LETTER

Manipulating badges of status only fools strangers

Alexis S. Chaine,[1,2]* Daizaburo Shizuka,[3] Theadora A. Block,[4] Lynn Zhang[4] and Bruce E. Lyon[4]

Abstract
Conflict is risky, but mechanisms that allow animals to assess dominance status without aggression can reduce such costs. Two different mechanisms of competitor assessment are expected to evolve in different contexts: badges of status are expected in larger, anonymous groups, whereas individual recognition is feasible in small, stable groups. However, both mechanisms may be important when social interactions occur both within and across stable social groups. We manipulated plumage in golden-crowned sparrows (*Zonotrichia atricapilla*) and found that two known badges of status – gold and black head plumage patch sizes – independently affect dominance among strangers but manipulations had no effect on dominance among familiar flockmates. Moreover, familiar flockmates showed less aggression and increased foraging relative to strangers. Our study provides clear experimental evidence that social recognition affects badge function, and suggests that variation in social contexts maintains coexistence and context-dependent use of these two dominance resolution mechanisms.

Keywords
Badge of status, individual recognition, social dominance, social recognition, social structure.

Ecology Letters (2018) **21**: 1477–1485

INTRODUCTION

Conflict over resources is ubiquitous among organisms. When the fitness stakes are high, conflict can entail extreme aggression and even lethal fights, for example, when males battle for access to females in polygynous mating systems. In contrast, when organisms compete over resources with modest effects on fitness, such as a meal, animals often settle conflicts with mechanisms that avoid the costs of overt aggression and injury (Maynard Smith & Price 1973; Maynard Smith & Harper 2003). Two such mechanisms for competitor assessment have been proposed: individual recognition and badges of status (Rohwer 1975; Barnard & Burk 1979; Maynard Smith & Harper 2003; Searcy & Nowicki 2005; Dale 2006; Tibbetts & Dale 2007; Bradbury & Vehrencamp 2011). Individual recognition allows an organism to tailor its behaviour in aggressive interactions with familiar opponents based on past interactions (Barnard & Burk 1979; Tibbetts & Dale 2007) or through observations of interactions between other individuals of known rank (i.e. transitive inference: Paz-Y-Miño C *et al.* 2004; Grosenick *et al.* 2007). In contrast, badges of status are signals that correlate with an individuals' fighting ability that allow any random pair of individuals to assess their chances of winning a given contest and therefore avoid costly fights (Rohwer 1975, 1977, 1982). Both mechanisms are supported by ample experimental evidence in the literature (Dale *et al.* 2001; Whiting *et al.* 2003; Senar 2006; Tibbetts & Dale 2007), but there is relatively little data on whether and under what circumstances these two mechanisms might co-occur (Sheehan & Bergman 2016b; but see Järvi & Bakken 1984; Whitfield 1988; Sheehan & Tibbetts 2010; Sheehan *et al.* 2014).

Social structure – particularly group size and stability – should affect whether individual recognition or badges of status are favoured as assessment strategies (Fig. 1) (Rohwer 1982; Whitfield 1987; Shultz & Gersick 2016). Individual recognition is thought most likely to occur in small, stable groups for two reasons: the number of individuals an animal can remember may be constrained (Dukas 1999; Burns *et al.* 2011), and the benefits of memorising individual identities and adjusting dominance behaviour accordingly are greater when the same individuals repeatedly interact (Fig. 1a) (Whitfield 1987; Ydenberg *et al.* 1988; Johnstone 1997; Tibbetts & Dale 2007). Conversely, badges of status do not require memory of past encounters, and thus could function in large groups where most interactions occur between newly interacting individuals (Rohwer 1975, 1982; Whitfield 1987) (Fig. 1c). Taken together, individual recognition should be favoured in societies that feature small stable groups, while badges of status should be favoured in systems with large groups where repeated contests are relatively rare, and also where individual recognition may not be feasible (Fig. 1a vs. c) (Rohwer 1982; Whitfield 1987; Shultz & Gersick 2016). This contrast in conditions has sometimes led to the perception that the two conflict resolution mechanisms are unlikely to coexist.

The debate over assessment strategies has often implied that the social organisation of a species can be neatly categorised into the two extremes – small and stable vs. large and fluid – that each favours different assessment strategies (Shields 1977; Rohwer 1982). However, the social structure of many animal populations cannot be characterised along a simple gradient of group size or group stability. Fission-fusion dynamics are pervasive in animals, even among species that were not

[1]Station d'Ecologie Théorique et Expérimentale du CNRS (UMR5321), Evolutionary Ecology Group, 2 route du, CNRS, 09200 Moulis, France
[2]Institute for Advanced Studies in Toulouse, Toulouse School of Economics, 21 allée de Brienne, 31015 Toulouse, France
[3]School of Biological Sciences, University of Nebraska-Lincoln, NE 68588, USA
[4]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95064, USA
*Correspondence: E-mail: alexis.chaine@cnrs.fr

(a) Small, stable social clusters: social recognition
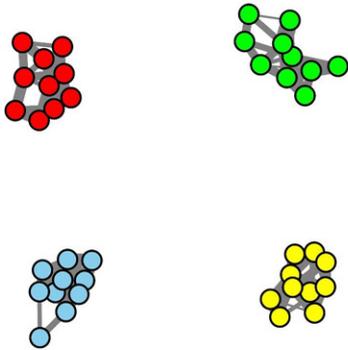
(b) Small, stable social clusters that interact with extra-group individuals: social recognition (familiar) and badges of Status (unfamiliar)
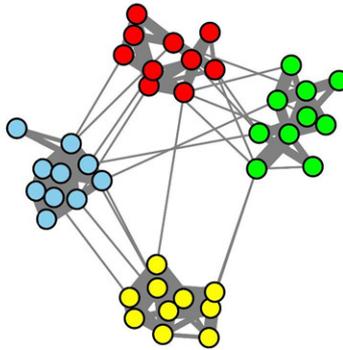
(c) Large anonymous group: badges of status
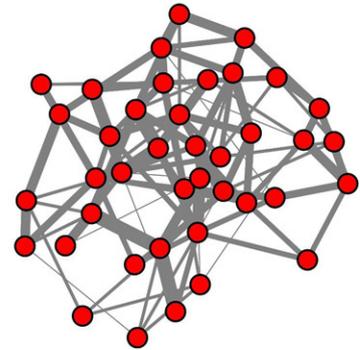
Figure 1 Schematic representation of social networks predicted to favour different dominance-resolving mechanisms. (a) Social (individual or group) recognition is expected to be favoured in smaller, stable groups where each individual only interacts with a few others on a regular basis but only rarely encounters new or unfamiliar individuals. (b) Intermediate networks that include both frequent, repeated interactions within a stable group and some interactions with extra-group individuals could favour the coexistence of both badges and social recognition. (c) Badges of status alone should be expected in fluid groups where individuals are likely to encounter a large number of other individuals so that the costs of remembering this large number outweighs the relative low benefit from rare repeat encounters. In each figure, circles represent individuals (e.g. nodes), lines represent the frequency of interaction between two individuals (e.g. edges) and colours represent social communities that are defined by high within-community interaction and low among-community interaction.

