques has emerged to document social structure. These techniques differ depending on species, ecology, and research questions and specific techniques are often chosen based on practicalities of the particular situation. While studying different species obviously requires different techniques, it has been extremely difficult to measure the efficacy of the different measures or to compare them. Few attempts have been made to test the validity, reliability, and power of the different methods.

Techniques used to measure association fall into three major categories: (1) group membership techniques, where association is measured by the frequency with which two individuals co-occur within the same group (e.g., Clapham, 1994 (whales); Frederick & Johnson, 1996 (marsupials); Holekamp et al., 1997 (hyenas); Newton-Fisher, 1999 (chimpanzees); Durrell et al., 2004 (pigs); Connor et al., 2006 (dolphins); Weinrich et al., 2006 (whales)), (2) nearest neighbour techniques, where the single closest individual to a focal animal is noted (e.g., Boinski & Mitchel, 1994 (monkeys); Dwyer & Lawrence, 1999 (sheep); Freeberg, 1999 (birds)) and (3) neighbourhood techniques, where all individuals within a fixed distance to a focal animal are noted (e.g., Boinski, 1994 (monkeys); Digby, 1995 (marmosets); Smith et al., 2002 (birds), Mitani et al., 2000 (chimpanzees), Silk et al., 2003 (baboons); Mollema et al., 2006 (cattle)). Analyses done on these data range from the mainly descriptive (e.g., Digby, 1995; Félix, 1997), to correlations using individuals' data (e.g., Holekamp et al., 1997; Freeberg, 1999), to comparisons of groups of animals (West et al., 2002). We concentrated on this last use, where the analyses take the form of either comparing the amount of association with different classes (which can be age, sex, kinship group, dominance group, etc.; e.g., Holekamp et al., 1997; Dwyer & Lawrence, 1999; Freeberg, 1999; West et al., 2002a,b), or how the strength of association differs between these classes (e.g., Holekamp et al., 1997).

Our interest in techniques of assessing social structure was predicated by our changing methods for studying social learning in brown-headed cowbirds (*Molothrus ater*). We have realized the need for exploring the effects of group-level processes on learning and development, requiring a shift in paradigms from studying individual subjects as isolates from the social group to studying groups of subjects. To characterize the nature of social groups, we have used a series of assortment tests that were modeled after other techniques used in the field (Smith et al., 2002; West et al., 2002b). These techniques have included measuring nearest neighbour and neighbourhood associations. This latter technique has taken several forms, from selecting the order of the target birds prior to observers entering the aviaries and then taking one data point from each target, to all-occurrence sampling all neighbourhood associations that the observer could record during an observation session. We have also used focal samples where we record all neighbourhood associations for randomly selected individuals for a set amount of time. We now use voice recognition software to collect data on social assortment, which has led to a fourfold increase in the amount we can take, but it has also led to us using less restricted rules of sampling (White et al., 2002a). Over the years we have believed that each technique has had its strengths and weaknesses in different situations, and practical considerations required the use of some of the techniques, but we had no true assessment of the accuracy of the various techniques.

The purpose of this work was to provide a new means for testing the performance of the rules for sampling social assortment. Measuring social behaviour has many difficulties. Since assortment is not behaviour per se, but the outcome of behaviour, it can be difficult to conceptualize or quantify. Judgments are required to decide how close individuals must be to one another to be considered associated and to determine why the individuals were associated. For example, individuals could be close together because they were drawn to a resource, as a result of random movement, or due to a preference to be near conspecifics or particular individuals. Furthermore, there are sampling issues. Any sample of assortment is usually a minuscule proportion of all assortment occurring over time in a group and particular rules of sampling may produce patterns that are not representative of the population. In sum, social assortment can be composed of a complex of rapidly moving individuals, providing nothing but experimenters' observations as a

quantifiable product. These factors make assortment difficult to document and difficult to determine if it was documented validly or reliably.

Most often, tests of social assortment have been designed to evaluate the patterns in the collected data to determine how probable it would be that the results would be produced given a null hypothesis of random assortment (Hinde, 1977; Manly, 1995; Whitehead, 1997; Bejder et al., 1998). While this is a critical analysis to detect if the patterns generated by the methods reveal systematic effects, it does not, however, test whether the measures used to collect the data did so accurately. Do measures of social assortment, for example, bias the observer to document some individuals at the expense of others? Do the movement patterns of one set of individuals influence the observed movement patterns of another set? What measurement techniques are most sensitive to differences in movement patterns of animals and under what conditions do the various techniques work best? To answer these questions, the actual underlying patterns of assortment (the entire population of assortment of all individuals) must be known; it is this actual underlying pattern against which the results of the sample measures can be compared.

We created a computer simulation wherein virtual organisms associated with each other using known rules and their association patterns were measured by virtual observers. The observers used several techniques for sampling. The three main types of association data they collected were group membership, nearest neighbour, and neighbourhood patterns. We assessed the virtual observers' accuracy by comparing their conclusions about assortment patterns with the true underlying distribution of assortment for each data type measured. Comparing true underlying distributions for the three data types provided us with a measure of how the techniques compared to one another that was independent of sampling issues or observer bias.

