The Journal of Wildlife Management; DOI: 10.1002/jwmg.21594

Research Article

Evaluating Response Distances to Develop Buffer Zones for Staging Terns

MELISSA A. ALTHOUSE, *Department of Environmental and Forest Biology, State University of New York College of Environmental Science and Forestry, 1 Forestry Drive, Syracuse, NY 13210, USA*
JONATHAN B. COHEN,[1] *Department of Environmental and Forest Biology, State University of New York College of Environmental Science and Forestry, 1 Forestry Drive, Syracuse, NY 13210, USA*
SARAH M. KARPANTY, *Department of Fish and Wildlife Conservation, Virginia Tech, Blacksburg, VA 24061, USA*
JEFFREY A. SPENDELOW, *U.S. Geological Survey Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708, USA*
KAYLA L. DAVIS, *Department of Fish and Wildlife Conservation, Virginia Tech, Blacksburg, VA 24061, USA*
KATHARINE C. PARSONS, *Coastal Waterbird Program, Mass Audubon, 208 S Great Road, Lincoln, MA 01773, USA*
CRISTIN F. LUTTAZI, *Coastal Waterbird Program, Mass Audubon, 208 S Great Road, Lincoln, MA 01773, USA*

ABSTRACT Buffer zones, calculated by flight-initiation distance (FID), are often used to reduce anthropogenic disturbances to wildlife, but FID can vary significantly across life-history stages. We examined the behavioral effect of potential natural (gulls and shorebirds) and anthropogenic (pedestrians) disturbance sources to staging roseate (*Sterna dougallii*) and common tern (*S. hirundo*) flocks from July to September in 2014 and 2015 at Cape Cod, Massachusetts, USA. We estimated the proportion of the flock exhibiting different responses to potential disturbance sources as a function of distance, flock size, percent roseate terns, and local disturbance rates, using Bayesian zero-and-one inflated beta regression. The proportion of tern flocks responding to the presence of shorebirds by flying was low (0.01 ± 0.001 [SE]) and did not vary by distance or other covariates, whereas the proportion flying in response to gulls increased as distance decreased, with smaller flocks, and with flocks with a larger proportion of roseate terns being more sensitive to gull presence. Prolonged flight response rapidly increased in probability from 0.0 to as much as 1.0 as distance from pedestrians to the flock decreased from 100 m and was much more likely with smaller flocks. Pedestrian activity levels also had an effect on flock responses; those engaged in active behaviors such as jogging were more likely to cause flushing than those engaged in passive behaviors. Terns seemed to view pedestrians as more of a threat than shorebirds and gulls, even though gulls are frequent kleptoparasites of terns. Pedestrians >120 m from a tern flock generally elicited the same probability of flight response as shorebirds and gulls. We recommend managers maintain anthropogenic disturbance levels at or below the intensity of those from natural sources at sites where recreation and wildlife values are both important. Because staging tern flocks may use a variety of areas within a site, we recommend instituting a 100-m buffer around areas potentially used by staging flocks at Cape Cod, where we studied every location roseate terns are known to use in large numbers. For other sites used by mixed-species tern flocks, we recommend the use of our field and analytical methods to develop appropriate buffer distances that will keep pedestrians far enough away to reduce the likelihood of flight and other non-locomotive anti-predator behaviors. These buffer zones will also benefit other species sensitive to human activity.

KEY WORDS anti-predator behavior, buffer zones, Cape Cod, common tern, disturbance, flight initiation distance, roseate tern, staging.

Predator evasion tactics vary by species but often culminate in a flight response designed to transport an organism out of a predator's reach, or to make continued capture attempts too costly to pursue (Yager 2010). However, there are costs associated with antipredator behaviors in the form of direct energy expenditure and forfeiting opportunities to acquire resources (Endler 1991, Fernandez-Juricic and Rodriguez-Prieto 2010). Ultimately, flight initiation is the result of a cost-benefit tradeoff between wasting energy on an unnecessary response and not fleeing soon enough to escape a predator (Burger and Gochfeld 1981, Blumstein 2003).

Energy demands increase during critical life-history stages, such as the post-natal growth or pre-migratory periods (Kvist and Lindström 2001), when the conservation of energy becomes particularly important (Drent et al. 2003, Inger et al. 2008). The tradeoff between anti-predator behavior, and gathering and conserving resources becomes even more

Received: 1 April 2017; Accepted: 2 September 2018

[1] *E-mail: jcohen14@esf.edu*

significant at these times (Pomeroy 2006). The post-breeding staging period of migratory birds is one such critical time for individuals to build energy reserves before migration (Schauroth and Becker 2008). Human disturbance has the potential to limit the use of resting areas (Trull et al. 1999) and alter energy budgets (Schummer and Eddleman 2003) of migrating waterbirds during staging periods. Newton (2006) reviewed several examples of conditions experienced during migration that can limit populations of migratory birds, including weather, resource availability, and human disturbance. In the cases of the pink-footed goose (*Anser brachyrhynchus*) in northern Norway and greater snow goose (*Anser caerulescens atlantica*) in North America, for example, continual anthropogenic disturbance by local farmers at stopover sites caused anti-predator behaviors that later reduced breeding success and survival rates.

Flight initiation distance (FID) is a metric of an anti-predator behavior commonly considered to be an indicator of a biological threshold for disturbance, and thus is often used by managers to develop buffer zones to reduce anthropogenic disturbances to wildlife (Blumstein 2003, Koch and Paton 2014). The goal of instituting buffer zones is to minimize negative effects of human activity on wildlife (Erwin 1989). Buffer zones are most effective when based on systematic field surveys, rather than anecdotal observations (Rodgers and Schwikert 2002, Whitfield et al. 2008). Recommendations for buffer zones can vary depending on the method used to quantify behavioral responses, and each method of estimating effective distances is based on a number of assumptions (Fernandez-Juricic et al. 2005).

Livezey et al. (2016) describe a number of individual-, species-, and site-specific factors that can cause FID to vary widely. For birds, the metabolic power required for flight initiation is one of the most energetically demanding means of locomotion (Alexander 2002). Flight initiation, a form of flapping flight, has a higher energy expenditure than soaring or sustained flight (Pennycuick 1989, Hedenstrom 1993, Møller 2010). Thus, frequent disturbances resulting in multiple flight initiations can result in excessive consumption of energy reserves. This energy expenditure can be particularly concerning for staging long-distance migratory birds (Schummer and Eddleman 2003). However, only 9% of the world's migratory bird habitat is adequately protected from development or disturbance, with the shortcomings usually occurring at staging and migration stopover sites (Runge et al. 2015). The institution of buffer zones around prime staging habitat in areas with excessive human activity can theoretically help rectify the energetic balance of sensitive species during this period.

The federally endangered northwest Atlantic population of the roseate tern (*Sterna dougallii dougallii*) becomes concentrated during the post-breeding, pre-migratory staging period (Jul–Sep) at a limited number of sites, primarily at Cape Cod, Massachusetts, USA, which makes this population vulnerable to local catastrophe and chronic stressors (Trull et al. 1999). This population declined by >20% between 2000 and 2008 (U.S. Fish and Wildlife Service [USFWS] 2010). As of 2015, the population contained 3,900 breeding pairs (C. S. Mostello, Massachusetts Division of Fisheries and Wildlife, personal communication). Though the breeding population has rebounded, reasons for the 8-year decline remain unclear (Spendelow et al. 2016). Survival rates for the roseate tern are unusually low for a seabird, with most mortality occurring away from the breeding colonies that are distributed from New York, USA, through Nova Scotia, Canada (Spendelow et al. 1995, 2002; Nisbet and Spendelow 1999). The most recent status review for the roseate tern identified the immediate need to determine factors causing a suspected decline in survival rates during the fall staging period to identify appropriate management decisions (USFWS 2010).

Sites currently used by roseate terns and sympatric species, such as the common tern (*S. hirundo*), during the fall staging period in Cape Cod, Massachusetts are subject to various forms of anthropogenic and natural disturbance (Trull et al. 1999), which may affect the activity budgets of pre-migratory terns (Jedrey et al. 2010). Regardless of sublethal or carryover effects to fitness, any form of anthropogenic disturbance resulting in a behavioral or physiological change in a protected species (e.g., roseate terns) can be considered as take under the Endangered Species Act (USFWS 1973), and thus human disturbance to these species should be quantified and minimized when possible. The Cape Cod National Seashore (CCNS) receives >4,000,000 visitors annually (National Park Service [NPS] 2016) within its 176.5 km^2, with peak recreational activity coinciding with the staging period of tern and shorebird migration. Balancing the management needs of migratory birds and other wildlife that depend on coastal habitats at Cape Cod with the area's intense recreational use is a formidable management challenge (R. Harris, Mass Audubon, unpublished report). Understanding the effect of human disturbance on staging flocks is complicated by the potential interactions of several factors that can influence responsiveness of birds, including flock size, species composition, and intensity of disturbance (Blumstein et al. 2003, 2005; Sabine et al. 2008; Koch and Paton 2014; Lilleyman et al. 2016). These factors vary among sites within CCNS and need to be recorded and controlled for in analyses of disturbance effects on terns in the area.