previously considered to be socially complex (Aureli *et al.* 2008; Mourier *et al.* 2012; Silk *et al.* 2014), and social network analysis can help uncover these complex patterns of social stability (Wey *et al.* 2008; Pinter-Wollman *et al.* 2014; Silk *et al.* 2014). For example, stable long-term associations can exist in clusters, or 'social communities' within social networks, but individuals in different social communities may nevertheless be connected weakly through transient interactions. In such multilevel societies, stability of social associations can be heterogeneous, where individuals often have encounters with both familiar *and* unfamiliar competitors – that is, contexts in which different dominance assessment mechanisms may be favoured (e.g. Fig. 1b) (Whitfield 1987; Shultz & Gersick 2016). Even within stable social communities, new recruits will be initially unfamiliar and may use badges of status to establish dominance relationships prior to making use of individual recognition once established in the group. In such cases, the function of badges of status would be context-dependent: they may play an important role in mediating interactions between unfamiliar individuals but may be less important between familiar individuals that can assess each other through individual recognition.

In this study, we investigated whether the use of badges of status depends on the familiarity of contestants in a winter population of the migratory golden-crowned sparrow (*Zonotrichia atricapilla*). In previous work, we provided evidence for multiple badges of status: contests with both natural variation (Chaine *et al.* 2011) and experimental manipulation (Chaine *et al.* 2013) showed that individuals with larger gold or larger black plumage patches win contests among strangers (i.e. individuals from different populations). However, social network analysis based on flock co-membership in a free-living population revealed that these sparrows live in small, stable winter

social groups that are connected together by relatively weaker associations across groups (Shizuka *et al.* 2014). Moreover, these groups are stable across years and many individuals form affiliations that last many years; these affiliations reform anew each year after migration to and from the breeding grounds. The fact that a significant amount of variation in flocking patterns is explained by the effect of the previous year's affiliation, even after controlling for the amount of overlap in home ranges (Shizuka *et al.* 2014), indicates that these birds likely show individual recognition in some aspects of their social behaviour. We already know that sparrows use badges in some dominance contexts; the golden-crowned sparrow social network may provide conditions under which both badges of status and individual recognition are beneficial in resolving social dominance.

Empirical tests of individual recognition and badges of status adopt somewhat different methods; showing that both occur in the same population requires careful experiments that control for familiarity, the context that should determine which mechanism the birds use. Badges of status are often documented through correlations between the size of a putative signal and dominance rank [e.g. lizards (Thompson & Moore 1991), insects (Tibbetts & Dale 2004), birds (Senar 2006; Tibbetts & Safran 2009) and mammals (Setchell & Wickings 2005)]. Confirmation that a specific trait is used as a badge requires manipulation of the presumed badge, which should trigger a change in the outcome of specific contests (Rohwer & Rohwer 1978; Pryke & Andersson 2003; Chaine *et al.* 2013). Experimental contests to test for badges of status typically involve unfamiliar contestants to remove the potential effect of prior information on the contest outcome. In contrast, individual recognition is documented in one of three ways: by comparing interactions between familiar and unfamiliar individuals (Hazlett 1969;

Caldwell 1985; Watt 1986), by measuring changes over time in the behaviours and interactions of a pair of individuals that become increasingly familiar with each other (Karavanich & Atema 1998; Tibbetts 2002; Vedder *et al.* 2010) or, more rarely, through manipulation of a putative identity signal (Whitfield 1988; Tibbetts 2002). Here, we investigated the potential co-occurrence of badges of status and social recognition by conducting experimental badge size manipulations in two types of contests – dyads of birds that are familiar with each other and dyads of unfamiliar birds. This design allowed us to determine if the badges that we have already shown to be effective among unfamiliar birds (Chaine *et al.* 2011, 2013) are also used among familiar birds or, alternatively, if these badges are ignored as would be expected if individuals use some form of social recognition (individual or group). We provide definitive evidence that sparrows use both badges of status and social recognition and demonstrate that variation in social context plays a key role in the coexistence of these two dominance-resolving mechanisms. Because our experimental approach was not designed specifically to discriminate between individual and group recognition, we could not directly test for individual recognition, the form of recognition most discussed in the extensive literature on badges of status (e.g. Rohwer 1975; Whitfield 1987). However, several considerations suggest that the social recognition we observed involved individual recognition, as we discuss below.

