Information content and testosterone dependence of animal signals: a case study with House Sparrows

Dissertation

an der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

durchgeführt am Max-Planck-Institut für Ornithologie

in Seewiesen

vorgelegt von

Silke Laucht

April 2011

Erstgutachter: Prof. Dr. Bart Kempenaers
Zweitgutachter: Prof. Dr. Matthias Starck

Eingereicht am: 07.04.2011
Tag der mündliche Prüfung: 20.09.2011

The English Sparrow

So dainty in plumage and hue,
A study in grey and in brown,
How little, how little we knew,
The pest he would prove to the town!
From dawn until daylight grows dim,
Perpetual chatter and scold.
No winter migration for him,
Not even afraid of the cold!
Scarce a song-bird he fails to molest,
Belligerent, meddlesome thing!
Wherever he goes as a guest
He is sure to remain as a King.

Mary Isabella Forsyth (1840-1914)

Table of Contents

Summary

Animals have evolved a variety of ornamental characteristics. Ornaments are conspicuous traits used as signals in a wide range of contexts. The most common types of ornaments are indicators or "signals of quality", traits that convey information about phenotypic and genetic constitution, status, parental care abilities, and other factors that are important for male-male competition and mate choice. In birds, signals of quality are often very colourful and can vary from inflexible plumage traits to rather dynamic bare parts (i.e. the bill, legs or wattles). Different colours stem from either structural or pigmentary underlying mechanisms. For pigmentation, melanins and carotenoids are by far the most common types and have been commonly argued to be linked to different information signalled. However recent research, including the results of my dissertation, challenges this view.

In order to remain evolutionarily stable (i.e. honest) quality signals need to be coupled with inescapable costs. For testosterone dependent ornaments, various costs associated with testosterone have been argued to provide the honesty enforcing mechanism. This could be either via social costs of challenging testosterone related aggressiveness for signals of dominance (the Badges of Status Hypothesis), or via physiological trade-offs such as immunosuppression or oxidative stress (the Testosterone Handicap Model). Immunosuppression and oxidative stress have especially attained widespread interest and acceptance. However, there are several assumptions that need to be tested before the generality of these hypotheses can be assumed.

For my dissertation I have addressed several of the underlying assumptions of the Testosterone Handicap Model in a model species for honest signalling research: the House Sparrow (*Passer domesticus*).

In chapter one I could show that there is seasonal variation in plasma testosterone levels and that individual levels in different seasons are very weakly correlated with each other. This has important implications on theories about signal honesty that are based on the correlations between testosterone levels during the time of ornament elaboration and ornament importance. This result also suggests that caution is

required when testosterone measurements for individuals are based on point samples. Finally, I demonstrated that badge size, a commonly argued testosterone dependent badge of status, was not related to day time testosterone levels. In contrast however, I provided strong evidence that bill colour could potentially function as a signal of behaviours and/or strategies related to average testosterone levels outside the breeding season.

In chapter two, I have shifted my emphasis from seasonal variation to diel variation, and I demonstrated that night time testosterone levels were significantly higher than day time levels during all seasons. I further argue that the expression of night time levels is considerably closer to individual maximal potential values. However, the reason why individuals have consistently elevated night time testosterone levels remains unclear. In contrast to day time levels, I found that badge size was positively correlated with night time levels (during the peak breeding season). This suggests that badge size has the potential to signal maximal testosterone levels and maximal potential aggression.

In chapter three, I broadened the number of House Sparrow ornaments I focused on to four, and studied the interplay of them over several years using a correlational approach. I found that House Sparrows have at least three ornaments that relate to testosterone levels and thus could potentially signal testosterone related behaviours (related to agonistic interactions). In addition, I demonstrate that at least two traits were strongly related to age.

In the last two chapters I used an experimental approach to test the developmental mechanisms underlying ornamentation. First, in chapter four, I tested whether social group composition (as determined by badge sizes) during moult influenced badge size. In one year I found that males housed in cages where every individual had similar badges tended to have stronger changes in their badges. However in two years there was no effect of group composition on badge size, indicating that the effect of social environment on badge size is weak and possibly dependent on as yet unknown factors.

Finally, in chapter five I examined three fundamental assumptions underlying recent theories on signal honesty of testosterone related ornaments. These were that (1)

testosterone is immunosuppressive, (2) carotenoids are immunoenhancing, and (3) testosterone increases the bioavailability of carotenoids. None of them were supported in either House Sparrows or Red-billed Queleas (*Quelea quelea quelea*) demonstrating that they are not of a general nature and that other costs than physiological trade-offs are needed to enforce signal honesty. This result provides compelling support for the important idea that developmental mechanisms for ornaments are best viewed as adaptations designed to optimize signalling, rather than as honesty-enforcing constraints.

To conclude, in this dissertation I have addressed several assumptions underlying the honesty of signals of quality and show that they are not as general as typically assumed. In combination, these results provide a consistent pattern that has major consequences for theories about signal honesty. I suggest that it is much more likely that testosterone related ornaments remain honest via costs of testosterone related behaviours such as social challenges from competitive rivals (as in the badge of status hypothesis). In addition I have also raised new questions that future research can profitably address to facilitate a full understanding of the enforcement mechanisms that keep quality signals honest.

Zusammenfassung

Im Laufe der Evolution wurde bei Tieren eine Vielzahl schmückender Merkmale, sogenannte Ornamente, selektiert, die in verschiedensten Zusammenhängen als Signale verwendet werden. Die häufigste Art von Ornamenten sind „Qualitätssignale“, die Informationen über phänotypische und genetische Verfassung, Dominanzstatus, Brutpflegeeigenschaften und andere Faktoren, die beim Konkurrenzkampf zwischen zwei Männchen oder bei der Partnerwahl eine Rolle spielen, preisgeben. Bei Vögeln sind Qualitätssignale oft in der Färbung der Tiere zu finden, bei der zwischen unveränderbaren Gefiedermerkmalen und variablen Haut- beziehungsweise Schnabelfarben unterschieden wird. Unterschiedliche Farben können entweder durch unterschiedliche Mikrostrukturen der Federn oder durch die Einlagerung verschiedener Farbpigmente (meist Karotenoide und Melanine) entstehen. Grundsätzlich wurden verschiedenen Pigmente mit unterschiedlicher Signalinformation in Zusammenhang gebracht. Aber neue Forschungsergebnisse, einschließlich meiner Dissertation, zweifeln diese Theorien an.

Um evolutionär stabil (d.h. also ehrlich) zu bleiben, müssen Qualitätssignale mit unausweichlichen Kosten verbunden sein. Bei Testosteron-regulierten Ornamenten wurden verschiedene Ehrlichkeit erzwingende Kosten, die in direktem Zusammenhang mit Testosteron stehen, vorgeschlagen. Diese können zum einen soziale Kosten bei Dominanzsignalen sein, die durch Herausforderung von Testosteron-regulierter Aggressivität entstehen (Hypothese über Statusmerkmale), oder zum anderen physiologische Kompromisse wie z.B. Immunsuppression oder oxidativer Stress (Testosteron Handicap Modell). Gerade Immunsuppression und oxidativer Stress stießen in der letzten Zeit auf viel Interesse und Akzeptanz. Aber bevor diese Hypothesen als allgemein gültig angenommen werden, sollten noch einige Annahmen getestet werden.

In meiner Dissertation habe ich einige Annahmen, die dem Testosteron Handicap Modell zu Grunde liegen, in einem Modellorganismus für Signal-Forschung untersucht: dem Haussperling (*Passer domesticus*).

In Kapitel eins konnte ich saisonale Schwankungen in Plasmatestosteronleveln aufzeigen. Außerdem korrelierten die Level von einzelnen Individuen zu verschiedenen Jahreszeiten nur schwach miteinander. Das hat bedeutende Auswirkungen auf Theorien zur Signalehrlichkeit, die darauf basieren, dass Testosteronlevel zur Zeit der Merkmalsausbildung mit denen zu der Zeit, wenn das Merkmal bedeutend ist, korrelieren. Dieses Ergebnis zeigt auch, dass man vorsichtig sein muss, wenn Testosteronmessungen für Individuen auf Einzelmessungen basieren. Schließlich konnte ich auch zeigen, dass die Größe des Brustflecks, der weithin als Testosteron-abhängiges Statusmerkmal betrachtet wird, nicht mit Testosteronleveln, die im Tagesverlauf gemessen wurden, korreliert. Im Gegensatz dazu konnte ich wichtige Hinweise erbringen, dass die Schnabelfarbe als potentielles Merkmal fungieren könnte, das Verhalten und Strategien in Abhängigkeit von mittleren Testosteronleveln außerhalb der Brutzeit signalisiert.

In Kapitel zwei habe ich meinen Schwerpunkt von saisonaler Variation auf Schwankungen im Tagesverlauf verlegt und konnte zeigen, dass Nachttestosteronlevel zu allen Jahreszeiten signifikant höher waren als Tagwerte. Des Weiteren denke ich, dass diese Nachtwerte deutlich näher an möglichen Maximalwerten eines Individuums liegen. Aber es bleibt weiterhin unklar, warum Nachtwerte so viel höher sind. Im Gegensatz zu den Ergebnissen mit den Tagwerten, konnte ich eine Abhängigkeit der Größe des Brustflecks von Nachtwerten (zur Brutzeit gemessen) feststellen. Das könnte bedeuten, dass der Brustfleck maximale Testosteronwerte und maximal mögliche Aggressivität signalisiert.

In Kapitel drei habe ich das Spektrum der untersuchten Ornamente auf vier erweitert und deren Zusammenspiel über mehrere Jahre in einem korrelativen Ansatz untersucht. Ich konnte feststellen, dass Haussperlinge mindestens drei Merkmale haben, die mit Testosteronleveln korrelieren und somit möglicherweise Testosteron-abhängiges Verhalten (im Zusammenhang mit agonistischen Interaktionen) signalisieren. Außerdem waren mindestens zwei Merkmale altersabhängig.

In den letzten beiden Kapiteln verwendete ich einen experimentellen Ansatz, um Entwicklungsmechanismen, die Ornamenten zu Grunde liegen, zu überprüfen. In Kapitel vier untersuchte ich, ob die soziale Gruppenzusammensetzung (über den

Brustfleck bestimmt) während der Mauser einen Einfluss auf die Größe des Brustflecks hat. In einem Jahr konnte ich einen Trend zu größeren Veränderungen in der Brustfleckgröße bei Männchen feststellen, die in Volieren mit Individuen mit ungefähr gleich großen Brustflecken waren. Aber in den anderen beiden Jahren gab es keinen solchen Effekt. Das bedeutet, dass der Einfluss von sozialer Umwelt auf Brustfleckgröße schwach und möglicherweise von noch unbekannten Faktoren abhängig ist.

In Kapitel fünf schließlich untersuchte ich drei grundlegende Annahmen, auf denen Theorien zur Ehrlichkeit von Testosteron-abhängigen Ornamenten basieren. Diese sind: (1) Testosteron ist immunosuppressiv, (2) Karotinoide sind immunverstärkend und (3) Testosteron vergrößert die Verfügbarkeit von Karotinoiden. Ich konnte keine Bestätigung für keine von ihnen bei Haussperlingen und Blutschnabelwebern (*Quelea quelea quelea*) finden, was zeigt, dass sie nicht allgemeingültig sind und dass andere Kosten, und nicht physiologische Kompromisse, benötigt werden, um die Signalehrlichkeit zu garantieren. Diese Ergebnisse bieten unmittelbare Beweise für die Idee, dass Entwicklungsmechanismen für Ornamente am besten als Anpassungen zur Signaloptimierung gesehen werden sollten, anstatt als Ehrlichkeit erzwingende Einschränkungen.

Abschließend stelle ich fest, dass ich in dieser Dissertation einige Annahmen, auf denen die Ehrlichkeit von Qualitätssignalen basiert, untersucht habe und zeigen konnte, dass sie nicht so allgemein gültig sind wie weithin angenommen. Zusammenfassend zeigen meine Ergebnisse ein konsistentes Muster, welches weit reichende Folgen für Theorien zur Signalehrlichkeit hat. Ich schlage deshalb vor, dass es wahrscheinlicher ist, dass Testosteron-regulierte Ornamente über Kosten durch Testosteron-abhängiges Verhalten wie soziale Herausforderungen von Rivalen (nach der Hypothese über Statusmerkmale) ehrlich bleiben. Des Weiteren habe ich neue Fragen aufgeworfen, die künftige Studien adressieren sollten, um somit ein tiefgreifendes Verständnis von Mechanismen, die die Ehrlichkeit von Qualitätsmerkmalen bedingen, zu erlangen.

General Introduction

"[…] we may conclude that weapons for battle, organs for producing sound, ornaments of many kinds, bright and conspicuous colours, have generally been acquired by the males through variation and sexual selection […]." (Darwin 1871, p. 560).

When watching animals one can observe that they use different traits such as colourful ornaments, sounds, or gestures to communicate with each other and also with individuals of other species. Also plants apply some traits to communicate information to potential herbivores or pollinators. Such general occurrence and consistency within and between species suggests general mechanisms underlying the evolution and development of such traits. But how can it be ensured that these traits provide honest information? How can lying, i.e. cheating, be prevented? For my dissertation I have addressed several assumptions underlying theories about the honesty of some traits – testosterone related colourful signals of quality – by using the House Sparrow (*Passer domesticus*) as a model species. In the following introduction I will briefly define and describe the nature of signals of quality, before going into details about different colour causing mechanisms underlying signals of qualities in birds and contrasting static and dynamic characters of such signals. Then, I will introduce theories on honesty-ensuring mechanisms, that is I will discuss costs associated with signals of quality. Subsequently, I will establish testosterone as one

of the links associating signals of quality with inescapable costs. Finally, I will give a short summary on House Sparrows and their ornaments before describing some general methods and an outline of my thesis.

Signals of quality

Animals have evolved an amazing number and diversity of ornamental traits as signals. Signals are traits that function in communication, that consist of one or several cues, and that are subject to selective pressures different than for other traits because of the social interactions between signaller and receiver (Candolin 2003). Although there is no universally accepted biological definition of ornaments that I am aware of, they are typically considered to be morphological characteristics that are particularly showy or elaborate, that are often sexually dimorphic, and that function as signals used in a wide range of contexts. For instance, ornamental coloration in birds can convey information about the signaller's quality, attractiveness, strategy, genetic compatibility, kinship, individual identity, and presence (Dale 2006).

Ornamental signals of quality are by far the most studied and are typically considered as signals of phenotypic and genetic constitution (Andersson 1994; reviewed in Dale 2006). More specifically signals of quality can provide information about a variety of traits related to the bearer's constitution such as social status, fighting ability, parental care abilities, disease status, good genes, ability to evade predators, etc. (Dale 2006). In addition, signals of quality can therefore be assessed in the context of male-male interactions (e.g. acquisition or defence of territories, mates, or other resources) or of male-female interactions (e.g. mate attraction) (summarized in Berglund et al. 1996).

Signals of quality are expected to have several characteristics in common, no matter what particular aspect of quality they convey (reviewed in Dale 2006): First, they demonstrate a high degree of between-individual variability, presumably because overall quality is also highly variable and affected by various developmental pathways. Second, they are largely influenced by environmental conditions (reviewed in Hill 2006). Third, like most quantitative traits, they generally display unimodal frequency distributions. However, bimodal distributions could occur in status signals

(reviewed in Dale 2006). Finally, interspecific variation between different species can be huge depending on the type of ornament and its particular information content.

Colourful signals of quality

Different animal species display a huge variety of ornamental traits that function as signals of quality. These include horns and antlers (e.g. a kudu's horns), modified or elongated fur (e.g. a lion's mane) or elaborate feather plumes (e.g. tail feathers of birds of paradise), bizarre forms of skin appendices (e.g. a cock's comb), or extremely colourful patches of fur, feathers, scales or skin, etc. (e.g. wing pattern of butterflies). In birds, bright coloration often functions as ornaments in general, and as signals of quality in particular.

Different colours are thought to be caused by different types of colour generating mechanisms, i.e. by differences in microstructure and from different pigments, typically melanins and carotenoids (Badyaev and Hill 2000, and references therein). Variation in microstructures is usually responsible for the blues and greens while variation in melanins or carotenoids is responsible for blacks/browns and yellow/reds, respectively. Because the mechanisms of different colours (i.e. peak wavelengths) differ, different coloured ornaments are commonly considered to potentially signal different information. Below, I describe in more detail each of the major types of colour generating mechanisms in birds.

Structure mediated colours are generally blue, purple, green, and ultraviolet, and often appear iridescent and glossy. They are caused when incoming light waves are scattered, reflected and constructively interfered by melanin granules and/or air vacuoles in the keratin that feathers are made of (Keyser and Hill 1999; Loyau et al. 2007). Structural colour changes considerably in the course of the year due to abrasion of feathers and microstructures, but this could be reduced by intensive preening and feather maintenance (summarized in Peters et al. 2006). Many plumage ornaments based on structural colours were found to signal condition and age, and to play a significant role in mate choice (e.g. Bennett et al. 1997; Keyser and Hill 1999; Delhey and Kempenaers 2006; Peters et al. 2006; Loyau et al. 2007).

Melanin based colours are generally black, grey, brown, rufous, and yellow, and can be summarized as earth tone colours (Badyaev and Hill 2000). Melanins are thought to be cheap because they can be synthesized by animals de novo during either (1) the amino acid catabolism from the nonessential amino acid tyrosine or (2) as non-toxic end products from the essential amino acid phenylalanine (summarized in Badyaev and Hill 2000; Griffith et al. 2006). Melanin based ornaments are also commonly considered to be more under genetic than environmental control (Badyaev and Hill 2000). Because they are thought to be cheap and genetically determined, melanin based ornaments are often argued to be less good indicators of genetic quality, but instead function as signals of dominances (reviewed in Badyaev and Hill 2000). However, recent research has challenged this conclusion by highlighting a lack of evidence for a significant difference between melanin and carotenoid based signals, and therefore suggests that more studies are needed (Griffith et al. 2006).

Carotenoid based colours are generally bright red, orange and yellow. Carotenoids are often argued to be limited in availability because they cannot be produced by the animal itself, but instead have to be taken up from food - potentially in rather high amounts for producing striking colours (reviewed in Badyaev and Hill 2000). Carotenoids are also thought to be required for the immune system - they are immunoenhancing and function as anti-oxidants (Chew and Park 2004; Alonso-Alvarez et al. 2008). Because carotenoids are potentially limited, they are often argued to be costly and thus good indicators of quality (Badyaev and Hill 2000). However, recent research suggests that carotenoids might not be as important for the immune system as widely assumed (Pérez-Rodríguez 2009; Vinkler and Albrecht 2010), and other factors such as haemoglobin or melatonin may have similar functions (Bertrand et al. 2006; Griffith et al. 2006; McGraw and Klasing 2006).

In sum, various colours of ornaments originate from different mechanisms. There is great interest in determining whether function can be implied from these underlying mechanisms, however to date it appears that we still have to definitively answer this question. Indeed, it seems quite likely that the signalling function of ornaments is context and species dependent more so than mechanism dependent. My view is that research on bird ornaments needs to compare ornaments with similar functions and contexts rather than ornaments with similar developmental mechanisms.

Static versus dynamic signals

In addition to different mechanisms and colours, visual signals of quality in birds can also be separated into plumage versus bare part ornaments. This contrast is quite important because these different kinds of ornaments vary in terms of their dynamism, that is in their capacity to change over time.

First, plumage ornaments are developed only once or twice per year during moult and then remain fixed (Pérez-Rodríguez 2008). Moreover, pigments used to colour feathers are irrepealably incorporated into the feathers and cannot be drawn-off for other purposes (Lozano 1994). They are therefore rather inflexible, and any information signalled needs to be consistent up until the next moult. Some species partly circumvent the inflexibility of plumage ornaments by adding gray or white edgings to the newly moulted feathers that conceal the ornament and that are worn off by preening until the ornament is fully needed (e.g. in male House Sparrows, *Passer domesticus*; (Møller and Erritzøe 1992)). Overall, plumage ornaments related to quality are generally suggested to function as long-term indicators of past condition, i.e. the condition that the individual was in when it moulted into the ornament.

Bare part ornaments include skin appendices such as wattles, eye rings or combs, other parts of skin such as the legs, and horned structures such as the bill. In contrast to plumage coloration, bare part coloration can change rather quickly and any time during the year (Karubian 2008). In addition, pigments used for coloration can be at least partly taken back and used for other purposes (Lozano 1994). Bare part coloration is often argued to function as a short-term indicator of current condition (Pérez-Rodríguez 2008; Ardia et al. 2010). Dynamic signals are considered especially important in environments that can change rapidly and where decisions (such as breeding) are based on current conditions (Bro-Jørgensen 2010).

Costly signals of quality

Presumably it would be advantageous, in terms of fitness gains, for individuals to signal the highest possible quality no matter what their actual quality is. What prevents low quality males from “cheating” and producing signals that convey that

they are actually high quality? In order to remain evolutionarily stable, signals of quality have to be inescapably costly (Zahavi 1975; Zahavi 1977; Grafen 1990). This means that they have to be correlated with handicaps they impose on the individuals so that an individual can only elaborate an ornament as much as it can withstand the costs associated with the handicap (Zahavi 1975). Classic examples of handicaps include the enormous tail of a peacock that constrains the bearer when moving around or when trying to escape from a predator, or conspicuous sexual displays that advertise an individual's location to predators as well as potential mates. Inter-individual variation in ornaments arises when higher quality males can afford the costs of elaborate ornamentation more than lower quality males, and when there is considerable variation between individuals in how many costs they can afford (Zahavi 2007).

Generally, honesty enforcing costs of signals of quality can be divided into social and physiological (e.g. production and maintenance) costs. Social costs of an ornament mean that an individual will be challenged by its conspecifics that have similarly sized ornaments and are trying to improve their rank (Jawor and Breitwisch 2003). Only when an ornament honestly signals the level of aggression and status, an individual can withstand attacks and energy costs imposed by fights (Maynard Smith and Harper 1988; Jawor and Breitwisch 2003; Tibbetts and Dale 2004). Therefore, cheaters will be selected against. Ornaments signalling status and level of aggression are also called 'badges of status'.

Physiological costs arise via trade-offs of resources that are used for ornament production or maintenance and other important physiological processes such as immune response (e.g Alonso-Alvarez et al. 2007; Peters 2007). They are based on the assumption that these resources are limited and trade-offs can form. One classic example are carotenoids that were suggested to be traded off between colourful ornamentation and oxidant defence (see above).

Hormone dependence of quality signals

Hormones play a major role in the regulation of diverse processes during an animal's life. Because of this diversity, they are thought to cause trade-offs by producing antagonistic effects on different traits (Ketterson and Nolan 1992) and therefore provide a potential solution to defining honesty enforcing costs of signalling. In this respect, the steroid hormone testosterone has gained increased interest. Testosterone is well recognized to impact physiological (e.g. sperm production), morphological (e.g. development of secondary sexual characters) and behavioural (e.g. courtship, aggression) characteristics of animals (Adkins-Regan 2005). Inter-individual variation in plasma testosterone levels might therefore be reflected in variation in these attributes (Laucht et al. 2011) and therefore in accordant testosterone dependent ornaments. Because testosterone has many costly behavioural and physiological effects on individuals (Wingfield et al. 2001) testosterone potentially provides a developmental mechanism whereby signal honesty can be enforced.

As discussed above, there are two kinds of costs that can ensure signal honesty: social and physiological costs. Testosterone can potentially regulate both kinds of costs. First, it can cause social costs via its effects on aggressive behaviour (reviewed in Soma 2006). Testosterone regulated aggression and thus dominance status can be advertised via badges of status. As described above testing of conspecifics will cause social costs, and an individual can only withstand these challenges when testosterone levels are high enough to cause the signalled level of aggression.

Second, testosterone can ensure the honesty of signals of quality by causing physiological costs via physiological trade-offs. This is known as the "Testosterone Handicap Model" (Adkins-Regan 2005) and it depends on physiological costs of high testosterone levels such as high energetic costs or the suppression of the immune system (Wingfield et al. 2001). Suppression of the immune system has especially achieved widespread attention.

Folstad and Karter (1992) argued that because testosterone tends to suppress the immune system, only individuals with good immune systems can withstand the costs of high testosterone levels and thus develop elaborate ornaments. This hypothesis is

known as the "Immunocompetence Handicap Hypothesis" (ICHH), and it has been highly influential. Alonso-Alvarez et al. (2007) extended this model, and suggested in their "Oxidation Handicap Hypothesis" that high testosterone levels cause oxidative stress (see also von Schantz et al. 1999) and that ornaments therefore signal the ability to counteract oxidative stress via antioxidant pigments. As carotenoids are thought to be traded off between colour pigments and antioxidants (see above) they were predicted to be the honest link between testosterone caused oxidative stress and carotenoid based ornamentation (Alonso-Alvarez et al. 2007; Alonso-Alvarez et al. 2008). The majority of research into the costs of ornamentation is currently based on these hypotheses. Support for the ICHH is weak however (Roberts et al. 2004), and Wedekind and Folstad (1994) questioned the evolutionary stability of the ICHH and suggested that immunosuppression by testosterone is better viewed as adaptive rather than as a cost.

Despite the large amount of studies and effort put into answering questions about signalling content and signal honesty of ornaments in birds and other species, there is still a huge number of questions that need to be addressed and answered. During my dissertation I have endeavoured to challenge current thinking by addressing some of the key assumptions of contemporary views of honest signalling. I have chosen one model species of ornamentation research, the House Sparrow (*Passer domesticus*) in order to do this.

The House Sparrow (*Passer domesticus*): model species for ornamentation research

"The house sparrow has several attributes that make it an ideal subject for many types of biological inquiry. These include its accessibility as a widespread commensal of urban and agricultural communities; its ready acceptance of nest-boxes as nest sites; its status as a pest of agriculture and as a disease vector for both humans and their livestock; its highly social behavior, which results in its foraging flocks and breeding semicolonially; its ready adaptation to laboratory conditions permitting extensive laboratory research (including captive breeding); and its lack of formal governmental protection in many places." (Anderson 2006, preface)

a) b)

Fig. 1: a) Male House Sparrow, b) Female House Sparrow

The House Sparrow is a small well-known passerine (~28g) that belongs to the genus *Passer* and the family Passeridae. The House Sparrow is probably the bird species with the widest distribution: it can be found nearly world-wide and in almost all kinds of habitats (Anderson 2006). It appears in close association with human-modified environments such as farms and surrounding farmland, residential, and urban areas (Lowther and Cink 2006). House Sparrows have a highly social life-style all year round with foraging flocks, breeding colonies, and roosting congregations (Anderson 2006). This sociality includes both males and females in mixed-sex groups.

The two sexes differ slightly in size and clearly in plumage characteristics. As House Sparrows have only one annual moult (after the breeding season in early fall) these differences remain constant over the year. While females are inconspicuously greyish brownish coloured, males have several distinct ornaments superimposed onto a grey and dark brown background (Fig. 1). The best studied male ornaments include the black breast bib –the so-called badge, bill coloration, the white wingbar, and the black area around the eye (see Fig. 2), but this list clearly does not include all conspicuous patches and potential ornaments.

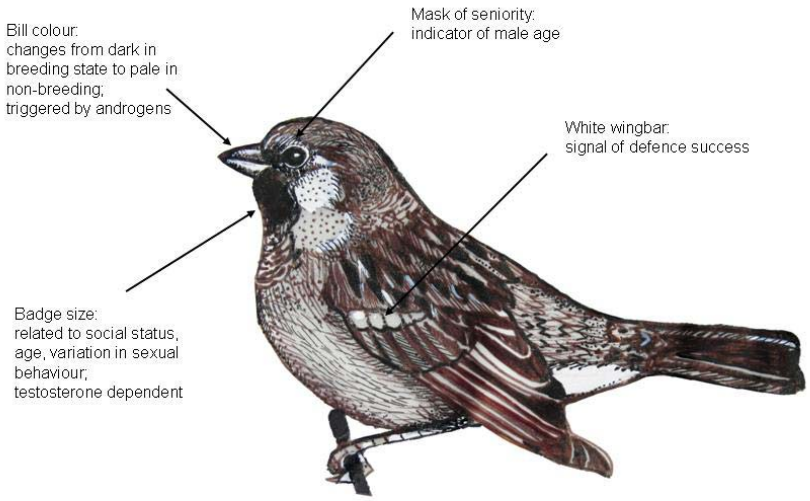


Fig. 2: Ornaments of a male House Sparrow. Indicated are all known ornaments and their general function. For references see General Introduction.
Drawing by Helga Gwinner.

The badge is by far the most studied ornament. Variation in badge size has been found to be related to social status (i.e. larger badges signal higher dominance), age (i.e. increase with age), and variation in sexual behaviour (mixed results for pairing and paternity, but larger badged males defend more their nest boxes and mates) (Møller 1987; Møller 1990; Veiga 1993; Reyer et al. 1998; Liker and Barta 2001; Václav and Hoi 2002; McGraw et al. 2003; Nakagawa et al. 2007; Morrison et al. 2008). In addition, several correlative and manipulative studies have found that badge size was related to testosterone levels around the time of the annual moult (Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001).

The signalling function of male bill colour, in contrast to badge size, has not been examined in any detail. Nevertheless, it has been long-known that bill colour changes from a pale horn colour in the non-breeding season to a blackish dark colour in the breeding season (Witschi and Woods 1936), and that this change is due to different pigmentation with eumelanins and linked to increases in male hormones (testosterone) (Keck 1933; Witschi 1936; Pfeiffer et al. 1944; Haase 1975; Donham et al. 1982). In addition, Hegner and Wingfield (1986) suggested that social competition might have caused bill darkening in fall. Furthermore, Václav (2006) proposed a possible effect of condition on bill colour; however, this has not yet been studied.

The white wingbar of the male House Sparrow is outlined by the white tips of the median coverts (Bókony et al. 2006). Bókony et al. (2006) found that the conspicuousness of this wingbar was related to defence success and suggested that it was a signal of defence ability of already occupied resources. Together with badge size it comprises one of multiple ornaments with different signalling functions during aggressive interactions (Bókony et al. 2006).

In contrast, the black area around the eyes was found to be a reasonably good indicator of male age and thus termed the "mask of seniority" (Nakagawa and Burke 2008). However, this has not been studied in any more detail, and correlations with other ornaments such as badge size are likely.

The above four ornaments have each been suggested to be signals of quality, and indeed they seem to largely fulfil the generally expected characteristics of such traits. However, in most cases they are not well studied. And in the case of the badge size results were inconsistent between different studies (on different populations). This leaves scope for more research and in-depth questions.

Aims of thesis

During my dissertation I have endeavoured to challenge current thinking by addressing some of the key assumptions underlying theories of honest signalling. More specifically, I have concentrated on several assumptions underlying the Testosterone Handicap Model explaining the honesty of testosterone related signalling. I will describe these assumptions that I have set out to challenge in more detail in the following.

First, plasma testosterone levels at different times in the course of the year and of the day need to correlate with each other to ensure signal honesty. Because plumage ornaments are often developed during the time of year when plasma testosterone levels are the lowest (i.e. moult), but ornaments are important during other seasons such as breeding when testosterone levels are high (e.g. Humphrey and Parkes 1959; Wingfield et al. 1990; Hahn et al. 1992), individuals with the highest levels during moult should also be the individuals with the highest levels during breeding season.

Similarly, testosterone levels fluctuate in the course of the day and therefore individuals with the highest levels at one time point should also be the ones with the highest levels at a different time point.

Second, testosterone is immunosuppressive. This is because testosterone causes physiological trade-offs and thus is thought to enforce signal honesty (see above). One physiological trade-off that has been widely discussed is immunosuppression.

Third, carotenoids are immunoenhancing. Because carotenoids can function as anti-oxidants (Chew and Park 2004; Alonso-Alvarez et al. 2008) they should be able to counteract negative endproducts that form during immune responses and are therefore immunoenhancing.

Fourth, testosterone increases the bioavailability of carotenoids. High quality individuals have high testosterone levels and will therefore try to circumvent the trade-offs of these high testosterone levels, i.e. immunosuppression. Because carotenoids are immunoenhancing, they could be used to counteract negative effects of high testosterone levels. Therefore, higher circulating testosterone levels should increase the bioavailability of carotenoids.

Because these assumptions are of a very general nature, signal elaboration and honesty should always be regulated in a similar way. This should also be the case, when testosterone levels vary under the influence of a variety of external and intrinsic factors. Therefore, I have chosen to examine these assumptions and the assumed influence on ornamentation in more detail and in different contexts: (1) in the course of the year, (2) in the course of the day, (3) as a function of age, (4) as the inter-play of different signals, (5) as a function of social environment, and (6) as function of changing physiology. In more detail, I have chosen the course of the year and of the day to represent changing external and internal factors, age to represent changes of internal factors and the outcome of current selection, social environment to represent the two-way relationship between testosterone and (aggressive) behaviour (e.g. Wingfield et al. 1990; Ketterson and Nolan 1992; Oyegbile and Marler 2005; Hau 2007), and the manipulation of circulating testosterone levels and immune challenges to represent changes in physiology.

General methods: study species

I used the House Sparrow as a study species because of its sociality and need of communication via ornaments, and because ornamentation is only based on one type of colour mediating mechanism, the melanin pigments. From the described ornaments I have focused mainly on badge size and bill colour: badge because of the variety of published studies, and bill colour because of the lack of information and high potential for being a short-term signal.

I studied a captive population of House Sparrows because this allowed me to collect long-term data on individuals, and to obtain repeated measures of individuals spread over the year and over a day. A captive environment also allowed to standardize environmental factors possibly influencing ornamentation and group life. For this purpose I kept more than 200 male and female House Sparrows in large semi-outdoor aviaries of 1.2 x 2.0 x 4.0 m in groups of six to ten. The aviaries, located in a barn-like building, were enclosed on one side only by chicken wire; hence, the birds were exposed to natural light and temperatures, but had *ad libitum* food and water. The fact that this population readily breeds in captivity, that most of the birds were wild-caught, and that housing conditions were very similar to where they had been caught provide some confidence that the results reflect patterns present in wild populations.

General methods: measurements of ornament colour and size

To measure ornament size and colour, I used digital photography. I demonstrated that this is a fairly accurate and repeatable measurement and that it is thus valid to be used in the context of bird ornamentation research (see chapter one). Out of the three possible methods for colour measurements (assessment of colour by eye and the classification into categories according to colour charts, photo spectrometry, and digital photography) I have chosen the latter because (1) pictures can be taken fairly quickly and easily, (2) pictures can be saved and scored multiple times, (3) photographs can be used for additional colour measurements such as size or pattern of ornaments, and (4) photos can be standardized afterwards. Despite its disadvantages (light conditions should be fairly standardized, and measurements cannot be

performed in the UV spectrum), this method has become a reasonable alternative to photo spectrometry (see chapter 1 and 2 in Hill and McGraw 2006).

To measure coloration and size of an ornament, I took two to four pictures of an individual bird in a standardized set-up: similar light conditions, constant distance between bird and camera, standardized camera settings, colour and size standard in the background, and birds always held the same way. These pictures were later analyzed using imaging processing software. To determine ornament size I encircled the ornament on the photograph and measured the encircled pixels using the program ImageJ 1.36b (Abramoff et al. 2004). To determined ornament coloration I used software written in R 2.4.0 (R Development Core Team 2006) to collect brightness values as measured in the standard HSB (i.e. hue, saturation, brightness) colour space. I measured the brightness of five pixels located at five randomly chosen positions on the ornament. I then calculated the mean of these five measures. Measurements were standardized by using the colour and size standards in the background of the photos. In addition, I used averages of all pictures taken per individual.

Thesis outline

Here, I address several assumptions underlying the Testosterone Handicap Model to explain honesty of testosterone related signals in a model species for ornamentation research, the House Sparrow. In particular, I have considered (1) how testosterone levels in the course of the year and of the day are related to the long-term changes of ornaments, (2) how age influences ornaments, (3) how different ornaments interact with each other, (4) how the social environment influences ornaments, and (5) how changes in physiology affect ornaments.

In chapter one, I have studied variation in badge size and bill colour in relation to plasma testosterone levels over the course of one full year, i.e. over four different seasons. First, I aimed to replicate results of other researchers about the relationship between badge size, testosterone levels and condition. Second, I tested the relationship between testosterone levels and changes in bill colour and I tested whether bill colour is a signal of quality. Third, and most importantly, I tested the basic underlying assumption of the testosterone handicap model of signal honesty -

that testosterone levels are correlated between the season when plumage ornaments are developed and the season when ornaments are most important.

In chapter two, I have compared testosterone levels at day and night in the course of one year and evaluated their impact on badge size. I first wanted to extend the assumption examined in chapter one about between seasonal correlations of testosterone levels to correlations between day and night testosterone levels. And second, I wanted to test how this variation is reflected in variation in the size of the badge.

In chapter three, I have studied the interrelationships of four different ornaments (badge size, bill colour, wingbar area and brightness, and leg colour) over several years and examined their information signalling potential. First, I wanted to confirm that all of them are potential signals of quality. Second, I aimed to distinguish between different possibilities of how multiple ornaments inter-play, i.e. if they signal different or similar information. And third, I examined the changes of these ornaments with age and over time.

In chapter four and five, I have used an experimental approach. In chapter four, I have studied how manipulations of the group composition of badge sizes, and thus of social environment, during moult affect the development of badge size across three different years. With this I have examined the variability of a fixed ornament and the influence of environment (in this case social environment) on mechanisms of ornament development.

In chapter five, I have studied the influence of testosterone implants and immune challenges on ornamentation (bill colour), plasma carotenoid levels and immune parameters in the House Sparrow and the Red-billed Quelea (*Quelea quelea quelea*). With this I wanted to test the key underlying assumptions of the "Testosterone Handicap Model" in one species with melanin based testosterone dependent ornaments and in one species with carotenoid based testosterone independent ornaments to examine the generality of these assumptions and the validity of these theories for explaining signal honesty. My results of this experiment have critical

implications to our understanding of how signal honesty can be maintained via developmental mechanisms.

References

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophoton. Int. 11:36-42

Adkins-Regan E (2005) Hormones and Animal Social Behavior. Princeton University Press, Princeton, NJ

Alonso-Alvarez C, Bertrand S, Faivre B, Chastel O, Sorci G (2007) Testosterone and oxidative stress: the oxidation handicap hypothesis. Proc. R. Soc. Biol. Sci. Ser. B 274:819-825

Alonso-Alvarez C, Pérez-Rodríguez L, Mateo R, Chastel O, Viñuela J (2008) The oxidation handicap hypothesis and the carotenoid allocation trade-off. J. Evol. Biol. 21:1789-1797

Anderson TR (2006) Biology of the ubiquitious house sparrow. Oxford University Press, New York

Andersson M (1994) Sexual Selection. Princeton University Press, Princeton, NJ

Ardia DR, Broughton DR, Gleicher MJ (2010) Short-term exposure to testosterone propionate leads to rapid bill color and dominance changes in zebra finches. Horm. Behav. 58:526-532

Badyaev AV, Hill GE (2000) Evolution of sexual dichromatism: contribution of carotenoid- versus melanin-based coloration. Biol. J. Linn. Soc. 69:153-172

Bennett ATD, Cuthill IC, Partridge JC, Lunau K (1997) Ultraviolet plumage colors predict mate preferences in starlings. Proc. Natl. Acad. Sci. U. S. A. 94:8618-8621

Berglund A, Bisazza A, Pilastro A (1996) Armaments and ornaments: An evolutionary explanation of traits of dual utility. Biol. J. Linn. Soc. 58:385-399

Bertrand S, Faivre B, Sorci G (2006) Do carotenoid-based sexual traits signal the availability of non-pigmentary antioxidants? J. Exp. Biol. 209:4414-4419

Bókony V, Lendvai AZ, Liker A (2006) Multiple cues in status signalling: The role of wingbars in aggressive interactions of male house sparrows. Ethology 112:947-954

Bro-Jørgensen J (2010) Dynamics of multiple signalling systems: animal communication in a world in flux. Trends Ecol. Evol. 25:292-300

Buchanan KL, Evans MR, Goldsmith AR, Bryant DM, Rowe LV (2001) Testosterone influences basal metabolic rate in male house sparrows: a new cost of dominance signalling? Proc. R. Soc. Biol. Sci. Ser. B 268:1337-1344

Candolin U (2003) The use of multiple cues in mate choice. Biol. Rev. (Camb.) 78:575-595

Chew BP, Park JS (2004) Carotenoid action on the immune response. J. Nutr. 134:257s-261s

Dale J (2006) Intraspecific variation in coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration Harvard University Press, MA, Cambridge, pp 36-86

Darwin C (1871) The descent of man, and selection in relation to sex. Penguin Books (2004), London, England

Delhey K, Kempenaers B (2006) Age differences in blue tit *Parus caeruleus* plumage colour: within-individual changes or colour-biased survival? J. Avian Biol. 37:339-348

Donham RS, Wingfield JC, Mattocks PW, Farner DS (1982) Changes in Testicular and Plasma Androgens with Photoperiodically Induced Increase in Plasma LH in the House Sparrow. Gen. Comp. Endocrinol. 48:342-347

Evans MR, Goldsmith AR, Norris SRA (2000) The effects of testosterone on antibody production and plumage coloration in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 47:156-163

Folstad I, Karter AJ (1992) Parasites, Bright Males, and the Immunocompetence Handicap. Am. Nat. 139:603-622

Gonzalez G, Sorci G, Smith LC, de Lope F (2001) Testosterone and sexual signalling in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 50:557-562

Grafen A (1990) Biological Signals as Handicaps. J. Theor. Biol. 144:517-546

Griffith SC, Parker TH, Olson VA (2006) Melanin-versus carotenoid-based sexual signals: is the difference really so black and red? Anim. Behav. 71:749-763

Haase E (1975) Effects of Testosterone Propionate on Secondary Sexual Characters and Testes of House Sparrows, *Passer domesticus*. Gen. Comp. Endocrinol. 26:248-252

Hahn TP, Swingle J, Wingfield JC, Ramenofsky M (1992) Adjustments of the Prebasic Molt Schedule in Birds. Ornis Scand. 23:314-321

Hau M (2007) Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. BioEssays 29:133-144

Hegner RE, Wingfield JC (1986) Gonadal Development during Autumn and Winter in House Sparrows. Condor 88:269-278

Hill GE (2006) Environmental Regulation of Ornamental Coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration. Harvard University Press, Cambridge, pp 507-560

Hill GE, McGraw KJ (2006) Bird Coloration: Function and Evolution, vol 2. Harvard University Press, MA, Camebridge

Humphrey PS, Parkes KC (1959) An approach to the study of molts and plumages. Auk 76:1-31

Jawor JM, Breitwisch R (2003) Melanin ornaments, honesty, and sexual selection. Auk 120:249-265

Karubian J (2008) Changes in breeding status are associated with rapid bill darkening in male red-backed fairy-wrens *Malurus melanocephalus*. J. Avian Biol. 39:81-86

Keck WN (1933) Control of the bill color of the male English sparrow by injection of male hormone. Proc. Soc. Exp. Biol. Med. 30:1140-1141

Ketterson ED, Nolan V (1992) Hormones and Life Histories - an Integrative Approach. Am. Nat. 140:S33-S62

Keyser AJ, Hill GE (1999) Condition-dependent variation in the blue-ultraviolet coloration of a structurally based plumage ornament. Proc. R. Soc. Biol. Sci. Ser. B 266:771-777

Laucht S, Dale J, Mutzel A, Kempenaers B (2011) Individual variation in plasma testosterone levels and its relation to badge size in House Sparrows *Passer domesticus*: It's a night-and-day difference. Gen. Comp. Endocrinol. 170:501-508

Liker A, Barta Z (2001) Male badge size predicts dominance against females in House Sparrows. Condor 103:151-157

Lowther PE, Cink CL (2006) House Sparrow (*Passer domesticus*) In: Poole A (ed) The Birds of North America Online. Cornell Lab of Ornithology, Ithaca. http://bna.birds.cornell.edu/bna/species/012.

