

## Impacts of acute and long-term vehicle exposure on physiology and reproductive success of the northern spotted owl

LISA S. HAYWARD,<sup>1,†</sup> ANN E. BOWLES,<sup>2</sup> JAMES C. HA,<sup>3</sup> AND SAMUEL K. WASSER<sup>1</sup>

<sup>1</sup>Center for Conservation Biology, University of Washington, Box 351800, Seattle, Washington 98195 USA

<sup>2</sup>Hubbs-SeaWorld Research Institute, 2595 Ingraham Street, San Diego, California 92109 USA

<sup>3</sup>Department of Psychology, University of Washington, Seattle, Washington 98195 USA

**Abstract.** Stress physiologists posit that multiple simultaneous demands faced by an organism may have non-additive effects on the magnitude of their response to disturbance. The environmental assessment literature emphasizes a similar phenomenon at the population level, arguing that populations can compensate for perturbations up to a threshold, beyond which disturbance impacts may be greatly magnified—and even cause system collapse. We integrated these two approaches to examine the roles of environment, life history stage, prior disturbance experience, and their interactions on vulnerability to disturbance in a free-living species. Specifically, we examined the effects of off-highway vehicle use on the federally threatened northern spotted owl (NSO), *Strix occidentalis caurina*, by measuring fecal glucocorticoid metabolites (fGCs), which reflect disturbance; fecal thyroid hormone metabolites (fT3), which reflect nutrition; and the number of offspring fledged within a season. We experimentally applied one hour of motorcycle exposure to NSOs during periods of incubation (May) and fledging (July), comparing fGC levels of treated NSO with those of non-exposed controls. Acute vehicle exposure generally increased fGCs in the short term. Males showed the highest glucocorticoid response to vehicle disturbance in May when they were typically solely responsible for feeding themselves, their mates and their nestlings. By contrast, response to motorcycle exposure among females depended on their level of fT3 and their number of young. Levels of fGCs were highest post treatment among females that lacked young and had high fT3 (good nutrition); fGC levels were lower in treated females compared to controls among females with two young and low fT3 (compromised nutrition), possibly reflecting allostatic overload. The correlational approach showed that NSO close to roads had higher levels of fT3, suggesting better nutrition. Surprisingly, fGC levels were unrelated to proximity of roads, irrespective of noise. Presumably, the tendency for traffic exposure to increase fGCs over the long-term was offset by nutritional gains (i.e., reduced fGCs and high fT3) associated with proximity to roads. Sound level meters enabled us to quantify road noise on a subset of NSO territories. NSO close to noisy roads fledged significantly fewer young than NSO near quiet roads, indicating that routine traffic exposure may decrease NSO reproductive success over time.

**Key words:** allostasis; conservation endocrinology; disturbance; northern spotted owl; physiology; *Strix*; traffic; wildlife.

**Received** 17 December 2010; revised 4 April 2011; accepted 5 April 2011; final version received 11 May 2011; **published** 3 June 2011. Corresponding Editor: D. P. C. Peters.

**Citation:** Hayward, L. S., A. Bowles, J. C. Ha, and S. K. Wasser. 2011. Impacts of acute and long-term vehicle exposure on physiology and reproductive success of the northern spotted owl. *Ecosphere* 2(6):art65. doi:10.1890/ES10-00199.1

**Copyright:** © 2011 Hayward et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits restricted use, distribution, and reproduction in any medium, provided the original author and sources are credited.

† **E-mail:** lhayward@u.washington.edu

## INTRODUCTION

### *Disturbance impacts on wildlife*

Organisms accommodate a dynamic suite of simultaneous and disparate pressures in pursuit of survival and reproduction. Stress physiologists have recently developed the concept of “allostatic load” to represent this condition (McEwen and Wingfield 2003, Sterling 2004). Every organism sustains some level of allostatic load associated with basic physiological and behavioral processes. Regulatory mechanisms such as the hypothalamic-pituitary-adrenal axis and thyroid hormone modulate the demands of this daily load. Pressures accumulating below a certain threshold can be accommodated by the regulatory systems in an adaptive manner (Wingfield and Kitaysky 2002). But when accumulating pressures tip the system toward “allostatic overload”, costs to reproduction and survival may ensue. Theoretically, the tipping point from load to overload is dictated by a combination of variables: life history (e.g., sex, breeding status), environment (e.g., weather, food availability, anthropogenic disturbance), body condition (e.g., fat reserve, parasite load) and central nervous system processing, such as prior experience of the disturbance and perception of threat. The effects of various pressures often interact to increase individual vulnerability to disturbance (McEwen and Wingfield 2003, Korte et al. 2005).

The cumulative effects concept in resource management was developed independently from the concept of allostasis; yet the two concepts share the perspective that disparate pressures and their interactions have non-additive effects on a system’s response to disturbance. Such interactions may be explained in part by the capacity of living organisms to accommodate environmental pressures below a certain threshold, often through behavioral and/or physiological plasticity (e.g., Jones et al. 2001). Once the threshold is reached, however, further disturbance may cause system collapse. Environmental impact assessments require consideration of cumulative effects, but these have been difficult to quantify (Cooper and Sheate 2002).

We integrated the concept of allostatic load with cumulative effects models to develop principles for determining vulnerability of wild-

life to disturbance, at both the individual and population level. Using these concepts, we designed a comprehensive test to determine the impacts of off-highway vehicle (vehicle) exposure on the northern spotted owl (NSO), *Strix occidentalis caurina*.

Roads and traffic noise are a growing concern to wildlife managers. Road traffic nearly tripled between 1971 and 2007, whereas the U.S. human population grew by only about one third during that 30-year period (<http://www.census.gov/compendia/statab>); (<http://www.fhwa.dot.gov/ohim/tvtw/tvtpage.cfm>); reviewed in Barber et al. 2009). Off-highway vehicle sales more than tripled between 1995 and 2003 (Cordell et al. 2005). Dale Bosworth (Chief, USFS) ranked unmanaged recreation as one of the four most serious threats to public lands (Bosworth 2004).

Several studies suggest that traffic exposure negatively impacts wildlife (White and Thurow 1985, Yarmoloy et al. 1988, Creel et al. 2002, Arlettaz et al. 2007, Barja et al. 2007, Thiel et al. 2008), but effects on survival and reproductive success have rarely been quantified (for a recent review see Barber et al. 2009). By contrast, physiological impacts of traffic exposure are relatively well-documented (Walker et al. 2006, Creel et al. 2002, Wasser et al. 1997). For example, Creel et al. (2002) showed an association between exposure to snowmobiles and elevated fecal glucocorticoid metabolites (fGCs) in wolves and elk but found no evidence of an effect of traffic on population dynamics. To our knowledge, no study has examined how the GC response to traffic varies with life history and environmental factors. Understanding such variation is important to conservation because it will help managers mitigate the impacts of anthropogenic activities on wildlife (Tarlow and Blumstein 2007).

### *Overview of study design*

From 2005 to 2008 we conducted a large-scale field study to quantify the effects of off-highway vehicle exposure on NSO in northern California near the southern limit of their range. Controlled experiments, wherein half of NSO pairs were exposed to an hour of motorcycle traffic, tested the following two primary hypotheses:

1. Acute vehicle exposure increases short-term

stress in NSO, indicated by elevated fGCs.

2. NSO GC response to acute vehicle exposure will vary with sex, season, breeding and nutritional status [as reflected by level of fecal thyroid hormone metabolites (fT3)].

In addition, correlational analysis tested an important third hypothesis:

3. Exposure to high levels of traffic noise will be associated with elevated baseline psychological (increased fGC) and nutritional stress (increased fGCs and diminished fT3) and reduced reproductive success.

As part of the correlational component we looked for associations of distance to closest road with baseline levels of fGCs, fT3 and number of offspring fledged. We used sound level meters (SLMs) to quantify noise on the roads closest to each site continuously for four week and weekend days during July of 2006 and 2007. After screening out continuous noise (e.g., wind, water, and persistent insect noise) we tested for an association of routine human-caused noise and other transient events on the road with NSO hormone levels and/ or reproductive success.

#### *Fecal hormone measures*

Circulating levels of steroids and thyroid hormones represent complementary mechanisms by which vertebrates integrate the immediate and long-term demands of survival and reproduction. These mechanisms hold enormous diagnostic potential for conservationists, not only to quantify disturbance impacts, but also to evaluate mitigation success. Still, we need a more basic understanding of the disturbance response in free-living organisms to help interpret these measures in a natural context. Although they require careful validation for each species (Touma and Palme 2005), fecal hormone metabolite measures provide an advantage over plasma measures because they can be obtained non-invasively. Non-invasive approaches are especially useful with threatened and endangered species; they reduce the need for multiple levels of controls, and sample acquisition is less likely to confound stress measures. Fecal measures also provide an integrated measure of hormone levels over time rather than the snapshot of a single time point represented by blood samples (Wasser

et al. 2000).

Glucocorticoids (GCs) are the hormones most often measured in studies of disturbance. Frequently used as index of the “stress” perceived by an organism, GCs show good potential as a diagnostic tool for conservationists (Busch and Hayward 2009). In all vertebrates studied to date, GCs increase as part of a generalized physiological stress response to challenge (Seyle 1946, Romero 2004). Similarly, relatively high baseline GC levels tend to be associated with disturbance. For example, high GCs have been associated with contamination from oil spill (e.g., Wikelski et al. 2002), increased predation pressure (e.g., Boonstra et al. 1998), decreased food availability (e.g., Clinchy et al. 2004), traffic exposure (e.g., Creel et al. 2002), tourism (e.g., Walker et al. 2006) and habitat fragmentation (e.g., Suorsa et al. 2003, for review see Busch and Hayward 2009). Elevated GCs tend to be associated with decreased fitness either due to reduced survival (e.g., Pride 2005) or reduced reproductive success (e.g., Bonier et al. 2007). However, the relationship between GCs and fitness is not always significant or consistent (Busch and Hayward 2009). Multiple stressors (e.g., psychological disturbance or reduced food availability) can also elevate GCs, complicating its interpretation.