MATERIALS AND METHODS

We used staged competition in aviaries to assess whether previous familiarity affected the role of plumage traits on social dominance. We caught golden-crowned sparrows at four different sites in January and February using Potter traps baited with millet seed. From 2007 to 2009, we trapped birds at two locations near Pogonip City Park (36°59′25.6″ N 122°02′12.9″ W and 36°59′04.2″ N 122°02′49.0″ W) and the University of California Santa Cruz (UCSC) campus (36°58′40.1″ N 122°03′04.8″ W). In 2012–2014, we added a site at the Santa Cruz Waldorf School (37°00′22.4″ N 122°04′26.7″ W). Birds were transported to the UCSC Arboretum soon after capture. We measured all birds (mass, wing chord, tarsus, beak length), banded them with colour bands and a USFWS metal band for identifying individuals within experiments, and photographed their crowns as in past studies (Chaine *et al.* 2011, 2013). All birds were then placed in one of two outdoor aviaries for habituation before use in experimental trials the following morning – these two aviaries were visually and acoustically isolated from each other, had shelter, and both food and water *ad libitum*. The two birds used in each trial were put in separate overnight aviaries to ensure a lack of contact immediately prior to the trial. Before running the trials, we measured size of gold and black crown patches (mm^2) from photos with Adobe Photoshop CS and these patch sizes were then used to assign birds to a trial such that the two birds had similar pre-manipulation patch size, as in previous work (Chaine *et al.* 2011, 2013). Sex was not used as a criterion for assigning birds to a trial because previous work showed that gender does not predict winter social dominance in this species (Chaine *et al.* 2011).

We conducted two different types of social dominance trials. 'Unfamiliar trials' ($N = 62$) used birds caught on the same day from two different populations > 800 m apart (often much greater distances). This distance is much larger than the average home range of individual sparrows (Shizuka *et al.* 2014) so the birds in the trial were very unlikely to have interacted with each other prior to the trial. The 'Familiar trials' involved birds that were trapped at the same site and within a short time period (< 30 min); based on our detailed understanding of flocking and feeding behaviour from our colour-banded population (Shizuka *et al.* 2014; unpubl. data), these were birds that would be very likely to have been in the same flock when captured (2012–2014 only; $N = 60$). Unfamiliar trials analysed here include some trials also used in a previous publication ($N = 32$) (Chaine *et al.* 2013) as well as new Unfamiliar trials ($N = 30$) conducted at the same time as Familiar trials to control for possible year effects and collect new data on distance between birds not gathered in the earlier experiments (see below).

Social dominance trials were conducted in the morning between 8 and 11 AM in a third contest aviary (same dimension as overnight aviaries) and each trial lasted 30 min. Food was removed from the overnight aviaries at sundown the evening before the trials so birds in dominance trials were motivated to access the single feeding tray at the centre of the contest aviary. Prior to the trial, one of the birds, chosen at random, had one of their colour patches (either black or gold) increased in size while the other received a sham manipulation using clear thinner (gold) or brown permanent marker (black). Gold patch size was increased using acrylic paint while black crown size was increased using a permanent ink pen; both colours were spectrally similar to natural colour patches (Chaine *et al.* 2013). As in our previous experiment (Chaine *et al.* 2013), the birds in each dyad did not differ in patch size before manipulation and the manipulation significantly increased badge size (Supporting Information). Because pre-manipulation patch size of the two birds was similar, we would expect manipulated birds to be dominant in half of the trials if patch size manipulation had no influence on dominance (Chaine *et al.* 2011, 2013). As in our previous experiment, gold ($N = 62$) and black ($N = 60$) manipulations showed similar outcomes, so we present combined results in the main text (results split by treatment in Fig. S1).

We scored social behaviour in two ways. First, an observer watched trials from ~30 m away with binoculars and scored social interactions as aggression (peck, supplant), avoidance (subordinate avoided challenging a dominant) or tolerance (no aggression while in close proximity), and we also tallied feeding behaviours (Chaine *et al.* 2011, 2013) for the first and second 15-min time periods separately. We used the sum of the social interactions to determine which bird was dominant or, in some cases, to conclude that dominance was not clear (Chaine *et al.* 2011, 2013). We concluded that dominance was clear when one bird consistently won encounters through supplants or avoidance behaviour (Chaine *et al.* 2011, 2013). In contrast, dominance was not clear when birds had the opportunity to interact (i.e. they moved around), but either no dominance behaviours occurred or both birds won an equal number of encounters (very rare). In a second approach, we
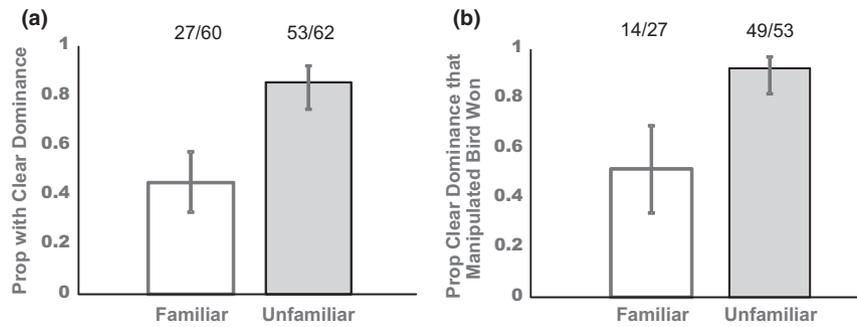
Figure 2 Comparisons between Familiar (open) and Unfamiliar (grey) trials in the proportion of trials where (a) a dominance was clear and (b) where the manipulated bird was dominant among trials where dominance was clear. Confidence intervals (95%) were calculated using Wilson's procedure described by Newcombe (1998).

scored tolerance behaviour as a measure of flocking and social affinity from videos taken from above the contest aviary. Generally, tolerance is a measure of whether birds socialise at close proximity without any aggression, and we quantified this as the number and duration of instances where the birds were within two body lengths of each other for more than 3 seconds and did not show aggressive behaviours (i.e. no dominance interactions; 2012–2014 only). The person who conducted the trials was not blind to treatment but dominance was evaluated as the sum of behaviours post data collection. In contrast, videos used to quantify tolerance and social clustering were scored blind to treatment which should eliminate observer bias. The results of the two approaches are concordant.