Our goal for this study was to determine the behavioral effects of potential disturbances to tern flocks dominated by common and roseate terns, and to provide recommendations on buffers to reduce disturbance. To this end, our research objectives were to estimate staging mixed-species tern flocks' behavioral response distances to recreational users (i.e., pedestrians) of staging sites and compare mixed-species tern flock responses to humans to that of flock responses to gulls, which may be kleptoparasites or predators of terns, and to flock responses to shorebirds, which would not be expected to be viewed as a threat by terns. We hypothesized that pedestrians represent a stimulus that would result in more severe responses, and at farther distances, than gulls and shorebirds.

STUDY AREA

The Cape Cod peninsula extends approximately 100 km into the Atlantic Ocean from eastern Massachusetts (41°54′N, 69°58′W). Its proximity to oceanic currents and its latitudinal position create a maritime climate, with seasonality but cooler summers and warmer winters than the nearby mainland and relatively even precipitation rates throughout the year. Mean monthly precipitation during our study was 6.5 cm, based on data from the Chatham Weather Station USC00191386 (National Climatic Data Center 2018). The first year of our study was cooler than the second, with an average mean monthly temperature of 18.9°C and 20.9°C, respectively, between July and September. The temperate climate of Cape Cod, along with a high diversity of ecosystems and proximity to productive foraging grounds, results in the Cape being an important area for a large number of terrestrial and marine species. Marine and estuarine systems include beaches, sand spits, tidal flats, salt marshes, and soft-bottom benthic areas (NPS 2014). The dominant plant in the supertidal beaches of our study sites was American beach grass (*Ammophila breviligulata*) and the marshes surrounding intertidal flats used by roosting terns were dominated by saltmarsh cordgrass (*Spartina alterniflora*). The fauna of the beaches and marshes included a suite of shorebirds, terns, and gulls, with haulout areas for abundant grey seals (*Halichoerus grypus*) and less commonly harbor seals (*Phoca vitulina*). Most of the land within our study sites was undeveloped national park, with scattered low-density human structures such as life-saving stations and visitor centers. A high proportion of the northwest Atlantic roseate tern population stages in dense concentrations in a few areas around Cape Cod (Jedrey et al. 2010) characterized by tidal flats, sandbars, or barrier island beaches. Many staging sites are within CCNS on the outer portion the peninsula. In these sites, recreational activities, such as walking, swimming, water sports, beach sports, fishing, and kayaking are common (M. E. Hake, NPS, personal communication). Beach driving, though a potential disturbance source, was banned from most of our survey sites and specifically from all intertidal areas where flocks were staging.

Survey site selection with CCNS was based on published information (Jedrey et al. 2010) and preliminary field observations that indicated significant and consistent use by staging roseate terns. We included all large and continuously used staging sites in our study: Wood End (PWE), Hatches Harbor (PHH), and Race Point North (PRP-N), Provincetown; Head of the Meadow (THM), Truro; Jeremy Point (WJP) and Marconi Beach (WMB), Wellfleet; Coast Guard Beach (ECG) and Nauset Marsh (ENM), Eastham; and North Beach (CNB), North Beach Island (CNB-I), and South Beach (CSB), Chatham (Fig. 1). All sites except CSB were within CCNS, though not all are managed by the NPS; CSB, accessed via Monomoy National Wildlife Refuge, was managed for low rates of human disturbance. Beach use (sports, kite-flying, grilling, and shade tents), bicycling, camping, clamming, horseshoe crab harvesting, fires and fireworks, jet skiing, kiteboarding, and over-sand vehicles were banned at this site (USFWS 2016).

Survey sites spanned 56 km of shoreline. Sites varied in topography, substrate, and vegetation (Table 1), all of which could influence tern behavior (Burger and Gochfeld 1981, Peters and Otis 2006, Murchison et al. 2016), but were similar enough that we did not include these environmental

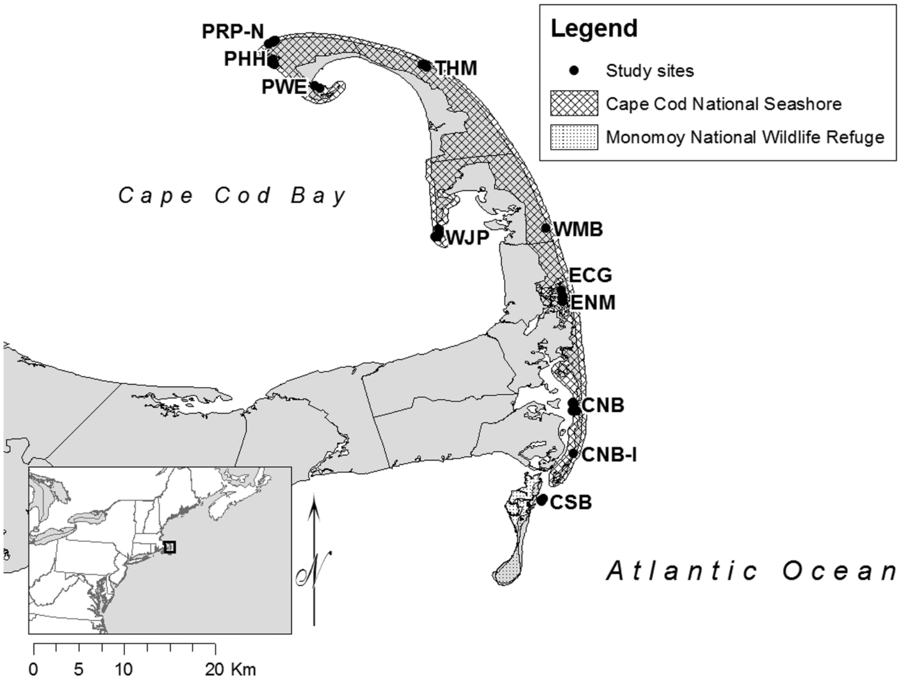

Figure 1. Survey sites within the Cape Cod National Seashore, Massachusetts, USA, 2014–2015. Sites include North Beach Island, Chatham (CNB-I); Head of the Meadow, Truro (THM); Jeremy Point, Wellfleet (WJP); North Beach, Chatham (CNB); Wood End, Provincetown (PWE); Race Point North, Provincetown (PRP-N); Hatches Harbor, Provincetown (PHH); Marconi Beach, Wellfleet (WMB); Coast Guard Beach, Eastham (ECG); the Nauset Marsh Complex, Eastham (ENM); and South Beach, Chatham (CSB).

Table 1. Coordinates, habitat descriptions, and non-research anthropogenic disturbance characteristics of sites used by staging mixed-species tern flocks in Cape Cod, Massachusetts, USA, 2014–2015.

Site	Latitude (N)	Longitude (W)	Habitat characteristics	Disturbance intensity[a]	Potential disturbance sources
PWE	42°2′3″	70°11′30″	Tidal flats and salt marsh	Low	Pedestrians, clammers
PHH	42°3′24″	70°13′57″	Cove with many types of intertidal zones	High	Pedestrians, watersports, dogs, vehicles
PRP-N	42°4′37″	70°14′2″	Beach with sandy spits, sandbars	High	Pedestrians, fishermen, watersports, dogs, vehicles
THM	42°3′10″	70°4′52″	Beach with sandbars	High	Pedestrians
WJP	41°53′6″	70°4′11″	Penninsula with tidal flats, sandbars, islands	Moderate	Pedestrians, boats
WMB	41°53′36″	69°57′47″	Beach with intertidal zone, sandbars	High	Pedestrians
ECG	41°49′53″	69°56′54″	Beach with salt marsh, tidal flats, sand bars, spit	Moderate	Pedestrians, watersports, boats
ENM	41°49′14″	69°56′46″	Spit, salt marsh, tidal flats	Low	Pedestrians, clammers, fishermen, boats, vehicles
CNB	41°43′9″	69°56′6″	Peninsula with many types of intertidal zones	Moderate	Pedestrians, vehicles, boats
CNB-I	41°40′14″	69°56′9″	Beach barrier island, sandbars	Moderate	Pedestrians, watersports, boats
CSB	41°37′24″	69°58′1″	Peninsula with tidal flats, intertidal beach	Low	None

[a] Anthropogenic disturbance intensity, categorized based on expert opinion as low (0–5 new human entries/hr), moderate (5–10/hr), or high (>10/hr).

covariates in our study. Off-road vehicles were allowed in restricted areas after the nesting period of protected shorebirds and seabirds ended at PHH, PRP-N, ENM, and CNB but were not permitted in the intertidal zones used by shorebirds and seabirds in our study. The WJP site was remote and was accessible to pedestrians only by a 7-km hike, but boat landings were fairly common. The WMB site contained a system of sandbars used by staging flocks in 2014, but these were scoured away by hurricanes before the 2015 field season. This site was abandoned by roseate terns in the second field season. The CSB site was the most remote in our study and the only one not located within CCNS. Recreational use of this site was severely limited, resulting in some of the most complex and intact habitat in the region.