Loyau A, Gomez D, Moureau BT, Théry M, Hart NS, Saint Jalme M, Bennett ATD, Sorci G (2007) Iridescent structurally based coloration of eyespots correlates with mating success in the peacock. Behav. Ecol. 18:1123-1131

Lozano GA (1994) Carotenoids, Parasites, and Sexual Selection. Oikos 70:309-311

Maynard Smith JFRS, Harper DGC (1988) The Evolution of Aggression - Can Selection Generate Variability? Philos. Trans. R. Soc. Lond. B Biol. Sci. 319:557-570

McGraw KJ, Dale J, Mackillop EA (2003) Social environment during molt and the expression of melanin-based plumage pigmentation in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 53:116-122

McGraw KJ, Klasing KC (2006) Carotenoids, immunity, and integumentary coloration in red junglefowl (*Gallus gallus*). Auk 123:1161-1171

Møller AP (1987) Variation in Badge Size in Male House Sparrows *Passer domesticus* - Evidence for Status Signaling. Anim. Behav. 35:1637-1644

Møller AP (1990) Sexual-Behavior Is Related to Badge Size in the House Sparrow *Passer domesticus*. Behav. Ecol. Sociobiol. 27:23-29

Møller AP, Erritzøe J (1992) Acquisition of Breeding Coloration Depends on Badge Size in Male House Sparrows *Passer domesticus*. Behav. Ecol. Sociobiol. 31:271-277

Morrison EB, Kinnard TB, Stewart IRK, Poston JP, Hatch MI, Westneat DF (2008) The links between plumage variation and nest site occupancy in male House Sparrows. Condor 110:345-353

Nakagawa S, Burke T (2008) The mask of seniority? A neglected age indicator in house sparrows *Passer domesticus*. J. Avian Biol. 39:222-225

Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T (2007) Assessing the function of house sparrows' bib size using a flexible meta-analysis method. Behav. Ecol. 18:831-840

Oyegbile TO, Marler CA (2005) Winning fights elevates testosterone levels in California mice and enhances future ability to win fights. Horm. Behav. 48:259-267

Pérez-Rodríguez L (2008) Carotenoid-based ornamentation as a dynamic but consistent individual trait. Behav. Ecol. Sociobiol. 62:995-1005

Pérez-Rodríguez L (2009) Carotenoids in evolutionary ecology: re-evaluating the antioxidant role. BioEssays 31:1116-1126

Peters A (2007) Testosterone and carotenoids: an integrated view of trade-offs between immunity and sexual signalling. BioEssays 29:427-430

Peters A, Delhey K, Goymann W, Kempenaers B (2006) Age-dependent association between testosterone and crown UV coloration in male blue tits (*Parus caeruleus*). Behav. Ecol. Sociobiol. 59:666-673

Pfeiffer CA, Hooker CW, Kirschbaum A (1944) Deposition of pigment in the sparrow's bill in response to direct applications as a specific and quantitative test for androgen. Endocrinology 34:389-399

R Development Core Team (2006) R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria

Reyer HU, Fischer W, Steck P, Nabulon T, Kessler P (1998) Sex-specific nest defense in house sparrows (*Passer domesticus*) varies with badge size of males. Behav. Ecol. Sociobiol. 42:93-99

Roberts ML, Buchanan KL, Evans MR (2004) Testing the immunocompetence handicap hypothesis: a review of the evidence. Anim. Behav. 68:227-239

Soma KK (2006) Testosterone and aggression: Berthold, birds and beyond. J. Neuroendocrinol. 18:543-551

Tibbetts EA, Dale J (2004) A socially enforced signal of quality in a paper wasp. Nature 432:218-222

Václav R (2006) Within-male melanin-based plumage and bill elaboration in male house sparrows. Zool. Sci. 23:1073-1078

Václav R, Hoi H (2002) Different reproductive tactics in house sparrows signalled by badge size: Is there a benefit to being average? Ethology 108:569-582

Veiga JP (1993) Badge Size, Phenotypic Quality, and Reproductive Success in the House Sparrow - a Study on Honest Advertisement. Evolution 47:1161-1170

Vinkler M, Albrecht T (2010) Carotenoid maintenance handicap and the physiology of carotenoid-based signalisation of health. Naturwissenschaften 97:19-28

von Schantz T, Bensch S, Grahn M, Hasselquist D, Wittzell H (1999) Good genes, oxidative stress and condition-dependent sexual signals. Proc. R. Soc. Biol. Sci. Ser. B 266:1-12

Wedekind C, Folstad I (1994) Adaptive or Nonadaptive Immunosuppression by Sex-Hormones. Am. Nat. 143:936-938

Wingfield JC, Hegner RE, Dufty AM, Ball GF (1990) The Challenge Hypothesis - Theoretical Implications for Patterns of Testosterone Secretion, Mating Systems, and Breeding Strategies. Am. Nat. 136:829-846

Wingfield JC, Lynn SE, Soma KK (2001) Avoiding the 'costs' of testosterone: Ecological bases of hormone-behavior interactions. Brain Behav. Evol. 57:239-251

Witschi E (1936) The bill of the sparrow as an indicator for the male sex hormone. I. sensitivity. Proc. Soc. Exp. Biol. Med. 33:484-486

Witschi E, Woods RP (1936) The bill of the sparrow as an indicator for the male sex hormone. II. structural basis. J. Exp. Zool. 73:445-459

Zahavi A (1975) Mate Selection - Selection for a Handicap. J. Theor. Biol. 53:205-214

Zahavi A (1977) Cost of Honesty - (Further Remarks on Handicap Principle). J. Theor. Biol. 67:603-605

Zahavi A (2007) Sexual selection, signal selection and the handicap principle. In: Jamieson BGM (ed) Reproductivev biology and phylogeny of birds: sexual selection, behavior, conservation, embryology, genetics Science Publishers, Enfield, NH, USA, pp 143-159

Published as: Laucht S, Kempenaers B, Dale J (2010) Bill color, not badge size, indicates testosterone-related information in house sparrows. Behavioral Ecology and Sociobiology 64: 1461-1471

Chapter One

Bill color, not badge size indicates testosterone-related information in house sparrows

Silke Laucht, Bart Kempenaers, James Dale

Abstract

The honesty of ornamental signals of quality is often argued to be enforced via costs associated with testosterone. It is still poorly understood, however, how seasonal variation of testosterone within individuals is related to the timing and extent of ornament development. Here we studied inter- and intra-individual variability of plasma testosterone levels in a population of 150 captive male House Sparrows (*Passer domesticus*) through the course of a full year. We further analyzed the relationship between plasma testosterone levels and two sexually dimorphic ornaments: badge size and bill coloration. Also, because of a known negative relation between molt and circulating testosterone levels, we analyzed the relationship between ornamentation and molt status during the fall. We found that testosterone levels increased towards the breeding season and decreased before the onset of annual molt. However, within individuals, relative testosterone titers demonstrated low repeatability between seasons. Plasma testosterone levels were not correlated with badge size in any season, but were correlated strongly with bill coloration during all periods, except the breeding season when variation in bill color was low. Finally, we found that bill coloration strongly correlated with molt status during fall. Our results indicate that bill coloration, not badge size, is the best ornamental indicator of a "running average" of male testosterone in House Sparrows, and therefore the best potential indicator of qualities and/or behavioral strategies associated with testosterone.

Introduction

Current interest in the relationship between the steroid hormone testosterone (T) and sexually selected ornamentation is high because testosterone is well known to be responsible for the elaboration of many sexually selected ornaments (reviewed in e.g. Candolin 2003; Roberts et al. 2004). The honesty of ornamental traits which specifically signal quality is argued to be enforced through inescapable costs associated with signaling (Grafen 1990; Johnstone and Norris 1993; Veiga 1993; Buchanan et al. 2003; Peters et al. 2004). Potential costs hypothesized to be associated with quality signals are testosterone-related suppression of the immune system (e.g. the Immunocompetence Handicap Hypothesis (ICHH) Folstad and Karter 1992)) or testosterone-related depression of resistance to oxidative stress (e.g. the Oxidation Handicap Hypothesis (OHH) (Alonso-Alvarez et al. 2007)). Another form of signaling cost (often suggested for melanin-based ornaments) is social enforcement via costs created by aggressive interactions with, or testing by, conspecifics (i.e., the badge of status hypothesis (Maynard Smith and Harper 1988; Jawor and Breitwisch 2003; Tibbetts and Dale 2004)). Testosterone-dependent ornaments can thus be expected to provide information to conspecifics about qualities associated with either physiological characteristics such as immunosuppression and/or oxidative stress or behavioral characteristics such as aggression (Folstad and Karter 1992; Johnstone and Norris 1993; Alonso-Alvarez et al. 2007; McGraw and Ardia 2007).

Since testosterone dramatically affects both behavior and physiology (reviewed in Wingfield et al. 1990; Ketterson and Nolan 1992; Adkins-Regan 2005; Hau 2007) maintaining elevated testosterone levels all year round can be costly (e.g. through increased risk of aggression-related injury or through increased metabolic costs (see Wingfield et al. 2001)). Temperate zone bird species, for example, show dramatic fluctuations in plasma testosterone levels over the course of the year. They rise towards the breeding season to levels needed for the physiological changes and behaviors associated with breeding and drop to baseline levels with the onset of prebasic (or post-nuptial) molt in early fall when birds are more vulnerable and less aggressive (Humphrey and Parkes 1959; Wingfield et al. 1990; Hahn et al. 1992). Even during periods of elevated testosterone, T-levels are generally kept at a certain breeding baseline and are then modulated on a short term basis as a result of social

interactions (e.g. aggression associated with reproductive behavior (Wingfield 1984; Wingfield 1985; Wingfield et al. 1990)). Overall, it is generally argued that testosterone levels are kept at an optimum that is well balanced with the various behavioral and physiological costs of maintaining them (Folstad and Karter 1992; Wingfield et al. 2001; Adkins-Regan 2005; Hau 2007).

To date, researchers have resolved many examples of an endocrine basis to variation in color-based ornaments occurring in birds (reviewed in Hill and McGraw 2006), monkeys (e.g. Setchell et al. 2008; Clough et al. 2009; Lewis 2009), fish (e.g. Dijkstra et al. 2007; Kurtz et al. 2007) and lizards (e.g. Thompson and Moore 1991; Salvador et al. 1996; Calisi and Hews 2007; Huyghe et al. 2009), and comprised of all of the three major mechanisms of coloration: structural (e.g. Peters et al. 2006), carotenoid-based (e.g. Hill 2002) and melanin-based (reviewed in Jawor and Breitwisch 2003; Bókony et al. 2008), in ornaments consisting of feathers (e.g. Evans et al. 2000; Safran and McGraw 2004), integuments (e.g. Verhulst et al. 1999; Mougeot et al. 2004; Blas et al. 2006) or bill (e.g. Keck 1933; Mundinger 1972; Murphy et al. 2009). Among these studies the House Sparrow (*Passer domesticus*) has become one of the model organisms for the study of testosterone and melanin-based ornaments.

Male House Sparrows have at least two noteworthy sexually dimorphic ornaments: the black breast bib or the "badge", and the black bill. Many studies have demonstrated that the size of the badge relates to social status, age, and variation in sexual behavior (Møller 1987; Møller 1990; Veiga 1993; Liker and Barta 2001; Vaclav and Hoi 2002; McGraw et al. 2003; Nakagawa et al. 2007; Morrison et al. 2008) and is subject to natural and sexual selection (Jensen et al. 2008). Furthermore, three studies found that badge size correlated positively with plasma testosterone levels around annual (prebasic) molt (Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001) when the new badge is formed. In contrast, the signaling function of the bill is still a mystery. Male bill color changes between seasons from a pale horn color in the non-breeding season to a blackish dark color in the breeding season (Witschi and Woods 1936), and the color change is due to pigmentation with eumelanins and is known to be mediated by male hormones (testosterone) (Keck 1933; Witschi 1936; Pfeiffer et al. 1944; Haase 1975; Donham et al. 1982). Although

Václav (2006) suggested the possibility of an effect of condition on bill color, so far there have been no tests of this.

Even though there appear to be clear relationships between testosterone levels and these two ornaments in House Sparrows, the relationship between within-individual variation in testosterone levels and the timing and extent of ornament development is still very poorly understood. For example, feather ornaments molted during times of lowest testosterone levels (such as the badge in the House Sparrow) can only remain honest indicators of testosterone levels during the breeding season if testosterone levels during different periods of the year are correlated. This, however, is surprisingly poorly studied (Kempenaers et al. 2008). To the best of our knowledge only one study has examined the relationship between breeding and post-breeding testosterone levels in House Sparrows (Buchanan et al. 2003). To address this gap, here we studied plasma testosterone levels, badge size, and bill color in a large captive population of male House Sparrows over the course of one year. The objectives of the study were to test the following four non-exclusive predictions.

1) Within individuals, plasma testosterone levels should be correlated between seasons. This was found by the only study comparing two different seasons (Buchanan et al. 2003) despite it being a fundamental assumption of theories trying to explain the testosterone-related honesty of signals that are developed at different times than when they are actually used (e.g. the ICHH (Folstad and Karter 1992; Alonso-Alvarez et al. 2007) and the OHH (Alonso-Alvarez et al. 2007)).

2) Badge size should be positively correlated with testosterone levels. As previously stated, badge size signals social status and sexual behavior (e.g. communal displays), both testosterone-related traits. Additionally, such a relation was previously found in correlational approaches and after artificial increases of testosterone levels during molt (Evans et al. 2000; Buchanan et al. 2001).

3) Bill color should be positively correlated with testosterone levels. As described above, seasonal changes in bill color are mediated by changes in

testosterone levels, and therefore we predicted that the extent of coloration in this dynamic ornament should be related to current testosterone levels.

4) Bill color should be correlated with molt status. In other species, molt status is strongly negatively correlated with plasma testosterone levels (Schleussner 1990; Hahn et al. 1992; Nolan et al. 1992). Therefore, because we expected a positive relation between bill color and testosterone levels, we also predicted a relation between bill color and molt status.

Although our study is correlational and we do not test the signaling role of the bill directly, we nevertheless assume here that bill coloration is an evolved signal (i.e. that it can influence decision-making in receivers), because it is a conspicuous, sexually dimorphic trait which has some apparent design in terms of the complex physiological processes responsible for its dynamic expression.

Material and Methods

Study population

We studied a population of 150 captive male House Sparrows held at the Max-Planck-Institute for Ornithology, Seewiesen, Germany. All males were after-hatching year birds. They were either caught in rural areas in Bavaria, Germany (under license: permit nr. 55.1-8642.3-3-2006 of the "Regierung Oberbayern", with several extensions) and held in captivity for at least eight months (n = 136) or born in captivity (n = 14; exclusion did not qualitatively change the results). From July 2006 until July 2007, individuals were kept in all-male groups of five or six in aviaries of size 1.2 x 2.0 x 4.0 m. After July 2007, we kept them in the same aviaries in groups of nine or ten (note that we did not measure testosterone in the males after we changed the group size). At all times, the birds had *ad libitum* access to food (wild seed mix for forest birds (Waldvogelfutter: RKW Sued, Universal Kraftfutterwerk, Kehl, Germany), sunflower seeds, crushed corn and wheat, oats, chicken starter, soybean meal extract, and mineral mix for birds), drinking and bathing water, and sand. The light-dark cycle and temperatures in the aviaries were close to natural conditions, as the aviaries were semi-outdoor with one side enclosed only by chicken

wire. Although our study is on a captive population of House Sparrows, we have confidence that our results reflect patterns present under wild populations because our sparrows were mostly wild-caught individuals that were housed in large semi-outdoor aviaries that were grouped together in a barn-like building that was very similar to the actual barns where we caught the sparrows. Moreover our House Sparrows readily bred under these conditions, indicating conditions were highly favorable for natural behavior. Finally our study replicates research by other groups working on captive sparrows, and so our results are highly comparable to theirs.

During five periods throughout the course of the year (Oct./Nov. 2006, Jan. 2007, March 2007, June 2007, and Sept./Oct. 2007) we caught all individuals and took biometric measurements, standardized photographs of the bill and badge, and blood samples (except in Sept./Oct. 2007, when no blood samples were taken). In fall 2007, we additionally scored molt status. During each blood sampling period, conducted between 07:00 and 10:00h and between 13:00 and 15:00h, we took 150-200 µl of blood from the wing vein within fifteen minutes after first starting to catch the birds. The time passed since first starting to catch birds did not have an influence on T levels (lme: $t = 1.13$, $p = 0.26$, $n = 551$ with season, day time, and bird ID as random effects). We collected the blood in 75 mm Na-heparinized micro haematocrit capillaries and centrifuged it at 13000 rpm for three minutes to separate the plasma. Plasma was stored at -80°C.

Determination of plasma T levels

Frozen plasma samples were sent to the endocrine laboratory of the Leibniz Institute for Zoo and Wildlife Research in Berlin, Germany, where testosterone (T) levels were determined by enzyme immunoassays (for details on the methods see Roelants et al. 2002). The inter-assay CV for the enzyme immunoassay was 12.3% and the intra-assay CV was 9.0%. Additionally, to calculate the true repeatability (intra-class correlation coefficient) of measuring serum T-levels, we split plasma samples of several males into duplicates right after centrifuging. Across the whole year, repeatability of these plasma T estimates was $R = 0.967 \pm 0.006$ (SE) ($p < 0.001$, $n = 2*122$). Note that this estimate includes additional non-assay sources of variation because of the immediate separation after centrifugation. It thus gives a conservative

estimate of measurement error in plasma T-levels. We assumed that all data points with a value of zero (29 out of 579) were actually below the detection limit (for each assay slightly different, but around 20 pg/ml) and thus assigned them the lowest value measured (15 pg/ml). Excluding those individuals entirely from the analysis did not qualitatively change the results. T levels are reported in pg per ml. Testosterone concentrations were natural log transformed to achieve a normal distribution, in order to fit standard least-squares models.

Determination of bill color

Immediately after blood samples were taken, we took two standardized photographs of each bird's bill. We used a Canon Power Shot S2 IS camera and took pictures at the highest resolution with flash. All males were held the same way (presenting the right side of the head and bill to the camera) in front of a gray card and color standard background at the same distance from the camera.

Digital photograph processing software written into R 2.4.0 (R Development Core Team 2006) was used to collect values of bill "brightness" as measured in the standard HSB (i.e. hue, saturation, and brightness) color space. Note that in this color-scoring scheme, "brightness" is an indicator of how dark or light a color is and correlates closely with mean total reflectance (Montgomerie 2006, also see Fig. 1). SL measured brightness of the individual pixels located at five randomly chosen positions each on the upper bill, lower bill and the gray background around the bill (used as a brightness standard). To standardize our bill brightness measurements between photos, we calculated overall mean gray card brightness of all photos of each season, we then determined the deviation of gray card brightness of a focal photo from the overall mean, and subtracted this deviation from mean bill brightness for each picture. This standardization renders a bill brightness score for each male that both compensates for any minor differences in overall brightness between photos and that also keeps our brightness variable as an actual color measurement (rather than a difference). For analyses, we used the mean of the standardized upper and lower bill brightnesses from both pictures. As expected for a trait with considerable phenotypic variation, these measurements were highly repeatable within individuals (repeatability (Lessells and Boag 1987): $R = 0.949 \pm 0.004$ (SE), $p < 0.001$, $n = 2*581$ for two

pictures). For repeatability measures between two consecutive years, we took additional photographs of the bills in January 2008.

To illustrate the correspondence between bill brightness (measured with digital photographs) and total reflectance, we measured the bills of two differently colored males with a hand-held spectrometer (Avantes, AvaSpec-2048, Eerbek, The Netherlands) with a deuteriume-halogen light source (Avantes, Ava-Light-D(H)-S). We measured five points each on the bill. In Fig. 1 we present averages over each 20nm spectral range averaging also the five measurements, and clearly demonstrate that total reflectance is dramatically different in bills with different brightness values.

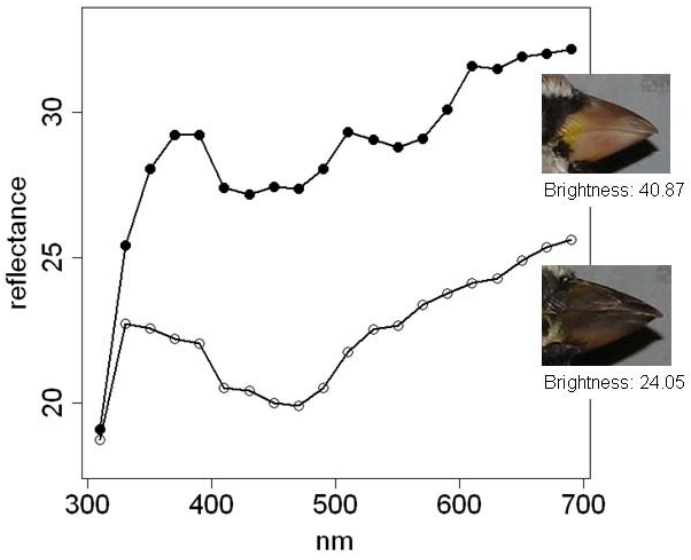


Fig. 1 Full spectrogram of two differently colored male House Sparrow bills. Our measures of bill brightness as measured from digital photographs corresponds closely with total reflectance as measured with spectrophotometry. Open circles represent a medium dark bill color, closed circles a pale bill color.

Determination of badge size

We took four pictures of the birds' breast badges during each season. For each picture, we held the birds ventrally such that the throat and badge was stretched and presented to the camera. Between each photograph of the badge, the bird was arranged into a different position, before rearranging it into the badge-exposed

position. SL measured the size of the badge from the photographs by encircling and measuring the area of the melanized badge in pixels using the program ImageJ 1.36b (Abramoff et al 2004). For standardization, we divided this area by a standard area present in each photograph and measured in the same way, and then converted the result into cm^2. For analyses, we used the average of all four pictures for each bird. In males with white tips to their badge feathers (Møller and Erritzøe 1992), badge size was estimated by outlining as best as possible the area occupied by any apparent melanized feathers underlying the white tips. The measurements were highly repeatable within individuals (R = 0.943, estimated according to Falconer and Mackay (1996) from repeatability of single pictures).

Determination of molt stage

From September 4th to 14th, we scored molt of 145 males. On October 9th and October 18th, we re-scored 30 males each. We determined molt status according to Ginn and Melville (2000, p 28 Fig. 7b) for primaries. When wing feathers were asynchronously molted on the right wing, we used the scores of the left wing. We scored old feathers as 0, fully grown new feathers as 5, and growing feathers as 1 to 4 according to their length. For analyses, we added up all scores of the single primaries of one bird (hereafter BTO (= British Trust for Ornithology) score). Additionally, we used a binary score, where a feather was simply determined as old (= 0) or molting and/or new (= 1). We then added up these scores (hereafter binary score). As the BTO score and the binary score produced very similar results, we only present the results of the BTO score in the results section.

Statistical analyses

We performed all statistical analyses using R 2.6.2 (R Development Core Team 2008; packages: ape, effects, nlme, RODBC, survival) at the significance level $\alpha = 0.05$. For overall analyses, we used linear mixed effect models (lmes) using individual ID as a random factor to account for repeated sampling. When using data from single seasons where each individual was represented only once, we used linear models (lms). Before analysis we checked for normal distribution, after analysis we assessed whether the assumption of lms and of lmes on the within-group errors and on the

random effects were violated. In the analyses reported here we did not remove any outliers (max. 4 per model) because we had large samples sizes and no a priori expectations of what constitutes normal ranges for our biological variables. In addition, all analyses yielded qualitatively similar results with or without outliers included.

For analyses of relationships between bill color and testosterone levels, we used the data of the four periods when plasma samples were taken. In addition, in the analysis of the relationship between badge size and testosterone or condition, we used the average of the March and June scores as a male's badge phenotype because at these times the white feather edges have mostly abraded off (Møller and Erritzøe 1992). For analyses of molt status, we used the data of fall 2007. For analyses with body condition we fitted three separate models, the first used the residuals of body mass regressed onto tarsus length, the second used body mass alone, and the third one used tarsus length alone. We calculated date as days passed since the beginning of measurements (Sept 25th 06 for analyses of testosterone, badge size, and bill color; Sept 4th 07 for analyses of molt and bill color). Therefore, date starts in fall (during molt) and ends in June (peak breeding season), and we expect close to linear relationships between date and T levels or bill color. We additionally accounted for day time in all models including T or used residuals of T in relation to date and day time (see indication in results). Note that differences in sample sizes are due to missing data points for single individuals because of sample loss during centrifugation or natural death of few individuals.

Results

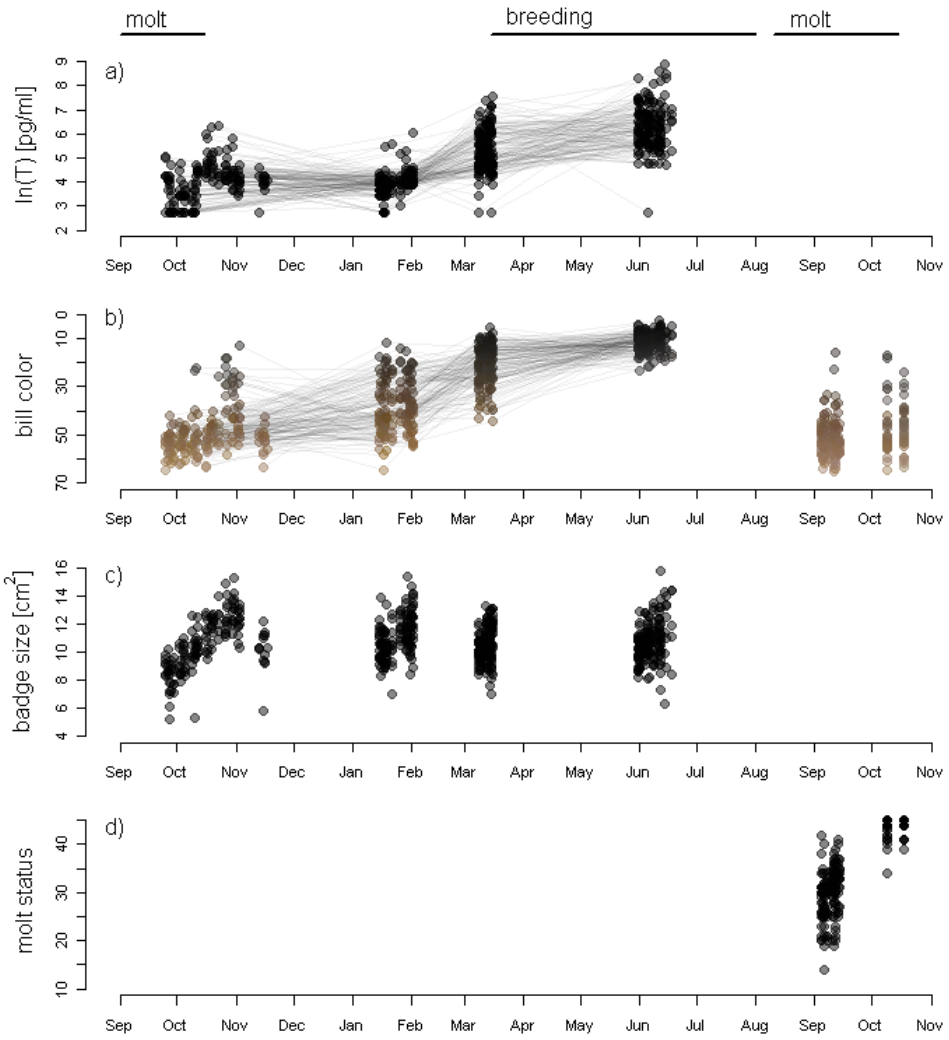


Fig. 2 Plasma testosterone levels (a, natural logarithm), bill color (b, brightness, actual bill color at sampling time presented), badge size (c, for each season the mean of March and June badge is plotted), and molt status (d, BTO score) of House Sparrows in the course of a year. We changed transparency of points and lines to improve visibility of overlapping points and lines.
The gray lines in a) and b) connect measurements of the same individuals at different seasons. Parallel lines would be expected if individuals' relative phenotypes are consistent between seasons.

Dynamics in plasma testosterone levels

Plasma T levels changed dramatically throughout the course of the year (Fig. 2a). We found a significant positive relationship between T and date (lme: $p < 0.0001$, $t_{424} = 24.67$, random effect: bird ID; after accounting for day time: $p < 0.0001$, $t_{414} = 23.58$).

However, within individuals, T levels were only weakly correlated between seasons (see lines in Fig. 2a connecting measurements of the same individuals): T levels were significantly correlated only between Sept/Oct and January and between January and March (Table1).

Table 1: Correlations of plasma testosterone levels of House Sparrows between seasons within individuals.
We used Pearson correlation tests of the residuals of ln testosterone in relation to date and day time for each of the two seasons compared.

periods compared	r	95% ci	statistic	p-value
fall - January	0.24	0.08, 0.39	$t_{137} = 2.88$	0.005
fall - March	0.07	-0.10, 0.23	$t_{136} = 0.78$	0.440
fall - June	0.02	-0.15, 0.19	$t_{134} = 0.26$	0.799
January - March	0.19	0.03, 0.35	$t_{136} = 2.29$	0.023
January - June	-0.05	-0.21, 0.12	$t_{134} = -0.54$	0.589
March - June	0.10	-0.07, 0.27	$t_{133} = 1.20$	0.232

We found no significant relationship between T and body condition when accounting for date: Neither tarsus length nor the residuals of body mass on tarsus length were related to T (lmes for overall analyses, lms for single seasons; all $p > 0.07$; body mass and tarsus length were positively correlated: $p < 0.001$ overall and for all seasons separately when accounting for date). However, when January data were analyzed separately, T was positively related to body mass after accounting for date ($t_{136} = 2.63$, $p = 0.009$, variance explained by the model $R^2 = 0.21$; other seasons and overall: $p > 0.06$). These results were very similar when using residuals of T in relation to date and day time instead of T except for the relation between T and body mass over all seasons ($t_{412} = 2.78$, $p = 0.006$).

Badge size

Badge size was not related to plasma testosterone levels in any of the four periods ($r < 0.09$, $t < 1.10$, $p > 0.27$; Fig. 2 & 3), even after accounting for tarsus length (variance explained by the model $R^2 < 0.06$, $t < 0.64$, $p > 0.53$). This did not change when using residuals of T in relation to date and day time. However, badge size was positively related to tarsus length ($r = 0.24$, $t_{133} = 2.90$, $p = 0.004$) and to body mass in March ($r = 0.17$, $t_{133} = 2.03$, $p = 0.04$) and June ($r = 0.18$, $t_{133} = 2.15$, $p = 0.03$), but not in fall and January ($t < 1.67$, $p > 0.09$). The residuals of body mass on tarsus length were not related to our badge scores for any season ($t < 0.86$, $p > 0.39$). Thus,

although there were weak positive correlations between badge size and overall size, there was no significant relationship between badge size and testosterone levels.

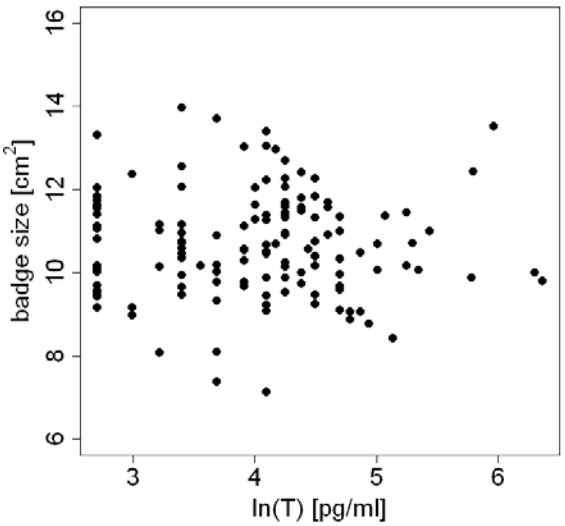


Fig. 3 Badge size in relation to plasma testosterone levels in House Sparrows during the fall (Sept.-Oct.).
We defined badge size as the mean of the March and June measurements and used the natural logarithm of plasma testosterone values. $p = 0.9$, $r = 0.01$

Dynamics in bill color and its relation to plasma T levels

Similar to testosterone, bill color changed dramatically throughout the course of the year (Fig. 2b). We found a highly significant relation between bill color and date (lme: $t_{423} = -42.31$, $p < 0.0001$, random effect: bird ID). From September until March, individuals varied substantially in bill color, ranging from pale horn-colored to almost black. Nevertheless, the distribution of bill color shifted from a majority of individuals with pale horn-colored bills in September to a majority of individuals with dark bills in March. During the breeding season between-individual variation in bill coloration was very low, i.e. all birds had dark bills (Bartlett Test of homogeneity of variances comparing breeding season with the other three seasons: Bartlett's $K^2_1 = 239.85$, $p < 0.001$). Within individuals, bill color was correlated between all seasons except the breeding season (Table 2). Bill color was also significantly repeatable between January measurements of two consecutive years ($r = 0.518 \pm 0.065$ (SE), $p < 0.001$, $n = 2*127$). Bill color was not related to body condition when accounting for

date: Neither body mass, nor tarsus length, nor the residuals of body mass on tarsus length were correlated with bill color (lmes for overall analyses, lms for single seasons; all $p > 0.18$).

Table 2: Correlations of bill color of House Sparrows between seasons within individuals.
We used Pearson correlation tests of residuals of bill color (brightness) in relation to date for each of the two seasons compared.

periods compared	r	95% ci	statistic	p-value
fall - January	0.49	0.35, 0.61	$t_{134} = 6.44$	< 0.001
fall - March	0.33	0.18, 0.47	$t_{137} = 4.14$	< 0.001
fall - June	0.05	-0.12, 0.22	$t_{134} = 0.63$	0.532
January - March	0.60	0.48, 0.69	$t_{134} = 8.59$	< 0.001
January - June	0.02	-0.15, 0.19	$t_{131} = 0.21$	0.838
March - June	-0.02	-0.18, 0.15	$t_{134} = -0.18$	0.856

We found a significant overall relationship between bill color and plasma T levels (lme: $t_{421} = -24.23$, $p < 0.001$, random effect: bird ID; Fig. 2a & b). Bill color and T were also correlated in all periods separately, except June when variation in bill color was very low (Table 3, Fig. 4). To exclude all date effects, we further used residuals of linear mixed effect models of bill color in relation to date and of T levels in relation to date: in this analysis bill color and T were still highly correlated for all seasons combined ($t_{574} = -6.94$, $p < 0.0001$).

Table 3: Correlations of bill color and plasma testosterone levels in House Sparrows during different seasons.
We used Pearson correlation tests of bill color (brightness) in relation to residuals of linear models of the natural logarithm of testosterone levels in relation to day time.

period	r	95% ci	statistic	p-value
fall	-0.25	-0.40, -0.09	$t_{137} = -3.06$	0.003
January	-0.24	-0.39, -0.08	$t_{134} = -2.88$	0.005
March	-0.25	-0.40, -0.09	$t_{136} = -3.02$	0.003
June	-0.09	-0.25, 0.07	$t_{148} = -1.13$	0.262

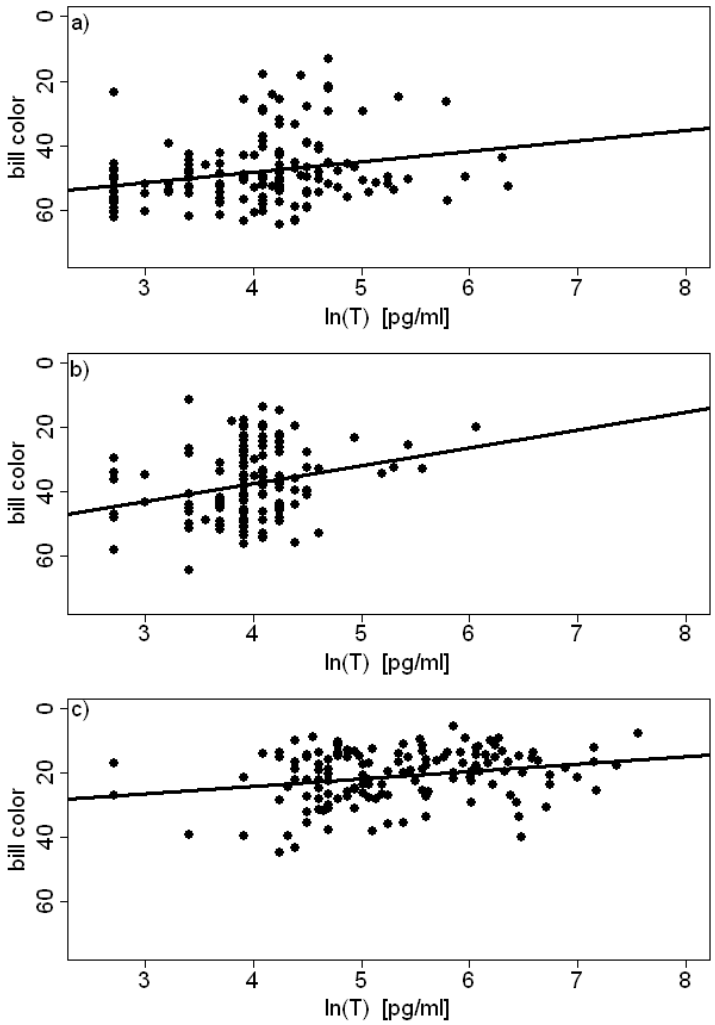


Fig. 4 Bill color in relation to plasma testosterone levels in House Sparrows (a: fall, b: January, and c: March data).
We measured bill color as brightness and used the natural logarithm of plasma testosterone values. (a: $p = 0.003$, $r = -0.25$, b: $p = 0.005$, $r = -0.24$, c: $p = 0.003$, $r = -0.25$).

Bill color and molt status

BTO molt score was highly significantly related to date (lme: $t_{57} = 39.23$, $p < 0.001$, random effect: bird ID). After accounting for date, the score was not related to body mass, tarsus length, or the residuals of mass on tarsus length (lmes for overall analyses, lms for September and October separately; all $p > 0.07$).

Bill color was independent of molt status in October and for September and October combined (lme for overall analysis, lm for October separately; all $p > 0.14$). However, when including date in the overall model and when looking at the September data

alone, we found a highly significant negative relationship between bill color and BTO molt score (lme for overall analysis after accounting for date: $t_{56} = 3.03$, $p = 0.004$, random effect: bird ID; lm for September: $r = 0.31$, $t_{142} = 3.84$, $p < 0.001$, Fig. 5). This means that birds got paler bills during the course of molt, but only as long as they were molting (hence no relation in October).

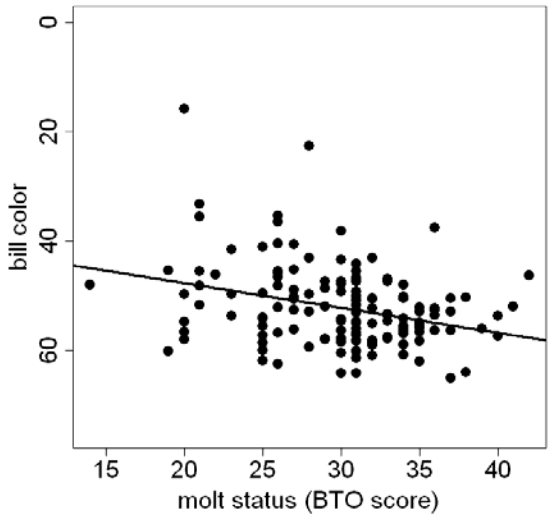


Fig. 5 Bill color in relation to molt status in House Sparrows (September data).
We measured bill color as brightness and molt status according to Ginn and Melville (2000, p 28 Fig. 7b). $p < 0.001$, $r = 0.31$.

Discussion

As expected and in accordance with many other studies (reviewed in Anderson 2006), plasma testosterone levels of captive male House Sparrows changed dramatically in the course of the year. We found the lowest values during the prebasic molt and the highest values during the breeding season, whereby there was only a slight increase between values in March, the beginning of breeding season, and June, the peak breeding season. The breeding season is the time with the most aggressive interactions between males establishing and defending territories, nest sites, and mates (e.g. Wingfield et al. 1990; Hau 2007). In contrast, molt is the time when aggression is lowest as molting birds are more vulnerable and their flight abilities are reduced (Swaddle and Witter 1997; Anderson et al. 2004). As House Sparrows show flocking behavior all year round, and tend to establish dominance ranks, inter-individual

aggression occurs most of the year (Anderson 2006) as reflected in increases of plasma testosterone levels after the termination of molt, but long before breeding commences.

In contrast to our first prediction, within-individual testosterone levels were only very weakly correlated between periods. Thus males with relatively high (or low) levels in one period did not have relatively high (or low) levels in another period. This result suggests that caution is required when interpreting point-sample testosterone measurements. Our result contrasts strongly with Buchanan et al (2003) who did find some consistency in individual testosterone levels between breeding and post-breeding season, albeit with a much lower sample size (n = 19) and medium effect (R^2 = 19.3, p = 0.027).

We did not find any relationship between badge size and testosterone levels in our population of House Sparrows during any of the studied periods, thus failing to support our second prediction. This is surprising as the result strongly contrasts with previous studies (Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001). However, our analysis is based on a much larger sample size, and four separate measurements of T per individual. Since we did find very strong correlations between testosterone levels and another ornament (bill color, see below), it seems unlikely that the lack of a correlation with badge size was a type II error. In our study population, badge size therefore does not seem to predict testosterone-related dominance and aggression (Ketterson and Nolan 1992). We therefore suggest that if badge size does reflect dominance in House Sparrows, then it is more likely that it indicates dominance related to both body size (this study, Buchanan et al. 2001) and age (see Nakagawa et al. 2007; Morrison et al. 2008) rather than dominance related to testosterone per se. For a better understanding of the potential signaling function of badge size in our population the determination of dominance ranks and of the degree of inter-individual aggression would be necessary. Note that testosterone could still be an important factor for the badge ornament despite the above results: For example, badge size could be related to true baseline testosterone levels or to true maximum levels (released during agonistic interactions), but these are hard to measure. Alternatively, testosterone could be negatively related to the length of the white tips of the badge feathers which may hide color signals during times when the signaling

function is less needed as suggested by Gonzalez et al. (2001). At any rate, the relationship between badge size and circulating testosterone in our population of House Sparrows is, at best, weak.

The lack of support for our first two predictions has important implications for the discussion of signal honesty in general (see also Kempenaers et al. 2008). The Immunocompetence Handicap Hypothesis (ICHH, Folstad and Karter 1992) and the Oxidation Handicap Hypothesis (OHH, Alonso-Alvarez et al. 2007) explain the honesty of signals of quality via the testosterone dependent immunosuppression and depressed resistance to oxidative stress, respectively (e.g. Evans et al. 2000; Gonzalez et al. 2001). However, both hypotheses assume that testosterone levels are correlated between seasons, so that a signal is not only honest at the time it is produced (i.e. molt for the badge), but also when it is used (i.e. all year round for the badge). Our observed inconsistency of testosterone levels between different seasons violates this basic assumption. Moreover, our results suggest only a weak (at best) correlation of testosterone levels with badge size, and therefore a relationship to physiological effects of testosterone such as immunosuppression is an unlikely explanation for the badge signal's honesty. Overall, we conclude that the honesty of the badge cannot be ensured via testosterone related physiological effects but rather via different costs, such as social costs.