To evaluate the impact of our plumage manipulation on social dominance, and particularly to be able to compare across the two sets of experiments, we examined two different outcomes: the proportion of trials that showed a clear dominance winner and then the proportion of this subset of trials where the manipulated (i.e. enlarged) bird was dominant. For the latter, the null expectation is that each bird has a 50% chance of winning so we evaluated the impact of the manipulation on dominance as the deviation from an expected 50 : 50 ratio using a binomial test. This null expectation applies because pre-manipulation badge sizes were matched for each dyad and because badge size predicted dominance in past studies (Chaine *et al.* 2011, 2013). In contrast, there is no clear null expectation of what proportion of trials should show clear dominance, so we simply provide the raw data for evaluation. We used Z-tests to compare the differences between Familiar and Unfamiliar treatments in the probability of observing clear dominance, as well as the probability of the bird with the experimentally enlarged patch winning the contest (Fleiss *et al.* 2003). We used ANOVA to compare rates of each behaviour (aggression, avoidance, tolerance and feeding) between Familiar and Unfamiliar trials. Data were analysed in R 3.1.3 (R Development Core Team 2009) and are available on Figshare.

All protocols were approved by the UCSC IACUC (Animal Welfare Permit Number A3859-01 to B. Lyon). Capture and banding was done under permits to B. Lyon: a California Department of Fish and Wildlife Scientific Collecting

permit, a Federal Fish and Wildlife Permit and a Federal Banding Permit.

RESULTS

Familiarity affected the likelihood of a trial showing clear dominance patterns. Trials with unfamiliar birds (i.e. birds caught at different sites) nearly always had a clear dominant individual (53 of 62 trials). In contrast, only about half of the familiar dyads (birds caught at the same site at the same time) had clear dominance (27 of 60 trials). The proportion of trials where dominance was clear was significantly greater in Unfamiliar trials than in Familiar trials (comparison of two probabilities: $Z = 4.71$, $P < 0.0002$, Fig. 2a).

The plumage manipulations influenced the outcome of dominance in trials involving unfamiliar birds but not familiar birds. In the subset of trials where dominance was clear, the manipulated bird was dominant over the sham individual in a significantly larger proportion of trials between unfamiliar birds relative to trials between familiar birds (comparison of two probabilities: $Z = 4.20$, $P < 0.0002$, Fig. 2b). Indeed, the manipulated bird was dominant in nearly all cases during Unfamiliar trials (49 of 53 trials, binomial test based on null expectation of 0.5 probability that an individual will be dominant based on pre-manipulation crown size: $P < 0.001$). The sample here includes trials conducted in our previous study using the same methods (Chaine *et al.* 2013), and thus simply increases the robustness of the previously reported patterns. In contrast, the manipulated bird was not more likely to be dominant than the sham individual among familiar dyads, which is what we would expect only if birds ignored the patch manipulation (14 of 27 trials, binomial test with null expectation of 0.5 probability that individual will be dominant based on pre-manipulation crown size: $P = 1$). These results are the same regardless of which plumage patch was manipulated (i.e. black or gold: Fig. S1).

Birds in Familiar trials had less costly social behaviours and spent more time feeding relative to unfamiliar dyads. Familiar trials had fewer aggressive interactions (ANOVA: $F_{1,120} = 3.46$, $P = 0.065$; Fig. S2a), fewer avoidance interactions (ANOVA: $F_{1,120} = 24.63$, $P < 0.0001$; Fig. 3a), and a higher rate of tolerance (a lack of aggression when close; ANOVA: $F_{1,120} = 96.50$, $P < 0.0001$; Fig. S2b). Familiar birds also spent significantly
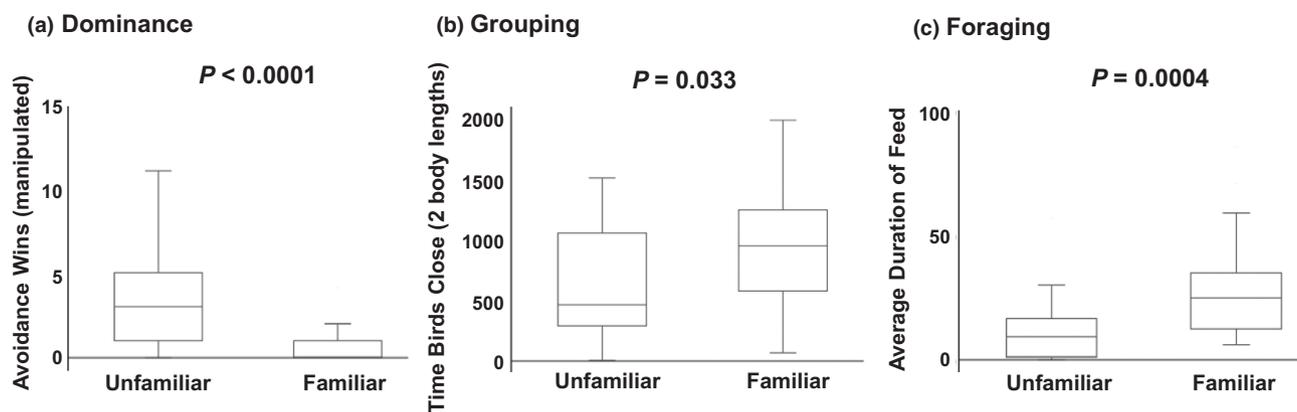
Figure 3 Comparison between Unfamiliar and Familiar trials in behaviours that reflect costs and benefits of grouping behaviour when dominance resolution occurs by recognition. (a) The number of dominance interactions where a subordinate avoided challenging a dominant (avoidance). (b) The total time the two birds spent very near each other without showing aggression (within two body lengths). (c) The average duration of feeding bouts. Figures show standard box and whisker plots.

more time in close proximity relative to unfamiliar dyads (ANOVA: $F_{1,57} = 4.77$, $P = 0.033$; Fig. 3b). Finally, while birds in both types of trials accessed food at similar rates (ANOVA: $F_{1,120} = 1.31$, $P = 0.25$; Fig. S2c), birds in Familiar trials spent a significantly longer time at the feeder on each visit compared to birds in Unfamiliar trials (ANOVA: $F_{1,101} = 13.07$, $P = 0.0004$; Fig. 3c).