Thyroid hormones are steroid-like hormones that have profound effects on metabolism, heart rate, blood pressure, and thermoregulation (Oppenheimer 1999). While thyroxine (T4) is the most common thyroid hormone in the body, triiodothyronine (T3) is the most biologically active, and therefore the most relevant for managers. Measures of T3 represent another valuable diagnostic tool for conservationists because T3 is closely associated with nutritional condition, generally found in high levels in well-fed individuals and in low levels during starvation (van der Heyden et al. 1986, Eales 1988, Hennemann et al. 1988, Blake et al. 1991, Flier et al. 2000, Douyon et al. 2002, Wasser et al. 2010). While GCs increase in response to both psychological and nutritional stress (Sapolsky et al. 2000, Wingfield and Romero 2001, Kitaysky et al. 2005, Wasser et al. 1997, 2000, 2010), T3 decreases in response to nutritional stress (Eales 1988, Flier et al. 2000, Douyon et al. 2002), but appears unaffected by psychological stress (Schew et al. 1996, Geris et al. 1999, see also Walpita et al.

2007). We can use T3, then, to help assess whether deleterious effects of a potential disturbance like traffic are direct or may be mediated through effects on food availability.

#### *Northern spotted owls*

The northern spotted owl (NSO) is listed as federally threatened under the Endangered Species Act. Long-term demographic studies show that NSO populations are declining at an average annual rate of 4% throughout their range (Courtney et al. 2004). Recent data, however, suggest that populations in northwestern California are not currently undergoing decline (Franklin et al. 2008). Wasser et al. (1997) found that, in Washington, male NSO within a quarter mile of logging roads had higher baseline fGCs than males farther from roads. Females showed no physiological effect of proximity to road or logging activity. At the same time Wasser et al. (1997) found that NSO baseline fGCs vary by sex and season, increasing in females through the course of the breeding season but decreasing in males (Wasser and Hunt 2005).

## METHODS

### *Study sites*

Researchers conducted fieldwork in Mendocino National Forest (Mendocino) and the South Fork Management Unit of Shasta-Trinity National Forest (Shasta) in Northern California. Both forests serve multiple uses including recreation, grazing, mining and logging. Forests are a mix of hardwood and deciduous trees, primarily Douglas fir (*Pseudotsuga menziesii*), Ponderosa pine (*Pinus ponderosa*), and multiple species of oak (*Quercus* spp).

Approximately 200,000 of Mendocino's 913,306 acres are designated as off-highway vehicle emphasis area and contain 170 miles of motorcycle, 112 miles of all-terrain vehicle and 187 miles of 4-wheel drive/ dune buggy routes. Mendocino annually hosts five enduro events, sanctioned by the American Motorcyclist's Association. Enduros are popular motorcycle events in which motorcyclists ride fast on forest trails for long distances. The Shasta encompasses approximately 400 miles of off-highway vehicle trails. Its 52,000 acre Chappie-Shasta off-highway vehicle emphasis area has about 250 miles of trails.

Though the area does not as yet host enduro events, it does host other annual competitive vehicle events. In both forests, off-highway vehicle recreation occurs year round (snowmobile use is prevalent in winter) with heaviest use from October to May, overlapping with the NSO breeding season (March–Aug).

### *Experimental design*

Active NSO sites (sites containing one or more breeding or non-breeding NSO adults) were identified through standard survey methods (Forsman 1983). On the days of experiment, sites were randomly assigned as either treatment or control and subjected to identical methods of data collection. Both control and treatment sites were located in areas ranging from high to low routine off-highway vehicle use. Experiments were conducted twice throughout the season (in May and in July) to test for seasonal variation in NSO responsiveness. The first round of experiments occurred annually between May 12 and May 30 (during incubation). The second round was conducted between July 8 and 24 (early fledgling period). Logistics prevented sampling each forest in both months during all four years of the study. Some NSO sites were included in experiments twice within a year and/or in multiple years. Sound level measures were taken on a subset of sites in 2006 and 2007 (Appendix A: Methods for sound level measurement collection and analysis).

### *Scat collection*

Once adults were located, researchers laid plastic sheets under the roosts and worked in pairs to keep the NSO under continuous observation, recording time of arrival, time of NSO detection (auditory), location (visual) and production of all scat. Scat was transferred to 0.25 oz polypropylene jars with unlined screw caps (similar to SKS Bottle and Packaging # 0611-01) as quickly as possible after defecation. Researchers used clean twigs collected on site, taking care to collect the feces in their entirety. Jars were labeled, sealed in ziploc bags and transferred to coolers with ice. At the end of the day samples were frozen at  $-20^{\circ}\text{C}$  and kept frozen until transfer to the University of Washington for extraction. See Appendix B: Laboratory methods.

### *Determining nesting status*

Early in the season, domestic mice were offered to adult NSO to determine nesting status. Usually nesting NSO take mice to their incubating mates or young, rather than eating or caching them (Forsman 1983). By following a NSO that flies with a mouse, one can often locate its nest. If a NSO took a maximum of eight mice and ate or cached them, the pair was assumed to be non-reproductive. Nesting status was further verified during the 15 plus hours of observation that each adult received during the days of experiment in May and July. Some failed nests may have gone undetected.

### *Simulated enduro events*

Acute off-highway vehicle exposure was applied by conducting simulated enduro events (enduros) in which experienced motorcycle riders rode their bikes back and forth for one hour along an approximately half mile stretch of road closest to the owl's roost or nest. The range of distances between road and NSO was 5–800 m. Simulated enduros involved between one and six riders on two-stroke or four-stroke bikes that met the legal criteria for noise production (less than 96 decibels). Volunteer riders were recruited from the local communities via fliers and an e-mail list generated by the Redding Dirt Riders, a local motorcycle club. Prior to the experiments, riders were instructed to ride at the same intensity as for an actual enduro. Experiments were conducted in accordance with University of Washington IACUC protocol # 2850-04.

On the days of experiment, teams of four researchers arrived at the NSO territories before first light (around 04:30) to maximize chances of locating pairs without having to simulate NSO calls. When possible, spontaneous vocalizations and prior knowledge of nests or favored roosts helped locate the NSO. In some cases, researchers mimicked NSO calls and used NSO response to help locate them. Once NSO were located, researchers placed plastic sheets under their roosts to facilitate scat collection and kept the adults under continuous observation until they left the site after 20:00. Juveniles were observed as circumstances allowed (not always continuously) and their scat was collected opportunistically.

### *Sample size*

Individuals included in the study comprised all NSO found on historic NSO territories in the South Fork Management Unit of Shasta and in Mendocino during mid April to late July of 2005–2008. Field work in 2007 was limited to only Shasta and only during the months of June and July due to a lawsuit against the California Off-Highway Motorized Vehicle Recreation commission. In July 2008 forest fires limited experiments to only a few sites in Mendocino. Over 170 separate historic territories were surveyed to locate treatment and control subjects during the four years. Using Forest Service data, we focused surveys on historic NSO territories and applied standard vocalization survey techniques to locate NSO nests and roosts (Forsman 1983). Briefly, surveys involved simulating NSO vocalizations at set call points. When NSO responded to these calls, field crews hiked to their location and recorded Universal Transverse Mercators (UTMs) for each roost and/ or nest using handheld GPS devices, GPS used in the field calculated distance to closest road; GIS software later validated these measurements. Only sites with nests or roosts within 800 m of a road were included in the study.

*Experimental component.*—Over the course of four field seasons (2005–2008) researchers conducted 103 experiments (treatment and control combined) on 51 separate NSO sites, 28 in Shasta and 23 in Mendocino. Sites contained single males, non-reproductive adult pairs or adult pairs with eggs or young. Not all owls produced scat samples that could be collected on the days of experiment (either because the owls failed to defecate or because samples landed on high branches). Additionally, 22 samples were excluded from final analysis because of their small size. Appendix D: Table D1 shows the sample size for final analysis.

*Correlational component.*—Over the course of the four years within-season individual breeding success was calculated 142 times, sometimes for two NSO on a single site within a year, sometimes for what may have been the same individual on the same site over multiple years (individuals were not banded). While we included data from all four years of the study in analyses of proximity to road, sound level measures were only taken in 2006 and 2007 (see

Appendix A: Methods for sound level measurement collection and analysis). Further, in those years only a subset of sites had sound measures for four consecutive days that could be used in the correlative analysis (23 in 2006 and seven in 2007), therefore sample sizes are considerably reduced for all analyses that involve noise measures.

#### *Testing for glucocorticoid response to simulated enduro*

Prior to testing for a GC response to simulated enduro we had to establish that the presence of researchers on each NSO territory would not cause a GC response sufficient to mask the effects of further disturbance (see Appendix C: Northern spotted owl response to presence of researchers).

Because some NSO produced multiple pre and/or post samples, treatment effect was analyzed using post average residuals. To calculate these values, residuals for each sample were grouped into “pre” and “post” categories. “Pre” samples comprised all samples collected on the day of experiment prior to the start of the simulated enduro or during an equivalent time period on control sites. “Post” samples consisted of all samples collected more than two hours after the start of the simulated enduro (or during equivalent time on control sites). GCs secreted into circulation are first detectable as fGCs in feces produced two hours later (Wasser and Hunt 2005). Although our original intent was to use repeated measures tests, the high number of individuals for whom pre or post samples were never obtained reduced the sample size enough to prohibit this analysis with the appropriate covariates. Instead, we used the average of all “post” residuals for each individual NSO as the response variable in all tests of GC response after determining no differences among pre treatment groups.