The model

We used an agent-based simulation because these types of models have been used effectively in studying social organization (Reynolds, 1987; Schank & Alberts, 2000; Bryson et al., 2007). We used C++ to create a two-dimensional confined world where moving agents (organisms) assorted and observers sampled the assortment of the organisms (code available on request).

Organisms

We created 10 male and 10 female organisms numbered for individual identification. Organisms existed in a world that measured 100 × 100 space units with no obstructions. During each time step, organisms could either be active (probability = 0.75) or inactive. If they were active, organisms moved from one x–y coordinate in the world to another at a random velocity that would move them from 1 to 10 space units. They could move in one of eight possible directions 45 degrees from each other. At each time step, each organism's movement could be either socially biased or not socially biased based on probability s, the social bias parameter. When movement was not socially biased, organisms moved in one of the eight directions with 0.125 probability. When movement was socially biased, organisms would move in one of the eight directions with a probability based on the number of individuals of the same sex occupying the quadrant of space bisected by that direction. However, only other organisms within the individual's 'local awareness' (set at 15 space units in all directions around the organism) were used to influence the individual's movement. For example, if there were organisms of the same sex at 10°, 50°, and 200° of a given organism (within its local awareness), the organism would be twice as likely to move in the 45° direction as the 225° direction because there were 2 individuals in the 0-90° quadrant and only one in the 180-270° quadrant. Similarly, 45° would be twice as likely as 90° for the same reason (see Figure 1). We pilot tested a number of different values of s and determined $s = 0.3$ to result in high levels of assortment by class. Thus, here we report on simu-

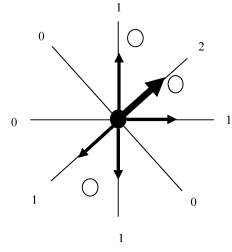

Figure 1. A graphical representation of the socially biased movement routine. The filled circle represents the individual making the movement decision; open circles indicate organisms within the centre individual's local awareness; numbers at the end of the eight directional lines indicate the number of organisms within the quadrant bisected by that line. Strength of social influence for moving in each direction is indicated by the thickness of arrows.

lations where s was modified between 0.3 and 0 ($s = 0$ results in movement that is random with respect to other organisms). Males and females could be assigned different values of s (s_m for males and s_f for females). Animated display of organism movement patterns under strong social bias ($s_m = s_f = 0.2$) and under weak social bias ($s_m = s_f = 0.02$) are available at http://www.psych.upenn.edu/~whitedj/lab/cowbird.htm

Observers

We programmed eight different types of observers; three took group membership data, one took nearest neighbour data, three took neighbourhood data (all occurrence and focal), and one took random data (each explained below). We attempted to make the behaviour of observers simulate real-world field observers. Observers could sample data once every 5 time steps in the simulation. We made observers slower than the movement rate of organisms because this is true for observing bird movements (but may not be true for other study species, see below). Some observers took data in blocks of time after which certain sampling rules were reset for the beginning of the next block (see description of observers, below). Blocks were composed of 60 samples of the organisms (300 time steps in the simulation).

Observers using group membership techniques began in a random location in the world and recorded all organisms within the group nearest the observer's location at that instant. The organism nearest the observer became the target individual for the group. Group membership was defined as all individuals within five spaces of the target, as well as any individuals within five spaces of any other group member. For the next sampled data point, observers moved to a new location and sampled again. This was done to simulate a situation where an observer may be moving through the environment, such as forest or the ocean, and recording membership of groups encountered. The three group membership observers differed only in the way they analyzed the resulting data distributions (see analyses below).

The nearest neighbour observer systematically selected each organism in a randomly generated order and noted the individual that was closest to it. If there were no other organisms within 30 spaces of the selected organism, the observer moved on to the next organism in the order. This 30-space 'giving up distance' was incorporated to simulate the real-world difficulty of determining a nearest neighbour for an organism that was a great distance

away from any other. The missed organism was placed at the end of the order to provide one more chance to collect data on it. At each sampling time point, the observer took data on the next organism in the order. When the observer reached the end of the random order, a new random order was generated, and the observer began at the beginning of the order. Thus, the nearest neighbour observer collected near-equal amounts of data on all organisms.

Neighbourhood data was taken both by ad lib (Altmann, 1974) scanning for all associations (two different observer types) and by focal sampling individuals (one observer type). For the scan method, two observers noted all organisms that were within 5 space units of a target individual. These two observers differed in how the target organism was selected. One observer randomly selected individuals that had at least one organism in its neighbourhood as the target individual. The target then could not be re-sampled until the data collection block was over. The second scan observer also randomly selected individuals in associations, but could reuse the target organism for other associations within the block. Both observers only counted an association as a single data point, that is, once a target was recorded as having another in its neighbourhood, the association could not be recorded again when the other organism became the target. After a data collection block was over, all information was re-set. The focal observer systematically chose each organism using a randomly generated order of focal targets. At each data sampling point, the focal observer noted all other organisms that came within 5 units of the focal target. At the start of a new block, the focal observer recorded associations from the next focal subject in the random order, re-randomizing and starting over when all organisms had been focal subjects.

As a control comparison, we created a 'random' observer who recorded two randomly selected organisms as associated at each data sampling point.