DISCUSSION

Our experiments revealed that plumage manipulations affect dominance status in staged contests among strangers (confirming the results of Chaine *et al.* 2013), but not among birds previously familiar with each other. These contrasting results provide evidence that the familiar birds recognised each other, either individually or at the group level, and that this recognition allowed them ignore experimental changes to plumage badges. This also implies that the crown plumage patches we manipulated were not the traits the birds use for social recognition. Finally, our experiments suggest how familiar birds can benefit from social recognition during such encounters – reduced aggression, tighter flocking and increased access to resources.

The differences we observed between the familiar and unfamiliar trials imply some form of recognition among familiar birds. We did not quantify the amount of prior social interaction leading to this social recognition, nor did our experiments directly test for individual recognition so there is a question of what type of recognition was involved – individual recognition or group-level recognition (e.g. familiar vs. unfamiliar)? Regardless of the level of social recognition, our results indicate that birds caught in the same place within a short time – that is, those that are typically very socially familiar due to close associations over long periods of time (Shizuka *et al.* 2014) – disregard manipulation of badges of status. However, distinguishing between these two forms of recognition can be difficult and direct evidence for individual recognition requires demonstrating that individuals vary their behaviour in response to an individual specific cue like an individually distinct vocalisation or visual features, and not some group-level attribute

(Tibbetts *et al.* 2008). For example, group recognition is common in social insects where individuals respond differently to familiar (same colony) and unfamiliar (different colony) individuals on the basis of colony specific olfactory cues (cuticular hydrocarbons, Hölldobler & Wilson 1990). Thus, simply showing that these insects respond differently to familiar and unfamiliar individuals may only show a group level response and not individual recognition. In birds, there is some evidence for group-specific calls in various contexts, including seasonally stable flocks (e.g. Nowicki 1983; Hausberger *et al.* 1995). However, we never observed vocalisations in any of our trials, so acoustic features cannot explain our results. Furthermore, our past work suggests that recognition between familiar sparrows generally goes beyond group membership and reflects true individual recognition. When winter social communities form anew each fall after fall migration, variation in the strength of pairwise affiliations is explained in large part by the degree of affiliations between those individuals in the previous year, even after controlling for shared space use by individuals, which strongly suggests individual recognition (Shizuka *et al.* 2014). Based on these considerations, we believe that individual recognition is the most likely explanation for the differences between familiar and unfamiliar trials. Identifying the trait or traits used in individual recognition remains an interesting, if challenging, question for future work.

Our study demonstrates how both badges of status and social recognition can mediate social competition within the same population. Prior work shows that badges of status and recognition (in particular, individual recognition) each provide a way to avoid costly conflict over resources, yet each has different benefits and constraints. For example, badges of status may allow rapid assessment of competitive ability and can function reliably even when groups are large and fluid and individual recognition is not feasible (Rohwer 1982). However, badges can, in theory, be susceptible to cheating (Rohwer & Ewald 1981). In contrast, individual recognition could allow for more reliable assessment of competitive ability, but it is only beneficial when there are repeated interactions between competitors (Ydenberg *et al.* 1988; Johnstone 1997; Tibbetts 2004; Tibbetts & Dale 2007), and it may

require investment in neural mechanisms to allow for memory of the outcomes of past interactions associated with each contestant (Dukas 1999; Tibbetts & Dale 2007; Sheehan & Tibbetts 2011). There also may be a limit to the size of the social group where individual recognition is cognitively feasible (Dunbar 1992). Traditionally, researchers have sought evidence for either badges of status or individual recognition for a given species, and evidence for one type of assessment strategy was held as evidence against the other (e.g. Shields 1977). Our study provides valuable empirical data to show that conditions favourable for both types of assessment strategies coexist in one social system, and that golden-crowned sparrows selectively use badges of status or social recognition depending on the social context. These findings should spur further progress in development of theory related to the evolution of competitor assessment strategies, which has been highlighted in a recent review by Sheehan & Bergman (2016b) and responding commentary (Barrett & Henzi 2016; Ridley *et al.* 2016; Sheehan & Bergman 2016a; Shultz & Gersick 2016).

We suggest that the context-dependent use of both badges of status and individual recognition in the same species may be more common than currently suspected. In a handful of species, both mechanisms of assessment have been tested but only one was found – for example, the various *Polistes* wasps studied in the impressive body of work by Tibbetts and colleagues (Tibbetts 2004; Tibbetts & Dale 2004; Sheehan & Tibbetts 2010). However, the results are more mixed in avian systems, with a few cases that hint at the co-occurrence of badges and recognition. For example, blue tits (*Parus caeruleus*) use badges of status (Rémy *et al.* 2010), but also show changes in the importance of badges for dominance across repeated interactions suggestive of individual recognition (Vedder *et al.* 2010), although winner/loser effects cannot be ruled out for this pattern (Hsu *et al.* 2006; Rutte *et al.* 2006). Likewise, white-crowned sparrows (*Zonotrichia leucophrys*), a congener of the sparrows we studied, use badges of status (Fugle *et al.* 1984), yet our reanalysis of experiments by Watt (1986) show that interactions between familiar birds show fewer aggressive interactions than between unfamiliar birds [reanalysis of data in Table 3 in (Watt 1986): comparison of the proportion of aggressive interactions between familiar and unfamiliar birds: $Z = -2.67$, $P = 0.008$]. This difference between familiar and unfamiliar birds in levels of aggression suggests that individual recognition may exist in white-crowned sparrows. Parsons & Baptista (1980) also showed that white-crowned sparrows with manipulated crowns maintain status within their flock, further suggestive of individual recognition. For these two sparrow species, careful experiments are now required to confirm that individual recognition occurs.