In clear contrast to badge size, we found that bill color was strongly correlated with testosterone levels in three of the four studied periods (no correlation in June, when variation in bill color was reduced). Furthermore, these relationships were independent of condition and repeatable in similar environmental conditions in consecutive years. These findings are in accordance with our third prediction and with other studies that have found that the degree of blackening in the bill coincided with an increase in testosterone levels (Keck 1933; Keck 1934; Witschi 1936; Pfeiffer et al. 1944; Haase 1975; Donham et al. 1982). This suggests that at any time outside the breeding season, bill color is a very good predictor of relative testosterone levels. A complete change of bill color takes about three and a half weeks in male House Sparrows (Anderson 2006). Therefore, bill color probably reflects average baseline testosterone levels (i.e. a running average) over a short period of time (a few weeks at most). Our results strongly suggest that any ornamental signaling function in male

House Sparrows that is directly related to testosterone-dependent information will be much more likely found with bill coloration rather than with the badge. More specifically we suggest that bill color might serve as a signal of "strategy" (see Dale 2006), rather than as a signal of quality per se. Although signals of strategy are distinct from signals of overall quality (see Dale 2006), strategy signals can still indirectly reveal relative quality provided only good quality males pursue more costly strategies (as might be associated with increased aggression associated with higher T levels, or higher T levels maintained over a longer period of time).

We thus hypothesize that between molt and breeding, some males in the population keep their testosterone levels relatively high, indicate this with darker bills, and are more dominant because of testosterone related aggressiveness, while other males maintain low testosterone levels, indicate this with paler bills, and are thus rather subdominant while avoiding costs of high testosterone levels. The former is similar to what some studies have described as 'autumn sexuality' which may be related to higher breeding success (reviewed in Hegner and Wingfield 1986) and to any inherent qualities associated with different durations of elevated testosterone levels (see e.g. Kempenaers et al. 2008). Such a potential signal for behavioral strategies could be used either in the context of mediating competitive interactions in feeding flocks (with potentially variable flock sizes and group members; as suggested for melanin based ornaments in new and old world sparrows (Tibbetts and Safran 2009)), and/or alternatively in the context of establishing early breeding territories/sites and acquiring a mate early (as suggested for Red Grouse (Mougeot et al. 2005)). Indeed, McGraw (2004) noted that the melanin based black cap of male American Goldfinches *Carduelis tristis* was particularly more variable in the non-breeding season and suggested that such "remnant" winter ornaments might be expected to evolve in animals that live in large non-breeding groups (e.g. status-signaling systems) or in those where mates begin associating before breeding onset. In both scenarios, a signal announcing the degree of aggressiveness a male is willing to engage in could help to resolve encounters without true fights as is often suggested for badge size (Møller 1987). Because bill color is a much more dynamic ornament and because badge size does not signal testosterone levels in our population we suggest that bill color signals testosterone related aggressiveness in the non-breeding season in House Sparrows. That is, we suggest that in House Sparrows bill color is the more likely

"badge-of-status" than is the badge. This is in accordance with another study that suggested bill darkening in fall to be a response to social competition (Hegner and Wingfield 1986).

We found a negative relation between bill color and molt in September - the peak period of molt. This reflects the general relationship between molt and decreased testosterone (Schleussner 1990; Nolan et al. 1992) when the majority of males had pale bills. The hypothesis that male bill color is a signal of strategy predicts that all molting males signal low aggression perhaps because they are more vulnerable at this time. Molt can therefore be considered the mirror image to breeding as in both seasons all males are pursuing the same strategy at the same time. Nevertheless, bill color variation during molt is larger than during breeding. This may reflect (1) that the timing and speed of molt are more flexible in response to external and internal factors (Hahn et al. 1992), and (2) that individuals do not need to synchronize to successfully complete molt (in contrast to breeding).

In summary, we found that individual plasma testosterone levels were highly variable and not repeatable. We furthermore found that badge size and bill color in male House Sparrows likely signal different kinds of information. Bill color is a strongly T-related trait, whereas badge size is not. We hypothesize that bill color indicates different aggression related strategies during the non-breeding season. More detailed studies on the function of bill color in the context of social interactions are needed for a better understanding of this testosterone related trait in male House Sparrows.

Acknowledgements

Many thanks to Ariane Mutzel and Andrea Wittenzellner for their help with taking blood samples and to Barbara Helm for help with the molt scoring. We are grateful to Mihai Valcu and Holger Schielzeth for statistical advice and programming, to Kevin McGraw and three anonymous reviewers for comments on the manuscript and to Magnus Rudorfer, Edith Bodendorfer, Annemarie Grötsch, and Frances Weigel for animal care. We thank the Behavioural Ecology and Evolutionary Genetics Group at the MPIO for discussion and comments. Funding for this research was provided by the Max Planck Society.

References

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophoton. Int. 11:36-42

Adkins-Regan E (2005) Hormones and Animal Social Behavior. Princeton University Press, Princeton, NJ

Alonso-Alvarez C, Bertrand S, Faivre B, Chastel O, Sorci G (2007) Testosterone and oxidative stress: the oxidation handicap hypothesis. Proc. R. Soc. Biol. Sci. Ser. B 274:819-825

Anderson KE, Davis GS, Jenkins PK, Carroll AS (2004) Effects of bird age, density, and molt on behavioral profiles of two commercial layer strains in cages. Poult. Sci. 83:15-23

Anderson TR (2006) Biology of the ubiquitious house sparrow. Oxford University Press, New York

Blas J, Pérez-Rodríguez L, Bortolotti GR, Vinuela J, Marchant TA (2006) Testosterone increases bioavailability of carotenoids: Insights into the honesty of sexual signaling. Proc. Natl. Acad. Sci. U S A 103:18633-18637

Bókony V, Garamszegi LZ, Hirschenhauser K, Liker A (2008) Testosterone and melanin-based black plumage coloration: a comparative study. Behav. Ecol. Sociobiol. 62:1229-1238

Buchanan KL, Evans MR, Goldsmith AR (2003) Testosterone, dominance signalling and immunosuppression in the house sparrow, *Passer domesticus*. Behav. Ecol. Sociobiol. 55:50-59

Buchanan KL, Evans MR, Goldsmith AR, Bryant DM, Rowe LV (2001) Testosterone influences basal metabolic rate in male house sparrows: a new cost of dominance signalling? Proc. R. Soc. Biol. Sci. Ser. B 268:1337-1344

Calisi RM, Hews DK (2007) Steroid correlates of multiple color traits in the spiny lizard, Sceloporus pyrocephalus. J. Comp. Physiol. B 177:641-654

Candolin U (2003) The use of multiple cues in mate choice. Biol. Rev. (Camb) 78:575-595

Clough D, Heistermann M, Kappeler PM (2009) Individual Facial Coloration in Male *Eulemur fulvus rufus*: A Condition-dependent Ornament? Int. J. Primatol. 30:859-875

Dale J (2006) Intraspecific variation in coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration vol 2. Harvard University Press, MA, Cambridge, pp 36-86

Dijkstra PD, Hekman R, Schulz RW, Groothuis TGG (2007) Social stimulation, nuptial colouration, androgens and immunocompetence in a sexual dimorphic cichlid fish. Behav. Ecol. Sociobiol. 61:599-609

Donham RS, Wingfield JC, Mattocks PW, Farner DS (1982) Changes in Testicular and Plasma Androgens with Photoperiodically Induced Increase in Plasma LH in the House Sparrow. Gen. Comp. Endocrinol. 48:342-347

Evans MR, Goldsmith AR, Norris SRA (2000) The effects of testosterone on antibody production and plumage coloration in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 47:156-163

Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics, 4th edition edn. Pearson Education Limited, Harlow, Essex

Folstad I, Karter AJ (1992) Parasites, Bright Males, and the Immunocompetence Handicap. Am. Nat. 139:603-622

Ginn HB, Melville DS (2000) Moult in birds, vol 19. British Trust for Ornithology, Norwich, England

Gonzalez G, Sorci G, Smith LC, de Lope F (2001) Testosterone and sexual signalling in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 50:557-562

Grafen A (1990) Biological Signals as Handicaps. J. Theor. Biol. 144:517-546

Haase E (1975) Effects of Testosterone Propionate on Secondary Sexual Characters and Testes of House Sparrows, *Passer domesticus*. Gen. Comp. Endocrinol. 26:248-252

Hahn TP, Swingle J, Wingfield JC, Ramenofsky M (1992) Adjustments of the Prebasic Molt Schedule in Birds. Ornis Scand. 23:314-321

Hau M (2007) Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. BioEssays 29:133-144

Hegner RE, Wingfield JC (1986) Gonadal Development during Autumn and Winter in House Sparrows. Condor 88:269-278

Hill GE (2002) A red bird in a brown bag: The function and evolution of colorful plumage in the House Finch. Oxford University Press, Inc., New York

Hill GE, McGraw KJ (2006) Bird Coloration: Function and Evolution, vol 2. Harvard University Press, MA, Camebridge

Humphrey PS, Parkes KC (1959) An approach to the study of molts and plumages. Auk 76:1-31

Huyghe K, Husak JF, Herrel A, Tadić Z, Moore IT, Van Damme R, Vanhooydonck B (2009) Relationships between hormones, physiological performance and immunocompetence in a color-polymorphic lizard species, *Podarcis melisellensis*. Horm. Behav. 55:488-494

Jawor JM, Breitwisch R (2003) Melanin ornaments, honesty, and sexual selection. Auk 120:249-265

Jensen H, Steinsland I, Ringsby TH, Sæther BE (2008) Evolutionary dynamics of a sexual ornament in the house sparrow (*Passer domesticus*): The role of indirect selection within and between sexes. Evolution 62:1275-1293

Johnstone RA, Norris K (1993) Badges of Status and the Cost of Aggression. Behav. Ecol. Sociobiol. 32:127-134

Keck WN (1933) Control of the bill color of the male English sparrow by injection of male hormone. Proc. Soc. Exp. Biol. Med. 30:1140-1141

Keck WN (1934) The control of the secondary sex characters in the English sparrow, *Passer domesticus* (Linnaeus). J. Exp. Zool. 67:315-347

Kempenaers B, Peters A, Foerster K (2008) Sources of individual variation in plasma testosterone levels. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363:1711-1723

Ketterson ED, Nolan V (1992) Hormones and Life Histories - an Integrative Approach. Am. Nat. 140:S33-S62

Kurtz J, Kalbe M, Langefors Å, Mayer I, Milinski M, Hasselquist D (2007) An experimental test of the immunocompetence handicap hypothesis in a teleost fish: 11-ketotestosterone suppresses innate immunity in three-spined sticklebacks. Am. Nat. 170:509-519

Lessells CM, Boag PT (1987) Unrepeatable Repeatabilities - a Common Mistake. Auk 104:116-121

Lewis RJ (2009) Chest staining variation as a signal of testosterone levels in male Verreaux's Sifaka. Physiol. Behav. 96:586-592

Liker A, Barta Z (2001) Male badge size predicts dominance against females in House Sparrows. Condor 103:151-157

Maynard Smith JFRS, Harper DGC (1988) The Evolution of Aggression - Can Selection Generate Variability? Philos. Trans. R. Soc. Lond. B Biol. Sci. 319:557-570

McGraw KJ (2004) Winter plumage coloration in male American goldfinches: Do reduced ornaments serve signaling functions in the non-breeding season? Ethology 110:707-715

McGraw KJ, Ardia DR (2007) Do carotenoids buffer testosterone-induced immunosuppression? An experimental test in a colourful songbird. Biol. Let. 3:375-378

McGraw KJ, Dale J, Mackillop EA (2003) Social environment during molt and the expression of melanin-based plumage pigmentation in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 53:116-122

Møller AP (1987) Variation in Badge Size in Male House Sparrows *Passer domesticus* - Evidence for Status Signaling. Anim. Behav. 35:1637-1644

Møller AP (1990) Sexual-Behavior Is Related to Badge Size in the House Sparrow *Passer domesticus*. Behav. Ecol. Sociobiol. 27:23-29

Møller AP, Erritzøe J (1992) Acquisition of Breeding Coloration Depends on Badge Size in Male House Sparrows *Passer domesticus*. Behav. Ecol. Sociobiol. 31:271-277

Montgomerie R (2006) Analyzing colors. In: Hill GE, McGraw KJ (eds) Bird Coloration, vol 1. Harvard University Press, Cambridge, MA, pp 90-147

Morrison EB, Kinnard TB, Stewart IRK, Poston JP, Hatch MI, Westneat DF (2008) The links between plumage variation and nest site occupancy in male House Sparrows. Condor 110:345-353

Mougeot F, Dawson A, Redpath SM, Leckie F (2005) Testosterone and autumn territorial behavior in male red grouse *Lagopus lagopus scoticus*. Horm. Behav. 47:576-584

Mougeot F, Irvine JR, Seivwright L, Redpath SM, Piertney S (2004) Testosterone, immunocompetence, and honest sexual signaling in male red grouse. Behav. Ecol. 15:930-937

Mundinger PC (1972) Annual Testicular Cycle and Bill Color Change in Eastern American Goldfinch. Auk 89:403-419

Murphy TG, Rosenthal MF, Montgomerie R, Tarvin KA (2009) Female American goldfinches use carotenoid-based bill coloration to signal status. Behav. Ecol. 20:1348-1355

Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T (2007) Assessing the function of house sparrows' bib size using a flexible meta-analysis method. Behav. Ecol. 18:831-840

Nolan V, Ketterson ED, Ziegenfus C, Cullen DP, Chandler CR (1992) Testosterone and Avian Life Histories - Effects of Experimentally Elevated Testosterone on Prebasic Molt and Survival in Male Dark-Eyed Juncos. Condor 94:364-370

Peters A, Delhey K, Goymann W, Kempenaers B (2006) Age-dependent association between testosterone and crown UV coloration in male blue tits (*Parus caeruleus*). Behav. Ecol. Sociobiol. 59:666-673

Peters A, Denk AG, Delhey K, Kempenaers B (2004) Carotenoid-based bill colour as an indicator of immunocompetence and sperm performance in male mallards. J. Evol. Biol. 17:1111-1120

Pfeiffer CA, Hooker CW, Kirschbaum A (1944) Deposition of pigment in the sparrow's bill in response to direct applications as a specific and quantitative test for androgen. Endocrinology 34:389-399

R Development Core Team (2006) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Roberts ML, Buchanan KL, Evans MR (2004) Testing the immunocompetence handicap hypothesis: a review of the evidence. Anim. Behav. 68:227-239

Roelants H, Schneider F, Goritz F, Streich J, Blottner S (2002) Seasonal changes of spermatogonial proliferation in roe deer, demonstrated by flow cytometric analysis of c-kit receptor, in relation to follicle-stimulating hormone, luteinizing hormone, and testosterone. Biol. Reprod. 66:305-312

Safran RJ, McGraw KJ (2004) Plumage coloration, not length or symmetry of tail-streamers, is a sexually selected trait in North American barn swallows. Behav. Ecol. 15:455-461

Salvador A, Veiga JP, Martin J, Lopez P, Abelenda M, Puerta M (1996) The cost of producing a sexual signal: Testosterone increases the susceptibility of male lizards to ectoparasitic infestation. Behav. Ecol. 7:145-150

Schleussner G (1990) Experimental Induction of Eccentric Primary Molt in Adult Male European Starlings (*Sturnus vulgaris*). J. Ornithol. 131:151-155

Setchell JM, Smith T, Wickings EJ, Knapp LA (2008) Social correlates of testosterone and ornamentation in male mandrills. Horm. Behav. 54:365-372

Swaddle JP, Witter MS (1997) The effects of molt on the flight performance, body mass, and behavior of European starlings (*Sturnus vulgaris*): An experimental approach. Can. J. Zool. 75:1135-1146

Thompson CW, Moore MC (1991) Throat Colour Reliably Signals Status in Male Tree Lizards, *Urosaurus ornatus*. Anim. Behav. 42:745-753

Tibbetts EA, Dale J (2004) A socially enforced signal of quality in a paper wasp. Nature 432:218-222

Tibbetts EA, Safran RJ (2009) Co-evolution of plumage characteristics and winter sociality in New and Old World sparrows. J. Evol. Biol. 22:2376-2386

Václav R (2006) Within-male melanin-based plumage and bill elaboration in male house sparrows. Zool. Sci. 23:1073-1078

Vaclav R, Hoi H (2002) Different reproductive tactics in house sparrows signalled by badge size: Is there a benefit to being average? Ethology 108:569-582

Veiga JP (1993) Badge Size, Phenotypic Quality, and Reproductive Success in the House Sparrow - a Study on Honest Advertisement. Evolution 47:1161-1170

Verhulst S, Dieleman SJ, Parmentier HK (1999) A tradeoff between immunocompetence and sexual ornamentation in domestic fowl. Proc. Natl. Acad. Sci. U S A 96:4478-4481

Wingfield JC (1984) Environmental and Endocrine Control of Reproduction in the Song Sparrow, *Melospiza-Melodia* .II. Agonistic Interactions as Environmental Information Stimulating Secretion of Testosterone. Gen. Comp. Endocrinol. 56:417-424

Wingfield JC (1985) Short-Term Changes in Plasma-Levels of Hormones during Establishment and Defense of a Breeding Territory in Male Song Sparrows, *Melospiza-Melodia*. Horm. Behav. 19:174-187

Wingfield JC, Hegner RE, Dufty AM, Ball GF (1990) The Challenge Hypothesis - Theoretical Implications for Patterns of Testosterone Secretion, Mating Systems, and Breeding Strategies. Am. Nat. 136:829-846

Wingfield JC, Lynn SE, Soma KK (2001) Avoiding the 'costs' of testosterone: Ecological bases of hormone-behavior interactions. Brain Behav. Evol. 57:239-251

Witschi E (1936) The bill of the sparrow as an indicator for the male sex hormone. I. sensitivity. Proc. Soc. Exp. Biol. Med. 33:484-486

Witschi E, Woods RP (1936) The bill of the sparrow as an indicator for the male sex hormone. II. structural basis. J. Exp. Zool. 73:445-459

Published as: Laucht S, Dale J, Mutzel A, Kempenaers B (2011) Individual variation in plasma testosterone levels and its relation to badge size in House Sparrows *Passer domesticus*: It's a night-and-day difference. General and Comparative Endocrinology 170: 501-508

Chapter Two

Individual variation in plasma testosterone levels and its relation to badge size in House Sparrows *Passer domesticus*: It's a night-and-day difference

Silke Laucht, James Dale, Ariane Mutzel, Bart Kempenaers

Abstract

The steroid hormone testosterone (T) plays a central role in the regulation of reproduction in animals. Although seasonal variation in T levels is well-studied, differences between day and night have only been described in relatively few species, and daily within-individual variation has been largely neglected when evaluating the relationship between T and the expression of sexual ornaments or behaviour. We measured plasma T levels during day and night in a captive population of House Sparrows, and analyzed their relationship with an important male ornament - badge size. T levels were on average twice as high at night than during daytime. This was true in all seasons, and in both males and females. Disturbance of the birds at night, but not during the day, led to significantly lower T levels, suggesting a rapid drop after an individual wakes up. The relationship between T levels and badge size depended on the time when T was measured. During the breeding season, badge size was strongly positively correlated with night-time, but not with daytime T levels. This suggests that badge size signals information related to an individual's maximum potential T level such as social dominance. Our study highlights that integrative research on the endocrine control of ornament expression needs to take diel variation in hormone levels into account.

Introduction

The steroid hormone testosterone (T) is well known to influence physiological, morphological and behavioural characteristics of animals (Adkins-Regan 2005), and individual variation in plasma T levels may reflect variation in these characteristics. Sources of individual variation in T include genetic, maternal, age, time-of-day, and social environment effects (Kempenaers et al. 2008). Variation in plasma T levels occurs seasonally (Wingfield et al. 1990), within seasons over different phases of reproduction (Wingfield et al. 1990), and over the 24h period (e.g. Aschoff 1979; Plant 1981), which may partly be due to pulsatile T secretion (Vizcarra et al. 2004). In contrast to seasonal variation, patterns of diel variation in T levels have been studied in comparatively fewer species (all known studies are summarized in Table 1).

In human males T levels show considerable diel variation with a maximum occurring in the early morning (Evans et al. 1971; Barberia et al. 1973; Walton et al. 2007). Similarly, in both diurnal mammals and birds, male testosterone levels were generally higher at night than during daytime (Table 1). The only study on female birds showed similar patterns (Hau et al. 2002), suggesting a general increase in night testosterone levels. Earlier studies on human males also suggested a relationship between increased T levels and sleep (Evans et al. 1971; Weitzman 1976; Aschoff 1979), and more recent studies describe a correlation of diel variation in T levels with REM sleep (Luboshitzky et al. 1999; Luboshitzky et al. 2001; Luboshitzky et al. 2003). The association of high T levels and sleep (rather than night-time per se) is strongly supported by patterns observed in nocturnal species, where T levels show the opposite pattern and peak during the day (Wilson et al. 1976; Hoffmann and Nieschlag 1977; Dixson and Gardner 1981; Perret 1985; Guchhait and Haldar 1999). The functional significance of the diel variation in T levels, if any, remains unknown.

In behavioural or evolutionary ecology, among the most studied roles of T is its influence on ornament elaboration (reviewed in e.g. Candolin 2003; Roberts et al. 2004; Kempenaers et al. 2008), and on behaviours related to male-male competition or female choice, such as aggression, mate guarding, song output and courtship (e.g. Hegner and Wingfield 1986; Wingfield et al. 1990; Adkins-Regan 2005). In most studies, T levels are measured in a single blood plasma sample taken from individuals

that were caught during the day. However, plasma T levels can show both pronounced daily rhythms (see above) and even shorter-term episodic pulses within the day (Vizcarra et al. 2004). Such short-term within-individual variation has not been taken into account when evaluating the relationship between individual T levels and ornament elaboration or behaviours. It also remains unclear whether the diel variation in T levels is more or less pronounced during the period when the influence of T on reproductive behaviour and ornament expression is most important. We are unaware of any studies that have specifically addressed these issues.

To examine circadian variation of plasma T levels, we studied a population of captive birds (House Sparrows, *Passer domesticus*). This allowed multiple sampling of the same individuals during day and night in different seasons. Our study had three general objectives: (1) to describe individual differences in night and day T levels in males and females during four periods covering an entire year; (2) to examine the effect of disturbance during day and night by sampling immediately after disturbance or 30-60 min later; (3) to investigate the relationship between day and night T levels and a male sexual ornament (badge size). Note that our study specifically focuses on broad-scale diel changes in T (i.e. night versus day values). Since we cannot sample T levels continually throughout the course of a day our study does not provide enough resolution to infer shorter-term T pulses (e.g. Vizcarra et al. 2004) although such episodic release of T may be an important additional source of individual variation.

The House Sparrow (*Passer domesticus*) is one of the model organisms for studies on endocrine control of breeding behaviour and ornamentation expression (Hegner and Wingfield 1986; Anderson 2006). House Sparrows have an obvious ornament, the black bib or badge, which is present in males but not in females. Previous studies showed that badge size is related to social status, age, and variation in sexual behaviour (Møller 1987; Møller 1990; Veiga 1993; Liker and Barta 2001; McGraw et al. 2003; Nakagawa et al. 2007; Morrison et al. 2008). The link between individual daytime plasma T levels and badge size has also been studied, but the results differ among studies. Some studies found a positive correlation between plasma T levels measured around the time of the annual (pre-basic) moult and the size of the new badge (Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001). However, Laucht et al. (2010) found no correlation between badge size and daytime plasma T

levels during any season. The difference between studies is difficult to explain, in particular because Evans et al. (2000) also demonstrated a causal effect of testosterone via implants. However, information about diel variation in T levels and its influence on badge size is lacking. This is important because differences in ornaments (here badge size) could reflect differences in individual variation of T levels (e.g. McGlothlin et al. 2008). This individual variation could be most prominent either in average T levels or in the increase of T levels above average (due to time, seasonal or social effects) and thus also leading to differences in maximal T levels.

Table 1: Review of studies that investigated diurnal variation in testosterone levels.

group	species	period of highest T	study
diurnal mammals	Human *Homo sapiens*	early morning[1]	Evans et al. (1971)
		early morning	Barberia et al. (1973)
		early morning[1]	Aschoff (1979) and references therein
		night	Schulz et al. (1995)
		night	Luboshitzky et al. (2001, 2003)
		early morning	Walton et al. (2007)
	Bonnet Monkey *Macaca radiata*	night	Kholkute et al. (1993)
	Rhesus Monkey *Macaca mulatta*	night[2]	Plant (1981)
		night	Sehgal et al. (1986)
	Ring-tailed Lemur *Lemur catta*	night	Van Horn et al. (1976)
	Stallion *Equus caballus*	early morning	Sharma (1976)
	Boar *Sus scrofa*	afternoon and night	Claus and Giménez (1977)
	Dog *Canis lupus*	several peaks in T, but no clear diurnal rhythm	DePalatis et al. (1978)
	NZ Rabbit *Oryctolagus cuniculus*	peaks every 4 - 5 hours	Moor and Younglai (1975)
diurnal birds	Domestic Fowl *Gallus gallus domesticus*	night	Schanbacher et al. (1974)
	Ring Dove *Streptopelia risoria*	night	Balthazart et al. (1981)
	Blue Tit *Cyanistes caeruleus*	night[3]	Kempenaers et al. (2008)
	Lapland Longspur *Calcarius lapponicus*	for some males[4]: night	Hau et al. (2002)

	European Stonechat *Saxicola torquata rubicolor*	higher T metabolite concentration in faeces at night and in early morning	Goymann and Trappschuh (2011)
nocturnal mammals	Owl Monkey *Aotus trivirgatus*	day	Dixson and Gardner (1981)
	Thick-tailed Galago *Galago crassicaudatus*	day	Wilson et al. (1976)
	Lesser Mouse Lemur *Microcebus murinus*	day	Perret (1985)
	Djungarian Hamster *Phodopus sungorus*	day	Hoffmann and Nieschlag (1977)
	Sprague-Dawley Rat *Rattus norvegicus*	day	Wilson et al. (1976)
nocturnal birds	Indian Spotted Owlet *Athene brama*	day	Guchhait and Haldar (1999)
reptiles	Tuatara *Sphenodon punctatus*	no daily cycle	Cree et al. (1990)
insects	Sand Field Cricket *Gryllus firmus*	cycles in Juvenile Hormone[5] in flight-capable[6] morph with peaks before dark	Zera et al. (2007)

1) suggested to be somewhat sleep dependent
2) drop with lights
3) one study showed the opposite
4) higher T levels at night in males kept in Alaska, but not in Seattle; no change in Alaska females, but a trend in May in Seattle females
5) JH is the insect analogue to T (Nijhout 1994; Tibbetts and Huang 2010 and references therein)
6) JH might regulate aspects of flight

Material and Methods

Study population

A population of 150 male and 9 female House Sparrows was held at the Max Planck Institute for Ornithology, Seewiesen, Germany, in 1.2 x 2.0 x 4.0 m aviaries. At all times, the birds had *ad libitum* access to food, drinking and bathing water, and sand for dust-bathing. The light-dark cycle and temperature regime in the aviaries were close to natural conditions, as the aviaries were semi-outdoor with one side enclosed only by chicken wire. All sparrows were after-hatch year at the time of the study, and

were either caught in rural areas in Bavaria, Germany (under license: permit nr. 55.1-8642.3-3-2006 of the "Regierung Oberbayern", with several extensions) and held in captivity for at least eight months (136 males, 4 females) or raised in captivity (14 males, 5 females). Males raised in captivity did not have different T levels compared to birds that were caught in the wild and held in captivity (linear mixed effect model: $z = 0.45$, n = 710 observations, $p = 0.65$, random effects: bird ID (n = 150), season (n = 4), time of day (n = 66). Males were kept in groups of five or six per aviary, and females were housed together in one aviary around the time of sampling. Aviary ID only explained 4.6 % of variation in T levels and including it in models did not qualitatively change our results. For simplicity, we do not further include it. Further details about the study population can be found in Laucht et al. (2010).

Individual sampling and measuring

We caught all individuals during four periods over the course of an entire year: 26 Sept - 3 Nov 2006 ("fall"), 15 Jan – 2 Feb 2007 ("winter"), 8 - 16 March 2007 ("spring"), and 31 May - 22 June 2007 ("summer"). During each season, all birds were blood sampled twice, once during the day and once during the night, with a range of 4 -21 days recovery time between bleeding events. For the daytime sampling, we captured all males either in the morning or in the afternoon at approximately the same times and took biometric measurements, standardized photographs of the badge, and a blood sample. For the nighttime samples, we caught a subsample of the males we sampled during the day (36 individuals in fall, 36 individuals in winter, 17 individuals in spring, and 53 individuals in summer). In fall, winter and spring all individuals were caught between midnight and 1:00. In summer, we caught six groups of individuals between 22:00-04:30. At night in fall, winter and spring each time we caught a different set of birds. In summer, about half of the birds had been caught at night in the previous season. Excluding these birds did not qualitatively change the results of the day-night comparison. All nighttime sampling was performed at 45 min (for first summer samples) or more than five hours (other seasons) after sunset and all birds were roosting during this time.

In winter, half of the birds (n = 18) were sampled first at night and second during the day, while the other half were sampled first during the day, and second at night. We

did this to check for an influence of previous bleeding on subsequent T-values, and found there was no such effect (linear model: $t_{33} = -1.30$, $p = 0.204$). In the other 3 season, all nighttime samples were taken after daytime bleeding.

During nighttime sampling in fall, winter and summer, we sampled half of the birds immediately after waking them up and the other half approximately 30-60 minutes after they woke up. For the daytime samples in spring and summer, we first sampled birds from one group of aviaries and 30-60 minutes later sampled individuals from a second group of aviaries.

In the summer, we additionally tested whether diel changes in testosterone occurs in a small sample of females (n = 9). These individuals were sampled once in the morning (at 07:45) and once at night (at 01:45, 8 days later).

During each sampling event, we took 150-200 µl of blood from the wing vein within 15 min after first starting to catch the birds. We collected the blood in 75 mm Na-heparinised micro-haematocrit capillaries, centrifuged it at 13000 rpm for three minutes, separated the plasma, and stored it at -80°C until analysis. The time passed since first starting to catch birds did not have an influence on T levels (linear mixed effect models with season, daytime, and bird ID as random effects: day: $z = 1.13$, $p = 0.26$, $n = 551$ (Laucht et al. 2010); night: $z = 0.41$, $p = 0.68$, $n = 159$).

Determination of plasma T levels

Frozen plasma samples were sent to the endocrine laboratory of the Leibniz Institute for Zoo and Wildlife Research in Berlin, Germany, where T levels were determined blindly by enzyme immunoassays (for further details on the methods see Roelants et al. (2002); see also Laucht et al. (2010)). The inter-assay coefficient of variation (CV) for the enzyme immunoassay was 12.3% and the intra-assay CV was 9.0%. To calculate the true repeatability (i.e. the intra-class correlation coefficient) of measuring serum T-levels, we split 122 plasma samples of several males into duplicates right after centrifugation. Based on samples from the whole year, the repeatability of these plasma T estimates was $R = 0.967 \pm 0.006$ (SE) ($F_{121, 122} = 59.24$, $p < 0.001$). We assumed that all values of zero were below the detection limit (9 out

of 282 cases) and assigned them the lowest value measured (15 pg/ml). T levels are reported in pg per ml, but were ln-transformed for statistical analyses. Daytime T levels are the same as reported in Laucht et al. (2010).

Determination of badge size

During each season, we took four pictures of the black bib of each male. For each picture, we held the birds ventrally such that the throat and bib were stretched and presented to the camera. We rearranged the bird's position between each photograph. SL measured the size of the badge from the photographs by encircling and measuring the area of the bib in pixels using the program ImageJ 1.36b (Abramoff et al. 2004) and later converting it into cm^2 using an area standard present in each photograph. For analyses, we used the average of all four pictures for each bird. These scores were highly repeatable within individuals ($R = 0.943$, estimated according to Falconer and Mackay (Falconer and Mackay 1996) from repeatability of single pictures). See Laucht et al. (2010) for additional details.

Statistical analyses

We performed all statistical analyses using R 2.8.0 (R Development Core Team 2008) (packages: lme4, nlme, RODBC) at the significance level $\alpha = 0.05$. We compared day and night T levels in each period both at the population level (including those individuals for which only day T-values were available; t-test) and at the individual level (paired t-test). To analyze the relationship between badge size or disturbance and plasma T levels, we used linear models and linear mixed effect models as indicated in the results. There was no need for model simplification. For analyses on badge size, we used the summer scores (means of the four photos), because badge size changed due to abrasion of the white feather edges in fall and early spring (Møller and Erritzøe 1992). The results did not qualitatively change when using means of spring and summer badge scores.

Note that sample sizes differ due to missing data points for single individuals and due to the exclusion of males in breeding aviaries for the summer samples.

Results

Day/night variation in T levels

In males, plasma T levels were on average 1.3 - 4.8 times higher during the night than during the day and this was significant during all four seasons (Fig. 1a & b, Table 2). However, the within-individual correlation between day and night levels was rather weak (Spearman rank correlation: fall: rho = 0.29, n = 36, p = 0.091; winter: rho = 0.38, n = 35, p = 0.023; spring: rho = -0.12, n = 17, p = 0.640; summer: rho = 0.27, n = 53, p = 0.054) indicating that males with the highest day levels did not necessarily have the highest night levels (Fig. 1b). Female plasma T levels were also significantly higher at night than during the day, at least in summer (Fig. 1a, Table 2).

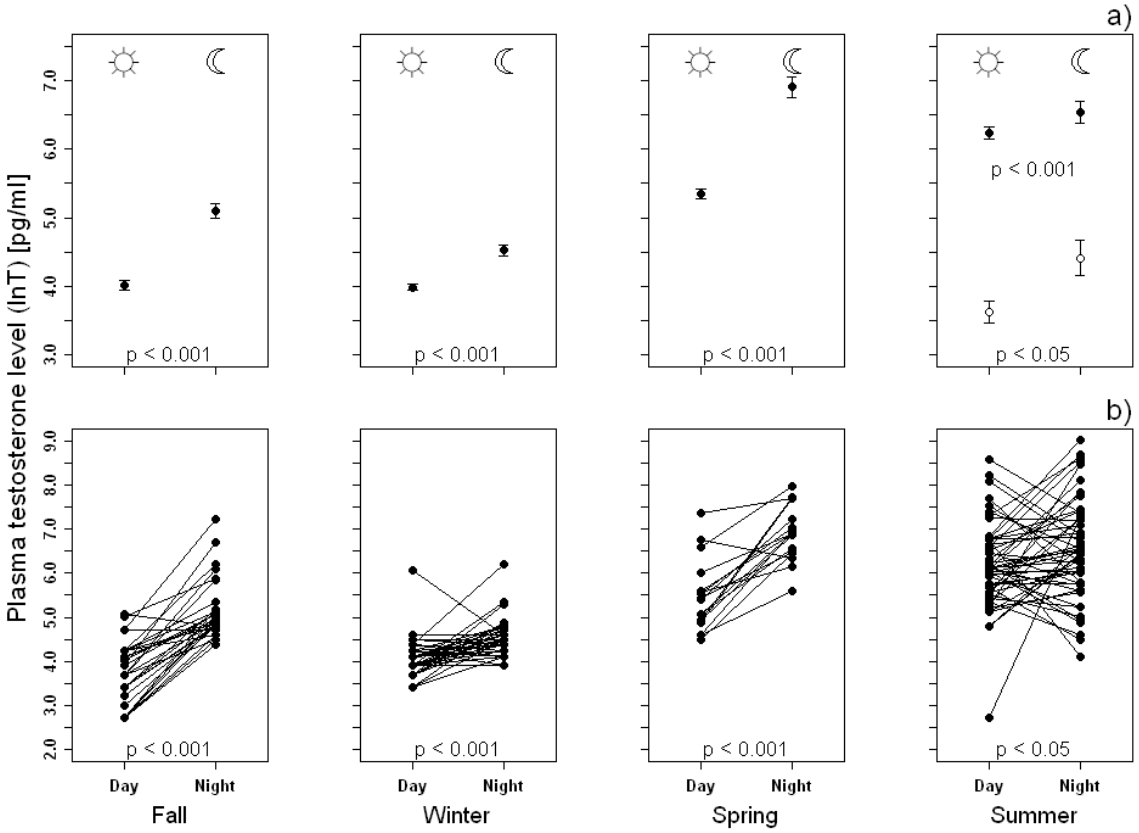


Fig. 1 Day and night plasma testosterone levels of male (closed circles) and female (open circles) House Sparrows at four different seasons.
a) at the population level: mean ± standard error
b) at the individual level
For statistics and sample sizes see Table 2.

During daytime there was no effect of previous disturbance on T levels (Fig. 2; linear models: in spring, in summer, or in both periods combined; all $p > 0.59$). However, at night T levels were significantly lower when birds had been awake for 30-60 min before being sampled (Fig. 2, Table 3; linear mixed effects model: all periods combined: $z = -3.12$, $n = 124$ observations, $p = 0.002$, crossed random effects: bird ID ($n = 100$ individuals) and season ($n = 3$)). The effect was observed in every season (Fig. 2), but it was only significant in winter (t-tests; fall: $t_{34} = -1.80$, $p = 0.082$; winter: $t_{33} = -3.22$, $p = 0.003$; summer: $t_{51} = -1.79$, $p = 0.081$).

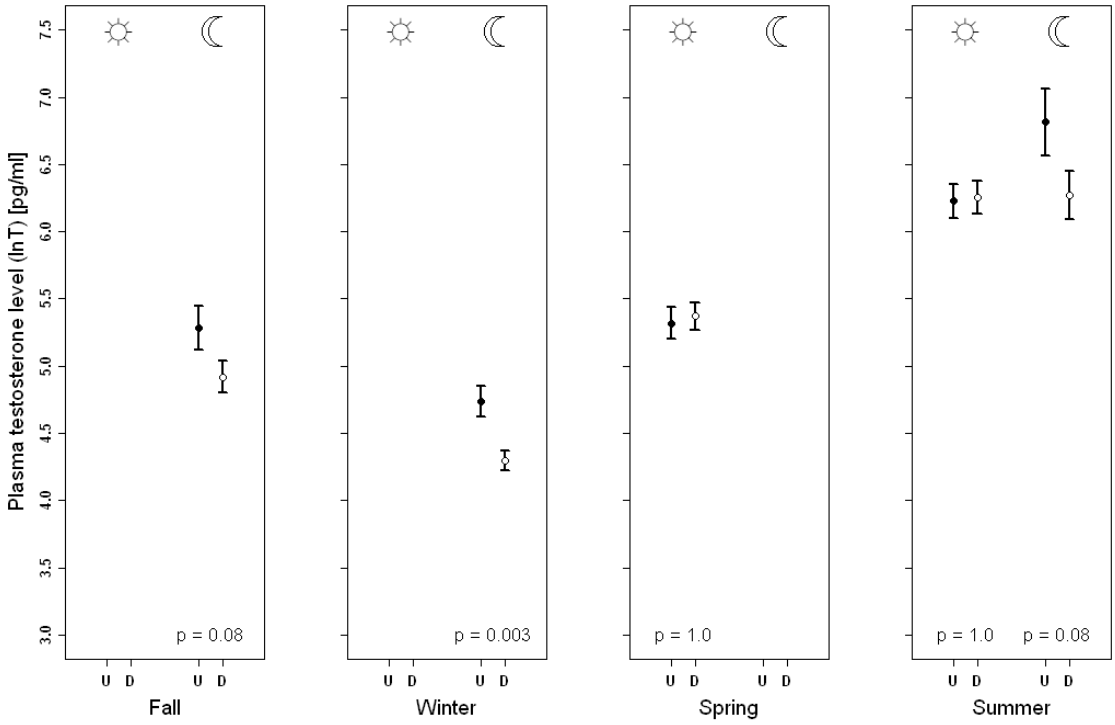


Fig. 2 Plasma testosterone levels of male House Sparrows at three different seasons in undisturbed (U, closed circles) and disturbed (D, open circles) groups. Presented are means ± standard errors. For statistics and sample sizes see Table 3.

Table 2: Seasonal and daily variation in plasma T levels of male and female House Sparrows. Shown are the geometric mean (± standard errors) values (in pg/ml) during the day and the night, population level t-test, and individual level (paired) t-tests comparing day and night values. Because all analyses were performed on log-transformed values, geometric means (the back-transformed mean of log-transformed values) are reported here. Standard errors were calculated as the back-transformed differences of means plus standard errors of log transformed T values. Overall mean was calculated as the means of seasonal means. Means population level statistics were calculated from all birds sampled but for the individual level statistics only those individuals that were sampled during day and night were included.

		means ± SE of T				**population level**		**individual level**	
		day	n	night	n	statistic	p-value	statistic	p-value
males	fall	55.22 ± 3.96	142	164.38 ± 18.18	36	t68.9 = 8.67	< 0.0001	t35 = 12.45	< 0.0001
	winter	53.66 ± 2.33	140	91.99 ± 7.35	35	t56.7 = 6.14	< 0.0001	t34 = 4.55	< 0.0001
	spring	209.46 + 16.62	139	1003.77 + 164.54	17	t24.9 = 9.22	< 0.0001	t16 = 6.23	< 0.0001
	summer	522.21 ± 39.84	103	690.43 ± 116.02	53	t74.4 = 13.2	< 0.0001	t52 = 2.17	0.035
females	summer	37.64 + 6.55	9	82.61 + 23.45	9	t13.6 = 2.65	0.019	t8 = 2.67	0.028

Table 3: Variation in night plasma T levels of male House Sparrows in disturbed and undisturbed groups. Shown are the geometric mean (± standard errors) values (in pg/ml), results of t-tests comparing levels of disturbed and undisturbed groups, and results of t-test comparing disturbed groups and daytime levels. Because all analyses were performed on log-transformed values, geometric means (the back-transformed mean of log-transformed values) are reported here. Standard errors were calculated as the back-transformed differences of means plus standard errors of log transformed T values.

	Mean ± SE of T				**disturbed-undisturbed**		**disturbed-day**	
	undisturbed	n	disturbed	n	statistic	p-value	statistic	p-value
overall	319.62 + 57.91	62	207.75 + 30.88	62	t118.12 = 1.99	0.049	t79.64 = 3.02	0.003
fall	197.41 + 35.52	18	136.87 + 17.20	18	t30.79 = 1.80	0.082	t30.23 = 6.62	< 0.0001
winter	113.83 + 13.77	18	73.42 + 5.37	17	t28.12 = 3.27	0.003	t29.28 = 3.80	0.0007
summer	911.84 + 258.79	26	528.19 + 102.22	27	t45.39 = 1.78	0.081	t35.55 = 0.06	0.953

Night T levels and badge size

Badge size was related to night plasma T levels in summer (linear model: $t_{43} = 2.54$, R = 0.36, p = 0.015; Fig. 3), but not significantly in any of the other sampled seasons (all p > 0.48). Because badge size was not related to day plasma T levels in June (Fig. 3, data from Laucht et al. (2010)), we further tested whether the relationship between T levels and badge size differed depending on the time when plasma samples were collected. This was indeed the case (interaction with time period: linear mixed effect model: z = 2.54, p = 0.015, n = 180 observations, random effect: bird ID (n = 135 individuals); Fig. 3).