METHODS

Field Methods

All field protocols were designed to minimize researcher disturbance to wild birds, per Fair et al. (2010), and Althouse et al. (2016) determined that flock responses to observers conducting leg band resightings during our study were minimal. We collected data on potential disturbance sources, and flock responses to these sources, across staging sites that covered a range of types and intensities of natural and anthropogenic potential disturbance. We visited 1–2 sites per day, 4 days per week, from early July through mid-September in 2014 and 2015, where we made observations from a distance $\geq$100 m, so we would not affect the behavior of tern flocks or potential disturbance sources. We sampled sites with different levels of potential human visitation rates (i.e., low: 0–5 new visits/hr, medium: 5–10 new visits/hr, high: >10 new visits/hr) based initially on expert opinion on human disturbance levels. We visited these sites at various times throughout day. We refined our classification of visitation rates in the second year based on results of our surveys of site-based disturbance intensity. We attempted to visit sites with different disturbance levels at least once per week.

Disturbance intensity.—We derived empirical disturbance estimates from 10-minute surveys of potential disturbances. The authors collected some data and attempted to visit sites representing a range of frequencies of tern use. Collaborators collected much of our data in conjunction with band resighting surveys that were simultaneously targeting color-banded roseate terns within CCNS, which consequently led to disproportionate sampling of sites used most frequently by staging tern flocks.

Disturbance surveys consisted of recording the presence of all potential disturbance sources within 100 m of a chosen point. If a staging flock was present, this point was centered within the flock; otherwise, we selected a point in the center of potential resting habitat or where flocks had been seen previously. We made observations from a distance of $\geq$100 m of this point so as not to affect the behavior of any terns present (Beale and Monaghan 2004, Tarr et al. 2010) and to minimize interference with potential disturbance sources. When disturbance surveys coincided with other researcher activities (e.g., another group of researchers resighting uniquely color-marked roseate terns), we performed our disturbance surveys $\geq$10 minutes after the observer had left the site to allow birds to resettle, or we included those researchers as potential disturbance sources.

During a 10-minute session, we instantaneously recorded the type of each potential disturbance source (including a range of human, and sympatric, predatory, and other non-predatory natural sources) entering the site (i.e., the 100-m radius circle), how long each source remained within the site, whether it was a unique or returning source, and the minimum distance reached between the source and the focal point. In instances where a large group entered or exited simultaneously (e.g., a group of pedestrians or a flock of shorebirds), we recorded the group as 1 source and noted the number of individuals within the group. Because humans often engaged in >1 activity throughout the survey duration, we also performed scans of the proportion of humans involved in different activities (i.e., walking, jogging, water sports [swimming, surfing, inner tubing, and stand-up paddle boarding], throwing or hitting objects [football, beach volleyball, badminton], kite flying, fishing, clamming, resting, and other [photography, bird watching, fat-tire bicycling, other researchers not associated with our project]) at the beginning and end of each disturbance survey (Altmann 1974).

Flock responses.—To quantify flock response to disturbance, we used flock scans to record the proportion of the flock exhibiting different behavioral responses to all potential disturbance sources (Altmann 1974). For either 2 hours or until terns left the site, depending on which condition first occurred, 1 observer continuously scanned the area within 200 m of focal flocks using 10 × 42 binoculars (Eagle Optics, Middleton, WI, USA) and reported all potential disturbance sources and their minimum distance to the flock, measured to the closest tern; another reported the proportion of the flock exhibiting different responses using an angled 85-mm 20–75× spotting scope (Zeiss, Oberkochen, Germany); and a third recorded the data and served as an additional lookout for the other 2 observers by reporting potential disturbance sources that may have been overlooked. One observer (MAA) was the same between years, and the other 2 were trained by the first. Observers underwent several days of training in species identification, distance estimation, behavioral classification, and counting techniques before initiating surveys. We verified all estimated distances with a laser rangefinder (Nikon, Tokyo, Japan) when possible (accurate to ±0.5 m when within 500 m in ideal conditions). We also recorded flock compositional data at 30-minute intervals. These data included flock size and the percent of roseate terns in the flock

Potential disturbance sources were gulls and shorebirds that walked or flew into or through our survey area, and all pedestrians. Other sources, such as birds of prey, dogs, and watercraft, did not occur frequently enough to include in our analyses. We recorded flock responses to disturbances as categorical behaviors with presumably increasing energetic demands. We categorized potential behavioral responses as no response, birds started or increased vocalizations and may have walked away but did not take flight, birds took flight but settled again within 1 minute, and birds took flight and remained aloft for >1 minute. The distance at which birds responded to disturbance but did not fly was termed the alertness distance (AD) and the distance at which they responded by flying was termed the flight initiation distance (FID). We recorded the proportion of the flock responding only when there were ≥10 terns present. We continued to record disturbance information and flock composition for an additional 30 minutes if the flock reduced to <10 individuals but then terminated our observations. We continuously recorded instances of kleptoparasitism of terns by gulls, including the gull species and tern species involved.

We found that it was not possible to accurately record species and flock behaviors from an observation distance >120 m, yet flocks began to respond to some disturbance sources much farther out than that. As such, observers were consistently present within the site. We implemented a 10-minute acclimation period following our entry into the survey site and before initiating our observations and remained as still as possible to minimize the effects of our presence.

Data Analysis

Disturbance intensity.—We calculated the number of minutes potential disturbance sources spent in the site by each unique potential disturbance source with the package lubridate (Grolemund 2011) in Program R (R Development Core Team 2017). We calculated the total disturbance minutes for each category (i.e., pedestrian, gull, shorebird) by adding the disturbance minutes within categories for each observation. For example, if 1 pedestrian were present for 5 minutes and a group of 3 pedestrians were present for 2 minutes, we would calculate 11 pedestrian-minutes for the observation. We then averaged the total disturbance minutes for each category across the study period by site and period of recreational activity (peak = 1000–1600, off peak = all other times, where determination of peak periods were based on our own experience at the sites).

Flock responses.—Flock scans of behavioral responses (proportion of the flock exhibiting different behavioral responses) resulted in proportional data, with response probabilities ranging from 0 to 1. The beta distribution is commonly used to model proportion data (Smithson and Verkuilen 2006), but the beta distribution does not contain 0 or 1 (Eskelson et al. 2011). Therefore, we used a zero-and-one inflated beta mixture model that uses a binomial distribution to model the discrete (0 and 1) responses, a beta distribution to model the continuous responses (>0 and <1), and another binomial distribution to model the probability that the responses came from the binomial rather than beta distribution (Ospina and Ferrari 2010).

In a Bayesian multimodel framework, we used JAGS (JAGS version 4.2.0, http://mcmc-jags.sourceforge.net/, accessed 4 Oct 2016) to model the proportion of the flock exhibiting response *i* (r^i) to a disturbance source, as a function of covariates. Except for an intercept-only null model, all models contained distance between the flock and the disturbance sources as a covariate, at a minimum. Additional models also contained the starting tern flock size, the proportion of the flock made up of roseate terns, the mean number of pedestrian disturbance minutes by site and period (i.e., peak and off-peak) combination, and an interaction between distance and the latter 3 variables. The form of the model was:

$$r^i = \mathrm{logit}^{-1}\left(\varepsilon_{\mathrm{bin}}{}^i\right) \times \mathrm{logit}^{-1}(\varepsilon_{\mathrm{disc}}{}^i) + \left[1 - \mathrm{logit}^{-1}(\varepsilon_{\mathrm{bin}}{}^i)\right] \times \mathrm{logit}^{-1}(\varepsilon_{\mathrm{cont}}{}^i),$$

where discrete responses, termed y.disc (i.e., where $y^i = 0$ or 1) were modeled as:

$$\mathrm{y.disc}^i \sim \mathrm{Bernoulli}\left(\mathrm{logit}^{-1}(\varepsilon_{\mathrm{disc}}{}^i)\right)$$

$$\varepsilon_{\mathrm{disc}}{}^{i} = X_{\mathrm{disc}}\beta^{\iota}_{\mathrm{disc}}$$

and continuous responses, termed y.cont (i.e., where $0 < y^{i} < 1$) were modeled as:

$$\mathrm{y.cont}^{i} \sim \mathrm{Beta}(p, q)$$

$$p = \mathrm{logit}^{-1}(\varepsilon_{\mathrm{cont}}{}^{i}) \times \tau$$

$$q = \left[1 - \mathrm{logit}^{-1}(\varepsilon_{\mathrm{cont}}{}^{i})\right] \times \tau$$

$$\varepsilon_{\mathrm{cont}}{}^{i} = X_{\mathrm{cont}}\beta^{\iota}_{\mathrm{cont}},$$

where $\tau = e^{t}$ was the scale parameter of the beta distribution, and the probability of a response being discrete was modeled as:

$$\mathrm{y.bin} \sim \mathrm{Bernoulli}\left(\mathrm{logit}^{-1}(\varepsilon_{\mathrm{bin}}{}^{i})\right)$$

$$\varepsilon_{\mathrm{bin}} = \beta^{i}_{0} + \beta^{i}_{1} \times \mathrm{distance} + \beta^{i}_{2} \times \mathrm{distance}^{2}$$