An open question is whether there are predictable conditions under which a species or population might use just one form of assessment or both assessment strategies. Complex social structure may favour the use of both assessment strategies. Four decades ago, Whitfield (1987) hypothesised that social organisation might be sufficiently variable in many species to favour both assessment mechanisms, and lamented the lack of understanding of social organisation as an impediment for understanding the evolution of badges and individual recognition. More recently, Shultz & Gersick (2016) suggested that the dual use of quality signalling and social recognition would be favoured in multilevel animal societies in which close-knit social clusters exist within a larger population because individuals would interact with both familiar and unfamiliar individuals (e.g. Fig. 1b). Our study system provides evidence to support this idea. Golden-crowned sparrows in our population live in a complex social network characterised by cohesive social clusters embedded within a larger population that interacts more loosely (Shizuka *et al.* 2014). This social structure is not driven by kinship (Arnberg *et al.* 2015), but rather preferences for familiar flockmates, even across years (Shizuka *et al.* 2014). Our experiments now show that familiarity not only affects flock associations, but also the way competitors use badges of status to mediate competition. In the natural social network context where interactions may occur with a mix of familiar and unfamiliar competitors, within-cluster (familiar) interactions may be mediated by recognition, whereas interactions among clusters (unfamiliar) may use badges of status, precisely as Whitfield (1987) predicted. The basic features of this social system – fission-fusion dynamics characterised by a mix of stable long-term associations and diffuse short-term associations – are likely widespread in avian systems (Silk *et al.* 2014). Even for systems in which long-term stable social relations dominate, we suggest that there are likely periods of time during which interactions among unfamiliar individuals occur and badge use would be favoured, such as during migration, the integration of new recruits into stable social groups, or periods of social instability following the breakdown of some groups. A better understanding of the natural patterns of social interactions throughout the annual cycle for birds generally may reveal common patterns that favour the use of one or both mechanisms of competitor assessment.

Our study demonstrates how controlling the social context of contests is paramount to accurately understand the dynamics of status signalling. This poses a challenge for researchers because it also requires a detailed understanding of the social structure of the population – that is, the extent to which social interactions are naturally confined to familiar vs. unfamiliar competitors. For example, earlier observational studies of badges of status suggested that the failure of putative badges to accurately predict dominance cast doubt on whether they served as badges (e.g. Whitfield 1987). However, it is also possible that those studies contained a large proportion of within group interactions where individual recognition rather than badges of status mediated dominance. Moreover, testing the function of putative badges through observations in the wild, rather than with experiments, may often produce weaker correlations between badge and dominance when interactions involve a mix of strangers and familiar birds. Our results suggest that the relationship between putative badges and dominance in the wild will depend critically on the frequency of interactions that are within groups of familiar individuals vs. between strangers, and highlight the importance of distinguishing between these two different contexts.

Our experimental results also have implications for the maintenance of honest signalling. A key question for badges of status has long been, why not cheat? Why do poor fighters

not produce badges that signal high fighting ability? In theory, the stability of badges of status as a signalling system relies on some mechanism to prevent cheating, such as condition dependence (Johnstone & Norris 1993; Tibbetts & Curtis 2007) or social punishment of cheaters (Rohwer 1982; Tibbetts & Dale 2004; Webster *et al*. in press). However, our plumage manipulation with familiar individuals apparently did not result in the rapid social punishment seen in similar experiments in other systems (Tibbetts & Izzo 2010; Ligon & McGraw 2016). This begs the question whether social recognition could change the dynamics of social enforcement of honesty in badges of status. We suggest that, given individual recognition, it may be possible for delayed social punishment to play out at timescales beyond the duration of a given dominance experiment. In theory, the value of a dishonest badge should be diminished in systems where both badges and individual recognition are used, because interactions within a social group no longer rely only on badges of status. Specifically, badges of status might be used to settle dominance when a bird initially joins a group, but recognition subsequently modulates interactions. If detection makes it less likely that cheaters can join a group, or if they subsequently develop weaker or more negative affiliations (i.e. individuals reduce interactions with cheaters or affiliations involve more hostility), then social costs linked to individual recognition could guard against cheating. For animals that live in stable long-lived groups, the presence of such 'punishment' could have dire long-term consequences such as reduced cooperation or even exclusion from a group. If there are important benefits to stable group membership the risks of dishonest status signalling in the presence of individual recognition may make it unlikely.

Finally, our experiments also demonstrate potential benefits of interacting with familiar individuals, such as reduced aggression and increased food intake which are key components of winter survival. Birds in Familiar trials spent much less time in conflict and more time in non-aggressive (tolerant) interactions relative to those in Unfamiliar trials. Familiar birds also spent much longer times at the feeder than unfamiliar birds, which in general would increase their food intake. Finally, familiar birds spent much more time close to each other than unfamiliar birds and such clustering could provide benefits such as reduced predation risk (Hamilton 1971; Lima *et al*. 1999). These benefits, which result from recognition rather than badges of status, may be key components that favour the stable community structure we see in golden-crowned sparrows (Shizuka *et al*. 2014). Thus, our experimental results may provide a key link between competitor assessment strategies and the social network structure of foraging flocks. The cues used for recognition – which we have not yet identified – may play an outsized role in the winter social system of golden-crowned sparrows by mediating flock membership as well as modulating the function of badges of status in contests over food.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

ASC, DS and BEL designed research. ASC, TAB, LZ and BEL conducted trapping and aviary experiments. ASC performed analyses. ASC and BEL wrote the initial manuscript and all authors contributed to revisions.

REFERENCES

Arnberg, N.N., Shizuka, D., Chaine, A.S. & Lyon, B.E. (2015). Social network structure in wintering golden-crowned sparrows is not correlated with kinship. *Mol. Ecol.*, 24, 5034–5044.

Aureli, F., Schaffner, C.M., Boesch, C., Bearder, S.K., Call, J., Chapman, C.A. *et al*. (2008). Fission-fusion dynamics: new research frameworks. *Curr. Anthropol.*, 49, 627–654.

Barnard, C.J. & Burk, T. (1979). Dominance hierarchies and the evolution of "individual recognition". *J. Theor. Biol.*, 81, 65–73.

Barrett, L. & Henzi, S.P. (2016). Understanding antagonism: a comment on Sheehan and Bergman. *Behav. Ecol.*, 27, 17–18.

Bradbury, J.W. & Vehrencamp, S.L. (2011). *Principles of Animal Communication*, 2nd edn. Sinauer Associates Inc, Sunderland, Mass.

Burns, J.G., Foucaud, J. & Mery, F. (2011). Costs of memory: lessons from 'mini' brains. *Proc. R. Soc. Lond. B Biol. Sci.*, 278, 923–929.