Although breeding outcome (unpaired, paired and non-nesting, failed nest attempt, fledged one or fledged two) is a significant main effect, we had few unpaired NSO in our study and fewer still that attempted nesting and failed. This would have precluded testing for interactions of breeding outcome with other salient variables. Instead, we coded pairs as having zero, one, or two offspring fledged for further analysis, lumping failed breeders with those that did not nest or

had no mate.

#### *General linear models and calculating effect size*

Before we created statistical models, we tested data for compliance with parametric assumptions, applying LOG transformation to hormone levels and distance to road. We then created general linear models (GLMs) using JMP (standard least squares fit) to test our predictions about traffic impacts on the NSO. We created multiple versions for each model, including all combinations of main effects and interactions that were biologically plausible, to identify the models that explained the most variance (i.e., yielded the highest adjusted  $R^2$  value). When the model was not significant we reported the F ratios and p values for each main effect and interaction and for the overall model. But when the model was significant we assigned each main effect and interaction its own adjusted  $R^2$  value by dropping variables one at a time to quantify their contribution to the overall  $R^2$  (their  $\Delta R^2$ ). Using these  $\Delta R^2$  values we then calculated  $\Delta$  effect sizes for each main effect and interaction as  $R^2 \text{ adj} / (1 - R^2 \text{ adj})$ . Delta effect size refers to the portion of overall effect size that can be attributed to each separate main effect. We used Cohen's guidelines to assign relative strength of effect size as negligible (<0.2) small (0.2–0.3), medium (around 0.5) and large (over 0.8). Given that the hormone values are LOG transformed and that small changes in hormone concentrations can lead to hugely significant physiological and/or behavioral effects, we consider these categorizations to be very conservative for effects on hormones and appropriate for effects on reproductive success. When interactions proved significant we conducted post-hoc subset analysis to better characterize associations among the variables.

#### *Testing for pseudoreplication*

Because we sometimes sampled from more than one NSO on a single site on the same day and from the same NSO more than once within a season and/or over multiple years, we paid particular attention to the issue of pseudoreplication (Hurlbert 1984). Our original intent was to include site number in all statistical models as a control for potential pseudoreplication. However we found that when we included site number for

models with hormonal response variables we over-partitioned the degrees of freedom and introduced imbalance into the models. Further examination showed that the problem of relative variance was not present in the hormonal data. Variation was higher within sites than among sites. After testing for the consequences of pseudoreplication, and not finding them, we used models without including site number for hormonal response variables. Adding site number to models that used number of young fledged as the response variable, however, did prove an important control for pseudoreplication and did not introduce instability.

## RESULTS

### *GC response to simulated enduro events varies with life history*

Life history variables impacted NSO GC response to the simulated enduro treatment. Month, sex and number of young ultimately fledged all influenced post-enduro average fGC levels. Treatment was significant as a main effect (Appendix D: Table D1; Fig. 1) and also interacted significantly with sex, and sex and month (Figs. 2, 3). Adult males and juveniles

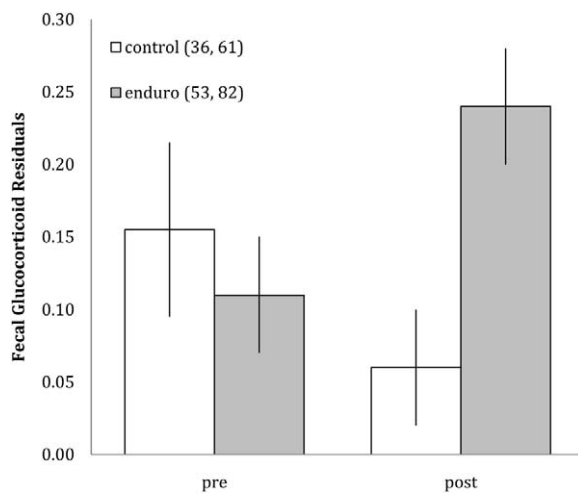


Fig. 1. Northern spotted owl fGC residuals before (pre) and after (post) one hour of experimentally applied motorcycle exposure on treated (enduro) and control sites, all subjects combined. Standard error is shown. Average residuals were adjusted to make values positive for graphing. Samples sizes are shown in parentheses.

showed a stronger response than did adult females (Fig. 2). Males had a higher GC response to treatment in May than in July (Appendix D: Table D3; Fig. 3). Female fGC level increased with number of young in May (Appendix D: Table D4), and there was a trend for treatment to interact with number of young for females in both May and July (Appendix D: Table D5). In both months, non-breeding females showed a higher GC response to enduro than breeders. Number of young did not influence fGC levels or interact with treatment among males.

### *T3 and breeding effort affect female GC response to enduro*

Among females, but not males, average morning fecal FT3 residuals interacted significantly with treatment and number of young to predict post treatment fGC levels (Appendix D: Table D6; Appendix E: Fig. E1). In general, enduro treatment *increased* fGC levels among females with high FT3 but *decreased* fGC levels relative to controls among females with low FT3. As the number of young increased from zero to two, the FT3 residual value representing an inflection point at which the treatment effect reversed directions increased from  $-0.28$  for non-breeders to  $0.23$  for females with one offspring and  $0.72$  for females with two young. Among non-breeding females with FT3 residuals higher than  $-0.28$  a strong, significant increase occurred in fGCs post enduro ( $n = 10$ ,  $R^2$  adj =  $0.51$ , effect size =  $1.04$ ,  $P = 0.01$ ). Among non-breeding females with FT3 residuals lower than  $-0.28$  there was a strong treatment effect in the opposite direction although the latter effect was not significant due to small sample size ( $n = 4$ ,  $R^2$  adj =  $0.52$ , effect size =  $1.08$ ,  $P = 0.18$ ). Low sample size prohibited separate analyses for each subset of breeding females.

### *Distance to road and road noise affect GC response to enduro*

Distance to road was not significant as a main effect or in interactions with treatment, sex and month when included in models of post-enduro GC response. However, subset analysis of males exposed to enduro showed a significant interaction of distance to road with month ( $R^2$  adj =  $0.22$ ,  $P = 0.006$ ,  $n = 41$ ). Further subset analysis showed that among males in May there was no

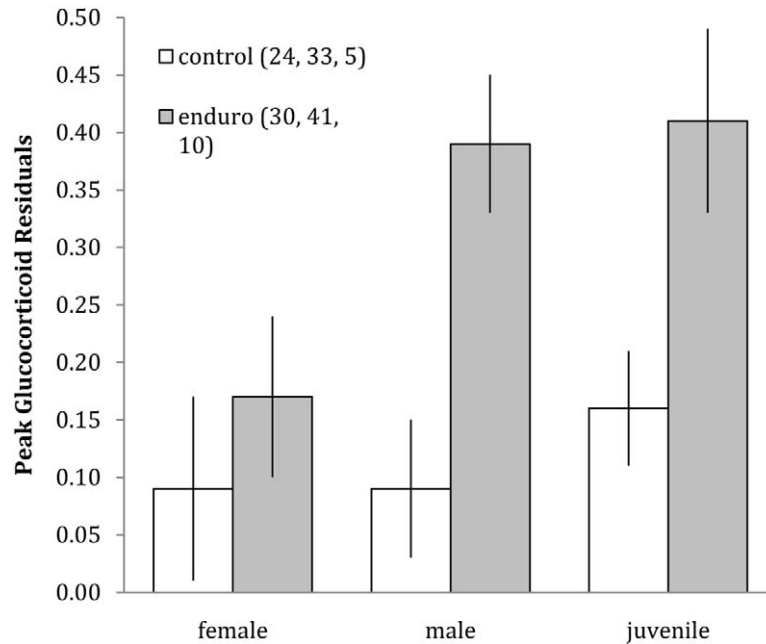


Fig. 2. Peak fGC residuals post treatment split by sex. Both months combined. Standard error is shown. Average residuals were adjusted to make values positive for graphing. Samples sizes are shown in parentheses.

significant relationship between distance to road and GC response ( $R^2$  adj = 0.06,  $P = 0.19$ ,  $n = 16$ ), while in July males closer to the road on which the enduro was conducted showed higher GC response ( $R^2$  adj = 0.14,  $P = 0.04$ ,  $n = 25$ , Fig. 4).

An analysis of the effects of routine road noise on GC response to enduro that included all adults exposed to enduros on sites with sound level recordings ( $n = 26$ ) showed a significant interaction of distance to road and road noise (Appendix D: Table D7). For adult NSO 50–800 m from the road, the higher the routine road noise on the road closest to NSO site, the lower the GC response to simulated enduro (Appendix D: Table D7). For NSO close to the road (within 50 m) routine road noise had no effect on the acute GC response to enduro.

#### *Correlative component I: distance to road and baseline hormone metabolites*

We found no relationship between distance to road and NSO baseline fGCs in this population (Appendix D: Table D8). Similarly, distance to road, road noise and their interaction had no effect on baseline fGCs, although sample size for this analysis was small (overall model  $R^2$  adj =

–0.12,  $P = 0.80$ ,  $n = 26$ ). In contrast, distance to road strongly predicted pre-experiment average fT3 (overall model  $R^2$  adj = 0.35,  $P = 0.03$ ,  $n = 77$ , Appendix D: Table D9), with fT3 highest close to roads. The strength of this correlation varied by month, and was stronger in July than in May (Appendix D: Table D10). After controlling for site number and year ( $F = 1.08$ ,  $P = 0.30$ ,  $n = 142$ ), we found no association between distance to road itself and number of fledged young.

#### *Correlative component II: road noise and reproductive success*

Distance to closest road and its interaction with road noise on that road were strongly and significantly associated with number of young fledged by mid July (overall model  $R^2$  adj = 0.93,  $P < 0.0001$ ,  $n = 42$ , Appendix D: Table D11). When noise levels on the road were low, NSO close to roads fledged more young; when noise levels on the road were high, the opposite pattern held (Fig. 5).