In the JAGS formulation, the data are supplied to the model in 3 subsets: disc (the *x* and *y* data for those observations where the proportion responding was 0 or 1), cont (the *x* and *y* data for those observations where the proportion responding was between 0 and 1), and bin (the entire dataset but with the *y* variable recoded as a 1 if the proportions were 0 or 1, and a 0 otherwise). Starting with the models for no response probability, we used the Deviance Information Criterion (DIC; Spiegelhalter et al. 2002) to compare the strength of evidence for models with different combinations of our covariates. We considered a model to have unequivocal support if ΔDIC for the next best model was >2. Once we identified the model with the most support, we used it to model the probabilities of the other response types. If 2 models had equivocal support, we based further modeling on the simpler model. Although modeling the proportional responses separately this way does not enforce the unit-sum constraint, our attempts to use zero-and-one adjusted Dirichlet distributions (Zadora et al. 2010) resulted in estimates that were not robust to initial values, apparently because of the large number of zeroes in the multivariate responses.

We ran our models using the package JagsUI (Kellner 2016), which interfaces with JAGS from the statistical software R (R Version 3.4.1, www.r-project.org, accessed 30 Jun 2017). We ran the models with 10,000 iterations and 3 chains, discarding the first 2,000 draws and retaining every 6th iteration. We defined non-informative priors and generated random initial values for all parameters. We assessed convergence by calculation of the Gelman-Rubin $\widehat{R}$ statistic (Gelman and Rubin 1992, Brooks and Gelman 1998) where we considered $\widehat{R} < 1.1$ to be evidence for convergence. We removed the response where birds increased vocalizations but did not take flight from models where shorebirds were the disturbance source because of a lack of observations.

To further examine the model fit, we calculated the mean observed responses and predicted responses within 10-m distance increments for each disturbance response type for the top model. We calculated 95% credible intervals on the predicted values of our model based on the 0.025 and 0.975 quantiles of predictions generated at each distance for every iteration of the 3 Markov chains. If the top model contained covariates in addition to distance, we generated prediction lines for realistic high and low values of those covariates. We generated figures with the R packages ggplot2 (Wickham 2009) and cowplot (Wilke 2016).

Minimum approach distance.—We used our model predictions to calculate suggestions for buffer zone distances around staging flocks to minimize human disturbance. We calculated a hypothetical management target for a buffer zone radius as the distance at which the proportion of a flock responding to pedestrians without flight was equal to the maximum predicted proportion of the flock responding to shorebirds without flight evaluated across all distances. We considered shorebirds to represent the lowest perceived threat to terns out of our 3 disturbance categories.

RESULTS

Disturbance Intensity

The number of potential disturbance minutes accrued by potential disturbance sources within a 10-minute period ranged from 0 to 4,500 minutes. The average number of minutes accrued per site-by-period combination ranged from 1 to 90.66 (Table 2). The most common human activity at all sites was resting, and the least common was fishing (Fig. 2).

Flock Responses

We recorded 407 observations of pedestrians, 292 observations of shorebirds, and 2,716 observations of gulls within 200 m of a mixed-species tern flock across 81 2-hour disturbance surveys (44 in 2014, 37 in 2015). Black-bellied plover (*Pluvialis squatarola*), sanderling (*Calidris alba*), willet (*Tringa semipalmata*), and semipalmated plover (*Charadrius semipalmatus*) were the most common shorebird species identified throughout the study; laughing gull (*Leucophaeus atricilla*), great black-backed gull (*Larus marinus*), herring gull (*L. argentatus*), and ring-billed gull (*L. delawarensis*) were the most common gulls. For shorebirds, the null model provided the best fit to the disturbance response data (ΔDIC of the next best model was 2.555). Thus, we estimated constant response probabilities for shorebirds. Mixed-species tern flocks generally exhibited no response (0.96 ± 0.0002) to the presence of shorebirds; a minor percentage (0.01 ± 0.0001) flew for >1 minute when disturbed by shorebirds. Tern flights caused by shorebirds were primarily reactions to shorebird alarm responses, such as nearby sandpiper flocks flushing or alarm calls from black-bellied plovers. Model predictions were very close to the mean observed proportion of tern flocks exhibiting different responses to shorebirds.

For all gulls, the model with distance interacting with tern flock size plus distance interacting with percent roseate terns provided the best fit (Table 3). The probability of each flock response to gulls decreased with increasing flock size, although the 95% credible intervals overlapped heavily across

Table 2. Disturbance intensities across 11 survey sites during periods of recreational activity (peak = between 1000 and 1600; off-peak = all other times) in Cape Cod, Massachusetts, USA, 2014–2015. We calculated the number of minutes that potential human disturbances were present within 10-minute observation periods[a], the size of staging tern flocks, and the percent of roseate terns within the flock and averaged these metrics within site-by-period units.

Site[b]	Period[b]	Disturbance min		Flock size		% Roseate tern	
		$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE
CSB	Off-peak	1.00	1.00	200	22	7.7	1.4
CSB	Peak	1.00	1.00	NA	NA	NA	NA
ENM	Off-peak	4.46	1.04	1,239	229	6.3	1.5
CNB	Off-peak	5.67	2.40	904	197	16.5	2.5
PWE	Off-peak	6.20	1.69	314	77	16.5	1.9
ENM	Peak	7.14	2.17	644	178	8.3	1.7
PHH	Peak	7.58	1.87	688	198	13.2	1.7
PRP-N	Off-peak	7.75	1.08	302	74	12.0	2.7
PRP-N	Peak	7.85	1.44	866	76	8.9	2.4
WJP	Off-peak	10.33	1.97	114	18	1.9	0.6
WJP	Peak	11.27	3.98	293	55	6.1	0.4
CNB-I	Peak	13.11	3.67	112	8	7.5	1.9
CNB	Peak	13.42	3.29	400	85	12.0	5.0
ECG	Off-peak	16.60	10.34	253	66	3.7	NA
PHH	Off-peak	16.87	5.51	404	103	12.3	2.5
PWE	Peak	17.60	13.26	1,796	220	9.7	1.5
THM	Peak	17.94	9.65	NA	NA	NA	NA
CNB-I	Off-peak	26.12	13.02	NA	NA	NA	NA
WMB	Off-peak	30.76	13.78	50	NA	0.0	NA
ECG	Peak	39.34	22.53	109	14	8.2	1.1
THM	Off-peak	57.32	39.16	330	55	6.0	1.3
WMB	Peak	90.66	70.41	NA	NA	NA	NA

[a] For example, if 1 pedestrian were present for 5 minutes and a group of 3 pedestrians were present for 2 minutes, the total disturbance minutes would be 11.

[b] Because of limitations in site access and use by staging flocks, no flock compositional data were available for the following 4 site-by-period sampling units: CNB-I off-peak, CSB peak, THM peak, and WMB peak.

most distances (Fig. 3). Proportion of the flock responding to gulls with a short (<1 min duration) flight increased with the percentage of the flock composed of roseate terns (Fig. 3). Model predictions for flock responses to gulls were also very close to the mean observed proportional responses. The maximum estimated proportion of the tern flock responding to gulls at any distance was <0.13 (0.02 ± 0.004).

We commonly observed gulls of various sizes near or within mixed-species tern flocks. Movements by the larger great black-backed and herring gulls tended to cause terns to relocate a short distance away (8.81 ± 0.99 m). Large gulls walking through a section of the tern flocks would cause those terns in the immediate vicinity to walk or fly for a brief moment; flapping flight by large gulls elicited responses of slightly longer duration, but the responses by terns still averaged <5 minutes. We frequently observed laughing gulls robbing prey items from common terns (laughing gulls made up 98% of total robberies, $n = 153$); we did not observe kleptoparasitism toward roseate terns. Ring-billed gulls also robbed prey from terns but to a lesser degree (<1%). Kleptoparasitic attacks by these medium gulls rarely affected >2 terns at a time, with the duration of pursuit usually lasting a few minutes.