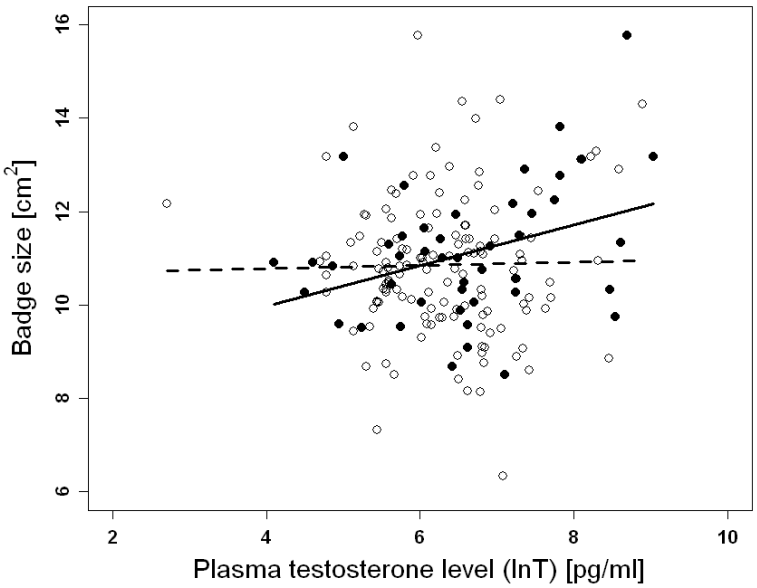


Fig. 3 Badge size in relation to June plasma night (closed circles, straight line; R = 0.35, t_{43} = 2.44, p = 0.019) and day (open circles, dashed line; R = 0.03, t_{133} = 0.31, p = 0.76) testosterone levels for male House Sparrows. We defined badge size as the June measurements.
The relationship between T levels and badge size differed between day and night: interaction with time period: linear mixed effect model: $t_{43} = 2.54$, p = 0.015, n = 180, random effect: bird ID.

Discussion

We found that in captive male and female House Sparrows plasma T levels were significantly higher during the night than during the day. T levels on average doubled, demonstrating dramatic daily within-individual fluctuations in plasma T levels. The higher night T levels were observed in males throughout the year. These results are in accordance with findings from diurnal mammals, and some diurnal bird species (Table 1). The only other study that has examined diel variation in T levels in female birds also found a trend towards higher night levels (Hau et al. 2002). However, the sample size in this as well as in our study was small and birds were only sampled at one time point during the night during one season. Overall, however, our results indicate strongly that nocturnal increases in T levels occur in diurnal species, independent of season and sex. Hence, there seems to be a general mechanism that causes T levels to increase at night.

Elevated nocturnal testosterone and sleep

Higher T levels may be associated with sleeping, as opposed to night-time per se. Studies on human males showed a correlation between the first REM sleep and the increase in T levels (Evans et al. 1971; Luboshitzky et al. 1999) and a significant delay in this T rise when sleep was fragmented (Luboshitzky et al. 2001) or shifted to daytime (Evans et al. 1971; Boyar et al. 1974). Consistent with the hypothesis that increased T levels are linked to sleep and not to night-time per se, is the observation that nocturnal animals show highest T levels during the day (Table 1). Our results are also consistent with this hypothesis: we found that T levels at night were lower when the birds were sampled 30-60 min after being disturbed (Fig. 2), suggesting a quick drop in T levels after waking up. An alternative explanation is that T levels dropped due to increased stress associated with disturbance. However, we would then have expected a decrease in T levels after disturbance during the day, which we did not observe (Fig. 2). Nevertheless, it remains possible that the House Sparrows experienced disturbance during the night as a much stronger stressor, and that this explains the difference in T level changes.

An alternative explanation for higher night T levels in diurnal animals, is that T levels are affected by social interactions, assuming that such interactions generally lead to a decrease in plasma T levels. Lower plasma T levels are expected in individuals that lose out in competitive interactions (e.g. subordinate individuals) (reviewed in Adkins-Regan 2005). Under this scenario, we would expect individuals with low day T levels (e.g. subordinate individuals) to show the strongest increase during the night, whereas those with the highest T levels should not show a further increase. One would then expect a lower variance in T levels at night than during the day, for which we found no evidence (F-test for equality of variances: each season analysed separately, $p = 0.31$-0.93). However, we note that particularly during the breeding season not all individuals had higher T levels at night (Fig. 1b). More detailed studies on the effects of sleep, (nocturnal) stress, and social interactions on within-individual variation in T levels are needed.

Function of elevated nocturnal testosterone

The patterns observed here and in other species (Table 1) raise an intriguing and important question. What is the functional significance – if any – of the nocturnal increase in plasma T? T may simply accumulate during the night (i.e. accumulation occurring due to diel changes in the half life of testosterone or of its binding proteins, in testosterone secretion or in the secretion of other hormones such as LH or prolactin), either because it is not used up (non-functional) or because individuals physiologically prepare for a (functionally) high T level in the early morning. The highest T levels should then be found when morning activity starts, which is indeed true for human males (highest levels found just before sunrise: Barberia et al. 1973; Schulz et al. 1995; Luboshitzky et al. 2003; Walton et al. 2007). However, other studies (Table 1) suggest that in some species T levels may fluctuate throughout the night, rather than gradually increase until early morning.

Alternatively, increased nocturnal T may itself be adaptive. A change in the secretion of T (rather than metabolic clearance rate) (Walton et al. 2007) could be functional, if higher T levels are required for short-term organizational functions such as neuronal development or memory consolidation, functions also suggested for sleep (Stickgold 2005). Additionally, the observation that nocturnal increase in T levels occurred

independently of season also suggests that T might play a role in regulating diurnal changes in physiology, i.e. function as a "sleep hormone".

An alternative organizational function of increased nocturnal T is that it is associated with increased spermatogenesis. Early studies on male House Sparrows found that spermatogenesis takes place at night with greatest activity between 02:00 and 04:00 (Foley 1929; Allender 1936; Riley 1937). It was further suggested that it is related to the drop of body temperature that occurs during sleep, which itself might be triggered by increased T (Feuerbacher and Prinzinger 1981). In Bonnet Monkeys (*Macaca radiata*), long term suppression of night-time T peaks had a negative influence on testis activity in general and on spermatogenesis in particular (Suresh and Moudgal 1995). However, we also found elevated night T levels in males during the fall when the testes are regressed, and in females. Although this does not refute the idea that one function of elevated night T is increased spermatogenesis; it is clear that other functional explanations cannot be excluded.

Night-time testosterone and ornamentation

We found a significant correlation between badge size of male House Sparrows and night-time plasma T levels in the peak breeding season (June). This contrasts with our previous finding (on the same population) that badge size was not correlated with daytime T levels in any season (Laucht et al. 2010), but is in agreement with other studies that found correlations between badge size and post-breeding or breeding daytime T levels (Buchanan et al. 2001; Gonzalez et al. 2001), and an effect of artificially increased T levels on badge size (Evans et al. 2000).

Previous studies suggest that – in House Sparrows and other bird species – agonistic interactions or challenges cause a short-term increase in T levels above breeding baseline levels (the "Challenge Hypothesis") (Hegner and Wingfield 1986; Wingfield et al. 1987; Wingfield et al. 1990; Hill 2002; McGlothlin et al. 2008). We suggest that night-time T levels reflect maximum T levels achieved during challenges and that these levels reflect competitive ability better than day levels do. Under the hypothesis that night-time T levels indeed reflect maximum T levels (see also below), the relationship between night-time T levels and badge size suggests that badge size could

be an "honest" signal of status. Maximum T levels during challenges could maintain signal honesty via social costs (i.e. keep the signal evolutionarily stable by preventing cheating via high costs; the badge of status hypothesis (Maynard Smith and Harper 1988; Jawor and Breitwisch 2003; Tibbetts and Dale 2004)). Badge size could therefore indicate an individual's dominance, level of aggression, and ability to defend itself in agonistic interactions during the breeding season.

Although the relationship between night-time T levels and maximal T levels during social challenges needs to be tested directly, there is indirect evidence that they may be similar. An experimental paradigm that is often used to estimate an individual's maximum T level is an injection with gonadotropin-releasing hormone (GnRH) (Wingfield and Farner 1993; Jawor et al. 2006). Previous studies showed that plasma T levels measured after GnRH challenge correlate positively with elevated T levels after social challenges (McGlothlin et al. 2008). Studies in several bird species and in several seasons showed that plasma T levels on average increased 2.1 fold (range: 1.0-5.7 fold) after GnRH challenge (Hirschenhauser et al. 2000; Moore et al. 2002; Jawor et al. 2006; Jawor et al. 2007). Our results indicate that plasma T levels increased on average 2.6 fold (range 1.3-4.8) from day to night (all seasons, males and females), which is similar to the effects shown by the GnRH challenge.

Conclusions

In summary, we found that plasma T levels of male and female House Sparrows were much higher at night than during the day, and we provide evidence that higher nocturnal T could be associated with sleep because disturbance at night, but not during the day, reduces T levels. Additionally, we found that male badge size, an ornament generally associated with dominance, was related to night-time T levels, but not to daytime levels. Overall, our results imply that diel cycles need to be considered in studies using measurements of plasma T levels.

Acknowledgements

Many thanks to the following people for their help with taking blood samples during day and night: Andrea Wittenzellner, Alexandra Hoffmann, Hendrik Reers, Elizabeth

Yohannes, Mihai Valcu, Emily DuVal, Silke Steiger, Elisabeth Bolund. We are grateful to Mihai Valcu for statistical advice, to Magnus Rudorfer, Edith Bodendorfer, Annemarie Grötsch, and Frances Weigel for animal care, and to the Behavioural Ecology and Evolutionary Genetics Group at the MPIO for discussion and comments. Wolfgang Goymann, Michaela Hau, and two anonymous reviewers provided helpful comments on an earlier version of the manuscript. Funding for this research was provided by the Max Planck Society.

References

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophotonics Int. 11:36-42

Adkins-Regan E (2005) Hormones and Animal Social Behavior. Princeton University Press, Princeton, NJ

Allender C (1936) Microscopical observations on the gonadal cycle of the English sparrow. Trans. Am. Microsc. Soc. 55:243-249

Anderson TR (2006) Biology of the ubiquitious house sparrow. Oxford University Press, New York

Aschoff J (1979) Circadian Rhythms: General Features and Endocrinological Aspects. In: Krieger DT (ed) Endocrine Rhythms. Raven Press, New York, pp 1-61

Balthazart J, Reboulleau C, Cheng MF (1981) Diurnal-Variations of Plasma Fsh, Lh, and Testosterone in Male Ring Doves Kept under Different Photoperiods. Gen. Comp. Endocrinol. 44:202-206

Barberia JM, Giner J, Cortes-Gallegos.V (1973) Diurnal-Variations of Plasma Testosterone in Men. Steroids 22:615-626

Boyar RM, Rosenfeld RS, Kapen S, Finkelstein JW, Roffwarg HP, Weitzman ED, Hellman L (1974) Human Puberty - Simultaneous Augmented Secretion of Luteinizing Hormone and Testosterone during Sleep. J. Clin. Investig. 54:609-618

Buchanan KL, Evans MR, Goldsmith AR, Bryant DM, Rowe LV (2001) Testosterone influences basal metabolic rate in male house sparrows: a new cost of dominance signalling? Proc. R. Soc. Biol. Sci. Ser. B 268:1337-1344

Candolin U (2003) The use of multiple cues in mate choice. Biol. Rev. (Camb.) 78:575-595

Claus R, Giménez T (1977) Diurnal Rhythm of 5alpha-Androst-16-En-3-One and Testosterone in Peripheral Plasma of Boars. Acta Endocrinol. 84:200-206

Cree A, Guillette LJ, Cockrem JF, Brown MA, Chambers GK (1990) Absence of Daily Cycles in Plasma Sex Steroids in Male and Female Tuatara (*Sphenodon punctatus*), and the Effects of Acute Capture Stress on Females. Gen. Comp. Endocrinol. 79:103-113

DePalatis L, Moore J, Falvo RE (1978) Plasma Concentrations of Testosterone and LH in the Male Dog. J. Reprod. Fertil. 52:201-207

Dixson AF, Gardner JS (1981) Diurnal-Variations in Plasma Testosterone in a Male Nocturnal Primate, the Owl Monkey (*Aotus trivirgatus*). J. Reprod. Fertil. 62:83-86

Evans JI, Maclean AM, Ismail AAA, Love D (1971) Circulating Levels of Plasma Testosterone during Sleep. Proc. R. Soc. Med. 64:841-842

Evans MR, Goldsmith AR, Norris SRA (2000) The effects of testosterone on antibody production and plumage coloration in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 47:156-163

Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics, 4th edn. Pearson Education Limited, Harlow, Essex

Feuerbacher I, Prinzinger R (1981) The Effects of the Male Sex-Hormone Testosterone on Body-Temperature and Energy-Metabolism in Male Japanese Quail (*Coturnix coturnix japonica*). Comp. Biochem. Physiol. A 70:247-250

Foley JO (1929) A note on the spermatogenetic wave in the testes of the adult English sparrow (*Passer domesticus*). Anat. Rec. 41:367-371

Gonzalez G, Sorci G, Smith LC, de Lope F (2001) Testosterone and sexual signalling in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 50:557-562

Goymann W, Trappschuh M (2011) Seasonal and diel variation of hormone metabolites in European stonechats -on the importance of high signal-to-noise-ratios in non-invasive hormone studies. J. Biol. Rhythms 26:44-54

Guchhait P, Haldar C (1999) Circadian rhythms of melatonin and sex steroids in a nocturnal bird, Indian spotted owlet *Athene brama* during reproductively active and inactive phases. Biol. Rhythm. Res. 30:508-516

Hau M, Romero LM, Brawn JD, Van't Hof TJ (2002) Effect of polar day on plasma profiles of melatonin, testosterone, and estradiol in high-Arctic Lapland Longspurs. Gen. Comp. Endocrinol. 126:101-112

Hegner RE, Wingfield JC (1986) Behavioral and Endocrine Correlates of Multiple Brooding in the Semicolonial House Sparrow *Passer domesticus*. 1. Males. Horm. Behav. 20:294-312

Hill GE (2002) A red bird in a brown bag: The function and evolution of colorful plumage in the House Finch. Oxford University Press, Inc., New York

Hirschenhauser K, Möstl E, Péczely P, Wallner B, Dittami J, Kotrschal K (2000) Seasonal relationships between plasma and fecal testosterone in response to GnRH in domestic ganders. Gen. Comp. Endocrinol. 118:262-272

Hoffmann K, Nieschlag E (1977) Circadian-Rhythm of Plasma Testosterone in Male Djungarian Hamster (*Phodopus sungorus*). Acta Endocrinol. 86:193-199

Jawor JM, Breitwisch R (2003) Melanin ornaments, honesty, and sexual selection. Auk 120:249-265

Jawor JM, McGlothlin JW, Casto JM, Greives TJ, Snajdr EA, Bentley GE, Ketterson ED (2006) Seasonal and individual variation in response to GnRH challenge in male dark-eyed juncos (*Junco hyemalis*). Gen. Comp. Endocrinol. 149:182-189

Jawor JM, McGlothlin JW, Casto JM, Greives TJ, Snajdr EA, Bentley GE, Ketterson ED (2007) Testosterone response to GnRH in a female songbird varies with stage of reproduction: implications for adult behaviour and maternal effects. Funct. Ecol. 21:767-775

Kempenaers B, Peters A, Foerster K (2008) Sources of individual variation in plasma testosterone levels. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363:1711-1723

Kholkute SD, Katkam RR, Puri CP (1993) Lack of Seasonality in Day and Night Serum Levels of Testosterone and Dihydrotestosterone, Testicular Size, and Body-Weight of Adult Bonnet Monkeys (*M. radiata*) Maintained in Captivity. Primates 34:71-76

Laucht S, Kempenaers B, Dale J (2010) Bill color, not badge size, indicates testosterone-related information in house sparrows. Behav. Ecol. Sociobiol. 64:1461-1471

Liker A, Barta Z (2001) Male badge size predicts dominance against females in House Sparrows. Condor 103:151-157

Luboshitzky R, Herer P, Levi M, Shen-Orr Z, Lavie P (1999) Relationship between rapid eye movement sleep and testosterone secretion in normal men. J. Androl. 20:731-737

Luboshitzky R, Shen-Orr Z, Herer P (2003) Middle-aged men secrete less testosterone at night than young healthy men. J. Clin. Endocrinol. Metab. 88:3160-3166

Luboshitzky R, Zabari Z, Shen-Orr Z, Herer P, Lavie P (2001) Disruption of the nocturnal testosterone rhythm by sleep fragmentation in normal men. J. Clin. Endocrinol. Metab. 86:1134-1139

Maynard Smith JFRS, Harper DGC (1988) The Evolution of Aggression - Can Selection Generate Variability? Philos. Trans. R. Soc. Lond. B Biol. Sci. 319:557-570

McGlothlin JW, Jawor JM, Greives TJ, Casto JM, Phillips JL, Ketterson ED (2008) Hormones and honest signals: males with larger ornaments elevate testosterone more when challenged. J. Evol. Biol. 21:39-48

McGraw KJ, Dale J, Mackillop EA (2003) Social environment during molt and the expression of melanin-based plumage pigmentation in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 53:116-122

Møller AP (1987) Variation in Badge Size in Male House Sparrows *Passer domesticus* - Evidence for Status Signaling. Anim. Behav. 35:1637-1644

Møller AP (1990) Sexual-Behavior Is Related to Badge Size in the House Sparrow *Passer domesticus*. Behav. Ecol. Sociobiol. 27:23-29

Møller AP, Erritzøe J (1992) Acquisition of Breeding Coloration Depends on Badge Size in Male House Sparrows *Passer domesticus*. Behav. Ecol. Sociobiol. 31:271-277

Moor BC, Younglai EV (1975) Variations in Peripheral Levels of LH and Testosterone in Adult Male Rabbits. J. Reprod. Fertil. 42:259-266

Moore IT, Perfito N, Wada H, Sperry TS, Wingfield JC (2002) Latitudinal variation in plasma testosterone levels in birds of the genus *Zonotrichia*. Gen. Comp. Endocrinol. 129:13-19

Morrison EB, Kinnard TB, Stewart IRK, Poston JP, Hatch MI, Westneat DF (2008) The links between plumage variation and nest site occupancy in male House Sparrows. Condor 110:345-353

Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T (2007) Assessing the function of house sparrows' bib size using a flexible meta-analysis method. Behav. Ecol. 18:831-840

Nijhout HF (1994) Insect hormones. Princeton University Press, Princeton, New Jersey

Perret M (1985) Diurnal-Variations in Plasma Testosterone Concentrations in the Male Lesser Mouse Lemur (*Microcebus murinus*). J. Reprod. Fertil. 74:205-213

Plant TM (1981) Time Courses of Concentrations of Circulating Gonadotropin, Prolactin, Testosterone, and Cortisol in Adult Male Rhesus-Monkeys (*Macaca mulatta*) Throughout the 24 H Light-Dark Cycle. Biol. Reprod. 25:244-252

R Development Core Team (2008) R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria

Riley GM (1937) Experimental studies on spermatogenesis in the house sparrow, *Passer domesticus* (Linnaeus). Anat. Rec. 67:327-351

Roberts ML, Buchanan KL, Evans MR (2004) Testing the immunocompetence handicap hypothesis: a review of the evidence. Anim. Behav. 68:227-239

Roelants H, Schneider F, Goritz F, Streich J, Blottner S (2002) Seasonal changes of spermatogonial proliferation in roe deer, demonstrated by flow cytometric analysis of c-kit receptor, in relation to follicle-stimulating hormone, luteinizing hormone, and testosterone. Biol. Reprod. 66:305-312

Schanbacher BD, Gomes WR, VanDemark NL (1974) Diurnal Rhythm in Serum Testosterone Levels and Thymidine Uptake by Testes in Domestic Fowl. J. Anim. Sci. 38:1245-1248

Schulz P, Chardon F, Agosti D, Schaad N, Rivest RW (1995) Parallel Nocturnal Secretion of Melatonin and Testosterone in the Plasma of Normal Men. J. Pineal Res. 19:16-22

Sehgal A, David GFX, Dubey AK, Pruthi JS, Kumar TCA (1986) Circadian and Circannual Changes in the Testicular Function of Adult Rhesus Monkeys. J. Biosciences 10:443-452

Sharma OP (1976) Diurnal-Variations of Plasma Testosterone in Stallions. Biol. Reprod. 15:158-162

Stickgold R (2005) Sleep-dependent memory consolidation. Nature 437:1272-1278

Suresh R, Moudgal NR (1995) A Role for Nocturnal Serum Testosterone Surge in Regulating Spermatogenesis in the Adult Non-human Primate. Endocrine 3:487-492

Tibbetts EA, Dale J (2004) A socially enforced signal of quality in a paper wasp. Nature 432:218-222

Tibbetts EA, Huang ZY (2010) The Challenge Hypothesis in an Insect: Juvenile Hormone Increases during Reproductive Conflict following Queen Loss in *Polistes* Wasps. Am. Nat. 176:123-130

Van Horn RN, Beamer NB, Dixson AF (1976) Diurnal Variations of Plasma Testosterone in two Prosimian Primates (*Galago crassicaudatus* and *Lemur catta*). Biol. Reprod. 15:523-528

Veiga JP (1993) Badge Size, Phenotypic Quality, and Reproductive Success in the House Sparrow - a Study on Honest Advertisement. Evolution 47:1161-1170

Vizcarra JA, Kreider DL, Kirby JD (2004) Episodic gonadotropin secretion in the mature fowl: Serial blood sampling from unrestrained male broiler breeders (*Gallus domesticus*). Biol. Reprod. 70:1798-1805

Walton MJ, Anderson RA, Kicman AT, Elton RA, Ossowska K, Baird DT (2007) A diurnal variation in testicular hormone production is maintained following gonadotrophin suppression in normal men. Clin. Endocrinol. 66:123-129

Weitzman ED (1976) Circadian Rhythms and Episodic Hormone Secretion in Man. Annu. Rev. Med. 27:225-243

Wilson MJ, McMillin JM, Seal US, Ahmed K (1976) Circadian Variation of Serum Testosterone in Adult Male Rat with a Late Morning Acrophase. Experientia 32:944-945

Wingfield JC, Ball GF, Dufty AM, Hegner RE, Ramenofsky M (1987) Testosterone and Aggression in Birds. Am. Sci. 75:602-608

Wingfield JC, Farner DS (1993) Endocrinology of reproduction in wild species. In: Farner DS, King JR, Parker KC (eds) Avian biology. Academic Press Ltd, New York, pp 163-327

Wingfield JC, Hegner RE, Dufty AM, Ball GF (1990) The Challenge Hypothesis - Theoretical Implications for Patterns of Testosterone Secretion, Mating Systems, and Breeding Strategies. Am. Nat. 136:829-846

Zera AJ, Zhao Z, Kaliseck K (2007) Hormones in the field: Evolutionary endocrinology of juvenile hormone and ecdysteroids in field populations of the wing-dimorphic cricket *Gryllus firmus*. Physiol. Biochem. Zool. 80:592-606

Non-submitted Manuscript

Chapter Three

Condition, testosterone and age correlations with multiple ornaments in male House Sparrows: patterns and implications

Silke Laucht, Bart Kempenaers, James Dale

Abstract

Why many animal species display several different kinds of ornaments has been a question of recent interest to researchers of animal communication. Originally, three key hypotheses that explain the evolution of multiple signals have been suggested: they can function as (1) multiple messages (different information for each ornament), (2) redundant messages or back-up signals (the same information for all ornaments), and (3) unreliable signals (no information). Even though these hypotheses were originally developed for signals of quality used in mate choice, they are also broadly applicable to other contexts of quality signalling such as rival competition. Here we tested the signal properties of ornamental traits in male House Sparrows (*Passer domesticus*) in order to better resolve the potential information that different ornaments reveal. We repeatedly measured badge size, wingbar conspicuousness, bill colour, and leg colour in 175 captive males over 3 years. We tested how these ornamental traits correlated with each other and to physical condition, plasma testosterone, age, and place of origin (i.e. wild or captive born). We found that (1) badge size and wingbar area were related to condition, (2) badge size, wingbar area, and bill colour were related to testosterone levels, (3) all ornaments except bill colour were related to age, and (4) origin had only minor effects. We conclude that badge size, wingbar area, and bill colour have the potential to function as multiple messages of different testosterone related behaviours, and that badge size, wingbar area, and leg colour have the potential to function as back-up signals of age. However, the empirical classification of ornaments as multiple messages or back-up signals is not clear, and we suggest instead it is best to view multiple ornaments as providers of varying degrees of various overlapping types of information.

Introduction

Many animals show several pronounced ornaments that are thought to be important in the context of mate choice (reviewed in Candolin 2003). Why do animals frequently employ more than one ornament when it would (presumably) be easier to provide a single distinct and easily discriminable signal to the receiver? In a groundbreaking paper, Møller and Pomiankowski (1993) proposed three hypotheses to explain the functional advantages of multiple ornaments: (1) The "multiple messages" hypothesis posits that different signals provide information of different aspects of signaller quality, such as different kinds of qualities (e.g. access to resources versus genetic immunity to current contagions) or condition on different time scales (Møller and Pomiankowski 1993; Johnstone 1996). Receivers can then either evaluate the signals together or selectively. Or alternatively the different signals could be directed to different receivers (Andersson et al. 2002). (2) The "back-up" (or redundant) signals hypothesis posits that multiple signals communicate the same aspect of quality. This assumes that each single ornament includes some error, and the receiver uses the combined stimuli of all ornaments to infer a more accurate estimate of this quality (Møller and Pomiankowski 1993; Johnstone 1996). Finally, (3) the "unreliable" (or uninformative) signal hypothesis predicts that multiple ornaments do not signal quality, but have instead evolved from an arbitrary preference, such as via the Fisherian runaway process (known as Fisherian cues), or to ease signal detection and assessment (Møller and Pomiankowski 1993; Candolin 2003). Overall, multiple ornaments could be advantageous over one single ornament as long as costs for the assessment of additional ornaments are not much higher than for a single ornament (Iwasa and Pomiankowski 1994; Johnstone 1995; Johnstone 1996; Candolin 2003). More recently, additional hypotheses have included other aspects of mate choice such as interference of males by hindering female choice or female resistance to male signals (Holland and Rice 1998; reviewed in Candolin 2003; Lozano 2009) and additional types of information such as behavioural strategy, or individual identity (Dale 2006).

Originally, all explanations for multiple ornaments were developed with mate choice in mind as the driving force. These hypotheses therefore neglected another important aspect of honest signalling: rival competition. Up to now, very few studies have

considered multiple ornaments in the context of pure competitive interactions (Balph et al. 1979; Bókony et al. 2006; Chaine et al. 2011). However, Berglund et al. (1996) suggested that single traits with a dual signalling function (i.e. signals that are used as both armaments in male-male competition and ornaments in mate choice) could evolve from signals that are kept honest only through competitive interactions. Additionally, Anderson et al. (2002) described a version of the multiple messages hypothesis called the "multiple receiver hypothesis" which states that multiple ornaments could be assessed by multiple receivers (i.e. mates and rivals) and thus signalling different contents to different receivers. Taken together, this suggests that multiple signals could reveal a huge variety of different information in addition to the ones originally described, important both in male-male and male-female interactions (Kekäläinen et al. 2010). This includes information about an individual's quality (genetic and phenotypic constitution), age, testosterone dependent behavioural traits (such as status, aggressiveness and breeding state), and combinations of these (Dale 2006).

In light of this, there are several reasons why it is of great value to perform detailed analyses of the characteristics that correlate with different ornaments within a single species over multiple years. Although correlational, this approach is valuable because it provides a detailed window into what different kinds of information are *potentially* available to any receivers of the ornaments at different time points over the year and at different life history stages of the signaller. For example, Freeman-Gallant et al. (2010) recently suggested that different ornaments in the Common Yellowthroat (*Geothlypis trichas*) signal redundant information at different life stages, demonstrating the importance of long-term studies and of the consideration of an individual's age. In addition, age can also reveal information about the role of different quality parameters on the development of traits that, once fully developed, signal the same information for a long time (e.g. song type (Rivera-Gutierrez et al. 2010)). Furthermore, for a powerful approach it is crucial to include all potential ornaments (Galván 2010) and to examine a large number of individuals. The examination of the relationship with hormone levels, in particular testosterone, is fundamental given that many ornaments are testosterone related (e.g. reviewed in Roberts et al. 2004; McGlothlin et al. 2008) or that many behaviours and physiological processes associated with male-male and male-female interactions are

testosterone dependent (summarized in Ketterson and Nolan 1992; Hau 2007). Finally, different parameters of body condition should be considered because they can give insights into the actual physical state and health of an individual (e.g. summarized in Schulte-Hostedde et al. 2005). Studies on multiple ornaments should therefore be based on a much broader framework and include as many parameters and individuals as possible.

Here, we studied four very different ornaments in a model species of visual communication research - the House Sparrow (*Passer domesticus*). Three of these ornaments (badge size, wingbar conspicuousness, and bill colour) are rather well-studied, and one of them (leg colour) is new. First, the size of the black breast bib or the "badge" has been found to correlate with social status, fighting success, age, variation in sexual behaviour, and plasma levels (Møller 1987; Møller 1990; Veiga 1993; Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001; Liker and Barta 2001; McGraw et al. 2003; Bókony et al. 2006; Nakagawa et al. 2007; Morrison et al. 2008; Laucht et al. 2011). Secondly, wingbar size was found to signal body condition (Poston et al. 2005), parasite resistance (Moreno-Rueda 2010), and competitive defence success (Bókony et al. 2006). Third, bill coloration is a conspicuous and dynamic testosterone dependent trait that changes from a pale horn coloration during the non-breeding season to black during the breeding season (Keck 1933; Witschi 1936; Witschi and Woods 1936; Pfeiffer et al. 1944; Haase 1975; Donham et al. 1982). Outside of breeding, variation in bill colour is correlated to plasma testosterone levels and therefore also a possible signal of testosterone related behaviours during these times (Laucht et al. 2010). Fourth, we include here a hitherto unstudied, putative ornament: leg colour. Leg colour is melanin based and ranges from a very pale horn colour to a brownish coloration. Because it is a bare part, it can not only be easily perceived by a receiver, but it could also change colour over a period of time by changing melanin content of the skin. Thus, leg colour has the hallmarks for a possible ornament.

Even though the function of male ornamentation in House Sparrows seems to be rather well known, only very few studies have looked at the interplay of multiple ornaments in a single population (e.g. Bókony et al. 2006; Laucht et al. 2010), and none have looked over multiple years. In order to distinguish between the three main

hypotheses about multiple ornaments we evaluate the properties of our candidate ornaments in a large captive population over three years. We test for (1) inter-correlations between the various ornaments and general correlations with (2) condition, (3) testosterone dependence, (4) age and (5) place of origin (captive bred or wild born).

Because we use a large data set of many different individuals, many putative ornaments, the same individuals over several years, and a variety of quality parameters, this study has great power to separate the three hypotheses for the function of multiple ornaments from each other and to better understand complex signalling in general. Under the general framework developed by Møller and Pomiankowski (1993), we made the following predictions for each of the three hypotheses: First, under the multiple messages hypothesis each ornament should convey different information, and so we predict low correlations between different ornaments and high correlations with different quality measures, respectively. Second, under the back-up signals hypothesis all ornaments should signal the same information, and so we predict high correlations between different ornaments and correlations of all ornaments with similar quality measures. And third, under the unreliable signal hypothesis ornaments should signal no information, and so we expect low correlations between ornaments and no correlations with the measures of quality. In addition to quality parameters directly important during mate choice and male-male interactions such as body condition, age and testosterone dependence, we also examined effects of origin (i.e. captive or wild born) because origin potentially provides information on population differences and on early development effects. Because we kept birds of both origins for several years together under the same conditions we expected interactive effects of origin and age.

Material and Methods

Study population

We studied a population of 175 captive male House Sparrows. Birds were held at the Max Planck Institute for Ornithology, Seewiesen, Germany. They were either caught in rural areas in Bavaria, Germany (under license: permit nr. 55.1-8642.3-3-2006 of

the "Regierung Oberbayern", with several extensions) and held in captivity for at least eight months (n = 142) or raised in captivity (n = 33). Housing was in unisexual groups (except for breeding in spring) ranging between five and ten individuals (depending on other experiments) in semi-outdoor (one side only enclosed by chicken wire) aviaries of 1.2 x 2.0 x 4.0 m. At all times, the birds had *ad libitum* access to food, sand for dust bathing, and drinking and bathing water. The light-dark cycle and temperatures in the aviaries were close to natural conditions. Further details about our study population can be found in Laucht et al (2010). Sample size changes over the years reflect natural deaths of some individuals and other individuals hatching into captivity during the study period.

During several periods in the course of four years (July 06, Oct./Nov. 2006, Jan. 2007, March 2007, June 2007, Jan. 08, June 08, and Jan. 09), we caught all individual birds, took biometric measurements, standardized photographs of the ornaments, and -in four seasons of one year- blood samples.

Blood sampling and determination of plasma T levels

During four periods throughout the course of one year (Oct./Nov. 2006, Jan. 2007, March 2007, and June 2007) we took blood samples at similar times either in the mornings or in the afternoons. During each blood sampling event, we took 150-200 µl of blood from the wing vein within fifteen minutes after first starting to catch the birds. We collected the blood in 75 mm Na-heparinized micro haematocrit capillaries and centrifuged it at 13000 rpm for three minutes to separate the plasma. Plasma was stored at -80°C.

Frozen plasma samples were sent to the endocrine laboratory of the Leibniz Institute for Zoo and Wildlife Research in Berlin, Germany, where testosterone (T) levels were determined blindly by enzyme immunoassays (for details on the methods see Roelants et al. 2002; see also Laucht et al. 2010). The inter-assay CV for the enzyme immunoassay was 12.3% and the intra-assay CV was 9.0%. Additionally, to calculate the true repeatability (intra-class correlation coefficient) of measuring plasma T levels, we split samples of several individuals into duplicates right after centrifuging. Across the whole year, repeatability of these plasma T estimates was $R = 0.967 \pm 0.006$ (SE)

($p < 0.001$, $n = 2*122$). We assumed that all data points with a value of zero were actually below the detection limit and thus assigned them the lowest value measured (15 pg/ml). T levels are reported in pg per ml, and we natural log transformed T values for statistical analyses. T levels are the same as reported in Laucht et al. (2010).

Determination of bill colour

During each sampling season we took two standardized photographs of each male's bill. For this, we used a Canon Power Shot S2 IS camera and took pictures at the highest resolution with flash. All males were held the same way (presenting the right side of the head and bill to the camera) in front of a gray card and colour standard background at the same distance from the camera. Digital photograph processing software written into R 2.4.0 (R Development Core Team 2006) was used to determine bill "brightness" as measured in the hue, saturation, and brightness colour space. SL measured brightness of the individual pixels located at five randomly chosen positions each on the upper bill, lower bill and the gray background around the bill (used as a brightness standard). To standardize our bill brightness measurements between photos, we calculated overall mean gray card brightness of all photos of each season, we then determined the deviation of gray card brightness of a focal photo from the overall mean, and subtracted this deviation from mean bill brightness for each picture. This standardization renders a bill brightness score for each male that both compensates for any minor differences in overall brightness between photos and that also keeps our brightness variable as an actual colour measurement (rather than a difference). For analyses, we used the mean of the standardized upper and lower bill brightnesses from both pictures. Using data of the first year of sampling, these measurements were highly repeatable within individuals (repeatability (Lessells and Boag 1987): $R = 0.949 \pm 0.004$ (SE), $p < 0.001$, $n = 2*581$ for two pictures). See Laucht et al. (2010) for additional details.

Determination of badge size

We took four pictures of the males' breast bibs during each season. For each picture, we held the birds ventrally such that the throat and bib was stretched and presented to

the camera. Between each photograph of the bib, the bird was rearranged into a different position. SL measured the size of the badge from the photographs by encircling and measuring the area of the bib in pixels using the program ImageJ 1.36b (Abramoff et al. 2004) and later converting it into cm^2 using an area standard present in each photograph. For analyses, we used the average of all four pictures for each bird. The measurements were highly repeatable within individuals within the first year ($R = 0.943$, estimated according to Falconer and Mackay (1996) from repeatability of single pictures). See Laucht et al (2010) for additional details.

Determination of wingbar area and brightness

We took two standardized pictures of the males' white wingbars during each season. For each picture, we held the birds dorsally with the right wing opened and the wingbar presented to the camera. Between each photograph of the wingbar, the bird was rearranged into a different position. Two students measured the size and the brightness of the wingbar by encircling and measuring the area in number of pixels and brightness of the white wingbar using the program ImageJ 1.36b (Abramoff et al. 2004). For standardization of the area, we divided this by a standard area present in each photograph and measured in the same way, and then converted the result into cm^2. For analyses, we used the average of both pictures for each bird. The measurements were highly repeatable within individuals (repeatability (Lessells and Boag 1987): area: $R = 0.883 \pm 0.006$ (SE), brightness: $R = 0.960 \pm 0.002$ (SE), $p < 0.001$, $n = 2*1289$ for two pictures).

Determination of leg colour

Each season, we took two standardized photographs of each bird's right leg. As for bill colour, we used digital photograph processing software written into R 2.4.0 (R Development Core Team 2006) to determine leg "brightness" (see 'Determination of bill colour'). We used brightness as a colour measurement because leg colour varies from a pale horn colour to a brown, i.e. apparent variation in total light reflectance. Additionally, leg colour is very similar to non-breeding bill colour for which we used brightness measurements. SL measured five points each on the leg and on the gray background around the leg (measured as a brightness standard). We then

standardized each picture as described above for bill colour. For analyses, we used the mean of both pictures. These measurements were highly repeatable within individuals (repeatability (Lessells and Boag 1987): brightness $R = 0.769 \pm 0.009$ (SE), $p < 0.001$, $n = 2*2074$ for two pictures).

Statistical analyses

We performed all statistical analyses using R 2.8.0 (R Development Core Team 2008); packages: nlme, lme4, multcomp, RODBC) at the significance level $\alpha = 0.05$. For overall analyses, we used linear mixed effect models (lme, and lmer for crossed random effects) using individual ID as a random factor to account for repeated sampling. When using data from single seasons where each individual was represented only once, we used linear models (lm). For correlations between ornaments we used the Pearson correlation test.

For analyses on badge size, we used the averages of the seasonal scores of each year (starting in autumn after moult and ending in July before the next moult), because badge size changed due to abrasion of the white feather edges in autumn and early spring (Møller and Erritzøe 1992). This score is highly correlated with the score we had previously used (Laucht et al. 2010; $p < 0.001$, $t_{153} = 28.05$, $r = 0.91$) but easier to perform with the set of sampled periods over the four years. Similarly, for analyses on wingbar area and brightness, we also used "yearly" averages of all seasonal scores. We decided to also use "yearly" averages as final leg colour scores because between-seasonal variation in leg colour was not consistent (i.e. not similar in the same seasons or of a regular fluctuation).

To test for correlations between any two ornaments, between ornaments and age, as well as ornaments and body condition we used the four yearly measurements for badge size, leg colour and wingbar area and brightness. For correlations with bill colour we used seasonal scores as bill colour changes regularly in the course of the year (Laucht et al. 2010). This means that we used the same yearly scores repeatedly when analyzing relationships with seasonal scores such as bill colour or T levels, and only once when analyzing relationships with yearly scores such as age or tarsus length.

For analyses with body condition we used body mass, tarsus length, and the interaction of both. For analyses with T levels, we controlled for time of day by fitting this in each model. For analyses with age, we used minimum age for wild caught birds calculated as if all wild caught birds were born in the summer before capture and true age for captive born birds. The time period from the beginning of the annual moult (August) until before the start of the next annual moult represented one year for these age measurements. Thus, age 1 is the time from moult into adult plumage until the completion of the first year, age 2 the following year and so on.

Results

The results of the various tests conducted in this study are qualitatively summarized in Table 1.

Table 1: Summary of results of the four different ornaments in relation to the four addressed test: ornament inter-correlation, condition dependence, testosterone dependence, and age dependence.

ornament	correlation with other ornaments	condition dependence	testosterone dependence	age dependence
badge size	in one year with wingbar area	tarsus, mass in March and June	in June	yes
wingbar area	in one year with badge size	mass in Jan and March	in fall	yes
wingbar brightness	none	no	no	
bill brightness	summer of 2 years with leg brightness	no	all seasons but June	no
leg brightness	summer of 2 years with bill brightness	no	no	yes

Inter-correlations between ornaments

We found no consistent patterns of correlations for any combination of two ornaments (Table 2). The few significant correlations seem rather arbitrary, and after Bonferroni correction the only one still significant is a correlation between bill colour and leg

colour in one of the eight studied seasons (June 08). Overall, this suggests that all ornaments are at best very weakly correlated with each other.