Analyses of observers' data

For each of the group observers, the data were summarized in the form of the number of times each organism was in a group with each other organism. Observers then calculated group membership association indices for all pairs of organisms in one of three different ways according to the most popular methods in the literature (see Cairns & Schwager, 1987). These three methods are all attempts to calculate from observations of groups the proportion

of total time an individual spends in a group with another, and are referred to as half-weight, twice-weight, and square root indices. Half weight was calculated as:

$$\frac{x}{x + y_{ab} + 1/2(y_a + y_b)}$$

twice-weight:

$$\frac{x}{x + y_{ab} + y_a + y_b}$$

square root:

$$\frac{x}{\sqrt{(x + y_a + y_{ab})(x + y_b + y_{ab})}},$$

where x represents instances where individuals a and b are located in the same group, y_a (or y_b) are instances where only individual a (or b) are located in a group, and y_{ab} represents instances where individuals a and b are located separately (after Cairns & Schwager, 1987; Whitehead & Dufault, 1999). Practically it was not possible to determine the value y_{ab} from the data collected from the observers, as they only recorded a single group at a time. Thus, this term was omitted in calculating association indices (as is often done in practice; Whitehead & Dufault, 1999). We then calculated for each organism a mean of the association indices for associations with males and with females.

Nearest neighbour and neighbourhood data were also summarized as the number of associations each organism had with all other individuals. For each individual, we determined sums of the number of associations with males and with females, and calculated the proportion of total associations these formed.

Using the above summarizing values for association with males and females for each individual, all observer types then performed statistical tests to determine whether there were significant levels of assortment within a sex, that is, whether males assorted more with males than with females, and vice versa for females, and whether there was a significant difference in the strength of association between the two sexes, that is, whether males associated with males more or less than females associated with females (hereafter known as the 'sex difference' in association). Thus, we had a within-subjects statistical analysis measuring for each organism whether it assorted more with members of its own sex or members of the opposite sex, and we had

a between-subjects statistical analysis comparing males to females in the strength of their association with members of their own sex.

We attempted to produce sample sizes comparable to what researchers in the field might produce. We were interested in differences of a magnitude that would be investigated in the real world and, thus, we subjected our virtual organisms to the standards of statistical inference to which real organisms would be held. Assortment was considered significant using $\alpha = 0.05$ (two-tailed). We used non-parametric analyses (Wilcoxon T- and Mann–Whitney U-tests) because these are commonly used for testing significance in the field. Thus, we were able to assess the statistical power for the different methods under similar circumstances as would be done in the real world.

True underlying distributions of association

The true underlying distributions represent the population of all associations. There were several true underlying distributions because the population differs depending on the data types; group membership, nearest neighbour, and neighbourhood. Each of these data types represent associations differently and, thus, require a separate population of associations. The true underlying distributions differ from the observers' data in several ways. All associations made by organisms are recorded at all time steps. At each time step, group membership was determined and recorded for all individuals using the same criteria as for the observers. We then calculated the group true association index as the proportion of time an individual spent in a group with another individual. To do so, we divided the number of time steps the two individuals appeared in the same group by the total number of time steps in the simulation. The nearest neighbour of every individual (regardless of absolute distance) was noted at each time step to provide a true underlying distribution of closest individuals. Individuals within the neighbourhood distance were recorded for all individuals every time step to provide a true underlying distribution of neighbourhood composition. Thus, we produced three population measures of associations from which the observers sampled a subset.

As with the observers' data, we calculated from the true underlying distributions the average association indices or proportion of associations of individuals with males and females. Significance for levels of assortment for males, females, and the sex difference were determined for the true underlying distributions in the same manner as for the observers' data.

Running the model

At the start of a trial of the model, organisms began scattered randomly throughout the world and preformed 50 movements prior to data collection to establish association patterns. Observers then collected data and the true underlying distribution was determined for movements in the next 4800 time steps. This trial size was use to create a dataset large enough to make statistical conclusions. To create a sampling distribution around each trial, we had thirty of each of the eight types of observers collect data simultaneously. Since the observers chose organisms randomly, they did not all produce the same sample. Thus, after one trial was done, the model produced the true underlying distributions of association with significance values for within-sex and between-sex differences for the three data types. In addition, it produced the number of observers out of 30 that found significant levels of assortment for males, females, and the sex difference for each of the eight observer types. We varied the s parameter across trials and for every level of assortment we tested (e.g., $s_m = 0.3$, $s_f = 0.3$, see below), we ran 20 replicate trials. To analyze the effects of data type independent of any observer sampling bias, we determined the number of trials out of the 20 that showed significant differences for the three true underlying distributions. To examine the accuracy of observers, we calculated across the 20 trials the mean number of the 30 observers for each trial that found significant differences in their collected data.

We tested assortment levels in two phases. In phase one, we tested situations where assortment levels differed for the sexes. Male assortment remained constant at $s_m = 0.1$ and females varied from $s_f = 0.3$ through 0 (specifically, 0.3, 0.2, 0.1, 0.05, 0.03, 0.01 and 0). Here we expected to see significant assortment within classes at all levels (except for females at $s_f = 0$) and a significant sex difference that varied from a large difference when females were assorting at the high and low levels, and that was reduced as female assortment approached male assortment ($s_f = 0.1$). In phase two we tested situations where males and females assorted at the same strength. Here we varied the strength of assortment of both s_m and s_f from 0.3 to 0 as above. We expected significant effects of assortment within classes (except at $s_m = s_f = 0$) and no significant sex difference.