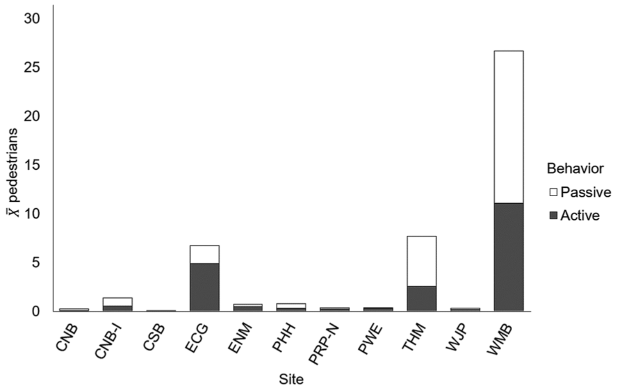

Figure 2. Mean number of pedestrians engaged in recreational activities at sites that experience significant and consistent use by staging mixed-species tern flocks in the Cape Cod National Seashore in Massachusetts, USA, 2014–2015. Active behaviors included walking, jogging, water sports (including swimming, surfing, inner tubing, and stand-up paddle boarding), throwing or hitting objects (including football, beach volleyball, and badminton), and kite flying; passive behaviors included fishing, clamming, resting, and other (photography, bird watching, other researchers not associated with our project). Sites include Marconi Beach, Wellfleet (WMB); Head of the Meadow, Truro (THM); Coast Guard Beach, Eastham (ECG); North Beach Island, Chatham (CNB-I); Hatches Harbor, Provincetown (PHH); the Nauset Marsh Complex, Eastham (ENM); Race Point North, Provincetown (PRP-N); Wood End, Provincetown (PWE); Jeremy Point, Wellfleet (WJP); North Beach, Chatham (CNB); and South Beach, Chatham (CSB).

Distance and flock size were strong predictors of the probability of flock response to pedestrians (Table 4). Probability of flock responses increased when pedestrians were within 100 m of a staging flock, with small tern flocks being especially sensitive (Fig. 4). Pedestrians within 65 m, for example, led to 0.05 of the flocks with an average flock

Table 3. Model selection results for variables explaining mixed-species tern flock behavioral responses to gulls during the staging period (Jul–Sep), Cape Cod, Massachusetts, USA, 2014–2015. Potential model covariates included distance to a potential disturbance source (measured from the closest tern), flock size, percent roseate terns in the flock, and average number of disturbance minutes accrued within 10 minutes for the survey's particular site and time relative to peak recreational use (peak = 1000–1600; off-peak = all other times).

Model	pD[a]	ΔDIC[b]	w_i[c]
Distance × flock size + distance × % roseate	16.1	0.00	0.657
Distance × flock size + % roseate	14.0	3.10	0.140
Distance + flock size + % roseate	11.9	3.42	0.119
Distance + flock size × % roseate	15.0	4.09	0.085
Distance + % roseate	10.5	17.65	0.000
Distance + flock size	9.4	19.64	0.000
Distance × % roseate	12.1	22.35	0.000
Distance × disturbance min	11.5	26.35	0.000
Distance × flock size	11.7	30.85	0.000
Distance + disturbance min	9.5	31.29	0.000
Distance	7.5	31.62	0.000
Null	3.5	873.77	0.000

[a] Effective number of parameters.

[b] Deviance Information Criterion difference between the best model (17,377.65) and the model for which the ΔDIC is given.

[c] Model weights, calculated as the relative likelihood of the given model ($e^{(-0.5 \times \Delta\mathrm{DIC})}$) divided by the sum of these values across all models.

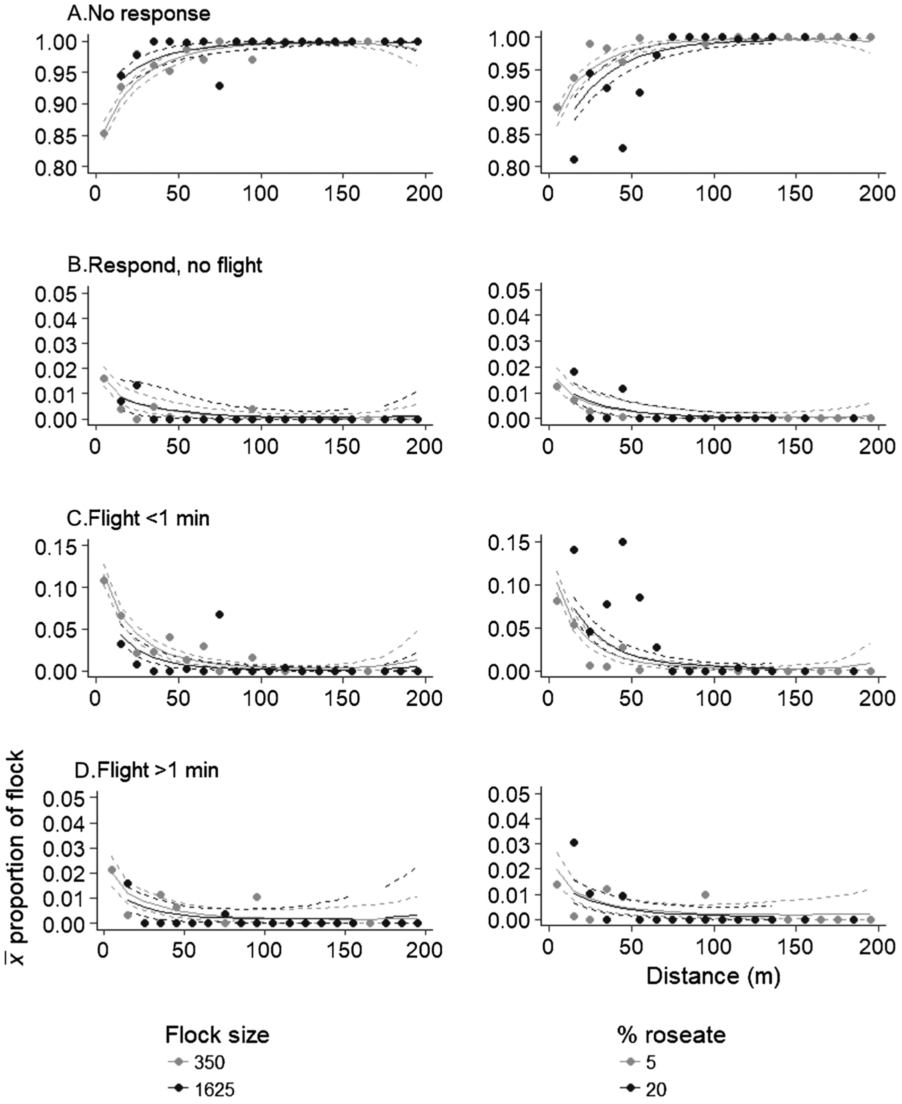

Figure 3. Mean proportion of a mixed-species tern flock exhibiting no response (A), response without flight (B), flight of <1 minute (C), and flight of ≥1 minute (D) versus minimum distance to gulls in Cape Cod, Massachusetts, USA, 2014–2015. Solid lines indicate zero-and-one inflated beta regression model predictions, with dashed 95% credible intervals. Solid points represent the mean observed response calculated within 10-m distance intervals. Graphs in the left column demonstrate flock responses for 2 different flock sizes while holding the percent roseate terns constant at 10% (the mean value in our dataset); graphs in the right column demonstrate flock responses for 2 different proportions of roseate terns, holding flock size constant at 640 (the mean value in our dataset).

size of 1,625 terns flushing, whereas 0.10 of flocks with a flock size of 350 terns flushed at this distance. The proportion responding with flights with durations of >1 minute, and potentially leaving the site, had the sharpest increase as pedestrians neared. The proportion responding with short flights (<1 min in length) began increasing at a farther distance than long flights, but this proportion fell as the proportion responding with long flights rapidly increased. This result generally reflects the trend of escalating flock responses we observed as pedestrians drew near. The proportion of the flock responding with alertness was generally low but also declined with distance to pedestrians. When pedestrians were >120 m from a tern flock, the probability of flight response was generally similar to that of shorebirds and gulls, regardless of the flock size and composition. Model predictions for pedestrian disturbance sources closely followed mean observations, with most observations within 95% credible intervals of the predicted values. Our results contained one visually obvious outlier (Fig. 4) so we reran our models without that observation to explore its influence on our results. Our model rankings were nearly identical (within 0.75 DIC) to our results that included the outlier, indicating that the distance + flock size model was still the best fitting.

Recreational activities involving more rapid movements (e.g., jogging, walking, playing sports) within 200 m of a flock caused at least part of a tern flock to react (proportion responding > 0) more frequently than less energetic activities (e.g., slowly searching for seashells, clamming, sunbathing). Active pedestrian activities elicited flight responses from the flock 0.31 of the time ($n = 110$ of 356 observations, distance = 115.3 m ± 7.2), whereas flight responses occurred

Table 4. Model selection results for variables explaining mixed-species tern flock behavioral responses to pedestrians during the staging period (Jul–Sep) in Cape Cod, Massachusetts, USA, 2014–2015. Potential model covariates included distance to a potential disturbance source (measured from the closest tern), flock size, percent roseate terns in the flock, and average number of disturbance minutes accrued within 10 minutes for the survey's particular site and time period (1000–1600 = peak, all other times = off-peak).