Table 2: Correlations between any two of the four studied ornaments of male House Sparrows. We used yearly means of measurements for badge size, wingbar area and brightness, and leg colour. Because of the changes in bill colour in the course of the year, we used seasonal measurements for correlations with bill colour.

		year	R	P
badge size	wingbar area	05/06	**0.220**	**0.021**
		06/07	0.063	0.511
		07/08	-0.037	0.697
		08/09	0.053	0.584
badge size	wingbar brightness	05/06	0.162	0.092
		06/07	-0.139	0.147
		07/08	-0.060	0.532
		08/09	0.037	0.698
bill colour	leg colour	July 06	**0.224**	**0.019**
		autumn 06	-0.010	0.921
		Jan 07	-0.061	0.529
		March 07	-0.049	0.612
		June 07	0.023	0.808
		Jan 08	0.028	0.77
		June 08	**0.356**	**< 0.001**
		Jan 09	0.069	0.476
bill colour	wingbar area	July 06	-0.027	0.782
		autumn 06	-0.163	0.088
		Jan 07	-0.014	0.886
		March 07	0.054	0.576
		June 07	0.090	0.349
		Jan 08	-0.164	0.087
		June 08	0.166	0.084
		Jan 09	-0.128	0.184
bill colour	wingbar brightness	July 06	0.014	0.886
		autumn 06	-0.014	0.887
		Jan 07	-0.027	0.781
		March 07	-0.017	0.858
		June 07	0.105	0.273
		Jan 08	0.049	0.614
		June 08	0.043	0.657
		Jan 09	-0.052	0.592
badge size	leg colour	05/06	-0.012	0.902
		06/07	-0.038	0.696
		07/08	0.039	0.683
		08/09	0.019	0.846
wingbar area	leg colour	05/06	-0.046	0.633
		06/07	0.039	0.684
		07/08	0.140	0.146
		08/09	-0.042	0.661
wingbar brightness	leg colour	05/06	-0.106	0.270
		06/07	0.042	0.663
		07/08	0.040	0.675
		08/09	0.097	0.313

Table 3: Results of overall linear mixed effect models of ornaments (badge size, bill brightness, wingbar area, wingbar brightness, leg brightness) in relation to body mass, tarsus length, or the interaction of both in all years combined. We fitted bird ID and year or season (for bill colour and mass), respectively, as random effects. For details on single seasons see Table 4.

ornament	**mass**			**tarsus**			**interaction**		
	t	p	n	t	p	n	t	p	n
badge size	1.49	0.137	1171	**3.33**	**< 0.001**	551	-1.22	0.221	1104
bill color	-0.62	0.533	1171	-0.64	0.525	1107	0.92	0.358	1104
wingbar area	**5.09**	**< 0.001**	1171	**3.41**	**< 0.001**	551	-0.58	0.565	1104
wingbar brightness	**4.35**	**< 0.001**	1171	0.45	0.656	551	**-2.14**	**0.032**	1104
leg color	**3.11**	**0.002**	1171	-0.004	0.996	551	**-1.98**	**0.048**	1104

Table 4: Results of linear models of ornaments (badge size, bill brightness, wingbar area, wingbar brightness, leg brightness) in relation to body mass or tarsus length in the single seasons. Note that the interaction of mass and tarsus length was not significant for any of the ornaments or seasons and we therefore do not present the results. We present results of badge size, wingbar area, wingbar brightness and leg brightness in relation to tarsus length only once per year as they do not change over seasons.

season	badge size		bill brightness		wingbar area		wingbar brightness		leg brightness	
	mass	tarsus	mass	tarsus	mass	tarsus	mass	tarsus	mass	tarsus
July 06	t_{140} = 3.15 p = **0.002**	t_{137} = 2.88 p = **0.005**	t_{140} = -1.59 p = 0.115	t_{137} = -1.27 p = 0.208	t_{140} = 3.78 p **< 0.001**	t_{137} = 3.06 p = **0.003**	t_{140} = 1.98 p = **0.050**	t_{137} = 0.29 p = 0.77	t_{140} = 0.50 p = 0.619	t_{137} = 0.03 p = 0.978
Oct/Nov 06	t_{150} = 0.80 p = 0.428	t_{150} = 2.70 p = **0.008**	t_{150} = -0.44 p = 0.662	t_{150} = -0.07 p = 0.947	t_{150} = 1.49 p = 0.138	t_{150} = 1.93 p = 0.056	t_{150} = 0.25 p = 0.086	t_{150} = 0.20 p = 0.838	t_{150} = 0.12 p = 0.905	t_{150} = -0.07 p = 0.948
Jan 07	t_{137} = 1.09 p = 0.277		t_{137} = -0.92 p = 0.360	t_{137} = -0.73 p = 0.468	t_{137} = 2.49 p = **0.014**		t_{137} = 0.73 p = 0.464		t_{137} = -0.65 p = 0.515	
March 07	t_{137} = 2.19 p = **0.030**		t_{137} = -0.68 p = 0.495	t_{137} = 0.02 p = 0.986	t_{137} = 2.69 p = **0.008**		t_{137} = -0.03 p = 0.974		t_{137} = 0.19 p = 0.848	
June 07	t_{148} = 2.17 p = **0.031**		t_{148} = 0.52 p = 0.606	t_{145} = -0.48 p = 0.629	t_{148} = 2.60 p = **0.010**		t_{148} = 0.74 p = 0.463		t_{148} = 0.36 p = 0.722	
Jan 08	t_{158} = 1.33 p = 0.187	t_{138} = 1.84 p = 0.068	t_{158} = -1.66 p = 0.099	t_{138} = -0.26 p = 0.799	t_{158} = 4.64 p **< 0.001**	t_{138} = 2.37 p = **0.019**	t_{158} = 0.14 p = 0.685	t_{138} = 0.63 p = 0.532	t_{158} = 2.36 p = **0.019**	t_{138} = 0.86 p = 0.392
June 08	t_{149} = 1.71 p = 0.089		t_{149} = 2.43 p = **0.016**	t_{129} = -0.23 p = 0.819	t_{149} = 5.46 p **< 0.001**		t_{149} = 2.50 p = **0.014**		t_{149} = 2.44 p = **0.016**	
Jan 09	t_{136} = 3.43 p **< 0.001**	t_{118} = 3.66 p **< 0.001**	t_{136} = -1.53 p = 0.130	t_{118} = -0.23 p = 0.819	t_{136} = 3.12 p = **0.002**	t_{118} = 3.17 p = **0.002**	t_{136} = 3.89 p **< 0.001**	t_{118} = 0.35 p = 0.726	t_{136} = 1.78 p = 0.078	t_{118} = -0.99 p = 0.326

Condition dependence of ornaments

We first analyzed the different ornaments in relation to condition in overall models including all data of the four years (Table 3), and then for each season separately (see Table 4 for results of the single seasons). In the overall models, we found that badge size and wingbar area were condition dependent, whereas bill colour was not (Table 3). Leg colour also seemed condition dependent, however, because there were no correlations in single seasons (see Table 4) this result seems rather arbitrary.

The seasonal relationships (Table 4) show an influence of condition on badge size, wingbar area and to some extent on wingbar brightness. Bill and leg colour seem to not be related to condition since the significant relationships in one year are rather weak and do not hold when applying Bonferroni corrections.

T dependence of ornaments

In a previous study we found that bill colour and T levels were correlated in an overall model and during all different seasons, except in June ($r \leq -0.24$, $t \leq -2.88$, $p \leq 0.005$; see also Table 3 in Laucht et al. (2010). Badge size was not related to day time T levels ($p > 0.27$; Laucht et al. 2010), but in a different study we found it to correlate with night time T levels in peak breeding season ($t_{42} = 2.54$, $r = 0.36$, $p = 0.015$; see also Fig 3 in Laucht et al. (2011)). Leg colour was not related to testosterone levels at any of the four studied seasons ($p > 0.25$, $n = 136$ per season). But wingbar area was significantly and positively related to T levels in autumn ($p = 0.005$, $t_{133} = 2.85$; Fig 1), but not in any of the other studied periods ($p > 0.12$, $n = 136$ per season). However, wingbar brightness was not related to testosterone levels in any season ($p > 0.44$, $n = 136$ per season).

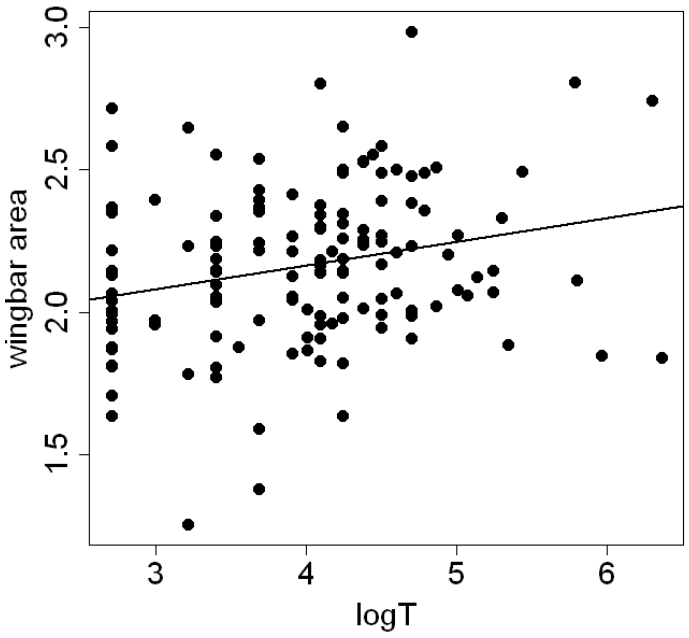


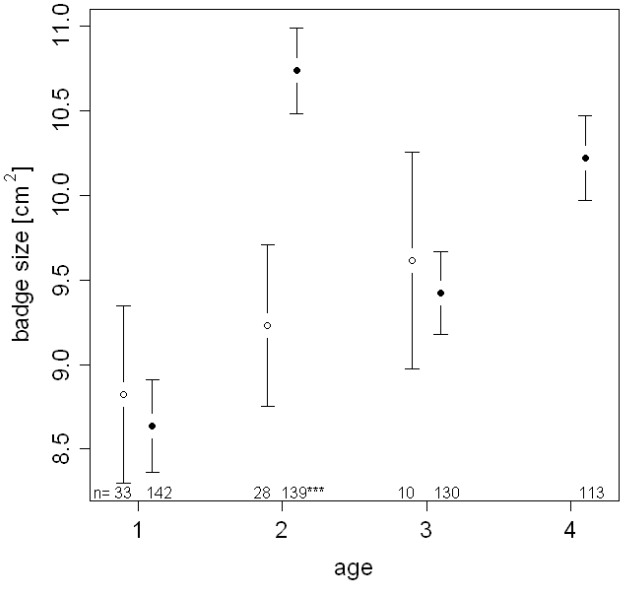
Fig. 1 Correlation between wingbar area and testosterone levels of male House Sparrows in autumn.
We defined wingbar area as the mean of the four studied seasons and used the natural logarithm of plasma T levels. $p = 0.005$, $t_{134} = 2.87$, $r = 0.24$

Age dependence of ornaments

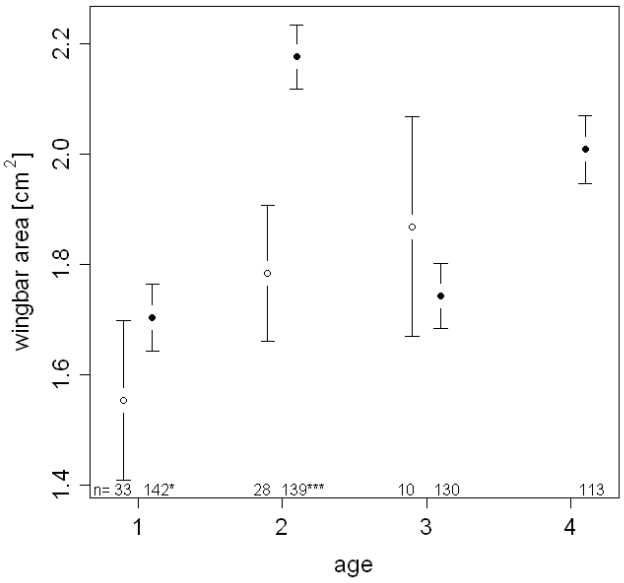
Leg colour, wingbar area and wingbar brightness, and badge size were all strongly correlated with age (Table 5, Fig 2 a-d). Bill colour was not related to age. We further examined the differences between individuals by modelling the relationship of each ornament to age for each individual (after accounting for season in bill colour analyses) and then looking at the coefficients of slope in relation to intercept. We found this to be highly significant for badge size, wingbar area, leg colour, and bill colour ($p < 0.001$, badge size: $t_{148} = -13.76$, wingbar area: $t_{150} = -15.88$, leg colour: $t_{150} = -16.85$, bill colour: $t_{168} = 5.50$). This means that birds with more elaborate ornaments at a younger age increased ornaments less strongly than birds with less elaborate ornaments at a younger age.

Table 5: Results of liner mixed effect models (lme) of ornaments in relation to age with bird ID as random effects. For bill colour we used also season as a random effect, and for the first model origin (lmer, hence no df). Sample sizes for badge size, wingbar conspicuousness, and leg colour were n = 595, for the analyses with bill colour n = 1174.

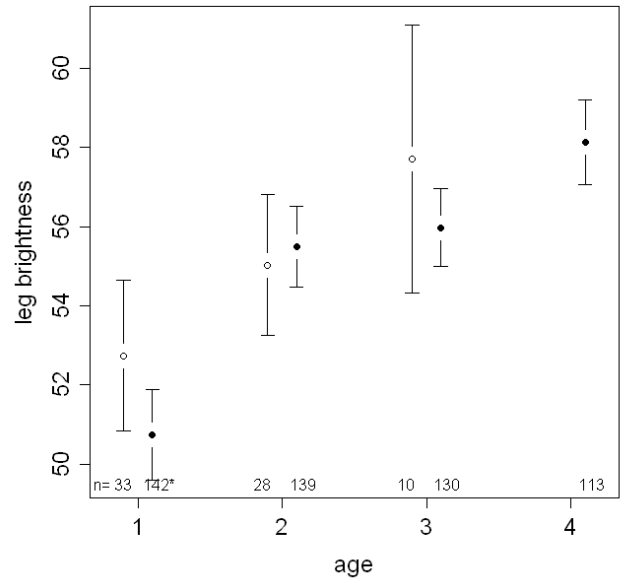
		origin as random effect, age: 1-4		age: 1-4		age: 1 or older	
ornament		**statistic**	**p**	**statistic**	**p**	**statistic**	**p**
badge size	age	t = 8.71	**< 0.001**	t_{415} = 1.99	**0.05**	t_{415} = 2.08	**0.04**
	origin			t_{176} = 1.00	0.32	t_{176} = -0.77	0.44
	age:origin			t_{415} = -0.08	0.93	t_{415} = 4.48	**< 0.001**
wingbar area	age	t = 4.93	**< 0.001**	t_{415} = 3.46	**< 0.001**	t_{415} = 3.93	**< 0.001**
	origin			t_{176} = 3.85	**< 0.001**	t_{176} = 2.37	**0.02**
	age:origin			t_{415} = -2.42	**0.02**	t_{415} = 0.38	0.70
wingbar brightness	age	t = -9.27	**< 0.001**	t_{415} = -2.37	**0.02**	t_{415} = -3.34	**< 0.001**
	origin			t_{176} = 1.60	0.11	t_{176} = 0.25	0.81
	age:origin			t_{415} = 0.23	0.81	t_{415} = 2.19	**0.03**
leg brightness	age	t = 15.91	**< 0.001**	t_{415} = 3.55	**< 0.001**	t_{415} = 3.27	**0.001**
	origin			t_{176} = -0.79	0.43	t_{176} = -2.02	**0.05**
	age:origin			t_{415} = -0.03	0.98	t_{415} = 2.86	**0.005**
bill brightness	age	t = -1.85	0.06	t = -0.89	0.38	t = 0.19	0.85
	origin			t = 1.20	0.23	t = -0.71	0.48
	age:origin			t = -1.01	0.31	t = 0.58	0.56



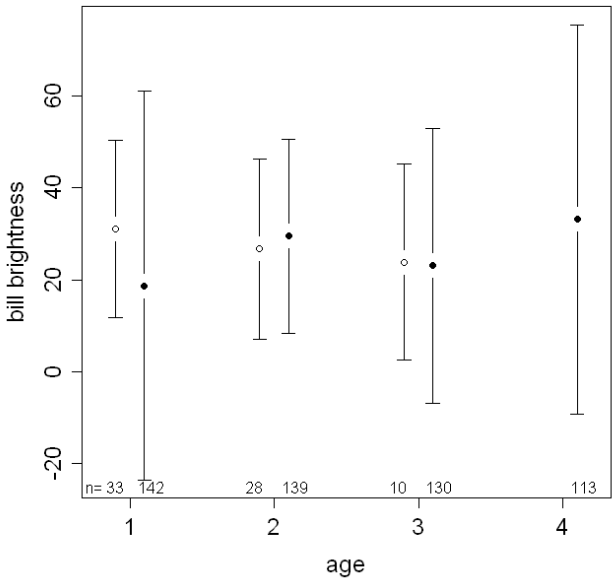
a)


wingbar area [cm^2]
1.4
1.6
1.8
2.0
2.2
n= 33
142*
28
139***
10
130
113
1
2
3
4
age

b)


leg brightness
50
52
54
56
58
60
n= 33
142*
28
139
10
130
113
1
2
3
4
age

c)



d)

Fig. 2 Correlations between badge size (a), wingbar area (b), leg colour (c), and bill colour (d) and age of male House Sparrows.
Open circles represent birds born in captivity, closed circles birds born in the wild. We present estimates and 95% confidence intervals of linear mixed effect models of ornament in relation to age accounting for bird ID (and season for bill colour). We also present sample sizes and significances for a t-test between captive and wild born birds for a given age and ornament (* $p < 0.05$, *** $p < 0.001$) except for bill colour.
We defined badge size, wingbar area and leg colour as the yearly means, and bill colour as seasonal values (as bill colour changes over the seasons). Age was defined as true age for captive born birds and as minimum age for wild born birds.

Interactive age and origin dependence of ornaments

We found moderate effects of origin (born in captivity or in the wild) on badge size, wingbar conspicuousness, and leg colour (Table 5, Fig 2 a-d). This was not the case in bill colour. The observed effect was mainly due to stronger fluctuations in ornament elaboration in wild born birds. This could have resulted from the fact that all wild born birds were the same age, and thus environmental influences were not levelled out.

Discussion

In order to test the three original hypotheses (multiple messages, back-up signal, and unreliable signal hypothesis (Møller and Pomiankowski 1993)) about the function of multiple ornaments in an extended (i.e. in a male-female and male-male) context we studied four different (putative) ornamental signals in male House Sparrows: badge size, wingbar area and brightness, bill colour, and leg colour. We found that none of the ornaments was consistently correlated with any of the others. In addition, we found that badge size and wingbar were related to body condition, but not bill and leg colour. Except for leg colour, all ornaments were testosterone dependent, but clearly differed in the timing of dependency (i.e. relations in autumn, breeding season, or all seasons but breeding). Additionally, all ornaments but bill colour were age dependent. Badge size, wingbar area, and leg colour were also weakly related to origin. Overall, this suggests that all four traits have the potential to function as quality signals and that each of them as a whole could signal a different, albeit overlapping, aspect of male quality.

Taken together, our results provide insight into the complexity of House Sparrow signalling because they demonstrate that male House Sparrows possess at least four ornamental traits that could function at providing information about phenotypic and age related quality. House Sparrows have a highly social lifestyle (foraging flocks, colony breeding, roosting congregations) where a multifaceted communication and thus signalling system could facilitate the complex dynamics of frequent interactions among individuals (Anderson 2006), i.e. to help avoid conflicts and to improve mate choice. Badge size has previously been suggested to signal social status, overall fighting success, age, variation in sexual behaviour, and (maximal) testosterone levels, and thus to play a role in male-male interactions as well as for female mate choice (Møller 1987; Møller 1990; Veiga 1993; Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001; Liker and Barta 2001; McGraw et al. 2003; Bókony et al. 2006; Nakagawa et al. 2007; Morrison et al. 2008; Laucht et al. 2011). Our results are in accordance with this. A different aspect of testosterone and age related behaviour and quality (this study) is conveyed by wingbar conspicuousness that has previously been suggested to signal defence success, body condition, and parasite resistance (Poston et al. 2005; Bókony et al. 2006; Moreno-Rueda 2010). In contrast, bill colour

potentially does not signal quality, but testosterone related behavioural strategies and readiness to breed (Laucht et al. 2010). In addition to age signalling integrated into these ornaments, leg colour might be a true signal of age, although more research is needed. These findings demonstrate support for the hypothesis that multiple ornaments can play a role as armaments and ornaments (Berglund et al. 1996) and be addressed to multiple receivers (i.e. mates and rivals) (Andersson et al. 2002). However, the actual function as a signal and the addressing of one or multiple receivers still remains to be tested.

Three of the four studied ornaments were relatively strongly related to age. Age has often been argued to influence aspects of quality and/or status (i.e. older males are in better condition and attain higher status) (Ketterson 1979; Cucco and Malacarne 1999; Komdeur et al. 2005; e.g. Delhey and Kempenaers 2006; Nakagawa et al. 2007; Nakagawa and Burke 2008; Botero et al. 2009; Galván and Møller 2009). This could apply for badge size and wingbar area that both most likely signal age integrated into quality or status. However, House Sparrows seem to also signal true age via leg colour and another ornament –the black eye mask (Nakagawa and Burke 2008). This suggests that it could be potentially beneficial for an individual male to signal true age in addition to the interaction of age and condition or status (Nakagawa and Burke 2008). This could be especially important in the context of mate choice when females choose mates of certain age classes. The benefits and disadvantages of mating with an older (as opposed to very old) partner have been discussed widely (Manning 1985; Kokko 1997; Kokko 1998; Brooks and Kemp 2001). In general, when trading off advantages and disadvantages (for a detailed list see Brooks and Kemp 2001) it seems likely that benefits of choosing older mates are context dependent (such as stability of the environment, spread of diseases, etc.). If this is true, a female could disentangle different age effects and decide context dependently for an older or a younger mate if male ornaments did not only signal quality but also true age. However, so far, in House Sparrows it has been only found that reproductive success increased with age (Hatch and Westneat 2007), but context dependent choice of older males has not been examined yet.

For three ornaments, badge size wingbar area and leg colour, we found a weak effect of origin (i.e. captive or wild born) at different age classes. This was mainly due to

stronger fluctuations in ornament elaboration in wild born birds. One reason for this could have been that all wild born birds were caught at the same time and thus were in the same age class throughout the study. In contrast, captive born birds were born in different years and thus not in the same age classes throughout the study. Therefore, effects other than age were more levelled out in captive born birds and could have thus led to fluctuations in wild born birds. Though housing conditions were the same over all years, other environmental conditions such as temperature, sun light, or humidity could have changed. In addition, the social environment for single individuals was not the same each year. Therefore, environmental factors could have had differently strong influences on ornamentation in different years and led to fluctuations. Taken together, this suggests that recent condition is more important for ornamentation rather than early developmental condition.

Ignoring the observed age-relationships for a moment, our results generally support the multiple messages hypothesis (Møller and Pomiankowski 1993; Candolin 2003) because the four ornaments did not correlate with each other and they co-varied differentially with our condition parameters and testosterone measures. For support of the back-up signal hypothesis we would have expected inter-correlations between ornaments and relationships with similar quality parameters, while for the unreliable signal hypothesis we would have expected no inter-correlations and no relationships with the condition measures (Møller and Pomiankowski 1993; Candolin 2003). This conclusion is in accordance with previous studies on House Sparrow ornamentation that have suggested that badge size and wingbar conspicuousness as well as badge size and bill colour were multiple messages (Bókony et al. 2006; Laucht et al. 2010). However when we take the age relationship into consideration it is more challenging to unequivocally categorize multiple ornaments in House Sparrows. This is because several ornaments were related to age in a similar way suggesting that they could function as back-up signals of age and age related qualities. Therefore, the strict separation of multiple ornaments into one of these three alternative functions seems contrived.

Instead, we suggest that all ornaments seem to signal different but overlapping aspects of quality. Similarly, a study on Great Tit (*Parus major*) song found that song as a multidimensional signal was signalling multiple and back-up information (Rivera-

Gutierrez et al. 2010) suggesting that the phenomenon observed in the House Sparrow might be more general. This means that quality signals, and potentially others as well, rather lie on a spectrum revealing varying degrees of varying types of information instead of in one of three categories (along the lines of the three hypotheses). Therefore, multiple ornaments are better described as a continuum between different classes or in context dependent classifications.

In summary, we found badge size, wingbar area, bill and leg colour of male House Sparrows to have potential for functioning as both multiple messages of various qualities and back-up messages of age. Overall, our results provide strong evidence for the signalling importance of multiple ornaments in birds (Calkins and Burley 2003; Doucet and Montgomerie 2003; Jawor et al. 2003; Jawor et al. 2004; Hegyi et al. 2007; Karubian 2008; Reudink et al. 2009). However our powerful analysis provides a significant challenge to the notion that multiple ornaments can be empirically classified into simple categories such as multiple, back-up, and unreliable signals.

Acknowledgements

Many thanks to Ariane Mutzel and Andrea Wittenzellner for their help with taking blood samples. We are grateful to Pernilla Borgström and Maryjane Heckel for scoring wingbar area and brightness, to Mihai Valcu and Holger Schielzeth for statistical advice and programming, and to Magnus Rudorfer, Frances Weigel, Edith Bodendorfer, and Annemarie Grötsch for animal care. We thank the Behavioural Ecology and Evolutionary Genetics Group at the MPIO for discussion and comments. Funding for this research was provided by the Max Planck Society. The study was performed in agreement with the current laws of Upper Bavaria, Germany (under license: permit nr. 55.1-8642.3-3-2006 of the “Regierung Oberbayern”, with several extensions).

References

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophoton. Int. 11:36-42

Anderson TR (2006) Biology of the ubiquitious house sparrow. Oxford University Press, New York

Andersson S, Pryke SR, Ornborg J, Lawes MJ, Andersson M (2002) Multiple receivers, multiple ornaments, and a trade-off between agonistic and epigamic signaling in a widowbird. Am. Nat. 160:683-691

Balph MH, Balph DF, Romesburg HC (1979) Social Status Signaling in Winter Flocking Birds: an Examination of a Current Hypothesis. Auk 96:78-93

Berglund A, Bisazza A, Pilastro A (1996) Armaments and ornaments: An evolutionary explanation of traits of dual utility. Biol. J. Linn. Soc. 58:385-399

Bókony V, Lendvai AZ, Liker A (2006) Multiple cues in status signalling: The role of wingbars in aggressive interactions of male house sparrows. Ethology 112:947-954

Botero CA, Rossman RJ, Caro LM, Stenzler LM, Lovette IJ, de Kort SR, Vehrencamp SL (2009) Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. Anim. Behav. 77:701-706

Brooks R, Kemp DJ (2001) Can older males deliver the good genes? Trends Ecol. Evol. 16:308-313

Buchanan KL, Evans MR, Goldsmith AR, Bryant DM, Rowe LV (2001) Testosterone influences basal metabolic rate in male house sparrows: a new cost of dominance signalling? Proc. R. Soc. Biol. Sci. Ser. B 268:1337-1344

Calkins JD, Burley NT (2003) Mate choice for multiple ornaments in the California quail, *Callipepla californica*. Anim. Behav. 65:69-81

Candolin U (2003) The use of multiple cues in mate choice. Biol. Rev. (Camb.) 78:575-595

Chaine AS, Tjernell KA, Shizuka D, Lyon BE (2011) Sparrows use multiple status signals in winter social flocks. Anim. Behav. 81:447-453

Cucco M, Malacarne G (1999) Is the song of Black Redstart males an honest signal of status? Condor 101:689-694

Dale J (2006) Intraspecific variation in coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration Harvard University Press, MA, Cambridge, pp 36-86

Delhey K, Kempenaers B (2006) Age differences in blue tit *Parus caeruleus* plumage colour: within-individual changes or colour-biased survival? J. Avian Biol. 37:339-348

Donham RS, Wingfield JC, Mattocks PW, Farner DS (1982) Changes in Testicular and Plasma Androgens with Photoperiodically Induced Increase in Plasma LH in the House Sparrow. Gen. Comp. Endocrinol. 48:342-347

Doucet SM, Montgomerie R (2003) Multiple sexual ornaments in satin bowerbirds: ultraviolet plumage and bowers signal different aspects of male quality. Behav. Ecol. 14:503-509

Evans MR, Goldsmith AR, Norris SRA (2000) The effects of testosterone on antibody production and plumage coloration in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 47:156-163

Freeman-Gallant CR, Taff CC, Morin DF, Dunn PO, Whittingham LA, Tsang SM (2010) Sexual Selection, Multiple Male Ornaments, and Age- and Condition-Dependent Signaling in the Common Yellowthroat. Evolution 64:1007-1017

Galván I (2010) Plumage coloration can be perceived as a multiple condition-dependent signal by Great Tits *Parus major*. Ibis 152:359-367

Galván I, Møller AP (2009) Different roles of natural and sexual selection on senescence of plumage colour in the barn swallow. Funct. Ecol. 23:302-309

Gonzalez G, Sorci G, Smith LC, de Lope F (2001) Testosterone and sexual signalling in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 50:557-562

Haase E (1975) Effects of Testosterone Propionate on Secondary Sexual Characters and Testes of House Sparrows, *Passer domesticus*. Gen. Comp. Endocrinol. 26:248-252

Hatch MI, Westneat DF (2007) Age-related patterns of reproductive success in house sparrows *Passer domesticus*. J. Avian Biol. 38:603-611

Hau M (2007) Regulation of male traits by testosterone: implications for the evolution of vertebrate life histories. BioEssays 29:133-144

Hegyi G, Szigeti B, Török J, Eens M (2007) Melanin, carotenoid and structural plumage ornaments: information content and role in great tits *Parus major*. J. Avian Biol. 38:698-708

Holland B, Rice WR (1998) Perspective: Chase-away sexual selection: Antagonistic seduction versus resistance. Evolution 52:1-7

Iwasa Y, Pomiankowski A (1994) The Evolution of Mate Preferences for Multiple Sexual Ornaments. Evolution 48:853-867

Jawor JM, Gray N, Beall SM, Breitwisch R (2004) Multiple ornaments correlate with aspects of condition and behaviour in female northern cardinals, Cardinalis cardinalis. Anim. Behav. 67:875-882

Jawor JM, Linville SU, Beall SM, Breitwisch R (2003) Assortative mating by multiple ornaments in northern cardinals (Cardinalis cardinalis). Behav. Ecol. 14:515-520

Johnstone RA (1995) Honest Advertisement of Multiple Qualities Using Multiple Signals. J. Theor. Biol. 177:87-94

Johnstone RA (1996) Multiple displays in animal communication: 'Backup signals' and 'multiple messages'. Philos. Trans. R. Soc. B-Biol. Sci. 351:329-338

Karubian J (2008) Changes in breeding status are associated with rapid bill darkening in male red-backed fairy-wrens *Malurus melanocephalus*. J. Avian Biol. 39:81-86

Keck WN (1933) Control of the bill color of the male English sparrow by injection of male hormone. Proc. Soc. Exp. Biol. Med. 30:1140-1141

Kekäläinen J, Valkama H, Huuskonen H, Taskinen J (2010) Multiple Sexual Ornamentation Signals Male Quality and Predicts Female Preference in Minnows. Ethology 116:895-903

Ketterson ED (1979) Status Signaling in Dark-Eyed Juncos. Auk 96:94-99

Ketterson ED, Nolan V (1992) Hormones and Life Histories - an Integrative Approach. Am. Nat. 140:S33-S62

Kokko H (1997) Evolutionarily stable strategies of age-dependent sexual advertisement. Behav. Ecol. Sociobiol. 41:99-107

Kokko H (1998) Good genes, old age and life-history trade-offs. Evol. Ecol. 12:739-750

Komdeur J, Oorebeek M, van Overveld T, Cuthill IC (2005) Mutual ornamentation, age, and reproductive performance in the European starling. Behav. Ecol. 16:805-817

Laucht S, Dale J, Mutzel A, Kempenaers B (2011) Individual variation in plasma testosterone levels and its relation to badge size in House Sparrows *Passer domesticus*: It's a night-and-day difference. Gen. Comp. Endocrinol. 170:501-508

Laucht S, Kempenaers B, Dale J (2010) Bill color, not badge size, indicates testosterone-related information in house sparrows. Behav. Ecol. Sociobiol. 64:1461-1471

Liker A, Barta Z (2001) Male badge size predicts dominance against females in House Sparrows. Condor 103:151-157

Lozano GA (2009) Multiple cues in mate selection: The sexual interference hypothesis. Bioscience Hypotheses 2:37-42

Manning JT (1985) Choosy Females and Correlates of Male Age. J. Theor. Biol. 116:349-354

McGlothlin JW, Jawor JM, Greives TJ, Casto JM, Phillips JL, Ketterson ED (2008) Hormones and honest signals: males with larger ornaments elevate testosterone more when challenged. J. Evol. Biol. 21:39-48

McGraw KJ, Dale J, Mackillop EA (2003) Social environment during molt and the expression of melanin-based plumage pigmentation in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 53:116-122

Møller AP (1987) Variation in Badge Size in Male House Sparrows *Passer domesticus* - Evidence for Status Signaling. Anim. Behav. 35:1637-1644

Møller AP (1990) Sexual-Behavior Is Related to Badge Size in the House Sparrow *Passer domesticus*. Behav. Ecol. Sociobiol. 27:23-29

Møller AP, Erritzøe J (1992) Acquisition of Breeding Coloration Depends on Badge Size in Male House Sparrows *Passer domesticus*. Behav. Ecol. Sociobiol. 31:271-277

Møller AP, Pomiankowski A (1993) Why Have Birds Got Multiple Sexual Ornaments. Behav. Ecol. Sociobiol. 32:167-176

Moreno-Rueda G (2010) Uropygial gland size correlates with feather holes, body condition and wingbar size in the house sparrow *Passer domesticus*. J. Avian Biol. 41:229-236

Morrison EB, Kinnard TB, Stewart IRK, Poston JP, Hatch MI, Westneat DF (2008) The links between plumage variation and nest site occupancy in male House Sparrows. Condor 110:345-353

Nakagawa S, Burke T (2008) The mask of seniority? A neglected age indicator in house sparrows *Passer domesticus*. J. Avian Biol. 39:222-225

Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T (2007) Assessing the function of house sparrows' bib size using a flexible meta-analysis method. Behav. Ecol. 18:831-840

Pfeiffer CA, Hooker CW, Kirschbaum A (1944) Deposition of pigment in the sparrow's bill in response to direct applications as a specific and quantitative test for androgen. Endocrinology 34:389-399

Poston JP, Hasselquist D, Stewart IRK, Westneat DF (2005) Dietary amino acids influence plumage traits and immune responses of male house sparrows, *Passer domesticus*, but not as expected. Anim. Behav. 70:1171-1181

R Development Core Team (2006) R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria

R Development Core Team (2008) R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria

Reudink MW, Marra PP, Boag PT, Ratcliffe LM (2009) Plumage coloration predicts paternity and polygyny in the American redstart. Anim. Behav. 77:495-501

Rivera-Gutierrez HF, Pinxten R, Eens M (2010) Multiple signals for multiple messages: great tit, *Parus major*, song signals age and survival. Anim. Behav. 80:451-459

Roberts ML, Buchanan KL, Evans MR (2004) Testing the immunocompetence handicap hypothesis: a review of the evidence. Anim. Behav. 68:227-239

Roelants H, Schneider F, Goritz F, Streich J, Blottner S (2002) Seasonal changes of spermatogonial proliferation in roe deer, demonstrated by flow cytometric analysis of c-kit receptor, in relation to follicle-stimulating hormone, luteinizing hormone, and testosterone. Biol. Reprod. 66:305-312

Schulte-Hostedde AI, Zinner B, Millar JS, Hickling GJ (2005) Restitution of mass-size residuals: Validating body condition indices. Ecology 86:155-163

Veiga JP (1993) Badge Size, Phenotypic Quality, and Reproductive Success in the House Sparrow - a Study on Honest Advertisement. Evolution 47:1161-1170

Witschi E (1936) The bill of the sparrow as an indicator for the male sex hormone. I. sensitivity. Proc. Soc. Exp. Biol. Med. 33:484-486

Witschi E, Woods RP (1936) The bill of the sparrow as an indicator for the male sex hormone. II. structural basis. J. Exp. Zool. 73:445-459

Non-submitted Manuscript

Chapter Four

The effects of intra-flock variability in badge size on ornament development in male House Sparrows, *Passer domesticus*

Silke Laucht, Bart Kempenaers, James Dale

Abstract

To better comprehend the signalling content of ornamental signals of quality, it is crucial to understand the underlying mechanisms that influence ornament development. Environmental conditions are especially known to affect quality signals, because the environment is strongly expected to impact overall phenotypic condition. However, among many different environmental factors that affect ornaments, the social environment (i.e. the composition of individuals a focal animal interacts with) has been largely neglected in research on ornamental signals of quality. The social environment is especially expected to play a strong role on the development of status-related signals, because status is a *relative* phenotype and thus highly dependent on the composition of an individual's competition. Here, we investigated the role of the social environment on the development of badge size in moulting male House Sparrows (*Passer domesticus*). House Sparrow badges are testosterone dependent ornaments that are typically argued to be "badges of status" – all-purpose signals of dominance and aggression. We experimentally manipulated the composition of individuals in groups by creating groups with similarly sized badges (i.e. presumably high social instability) or variable badges (i.e. presumably low social instability). We tested the effect of this treatment on the individuals' new badges, and we predicted in groups with similarly sized badges both (1) larger changes (on average) between pre- and post-moult badge and (2) increases in between-individual variation. We found support for prediction 1 in one year of three years and no support for prediction 2. We provide possible reasons why we found inconsistent patterns between years, and we conclude that the social environment can have an influence on ornament development, albeit a complex one. More research on this particularly poorly known area of honest signalling is badly needed.

Introduction

In the broad sense, ornamental signals of quality can potentially provide information about various components of overall genetic and phenotypic constitution (such as good genes, physical condition, freedom from disease, parental care abilities and social status) (Dale 2006). However, to better resolve a signal's specific information content it is critical to understand the various factors that influence signal development. All phenotypic development is determined by genetic and environmental parameters to varying degrees, and in the case of signals of quality, environmental conditions are especially well known to largely affect their expression (Hill 2006). This is presumably because "quality" or overall constitution is generally expected to be largely environmentally dependent. Using plumage colour of birds as a model, Hill (2006) described four broad classes of environmental factors that can influence quality signals: (1) pigment access, (2) nutritional condition, (3) parasites, and (4) social environment. While the former three have been largely studied in a variety of different bird species, to date very little is known about the influence of the social environment on ornament elaboration (Hill 2006).

Hill (2006) further described three major pathways how the social environment can influence ornamentation: (1) competition for food and thus access to important nutrients, (2) effects of parasites exchanged between group members, and (3) effects of agonistic interactions on hormone levels. So far, most research has focused on the effects of the agonistic behaviour on hormone levels of group members. There is a complex interplay between hormones and aggression, where hormone levels are known to affect aggression, while the outcome of aggressive interactions in turn can strongly affect hormone levels. For example, under the "Challenge Hypothesis" (Wingfield et al. 1990) circulating testosterone levels are activationally increased above baseline levels during agonistic interactions. In addition, the outcome of earlier contests can influence hormone levels which then influence the outcome of later interactions, so-called "winner-loser" effects (e.g. Oyegbile and Marler 2005). In unstable social groups, increased aggression might occur regularly and thus should consequently affect testosterone related ornaments during their development (McGraw et al. 2003). In agreement with this, it was found in male Mandrills (*Mandrillus sphinx*) and Red Junglefowl (*Gallus gallus*) that ornaments were related

to social rank within groups and circulating androgen levels at the same time (Parker et al. 2002; Setchell et al. 2008). Additionally, a study on House Sparrows (*Passer domesticus*) suggested that the size of the black breast patch – a male status signal – increased more in beta males that experienced higher aggression during moult (McGraw et al. 2003).

The above empirical examples involve aggression related "badges of status" traits which are strongly testosterone dependent and generally thought to indicate aspects of fighting ability, willingness to take risks and age related status (reviewed in Jawor and Breitwisch 2003; Tibbetts and Dale 2004). In general, such types of status related signals should make good candidates for indicators that are strongly influenced by conditions of the social environment because status is basically a *relative* characteristic of an individual. That is, social status is determined not just by absolute phenotypic characteristics such as health and physical condition, but also by relative traits such as how good you fair against the current competition. For example, a male with an average ornament competing in a group comprised of below average individuals is of *relatively higher quality* than he would be in a group comprised of above-average individuals. As such, one could in theory manipulate the relative quality of individuals simply by re-organizing the ornamental phenotypic variability of the population. To date, no study has manipulated the actual spectrum of ornamentation in group members that individuals interact with, and then evaluated how this affected subsequent ornament development.

In this study, we experimentally tested the influence of the social environment (as determined by group ornament distribution) on badge size in House Sparrows (*Passer domesticus*). House Sparrows are well suited for testing the effects of social parameters on phenotypes because of their highly social lifestyle with foraging flocks, breeding colonies, and roosting congregations (Anderson 2006). In addition, the size of the black breast bib or badge has been widely studied and is commonly accepted as an all-purpose badge of status (Møller 1987a; Møller 1987b; Veiga 1995; Hein et al. 2003). As mentioned earlier, McGraw et al. (2003) have demonstrated that the amount of aggression within groups during the annual moult had a positive effect on the size of the moulted badge. Additionally, across various studies badge size has been found to correlate with overall fighting success, age, and testosterone levels

(Møller 1987b; Møller 1990; Veiga 1993; Evans et al. 2000; Buchanan et al. 2001; Gonzalez et al. 2001; Liker and Barta 2001; McGraw et al. 2003; Bókony et al. 2006; Nakagawa et al. 2007; Morrison et al. 2008; Laucht et al. 2011).

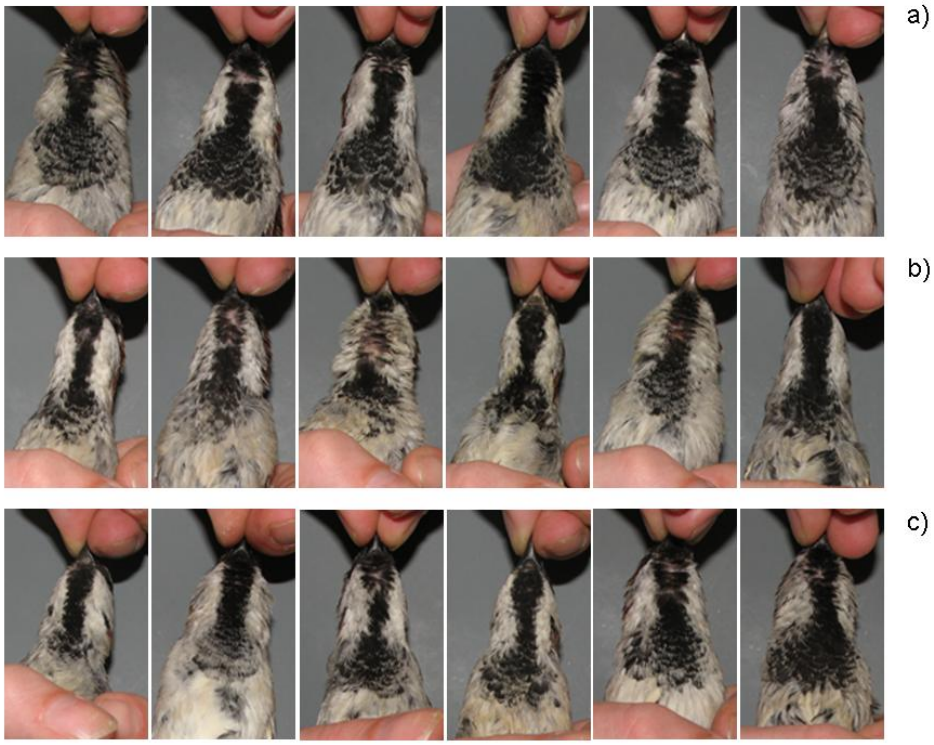


Fig. 1 Examples of the two treatment groups: a) uniform badges group, large badges, b) uniform badges group, small badges, and c) variable badges group

Over three annual moults, we monitored a large captive population of male House Sparrows (≥117 individuals per year), kept in flocks of 6 to 10 individuals in large semi-outdoor aviaries. In each flock, we manipulated group composition of badge sizes just prior to each annual moult and then examined its effects on post-moult badge size. We created two different treatment groups: birds of the first treatment group shared an aviary with only individuals with similarly sized badges (see Fig 1 a and b; hereafter "uniform" badges). Birds of the second group shared an aviary with individuals with badge sizes in a continuum from large to small (see Fig 1 c, hereafter "variable" badges). We reasoned that in uniform groups, individuals would be competing with relatively similar rivals with respect to badge related qualities (Maynard Smith and Harper 1988; Jawor and Breitwisch 2003; Tibbetts and Dale 2004; Tibbetts and Safran 2009). In contrast, individuals in variable groups would be

competing with relatively dissimilar rivals with respect to badge related qualities. Thus, a small-badged male in a uniform group, for example, would be *more likely* to increase in relative quality than a small-badged male in a variable group (who would be competing mostly with larger-badged individuals). Therefore, from the perspective of any individual in the experiment, relative (badge related) quality is expected *on average* to change more for individuals in uniform groups.

Based on the reasoning outlined above we made two a priori predictions about the effect of our treatment on the new badges males moulted into. First, we predicted that birds in the uniform groups would have on average stronger changes in their badges, while birds in the variable groups should not change badge size as dramatically (see above). Second, we predicted as a consequence of this, an increase in between-individual differences in badge size in uniform groups. To test these predictions we need to control for regression toward the mean, where we generally expect larger changes, in opposite directions, in birds at either ends of the continuum. Thus, irrespective of treatment, individuals with a small pre-moult badge will tend to increase badge size while individuals with a large pre-moult badge will tend to decrease badge size. Thus, in our tests, we control for pre-moult conditions, and examine the slopes (i.e. interactions between treatment and pre-moult conditions).