Caldwell, R.L. (1985). A test of individual recognition in the stomatopod Genodactylus festae. *Anim. Behav.*, 33, 101–106.

Chaine, A.S., Tjernell, K.A., Shizuka, D. & Lyon, B.E. (2011). Sparrows use multiple status signals in winter social flocks. *Anim. Behav.*, 81, 447–453.

Chaine, A.S., Roth, A.M., Shizuka, D. & Lyon, B.E. (2013). Experimental confirmation that avian plumage traits function as multiple status signals in winter contests. *Anim. Behav.*, 86, 409–415.

Dale, J. (2006). Intraspecific variation in coloration. In *Bird coloration 2: function and evolution* (eds Hill, G., McGraw, K.). Harvard University Press, Cambridge, MA, pp. 36–86.

Dale, J., Lank, D.B. & Reeve, H.K. (2001). Signaling individual identity versus quality: a model and case studies with ruffs, queleas, and house finches. *Am. Nat.*, 158, 75–86.

Dukas, R. (1999). Costs of memory: ideas and predictions. *J. Theor. Biol.*, 197, 41–50.

Dunbar, R.I.M. (1992). Neocortex size as a constraint on group size in primates. *J. Hum. Evol.*, 22, 469–493.

Fleiss, J.L., Levin, B. & Paik, M.C. (2003). Chapter 3: Assessing Significance in a Fourfold Table. In: *Statistical Methods for Rates and Proportions*. John Wiley & Sons, Inc., Hoboken, New Jersey, pp. 50–63.

Fugle, G.N., Rothstein, S.I., Osenberg, C.W. & McGinley, M.A. (1984). Signals of status in wintering white-crowned sparrows, Zonotrichia leucophrys gambelii. *Anim. Behav.*, 32, 86–93.

Grosenick, L., Clement, T.S. & Fernald, R.D. (2007). Fish can infer social rank by observation alone. *Nature*, 445, 429–432.

Hamilton, W.D. (1971). Geometry for the selfish herd. *J. Theor. Biol.*, 31, 295–311.

Hausberger, M., Richard-Yris, M.-A., Henry, L., Lepage, L. & Schmidt, I. (1995). Song sharing reflects the social organization in a captive group of European starlings (Sturnus vulgaris). *J. Comp. Psychol.*, 109, 222–241.

Hazlett, B.A. (1969). "Individual" recognition and agonistic behaviour in pagurus bernhardus. *Nature*, 222, 268–269.

Hölldobler, B. & Wilson, E.O. (1990). *The Ants*. Harvard University Press, Harvard, MA.

Hsu, Y., Earley, R.L. & Wolf, L.L. (2006). Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biol. Rev. Camb. Philos. Soc.*, 81, 33–74.

Järvi, T. & Bakken, M. (1984). The function of the variation in the breast stripe of the great tit (Parus major). *Anim. Behav.*, 32, 590–596.

Johnstone, R.A. (1997). Recognition and the evolution of distinctive signatures: when does it pay to reveal identity? *Proc. R. Soc. B-Biol. Sci.*, 264, 1547–1553.

Johnstone, R.A. & Norris, K. (1993). Badges of status and the cost of aggression. *Behav. Ecol. Sociobiol.*, 32, 127–134.

Karavanich, C. & Atema, J. (1998). Individual recognition and memory in lobster dominance. *Anim. Behav.*, 56, 1553–1560.

Ligon, R.A. & McGraw, K.J. (2016). Social costs enforce honesty of a dynamic signal of motivation. *Proc. R. Soc. B*, 283, 20161873.

Lima, S.L., Zollner, P.A. & Bednekoff, P.A. (1999). Predation, scramble competition, and the vigilance group size effect in the dark-eyed junco (Junco hyemalis). *Behav. Ecol. Sociobiol.*, 46, 110–116.

Maynard Smith, J. & Harper, D. (2003). *Animal Signals*. Oxford University Press, New York, NY.

Maynard Smith, J. & Price, G.R. (1973). The logic of animal conflict. *Nature*, 246, 15–18.

Mourier, J., Vercelloni, J. & Planes, S. (2012). Evidence of social communities in a spatially structured network of a free-ranging shark species. *Anim. Behav.*, 83, 389–401.

Newcombe, R.G. (1998). Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat. Med.*, 17, 873–890.

Nowicki, S. (1983). Flock-specific recognition of chickadee calls. *Behav. Ecol. Sociobiol.*, 12, 317–320.

Parsons, J. & Baptista, L.F. (1980). Crown color and dominance in the white-crowned sparrow. *Auk*, 97, 807–815.

Paz-Y-Miño C, G., Bond, A.B., Kamil, A.C. & Balda, R.P. (2004). Pinyon jays use transitive inference to predict social dominance. *Nature*, 430, 778–781.

Pinter-Wollman, N., Hobson, E.A., Smith, J.E., Edelman, A.J., Shizuka, D., de Silva, S. *et al.* (2014). The dynamics of animal social networks: analytical, conceptual, and theoretical advances. *Behav. Ecol.*, 25, 242–255.

Pryke, S.R. & Andersson, S. (2003). Carotenoid-based epaulettes reveal male competitive ability: experiments with resident and floater red-shouldered widowbirds. *Anim. Behav.*, 66, 217–224.

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rémy, A., Grégoire, A., Perret, P. & Doutrelant, C. (2010). Mediating male–male interactions: the role of the UV blue crest coloration in blue tits. *Behav. Ecol. Sociobiol.*, 64, 1839–1847.

Ridley, A.R., Humphries, D.J. & Wiley, E.M. (2016). Is information from both quality signaling and social recognition really redundant? A comment on Sheehan and Bergman. *Behav. Ecol.*, 27, 14–15.

Rohwer, S. (1975). The social significance of avian winter plumage variability. *Evolution*, 29, 593–610.

Rohwer, S. (1977). Status signaling in Harris Sparrows some experiments in deception. *Behaviour*, 61, 107–129.