Model	pD[a]	ΔDIC[b]	w_i[c]
Distance + flock size	9.6	0.00	1
Distance × flock size	12.1	2.50	0
Distance	6.9	10.69	0
Distance + disturbance min	9.3	10.83	0
Distance + % roseate	9.5	11.89	0
Distance × % roseate	11.4	14.64	0
Distance × disturbance min	11.8	15.95	0
Null	3.7	197.24	0

[a] Effective number of parameters.

[b] Deviance Information Criterion difference between the best model (457.64) and the model for which the ΔDIC is given.

[c] Model weights, calculated as the relative likelihood of the given model ($e^{(-0.5 \times \Delta\mathrm{DIC})}$) divided by the sum of these values across all models.

in only 0.07 of observations of passive activities ($n = 4$ of 51 observations, distance = 83.1 m ± 2.7). A *post hoc* Pearson's chi-squared test with Yates' continuity correction without regard to distance indicated a significantly greater likelihood of flocks responding to more active pedestrians than passive pedestrians ($\chi^2 = 10.64$, $P \leq 0.001$).

Minimum Approach Distance

Our results for response distances to shorebird and gull disturbance sources could be used as the upper threshold of acceptable pedestrian disturbance rates at our survey sites. Based on this threshold, a buffer distance could be calculated for protecting terns from anthropogenic disturbance (Rodgers and Schwikert 2002). For example, managers at CCNS might decide that pedestrians should be kept at a distance from tern flocks at which 96% of the terns show no behavioral response to them, which would be similar to the flock response to shorebirds (a non-threat) that we observed. Based on the no response curve for pedestrians, the managers could select a distance of approximately 100 m to minimize disturbance.

DISCUSSION

Although site-level anthropogenic disturbance intensity did not appear to influence how staging mixed-species tern flocks responded to potential disturbance sources, the type of disturbance source and flock composition were important determinants of flock response. Tern flocks responded strongly to approaching pedestrians but did not exhibit the

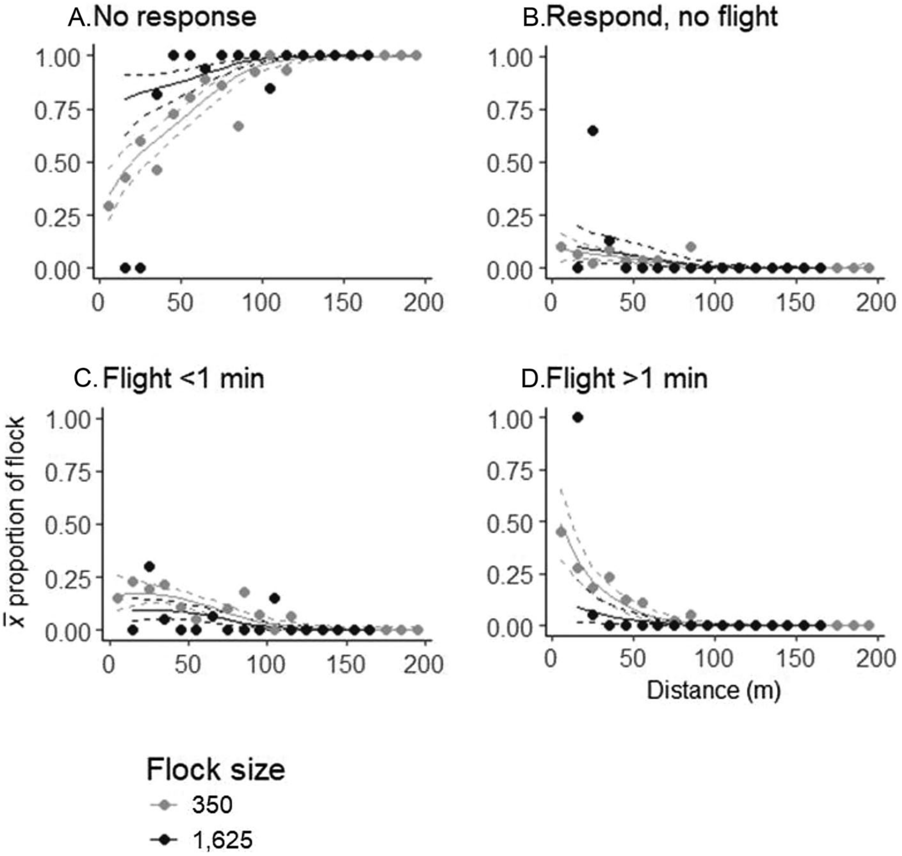

Figure 4. Mean proportion of a mixed-species tern flock exhibiting no response (A), response without flight (B), flight of <1 minute (C), and flight of ≥1 minute (D) versus minimum distance to pedestrians in Cape Cod, Massachusetts, USA, 2014–2015. Predictions are shown for 2 flock sizes. Solid lines indicate zero-and-one inflated beta regression model predictions, with dashed 95% credible intervals. Solid points represent the mean observed response calculated within 10-m distance intervals.

same flush behaviors upon the approach of shorebirds or gulls. For example, flocks were 4.1 times more likely to flush from approaching pedestrians than to gulls at 50 m, and 15.5 times more than to shorebirds. Moreover, terns were 3 times more likely to flush from gulls than to shorebirds at 50 m.

Threat perception likely played a role in the differential response of terns to different sources. Shorebirds are not predators or competitors of terns, greater black-backed gulls can prey on terns, and gulls of many species act as kleptoparasites on terns. We frequently observed kleptoparasitism by laughing and ring-billed gulls of adult terns returning with prey items for hatch-year terns, though nearly all kleptoparasitic attacks were made by laughing gulls and targeted common rather than roseate terns. Tern responses to kleptoparasitism are likely moderated to allow for facultative adjustment of behaviors to optimize energy expenditure (Ridley and Raihani 2007). This could explain why large flocks were less likely than small flocks to respond to the presence of gulls with short flights; the risk of an individual being targeted is diluted in a larger flock (Goodale and Ruxton 2010). The probability of terns demonstrating alertness and long flights were generally similar between flock sizes, perhaps because kleptoparasitism attempts generally result in short aerial chases. Longer chases would consume more energy, and the gain from a single prey item would likely not be worth the additional cost. An increase in the average flight duration from 1 minute to 5 minutes results in an additional 16.61 kJ of energy used per flight, as calculated by equations 1 and 2 in Lilleyman et al. (2016) using the average mass of 115 g. Even if terns repeatedly flushed 5 times from a gull for $\leq$1 minute of flight, their energy expenditure would hypothetically still be at least 7.55 kJ less than 1 5-minute flight.

We found evidence for intraspecific differences in flushing responses to gulls, where flocks with a large proportion of roseate terns were more likely than flocks with a small proportion to respond to approaching gulls and at greater distances. The fields of anti-predator behavior and disturbance ecology have consistently reported many external variables to influence FID, but exploration of species-specific factors has been limited. Notable exceptions include controlled approaches to quantify FID variation among 8 species of shorebirds in Australia by Blumstein et al. (2003) and simulation models by Blumstein et al. (2005) and Blumstein (2006). Roseate terns are less aggressive than common terns, and benefit from the defensive behavior of common tern colonies during the breeding period (Burger and Gochfeld 1988). This benefit may extend beyond this period. Our results suggest that roseate terns perceive less risk of gull predation or kleptoparasitism when in flocks with a high proportion of common terns than when there are fewer common terns in the flock.

The magnitude of response to approaching humans was much larger than that for the 2 natural sources of potential disturbance (i.e., gulls and shorebirds), and still varied as a function of flock size but was not sensitive to the percent of roseate terns in the flock. Humans are one of the largest organisms present within these staging areas, and thus they may be perceived as one of the greatest threats to tern flocks. Moreover, in the South American wintering areas, these terns may be subject to hunting pressure (Buckley and Buckley 1984, Kaufman 1996), which can sensitize individuals to human presence (Clucas and Marzluff 2012).

Recreational activities involving more rapid movements (e.g., jogging) seem to create more disturbance than conservative activities (e.g., searching for seashells). Active pedestrians move about in fully erect posture, often accompanied by frequent motions and loud conversations. Those at rest presented a smaller profile, tended to move less, and were generally quieter than their counterparts, and thus may have appeared less threatening to staging flocks. Energetic pedestrian activities may be perceived as predatory behaviors by tern flocks; unlike the phenomenon of kleptoparasitism, terns may not have had the time to adapt via facultative adjustment of their energy output in their response when there is no true direct threat to their survival. Staging tern flocks respond to different types of human activities in different ways, including habituation to frequent field activities (Zingo and Spendelow 1998, Nisbet 2000). Our large team of rotating observers (which may not present the same opportunity for habituation as a consistent team) documented significantly lower flock responses to researchers than pedestrians, a result attributed to the differential manner in which flocks were approached (Althouse et al. 2016). Passive activities may still cause subtle disturbances such as reduced local carrying capacity (Bednar-Friedl et al. 2012), or carryover effects (Harrison et al. 2011, O'Connor et al. 2014, O'Connor and Cooke 2015), but we did not attempt to discern these effects in our study.