Using a large captive population, three different years, a suitable signal, and two treatment groups we created a powerful setup to test the effects of group composition (i.e. social environment) on ornament development. Note that in this experiment we do not monitor individual-level changes in dominance, behaviour or hormone levels. Although such data would be extremely insightful, for the scale involved in this experiment we did not have the resources required to measure such variables accurately during the short time frame available of annual moult. Moreover, accurate assessment of hormone levels are difficult at this time of year because they are particularly low (Laucht et al. 2010), point-samples of hormones offer very limited measures of baseline T-levels (Laucht et al. 2010), and we wanted to leave the birds relatively undisturbed and free to interact in their social groups. Thus, instead of getting detailed data on each individual, we elected to pursue a large-scale manipulation powerful enough to detect broad scale population-level changes based

on two logically derived predictions based on current views of status related ornaments.

Material and Methods

Study population

We studied a population of 169 captive male House Sparrows held at the Max Planck Institute for Ornithology, Seewiesen, Germany. At the time of study, all birds were after hatching year. They were either caught in rural areas in Bavaria, Germany (under license: permit nr. 55.1-8642.3-3-2006 of the "Regierung Oberbayern", with several extensions) and held in captivity for at least 8 months (n = 139) or raised in captivity (n = 30). We kept birds in groups of varying sizes (see below) in aviaries of size 1.2 x 2.0 x 4.0 m equipped with perches, nest boxes and big branches. At all times, the birds had *ad libitum* access to food, drinking and bathing water, and sand for dust-bathing. The light-dark cycle and temperatures in the aviaries were close to natural conditions, as the aviaries were semi-outdoor with one side enclosed only by chicken wire. The sides towards neighbouring aviaries were covered with hessian to prevent interactions with non-group members. Further details about our study population can be found in Laucht et al (2010).

Experimental procedures

We performed each experiment during annual moult in three different years (2006, 2007, and 2009) with roughly the same birds in each year (some birds were born later and some birds died of natural causes before the end of the study). Prior to each experiment we scored male badge sizes with digital photography (see below) and used these scores to assign individuals to treatments and groups. In groups with *uniform* badges individuals had relatively similarly sized badges; in groups with *variable* badges individuals had a diverse range of badges from small to large (e.g. see Fig. 1). To assign males to groups, male badge scores were ranked, and then individuals were assigned randomly to either the uniform or variable treatment, such that the mean and variance of badge sizes were similar across each treatment. Males were then randomly assigned into new social groups (i.e. aviaries), and we made sure new

groups were comprised of males that had not been housed together in the same aviary previously. Group sizes in each aviary were six in 2006, ten in 2007, and seven to eight in 2009. In 2006 and 2009, birds were moved into new groups just at the start of their annual moult (early August). In 2007 birds were arranged into new groups about three weeks before their annual moult (mid-July). In all years, birds were kept in their groups until long after the completion of moult at which point we scored the males' new badge sizes.

Determination of badge size

In each year we took four pictures of the males' black bibs prior to the experiment and then again after the completion of moult. For each picture, we held the birds ventrally such that the throat and bib were stretched and presented to the camera (as in Fig 1). Between each photograph of the bib, the bird was rearranged into a different position. SL measured the size of the badge from the photographs by encircling and measuring the area of the bib in pixels using the program ImageJ 1.36b (Abramoff et al. 2004) and later converting it into cm^2 using an area standard present in each photograph. For analyses we used the mean of the scores of each of the four photographs. Using photographs taken at four different seasons in the course of one year, these measurements were highly repeatable within individuals ($R = 0.943$, estimated according to Falconer and Mackay (1996) from repeatability of single pictures). See Laucht et al (2010) for additional details.

Statistical analyses

We performed all statistical analyses using R 2.11.0 (R Development Core Team 2010; packages: nlme, RODBC) at the significance level $\alpha = 0.05$. Generally, for overall analyses, we used linear mixed effect models (lme) including year as a fixed effect and individual ID as a random factor to account for repeated sampling. For analyses on single years (each individual was represented only once) we used linear models (lm). When controlling for body condition, we fitted tarsus length and body mass in the model, because the use of residuals has been criticized (e.g. Green 2001; Freckleton 2002). .

First we tested at the population level whether there was a general effect of treatment on badge size. A priori, we did not expect any effect because individuals in both treatments are expected to express similar changes in badge size *on average* (it is the slopes that should differ – see below). Thus for this analysis, we tested for a relationship between post-moult badge size and treatment after accounting for pre-moult badge size and body condition in each year separately and in all years combined (individual ID included as random factor). Using absolute or relative change in badge size instead of final badge size accounted for pre-moult badge size did not qualitatively change the results. Including social group ID (i.e. aviary ID) as a random factor also did not qualitatively change the results.

We tested our first prediction using linear mixed effect models of post-moult badge size in relation to the interaction of pre-moult badge size and treatment after accounting for body condition and with social group ID as random effect, both overall and in the three years separately. This is the best way of testing the first prediction because we expect a high correlation between post-moult and pre-moult badge size, but with different slopes in the two treatment groups. More specifically, in the uniform groups we expect a *more-shallow* slope of the post-moult badge versus pre-moult badge regression. This shallower slope should manifest from the stronger average changes in badge size in this treatment. Note that for graphical presentation we used centred (i.e. yearly means subtracted from individual badges) and scaled (divided by the yearly standard deviation) post- and pre-moult badges to exclude confounding year effects.

We tested our second prediction by examining post-moult group standard deviation (SD) of badge sizes in relation to treatment. We *a priori* expected higher intra-flock variance of badge sizes in the uniform treatment groups after moult (i.e. an increase), but not in the variable treatment groups. Again due to regression toward the mean, a key predictor of change in SD is expected to be pre-moult cage SD. In this particular test, we expected an effect of treatment, but no significant interaction between treatment x pre-moult SD (i.e. we expected the model line fits to be parallel for each treatment, but higher (a greater post-moult SD) in the uniform groups). We only performed this analysis on all years pooled together since, because group values are used, our power is too low to look for meaningful trends within-years. Because of

this, each individual bird is included in the analysis up to three times. However because group composition changed consistently from year to year and because we use group values as data points (not individual values) there is no pesudoreplication (i.e. each group is unique).

Results

First, we analyzed changes in badge size on the individual level for the whole population. Badge size did not change differently in the two experimental groups: neither overall (lme: $t_{214} = -1.55$, $p = 0.123$ after accounting for year and condition, random effect: bird ID) nor in single years (lm: $|t| \leq 1.66$, $p \geq 0.099$). This means that there were no overall population effects.

To test prediction 1 we tested for an effect of the interaction of pre-moult badge size and treatment on post-moult badge size. In all years pooled together there was no significant effect (Table 1), however in 2007 (but not in 2006 and 2009) there was a significant effect of treatment with the predicted shallower slope in the uniform groups (Fig. 2, Table 1).

Third, as expected the change in intra-flock badge SD was strongly negatively related to pre-moult badge SD (lm: $t_{51} = -12.05$, $p < 0.001$, after accounting for year, Fig 3a). Thus, groups with high pre-moult SD tended to have reductions in post-moult SD, while groups with low pre-moult SD tended to have increases in post-moult SD (i.e. regression toward the mean). However there was no indication that experimental treatment had any additive impact on the nature of this relationship. Then, we analyzed post-moult SD in relation to treatment after accounting for pre-moult SD (Fig. 3b). As expected, the slopes of the relationship were similar between treatments after accounting for year (lm: $t_{49} = 0.69$, $p = 0.491$ after accounting for year). Contrary to our prediction however, there was no effect overall for treatment (lm: $t_{50} = 0.11$, $p = 0.914$ after accounting for year).

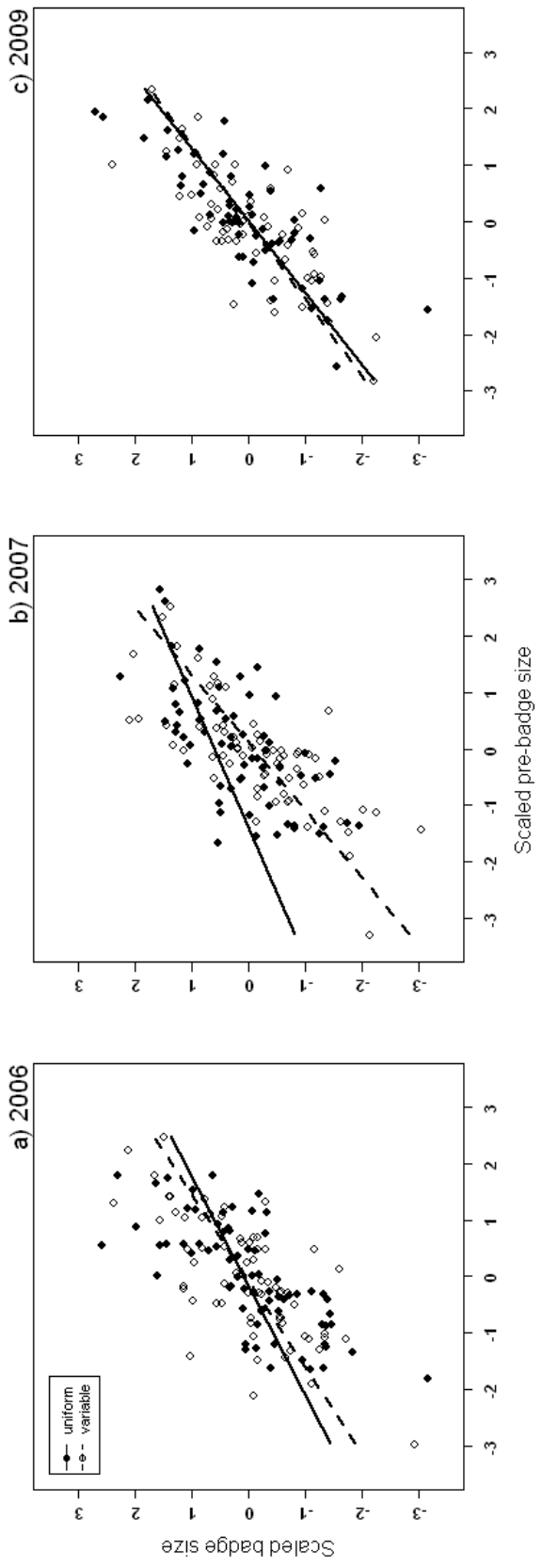


Fig. 2 Scaled post-moult badge size in relation to scaled pre-moult badge size of male House Sparrows in the two experimental groups in the three different years a), b), and c). We centred and then scaled badge sizes separately for post- and pre-moult badges by year, i.e. we subtracted the yearly mean badge size from an individual's badge size and divided the results by the standard deviation. We did so for this graph in order to exclude confounding year effects. Closed circles and straight lines represent birds in the uniform-badges-treatment group, open circles and dashed lines represent birds in the variable-badges-treatment group.

Table 1: Results of linear mixed effect models of badge size in relation to the interaction of pre-moult badge size and experimental group after accounting for condition and with social group ID (i.e. aviary ID) as a random factor. In the overall model we also accounted for year and included bird ID as a random factor.

Year	Pre-moult badge		Treatment		Interaction		
overall	$z = 6.33$	$p < 0.001$	$z = -0.80$	$p = 0.422$	$z = 0.81$	$p = 0.422$	n = 370 observations
2006	$t_{111} = 6.18$	$p < 0.001$	$t_{22} = 0.95$	$p = 0.354$	$t_{111} = -0.82$	$p = 0.412$	
2007	$t_{112} = 5.61$	$p < 0.001$	$t_{12} = -2.70$	$p = 0.019$	$t_{112} = 2.55$	$p = 0.012$	
2009	$t_{81} = 8.12$	$p < 0.001$	$t_{14} = 1.26$	$p = 0.230$	$t_{81} = -1.23$	$p = 0.220$	

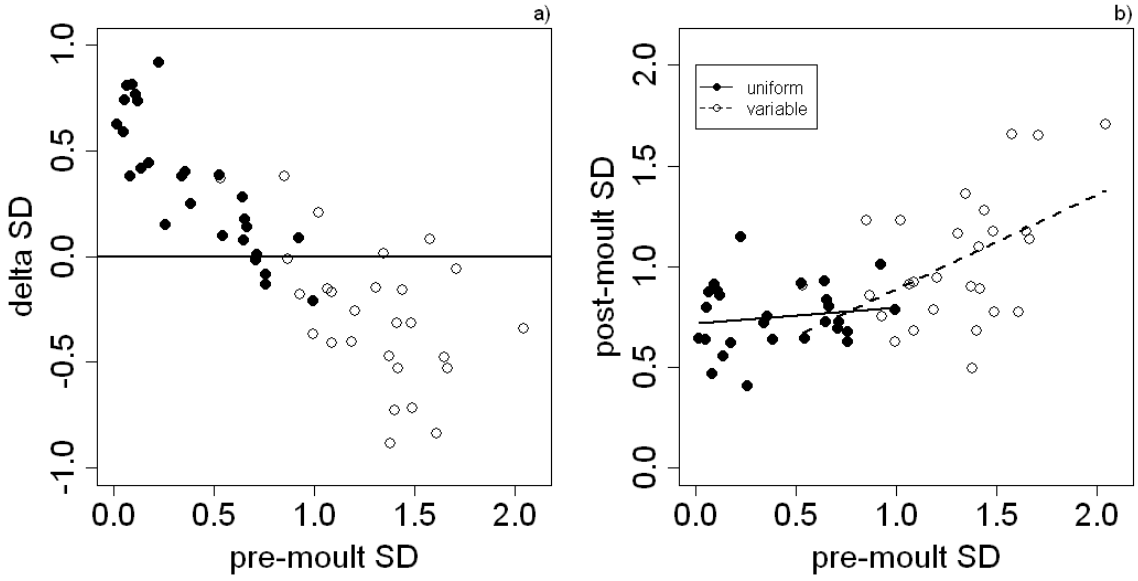


Fig. 3 a) Delta (i.e. post-moult minus pre-moult) group standard deviation of badge sizes in relation to pre-moult group standard deviation in the two experimental groups in the three years combined. b) Post-moult group standard deviation in relation to pre-moult group standard deviation in the two experimental groups in the three years combined.
Closed circles and straight lines represent birds in the uniform-badges-treatment group, open circles and dashed lines represent birds in the variable badges treatment group.

Discussion

We expected that a uniform group composition would on average change an individual's quality more (relative to its flock mates) than in variable badges groups. We predicted that this would be reflected in stronger changes in badge size and more intra-flock variability in uniform groups. However, we found that in our two treatment groups badge sizes changed differently in only one of the three years. In addition, between individual variation in badge size was not affected by treatment. Note that the analyses for group variation were not very powerful and new statistical methods such as double GLMs (D. Westneat personal communication) would be highly recommended for the future. Overall, our results demonstrate that group composition of badge sizes during moult had only a weak effect on male ornamentation in House Sparrows.

The above results suggest that either our treatment only weakly influenced levels of aggressive interactions and hence social environment or that there is no general effect of group composition on badge size. Aggression during moult is generally low (Hahn

et al. 1992), and thus the manipulation of social environment might be hard to perform or might only cause weak effects. In addition, we used large aviaries, where the birds could have had enough space to escape from each other and to establish social structures without much aggressive interactions. However, group sizes were rather large (six to ten individuals) and close to the maximal capacity of these aviaries which led to massive interactions during breeding season (personal observations).

Nevertheless, we did find a relatively clear significant effect in the predicted direction in 2007 (Fig 2b). It is noteworthy that group sizes were larger in this year than in the other two years, which potentially led to more social interactions between flock mates. Additionally, in 2007, we started the experiment a few weeks before the onset of moult in contrast to the other two years when we formed the groups immediately before the beginning of moult. (Note that the group sizes we used and the dates we formed the groups were largely dependent on other research protocols operating concurrently to this experiment). Therefore in 2007, the birds placed in new groups still had presumably elevated testosterone levels and higher motivation to engage in aggressive interactions because testosterone levels do not drop until the onset of moult (Laucht et al 2010). Moreover, it is unknown exactly when any physiological "programming" of badge size can or does occur. It could be that a critical window of physiological and cellular badge size programming occurs in advance of the oncoming moult. In 2007, it could have therefore been that this programming occurred after the males were placed in their new groups, but in 2006 and 2009 this programming had already occurred prior to new group formation. If so, then this suggests that the social environment could indeed have a general effect on ornament development, but that the critical window for this effect to manifest occurs before moult rather than during moult.

Our results differed from the only other study that examined social effects during moult on House Sparrow badge sizes (McGraw et al. 2003). This study, which looked at the effects of aggression levels and dominance in triads of House Sparrows kept in cages, suggests clear effects of the social environment while our study suggests rather small carry-over effects from before moult. These differences could be due to differences in the two experimental set-ups. First, we used large aviaries as opposed to small cages. Thus, escalated aggressive interactions in these small cages

could have led to a larger increase of testosterone. Additionally, interactions could have been more severe because contestants could not escape from each other. Second, McGraw et al. (2003) only found effects of social environment on badge sizes of beta males. Because we did not score dominance ranks we cannot exclude the possibility that beta males in our groups had more pronounced effects on badge size changes. However, dominance structures in large groups are rarely linear (reviewed in Anderson 2006) suggesting that more than one bird per group should have been affected which we should have been able to observe with our set-ups. Third and most importantly, McGraw et al. (2003) arranged groups of birds long before the onset of moult. Therefore, the observed effect in their study could have easily resulted from social interactions occurring during a badge programming window that occurs prior to moult rather than during moult. If so, then McGraw et al's (2003) results would be in agreement with ours.

To conclude, we found that group composition during moult had probably no direct influence on badge size. However, social group composition in the period just before moult potentially had an effect. Our results suggest the possibility of a critical "programming" window just prior to moult where social interactions contribute to future badge size. Future research aimed at identifying the exact timing of when badge size is determined relative to moult would be extremely valuable. The fact that there are so many unknowns here is testimony to the exploratory nature of this research and to how little we actually know about the mechanisms of ornament development in general.

Acknowledgements

Many thanks to Katja Temnow and Andrea Wittenzellner for help with rearranging birds. We are grateful to Mihai Valcu for statistical advice, to Magnus Rudorfer, Frances Weigel, Edith Bodendorfer, and Annemarie Grötsch for animal care, and to the Behavioural Ecology and Evolutionary Genetics Group at the MPIO for discussion and comments. Funding for this research was provided by the Max Planck Society.

References

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophoton. Int. 11:36-42

Anderson TR (2006) Biology of the ubiquitious house sparrow. Oxford University Press, New York

Bókony V, Lendvai AZ, Liker A (2006) Multiple cues in status signalling: The role of wingbars in aggressive interactions of male house sparrows. Ethology 112:947-954

Buchanan KL, Evans MR, Goldsmith AR, Bryant DM, Rowe LV (2001) Testosterone influences basal metabolic rate in male house sparrows: a new cost of dominance signalling? Proc. R. Soc. Biol. Sci. Ser. B 268:1337-1344

Dale J (2006) Intraspecific variation in coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration Harvard University Press, MA, Cambridge, pp 36-86

Evans MR, Goldsmith AR, Norris SRA (2000) The effects of testosterone on antibody production and plumage coloration in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 47:156-163

Freckleton RP (2002) On the misuse of residuals in ecology: regression of residuals vs. multiple regression. J. Anim. Ecol. 71:542-545

Gonzalez G, Sorci G, Smith LC, de Lope F (2001) Testosterone and sexual signalling in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 50:557-562

Green AJ (2001) Mass/length residuals: Measures of body condition or generators of spurious results? Ecology 82:1473-1483

Hahn TP, Swingle J, Wingfield JC, Ramenofsky M (1992) Adjustments of the Prebasic Molt Schedule in Birds. Ornis Scand. 23:314-321

Hein WK, Westneat DF, Poston JP (2003) Sex of opponent influences response to a potential status signal in house sparrows. Anim. Behav. 65:1211-1221

Hill GE (2006) Environmental Regulation of Ornamental Coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration. Harvard University Press, Cambridge, pp 507-560

Jawor JM, Breitwisch R (2003) Melanin ornaments, honesty, and sexual selection. Auk 120:249-265

Laucht S, Dale J, Mutzel A, Kempenaers B (2011) Individual variation in plasma testosterone levels and its relation to badge size in House Sparrows *Passer domesticus*: It's a night-and-day difference. Gen. Comp. Endocrinol. 170:501-508

Laucht S, Kempenaers B, Dale J (2010) Bill color, not badge size, indicates testosterone-related information in house sparrows. Behav. Ecol. Sociobiol. 64:1461-1471

Liker A, Barta Z (2001) Male badge size predicts dominance against females in House Sparrows. Condor 103:151-157

Maynard Smith JFRS, Harper DGC (1988) The Evolution of Aggression - Can Selection Generate Variability? Philos. Trans. R. Soc. Lond. B Biol. Sci. 319:557-570

McGraw KJ, Dale J, Mackillop EA (2003) Social environment during molt and the expression of melanin-based plumage pigmentation in male house sparrows (*Passer domesticus*). Behav. Ecol. Sociobiol. 53:116-122

Møller AP (1987a) Social Control of Deception among Status Signaling House Sparrows *Passer domesticus*. Behav. Ecol. Sociobiol. 20:307-311

Møller AP (1987b) Variation in Badge Size in Male House Sparrows *Passer domesticus* - Evidence for Status Signaling. Anim. Behav. 35:1637-1644

Møller AP (1990) Sexual-Behavior Is Related to Badge Size in the House Sparrow *Passer domesticus*. Behav. Ecol. Sociobiol. 27:23-29

Morrison EB, Kinnard TB, Stewart IRK, Poston JP, Hatch MI, Westneat DF (2008) The links between plumage variation and nest site occupancy in male House Sparrows. Condor 110:345-353

Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T (2007) Assessing the function of house sparrows' bib size using a flexible meta-analysis method. Behav. Ecol. 18:831-840

Oyegbile TO, Marler CA (2005) Winning fights elevates testosterone levels in California mice and enhances future ability to win fights. Horm. Behav. 48:259-267

Parker TH, Knapp R, Rosenfield JA (2002) Social mediation of sexually selected ornamentation and steroid hormone levels in male junglefowl. Anim. Behav. 64:291-298

R Development Core Team (2010) R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria

Setchell JM, Smith T, Wickings EJ, Knapp LA (2008) Social correlates of testosterone and ornamentation in male mandrills. Horm. Behav. 54:365-372

Tibbetts EA, Dale J (2004) A socially enforced signal of quality in a paper wasp. Nature 432:218-222

Tibbetts EA, Safran RJ (2009) Co-evolution of plumage characteristics and winter sociality in New and Old World sparrows. J. Evol. Biol. 22:2376-2386

Veiga JP (1993) Badge Size, Phenotypic Quality, and Reproductive Success in the House Sparrow - a Study on Honest Advertisement. Evolution 47:1161-1170

Veiga JP (1995) Honest Signaling and the Survival Cost of Badges in the House Sparrow. Evolution 49:570-572

Wingfield JC, Hegner RE, Dufty AM, Ball GF (1990) The Challenge Hypothesis - Theoretical Implications for Patterns of Testosterone Secretion, Mating Systems, and Breeding Strategies. Am. Nat. 136:829-846

Non-submitted Manuscript

Chapter Five

No evidence for an interaction between testosterone, immune function, and carotenoid availability in House Sparrows and Red-billed Queleas: implications to testosterone handicap models of honest signaling

Silke Laucht, Bart Kempenaers, Kevin McGraw, Elena Arriero, Wolfgang Goymann, Dennis Hasselquist, James Dale

Abstract

It is poorly understood how different factors of the signalling system inter-play and ensure signal honesty with respect to individual quality. The "testosterone handicap model" (THM) is probably the most widely accepted model to explain the honesty of testosterone (T) dependent signals of quality. It posits inescapable physiological trade-offs between elevated T levels, the degree of ornament elaboration, and various physiological costs of T (especially immunosuppression and/or oxidative stress). Putatively strong support for the THM has come from research on carotenoid based ornamentation. Because carotenoids are assumed to be limited in availability, are required for signal development, and are thought to be immunoenhancers and powerful antioxidants they provide an intuitive mechanism whereby T related handicaps can manifest. In this study we test three fundamental assumptions of the THM as applied to carotenoid based ornamentation: (1) T is immunosuppressive, (2) carotenoids are immunoenhancing, and (3) T increases the bioavailability of carotenoids. Unlike previous research, however, we test these assumptions using one species with a T dependent melanin based ornament (the House Sparrow, *Passer domesticus*), and one species with a non-T dependent carotenoid based ornament (the Red-billed Quelea, *Quelea quelea quelea*). Although the predicted trade-offs in these species should also occur, we show that they do not; which raises the critical question of whether they can contribute to signal honesty in any species. Our results strongly support the view that the observed effects of T on ornaments and various physiological processes are the adaptive outcome of selection on the signalling system, rather than proximate constraints which can maintain signal honesty.

Introduction

Despite widespread animal ornamentation research over the past three decades, it is still very poorly understood how different factors of the signalling system inter-play and ensure signal honesty with respect to individual quality. The fundamental assumption underlying the vast majority of research is that ornamental traits require inescapable costs (or handicaps) in order to remain honest indicators of quality (Zahavi 1975; Zahavi 1977; Grafen 1990; Johnstone and Norris 1993). A great majority of studies about costly signals of quality have reasoned that signal honesty might be enforced through the "testosterone handicap model" (hereafter referred to as the THM) (Adkins-Regan 2005) which posits inescapable physiological trade-offs between elevated testosterone (T) levels, the degree of ornament elaboration, and various physiological costs of T – most notably decreased immune function and increased oxidative-stress (Folstad and Karter 1992; Wedekind and Folstad 1994; von Schantz et al. 1999; Poiani et al. 2000; McGraw and Ardia 2003; Alonso-Alvarez et al. 2004; Mougeot et al. 2004; Owen-Ashley et al. 2004; Bertrand et al. 2006; Blas et al. 2006; Alonso-Alvarez et al. 2007; McGraw and Ardia 2007; Mougeot et al. 2007; Peters 2007; Roberts et al. 2007; Alonso-Alvarez et al. 2008; Mougeot et al. 2009; Martínez-Padilla et al. 2010; Vinkler and Albrecht 2010).

The THM has received particularly strong support in studies focused on carotenoid based T dependent ornaments. Carotenoids are bioactive yellow-red pigments that need to be acquired from the diet (in vertebrates) and are also argued to have important immunoenhancing and anti-oxidant physiological functions (reviewed in Peters 2007). Carotenoid based T dependent ornaments are hypothesized, under the THM, to therefore honestly signal the quality of the immune system and/or the ability to withstand oxidative stress to potential mates. This view critically assumes that the costs that keep ornaments honest are physiological trade-offs stemming from limited carotenoid bioavailability: carotenoids can either be used for signal elaboration or to physiologically buffer the immunosuppression and/or oxidative stress caused by elevated T levels (Peters 2007). The ability to negotiate this trade-off is expected to be positively correlated with the degree of ornament elaboration.

Studies that have evaluated the THM in species with carotenoid based ornamentation typically test for (1) the T dependency of the ornament (e.g. Mougeot et al. 2007; Cassagrande 2010; Martínez-Padilla et al. 2010), (2) the immunosuppressive effect of T (e.g. Hasselquist et al. 1999; Duffy et al. 2000; Owen-Ashley et al. 2004), (3) a positive correlation between carotenoids and enhanced immune function and/or anti-oxidant capacity (e.g. Alonso-Alvarez et al. 2004; McGraw and Klasing 2006; Biard et al. 2009), and (4) a T related increase of the bioavailability of carotenoids (e.g. Blas et al. 2006; McGraw et al. 2006; McGraw and Ardia 2007). Support for any of these mechanistic predictions is usually taken as positive evidence that the ornament functions as an honest signal of immune system quality and/or the ability to withstand oxidative stress to choosy females (e.g. in Faivre et al. 2003).

In this approach, the hypothesized function (inter-sexual signalling) is used to predict expected physiological mechanisms. Although elegant, this approach is conceptually problematic because it is an example of "mechanistic just-so story-telling": the observed mechanisms are used to infer putative function (e.g. if there is a link between ornamentation and immune response, the ornament must therefore signal immune function to receivers (Faivre et al. 2003)). Indeed, it is very rarely considered whether alternative costs for ornamentation are also consistent with the predicted mechanisms of the THM (Martin et al. 2006c). For example, although a simple relationship between immune response and ornamentation is consistent with an inter-sexual immune signalling function, it is also predicted by signalling social dominance to competitive rivals (dominant individuals should have better immune systems (Hasselquist et al. 1999; Verhulst et al. 1999; Poiani et al. 2000; Roberts et al. 2007)), or even traditional Zahavian-type handicaps such as advertising the ability to avoid predators (individuals with compromised immune systems should decrease their ornamentation because they are more vulnerable to predators (Møller and Erritzøe 2000)). Indeed all of the putative mechanistic links of the THM are consistent with multiple alternative signalling functions and costs for ornaments. However, the THM does critically assume that the underlying mechanisms provide inescapable physiological constraints that ensure the signalling honesty of the ornament. Importantly, other functional hypotheses (such as signalling social dominance or predator-avoidance abilities) do not make this assumption - they instead argue that

these underlying mechanisms are byproducts (i.e. physiological adaptations) that are the consequences of (rather than constraints on) signalling function.

In this study we test the THM by evaluating whether the underpinning physiological mechanisms are general. That is, we test whether these mechanisms have the general potential to provide mechanistic constraints that can enforce honest signalling of immune function and/or oxidative stress per se. Therefore, a counter-intuitive but *strong* test for the potential signal honesty via the THM is to evaluate these mechanisms in species that either (1) do not have carotenoid based ornaments or (2) do not have T dependent ornaments. This approach removes one confounding factor (i.e. the carotenoid dependency or the T dependency of the ornament, respectively) of this complex inter-play in order to test for the generality of the required critical assumptions for signal honesty. In the first case, for an ornament that is not carotenoid based, T should increase ornament elaboration as much as it suppresses the immune system/causes oxidative stress, and T should enhance carotenoid availability for the immune and anti-oxidant system (since carotenoids are not needed for the ornament). In the second case, T should be immunosuppressive even though the ornament is not T dependent and T should enhance bioavailability of the carotenoids in order to buffer these immunosuppressive effects and oxidative stress.

Here we evaluate these links between T, carotenoids, and immune function in one species with a T dependent but not carotenoid based ornament (the House Sparrow, *Passer domesticus*) and in one species with a carotenoid based ornament that is not T dependent (the Red-billed Quelea, *Quelea quelea quelea*). We measured plasma T and carotenoid levels, ornamentation, and immune response in T and placebo-implanted birds before and after two different immune challenges. Both species have dynamic traits, their bill colours, that can change within a few days to weeks (Anderson 2006) and thus, respond faster to changes in underlying factors than plumage ornaments would.

The House Sparrow has a melanin based bill colour that changes from a pale horn colour in non-breeding season to a black in the breeding season (Witschi and Woods 1936; Laucht et al. 2010). Outside the breeding season, variation in bill colour is strongly T dependent (Laucht et al. 2010). The T bill colour relationship reaches a

plateau in peak breeding season (i.e. higher T levels do not increase coloration further, but are needed to maintain the coloration) when all males have dark bills. We evaluated birds in breeding condition for the sexual selection context (Blas et al. 2006), and to provide a maximally natural physiological T treatment to show maximal suppressive effects of T (Duffy et al. 2000). This ensures our study captures all potential seasonal physiological cascades and mechanisms associated with the immune system, carotenoids and T, and in order to have comparable values to other studies (e.g. Poiani et al. 2000; Westneat et al. 2003; Blas et al. 2006; McGraw and Ardia 2007).

The Red-billed Quelea has a carotenoid based red bill that is correlated to social dominances (Shawcross and Slater 1984) and phenotypic condition (Dale 2000). Bill redness in this species is not T dependent, but rather estrogen inhibited (Witschi 1961). Males maintain red bills all year round (despite massive seasonal fluctuations in T), whereas females only develop the red coloured bills in the non-breeding season when their ovaries are regressed (Witschi 1961; Owens and Short 1995). Castrated males maintain their red bill colour, whereas ovariectomized females develop red bills (Witschi 1961; Owens and Short 1995). Queleas do not have any known T dependent ornaments.

In this study, we test three general predictions underlying the THM's expected relationships between carotenoids, testosterone, and the immune system: (1) T is immunosuppressive, (2) carotenoids are immunoenhancing, and (3) T increases the bioavailability of carotenoids. Our reasoning is that since the THM assumes these relationships provide inescapable physiological constraints that enforce signal honesty, then they should occur independently of whether the ornament is carotenoid dependent (e.g. House Sparrow), or T dependent (e.g. Red-billed Quelea). If, on the other hand, these relationships are not supported in these species, then this suggests that when they do occur in other species, they are best interpreted as physiological adaptations that are *consequences* of signal function, rather than mechanistic *constraints* that enforce signal honesty.

To test these three general predictions we performed repeated measures of ornamentation and T levels in T- and placebo-implanted males. We assessed immune

response by challenging birds with (1) PHA, and (2) LPS or PBS (as controls). PHA (phytohaemagglutinin) injections are very widely used to assess a bird's immunocompetence, and specifically they provide a multifaceted measure of cutaneous immune activity (Martin et al. 2006b). LPS (lipopolysaccharides of a gram-negative bacterium) injections as immune challenges have recently become more common and induce an immediate inflammatory response and the production of specific antibodies within a few days (summarized in e.g. Bonneaud et al. 2003; Owen-Ashley et al. 2006). We measured various immune parameters such as skin swelling after PHA, haematocrit, activity of natural antibodies and complement proteins, lysozyme concentration, haptoglobin concentration, total IgG antibodies, and specific LPS antibodies. We assessed carotenoid levels by repeated samples of blood serum carotenoids. In addition we also measured dietary intake, to see if carotenoid requirements were affecting the amount of food eaten. For each general prediction, we performed multiple specific tests listed in Table 1.

Table 1: Summary of hypotheses and observations in our study.

General Prediction	Test	Observation
T is immunosuppressive	negative effect of T implant on PHA skin swelling	no effect
	negative effect of T implant on immune measures	no effect
Carotenoids are immunoenhancing	PHA response should be positively correlated with carotenoids	no correlation
	immune measure should be positively correlated with carotenoids	no correlation
	total carotenoids should decrease after immune challenge because they were used up	observed, but differently in the two species
	higher food intake after immune challenges to counteract usage of carotenoids	stress- and sickness-related lower intake
T increases bioavailability of carotenoids	positive relation between plasma carotenoids and T levels	no relationship
	negative effect of T implant and LPS challenge on carotenoids (2-way-interaction)	no interaction
	combined effect of implant, immune challenge and carotenoids on immune response (3-way-interaction)	no interaction
	higher food intake in T implanted birds to counteract usage of carotenoids	no difference

Material and Methods

Study population

a) House Sparrows (Passer domesticus)

We studied 34 randomly chosen male House Sparrows from our captive population held at the Max Planck Institute for Ornithology, Seewiesen, Germany (see Laucht et al. 2010). All males were wild caught in rural areas in Bavaria, Germany (under license: permit nr. 55.1-8642.3-3-2006 of the "Regierung Oberbayern", with several extensions) and held in captivity for 28 months. Before the experiment all individuals were kept in groups of five to ten in aviaries of size 1.2 x 2.0 x 4.0 m. At all times, the birds had *ad libitum* access to food (wild seed mix for forest birds (Waldvogelfutter: RKW Sued, Universal Kraftfutterwerk, Kehl, Germany), sunflower seeds, crushed corn and wheat, oats, chicken starter, soybean meal extract, and mineral mix for birds), drinking and bathing water, and sand. The light-dark cycle and temperatures in the aviaries were close to natural conditions, as the aviaries were semi-outdoor with one side enclosed only by chicken wire.

b) Red-billed Queleas (Quelea quelea quelea)

We studied 34 randomly chosen male Red-billed Queleas from our captive population held at the Max Planck Institute for Ornithology, Seewiesen, Germany. The birds had been wild caught at unknown specific locations in Senegal and then transported to Germany for sale in the pet trade. The birds had been at the institute for over two years prior to the experiment. All birds were held in a single mixed-sexed group in an indoor/outdoor flight pen of size 7.0 x 2.5 x 14.7 m before the experiment (during the winter the birds were excluded from the outdoor portion of the enclosure; aviary size was then 7.0 x 2.5 x 3.7 m). At all times, the birds had *ad libitum* access to food (tropical seed mixture (Supravit GmbH, Bruckmühl, Germany): five different kinds of millet, 10% canary grass, 2% blackseed), drinking and bathing water. In the indoor-aviary, the daily light dark cycle was maintained at 12:12, and the temperature was maintained at 24 ºC.

Experimental approach

The experiment was performed in 2008 starting in May with the House Sparrows and August with the Queleas. Three weeks prior to the experiment we moved the birds into cages of size 0.61 x 0.5 x 0.4 m for acclimatization. Each bird was kept alone in one cage to facilitate experimental procedures and to avoid social modulation of T levels and ornaments. However, since both species are social, we kept all experimental birds in the same room in such a way that each male could interact with another male (visually and acoustically) through a grid and with the rest of the males acoustically. Each cage was equipped with three perches, food (as described above for each species), drinking water *ad libitum,* and bathing water twice a week. Birds were exposed to natural light and temperature fluctuations at the times of year of their experiment.

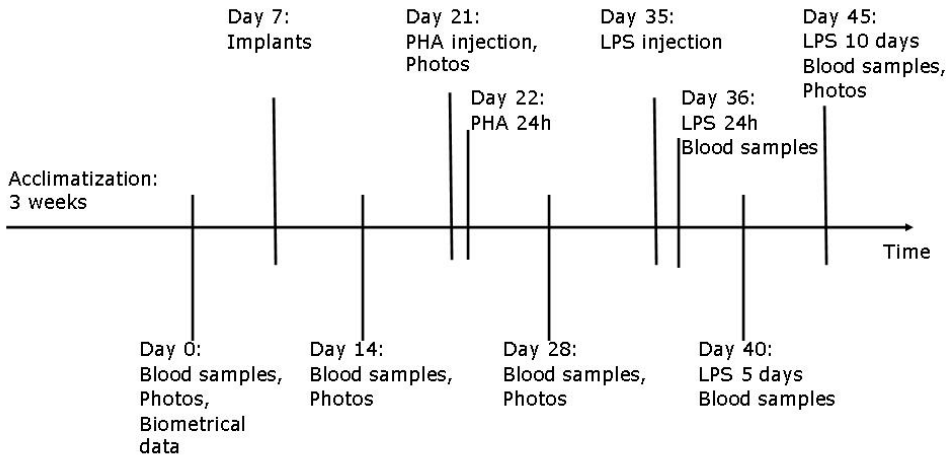


Fig. 1 Timeline of experiment

Fig. 1 provides an overview of the experimental timeline. On day zero, 14 and 28 of the experiment we caught all birds, took blood samples, photographs of the bill and measured body mass. On day seven we implanted half of the birds (n = 17) with testosterone pellets (1.5 mg testosterone, 60 day release, Innovative Research of

America, Sarasota, Florida) and half of them (n = 17) with placebo pellets consisting of pure binding material (60 day release, Innovative Research of America, Sarasota, Florida) subcutaneously through a small incision on the upper back. For more details on release rates of pellets and a comparison to silastic tubes see Fusani (2008). The incision was sealed with VetGlu tissue adhesive (Heiland Vet GmbH, Germany). On day 21 of the experiment, we measured the thickness of the right wing web with a metric mechanical thickness gage (model 304-196, The Dyer Company, Lancaster PA, USA) and injected 50 µg PHA (Lectin from *Phaseolus vulgaris*, L8754, Sigma-Aldrich Co., Missouri, USA) diluted in 50 µl of PBS buffer on the same spot. We used the simplified method without control injection of PBS in the other wing web as suggested by Smits et al. (1999) and supported by Martin II et al. (2006b). After 24h (Martin et al. 2003; Biard et al. 2009) (day 22 of the experiment), we re-measured thickness of the wing web. On day 35 of the experiment, we injected half of the T-implanted (n = 9) and half of the placebo-implanted (n = 9) birds intraperitoneally with 0.001 mg LPS (Lipopolysaccharides from *Salmonella enterica* serotype *typhimurium*, L7261, Sigma-Aldrich Co., Missouri, USA) per gram body mass (i.e. 0.028 mg for the House Sparrows and 0.018 mg for the Queleas) diluted in 0.1 ml PBS buffer. The other halves of the birds were injected with the same amount of PBS buffer for control. We took additional blood samples 24 hours (day 36), five (day 40) and ten (day 45) days after the LPS injection. For logistical feasibility of the above procedures we divided all birds into two groups and staggered day 0 of the experiment by two days. All treatments were evenly distributed among these two groups and among cages and neighbours.

During blood sampling we always caught and bled birds at the same time of day in the afternoon, and in the same order to keep effects of stress and time since first starting to catch birds constant. We took 150-200 µl of blood from the wing vein, collected the blood in 75-mm Na-heparinized micro haematocrit capillaries, and centrifuged it at 13,000 rpm for 3 min to separate the plasma. Plasma was stored at -70 ºC.

All experimental procedures were approved by the Ethical Committee of the government of Upper Bavaria (permit nr: 55.2-1-54-2531-118-07 of the "Regierung von Oberbayern").

Determination of food intake

We measured individual food intake two days of each week of the experiment. On the first day the birds were unstressed, i.e. not captured for measurements. At time 0, we cleaned all cages and equipped them with fresh paper on the floor, and weighed the amount of food that was provided. After 24 hours we weighed the food that was remaining in the food dish and scattered uneaten on the cage floor, and calculated the difference between food weights as total intake. We repeated the procedure on the second day, however on this day the birds were stressed at some point because of blood sampling, implanting or immune challenges. So, each week we could compare food intake on a day they were unstressed and a day they were stressed.

In addition, immediately after the LPS challenge we measured food intake daily for as many days needed for the birds to get back to pre-challenge levels of intake (three days for House Sparrows and four for Queleas).

Determination of plasma T levels

We measured T levels from blood samples taken on day zero, 14 and 28. We determined plasma T levels via direct radio-immunoassay using testosterone antiserum T3-125 (Esoterix Endocrinology, Calabasas, CA, USA) following the protocols described in Goymann et al. (2002; 2006). Cross reactivities of this antiserum are as follows: testosterone 100%, 5a-dihydrotestosterone 44%, d-1-testosterone 41%, d-1-dihydrotestosterone 18%, 5a-androstan-3b,17b-diol 3%, 4-androsten-3b,17b-diol 2.5%, d-4-androstenedione 2%, 5b-androstan-3b,17b-diol 1.5%, estradiol 0.5%, and less than 0.2% with 23 other steroids tested. Plasma samples were equilibrated with 1500 dpm of tritiated testosterone (Perkin Elmer, Wellesley, MA, USA) for the calculation of recoveries. Mean $\pm$ SD extraction efficiency for plasma T was 91 $\pm$ 0.4 % for House Sparrows and 88 $\pm$ 1.3 % for Queleas. We measured standard curves and sample concentrations in duplicates and calculated them with Immunofit 3.0 (Beckman Inc., Fullerton, CA, USA) using a four parameter logistic curve fit. We defined the lower detection limits of the standard curves as the first value outside the 95% confidence intervals for the zero standard (B_{max}); it was 0.3 (House Sparrow samples) and 0.4 pg/tube (Quelea samples). The intra-assay

coefficient of variation was 6.9 % (determined from standard testosterone) and 7.7 %. The intra-extraction coefficient of variation of extracted plasma pools was 1.3 and 0.2 %. Because of the significant cross-reactions of the used testosterone antibodies with 5a-dihydrotestosterone (44%) our T measurements may include a fraction of 5a-DHT.