As well as testing different degrees of assortment by varying s, we also tested the performance of the methods under different movement parameters. We tested performance under four different conditions: (1) under lower

activity than was used in the original simulations, (2) under higher density, (3) under lower density and (4) in a bigger world. For the lower activity test we reduced the probability that organisms were active from 0.75 to 0.25. For the higher density test we doubled the number of organisms in both classes. For the lower density test we doubled the area of the world by increasing it to 144×144 space units. For the bigger world test, we increased the size of the world and increased the number of organisms of each class to 20 each, thus maintaining the same density of our original simulations. We tested these four types under two conditions. First with s_m and s_f both $= 0.1$ and second with $s_m = 0.1$ and $s_f = 0.05$.

Results

True underlying distributions

The true underlying distribution was based on an average of 754.97 (± 1 SEM $= \pm 6.62$) data points per organism per trial using the group membership method. Neighbourhood measures were based on 800.36 ± 6.23 data points per organism per trial. The nearest neighbour method produced 4800 data points per organism per trial.

Figure 2a-c depicts the proportion of trials in which the true underlying distributions found significant assortment for conditions in phase 1, where male and female social movement biases differed. Group membership and neighbourhood measures were highly accurate in determining significant levels of assortment for males (Figure 2a). For female assortment, as expected, as s_f was reduced, the number of trials showing significant female assortment decreased. Nearest neighbour became less accurate at finding significant levels of male assortment as female assortment was reduced. The nearest neighbour measure was also less accurate than the group membership or neighbourhood measure at detecting female assortment (Figure 2b). In addition, when females were moving at random ($s_f = 0$), nearest neighbour found significant female assortment more often than chance (6/20; Binomial test, $p < 0.0001$).

For the sex difference, the group membership measure was most sensitive at finding a significant result (Figure 2c). Group membership outperformed neighbourhood in 10 out of 13 conditions (with 2 ties) and both group mem-

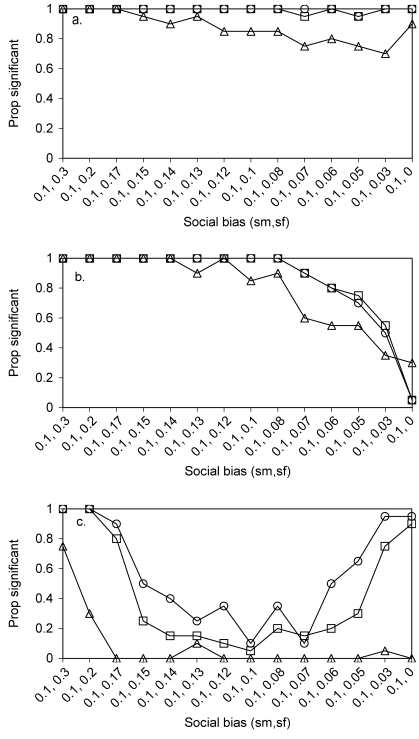

Figure 2. Proportion of trials in which the three true underlying distributions ($\circ$ = group method, $\square$ = neighbourhood method, Δ = nearest neighbour method) found significant levels of (a) male assortment, (b) female assortment and (c) a difference in assortment strength between the sexes for phase 1 conditions where $s_{\mathrm{m}} = 0.1$ and s_{f} varied from 0.3 through 0.

bership membership and neighbourhood outperformed nearest neighbour in every condition.

Figure 3a-c depicts the proportion of trials that the underlying distribution reflected significant effects for conditions in phase 2, where the social movement bias was the same for males and females. Again both group membership and neighbourhood measures were more sensitive in determining assortment within classes than nearest neighbour (Figure 3a,b). Surprinsly, the group membership method produced significant results for the difference in the magnitude of assortment between the sexes more often than chance when no such difference existed (one-sample t-test: $t_6 = 4.40$, $p < 0.005$ combining across the seven conditions).

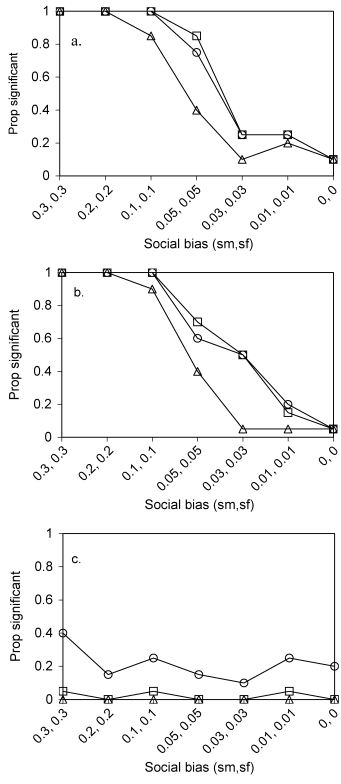

Figure 3. Proportion of trials in which the three true underlying distributions (∘ = group method, □ = neighbourhood method, Δ = nearest neighbour method) found significant levels of (a) male assortment, (b) female assortment and (c) a difference in assortment strength between the sexes for phase 2 conditions where s_m and s_f were equal and varied from 0.3 through 0.