Rohwer, S. (1982). The evolution of reliable and unreliable badges of fighting ability. *Am. Zool.*, 22, 531–546.

Rohwer, S. & Ewald, P.W. (1981). The cost of dominance and advantage of subordination in a badge signaling system. *Evolution*, 35, 441–454.

Rohwer, S. & Rohwer, F.C. (1978). Status signalling in Harris' sparrows: experimental deceptions achieved. *Anim. Behav.*, 26, 1012–1022.

Rutte, C., Taborsky, M. & Brinkhof, M.W.G. (2006). What sets the odds of winning and losing?. *Trends Ecol. Evol.*, 21, 16–21.

Searcy, W.A. & Nowicki, S. (2005). *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems*. Princeton University Press, Princeton.

Senar, J.C. (2006). Color displays as intrasexual signals of aggression and dominance. In *Bird Coloration 2: Function and Evolution* (eds Hill, G., McGraw, K.). Harvard University Press, Cambridge, MA, pp. 87–136.

Setchell, J.M. & Wickings, E.J. (2005). Dominance, status signals and coloration in male Mandrills (Mandrillus sphinx). *Ethology*, 111, 25–50.

Sheehan, M.J. & Bergman, T.J. (2016a). A quality signaling–recognition trade-off at the level of the type of interaction not species: a response to comments on Sheehan and Bergman. *Behav. Ecol.*, 27, 18–19.

Sheehan, M.J. & Bergman, T.J. (2016b). Is there an evolutionary trade-off between quality signaling and social recognition? *Behav. Ecol.*, 27, 2–13.

Sheehan, M.J. & Tibbetts, E.A. (2010). Selection for individual recognition and the evolution of polymorphic identity signals in Polistes paper wasps. *J. Evol. Biol.*, 23, 570–577.

Sheehan, M.J. & Tibbetts, E.A. (2011). Specialized face learning is associated with individual recognition in paper wasps. *Science*, 334, 1272–1275.

Sheehan, M.J., Straub, M.A. & Tibbetts, E.A. (2014). How does individual recognition evolve? Comparing responses to identity information in polistes species with and without individual recognition. *Ethology*, 120, 169–179.

Shields, W.M. (1977). The social significance of avian winter plumage variability: a comment. *Evolution*, 31, 905–907.

Shizuka, D., Chaine, A.S., Anderson, J., Johnson, O., Laursen, I.M. & Lyon, B.E. (2014). Across-year social stability shapes network structure in wintering migrant sparrows. *Ecol. Lett.*, 17, 998–1007.

Shultz, S. & Gersick, A.S. (2016). The evolution of signaling complexity: a comment on Sheehan and Bergman. *Behav. Ecol.*, 27, 16–17.

Silk, M.J., Croft, D.P., Tregenza, T. & Bearhop, S. (2014). The importance of fission–fusion social group dynamics in birds. *The Ibis*, 156, 701–715.

Thompson, C.W. & Moore, M.C. (1991). Throat color reliably signals status in male tree lizards, Urosaurus ornatus. *Anim. Behav.*, 42, 745–753.

Tibbetts, E.A. (2002). Visual signals of individual identity in the wasp Polistes fuscatus. *Proc. R. Soc. B Biol. Sci.*, 269, 1423–1428.

Tibbetts, E.A. (2004). Complex social behaviour can select for variability in visual features: a case study in Polistes wasps. *Proc. Biol. Sci.*, 271, 1955–1960.

Tibbetts, E.A. & Curtis, T.R. (2007). Rearing conditions influence quality signals but not individual identity signals in Polistes wasps. *Behav. Ecol.*, 18, 602–607.

Tibbetts, E.A. & Dale, J. (2004). A socially enforced signal of quality in a paper wasp. *Nature*, 432, 218–222.

Tibbetts, E.A. & Dale, J. (2007). Individual recognition: it is good to be different. *Trends Ecol. Evol.*, 22, 529–537.

Tibbetts, E.A. & Izzo, A. (2010). Social punishment of dishonest signalers caused by mismatch between signal and behavior. *Curr. Biol.*, 20, 1637–1640.

Tibbetts, E.A. & Safran, R.J. (2009). Co-evolution of plumage characteristics and winter sociality in New and Old World sparrows. *J. Evol. Biol.*, 22, 2376–2386.

Tibbetts, E.A., Sheehan, M.J. & Dale, J. (2008). A testable definition of individual recognition. *Trends Ecol. Evol.*, 23, 356.

Vedder, O., Schut, E., Magrath, M.J.L. & Komdeur, J. (2010). Ultraviolet crown colouration affects contest outcomes among male blue tits, but only in the absence of prior encounters. *Funct. Ecol.*, 24, 417–425.

Watt, D.J. (1986). A comparative study of status signaling in sparrows (genus Zonotrichia). *Anim. Behav.*, 34, 1–15.

Webster, M.S., Ligon, R.A. & Leighton, G.M. (in press). Social costs are an underappreciated force for honest signalling in animal aggregations. *Anim. Behav.*, https://doi.org/10.1016/j.anbehav.2017.12.006.

Wey, T., Blumstein, D.T., Shen, W. & Jordán, F. (2008). Social network analysis of animal behaviour: a promising tool for the study of sociality. *Anim. Behav.*, 75, 333–344.

Whitfield, D.P. (1987). Plumage variability, status signalling, and individual recognition in avian flocks. *Trends Ecol. Evol.*, 2, 13–18.

Whitfield, D.P. (1988). The social significance of plumage variability in wintering Turnstone Arenaria interpres. *Anim. Behav.*, 36, 408–415.

Whiting, M.J., Nagy, K.A. & Bateman, P.W. (2003). Evolution and maintenance of social status signalling badges: experimental manipulations in lizards. In: *Lizard Social Behaviour* (ed Fox, S.F., Kelly McCoy, J. & Baird, T.A.). Johns Hopkins University Press, Baltimore, pp. 47–82.

Ydenberg, R.C., Giraldeau, L.A. & Falls, J.B. (1988). Neighbours, strangers, and the asymmetric war of attrition. *Anim. Behav.*, 36, 343–347.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.