As the amount of time passed since first starting to catch the birds had a slight negative effect on T levels in House Sparrows (lme: $t = -2.52$, $df = 67$, $p = 0.014$, random effect bird ID) and as body mass had a slight positive effect in Queleas (lme after accounting for time since starting to capture: House Sparrows: $t = 0.35$, $df = 66$, $p = 0.275$; Queleas: $t = 2.25$, $df = 64$, $p = 0.028$; random effect: bird ID) we included time and body mass in the statistical models of both species.

Determination of bill color

Before starting the experiment and on days zero, 14, 21, 28, and 45 we photographed bill coloration under standardized flash-photography conditions with a Canon Power Shot S2 IS camera. We took two photos of each House Sparrow's right profile and one photo of each Quelea's right profile, left profile and top of the head respectively (i.e. three photos of the bill in total).

We quantified bill coloration in both species by measuring bill "brightness" on the HSB (i.e. hue, saturation, brightness) colour scale as measured with image processing software written into R 2.4.0 (R Development Core Team 2006). Bill brightness is analogous to total reflectance and we have previously demonstrated this measure to be an intuitive quantification of House Sparrow bill colour variation (Laucht et al. 2010). In Queleas, brightness also best represented apparent colour variation because bill colour in this subspecies varies from very dark red to brighter orange. Note that analyses on other measures of bill coloration in Queleas (i.e. hue, saturation, and PC1 & PC2 of RGB scores) yielded qualitatively similar results to those we report here for brightness.

On each photo, SL measured the brightness at five randomly chosen positions each on the upper bill, the lower bill and the gray colour background immediately adjacent to the bill (in the case of the top of the head Quelea photos, only the upper bill and grey

card were measured). To standardize slight variation between photos, we calculated the mean gray card brightness of all photos, determined the deviation of the grey card brightness of a focal photo from this overall mean, and subtracted this deviation from the mean bill brightness for focal photo (for more details see Laucht et al. 2010). For analyses, we used the total means of upper and lower brightness scores of the two or three photos for House Sparrows and Queleas respectively. These measurements were highly repeatable within individuals (repeatability (Lessells and Boag 1987): for House Sparrows: $R = 0.805 \pm 0.025$ (SE), $p < 0.001$, $n = 2 \times 204$ for two pictures; for Queleas: $R = 0.656 \pm 0.040$ (SE), $p < 0.001$, $n = 2 \times 201$ for two pictures).

Determination of plasma carotenoid levels

Whenever we took blood samples we collected some plasma for carotenoid analyses. We extracted plasma carotenoids according to the protocol for the ethanol-TBME extraction method described in McGraw et al. (2008). In brief, solvent was added to the plasma, and after centrifugation evaporated to dryness with nitrogen. Extracted carotenoids were immediately frozen and sent to Tempe, Arizona, USA for further analyses via high-performance liquid chromatography (McGraw et al. 2008).

For both species, the different carotenoid measures (lutein, zeaxanthin, lutein cis isomers, and total carotenoids for House Sparrows; lutein, zeaxanthin, alpha-doradexanthin, astaxanthin and total carotenoids for Queleas) were highly correlated ($R > 0.512$). Therefore, we used total carotenoids for the analyses.

As total carotenoids were slightly correlated with time passed since first starting to capture birds (lme: House Sparrows: $t = 2.04$, $df = 169$, $p = 0.043$; Queleas: $t = -3.08$, $df = 150$, $p = 0.003$; random effect: bird ID), we included time in the models.

Determination of immune responses

At various times of the experiment (Fig. 1) we measured haematocrit, activity of natural antibodies and complement proteins, lysozyme concentration, haptoglobin concentration, total IgG antibodies, and specific LPS antibodies (each described in more detail below). Because results were very similar for single immune measures,

we also combined these different measures (after z-transformation) into a single immune-response index in order to facilitate detection of broad patterns.

a) Haematocrit

We measured this as the ratio of red blood cells to overall blood after centrifugation in the capillary tube.

b) Activity of natural antibodies and complement proteins

We used this measure as an estimation of constitutive immune defences (i.e. expressed at all times). We followed the protocol by Matson et al. (2005), however scaled down to 15µl of plasma with the blood samples taken on day 28, 40 and 45 in the House Sparrows and on day 40 and 45 in the Queleas. Natural antibody (Nab) titres and complement activity were scored as $-\log_2$ of the highest dilution exhibiting agglutination (Nab) or lysis (complement). Samples in which no agglutination or no lysis was observed were scored as 0. Samples with less than 15µl of plasma were not used in the assay.

c) Lysozyme concentration

We measured this with blood samples taken on day zero and 36 in House Sparrows and with the samples of day 28 and 36 in Queleas. We followed the protocol of Millet et al. (2007), except we used 300 µl agar suspension and 15 µl of plasma per well. Note that we excluded some samples because of too little plasma.

d) Haptoglobin concentration

We measured this with blood samples taken on day zero and 36 in House Sparrows and day 28 and 36 in Queleas. We used a commercial haptoglobin test kit (TP-801, Tridelta Development Limited, Maynooth, Co. Kildare, Ireland). Briefly, we measured peroxidase activity of haemoglobin that is directly proportional to the concentration of haptoglobin in the sample. Concentration was determined via a standard curve and readings with a plate reader (VersaMax ELISA Microplate Reader, Molecular Devices, Inc., Sunnyvale, CA, USA) at 630 nm.

e) Total IgG and anti-LPS antibodies

We measured this with blood samples taken on day 28, 40 and 45 (and day zero for some House Sparrow samples). Note that we excluded some samples because of too little plasma. We analyzed samples in Lund, Sweden. Generally, we followed the protocol described in Westneat et al. (2003). Briefly, we coated plates with donkey-anti-chicken-IgG for total IgG antibodies and with rabbit-anti-redwing-LPS antibodies for specific LPS antibodies. We diluted plasma samples 1:200 for LPS antibodies and 1:400 for total IgG antibodies. On each plate we ran duplicates of each sample and seven different dilutions of a known standard plus blanks for the calculation of the standard curve. For analyses, we used the means of the calculated concentrations of the two duplicates of each sample.

Statistical analyses

We performed all statistical analyses using R 2.11.0 (R Development Core Team 2010; packages: effects, gdata, lattice, lme4, nlme, RODBC) at the significance level $\alpha = 0.05$. To account for multiple measures of the same individuals we used linear mixed effect models (lme) with bird ID as a random effect. When only one measurement per bird was needed we used linear models (lm).

Unless indicated, results do not qualitatively change when analyzed with delta values (i.e. between the beginning of the experiment and the focal time point) for carotenoids, testosterone levels, bill color, or food intake. Results also do not change when using food intake per unit body size.

Results

Effect of implants on T levels

a) House Sparrows

T levels were higher in T-implanted birds than in control birds (lme: $p = 0$, $t = 6.72$, $df = 32$, random effect: bird ID). One week after implanting T levels of C birds ranged between 189 and 4938 pg/ml with a mean of 1051 pg/ml. Thos of T birds ranged

between 3077 and 10760 pg/ml with a mean of 6722 pg/ml. The natural range of our population was between 110 and 7260 pg/ml in a previous summer (Laucht et al. 2010).

b) Queleas

T levels were higher in T-implanted birds than in control birds (lme: $p = 0$, $t = 11.87$, $df = 31$, random effect: bird ID). One week after implanting T levels of C birds ranged between 95 and 1732 pg/ml with a mean of 510 pg/ml. Those of T birds ranged between 248 and 11840 pg/ml with a mean of 7296 pg/ml. The natural range of our population was between 60 and 1900 pg/ml in a previous summer (unpublished data).

Relationship of bill colour with testosterone, plasma carotenoids, and immune function

a) House Sparrow

Bill colour was not related to T levels in an overall model (lme: after accounting for time since capture and body mass $t = 1.73$, $df = 65$, $p = 0.089$, random effect: bird ID), but positively related to T levels when looking at T and placebo-implanted birds separately (lme: after accounting for time since capture and body mass $t > 2.77$, $p < 0.009$, random effect: bird ID). A positive relation means higher T levels were correlated with less black bills and is in the opposite direction to the expected relationship. However, bill colour was not related to implant, LPS or PBS injection or the interaction of the two ($t < 1.53$, $p > 0.135$) indicating that the observed correlation between T and bill colour in summer (when all males have relatively dark bills) is very weak, and possibly spurious.

Bill colour was not correlated with total carotenoids after controlling for time (lme: $t = 0.36$, $df = 100$, $p = 0.722$). Bill colour did not change in relation to LPS challenge (lm: $t = -1.06$, $df = 32$, $p = 0.297$). Similarly, bill colour was not related to immune response (lmer: $z = -0.217$, $p = 0.828$, random effects: bird ID, type of immune measure and experiment action; after accounting for T levels).

Table 2: Mean + SE of different immune measures in the four treatment groups of a) House Sparrows and b) Red-billed Queleas. Presented are also the results of linear mixed effect models of these immune measures in relation to implant, LPS/PBS, or the interaction of both.
The used abbreviations are as following: C = control-implanted, T = T-implanted, L = LPS injected, P = PBS injected

HOSP

	mean ± SE				implant			LPS/PBS			interaction implant*LPS/PBS		
immune measure	**CL**	**TL**	**CP**	**TP**	**t**	**p**	**df**	**t**	**p**	**df**	**t**	**p**	**df**
haematocrit	0.531 ± 0.007	0.513 ± 0.007	0.538 ± 0.007	0.536 ± 0.007	-0.656	0.516	32	1.361	0.183	32	0.733	0.469	30
haptoglobin [mg/ml]	0.101 ± 0.008	0.108 ± 0.014	0.099 ± 0.023	0.102 ± 0.025	-1.782	0.095	15	-0.773	0.446	28	-0.028	0.978	26
lysozyme log[mg/l]	6.435 ± 0.295	6.407 ± 0.283	7.132 ± 0.392	6.315 ± 0.105	0.158	0.876	27	0.944	0.352	31	-1.389	0.176	29
lysis	2.1 ± 0.180	2.464 ± 0.170	2.6 ± 0.208	2.1 ± 0.067	-0.173	0.865	24	0.215	0.832	24	-2.484	**0.021**	22
agglutination	4.767 ± 0.483	4.967 ± 0.446	5.682 ± 0.365	5.423 ± 0.525	-0.189	0.852	30	0.86	0.397	30	-0.648	0.523	28
IgG	45.310 ± 1.658	48.164 ± 1.702	44.538 ± 2.170	44.945 ± 3.067	0.105	0.917	32	-0.891	0.379	32	-0.608	0.548	30
LPS antibodies	9.953 ± 1.248	12.735 ± 1.228	9.933 ± 0.791	8.457 ± 0.593	0.547	0.589	32	-1.535	0.135	32	-1.549	0.132	30

RBQU

	mean ± SE				implant			LPS/PBS			interaction implant*LPS/PBS		
immune measure	**CL**	**TL**	**CP**	**TP**	**t**	**p**	**df**	**t**	**p**	**df**	**t**	**p**	**df**
haematocrit	0.586 ± 0.123	0.544 ± 0.004	0.579 ± 0.009	0.575 ± 0.007	-1.841	0.068	132	0.717	0.479	31	1.641	0.112	29
haptoglobin [mg/ml]	0.110 ± 0.053	0.055 ± 0.013	0.100 ± 0.011	0.053 ± 0.014	-0.123	0.903	25	1.404	0.172	27	-0.469	0.643	25
lysozyme log[mg/l]	17.557 ± 2.048	17.072 ± 0.954	17.870 ± 1.976	15.398 ± 0.830	-0.14	0.89	32	-0.441	0.662	31	-0.622	0.539	29
lysis	2.455 ± 0.196	2.4 ± 0.256	2.444 ± 0.317	2.594 ± 0.139	0.354	0.726	25	0.475	0.639	25	0.568	0.575	23
agglutination	3.967 ± 0.404	3.467 ± 0.501	3.607 ± 0.454	5.188 ± 0.655	0.844	0.406	30	1.126	0.269	30	1.665	0.107	28
IgG	44.534 ± 1.613	45.706 ± 2.811	42.378 ± 1.918	44.023 ± 1.996	0.262	0.795	32	-0.655	0.518	31	0.08	0.937	29
LPS antibodies	8.730 ± 0.806	8.221 ± 0.836	9.997 ± 0.571	10.544 ± 0.553	-0.008	0.994	32	1.887	0.069	31	0.543	0.591	29

b) Queleas

Bill colour was not correlated with T levels in an overall model (lme: after accounting for time since capture and body mass $t = -1.44$, $df = 63$, $p = 0.156$, random effects: bird ID) nor when splitting the data into T and placebo implanted birds (lme: after accounting for time since capture and body mass $|t| < 1.05$, $p > 0.311$, random effects: bird ID). Bill colour was also not related to implant, LPS or PBS injection or the interaction of the two ($|t| < 1.50$, $p > 0.145$; using delta bill colour makes the interaction significant $p = 0.045$).

Bill colour was not correlated with total carotenoids after controlling for time (lme: $t = 1.03$, $df = 84$, $p = 0.308$). Bill colour did not change in relation to LPS challenge (lm: $t = 1.50$, $df = 31$, $p = 0.145$). Bill colour was not related to immune response (lmer: $z = -1.03$, $p = 0.305$, random effects: bird ID, type of immune measure and experiment action; after accounting for time since capture and T levels).

Prediction 1: testosterone is immunosuppressive

Table 2 provides details on the various immune responses measured in both species.

a) House Sparrows

Skin swelling after PHA injection was not correlated with implant or T level (after accounting for time since capture) (lm: $|t| < 1.37$, $p > 0.180$; Fig 2). Overall immune response was not related to T levels (lme: $t = 1.63$, $p = 0.103$, $df = 243$, random effect: bird ID; after accounting for time since capture and body mass; Fig 2), nor to the interaction of implant and LPS/PBS injection (lme: $t = -1.45$, $p = 0.157$, $df = 30$, random effect: bird ID; after accounting for time since capture and body mass).

b) Queleas

Skin swelling after PHA injection was not correlated with implant or T level (after accounting for time since capture) (lm: $|t| < 0.56$, $p > 0.577$; Fig 2). Overall immune response was not related to testosterone level (lme: $t = -0.63$, $df = 178$, $p = 0.529$, random effect: bird ID; after accounting for time since capture and body mass; Fig 2).

There was a significant marginal effect of the interaction of implant (testosterone or placebo) and LPS/PBS injection on immune response (lme: t = 2.01, df = 29, p = 0.044, random effect: bird ID; after accounting for time since capture and body mass).

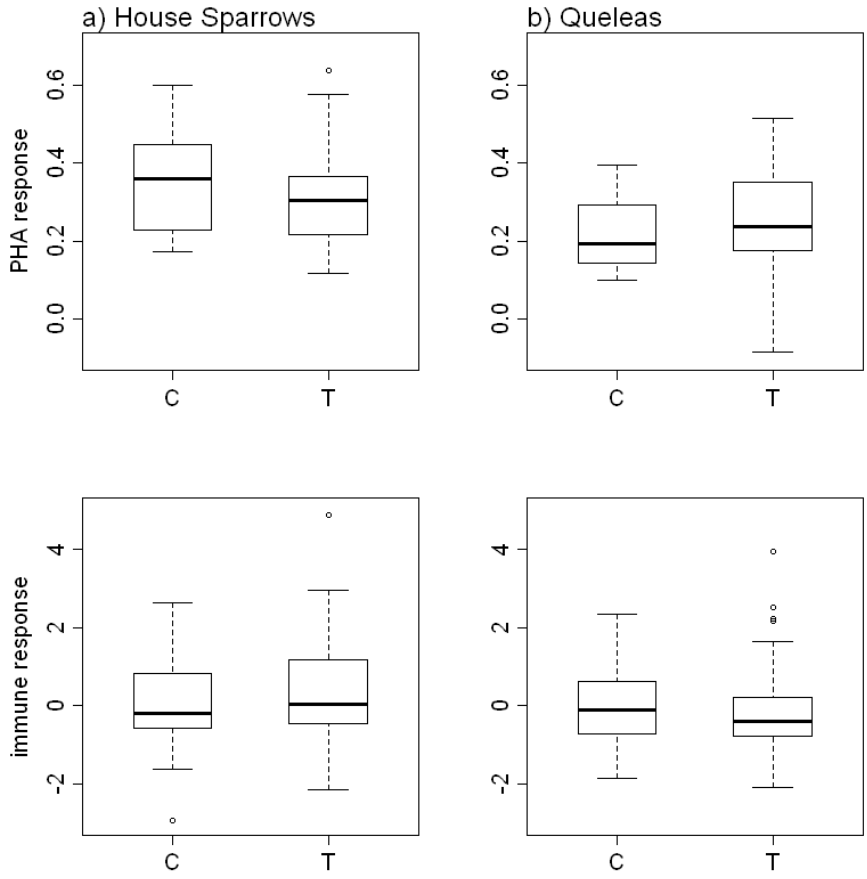


Fig. 2 Effects of implants (C = placebo, T = testosterone) on PHA response and on other immune measures in a) House Sparrows and b) Red-billed Queleas

Prediction 2: carotenoids are immunoenhancing

a) House Sparrows

Total carotenoids were not related to body mass, tarsus length or the interaction of both after accounting for time ($|t| < 0.60$, $p > 0.55$). PHA response was not related to total carotenoids (lm: after accounting for time since starting to capture birds: t = -1.09, df = 31, p = 0.282, Fig 3). Similarly, immune response was not related to total

carotenoids (lme: t = 0.43, df = 640, p = 0.67, random effect: bird ID; after accounting for type of immune measure, time since capture, and body mass, Fig 4).

Total carotenoids did not change between before and after PHA injection, neither overall nor when splitting the data into T and placebo implanted birds (paired t-tests: $|t| < 0.94$, $p > 0.36$). However, there was a significant decrease in total carotenoids in LPS injected birds (after controlling for time and implant) 24h after LPS injection (lme: t = -3.08, df = 50, p = 0.003, random effect: bird ID).

Food intake after PHA injection was lower than before PHA injection (paired t-test: t = 3.06, df = 33, p-value = 0.004). However, this could be only an effect of stress as food intake was lower when birds were stressed (t = -6.73, df = 237, p = 0). There was an influence of LPS on food intake within 24h after LPS/PBS injection (t = 4.57, df = 31, $p < 0.001$), but not of implant or the interaction of implant and LPS/PBS ($t < 0.83$, $p > 0.415$). This effect disappeared after one day.

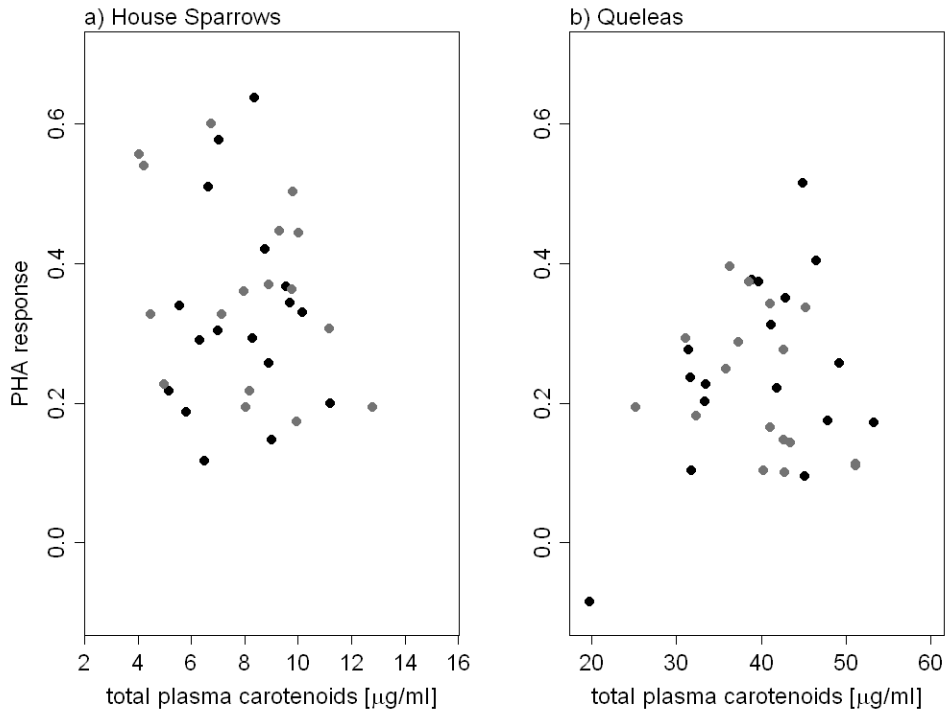


Fig 3: Relationship between total plasma carotenoids and PHA response in a) House Sparrows and b) Red-billed Queleas. Black dots represent T-implanted birds; gray dots represent placebo-implanted birds.

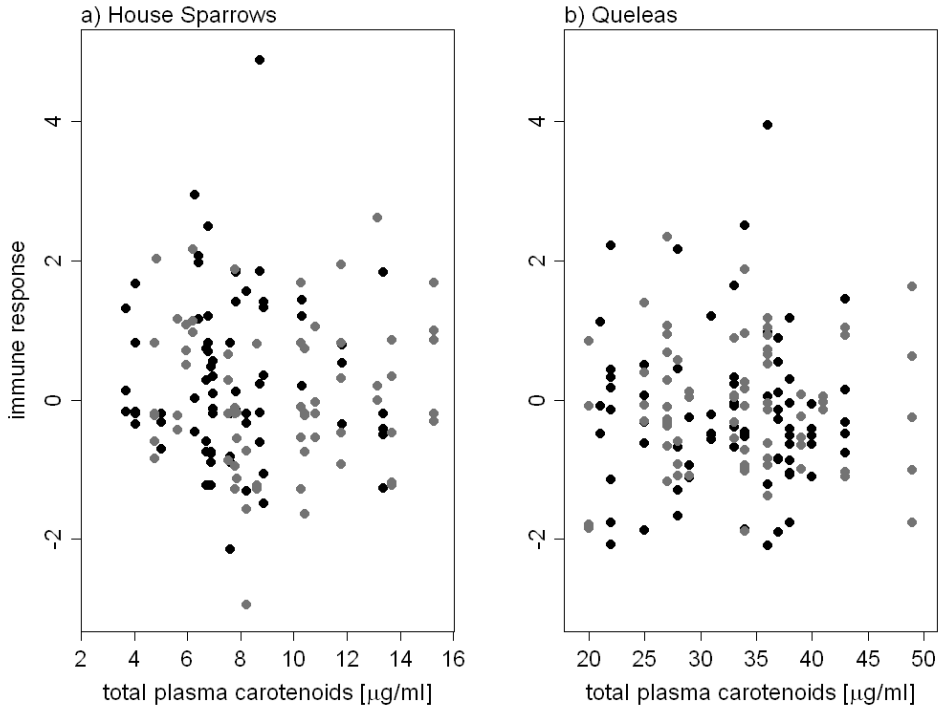


Fig 4: Relationship between total plasma carotenoids and immune response in a) House Sparrows and b) Red-billed Queleas. Black dots represent T-implanted birds; gray dots represent placebo-implanted birds.

b) Queleas

Total carotenoids were not related to body mass, tarsus length or the interaction of both after accounting for time ($|t| < 1.73$, $p > 0.093$). PHA response was not related to total carotenoids (lm: after accounting for time since starting to capture birds: $t = 0.78$, $df = 31$, $p = 0.441$, Fig 3). There was a trend for a relationship between immune response and total carotenoids (lme: $t = 1.66$, $df = 515$, $p = 0.097$, random effect: bird ID; after accounting for type of immune measure, time since capture, and body mass, Fig 4).

Total carotenoids were significantly lower after PHA injection than before, both, overall or when splitting the data into T and placebo implanted birds (paired t-tests: $t < 2.21$, $p < 0.043$). However, there was a significant increase in total carotenoids in LPS injected birds (after controlling for time and implant) after LPS injection (lme: $t = -3.08$, $df = 50$, $p = 0.003$, random effect: bird ID) that is due to the low levels after

PHA test and that does not hold when comparing carotenoids after LPS injection and at the start of the experiment ($t < 1.25$, $p > 0.240$).

Food intake after PHA injection was lower than before PHA injection (paired t-test: $t = 5.54$, $df = 33$, $p < 0.001$). However, this could be only an effect of stress as food intake was lower when birds were stressed ($t = -9.84$, $df = 334$, $p = 0$). There was an influence of LPS on food intake within 24h and 48h after LPS/PBS injection ($t > 2.40$, $df = 31$, $p < 0.023$), but not of implant or the interaction of implant and LPS/PBS ($t < 1.25$, $p > 0.221$). This effect disappeared after two days.

Prediction 3: T increases bioavailability of carotenoids

a) House Sparrows

There was a trend for a negative correlation between total carotenoids and T levels after accounting for time in an overall model (lme: $t = -1.98$, $df = 66$, $p = 0.052$, random effects: bird ID, Fig 5), but this did not hold when splitting data into T and placebo-implanted birds (lme: $|t| < 1.09$, $p > 0.21$, random effects: bird ID). Total carotenoids were also not related to implant, LPS/PBS injection or the interaction of both ($|t| < 1.49$, $p > 0.146$).

Immune measures were not correlated with the three-way-interaction of implant, LPS/PBS injection, and total carotenoids (lme: after accounting for type of measure, time since capture and body mass: $t = -0.96$, $df = 357$, $p = 0.336$; random effect: bird ID).

Food intake was not different in T than in C implanted birds (lme: $t = 0.75$, $df = 32$, $p = 0.460$, random effect: bird ID, Fig 5).

b) Queleas

There was a trend for a negative correlation between total carotenoids and T levels after accounting for time in an overall model (lme: $t = -1.71$, $df = 64$, $p = 0.092$, random effects: bird ID, Fig 5), but this did not hold when splitting data into T and placebo-implanted birds (lme: $|t| < 0.93$, $p > 0.360$, random effects: bird ID). Total

carotenoids were also not related to implant, LPS/PBS injection or the interaction of both ($|t| < 1.17$, $p > 0.244$).

Immune measures were not correlated with the three-way-interaction of implant, LPS/PBS injection, and total carotenoids (lme: after accounting for type of measure, time since capture and body mass: $t = 0.64$, $df = 296$, $p = 0.522$; random effect: bird ID).

Food intake was not different in T than in C implanted birds (lme: $t = 0.45$, $df = 32$, $p = 0.642$, random effect: bird ID, Fig 5).

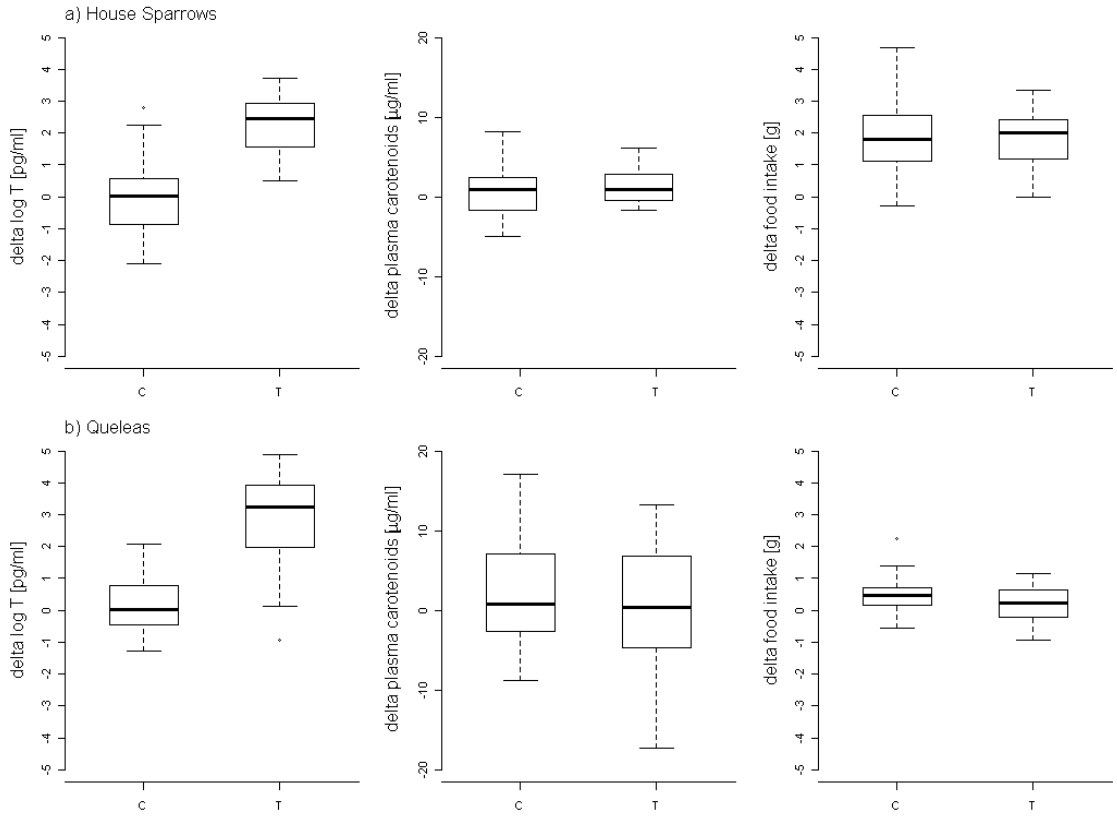


Fig 5: Effects of implants (C = placebo, T = testosterone) on testosterone levels, total plasma carotenoids, and food intake in a) House Sparrows and b) Red-billed Queleas.

Discussion

Our study, which is comprised of experimental T increases in two species (each with different mechanisms of ornament coloration), two forms of immune challenges, a

thorough arsenal of immunocompetence measures, and repeated measurements of ornament elaboration, plasma carotenoids and T levels, represents the most thorough test of the THM to date. In brief, we found no support for any of the predictions about carotenoid, testosterone, and immune system interrelationships underlying testosterone handicap models of signal honesty (Table 1). In particular, we first found no support for an immunosuppressive effect of T. Second, we found no support for an immunoenhancing effect of carotenoids. Third, and perhaps most surprisingly, we found absolutely no effect of increased T on the bioavailability of carotenoids in plasma. Therefore, we can reject all three predictions. We discuss the consequences of this below.

First, there was no suggestion that T was immunosuppressive in House Sparrows or Queleas (Table 1) implying that the proposed immunosuppressive effect of T is not general. Although this stands in contrast to many studies (e.g. Hasselquist et al. 1999; Duffy et al. 2000; Owen-Ashley et al. 2004) several others have also found that T is not immunosuppressive (Braude et al. 1999; Hasselquist et al. 1999; Westneat et al. 2003; Roberts et al. 2004; McGraw and Ardia 2007; Garvin et al. 2008). Taken together all this suggests that T is immunosuppressive only in certain circumstances, and that models that assume a general (i.e. constraining) immunosuppressive effect of T need to be reassessed. We suggest that when T is immunosuppressive the immunosuppression is best viewed as an adaptive response to elevated T levels rather than as a potential honesty enforcing constraint on T dependent ornaments – i.e. that the elevated T effectively functions as an internal signal that modulates investment into various physiological processes. In agreement with this hypothesis several other studies testing costs of elevated T levels have also found only minor modulations of physiological processes that are best seen as flexible adaptations rather than constraints. Along this line, the findings that suppression of one immune parameter is not necessarily an indication of an overall suppression of the immune system or a general trade-off with other physiological functions (Adamo 2004), and that resources within the immune system are rather redistributed (Braude et al. 1999; Martin et al. 2006c; Garvin et al. 2008) do not suggest that T is generally immunosuppressive but rather that these are special circumstances and signs of flexible adaptations. This is additionally supported by effects of environment (Buehler et al. 2009; Martínez-Padilla et al. 2010), captivity (Kuhlman and Martin 2010), or season (Martin et al.

2006a; Owen-Ashley and Wingfield 2006) that should not exist if T was generally immunosuppressive.

Second, our study provides little support for an immunoenhancing effect of carotenoids. Although we did observe decreases in carotenoid levels after immune challenges (Table 1), these decreases were rather small and they occurred inconsistently (i.e. after LPS injection in the House Sparrows and after PHA injection in the Queleas). For most of the tests we conducted, there was no relationship between carotenoid levels and immunocompetence. Although we did not test for an enhancement of the anti-oxidant capacity by carotenoids per se, it is generally argued that the immune system and anti-oxidant system are coupled (Alonso-Alvarez et al. 2008). Indeed, the immunoenhancing function of carotenoids is suggested to work via anti-oxidation processes because the activation of immune cells and killing of pathogens produces reactive oxygen species (ROS) that cause oxidative stress (Chew and Park 2004). Considering the widely assumed generality of the immunoenhancing properties of carotenoids, then immunoenhancement in both our study species was strongly predicted (e.g. Alonso-Alvarez et al. 2004; McGraw and Klasing 2006; Biard et al. 2009). However, we found no clear relationships between immune response and carotenoid levels, suggesting they do not function generally as immunoenhancers and/or anti-oxidants (Pérez-Rodríguez 2009).

One reason why carotenoids did not function as immunoenhancers in this study is that carotenoid availability might not be particularly limited in our two species, and so no trade-offs occur between immunoenhancement/anti-oxidant function and other functions (such as signalling). Indeed, this could be the case in Queleas, where Dale (2000) demonstrated that a high degree of intraspecific variation in carotenoid based plumage coloration was not related to male quality. Therefore it seems likely that the carotenoids required for plumage in Queleas are not limited (note that this is not necessarily the case for Quelea bill coloration, which probably requires considerably higher amounts of carotenoids than plumage). However, this situation is improbable in House Sparrows, because plasma carotenoid levels are low compared to other species (e.g. ~25% to 50% of levels in Zebra Finches, *Taeniopygia guttata* (McGraw et al. 2006), or Red-legged Partridges, *Alectoris rufa* (Blas et al. 2006), respectively).

Third, there was no indication that T is involved in the modulation of plasma carotenoids in our two species. This result contrasts strongly with patterns in other species where there are clear increases of carotenoid bioavailability in response to T augmentation (e.g. Blas et al. 2006; McGraw et al. 2006; McGraw and Ardia 2007; Alonso-Alvarez et al. 2008). Because our T implants caused considerably higher T levels, they should have had large effects on overall metabolism and increased oxidative stress. If T is generally associated with increased oxidative stress, and decreased immune function, then we should have observed a change in carotenoids as a function of T. Given that carotenoids did not have an immunoenhancing effect in our study species, that T was not immunosuppressive, and that in House Sparrows there is no carotenoid based ornament and in Queleas the carotenoid ornament is not T dependent, we propose that there is no need to increase the bioavailability of carotenoids in response to elevated T in these two species. We conclude that other species demonstrate an increase in plasma carotenoids in response to T augmentation because of a need to modulate the elaboration of T dependent carotenoid based ornaments, rather than to negotiate immunosuppression/oxidative stress per se.

A fundamental question that is raised by our study is why are House Sparrows and Red-billed Queleas different from other species that have been looked at so far? In particular, Zebra Finches and Red-legged Partridges provide strong contrasts because in these species, much clearer links occur between T levels, carotenoids, the immune/anti-oxidation systems and ornamentation. This question represents a considerable challenge for future research to address. However, in general, we contend that interspecific differences in the physiological processes that underlie ornament development will be the outcome of selection for different mechanisms that facilitate and/or optimize signalling different kinds of information. That is, we view different mechanisms as the product of differences in the evolved function of the ornament. This view contrasts strongly with much of the contemporary research on ornamentation, where mechanisms are instead often used to infer function.

In order to understand the complexities of any mechanisms of signal development, we need a thorough understanding of the social function of the signal. Integrative biology research typically assumes that ornaments function as costly sexually selected indicators of overall genetic and phenotypic quality directed towards females during

mate choice. However, support for this assumption is rarely provided and alternative signalling functions are rarely considered. Indeed, in House Sparrows it is unlikely that variation in bill coloration is used in mate choice, given that during the breeding season there is virtually no variance in its expression (Laucht et al. 2010). Given its strong seasonal dependency and threshold relationship with T, it is more likely that bill colour signals something to do with different behavioural strategies (Dale 2006) associated with being in non-breeding versus breeding condition (such as aggression, Laucht et al. 2010). In Queleas, it is also unlikely that variance in bill coloration is particularly important during mate choice given that females also develop red bills only during the *non-breeding* season, when competition for limited food resources is very high. Even Zebra Finches and Red-legged Partridges have mating systems where sexual selection is not expected to be particularly strong: monogamy with long-term pair bonds. In Zebra Finches, demonstrating a strong mate choice for redder bills (let alone a benefit to females for having such preferences) has been notoriously difficult (Sullivan 1994; Balzer and Williams 1998; Forstmeier and Birkhead 2004) in comparison to other systems. In Red-legged Partridges, we are unaware of any research directed towards understanding the social contexts and decision making influences their carotenoid based ornaments are used for. Considering the current state of knowledge, we argue that it is just as likely that these ornaments signal aggression and dominance to competitive rivals as it is that they signal quality of the immune/anti-oxidation system to potential mates. Indeed, in Zebra Finches, Ardia et al. (2010) have recently demonstrated that a short term exposure to T produces a rapid change in bill colour and dominance. Moreover, the ornaments of Red-legged Partridges are deliberately displayed to rivals during aggressive interactions (Cramp and Simmons 1977-1994) suggesting a prominent role for them as signals used during competitive interactions.

In summary, none of the three general predictions was supported by our study. Since the assumptions of the THM need to be of a general nature (i.e. provide unavoidable costs of ornamentation), we expected them to apply to all species. However, our results demonstrate rather that these three assumptions are species and/or situation specific. As such they suggest that they are evolutionarily dynamic, and therefore avoidable as honesty-enforcing costs of ornamentation (i.e. in principle, easily cheatable). Our study provides a broader window into alternative cost-bearing

mechanisms of honest signalling systems. One possibility is enforcement via social costs according to the "badge of status hypothesis" (Maynard Smith and Harper 1988; Jawor and Breitwisch 2003; Tibbetts and Dale 2004). Unlike the THM, social costs for ornamentation do not require cheatable physiological constraints to enforce signal honesty, but instead argue that physiological mechanisms function as fine-tuned adaptations which facilitate optimal signalling.

Acknowledgements

Many thanks to Rebecca Wutzler, Julia Wickjürgen, Maryjane Heckel, Magnus Ruhdorfer, and Edith Bodendorfer for their help with experimental procedures and animal care. We are grateful to Ingrid Schwabl, Monika Trappschuh, Evi Fricke, Sandra Chiriac, and Christine Hermannstätter for their help in the lab, to Lisa Trost for her help with obtaining the permit for the experiment, to Barbara Helm, Helga Gwinner, and Albertine Leitao for teaching experimental procedures, and to Mihai Valcu and Wolfgang Forstmeier for statistical advice. We thank the Behavioural Ecology and Evolutionary Genetics Group at the MPIO for discussion and comments. Funding for this research was provided by the Max Planck Society. The study was performed in agreement with the current laws of Upper Bavaria, Germany (under licenses: permit nr. 55.1-8642.3-3-2006 with several extensions; and permit nr: 55.2-1-54-2531-118-07 of the "Regierung von Oberbayern").