Observer data patterns

Group membership observers took on average 198.08 ± 4.74 data points. Nearest neighbour observers averaged 232.84 ± 4.64 data points per trial. For neighbourhood observers, scan sampling produced 219.90 ± 4.47 data points per trial. Focal observers collected 165.88 ± 3.19 data points. Random observers collected 960 data points per trial.

Figure 4a-c depicts the mean number of the 30 observers that found significant effects across the 20 trials under the different *s* levels in phase 1. Figure 5a-c depicts the mean number of 30 observers that found significant effects across the 20 trials under the different *s* levels in phase 2. In calculat-

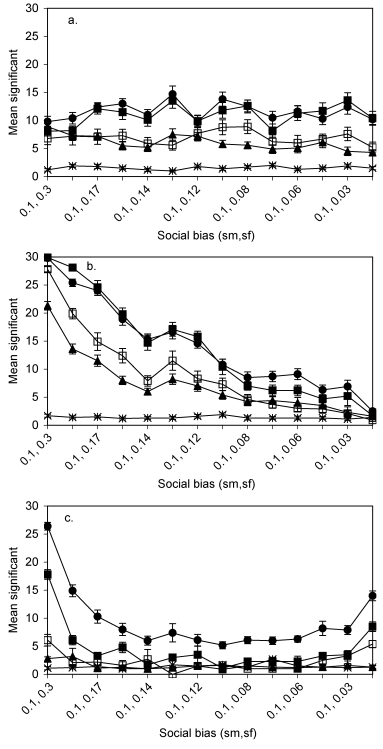

Figure 4. Mean number of observers out of 30 for each observer method across 20 trials that found significant assortment for (a) male assortment, (b) female assortment and (c) a difference in assortment strength between the sexes for phase 1 conditions where $s_m = 0.1$ and s_f varied from 0.3 through 0. • = group observer, ■ = neighbourhood scan observer, □ = neighbourhood focal observer, ▲ = nearest neighbour observer, X = random observer.

ing means, we only analyzed the data from runs where the true underlying distribution correctly reflected the patterns of assortment as given by the social bias parameters. That is, if the true underlying distribution did not find significant assortment where the social bias parameters were set such that there should be assortment (or found assortment where there should have been none), we did not use the observers' data from that trial. Otherwise, it would have been impossible to determine whether it was the measure itself, as reflected by the underlying distribution, or the method of sampling which was responsible for the outcome. We used data from all trials for nearest neighbour measures of the sex difference because the true underlying distribution for nearest neighbour never found significant sex differences where it should have.

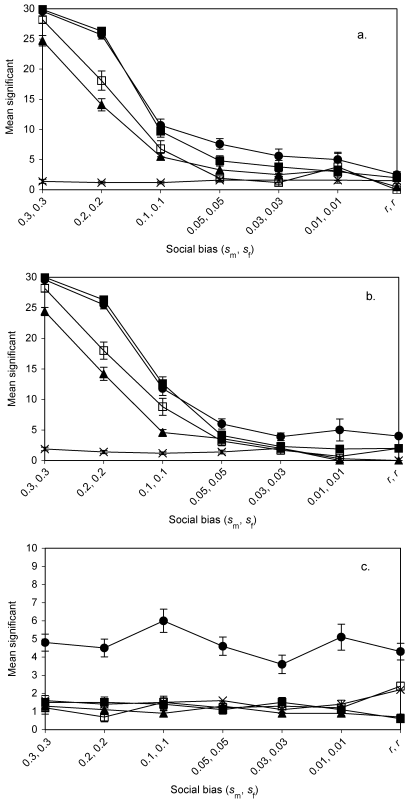

Figure 5. Mean number of observers out of 30 for each observer method across 20 trials that found significant assortment for (a) male assortment, (b) female assortment and (c) a difference in assortment strength between the sexes for phase 2 conditions where s_{m} and s_{f} were equal and varied from 0.3 through 0. • = group observer, ■ = neighbourhood scan observer, □ = neighbourhood focal observer, ▲ = nearest neighbour observer, X = random observer.

Group membership

We found no significant or persistent differences among the results of the twice weight, half-weight, or square root observers (see Cairns & Schwager, 1987). For simplicity, we report only the results from the twice weight observer here. For phase 1 trials, group observers were highly accurate at detecting significant patterns of assortment for males (Figure 4a), females (Figure 4b) and the sex difference (Figure 4c). In this last category, group observers outperformed all other methods at every level of assortment tested.

Similar to the true underlying distribution, group observers in phase 2 found significant differences in assortment between sexes more often than chance when no difference actually existed (Figure 5c; single sample t-tests: all $t_{s\geqslant 12} > 4.12$ all $p_s < 0.0001$). This bias was in addition to the underlying distribution's bias, because we used only the data from the trials where the true underlying group distribution found no significant difference in assortment between the sexes.