## DISCUSSION

### *GC response to simulated enduro events*

Overall, northern spotted owls increased fGC

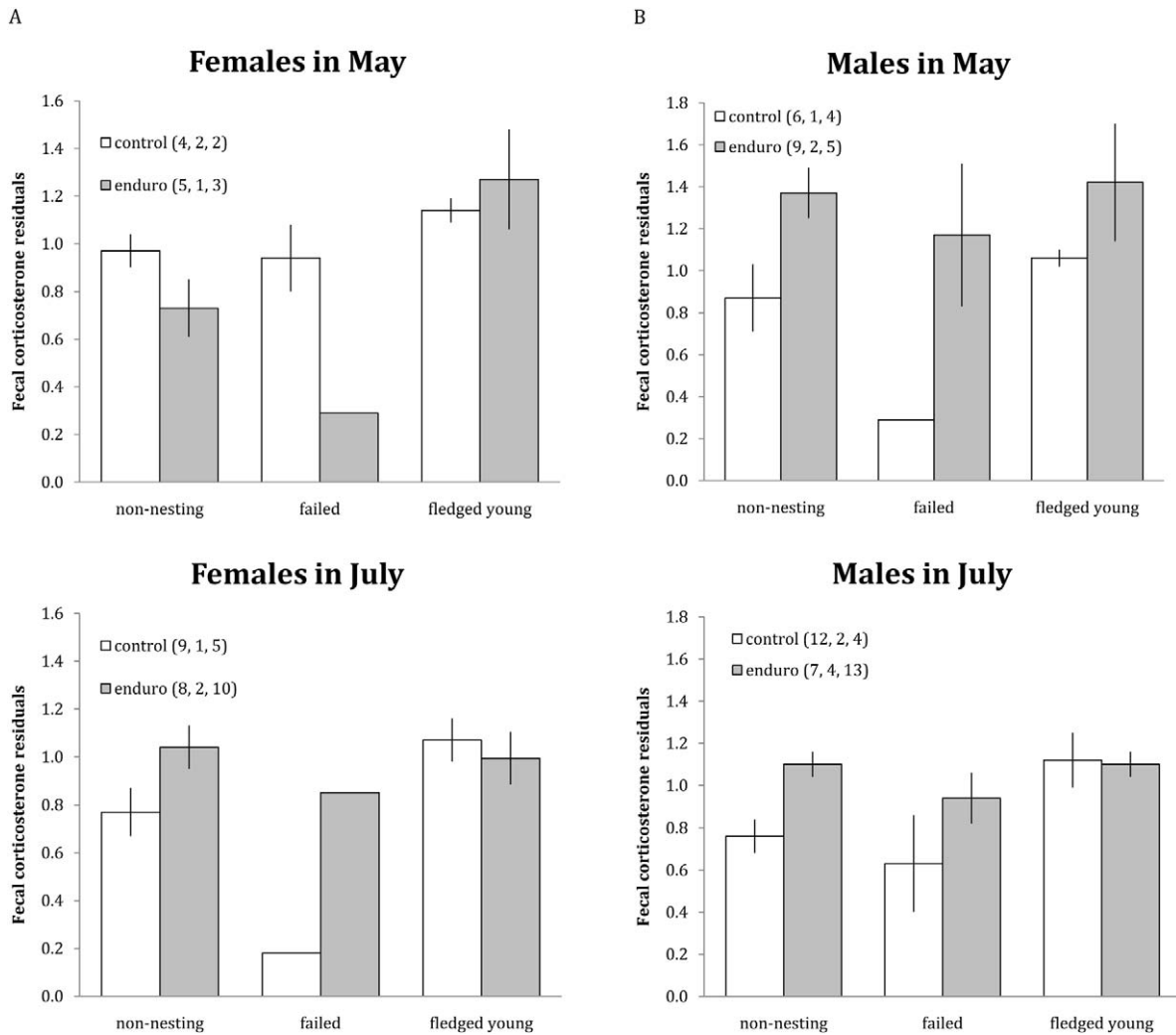


Fig. 3. Peak fGC residuals post treatment split by sex, treatment, month and breeding outcome. Standard error is shown. Average residuals were adjusted to make values positive for graphing. Samples sizes are shown in parentheses.

levels in response to acute vehicle exposure (Fig. 1). However, as predicted, the degree of GC responsiveness depended on sex, month (breeding stage), level of ft3 (nutritional status) and number of young (Appendix D: Table D2; Figs. 2, 3). The highest sensitivity to acute vehicle exposure occurred among males in May when females incubate almost constantly and males are solely responsible for feeding themselves, their mates and their young (Fig. 3). At that time, distance to the road on which the enduro occurred did not affect male GC response. But in July, when males showed less overall GC

response to enduro, the strength of the response diminished as distance to road increased (Fig. 4). Fledglings seemed to show high GC responsiveness to acute vehicle exposure (Fig. 3). Their sensitivity to a novel disturbance makes sense, but sample size was too low to rule out chance.

Among females, GC responsiveness to acute vehicle exposure depended heavily on number of young and nutritional condition as measured by level of ft3 (Appendix D: Table D6). The lower the T3 in a female, the more she reduced fGCs relative to controls. In particular, non-breeding females in good nutritional condition (high ft3)

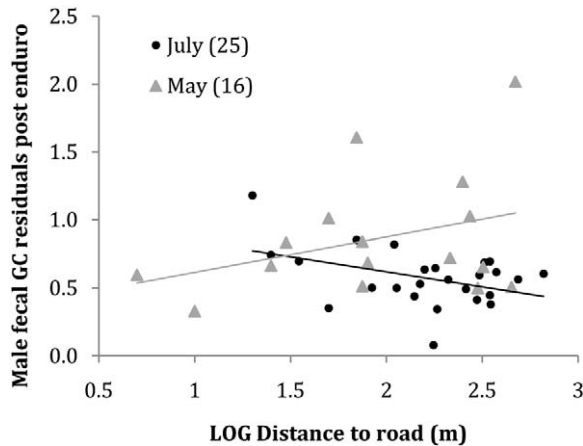


Fig. 4. Correlation between distance to road and peak GC post treatment for males split by month. Standard error is shown. Average fGC residuals were adjusted to make values positive for graphing. Samples sizes are shown in parentheses.

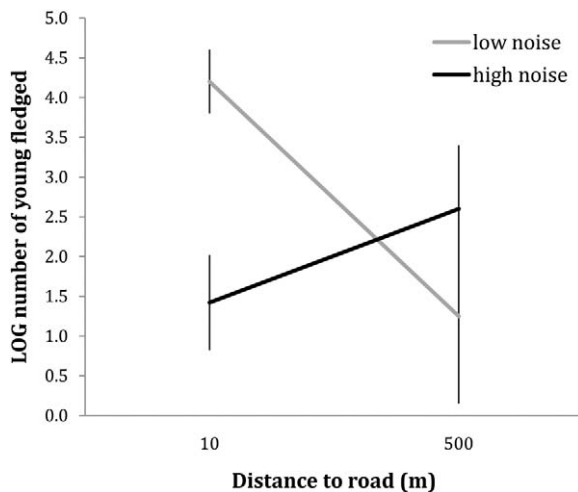


Fig. 5. Correlations between distance to road and NSO reproductive success on roads with relatively low and high levels of noise. Standard error is shown.

showed a very strong, significant increase in GCs in response to enduros (effect size = 0.53). In contrast, non-breeding females in poor nutrition (low fT3) showed an equally strong reversed treatment effect, with fGC levels lower in enduro-treated individuals than in controls (effect size = 0.54). These responses make sense when we consider that GC is low under good nutrition (reflected by high fT3) and high under

poor nutrition (reflected by low fT3). Thus, fGC should have been, and was, low in high fT3 pretreatment individuals (based on controls), allowing for a robust post-treatment stress response. By contrast, fGC should have been, and was, already high in pretreatment females with poor nutrition (low fT3). Further GC elevation under those conditions could have pushed the female into allostatic overload (McEwen and Wingfield 2003), whereby physiological coping mechanisms would be overwhelmed. Dampening of the GC response was even more dramatic in females with two young. Pre-treatment GCs should have been especially high in females protecting and provisioning two young under the constraints of low fT3, resulting in even more likelihood of allostatic overload among this subset. It is also possible that females with two young reacted more strongly than females with one or no young to the presence of researchers (Appendix C: Northern spotted owl glucocorticoid response to presence of researchers) and this too could have contributed to allostatic overload.

Previous studies have observed decreased GCs in response to perturbation among chronically stressed birds (Cyr and Romero 2008). Short-term elevations of GC levels are generally considered to be part of an adaptive emergency response to threat, [although this assumption lacks good empirical support; for review see Breuner et al. (2009)]. However, the GC response mobilizes energy from stores by promoting gluconeogenesis and may be maladaptive for individuals that need to conserve their few remaining reserves (Busch and Hayward 2009). Individuals with high allostatic load may reduce GCs so as to conserve energy while waiting out the perturbation. Interestingly, while high levels of GC are sometimes associated with reduced reproductive success (Bonier et al. 2009), in our study, some of the lowest average fGC residuals were seen among individuals with failed breeding attempts (Fig. 1).

#### Effects of road noise on GC response

There was no evidence that GC response to enduro diminished with exposure to routine road noise in May or among NSO within 50 m of a road in July. Traffic appeared always to be associated with high GC response in these NSO. The fact that male NSO 50–800 m from loud

roads showed lower GC response to acute motorcycle exposure than males 50–800 m from quiet roads in July suggests that partial habituation to traffic may occur in this species among males a sufficient distance (>50 m) from the road. Further study, however, is needed to verify these tentative findings. First, to better quantify routine traffic exposure for each pair, sound level meters should be set up on the road and under each favored roost or nest to account for sound propagation across steep, heavily vegetated terrain and to distinguish air traffic from road traffic. Video cameras along the road would also help to distinguish noise from road traffic from noise created by other sources. Second, we must determine whether a dampened GC response to acute motorcycle exposure represents true “habituation”, or a down-regulation of the entire hypothalamic-pituitary adrenal axis (Cyr and Romero 2009). Regardless of whether NSO habituate to disturbance from road traffic, we found exposure to road noise associated with a strong decrease in NSO reproductive success.