Current methods of calculating buffer zones to minimize flock responses to disturbance use the 95% confidence interval for mean FID plus the mean alert distance (Rodgers and Schwikert 2002, Fernandez-Juricic et al. 2005). Our no response curves integrate flight and alertness distances and can be used to set buffer zones. Recommended buffer zones for roseate and common terns vary between 140 m (Rodgers and Schwikert 2002) and 200 m (Erwin 1989). These recommended areas, however, were not calculated for staging terns. Their applicability to the staging period is based on a number of assumptions identified by Fernandez-Juricic et al. (2005), notably that FID does not vary with time of the year, and bird use of the buffer area is constant and equal to the carrying capacity of the system. The efficacy of minimum approach distances from breeding period studies in the migration staging period and other distinct life stages is questionable. We have satisfied both above assumptions by evaluating FID within the post-breeding staging period only.

For roseate terns in the northeast United States, prevention of disturbance is a legal and ecological requirement for management and we were able to provide baseline response distances for staging terns to multiple sources of disturbance. Moreover, our analysis of flock responses to gulls and shorebirds provide some potential targets for flock response levels to humans.

MANAGEMENT IMPLICATIONS

Managers interested in applying buffer zones based on FID should apply buffer zones as small as 100 m. Managers may also want to consider more conservative buffers in areas frequented by smaller tern flocks because terns in small flocks may be more sensitive to disturbance than when in larger flocks. Because other factors known to affect FID can vary among locations and times, such as exposed area at high tide (Lilleyman et al. 2016) and average intruder group size and orientation (Burger and Gochfeld 1981, Geist et al. 2005), we recommend that managers adopting buffer zones monitor their effectiveness at their own sites, and adjust them if conditions warrant.

Given the dynamic nature of coastal systems and the species using them, we also recommend annual monitoring to determine prominent staging sites early in the staging period, with the application of the chosen buffer around areas potentially used by staging flocks. In addition, we recommend that where practicable, managers consider applying a single protected area (e.g., $\geq$100-m radius around all intertidal zones) within a site. A single larger protected area will likely be more effective at reducing the effects of human disturbance on terns than several smaller buffers because terns in a site are often dispersed in multiple smaller flocks with temporally fluid locations throughout the day. Roping off these protected areas and posting signage can be an effective management tool (Ikuta and Blumstein 2003), although how much so often depends on local support by stakeholders (Glover et al. 2011).

ACKNOWLEDGMENTS

The use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the United States Government. The USFWS provided necessary resources to complete the study. M. E. Hake, R. P. Cook, J. J. Taylor, and K. E. Iaquinto provided logistical support and site access. J. L. Correia, J. C. May, S. S. Brady, and P. L. Gallo assisted with data collection. The Cape Cod National Seashore, United States Geological Survey, CNY Wildfowlers Association, and Edna Baily Sussman Foundation provided funding.

LITERATURE CITED

Alexander, R. M. 2002. The merits and implications of travel by swimming, flight and running for animals of different sizes. Integrative and Comparative Biology 42:1060–1064.

Althouse, M. A., J. B. Cohen, J. A. Spendelow, S. M. Karpanty, K. D. Davis, K. C. Parsons, and C. F. Luttazi. 2016. Quantifying the effects of research band resighting activities on staging terns in comparison to other disturbances. Waterbirds 39:415–419.

Altmann, J. 1974. Observational study of behavior: sampling methods. Behavior 49:227–267.

Beale, C. M., and P. Monaghan. 2004. Behavioral responses to human disturbance: a matter of choice? Animal Behavior 68:1065–1069.

Bednar-Friedl, B., D. A. Behrens, and M. Getzner. 2012. Optimal dynamic control of visitors and endangered species in a National Park. Environmental and Resource Economics 52:1–22.

Blumstein, D. T. 2003. Flight-initiation distance in birds is dependent on intruder starting distance. Journal of Wildlife Management 67:852–857.

Blumstein, D. T. 2006. Developing an evolutionary ecology of fear: how life history and natural history traits affect disturbance tolerance in birds. Animal Behaviour 71:389–399.

Blumstein, D. T., L. L. Anthony, R. Harcourt, and G. Ross. 2003. Testing a key assumption of wildlife buffer zones: is flight initiation distance a species-specific trait? Biological Conservation 110:97–100.

Blumstein, D. T., E. Fernández-Juricic, P. A. Zollner, and S. C. Garity. 2005. Inter-specific variation in avian responses to human disturbance. Journal of Applied Ecology 42:943–953.

Brooks, S. P., and A. Gelman. 1998. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7:434–455.

Buckley, P. A., and F. G. Buckley. 1984. Seabirds of the north and middle Atlantic coasts of the United States: their status and conservation. Pages 101–133 *in* J. P. Croxall, P. G. H. Evans, and R. W. Schreiber, editors. Status and conservation of the world's seabirds. International Council for Bird Preservation, Cambridge, United Kingdom.

Burger, J., and M. Gochfeld. 1981. Discrimination of the threat of direct versus tangential approach to the nest by incubating herring and great black-backed gulls. Journal of Comparative and Physiological Psychology 95:676–684.

Burger, J., and M. Gochfeld. 1988. Defensive aggression in terns: effect of species, density, and isolation. Aggressive Behavior 14:169–178.

Clucas, B., and J. M. Marzluff. 2012. Attitudes and actions toward birds in urban areas: human cultural differences influence bird behavior. Auk 129:8–16.

Drent, R., C. Both, M. Green, J. Madsen, and T. Piersma. 2003. Pay-offs and penalties of competing migratory schedules. Oikos 103:274–292.

Endler, J. A. 1991. Interactions between predators and prey. Pages 169–196 *in* J. R. Krebs and N. B. Davies, editors. Behavioral ecology: an evolutionary approach. Third edition. Blackwell Scientific Publications, Oxford, United Kingdom.

Erwin, R. M. 1989. Responses to human intruders by birds nesting in colonies: experimental results and management guidelines. Colonial Waterbirds 12:104–108.

Eskelson, B. N., L. Madsen, J. C. Hagar, and H. Temesgen. 2011. Estimating riparian understory vegetation cover with beta regression and copula models. Forest Science 57:212–221.

Fair, J., E. Paul, and J. Jones, editors. 2010. Guidelines to the use of wild birds in research. Ornithological Council, Washington, D.C., USA.

Fernandez-Juricic, E., and I. Rodriguez-Prieto. 2010. Risk allocation in anti-predator behavior. Pages 75–78 *in* M. D. Breed and M. Janice, editors. Encyclopedia of animal behavior. Volume 1. Elsevier Science & Technology, London, United Kingdom.

Fernandez-Juricic, E., M. P. Venier, D. Renison, and D. T. Blumstein. 2005. Sensitivity of wildlife to spatial patterns of recreationist behavior: a critical assessment of minimum approaching distances and buffer areas for grassland birds. Biological Conservation 125:225–235.

Geist, C., J. Liao, S. Libby, and D. T. Blumstein. 2005. Does intruder group size and orientation affect flight initiation distance in birds? Animal Biodiversity and Conservation 28:69–73.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–472.

Glover, H. K., M. A. Weston, G. S. Maguire, K. K. Miller, and B. A. Christie. 2011. Towards ecologically meaningful and socially acceptable buffers: response distances of shorebirds in Victoria, Australia, to human disturbance. Landscape and Urban Planning 103:326–334.

Goodale, E., and G. D. Ruxton. 2010. Antipredator benefits from heterospecifics. Pages 94–99 *in* M. D. Breed and J. Moore, editors. Encyclopedia of animal behavior. Volume 1. Elsevier Science & Technology, London, United Kingdom.

Grolemund, G. 2011. Dates and times made easy with lubridate. Journal of Statistical Software 40:1–25.

Harrison, X. A., J. D. Blount, R. Inger, D. R. Norris, and S. Bearhop. 2011. Carry-over effects as drivers of fitness differences in animals. Journal of Animal Ecology 80:4–18.

Hedenstrom, A. 1993. Migration by soaring or flapping flight in birds: the relative importance of energy cost and speed. Philosophical Transactions: Biological Sciences 342:353–361.

Ikuta, L. A., and D. T. Blumstein. 2003. Do fences protect birds from human disturbance? Biological Conservation 112:447–452.