Only the bias in the sex difference was influenced by the different movement parameters (Oneway ANOVA conducted on the four different movement parameters and the original: $F_{4,83} = 3.165$, $p < 0.05$). Tukey post hoc analyses revealed that significantly more group membership observers found significant assortment between the sexes (when none actually existed) under high density ($p < 0.05$).

We hypothesized that the increased likelihood of the false positives in the sex differences for the group method was influenced by the size of the groups. The bias increased when the organisms were assorting more strongly and under higher density. If this hypothesis is correct, then the error should be larger when individuals were more clustered. To test this hypothesis, we increased local awareness to a radius of 144 spaces around each individual such that the location of organisms anywhere in the world influenced individual's movement. This served to produce aggregations of larger sizes than the many small groups produced when local awareness was lower. We ran the simulation at $s_m = 0.2$ $s_f = 0.2$ and compared the resulting data to the results with standard local awareness under the same s values. Significantly more observers made the error of finding sex differences when the local awareness was 144 (8.6 ± 1.15) than when the local awareness was 15 (4.5 ± 0.49; $t_{22} = 3.895$, $p < 0.0001$).

Nearest neighbour

Nearest neighbour measures were less accurate than group or neighbourhood scanning methods at detecting assortment within classes (Figures 4a,b and 5a,b) and the sex difference (Figure 4c).

The only influence of different movement parameters on the performance of the nearest neighbour measures came in the higher density simulations. Nearest neighbour measures were significantly more effective at detecting male and female assortment under high density than under normal density for simulations when $s_m = s_f = 0.1$ and when $s_m = 0.1$ and $s_f = 0.05$ (Oneway ANOVAs: all $F_{s\geqslant 4,65} \geqslant 7.61$, all $p_s < 0.001$).

Neighbourhood

There were no differences in performance of the two types of neighbourhood scan observers. Here we only report the results from the scan sampling observer that could not resample target organism within blocks. The scan sampling methods performed similarly to the group method in documenting within sex effects (Figures 4a,b and 5a,b). While the group method outperformed the scan method in detecting sex differences when they existed (13/13 times; Figure 4c), the scan method, unlike the group method, showed no bias in detecting a sex difference when one did not exist (all one-sample t-tests, $t_{s\leqslant 19} \leqslant 1.67$, NS; Figure 5c).

Changing activity levels of the organisms influenced performance of the scan neighbourhood method. Under lower activity level, the scan neighbourhood method was significantly better at detecting male and female assortment at $s_m = s_f = 0.1$ and $s_m = 0.1$, $s_f = 0.05$ (ANOVAs: all $F_{s\geqslant 4,71} \geqslant 2.67$, all $p_s < 0.05$).

Focal sampling was significantly less sensitive at detecting any significant effects than was scan sampling. Summing across all trials in phases 1 and 2 where there was significant assortment, focal observers performed worse at detecting the assortment than scanning 50 times and better than scanning only twice. The focal sampling method was not influenced by any of the different movement parameters.

While the focal method's lack of sensitivity in finding effects compared to scanning could be related to the amount of data collected, we hypothesized focal sampling may have performed worse than scan sampling in part because it was less likely that a focal sample of any organism at a point in time would be representative of the patterns of all organisms compared to a scan sample of all organisms. To test this hypothesis, we varied the block length, the amount of time the focal observer spent sampling an individual target. We varied the block length from 60 data collection time steps (standard), up to 240 time steps, and down to 30 time steps, and 5 time steps. Thus, at a block length of 5, focal observers were shifting between target individuals as rapidly as were scanning observers. Data are shown in Figure 6. Amount of data collected did not change significantly ($F_{3,76} = 0.051$, NS), but accuracy improved as block length was reduced. Accuracy of detecting male assortment, female assortment and the sex difference all increased as block length decreased (males: $F_{3,75} = 3.97$, $p < 0.05$; females: $F_{3,58} = 4.894$,

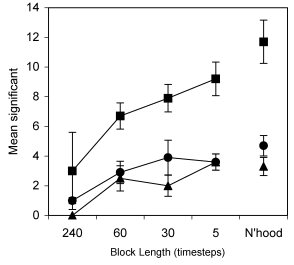

Figure 6. Mean number of neighbourhood focal observers out of 30 across 20 trials that found significant levels of (■) male assortment, (●) female assortment, and (▲) a difference between the sexes in assortment strength at $s_m = 0.1$, $s_f = 0.05$ when the block length varied from 240 time steps to 5 time steps. Neighbourhood scan data presented (N'hood) for comparison.

$p < 0.005$; sex difference: $F_{3,26} = 9.42$, $p < 0.001$). All showed significant linear trends (males: $F_{1,75} = 14.45$, $p < 0.001$, females: $F_{1,58} = 9.06$, $p < 0.005$, sex difference: $F_{1,26} = 20.98$, $p < 0.001$).

Discussion

As expected, the true underlying distributions were comprised of more data and were more sensitive in detecting assortment patterns of the organisms than were the observers' samples. Both the populations (true underlying distributions) and samples (observers) of social assortment reflected the social bias parameter across most levels of assortment, thus indicating that the measures were valid overall in measuring social assortment, though there were differences in power across measurement types. In addition, some systematic inaccuracies did emerge at both the population and sample levels revealing some problems in measurement and sampling rules.