References

Adamo SA (2004) How should behavioural ecologists interpret measurements of immunity? Anim. Behav. 68:1443-1449

Adkins-Regan E (2005) Hormones and Animal Social Behavior. Princeton University Press, Princeton, NJ

Alonso-Alvarez C, Bertrand S, Devevey G, Gaillard M, Prost J, Faivre B, Sorci G (2004) An experimental test of the dose-dependent effect of carotenoids and immune activation on sexual signals and antioxidant activity. Am. Nat. 164:651-659

Alonso-Alvarez C, Bertrand S, Faivre B, Chastel O, Sorci G (2007) Testosterone and oxidative stress: the oxidation handicap hypothesis. Proc. R. Soc. Biol. Sci. Ser. B 274:819-825

Alonso-Alvarez C, Pérez-Rodríguez L, Mateo R, Chastel O, Viñuela J (2008) The oxidation handicap hypothesis and the carotenoid allocation trade-off. J. Evol. Biol. 21:1789-1797

Anderson TR (2006) Biology of the ubiquitious house sparrow. Oxford University Press, New York

Ardia DR, Broughton DR, Gleicher MJ (2010) Short-term exposure to testosterone propionate leads to rapid bill color and dominance changes in zebra finches. Horm. Behav. 58:526-532

Balzer AL, Williams TD (1998) Do female zebra finches vary primary reproductive effort in relation to mate attractiveness? Behaviour 135:297-309

Bertrand S, Faivre B, Sorci G (2006) Do carotenoid-based sexual traits signal the availability of non-pigmentary antioxidants? J. Exp. Biol. 209:4414-4419

Biard C, Hardy C, Motreuil S, Moreau J (2009) Dynamics of PHA-induced immune response and plasma carotenoids in birds: should we have a closer look? J. Exp. Biol. 212:1336-1343

Blas J, Pérez-Rodríguez L, Bortolotti GR, Vinuela J, Marchant TA (2006) Testosterone increases bioavailability of carotenoids: Insights into the honesty of sexual signaling. Proc. Natl. Acad. Sci. U. S. A. 103:18633-18637

Bonneaud C, Mazuc J, Gonzalez G, Haussy C, Chastel O, Faivre B, Sorci G (2003) Assessing the cost of mounting an immune response. Am. Nat. 161:367-379

Braude S, Tang-Martinez Z, Taylor GT (1999) Stress, testosterone, and the immunoredistribution hypothesis. Behav. Ecol. 10:345-350

Buehler DM, Tieleman BI, Piersma T (2009) Age and environment affect constitutive immune function in Red Knots (*Calidris canutus*). J. Ornithol. 150:815-825

Chew BP, Park JS (2004) Carotenoid action on the immune response. J. Nutr. 134:257s-261s

Cramp S, Simmons KEL (1977-1994) Handbook of the birds of Europe, the Middle East and North Africa : the birds of the western Palearctic In. Oxford University Press, Oxford, UK

Dale J (2000) Ornamental plumage does not signal male quality in red-billed queleas. Proc. R. Soc. Biol. Sci. Ser. B 267:2143-2149

Dale J (2006) Intraspecific variation in coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration Harvard University Press, MA, Cambridge, pp 36-86

Duffy DL, Bentley GE, Drazen DL, Ball GF (2000) Effects of testosterone on cell-mediated and humoral immunity in non-breeding adult European starlings. Behav. Ecol. 11:654-662

Faivre B, Grégoire A, Préault M, Cézilly F, Sorci G (2003) Immune activation rapidly mirrored in a secondary sexual trait. Science 300:103-103

Folstad I, Karter AJ (1992) Parasites, Bright Males, and the Immunocompetence Handicap. Am. Nat. 139:603-622

Forstmeier W, Birkhead TR (2004) Repeatability of mate choice in the zebra finch: consistency within and between females. Anim. Behav. 68:1017-1028

Fusani L (2008) Endocrinology in field studies: Problems and solutions for the experimental design. Gen. Comp. Endocrinol. 157:249-253

Garvin JC, Dunn PO, Whittingham LA, Steeber DA, Hasselquist D (2008) Do male ornaments signal immunity in the common yellowthroat? Behav. Ecol. 19:54-60

Goymann W, Mostl E, Gwinner E (2002) Non-invasive methods to measure androgen metabolites in excrements of European stonechats, *Saxicola torquata rubicola*. Gen. Comp. Endocrinol. 129:80-87

Goymann W, Trappschuh M, Jensen W, Schwabl I (2006) Low ambient temperature increases food intake and dropping production, leading to incorrect estimates of hormone metabolite concentrations in European stonechats. Horm. Behav. 49:644-653

Grafen A (1990) Biological Signals as Handicaps. J. Theor. Biol. 144:517-546

Hasselquist D, Marsh JA, Sherman PW, Wingfield JC (1999) Is avian humoral immunocompetence suppressed by testosterone? Behav. Ecol. Sociobiol. 45:167-175

Jawor JM, Breitwisch R (2003) Melanin ornaments, honesty, and sexual selection. Auk 120:249-265

Johnstone RA, Norris K (1993) Badges of Status and the Cost of Aggression. Behav. Ecol. Sociobiol. 32:127-134

Kuhlman JR, Martin LB (2010) Captivity affects immune redistribution to skin in a wild bird. Funct. Ecol. 24:830-837

Laucht S, Kempenaers B, Dale J (2010) Bill color, not badge size, indicates testosterone-related information in house sparrows. Behav. Ecol. Sociobiol. 64:1461-1471

Lessells CM, Boag PT (1987) Unrepeatable Repeatabilities - a Common Mistake. Auk 104:116-121

Martin LB, Han P, Kwong J, Hau M (2006a) Cutaneous immune activity varies with physiological state in female house sparrows (*Passer domesticus*). Physiol. Biochem. Zool. 79:775-783

Martin LB, Han P, Lewittes J, Kuhlman JR, Klasing KC, Wikelski M (2006b) Phytohemagglutinin-induced skin swelling in birds: histological support for a classic immunoecological technique. Funct. Ecol. 20:290-299

Martin LB, Scheuerlein A, Wikelski M (2003) Immune activity elevates energy expenditure of house sparrows: a link between direct and indirect costs? Proc. R. Soc. Biol. Sci. Ser. B 270:153-158

Martin LB, Weil ZM, Nelson RJ (2006c) Refining approaches and diversifying directions in ecoimmunology. Integr. Comp. Biol. 46:1030-1039

Martínez-Padilla J, Mougeot F, Webster LMI, Pérez-Rodríguez L, Piertney SB (2010) Testing the interactive effects of testosterone and parasites on carotenoid-based ornamentation in a wild bird. J. Evol. Biol. 23:902-913

Matson KD, Ricklefs RE, Klasing KC (2005) A hemolysis-hemagglutination assay for characterizing constitutive innate humoral immunity in wild and domestic birds. Dev. Comp. Immunol. 29:275-286

Maynard Smith JFRS, Harper DGC (1988) The Evolution of Aggression - Can Selection Generate Variability? Philos. Trans. R. Soc. Lond. B Biol. Sci. 319:557-570

McGraw KJ, Ardia DR (2003) Carotenoids, immunocompetence, and the information content of sexual colors: An experimental test. Am. Nat. 162:704-712

McGraw KJ, Ardia DR (2007) Do carotenoids buffer testosterone-induced immunosuppression? An experimental test in a colourful songbird. Biol. Letters 3:375-378

McGraw KJ, Correa SM, Adkins-Regan E (2006) Testosterone upregulates lipoprotein status to control sexual attractiveness in a colorful songbird. Behav. Ecol. Sociobiol. 60:117-122

McGraw KJ, Klasing KC (2006) Carotenoids, immunity, and integumentary coloration in red junglefowl (*Gallus gallus*). Auk 123:1161-1171

McGraw KJ, Tourville EA, Butler MW (2008) A quantitative comparison of the commonly used methods for extracting carotenoids from avian plasma. Behav. Ecol. Sociobiol. 62:1991-2002

Millet S, Bennett J, Lee KA, Hau M, Klasing KC (2007) Quantifying and comparing constitutive immunity across avian species. Dev. Comp. Immunol. 31:188-201

Møller AP, Erritzøe J (2000) Predation against birds with low immunocompetence. Oecologia 122:500-504

Mougeot F, Irvine JR, Seivwright L, Redpath SM, Piertney S (2004) Testosterone, immunocompetence, and honest sexual signaling in male red grouse. Behav. Ecol. 15:930-937

Mougeot F, Martínez-Padilla J, Webster LMI, Blount JD, Pérez-Rodriguez L, Piertney SB (2009) Honest sexual signalling mediated by parasite and testosterone effects on oxidative balance. Proc. R. Soc. Biol. Sci. Ser. B 276:1093-1100

Mougeot F, Pérez-Rodríguez L, Martínez-Padilla J, Leckie F, Redpath SM (2007) Parasites, testosterone and honest carotenoid-based signalling of health. Funct. Ecol. 21:886-898

Owen-Ashley NT, Hasselquist D, Wingfield JC (2004) Androgens and the immunocompetence handicap hypothesis: Unraveling direct and indirect pathways of immunosuppression in song sparrows. Am. Nat. 164:490-505

Owen-Ashley NT, Turner M, Hahn TP, Wingfield JC (2006) Hormonal, behavioral, and thermoregulatory responses to bacterial lipopolysaccharide in captive and free-living white-crowned sparrows (Zonotrichia leucophrys gambelii). Horm. Behav. 49:15-29

Owen-Ashley NT, Wingfield JC (2006) Seasonal modulation of sickness behavior in free-living northwestern song sparrows (Melospiza melodia morphna). J. Exp. Biol. 209:3062-3070

Owens IPF, Short RV (1995) Hormonal Basis of Sexual Dimorphism in Birds - Implications for New Theories of Sexual Selection. Trends Ecol. Evol. 10:44-47

Pérez-Rodríguez L (2009) Carotenoids in evolutionary ecology: re-evaluating the antioxidant role. BioEssays 31:1116-1126

Peters A (2007) Testosterone and carotenoids: an integrated view of trade-offs between immunity and sexual signalling. BioEssays 29:427-430

Poiani A, Goldsmith AR, Evans MR (2000) Ectoparasites of house sparrows (Passer domesticus): an experimental test of the immunocompetence handicap hypothesis and a new model. Behav. Ecol. Sociobiol. 47:230-242

R Development Core Team (2006) R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria

R Development Core Team (2010) R: A language and environment for statistical computing. In. R Foundation for Statistical Computing, Vienna, Austria

Roberts ML, Buchanan KL, Evans MR (2004) Testing the immunocompetence handicap hypothesis: a review of the evidence. Anim. Behav. 68:227-239

Roberts ML, Buchanan KL, Hasselquist D, Evans MR (2007) Effects of testosterone and corticosterone on immunocompetence in the zebra finch. Horm. Behav. 51:126-134

Shawcross JE, Slater PJB (1984) Agonistic Experience and Individual Recognition in Male *Quelea quelea*. Behav. Process. 9:49-60

Smits JE, Bortolotti GR, Tella JL (1999) Simplifying the phytohaemagglutinin skin-testing technique in studies of avian immunocompetence. Funct. Ecol. 13:567-572

Sullivan MS (1994) Discrimination among males by female zebra finches based on past as well as current phenotype. Ethology 96:97-104

Tibbetts EA, Dale J (2004) A socially enforced signal of quality in a paper wasp. Nature 432:218-222

Verhulst S, Dieleman SJ, Parmentier HK (1999) A tradeoff between immunocompetence and sexual ornamentation in domestic fowl. Proc. Natl. Acad. Sci. U. S. A. 96:4478-4481

Vinkler M, Albrecht T (2010) Carotenoid maintenance handicap and the physiology of carotenoid-based signalisation of health. Naturwissenschaften 97:19-28

von Schantz T, Bensch S, Grahn M, Hasselquist D, Wittzell H (1999) Good genes, oxidative stress and condition-dependent sexual signals. Proc. R. Soc. Biol. Sci. Ser. B 266:1-12

Wedekind C, Folstad I (1994) Adaptive or Nonadaptive Immunosuppression by Sex-Hormones. Am. Nat. 143:936-938

Westneat DF, Hasselquist D, Wingfield JC (2003) Tests of association between the humoral immune response of red-winged blackbirds (*Agelaius phoeniceus*) and

male plumage, testosterone, or reproductive success. Behav. Ecol. Sociobiol. 53:315-323

Witschi E (1961) Sex and secondary sexual characters. In: Marshall AJ (ed) Biology and Comparative Physiology of Birds. Academic Press, New York and London, pp 115-168

Witschi E, Woods RP (1936) The bill of the sparrow as an indicator for the male sex hormone. II. structural basis. J. Exp. Zool. 73:445-459

Zahavi A (1975) Mate Selection - Selection for a Handicap. J. Theor. Biol. 53:205-214

Zahavi A (1977) Cost of Honesty - (Further Remarks on Handicap Principle). J. Theor. Biol. 67:603-605

General Discussion

I have investigated a number of critical questions and assumptions of theories to explain the honesty of ornamental signals of quality. I have used the House Sparrow as a model, and so these results have important implications for our understanding of the ecology and ornamentation of this species. However House Sparrows also provide a generalized model system which can be extended to other species and other types of ornaments. The key conclusions that can be drawn from this work are as follows.

In chapters one and two I have demonstrated that House Sparrow plasma testosterone levels vary considerably in the course of the year. Correlations between different seasons were absent or very weak. Testosterone also varied considerably over the course of the day, with night time values being consistently higher than day time values. I hypothesized that night time testosterone levels approach similar levels as an individual's potential maximal levels. I demonstrated that bill colour is strongly correlated to current testosterone levels, and I argued that bill colour has strong potential to function as a short-term signal of testosterone activated behaviours and strategies. In contrast, I found that badge size was not related to day time testosterone levels, but was positively correlated with night time levels. Badge size is therefore a

potential signal of maximal testosterone levels and related behaviours such as aggression.

In chapter three I have investigated in detail some of the aspects of quality related information that badge size, bill colour, and wingbar area potentially signal. I concluded that these traits can possibly function as signals of multiple messages of different testosterone related behaviours. I also provided, for the first time, evidence for the potential signalling function of a previously unrecognized ornament –leg colour. Leg colour appears to be strongly related to age and contrasts with badge size and wingbar area which are also correlated with age but in an interactive fashion with other physiological parameters.

In the last two chapters I used an experimental approach to investigate the mechanisms of ornament development by manipulating the social environment and physiology. In chapter four I demonstrated that the development of badge size is largely independent of group composition and thus of social environment during moult, although in one year there was an apparent effect of the inter-male variability in badge size on later badge development. While my experiment yielded complex results, overall the patterns suggested the existence of a "programming window" for badge size already before the onset of moult.

In chapter five I investigated the generalizability of three key assumptions of the "Testosterone Handicap Model" to explain the honesty of carotenoid based ornamental quality signals: (1) testosterone is immunosuppressive, (2) carotenoids are immunoenhancing, and (3) testosterone increases the bioavailability of carotenoids. Because of the very general nature of these assumptions, they provide inescapable physiological constraints that should occur independently of the carotenoid or testosterone dependence of the ornament. Therefore, I tested them in one species with testosterone dependent carotenoid independent ornaments, the House Sparrow, and in one species with testosterone independent carotenoid dependent ornaments, the Red-billed Quelea. I found that these assumptions were not supported. These challenging findings suggest that other explanations of signal honesty may therefore be more likely.

Variation in plasma testosterone levels

Seasonal and probably also diel variation in hormone levels reflect the need and usage of these hormones to regulate different life history stages and physiology. Indeed, and in agreement with other studies and other species (Dawson 1983; Wingfield et al. 1990; Cockrem and Seddon 1994; Kellam et al. 2004; Anderson 2006; Jawor 2007), I found that House Sparrow plasma testosterone levels vary considerably in the course of the year and between seasons (chapter one). In addition, plasma testosterone levels also differ between day and night (chapter two) and therefore fluctuate considerably over the 24 hour period as has been suggested for a variety of other species (see Table 1 in chapter two). Because high levels of hormones, especially testosterone, are linked with numerous physiological and behavioural costs (reviewed in Wingfield et al. 2001) hormone levels cannot be kept at maximum levels at all times. The more efficient strategy is to flexibly adapt release rates. This adaptation needs to be so plastic and quick that hormone levels can be modulated to stochastic short-term changes of the environment (including social environment) such as has been suggested for testosterone during agonistic interactions (Wingfield et al. 1990).

As an important consequence, the high flexibility of release rates makes it extremely difficult to accurately measure (baseline) hormone levels. Nevertheless, the common approach in behavioural ecology is to take one blood sample at one time point. There are several good reasons for this: small animals cannot be sampled multiply in short time intervals, and wild animals often cannot be recaptured. Furthermore, it is commonly assumed that despite all this variation hormone levels at different time points are still correlated with each other and repeatable (Romero and Reed 2008). However, my research shows that day and night testosterone levels of the same individuals measured a few days apart were only weakly correlated (chapter two) invalidating the latter assumption in House Sparrows. In addition, when comparing plasma testosterone levels across the four different seasons, I found only low correlations (chapter one). This strongly suggests that testosterone levels at different seasons are either not correlated, or that environmental conditions mask these correlations.

Because I used captive birds in the same aviary set-up over all four seasons I greatly reduced the possibility that environmental conditions differently influenced individuals. It is therefore more likely that individuals pursue different strategies concerning baseline and elevated testosterone levels in one season that are rather independent from strategies in a different season. However, seasonal average or maximum levels could still be correlated. Correlations of average testosterone levels in our House Sparrows population are likely because male bill colour is correlated between seasons (chapter one), and bill colour reflects an approximately 3.5-week running average of testosterone levels. To test for a correlation between maximal testosterone levels, GnRH (gonadotropin releasing hormone) injections that cause a maximal release of testosterone would be useful (Wingfield and Farner 1993; Jawor et al. 2006). GnRH induced maximal testosterone levels were highly repeatable in Dark-eyed Juncos (*Junco hyemalis*) whereas unchallenged testosterone levels were not (Jawor et al. 2006). Taken together, this suggests that point samples of testosterone levels have only very limited value for the interpretation of running hormone levels. Other methods such as GnRH induced testosterone levels might be more useful.

The flexible release rates of testosterone and inconsistencies between different time points pose challenges for theories about testosterone dependent ornaments. Testosterone Handicap Models of signal honesty (see introduction) have assumed that testosterone levels - and hence honesty enforcing costs - are correlated between seasons. More specifically, how can an ornament be developed via a testosterone dependent mechanism in one season and function as an honest signal of testosterone related traits (such as immunocompetence) in a different season? Above all, when ornament elaboration takes place, testosterone levels are minimal (annual moult of plumage) (Wingfield et al. 1990; Hahn et al. 1992), even barely measurable in some species, and thus at this time putative honesty-enforcing costs linked with testosterone levels may be low or absent. Therefore the required assumption of Testosterone Handicap Models does not hold (see above) and the link between ornaments, signal honesty and testosterone clearly needs revision.

Ornaments in House Sparrows

House Sparrows have a number of ornaments. In our captive population, I demonstrated that badge size, bill colour, wingbar size and leg colour could each potentially signal a different aspect of quality. These qualities include testosterone related characters, body condition and age (all chapters, mainly chapter three). Synthesizing my results with those of other studies (e.g. Møller 1987; Bókony et al. 2006; Nakagawa and Burke 2008) we can now argue that male House Sparrows have at least two potential signals of age (the black area around the eye and leg colour) and three potential signals of fighting ability, testosterone related behaviours and strategies (badge size, bill colour, and wingbar area) (Fig. 1). However, the transition between signals of age and of testosterone related information is continuous, because some of the testosterone related ornaments (badge size and wingbar area) were also related to age, and because condition had an additional influence on some ornaments (badge size and wingbar area). My data suggest that ornaments in general are correlated with multiple aspects of quality, with different ornaments having slightly different weightings for each aspect. This makes sense, given the overall expectation of these traits to incur costs. However, more research on receiver behaviour in response to variation in these putative signals is necessary.

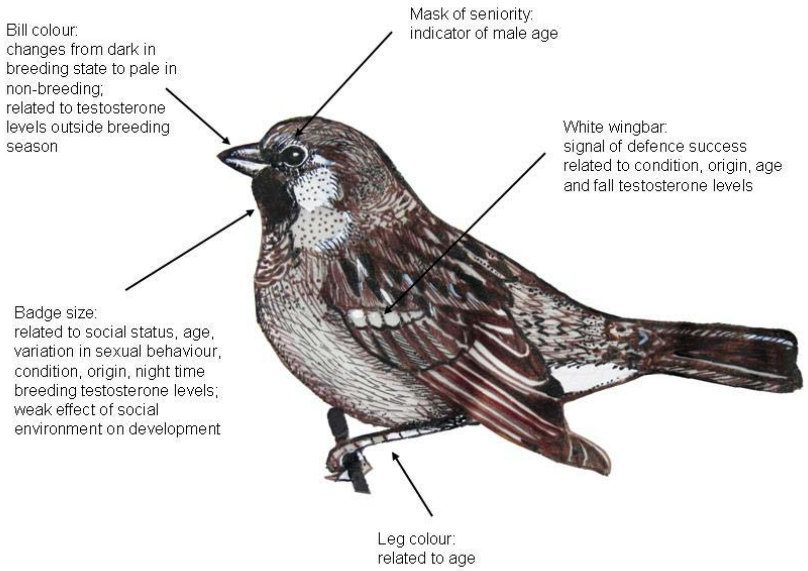


Fig. 1: Ornaments of male House Sparrows. Indicated are all known ornaments and their general function including the results of my dissertation. This figure is adapted from Fig. 2 of the introduction.
Drawing by Helga Gwinner.

In general, signals of age could be important in cases where young adults pursue different strategies than older birds, i.e. after the second adult moult (Lyon and Montgomerie 1986; Karubian 2002; Hatch and Westneat 2007; Karubian 2008). Different strategies could simply evolve when yearlings are less experienced and not able to defend themselves, their territories, or their mates as well as older birds, or when they are of inferior condition. This has been studied extensively in the context of delayed plumage maturation, whose main function seems to be to reduce aggression of adults towards yearlings (summarized in Dale 2006). House Sparrows do not have delayed plumage maturation per se, but they do show several continuous age related traits. All ornaments combined could provide receivers with even more detailed information on age related status. This could be essential in large feeding flocks (Ketterson 1979; Porter and Sealy 1982; Lyon and Montgomerie 1986). In addition, age related ranks and strategies could be of importance during breeding given coloniality, no territoriality, multiple broods, facultative extra-pair matings, and facultative polygyny (Anderson 2006).

In House Sparrows, badge size and wingbar conspicuousness potentially signal different aspects of fighting ability (Bókony et al. 2006). Badge size and bill colour also signal different relationships with testosterone levels (maximal and average, respectively; chapter one and two). In addition, bill colour is a dynamic ornament that can change within a few days to weeks and flexibly signal current states, whereas badge size and wingbar conspicuousness are fixed ornaments that signal the same information all year round. Bill coloration most importantly is strongly related to breeding condition, and therefore presumably signals aspects of behaviours related to being in breeding state or not.

Diverse signals of testosterone related behaviours and strategies, such as aggression and fighting abilities, are important to reduce serious agonistic interactions. For example, if two contestants differ largely in their underlying abilities, both participants would be better off to avoid the violent conflict. The assessment of testosterone dependent ornaments could reduce conflicts and minimize escalation because the inferior individual can quickly assess that he has no chance in winning and avoid injury or even death (Maynard Smith and Harper 1988; Jawor and

Breitwisch 2003; Tibbetts and Safran 2009). Actual fights could then be reduced to situations where differences in ornaments are too small to be assessed at first glance.

I have concentrated here on ornamentation in the context of male-male competition which is only one aspect of sexual selection and signals of quality. Another important aspect of sexual selection and signals of quality is mate choice. Especially the badge has been widely studied in this context. It was found that males with larger badges were more engaged in communal displays and both forced and unforced extra-pair and within-pair copulations (Møller 1990), whereas males with smaller badges invested more in their current broods (Griffith 2000), were preferred as social fathers (Griffith et al. 1999) and were more likely to be polygynous (Griffith et al. 1999). However, being cuckolded was independent of badge size (Cordero et al. 1999; Whitekiller et al. 2000; Stewart et al. 2006). In our population we have experimentally set up small breeding colonies, and have found that males with smaller badges were more likely to be polygynous and to gain extra-pair paternity (trend, S. Laucht, unpublished data). Because findings between different studies and thus populations are inconsistent and because the badge is so far the only male ornament of House Sparrows studied in the context of mate choice it would be fruitful to examine more ornaments in the context of mate choice and to move from signals of status to signals of other qualities.

Broadening the scope: other species, other ornaments

In the last chapter of my thesis I broadened the scope and moved from ornamentation and signal honesty in House Sparrows as a model system to other species, and I addressed the assumed generality of theories explaining signal honesty. This study also represents a step from melanin based ornaments to carotenoid based ornaments. More specifically, I have tested several general assumptions of the Testosterone Handicap Model and its further developments to explain the honesty of testosterone dependent carotenoid based signals of quality. Because of the very general nature of these assumptions, I expected them to provide inescapable physiological constraints that should occur independently of the carotenoid or testosterone dependence of the ornaments. However, I found that none of them applied for House Sparrows (melanin based testosterone dependent ornament) and Red-billed Queleas (carotenoid based

testosterone independent ornament). I showed that testosterone was not immunosuppressive, that carotenoids were not immunoenhancing, and that testosterone did not increase the bioavailability of carotenoids. Taken together, this suggests that these generally accepted assumptions about signalling honesty might not be as general but rather species and/or situation specific (see discussion of chapter five). The suggested physiological trade-offs therefore seem to be evolutionarily dynamic and for this reason unsuitable as honesty-enforcing costs of ornamentation (see discussion of chapter five).

My results represent a challenge to the results and interpretation of other studies that have suggested support for some or all of these assumptions (see chapter five) and more importantly a challenge of current theories that posit a testosterone based cost to honest ornaments. However, they also demonstrate that we are far from understanding signal honesty in animals, and that it is important to develop theories that can be tested in a variety of species and circumstances to prove their generality.

If high testosterone levels do not cause costs via inescapable trade-offs but rather adaptive physiological changes, the question what keeps signals honest remains. In the House Sparrow it is very likely that the honesty of testosterone related ornaments can be enforced via social costs of testosterone related behaviour via challenges by and defence against conspecifics (all chapters). A main argument for this theory (known as the Badges of Status Hypothesis, see general introduction) is that there is no possibility to cheat, hence it is evolutionarily stable. Furthermore, it is independent of mechanisms underlying ornament development which makes it applicable to a wide array of ornaments, assuring its generality. The hypothesis is also in agreement with recent suggestions about the similarity of carotenoids and melanins (Griffith et al. 2006). It has been suggested as an explanation for signal honesty in several other species (Rohwer 1975; McGraw 2004; Tibbetts and Safran 2009) and might also be true for species that were studied in the context of the Immunocompetence Handicap Hypothesis such as Zebra Finches (*Taeniopygia guttata*) or Red-legged Partridges (*Alectoris rufa*) (see discussion of chapter five). However, the examination of social costs and a critical test of the badges of status hypothesis in different species are still needed before this hypothesis can be

considered as a generally applicable explanation for the honesty of testosterone related signals.

In addition, the Badge of Status Hypothesis was created to only explain the honesty of signals of status, that is only of a subset of signals of quality. The question arises therefore of how quality signals other than signals of status (e.g. those important in mate choice) remain honest. As it can be seen from the badge of the male House Sparrow, many signals of status might also play an important role in mate choice (see above). Therefore, the separation into different kinds of testosterone related quality signals might, in fact, not be very meaningful. This is likely because testosterone affects not only aggressiveness but also other reproductive behaviours (e.g. Wingfield et al. 1990). As a consequence, the honesty of ornaments signalling different kinds of behaviours, such as aggression and breeding behaviour, that are linked with each other via testosterone levels, need to be only enforced for one of them, e.g. for aggression. As aggression or status can be easily tested and thus honesty can be enforced via social costs (see above), and as choosing females could potentially eavesdrop on these encounters (or at least their outcome, i.e. achieved status as inferred by phenotype), signals of any kind of testosterone related quality (no matter what type of ornament or what species) could be kept honest via social costs (discussed in Berglund et al. 1996). However, this remains to be experimentally tested in future studies.

Besides signals of quality there is a range of other types of ornaments such as signals of kinship, of individual identity, or for species recognition (summarized in Dale 2006). For some of them, honesty enforcing costs could still be social costs in the broader sense, simply if individuals with a certain phenotype are avoided by conspecifics. For others, such as Fisherian traits that evolve from female preferences, no costs are needed because such ornaments do not signal any true information. Generally, no matter what information is conveyed signal honesty can be enforced as long as costs for cheating are higher than benefits gained from cheating (reviewed in Számadó 2011).

Conclusion

I have addressed and challenged a range of questions concerning ornamentation and signal honesty using the House Sparrow as a model species. More specifically, I have shown that several key assumptions of the Testosterone Handicap Model about signalling honesty are not of a general nature. Instead, I provided indirect support for the honesty enforcement of testosterone related signals of quality via alternative costs such as social costs (according to the Badges of Status Hypothesis). For future studies it would be therefore important to critically test the social costs hypothesis not only in House Sparrows but in a range of different species. This should be done by testing whether cheating is costly, i.e. whether it would be punished by conspecifics via aggressive challenges. Cheating in this context means more elaborate ornaments than one would expect from underlying testosterone levels and testosterone related behaviours. This could be achieved by manipulations of ornaments of individuals that are known or unknown to conspecifics. Several studies have used such ornament manipulations to test signal honesty (Parsons and Baptista 1980; Fugle et al. 1984; Veiga 1995; Tibbetts and Dale 2004; Tibbetts and Izzo 2010), however, all of them have focused on one ornament alone and have therefore neglected the interplay of multiple ornaments. As multiple ornaments are commonly used to signal quality related information (chapter three) this should, despite its difficulties, be taken into account when manipulating ornamentation. In addition, it could be also fruitful to compare individuals whose ornaments have been manipulated with other individuals whose plasma testosterone levels were simultaneously manipulated, because in the latter one would expect an honest relationship between signal and behaviour. A mismatch between signal and behaviour was found in Paper Wasps (*Polistes dominulus*) to cause social punishment according to the incongruence hypothesis (Tibbetts and Izzo 2010). This hypothesis predicts aggression towards any type of inaccurate signalling (behavioural or ornamental) as a function of self interest, i.e. of testing a rival in order to obtain a higher rank or better resources (reviewed in Tibbetts and Izzo 2010). Therefore, a careful test could potentially explain signal honesty in a wide range of contexts.

In addition to the direct examination of social costs, it could be also useful for future studies to concentrate on maximal testosterone levels as enforced via GnRH injections

or to focus on external factors that modulate circulating testosterone levels. It has been suggested for example, that testosterone levels are suppressed by immune activation (Verhulst et al. 1999; Boonekamp et al. 2008). The manipulation of environmental and physiological factors and the examination of their effects on natural testosterone levels could potentially reveal (1) how individual variation of testosterone levels can exist when the main costs of high testosterone levels are behavioural costs and (2) how other studies have found relationships with testosterone levels and e.g. parasite load. Additionally, it could explain why studies on different populations of the same species (e.g. the House Sparrow) or on different species with similar ecologies often find different and contradictory results.

Finally, another very important aspect for future studies is the detailed examination of elevated testosterone levels at night. More specifically, the question still remains why night time plasma testosterone levels are much higher than day time levels (discussion of chapter two). Furthermore, details of testosterone levels of focal individuals in the course of the day and the night and over several seasons or even years could uncover not only individual differences and generalities in secretion patterns but also potentially explain different behavioural strategies related to testosterone levels. Last but not least the comparison of night time testosterone levels with levels after social challenges (according to the Challenge Hypothesis (Wingfield et al. 1990)) and after GnRH challenges could prove if night time testosterone levels are close to potential maximal levels and hence especially important for the development of signals used in the context of male-male competition and aggression.

References

Anderson TR (2006) Biology of the ubiquitious house sparrow. Oxford University Press, New York

Berglund A, Bisazza A, Pilastro A (1996) Armaments and ornaments: An evolutionary explanation of traits of dual utility. Biol. J. Linn. Soc. 58:385-399

Bókony V, Lendvai AZ, Liker A (2006) Multiple cues in status signalling: The role of wingbars in aggressive interactions of male house sparrows. Ethology 112:947-954

Boonekamp JJ, Ros AHF, Verhulst S (2008) Immune activation suppresses plasma testosterone level: a meta-analysis. Biol. Letters 4:741-744

Cockrem JF, Seddon PJ (1994) Annual Cycle of Sex Steroids in the Yellow-Eyed Penguin (Megadyptes Antipodes) on South-Island, New-Zealand. Gen. Comp. Endocrinol. 94:113-121

Cordero PJ, Wetton JH, Parkin DT (1999) Extra-pair paternity and male badge size in the House Sparrow. J. Avian Biol. 30:97-102

Dale J (2006) Intraspecific variation in coloration. In: Hill GE, McGraw KJ (eds) Bird Coloration Harvard University Press, MA, Cambridge, pp 36-86

Dawson A (1983) Plasma Gonadal-Steroid Levels in Wild Starlings (Sturnus-Vulgaris) during the Annual Cycle and in Relation to the Stages of Breeding. Gen. Comp. Endocrinol. 49:286-294

Fugle GN, Rothstein SI, Osenberg CW, Mcginley MA (1984) Signals of Status in Wintering White-Crowned Sparrows, *Zonotrichia leucophrys gambelii*. Anim. Behav. 32:86-&

Griffith SC (2000) A trade-off between reproduction and a condition-dependent sexually selected ornament in the house sparrow *Passer domesticus*. Proc. R. Soc. Biol. Sci. Ser. B 267:1115-1119

Griffith SC, Owens IPF, Burke T (1999) Female choice and annual reproductive success favour less-ornamented male house sparrows. Proc. R. Soc. Biol. Sci. Ser. B 266:765-770

Griffith SC, Parker TH, Olson VA (2006) Melanin-versus carotenoid-based sexual signals: is the difference really so black and red? Anim. Behav. 71:749-763

Hahn TP, Swingle J, Wingfield JC, Ramenofsky M (1992) Adjustments of the Prebasic Molt Schedule in Birds. Ornis Scand. 23:314-321

Hatch MI, Westneat DF (2007) Age-related patterns of reproductive success in house sparrows *Passer domesticus*. J. Avian Biol. 38:603-611

Jawor JM (2007) Testosterone in Northern Cardinals (Cardinalis cardinalis): Possible influence of prolonged territorial behavior. Auk 124:331-338

Jawor JM, Breitwisch R (2003) Melanin ornaments, honesty, and sexual selection. Auk 120:249-265

Jawor JM, McGlothlin JW, Casto JM, Greives TJ, Snajdr EA, Bentley GE, Ketterson ED (2006) Seasonal and individual variation in response to GnRH challenge in male dark-eyed juncos (*Junco hyemalis*). Gen. Comp. Endocrinol. 149:182-189

Karubian J (2002) Costs and benefits of variable breeding plumage in the red-backed fairy-wren. Evolution 56:1673-1682

Karubian J (2008) Changes in breeding status are associated with rapid bill darkening in male red-backed fairy-wrens *Malurus melanocephalus*. J. Avian Biol. 39:81-86

Kellam JS, Wingfield JC, Lucas JR (2004) Nonbreeding season pairing behavior and the annual cycle of testosterone in male and female downy woodpeckers, Picoides pubescens. Horm. Behav. 46:703-714

Ketterson ED (1979) Status Signaling in Dark-Eyed Juncos. Auk 96:94-99

Lyon BE, Montgomerie RD (1986) Delayed Plumage Maturation in Passerine Birds - Reliable Signaling by Subordinate Males. Evolution 40:605-615

Maynard Smith JFRS, Harper DGC (1988) The Evolution of Aggression - Can Selection Generate Variability? Philos. Trans. R. Soc. Lond. B Biol. Sci. 319:557-570

McGraw KJ (2004) Winter plumage coloration in male American goldfinches: Do reduced ornaments serve signaling functions in the non-breeding season? Ethology 110:707-715

Møller AP (1987) Variation in Badge Size in Male House Sparrows *Passer domesticus* - Evidence for Status Signaling. Anim. Behav. 35:1637-1644

Møller AP (1990) Sexual-Behavior Is Related to Badge Size in the House Sparrow *Passer domesticus*. Behav. Ecol. Sociobiol. 27:23-29

Nakagawa S, Burke T (2008) The mask of seniority? A neglected age indicator in house sparrows *Passer domesticus*. J. Avian Biol. 39:222-225

Parsons J, Baptista LF (1980) Crown Color and Dominance in the White-Crowned Sparrow. Auk 97:807-815

Porter JM, Sealy SG (1982) Dynamics of Seabird Multispecies Feeding Flocks: Age-Related Feeding-Behavior. Behaviour 81:91-109

Rohwer S (1975) Social Significance of Avian Winter Plumage Variability. Evolution 29:593-610

Romero LM, Reed JM (2008) Repeatability of baseline corticosterone concentrations. Gen. Comp. Endocrinol. 156:27-33

Stewart IRK, Hanschu RD, Burke T, Westneat DF (2006) Tests of ecological, phenotypic, and genetic correlates of extra-pair paternity in the House Sparrow. Condor 108:399-413

Számadó S (2011) The cost of honesty and the fallacy of the handicap principle. Anim. Behav. 81:3-10

Tibbetts EA, Dale J (2004) A socially enforced signal of quality in a paper wasp. Nature 432:218-222

Tibbetts EA, Izzo A (2010) Social Punishment of Dishonest Signalers Caused by Mismatch between Signal and Behavior. Curr. Biol. 20:1637-1640

Tibbetts EA, Safran RJ (2009) Co-evolution of plumage characteristics and winter sociality in New and Old World sparrows. J. Evol. Biol. 22:2376-2386

Veiga JP (1995) Honest Signaling and the Survival Cost of Badges in the House Sparrow. Evolution 49:570-572

Verhulst S, Dieleman SJ, Parmentier HK (1999) A tradeoff between immunocompetence and sexual ornamentation in domestic fowl. Proc. Natl. Acad. Sci. U. S. A. 96:4478-4481

Whitekiller RR, Westneat DF, Schwagmeyer PL, Mock DW (2000) Badge size and extra-pair fertilizations in the House Sparrow. Condor 102:342-348

Wingfield JC, Farner DS (1993) Endocrinology of reproduction in wild species. In: Farner DS, King JR, Parker KC (eds) Avian biology. Academic Press Ltd, New York, pp 163-327

Wingfield JC, Hegner RE, Dufty AM, Ball GF (1990) The Challenge Hypothesis - Theoretical Implications for Patterns of Testosterone Secretion, Mating Systems, and Breeding Strategies. Am. Nat. 136:829-846

Wingfield JC, Lynn SE, Soma KK (2001) Avoiding the 'costs' of testosterone: Ecological bases of hormone-behavior interactions. Brain Behav. Evol. 57:239-251

Acknowledgements

First of all I would like to thank my day-to-day supervisor James Dale for sparking my interests in male ornamentation and signal honesty and for always reminding me that there are other species than House Sparrows. Also many thanks for trying to teach me independence, accuracy and patience, for always motivating me, and for help, support and advice. I am also very grateful to my other supervisor Bart Kempenaers for giving me the opportunity to do a PhD in his group, for letting me participate in international conferences, but most importantly for advice and constructive critique. In addition, I would like to thank my PAC members Michaela Hau and Wolfgang Goymann for discussion and helpful comments.

These studies would have not been possible without the endless help of the animal care takers (Frances Weigel, Magnus Ruhdorfer, Edith Bodendorfer, Annemarie Grötsch, and several others), technical assistants (Andrea Wittenzellner, Agnes Türk, and Katja Temnow), interns and students (Maryjane Heckel, Julia Wickjürgen, Pernilla Borgström, and others), and many colleagues who volunteered during day and night times (just to name a few examples: Barbara Helm, Rebecca Wutzler). I am also very grateful to Martin Dehnhard, Wolfgang Goymann, Kevin McGraw, Dennis Hasselquist, Elena Arriero, and their lab teams who analyzed different samples for me and who taught me lab techniques, our lab team (Sylvia Kuhn, Kim Teltscher, Alexander Girg, and Christine Hermannstädter), Lisa Trost for help with getting the animal experiment permit, Hendrik Reers for little tipps and tricks, the workshops and library teams in Seewiesen, and most importantly Mihai Valcu and Holger Schielzeth for patiently answering all my statistical questions. I would also like to thank the Behavioural Ecology and Evolutionary Genetics Group of the MPIO and especially Elisabeth Bolund, Emmi Schlicht, Rebecca Wutzler, and John Lesku for fruitful discussions and ideas.

Also many thanks to Helga Gwinner, Alain Jacot, Stefan Greif, and Emmi Schlicht for interesting hours and discussions when collecting data for interesting side projects that haven't made it into this thesis.

Last but not least, I would like to thank my family and friends (scientists and non-scientists) for always being interested in what I was doing and for always supporting me.

Silke Laucht
Seewiesen, April 2011

Author Contributions

Chapter One

S.L. contributed to the study concept and design, practical work, ornament scoring, data analyses, and writing.
B.K. contributed to the study concept, discussion, and provided comments on the manuscript.
J.D. contributed to the study concept and design, practical work, discussion, and writing.

Chapter Two

S.L. contributed to the study concept and design, practical work, ornament scoring, data analyses, and writing.
J.D. contributed to the study concept and design, practical work, discussion, and writing.
A.M. helped with practical work, and provided comments on the manuscript.
B.K. contributed to the study concept, discussion, and writing.

Chapter Three

S.L. contributed to the study concept and design, practical work, ornament scoring, data analyses, and writing.
B.K. contributed to the study concept, discussion, and provided comments on the manuscript.
J.D. contributed to the study concept and design, practical work, discussion, and writing.

Chapter Four

S.L. contributed to the study concept and design, practical work, ornament scoring, data analyses, and writing.
B.K. contributed to the study concept, discussion, and provided comments on the manuscript.

J.D. contributed to the study concept and design, practical work, discussion, and writing.

Chapter Five

S.L. contributed to the study concept and design, practical work, ornament scoring, data analyses, and writing.

B.K. contributed to the study concept, discussion, and provided comments on the manuscript.

K.J.M. analyzed plasma carotenoid levels, and provided comments on the manuscript.

E.A. helped with analyzing immune parameters, and provided comments on the manuscript.

W.G. analyzed plasma testosterone levels, contributed to discussion, and provided comments on the manuscript.

D.H. helped with analyzing immune parameters, and provided comments on the manuscript.

J.D. contributed to the study concept and design, practical work, discussion, and writing.

Bart Kempenaers	Silke Laucht
Doktorvater	Doktorandin

Present Addresses of Co-authors

James Dale (previously at the MPIO Seewiesen)
Institute of Natural Sciences
Massey University
Private Bag 102 904
North Shore City 0745, Auckland, New Zealand

Bart Kempenaers
Max Planck Institute for Ornithology
Department of Behavioural Ecology and Evolutionary Genetics
Eberhard-Gwinner-Str. 7/8
D-82319 Seewiesen (Starnberg), Germany

Elena Arriero (previously at the MPIO Radolfzell)
School of Biology, University Complutense of Madrid
Department of Zoology and Physical Anthropology
Jose Antonio Novais 2
28040 Madrid, Spain

Wolfgang Goymann
Max Planck Institute for Ornithology
Department of Behavioural Neurobiology
Eberhard-Gwinner-Str. 6
D-82319 Seewiesen (Starnberg), Germany

Dennis Hasselquis
Lund University
Department of Animal Ecology
Ecology Building
22362 Lund, Sweden

Kevin J. McGraw
Arizona State University
School of Life Sciences
427 E. Tyler Mall
Tempe, AZ 85287-4501, USA

Ariane Mutzel
Max Planck Institute for Ornithology
Department of Behavioural Ecology and Evolutionary Genetics
Eberhard-Gwinner-Str. 7/8
D-82319 Seewiesen (Starnberg), Germany

Curriculum Vitae

Name	Silke Laucht
Present address	Max Planck Institute for Ornithology Department of Behavioural Ecology and Evolutionary Genetics Eberhard-Gwinner Str. 7/8 D-82319 Seewiesen (Starnberg), Germany phone: +49-(0)8157-932 320 fax: +49-(0)8157-932-400 email: laucht@orn.mpg.de
Date of birth	20.12.1980
Place of birth	Munich, Germany

Education

1991-2000	Secondary school, Gymnasium Bad Aibling Final grade: 1.4
2000-2006	Biology studies at the University of Ulm, Germany Diploma with final grade 1.2
	Diploma thesis: "Differences in age- and sex ratios of migrating landbirds in Long Point, Ontario, Canada, in relation to location and light sources" (supervisors Prof. Elisabeth Kalko, University of Ulm, and Jon McCracken, Bird Studies Canada)
2003-2004	Studies at Laurentian University, Sudbury, Ontario, Canada
	4th year thesis: "Taxonomic patterns of diversity in recovering plant communities: a comparison of the recovery processes in Sudbury, Ontario and Mt St Helens, Washington" (supervisor Prof. Madhur Anand)
2006	6-months internship at the Max Planck Institute for Ornithology, Seewiesen, Germany
2006-2007	Technical Assistant at the Max Planck Institute for Ornithology
2007-present	PhD student at the Max Planck Institute for Ornithology **Doctoral thesis**: "Information content and testosterone dependence of animal signals: a case study with House Sparrows" (supervisors: Dr. James Dale, Prof. Bart Kempenaers)

Publications in Peer-reviewed Journals

Mutzel A, Kempenaers B, Laucht S, Dingemanse NJ, Dale J:
Circulating testosterone levels do not affect exploration in house sparrows: observational and experimental tests.
Animal Behaviour 81: 731-739 (2011).

Laucht S, Dale J, Mutzel A, Kempenaers B:
Individual variation in plasma testosterone levels and its relation to badge size in House Sparrows *Passer domesticus*: It's a night-and-day difference.
General and Comparative Endocrinology 170: 501-508 (2011).

Ritschard M, Laucht S, Dale J, Brumm H:
Enhanced testosterone levels affect singing motivation but not song structure and amplitude in Bengalese finches.
Physiology & Behavior 102: 30-35 (2011).

Laucht S, Kempenaers B, Dale J:
Bill color, not badge size, indicates testosterone-related information in house sparrows.
Behavioral Ecology and Sociobiology 64: 1461-1471 (2010).

Laucht S, DuVal EH, Kempenaers B:
Maternal correlates of brood sex ratio variation in the lekking lance-tailed manakin.
Journal of Avian Biology 39: 198-205 (2008).

Conference Presentations

While you were sleeping: Diurnal variation of testosterone levels and its relationship to male ornamentation in House Sparrows. 13th International Behavioral Ecology Congress (ISBE); 26.09. – 01.10.10 in Perth, Australia (Talk)

Circadian variation of testosterone levels and its impact on male ornaments in House Sparrows, *Passer domesticus.* Graduiertentreffen Verhaltensbiologie (DZG and Ethologische Gesellschaft); 11. – 13.11.09 in Seewiesen, Germany (Poster)

Circadian variation of testosterone levels and its impact on male ornaments in House Sparrows, *Passer domesticus.* ASAB summer comference 2009; 02. – 04.09.09, St. John's College, Oxford, England (Poster)

Dynamics in testosterone, bill coloration, and badge size in male House Sparrows, *Passer domesticus*. 12th International Behavioral Ecology Congress (ISBE); 9. – 15.08.08 Cornell University, Ithaca, New York, USA (Talk)

Dynamics in testosterone and ornaments: A study on male House Sparrows. 13th Annual European Meeting of PhD Students in Evolutionary Biology; 13.- 18.08.07 in Lund, Sweden (Talk)

Declaration

Ehrenwörtliche Versicherung

Ich versichere hiermit ehrenwörtlich, dass die von mir vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

München, den _____ _____

Silke Laucht

Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist, und dass ich mich anderweitig einer Doktorprüfung ohne Erfolg nicht unterzogen habe.

München, den _____ _____

Silke Laucht