#### *Associations between proximity to road and baseline hormone metabolites*

The positive association of proximity to roads and level of fT3 suggests that NSO in our study population eat better when close to a road, particularly during July (Appendix D: Table D9). Prey (e.g., wood rats) may be more abundant, or easier to locate and capture near roads. This suggestion is consistent with work by Franklin et al. (2000) on a neighboring population of NSO, which showed that reproductive success is highest on sites with early successional forest. A similar pattern is not expected farther north where younger forest provides no fitness advantage for NSO (Dugger et al. 2005).

In contrast to Wasser et al. (1997), we found no association between proximity to road and baseline fGC levels in this population of NSO. Wasser et al. (1997) conducted their study on NSO in Washington and Oregon where NSO depend primarily on flying squirrels for food. Flying squirrels are old growth specialists and likely avoid edge. The difference in our results may be due to an advantage of proximity to road in our study population (implied by higher fT3) that is not conferred to NSO farther north. This advantage of proximity to road may offset other

potential disadvantages that would result in increased baseline glucocorticoids.

#### *Proximity to road, road noise and reproductive success*

While road proximity alone showed no association with number of young fledged, once noise level on that road and the interaction of proximity and noise were included in analysis (Appendix D: Table D11; Fig. 5) proximity to road noise significantly affected reproductive success. NSO within 100 m of quiet roads fledged more young than NSO further from roads (within a range of 800 m). In contrast, NSO within 100 m of loud roads fledged fewer young. For owls within 100 m of a road the association of high noise and reduced reproductive success is strong.

Although our original intent was to use sound level meters to quantify routine vehicle traffic on all roads, limited resources prevented analysis to exclude all other noises from the sound level metric. Therefore, our measure (other  $AC_{Aeq}$ ) was not completely cleared of aircraft noise and intermittent noises from natural sources, although most animal noises (sustained cicada song and passerine choruses) were excluded. Thus, our measure consists of mostly road traffic, though it does include other random man-made sounds. The fact that it has such significant effects on number of young fledged by NSO close to roads clearly suggests that traffic has a negative impact on NSO reproductive success.

Causality is impossible to determine with a correlative study. However, the combination of associations we found (e.g., proximity to road and high fT3; increased reproductive success close to quiet roads; reduced reproductive success near loud roads; increased fGCs in response to acute vehicle exposure) suggests that the NSO in our study population may be drawn to roads for better forage, thereby becoming more susceptible to the deleterious effects of vehicle exposure. Elucidating the mechanisms of impact was beyond the scope of this study. It may be that traffic noise interferes with the NSO's ability to detect predators. Great horned owls (*Bubo virginianus*), a primary predator of juvenile NSO, prefer forest edge, like that provided by roads. The impacts of road noise on Great horned owls are as yet unknown.

### Implications

Our results have several implications for conservation management and for the theoretical framework of stress physiology. First, we found that, overall, spotted owls increase fGCs in response to acute traffic exposure. This is consistent with the elevated baseline fGCs found among male NSO with territories close to roads in Washington (Wasser et al. 1997), and suggests that traffic acts as a stressor on the NSO. Further, the highest sensitivity appears to occur among males in May when they are the sole providers for their mates and offspring. This suggests that spring may be a particularly important time to limit motorized recreation near NSO territories. Finally, in May, male GC response to enduro did not diminish with distance to road within the 800 m range in which our experiments occurred (Fig. 4). This outcome is surprising given that vehicle noise is often not detectable over such a range, and certainly would not have been high at ranges greater than 200–400 m. In our sample, the number of roosts far from roads was limited, so the lack of a clear cutoff in the GC response could be the result of a sampling problem. Nevertheless, it may be that the owls reacted to road noise at the limits of detection, i.e., that their response was the result of detecting the disturbance rather than reacting to the noise itself. This suggests that managers should take a precautionary approach to motorized traffic within 800 m of NSO territories until better information is obtained. Although our study did not include many juveniles, our results also suggest that juveniles may be especially vulnerable to disturbance. Thus, limiting human activity near territories of NSOs with young may be wise.

Conservation endocrinology is complicated by the fact that elevated baseline GCs can be positively, negatively or not associated with survival and/or reproduction (Busch and Hayward 2009). An acute elevation of GCs in response to an experimentally applied stimulus is less ambiguous and generally acknowledged a marker of disturbance (McEwan and Wingfield 2003). Our results, however, emphasize that the magnitude and direction of response may vary with life history stage and body condition. Decreased GCs in response to experimentally applied stimuli has been occasionally documented (Cyr and Romero 2007) and consistently

associated with chronic, debilitating stress. The fact that females with two young and low T3 showed a decrease in fGCs in response to acute traffic exposure indicates that stress from traffic may cause serious detriment to NSO with high allostatic load. The relationship between T3 and GC responsiveness also suggests that impacts from roads may be lessened when habitat quality is higher, consistent with the results of Franklin et al. (2000) and with associations of reduced GC responsiveness in individuals with high body condition—at least in the non-breeding season (Wingfield et al. 1994). The results of our study are the first that we know of to reveal the full spectrum of GC response to disturbance within a single population. They underline the importance of context when interpreting endocrine data with management applications. Had we not had information about T3, we would have concluded that there was no GC response to acute traffic exposure among female NSO. However, more careful analysis revealed both a strong increase among females with high nutritional reserves and a strong decrease among females with low reserves. We recommend that future studies of GC response to disturbance fully consider other proxies of allostatic load.

For any study with conservation or management implications the question of how disturbance impacts lifetime fitness and population health is critical (Tarlow and Blumstein 2007). We found that acute vehicle exposure increased fGCs in NSO. This finding, together with the strong association of decreased reproductive success near loud roads, suggests that disturbance from traffic can have a significant impact on the fitness of the NSO.

### ACKNOWLEDGMENTS

Don Amador of the Blue Ribbon Coalition helped with experiments. Rebecca Nelson Booth, Lynn Erkman and Betsy Walker helped with lab work. Crew leaders for field work included Zachary Folk, Jennifer Hartman, Erica Capuana, Todd Bush, Aimee Reiss, Eric DeRoche, Jay Linke, Yoav Bar-Ness, Heath Smith, Chris Zeiminski and volunteer interns Jordan Aldersley, Jordan Birch, Ann Brunelle, Chase Cammarato, Rayma Cooley, Ben Cutler, Dustin Detweiler, Garrett Dowgiallo, Blake Etchison, Matt Gilbert, Shane Gillett, Ruben Gutstein, Melissa Habenict, Ryan Hampshire, Cameron Hastie, Andrew Henderson, Todd Histan, Wes Hoyer, Steffani Jijon, Nicky Johnston, Geneva

Karwoski, Nicole Koeltzow, Bud Marks, Ben McClain, Catherine Nitz, Takahiro Ota, Nick Politte, Tatiana Ramirez, Kurt Rosin, Melissa Rothrock, Kate Rowe, Kristen Sani, Laura Walls, Michelle Welsley and David Wolfson. HSWRI staff members Tina Yack and Samuel Denes helped with noise measurements. The Forest Service funded our work through a grant from the California Off-Highway Motorized Vehicle Commission under agreement number 05-JV-11052007-028. Support also came from NSF award # IOS-0719589, as well as from a grant from the American Motorcycle Association. USDA Forest Service support came from Diana Craig, Peggy O'Connell, Kelly Wolcott, Dennis Garrison, Mark Goldsmith, Paula Crumpton, Becky Rogers, Jim Ruhl, Helayna Cooley, Lee Morgan and Marcus Spivey. US Fish and Wildlife Service support came from Lyle Lewis, Rob Emgee, Ron Clemenston, Doug Powers, Heidi Crowell and Jen Ballard.