Group membership

The three group membership indices did not differ substantially in our simulation. Cairns & Schwager (1987) suggest that these indices tend to vary due to biases in observers' ability to locate individuals in a population. Since our model did not have such biases in the virtual observers, it is reasonable that the indices did not differ.

Overall, group membership observers were highly accurate at detecting sex differences in assortment strength when they existed but were more likely

to make a false positive in assortment between the sexes. This bias was apparent in the true underlying distribution, and more pronounced in the observers' sampling data. The bias was more apparent when the organisms were assorting by class under strong social bias and also when we manipulated movement patterns to make them assort in larger groups. It seems that in large groups, associations among individuals compound rapidly. Thus, small differences in group size occurring by chance can produce pronounced differences in individuals' association strengths. For example, in a large group, two individuals can be separated by a large distance, but still be considered in the same group through linkages to others. Perhaps this leads to higher levels of assortment across individuals than are truly represented. For most tests in the simulation, the group method proved to be highly accurate. Care should be taken however, when making comparisons across classes (age, sex, kinship, dominance, etc). This is especially the case for subjects that assort in large groups. This bias is not a sampling issue, but is inherent in the measure itself. We are currently experimenting with modifications to the group method to remove this bias.

Nearest neighbour

The nearest-neighbour method produced inaccuracies in several areas. The underlying distribution was insensitive to detecting differences in affiliation strength across classes. In fact, the underlying distribution measure was worse than the nearest neighbour observers' sample at detecting the sex difference. The nearest neighbour observers did not take a measure from a target if there was no other organism within 30 spaces. There was no such modifier in the underlying distribution method. When there was a difference in assortment between the classes where one group was clustered and one was more dispersed it was more likely for the dispersed class to have another member of the same class as a nearest neighbour. This also caused inaccuracies when one class was clustered and the other was moving randomly ($s_m = 0.1$, $s_f = 0$). In this case, the nearest neighbour underlying distribution measured the random moving class as assorting. In sum, the critical variable to be considered when using the nearest neighbour technique is the giving up distance, the distance at which neighbours are considered to be assorting (see Hinde, 1977).

Neighbourhood

For neighbourhood methods, focal sampling was systematically less sensitive to detecting effects both within and between classes than was scan sampling. The focal method was influenced by the amount of time spent per individual. Reducing the amount of time spent sampling each individual, while not increasing the overall amount of data collected, increased the focal method's accuracy. The focal observers also took fewer data, which is an issue when recording assortment must be done quickly, as when the subjects are changing over time (Altmann, 1974).

This work stemmed from a need to assess our methods of data collection for studying social behaviour in cowbirds; thus, the virtual organisms tended to move in birdlike patterns. We attempted to make the model more generalizable to other study species by changing the density and activity levels of the organisms. These changes had varying effects on the observed assortment patterns for the different methods, with focal sampling being the least influenced by the different movement parameters.

This simulation did not investigate human biases, which are major concerns for taking social assortment data and some methods are more susceptible to human biases than others. For example the group and scan method would be most influenced by a difference in visibility among subjects. Nearest neighbour and group methods require more assessment of spatial dynamics among individuals and, thus, may be most likely to be impacted by errors in judgment. There are several important works on these subjects already (Altmann, 1974; Hinde, 1977; Cairns & Schwager, 1987; Fragazsy et al., 1992; Whitehead & Dufault, 1999). Future directions of our model will attempt to test some components of human observational biases, as well as issues relating to observer agreement.

Also the model does not assess why the organisms were associating. Our movement functions required organisms to move toward other organisms. It could be that moving to a location produces differing patterns of assortment under the different methods (Whitehead, 1999). Social versus demographic movement will be tested in future simulations, as well as more complex interactions among the classes.

The model has proved to be heuristic in allowing us to think about the strengths and weaknesses of the measures we use in the lab and to begin to create new methods of sampling that will be less biased and most sensitive

to detecting patterns of social behaviour. We recommend the use of agent-based simulations for testing data in situations where real-world control over or knowledge of underlying distributions is impossible to attain.

Acknowledgements

We wish to acknowledge Meredith J. West, Andrew P. King for their support and insights during the development of this work. Jeff Schank provided comments during the creation of the simulation. Research was supported by NSF, an HHMI predoctoral fellowship for V.A.S. and an NSERC post-doctoral fellowship for D.J.W.

References

Altmann, J. (1974). Observational study of behavior: sampling methods. — Behaviour 48: 227-265.

Bejder, L., Fletcher, D. & Bräger, S. (1998). A method for testing association patterns of social animals. — Anim. Behav. 56: 719-725.

Boinski, S. (1994). Affiliation patterns among male Costa Rican squirrel monkeys. — Behaviour. 130: 191-209.

Boinski, S. & Mitchell, C.L. (1994). Male residence and association patterns in Costa Rican squirrel monkeys (*Saimiri oerstedi*). — Am. J. Primatol. 34: 157-169.

Bryson, J.J., Ando, Y. & Lehmann, H. (2007). Agent-based modeling as scientific method: a case study analysing primate social behaviour. — Philos. Trans. Roy. Soc. Lond. B Biol., Epub ahead of print.