Inger, R., G. A. Gudmundsson, G. D. Ruxton, J. Newton, K. Colhoun, S. Auhage, and S. Bearhop. 2008. Habitat utilisation during staging affects body condition in a long distance migrant, *Branta bernicla hrota*: potential impacts on fitness? Journal of Avian Biology 39:704–708.

Jedrey, E. L., R. J. Harris, and E. A. Ray. 2010. Roseate terns—citizens of the world: the Canada to Cape Cod connection. Bird Observer 38:146–150.

Kaufman, K. 1996. Lives of North American Birds. Peterson natural history companions, Houghton Mifflin Co., University of Michigan. http://www.audubon.org/field-guide/bird/roseate-tern. Accessed 4 Dec 2016.

Kellner, K. 2016. jagsUI: a wrapper around "rjags" to streamline "jags" analyses. https://CRAN.R-project.org/package=jagsUI. Accessed 31 Oct 2016.

Koch, S. L., and P. W. C. Paton. 2014. Assessing anthropogenic disturbances to develop buffer zones for shorebirds using a stopover site. Journal of Wildlife Management 78:58–67.

Kvist, A., and Å. Lindström. 2001. Basal metabolic rate in migratory waders: intra-individual, intraspecific, interspecific and seasonal variation. Functional Ecology 15:465–473.

Lilleyman, A., D. C. Franklin, J. K. Szabo, and M. J. Lawes. 2016. Behavioural responses of migratory shorebirds to disturbance at a high-tide roost. Emu 116:111–118.

Livezey, K. B., E. Fernandez-Juricic, and D. T. Blumstein. 2016. Database of bird flight initiation distances to assist in estimating effects from human disturbance and delineating buffer areas. Journal of Fish and Wildlife Management 7:181–191.

Møller, A. P. 2010. Up, up, and away: relative importance of horizontal and vertical escape from predators for survival and senescence. Journal of Evolutionary Biology 23:1689–1698.

Murchison, C. R., Y. Zharikov, and E. Nol. 2016. Human activity and habitat characteristics influence shorebird habitat use and behavior at a Vancouver island migratory stopover site. Environmental Management 58:386–398.

National Climatic Data Center. 2018. Climate data online search. https://www.ncdc.noaa.gov/cdo-web/search. Accessed 2 Apr 2018.

National Park Service [NPS]. 2014. Nature & science—Cape Cod National Seashore. http://www.nps.gov/caco/naturescience/index.htm. Accessed 4 Apr 2014.

National Park Service [NPS]. 2016. National Park Service visitor use statistics for Cape Cod National Seashore. NPS Stats. National Park Service visitor use statistics. https://irma.nps.gov/Stats/Reports/Park. Accessed 20 Sep 2016.

Newton, I. 2006. Can conditions experienced during migration limit the population levels of birds? Journal of Ornithology 147:146–166.

Nisbet, I. C. T. 2000. Disturbance, habituation, and management of waterbird colonies. Waterbirds 23:312–332.

Nisbet, I. C. T., and J. A. Spendelow. 1999. Contribution of research to management and recovery of the roseate tern: review of a twelve-year project. Waterbirds 22:239–252.

O'Connor, C. M., and S. J. Cooke. 2015. Ecological carryover effects complicate conservation. Ambio 44:582–591.

O'Connor, C. M., D. R. Norris, G. T. Crossin, and S. J. Cooke. 2014. Biological carryover effects: linking common concepts and mechanisms in ecology and evolution. Ecosphere 5:art28.

Ospina, R., and S. L. P. Ferrari. 2010. Inflated beta distributions. Statistical Papers 51:111–126.

Pennycuick, C. J. 1989. Bird flight performance: a practical calculation manual. Oxford University Press, Oxford, United Kingdom.

Peters, K. A., and D. L. Otis. 2006. Shorebird roost-site selection at two temporal scales: is human disturbance a factor? Journal of Applied Ecology 44:196–209.

Pomeroy, A. C. 2006. Tradeoffs between food abundance and predation danger in spatial usage of a stopover site by western sandpipers, *Calidris mauri*. Oikos 112:629–637.

R Development Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. Accessed 30 Jun 2017.

Ridley, A., and N. Raihani. 2007. Facultative response to a kleptoparasite by the cooperatively breeding pied babbler. Behavioral Ecology 18:324–330.

Rodgers, J. A., and S. T. Schwikert. 2002. Buffer-zone distances to protect foraging and loafing waterbirds from disturbance by personal watercraft and outboard-powered boats. Conservation Biology 16:216–224.

Runge, C. A., J. E. Watson, S. H. M. Butchart, J. O. Hanson, H. P. Possingham, and R. A. Fuller. 2015. Protected areas and global conservation of migratory birds. Science 350:1255–1258.

Sabine, J. B., J. M. Meyers, C. T. Moore, and S. H. Schweitzer. 2008. Effects of human activity on behavior of breeding American oystercatchers, Cumberland Island National Seashore, Georgia, USA. Waterbirds 31:70–82.

Schauroth, C., and P. H. Becker. 2008. Post-fledging body mass increase in common terns *Sterna hirundo*: influence of age, sex and year. Ibis 150:50–58.

Schummer, M. L., and W. R. Eddleman. 2003. Effects of disturbance on activity and energy budgets of migrating waterbirds in south-central Oklahoma. Journal of Wildlife Management 67:789–795.

Smithson, M., and J. Verkuilen. 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychological Methods 11:54–71.

Spendelow, J. A., D. Monticelli, J. D. Nichols, J. E. Hines, I. C. T. Nisbet, G. Cormons, H. Hays, J. J. Hatch, and C. S. Mostello. 2016. Roseate tern breeding dispersal and fidelity: responses to two newly restored colony sites. Ecosphere 7:e01510.

Spendelow, J. A., J. D. Nichols, J. E. Hines, J.-D. Lebreton, and R. Pradel. 2002. Modelling postfledging survival and age-specific breeding probabilities in species with delayed maturity: a case study of roseate terns at Falkner Island, Connecticut. Journal of Applied Statistics 29:385–405.

Spendelow, J. A., J. D. Nichols, I. C. T. Nisbet, H. Hays, G. D. Cormons, J. Burger, C. Safina, J. E. Hines, and M. Gochfeld. 1995. Estimating annual survival and movement rates of adults within a metapopulation of roseate terns. Ecology 76:2415–2428.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde. 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64:583–639.

Tarr, N. M., T. R. Simons, and K. H. Pollock. 2010. An experimental assessment of vehicle disturbance effects on migratory shorebirds. Journal of Wildlife Management 74:1776–1783.

Trull, P., S. Hecker, M. J. Watson, and I. C. T. Nisbet. 1999. Staging of roseate terns *Sterna dougallii* in the post-breeding period around Cape Cod, Massachusetts, USA. Atlantic Seabirds 1:145–158.

U.S. Fish and Wildlife Service [USFWS]. 1973. Endangered Species Act of 1973 as amended through the 108th Congress. USFWS, Washington, D.C., USA.

U.S. Fish and Wildlife Service [USFWS]. 2010. Caribbean roseate tern and north Atlantic roseate tern (*Sterna dougallii dougallii*) 5-year review: summary and evaluation. USFWS southeast region, Caribbean ecological services field office, Boquerón, Puerto Rico; and northeast region, New England field office, Concord, New Hampshire, USA.

U.S. Fish and Wildlife Service [USFWS]. 2016. Appendix D—findings of appropriateness and compatibility determinations. Monomoy National Wildlife Refuge Comprehensive Conservation Plan. https://www.fws.gov/uploadedFiles/04w_Appendix_D_CDs_and_FAs(4036KB.pdf. Accessed 12 Sep 2017.

Whitfield, D. P., M. Ruddock, and R. Bullman. 2008. Expert opinion as a tool for quantifying bird tolerance to human disturbance. Biological Conservation 141:2708–2717.

Wickham, H. 2009. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, New York, USA. http://ggplot2.org. Accessed 29 Sep 2016.

Wilke, C. O. 2016. cowplot: streamlined plot theme and plot annotations for "ggplot2." https://CRAN.R-project.org/package=cowplot. Accessed 3 Oct 2016.

Yager, D. D. 2010. Predator evasion. Pages 765–773 *in* M. D. Breed and J. Moore, editors. Encyclopedia of animal behavior. Volume 1. Elsevier Science & Technology, London, United Kingdom.

Zadora, G., T. Neocleous, and C. Aitken. 2010. A two-level model for evidence evaluation in the presence of zeros. Journal of Forensic Sciences 55:371–384.

Zingo, J. A., and J. A. Spendelow. 1998. Frequent handling of roseate tern chicks: is more better? Pages 78–79 *in* Meeting Program, 1998 Meeting of the Colonial Waterbird Society and the Flamingo Specialist Group. Florida International University, North Miami, Florida, USA.

Associate Editor: Bev Gingras.