## LITERATURE CITED

- Arlettaz, R., P. Patthey, M. Baltic, T. Leu, M. Schaub, R. Palme, and S. Jenni-Eirmann. 2007. Spreading free-riding snow sports represent a novel serious threat for wildlife. *Proceedings of the Royal Society Biological Sciences Series B* 274:1219–1224.
- Barber, J. R., K. R. Crooks, and K. M. Fristrup. 2009. The costs of chronic noise exposure for terrestrial organisms. *Trends in Ecology and Evolution* 25:180–189.
- Barja, I., G. Silván, S. Rosellini, A. Piñeiro, A. González-Gil, L. Camacho, and J. C. Illera. 2007. Stress physiological responses to tourist pressure in a wild population of European pine marten. *Journal of Steroid Biochemistry and Molecular Biology* 104:136–142.
- Bonier, F., P. R. Martin, I. T. Moore, and J. C. Wingfield. 2009. Do baseline glucocorticoids predict fitness? *Trends in Ecology and Evolution* 24:634–642.
- Bonier, F., P. Martin, K. S. Sheldon, J. P. Jenson, S. L. Foltz, and J. C. Wingfield. 2007. Sex-specific consequences of life in the city. *Behavioral Ecology* 18:121–129.
- Boonstra, R., D. Hik, G. R. Singleton, and A. Tinnikov. 1998. The impact of predator-induced stress on the snowshoe hare cycle. *Ecological Monographs* 68:371–394.
- Bosworth, D. 2004. Four threats to the Nation's forests and grasslands. Speech. Idaho Environmental Forum. Boise, Idaho. ([www.fs.fed.us/news/2004/speeches/01/idaho-four-threats.shtml](http://www.fs.fed.us/news/2004/speeches/01/idaho-four-threats.shtml))
- Breuner, C. W., S. H. Patterson, and T. P. Hahn. 2009. In search of relationships between the acute adrenocortical response and fitness. *General and Comparative Endocrinology* 157:288–295.
- Busch, D. S. and L. S. Hayward. 2009. Stress in a conservation context: A discussion of glucocorticoid actions and how levels change with conservation-relevant variables. *Biological Conservation* 142:2844–2853.
- Busch, D. S., E. A. Addis, A. D. Clark, and J. C. Wingfield. 2010. Disentangling the effects of environment and life-history stage on corticosterone modulation in Costa Rican Rufous-collared sparrows, *Zonotrichia capensis costaricensis*. *Physiological and Biochemical Zoology* 83:87–96.
- Clinchy, M., L. Zanette, R. Boonstra, J. C. Wingfield, and J. N. M. Smith. 2004. Balancing food and predator pressure induces chronic stress in songbirds. *Proceedings of the Royal Society of London: B* 271:2473–2479.
- Cordell, K. H., C. J. Betz, G. Green, and M. Owens. 2005. Off-Highway Vehicle Recreation in the United State, Regions and States. A National Report from the National Survey on Recreation and the Environment (NSRE). ([www.fs.fed.us/recreation/programs/ohv/OHV\\_final\\_report.pdf](http://www.fs.fed.us/recreation/programs/ohv/OHV_final_report.pdf))
- Courtney, S. P., J. A. Blakesley, R. E. Bigley, M. L. Cody, J. P. Dumbacher, R. C. Fleischer, A. B. Franklin, J. F. Franklin, R. J. Gutierrez, J. M. Marzluff, and L. Sztukowski. 2004. Scientific evaluation of the status of the northern spotted owl. Sustainable Ecosystems Institute, Portland, Oregon, USA.
- Creel, S., J. E. Fox, A. Hardy, J. Sands, B. Garrott, and R. O. Peterson. 2002. Snowmobile activity and glucocorticoid stress responses in wolves and elk. *Conservation Biology* 16:809–814.
- Cyr, N. E., and L. M. Romero. 2007. Chronic stress in free-living European starlings reduces corticosterone concentrations and reproductive success. *General and Comparative Endocrinology* 151:82–89.
- Cyr, N. E., and L. M. Romero. 2008. Fecal glucocorticoid metabolites of experimentally stressed captive and free-living starlings: Implications for conservation research. *General and Comparative Endocrinology* 158:20–28.
- Cyr, N. E., and L. M. Romero. 2009. Identifying hormonal habituation in field studies of stress. *General and Comparative Endocrinology* 161:295–303.
- Douyon, L. and D. E. Scheingart. 2002. Effect of obesity and starvation on thyroid hormone, growth hormone, and cortisol secretion. *Endocrinology and Metabolism Clinics of North America* 31:173–189.
- Dugger, K. M., F. Wagner, R. G. Anthony, and G. S. Olson. 2005. The relationship between habitat characteristics and demographic performance of northern spotted owls in southern Oregon. *The Condor* 107:863–878.
- Eales, J. G. 1988. The influence of nutritional state on thyroid function in various vertebrates. *American Zoology* 28:351–362.

- Flier, J., M. Harris, and A. N. Hollenberg. 2000. Leptin, nutrition, and the thyroid: the why, the wherefore, and the wiring. *Journal of Clinical Investigation* 105:859–861.
- Forsman, E. D. 1983. Methods and materials for locating and studying spotted owls. General Technical Report PNW-162. United States Forest Service, Corvallis, Oregon, USA.
- Franklin, A. B., D. R. Anderson, R. J. Gutierrez, and K. P. Burnham. 2000. Climate, habitat quality, and fitness in northern spotted owl populations in northwestern California. *Ecological Monographs* 70:539–590.
- Franklin, A. B., P. C. Carlson, J. T. Rockweit, A. N. Canter, A. Roberts, N. Van Lanen, J. Lux, E. J. Sisson, A. Carrington Knox, and J. C. White. 2008. Monitoring the population ecology of spotted owls (*Strix occidentalis caurina*) in Northwestern California. Annual results, 2007. Progress report. U.S. Forest Service, Region 5, San Francisco, California, USA.
- Geris, K. L., L. R. Berghman, E. R. Kuhn, and V. M. Darras. 1999. The drop in plasma thyrotropin concentrations in fasted chickens is caused by an action at the level of the hypothalamus: The role of corticosterone. *Domestic Animal Endocrinology* 16:231–237.
- Gutzwiller, K. J. 1991. Assessing recreational impacts on wildlife: The value and design of experiments. Pages 249–255. *in* Transactions of the 56th National Academy of Wildlife and Natural Resources Conference.
- Hayward, L. H., R. K. Booth, and S. K. Wasser. 2010. Eliminating the artificial effect of sample mass on avian fecal hormone metabolite concentration. *General and Comparative Endocrinology* 169:117–122.
- Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–212.
- JMP Development Team. 2007. JMP. Version 7. SAS Institute Inc., Cary, North Carolina, USA.
- Kitaysky, A. S., M. D. Romano, J. F. Piatt, J. C. Wingfield, and M. Kikuchi. 2005. The adrenocortical response of tufted puffin chicks to nutritional deficits. *Hormones and Behavior* 47:606–619.
- Korte, S. M., J. M. Koolhaas, J. C. Wingfield, and B. S. McEwen. 2005. The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. *Neuroscience and Biobehavioral Reviews* 29:3–38.
- McEwen, B. and J. C. Wingfield. 2003. The concept of allostasis in biology and biomedicine. *Hormones and Behavior* 43:2–15.
- Norris, D. O. 1997. *Vertebrate endocrinology*. Third Edition. Academic Press, San Diego, California, USA.
- Pride, R. E. 2005. High faecal glucocorticoid levels predict mortality in ring-tailed lemurs (*Lemur catta*). *Biology Letters* 1:60–63.
- Romero, L. M. 2004. Physiological stress in ecology: Lessons from biomedical research. *Trends in Ecology and Evolution* 19:249–255.
- Sapolsky, R. M., L. M. Romero, and A. U. Munck. 2000. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrinology Review* 21:55–89.
- Schew, W. A., F. M. A. McNabb, and C. G. Scanes. 1996. Comparison of the ontogenesis of thyroid hormones, growth hormone, and insulin-like growth factor-I in ad libitum and food-restricted (altricial) European starlings and (precocial) Japanese quail. *General and Comparative Endocrinology* 101:304–316.
- Suorsa, P., E. Huhta, A. Nikula, M. Nikinmaa, A. Jantti, H. Helle, and H. Hakkarainen. 2003. Forest management is associated with physiological stress in an old-growth forest passerine. *Proceedings of the Royal Society B* 270:963–969.
- Tarlow, E. M. and D. T. Blumstein. 2007. Evaluating methods to quantify anthropogenic stressors on wild animals. *Applied Animal Behaviour Science* 102:429–451.
- Thiel, D., S. Jenni-Eirmann, V. Braunisch, R. Palme, and L. Jenni. 2008. Ski tourism affects habitat use and evokes a physiological stress response in capercaillie *Tetrao urogallus*: a new methodological approach. *Journal of Applied Ecology* 45:845–853.
- Walker, B. E., P. D. Boersma, and J. C. Wingfield. 2006. Habituation of adult magellanic penguins to human visitation as expressed through behavior and corticosterone secretion. *Conservation Biology* 20:146–154.
- Walpita, C. N., S. V. H. Grommen, V. M. Darras, and S. Van der Geyten. 2007. The influence of stress on thyroid hormone production and peripheral deiodination in the Nile tilapia. *General and Comparative Endocrinology* 152:18–25.
- Wasser, S. K., J. C. Azkarate, R. K. Booth, L. S. Hayward, K. Hunt, K. Ayres, C. Vynne, K. Gobush, D. Canales-Espinosa, and E. Rodríguez-Luna. 2010. Non-invasive measurement of thyroid hormone in feces in a diverse array of avian and mammalian species. *General and Comparative Endocrinology* 168:1–7.
- Wasser, S. K., K. Bevis, G. King, and E. Hanson. 1997. Noninvasive physiological measures of disturbance in the Northern Spotted Owl. *Conservation Biology* 11:1019–1022.
- Wasser, S. K. and K. E. Hunt. 2005. Noninvasive measures of reproductive function and disturbance in the Barred Owl, Great Horned Owl, and Northern Spotted Owl. *Annals of the N.Y. Acad-*

- emy of Sciences 1046:109–137.
- Wasser, S. K., K. E. Hunt, J. L. Brown, K. Cooper, C. M. Crockett, U. Bechert, J. J. Millspaugh, S. Larson, and S. L. Monfort. 2000. A generalized fecal glucocorticoid assay for use in a diverse array of nondomestic mammalian and avian species. *General and Comparative Endocrinology* 120:260–275.
- White, C. M. and T. L. Thurow. 1985. Reproduction of ferruginous hawks exposed to controlled disturbance. *The Condor* 87:14–22.
- Wikelski, M., V. Wong, B. Chevalier, N. Rattenborg, and H. L. Snell. 2002. Marine iguanas die from trace oil pollution. *Nature* 417:607–608.
- Wingfield, J. C. and L. M. Romero. 2001. Adrenocortical responses to stress and their modulation in free-living vertebrates. Pages 211–234 in B. S. McEwen, editor. *Handbook of physiology*. Oxford University Press, New York, New York, USA.
- Wingfield, J. C., P. Deviche, S. Sharbaugh, L. B. Astheimer, R. Holberton, R. Suydam, and K. Hunt. 1994. Seasonal changes in the adrenocortical responses to stress in redpolls, *Acanthis flammea*, in Alaska. *Journal of Experimental Zoology* 270:372–380.
- Yarmoloy, C., M. Bayer, and V. Geist. 1988. Behavior responses and reproduction of mule deer, *Odocoileus hemionus*, does following experimental harassment with an all-terrain vehicle. *Canadian Field Naturalist* 102:425–429.