Cairns, S.J. & Schwager, S.J. (1987). A comparison of association indices. — Anim. Behav. 35: 1454-1469.

Clapham, P.J. (1994). Maturational changes in patterns of association in male and female humpback whales, *Megaptera novaeangliae*. — J. Zool. 234: 265-274.

Connor, R.C., Smolker, R. & Bejder, L. (2006). Synchrony, social behaviour, and alliance information in Indian Ocean bottlenose dolphins, *Tursiops aduncus*. — Anim. Behav. 72: 1371-1378.

Digby, L.J. (1995). Social organization in a wild population of *Callithrix jacchus*: II. Intragroup social behavior. — Primates 36: 361-375.

Durrell, J.L., Sneddon, I.A., O'Connell, N.E. & Whitehead, H. (2004). Do pigs form preferential associations? — Appl. Anim. Behav. Sci. 89: 41-52.

Dwyer, C.M. & Lawrence, A.B. (1999). Ewe-ewe and ewe-lamb behaviour in a hill and lowland breed of sheep: A study using embryo transfer. — Appl. Anim. Behav. Sci. 61: 319-334.

Félix, F. (1997). Organization and social structure of the coastal bottlenose dolphin *Tursiops truncatus* in the Gulf de Guayaquil, Ecuador. — Aquat. Mammal. 23: 1-16.

Fragaszy, D.M., Boinski, S. & Whipple, J. (1992). Behavioral sampling in the field: comparison of individual and group sampling methods. — Am. J. Primatol. 26: 259-275.

Frederick, H. & Johnson, C.N. (1996). Social organization in the Rufous bettong, *Aepyprymnus rufescens.* — Aust. J. Zool. 44: 9-17.

Freeberg, T.M. (1999). Spatial associations provide a context for social learning of courtship patterns in brown-headed cowbirds (*Molothrus ater*). — J. Comp. Psychol. 113: 327-332.

Hinde, R.A. (1976). Interactions, relationships and social structure. — Man 11: 1-17.

Hinde, R.A. (1977). On assessing the bases of partner preferences. — Behaviour 62: 1-9.

Holekamp, K.E., Cooper, S.M., Katona, C.I., Berry, N.A., Frank, L.G. & Smale, L. (1997). Patterns of association among female spotted hyenas (*Crocuta crocuta*). — J. Mammal. 78: 55-64.

Manly, B.F.J. (1995). A note on the analysis of species co-occurrences. — Ecology 76: 1109-1115.

Mitani, J.C., Merriwether, D.A. & Zhang, C. (2000). Male affiliation, cooperation and kinship in wild chimpanzees. — Anim. Behav. 59: 885-893.

Mollema, L., Koene, P. & de Jong, M.C.M. (2006). Quantification of the contact structure in a feral cattle population and its hypothetical effect on the transmission of bovine herpesvirus 1. — Prevent. Vet. Med. 77: 161-179.

Newton-Fisher, N.E. (1999). Association by male chimpanzees: A social tactic? — Behaviour 136: 705-730.

Reynolds, C.W. (1987). Flocks, herds, and schools: a distributed behavioral model — Comput. Graph. 21: 25-34.

Schank, J.C. & Alberts, J.R. (2000). The developmental emergence of coupled activity as cooperative aggregation in rat pups. — Proc. Roy. Soc. Lond. B Biol. 267: 2307-2315.

Silk, J.B., Alberts, S.C. & Altmann, J. (2003). Social bonds of female baboons enhance female survival. — Science 302: 1231-1234.

Smith, V.A., King, A.P. & West, M.J. (2002). The context of social learning: association patterns in a captive flock of brown-headed cowbirds (*Molothrus ater*). — Anim. Behav. 63: 32-35.

Weinrich, M.T., Rosenbaum, H., Baker, C.S., Blackmer, A.L. & Whitehead, H. (2006). The influence of maternal lineages on social affiliations among humpback whales (*Megaptera novaeangliae*) on their feeding grounds in the southern Gulf of Maine. — J. Hered. 97: 226-234.

West, M.J., White, D.J. & King, A.P. (2002). Female brown-headed cowbirds' (*Molothrus ater*) organization and behaviour reflects male social dynamics. — Anim. Behav. 64: 377-384.

West-Eberhard, M.J. (1983). Sexual selection, social competition and speciation. — Q. Rev. Biol. 58: 155-183.

White, D.J., King, A.P. & West, M.J. (2002a). Facultative development of courtship and communication in cowbirds, *Molothrus ater*. — Behav. Ecol. 13: 487-496.

White, D.J., King, A.P. & Duncan, S.D. (2002b). Voice recognition technology as a tool for behaviour research. — Behav. Res. Methods Instr. Comput. 34: 1-5.

Whitehead, H. (1997). Analysing animal social structure. — Anim. Behav. 53: 1053-1067.

Whitehead, H. (1999). Testing association patterns of social animals. — Anim. Behav. 57: F26-F29. Available online at: http://www.academicpress.com/anbehav

Whitehead, H. & Dufault, S. (1999). Techniques for analyzing vertebrate social structure using identified individuals: Review and recommendations. — Adv. Study Behav. 28: 33-74.