## APPENDIX A

### METHODS FOR SOUND LEVEL MEASUREMENT COLLECTION AND ANALYSIS

#### *Collecting sound level measurements*

We used Larson-Davis 820 Type I sound level meters (SLMs) connected to ACO 7013 microphones to record noise on the roads closest to each NSO territory. Data were collected as continuous 2-second A-weighted time histories. Instruments were calibrated at the beginning and end of each recording session using Larson-Davis CAL200 calibrators. To quantify routine road noise, SLM equipment was set to run for a period of approximately 96 hours, encompassing two week and two weekend days. SLMs were set up within 5–15 meters of the road at the point determined by global information system (GIS) and global positioning system (GPS) to be closest to the owls' favored roost or nest. The receiving microphone was mounted at 1.25 m and oriented in the direction of the road. Equipment was obscured in naturally occurring foliage to prevent tampering, but care was taken to ensure a clear line of sight between the microphone and the road.

A number of acoustic metrics were collected from the SLM data, including event-related metrics (e.g., maximum 2-second equivalent continuous level [ $L_{Aeq}$ ] and sound exposure level [SEL]) and measures of cumulative exposure (e.g., observation  $L_{Aeq}$ , SEL, level percentiles, and duration of exposure above a criterion level). There is currently no agreement about the

metrics that provide the best estimators of exposure. Therefore, a secondary goal of the study was to determine which sound metrics explained the most variance in hormonal response.

Unfortunately, limited funds and access to SLMs meant that noise from traffic could not be measured both on the road and at the nest site, nor could observers be maintained at the SLM sites to identify noise sources. Noise sources were identified using characteristic time-history profiles based on experience with ground-truthed sources (A. Bowles, *personal observation*). This method is reliable when vehicles pass close to the SLM and background noise is low (i.e., when the signal-to-noise ratio [SNR] is high), but vehicle traffic could not always be completely isolated from non-traffic noise when the signal to noise ratio (SNR) was low. Under windy conditions or at sites where insects and water noise were dominant, vehicle passes were likely to be missed. Also, in the absence of spectral data, discrete noise sources were sometimes difficult to differentiate from one another. For example, commercial jet aircraft traffic was a substantial source at most locations, but could not be reliably distinguished from distant vehicles, especially if the background levels were already relatively high. Sources that were identifiable like commercial jet overflights and short bouts of biotic noise were marked and eliminated from the background noise calculation to the extent practicable, given limited resources for processing the data. Regardless, the remaining “ambient” data were probably contaminated with some natural

and human-made sources. Because the background noise measure is the best available measure of routine traffic on each site and because it turned out to be a highly significant predictor of NSO physiology and reproductive output, we have included these results in our final report, but chose to be conservative in our conclusions.

#### *Sound level measurement statistics*

After preliminary examination of the collected noise metrics, three were included in our statistical models as follows:

- $L_{Aeq}$ , the equivalent continuous sound level of the simulated enduro and control periods, a measure of the average sound level at the site.
- SEL, the sound exposure level of the cumulative events, a measure of total energy over a given measurement period.

Because sound level is measured on a logarithmic scale (the decibel [dB]), both  $L_{Aeq}$  and SEL are disproportionately influenced by high-amplitude sounds. A single event lasting a few seconds can produce an hourly value that is similar to the cumulative influence of many low amplitude events. In order to obtain a measure that captured the total duration of exposure to moderately high levels, a measure of exceedence time was also calculated.

- TA60 the proportion of the total time sound level exceeded a threshold of 60 dB.

The TA60 criterion was chosen as a measure of the proportion of high amplitude noise present at each site in the absence of experimental exposures. The criterion was chosen after conducting

a sensitivity analysis with criterion levels ranging from 55 to 66 dBA. All exceedence durations calculated using these levels were very highly correlated with each other. The 60 dBA value was eventually selected because it was in the middle of the range tested and was exceeded by only a few percent of the 2-second samples at most sites in the absence of vehicle noise (less than 1% at 22 of 23 sites in 2006 and 13 of 19 sites in 2007). Average ambient levels were generally between 25 dBA and 50 dBA (ambient  $L_{Aeq}$  for the observation period). In the exceptional cases, persistent insect noise was a likely factor, but observers were rarely present to identify the sources.

We used non-enduro transient noise events (Other  $L_{Aeq}$ ) as our best measure of routine road noise from discrete traffic passes isolated as much as possible from background noise produced by sources like water, wind and cicadas. As noted previously, it was not a perfect measure of typical traffic exposure for a site because it could not be completely separated from airplane passes or other discrete noise sources. Hereafter, it will be referred to as ‘road noise’, without specific reference to the sources.

Only a subset of sites had SLM measures for four consecutive days that could be used in the correlative analysis (23 in 2006 and 7 in 2007). Year and site number were included in the analysis to control for pseudoreplication. Distance to road was also included as an important co-variate, given that an owl with a roost distant from a road would experience substantially lower exposure to road noise than an owl with a roost close to the road.

## APPENDIX B

### LABORATORY METHODS

#### *Laboratory methods*

Measuring hormones from avian scat is a relatively new technique and our study involved extensive laboratory validation and associated modifications to the sample processing methodology (Hayward et al. 2010). Measuring hormone levels in small avian scats like those from the NSO has been frequently complicated by a strong

artificial negative correlation between scat mass and hormone concentration (Cyr and Romero 2008, Goymann 2005, Goymann et al. 2002, Millspaugh and Washburn 2004, Tempel and Gutierrez 2004, Washburn et al. 2004). We were able to apply an optimized methodology to samples extracted in 2007 and 2008. However, mass effects could only be statistically controlled in samples extracted in 2005 and 2006, due to the fact that these samples were extracted with a lower volume of extractant solution (Hayward et

al. in press). Therefore, after testing distribution and applying appropriate LOG transformations, we calculated residuals of both fGCs and fT3 for all samples from a linear regression of sample mass and hormone metabolite concentration split by year in order to control for extraction differences and associated mass effects across years. Each regression excluded all samples collected post enduro treatment. This methodology disallowed testing for year effects but did allow multiple years to be combined.

#### *Fecal steroid extraction and assay*

Fecal samples were cleared of urates as much as possible, weighed, extracted and assayed using methods described in Wasser et al. (2000),

Wasser and Hunt (2005), Wasser et al. (2010), and Hayward et al. (2010). We measured fGCs using an I<sup>125</sup> corticosterone kit from ICN Biomedicals, Costa Mesa, California (ICN # 07-120103) following supplier's instructions, but halving the volume of all reagents (Wasser et al. 2000). This kit was previously validated for spotted owl feces (Wasser et al. 1997, Wasser and Hunt 2005). Samples were assayed for fT3 using the I<sup>125</sup> T3 radioimmunoassay kit by MP Biomedicals (# 06-254282) and standards made in the Wasser lab, previously validated for use with the NSO (Wasser et al. 2010). Interassay variation was 6.3% for fGCs and 12.0% for fT3. Intra-assay variation was 3.2% and 2.4% respectively.

## APPENDIX C

### NORTHERN SPOTTED OWL GLUCOCORTICOID RESPONSE TO PRESENCE OF RESEARCHERS

#### *Methods for test of researcher effect*

Sample collections required researchers to spend up to 16 hours within approximately 15 m of occupied roosts or nests in order to acquire pre- and post-treatment samples on the days of experiments. Thus, an important first question was whether NSO showed a physiological response to prolonged researcher presence on their territory, and if so, whether the response was large enough to swamp any effect of experimental treatment. Final analysis of researcher effect included 136 observations for which time of researcher arrival and time of scat production were precisely recorded. Although we included samples collected on the days of experiments in this analysis, all samples produced post-enduro were excluded to eliminate the effects of vehicle exposure. Many samples were collected on non-experiment days and time of arrival on territory varied from pre-dawn to just before sundown. Time spent by researchers on territory before scat collection varied from five to 972 minutes (over 16 hours). Biologically relevant covariates (e.g., sex, number of young) and interactions were included as covariates in a general linear model of LOG fecal GCs with minutes of researcher on territory as a main effect and dropped when non-significant.

#### *Results*

There was a significant interaction of number of young and LOG minutes of researcher presence on territory, with a very small effect size (Table C1). A linear regression of LOG fGCs and time on territory split by number of young showed that adults with two offspring had a significant increase in fGCs with time of researcher presence (Table C2). However, the effect size was small enough to be considered negligible by Cohen's criteria and, more importantly, too small to swamp the treatment effect. There was no effect of researcher presence on males, non-breeding females or females with one juvenile. Juveniles themselves showed a moderate GC

Table C1. Results of general linear model testing for effect of researcher presence on territory (n = 136). Small p values suggest that there is a significant effect of number of young on level of fGCs and that there may be trend for number of young to interact with time of researcher presence on territory to influence fGC level. However, very small effect sizes indicate that these effects are negligible.

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Sex	0.07	0.02	0.02	0.16
Number of young	<b>0.02</b>	<b>0.07</b>	<b>0.08</b>	<b>0.002</b>
LOG Researcher time on territory	0.09	0	0	0.24
Number of young × time on territory	0.07	0.02	0.02	0.07
Overall model	<b>0.09</b>		<b>0.10</b>	<b>0.005</b>

Table C2. Level of fGC: a summary of a linear regression of LOG fGC residuals and time of researcher presence on territory split by number of young. Adult NSO with two young showed a slight significant positive relationship, while NSO with one or zero young did not. There was also a trend for juveniles to show a small increase in fGCs in response to researcher presence.

Factor	N	R <sup>2</sup>	Effect size	P
Zero young	72	-0.01	-0.01	0.73
One young	15	-0.06	-0.06	0.65
Two young	<b>36</b>	<b>0.13</b>	<b>0.15</b>	<b>0.02</b>
Juveniles	<b>13</b>	<b>0.22</b>	<b>0.28</b>	<b>0.06</b>

response to researcher presence that bordered on significance despite the small sample size (n = 13).

*Discussion of researcher effect*

The lack of an effect of researchers on GC level in all adult NSO except females with two young (Table C2) and the very small effect size in that subset suggest that we were not likely to swamp the GC response to motorcycles with a strong GC response to researchers. Nonetheless, female

NSO with two young and juveniles did show signs of disturbance from human presence on their territory. Although the effect size was very small, hormone values were LOG transformed. Additionally, small changes in levels of circulating hormone can have large effects on physiology and behavior (Norris 1997). From an evolutionary perspective it makes sense that the limited mobility of juveniles might make them and the females that care for them most sensitive to unfamiliar, ground-based noises.

Researchers in this study took care to stay as still and work as silently as possible. More disruptive human activities could well have a strong effect on NSO GC levels. A precautionary approach would minimize disturbance on this threatened species so we recommend that managers take care to limit human activity on NSO territories with young. In the future, the method we have described could be used to more precisely measure the range at which GC responses become detectable. These methods could also be used to develop valid behavioral measures of disturbance intensity, making it easier for researchers and managers to prevent disturbance.

APPENDIX D

Table D1. Sample size.

Year	Group	May		July	
		Control	Enduro	Control	Enduro
2005	Males	1	6	8	8
	Females	2	3	4	8
	Juveniles	0	0	1	0
2006	Males	6	4	10	9
	Females	4	2	8	8
	Juveniles	0	0	4	4
2007	Males	0	0	6	8
	Females	0	0	5	4
	Juveniles	0	0	1	3
2008	Males	6	8	2	4
	Females	5	5	1	2
	Juveniles	0	0	0	3
Total		24	28	50	62
Grand total					165

Table D2. Average fGCs Post treatment: exposure to an hour-long simulated enduro resulted in a slight but significant increase in NSO fGC levels. However, post treatment fGC level varied with month, sex and number of young. n = 123. Statistical significance is indicated in bold.

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Month	<b>0.20</b>	<b>0.02</b>	<b>0.03</b>	<b>0.05</b>
Sex	<b>0.19</b>	<b>0.03</b>	<b>0.05</b>	<b>0.02</b>
Number of young	<b>0.18</b>	<b>0.04</b>	<b>0.06</b>	<b>0.02</b>
Treatment	<b>0.18</b>	<b>0.04</b>	<b>0.06</b>	<b>0.02</b>
Month × treatment	0.22	0	0	0.85
Sex × treatment	<b>0.18</b>	<b>0.04</b>	<b>0.06</b>	<b>0.009</b>
Month × sex	0.21	0.01	0.02	0.15
Month × treatment × sex	<b>0.19</b>	<b>0.03</b>	<b>0.05</b>	<b>0.03</b>
Overall model	<b>0.22</b>		<b>0.28</b>	<b>&lt;0.0001</b>

Table D3. Treatment effect on average post enduro fGCs in males: the effect of simulated enduro treatment was stronger on males in May than in July. Statistical significance is indicated in bold. Number of young was not significant as a main effect nor as an interaction with treatment.

Month	N	R <sup>2</sup>	Effect size	P
May	<b>28</b>	<b>0.21</b>	<b>0.27</b>	<b>0.008</b>
July	<b>44</b>	<b>0.16</b>	<b>0.19</b>	<b>0.004</b>

Table D4. Post enduro fGCs among females in May. Number of young strongly predicted post treatment fGC levels among females in May while treatment did not. Females with young had higher levels of fGCs than non-breeding females. Additionally, there was a slight trend for females with young to respond more strongly to simulated enduro than females without young. n = 17. Statistical significance is indicated in bold.

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	p
Number of young	<b>0.07</b>	<b>0.38</b>	<b>0.74</b>	<b>0.006</b>
Treatment	0.42	0.03	0.09	0.20
Number of young × treatment	0.34	0.11	0.30	0.08
Overall model	<b>0.45</b>		<b>0.82</b>	<b>0.01</b>

Table D5. There was no effect of hour long simulated enduro on average fGC levels post treatment among females in July. Neither number of young nor treatment predicted post treatment fGC levels. Although there was a trend for females without young to have lower fGCs post-enduro than females with young, the effect size was negligible. n = 35.

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Number of young	0.07	0.02	0.02	0.23
Treatment	0.08	0.01	0.01	0.28
Number of young × treatment	0.02	0.07	0.08	0.07
Overall model	0.09		0.10	0.12

Table D6. Female average fGC post enduro: while treatment alone did not affect fGCs among females, level of ft3 and number of young interacted significantly with treatment to predict fGC level. Post-enduro fGC levels were highest among females with high ft3 and two young. Females with young and low ft3 actually showed a reversed treatment effect with lower fGCs post enduro than controls (n = 22).

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Month	<b>0.52</b>	<b>0.10</b>	<b>0.55</b>	<b>0.04</b>
Treatment	0.64	-0.02	-0.15	0.71
T3 pre average	<b>0.50</b>	<b>0.12</b>	<b>0.63</b>	<b>0.03</b>
T3 × treatment	<b>0.06</b>	<b>-0.68</b>	<b>1.69</b>	<b>&lt;0.0001</b>
Number young	<b>0.14</b>	<b>0.48</b>	<b>1.47</b>	<b>0.0003</b>
Number young × treatment	<b>0.28</b>	<b>0.34</b>	<b>1.24</b>	<b>0.001</b>
Overall model	<b>0.62</b>		<b>1.63</b>	<b>0.001</b>

Table D7. Post enduro average fGC: on sites between about 50–800 m from a road GC response to simulated enduro decreased as level of routine road noise on the site increased. Within 50 m of the road, level of routine noise had no effect on GC response to enduro (n = 26).

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Distance to road	0.24	0.00	0.00	0.36
Other AC <sub>Aeq</sub>	0.16	0.08	0.13	0.08
Distance from road × other AC <sub>Aeq</sub>	<b>0.09</b>	<b>0.15</b>	<b>0.22</b>	<b>0.03</b>
Overall model	<b>0.24</b>		<b>0.32</b>	<b>0.03</b>

Table D8. Baseline fGCs: there was no effect of distance to road on baseline fGCs (n = 82).

Factor	R <sup>2</sup>	F	P
Site number		0.99	0.51
Month		0.001	0.98
Sex		0.07	0.80
Distance from road		0.43	0.52
Sex × distance from road		0.64	0.43
Month × distance from road		0.40	0.53
Overall model	-0.05	0.92	0.61

Table D9. Baseline fT3: as distance from road decreased NSO fT3 levels increased. Levels of fT3 increased with proximity to road more steeply in July than in May (n = 77).

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Site number	0.08	0.27	0.45	0.05
Month	0.34	0.01	0.02	0.25
Sex	0.32	0.03	0.07	0.12
Distance from road	<b>0.25</b>	<b>0.10</b>	<b>0.21</b>	<b>0.04</b>
Sex × distance from road	0.33	0.02	0.05	0.35
Month × distance from road	<b>0.05</b>	<b>0.30</b>	<b>0.49</b>	<b>&lt;0.0001</b>
Sex × month × distance from road	0.28	0.08	0.15	0.07
Overall model	<b>0.35</b>		<b>0.54</b>	<b>0.03</b>

Table D11. Number of young fledged: distance to closest road and its interaction with noise on that road (Other AC<sub>Aeq</sub>) in July were strongly associated number of young fledged. Proximity to quiet roads correlated with relatively high reproductive success while proximity to loud roads correlated with low success (n = 42).

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Site number	<b>0.10</b>	<b>0.83</b>	<b>13.17</b>	<b>&lt;0.0001</b>
Year	<b>0.91</b>	<b>0.02</b>	<b>3.17</b>	<b>0.04</b>
Other AC <sub>Aeq</sub>	0.93	0	0	0.59
Distance from closest road	<b>0.58</b>	<b>0.35</b>	<b>11.9</b>	<b>&lt;0.0001</b>
Distance from road × other AC <sub>Aeq</sub>	<b>0.78</b>	<b>0.15</b>	<b>9.74</b>	<b>&lt;0.0001</b>
Overall model	<b>0.93</b>		<b>13.29</b>	<b>&lt;0.0001</b>

Table D10. Baseline fT3 in July: levels of fT3 in adult NSO increased with proximity to road most steeply during July (both sexes combined; n = 56).

Factor	R <sup>2</sup>	Δ R <sup>2</sup>	Δ Effect size	P
Site number	<b>0.04</b>	<b>0.35</b>	<b>0.60</b>	<b>0.06</b>
Distance from road	<b>-0.02</b>	<b>0.41</b>	<b>0.56</b>	<b>0.0007</b>
Overall model	<b>0.39</b>		<b>0.64</b>	<b>0.047</b>

## APPENDIX E

### Females

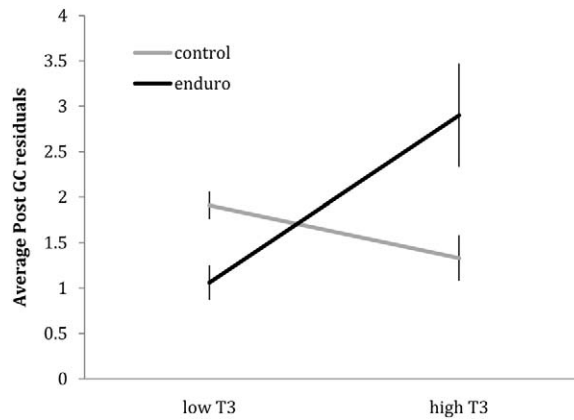


Fig. E1. Female GC response to simulated enduro split by relative level of T3. Standard error is shown. Average residuals were adjusted to make values positive for graphing.