Animal Welfare 2012, 21: 101-111
ISSN 0962-7286

*The effect of investigator disturbance on egg laying, chick survival and fledging mass of short-tailed shearwaters (*Puffinus tenuirostris*) and little penguins (*Eudyptula minor*)*

C Vertigan†, CR McMahon†‡, V Andrews-Goff† and MA Hindell†*

† Marine Predator Unit, Institute for Marine and Antarctic Studies (IMAS), University of Tasmania, Hobart, Tasmania 7001, Australia
‡ School of Environmental Research, Charles Darwin University, Darwin, Northern Territory 0909, Australia
* Contact for correspondence and requests for reprints: caitlin.vertigan@gmail.com

Abstract

*Field-based animal researchers need to balance the potential adverse effects of their research activities against the benefits of research outcomes, but the data required to do this are often lacking. Assessing, and subsequently reporting the effects of researcher activities on wild animal populations can be difficult, so that studies to detect these effects sometimes lack rigour or fail to encompass sufficient time to ensure that the effects are tested under a range of environmental stresses. We monitored the effect of investigators working in colonies of two seabirds, the short-tailed shearwater (*Puffinus tenuirostris*) and the little penguin (*Eudyptula minor*). Disturbance of breeding birds while checking nests or the weighing of chicks to monitor growth are very common activities for demographic and ecological studies, but how these activities may influence the birds is rarely measured. We investigated differing levels of disturbance during both activities between 2002–03 and 2008–09 to assess their effect on egg laying, chick survival and growth rate and observed no effect for nest checking or handling of short-tailed shearwaters and indeterminate effects for handling in little penguins. Over a period of several years the study has observed a large-scale decline in the number of breeding shearwaters and includes years when control nests had above and below average breeding success.*

Keywords: *animal welfare,* Eudyptula minor*, investigator disturbance, little penguin,* Puffinus tenuirostris*, short-tailed* shearwater

Table 1 Field schedule detailing year, species and activity.

Year	Short-tailed shearwater		Little penguin	
	Nest checks	Handled	Nest checks	Handled
2003	Y	Y	N	Y
2004	N	N	N	N
2005	Y	Y	N	Y
2006	Y	Y	N	Y
2007	Y	Y	N	N
2008	Y	N	N	N
2009	Y	N	N	N

has declined from 36,569 (± 7,291) to 12,602 (± 1,697) breeding pairs between 2003 and 2010 (Vertigan 2010), representing a loss in the order of 6,848 birds per year.

Breeding sympatrically with the short-tailed shearwater on many offshore islands in Tasmania is the little penguin. The global population of little penguins is between 350–600,000 breeding pairs (Marchant *et al* 1990) with 1,020 (± 228) breeding pairs present on Wedge Island (Vertigan 2010). The use of sympatric species allows us to monitor multiple trophic levels (as the main prey is different for both species) and multiple spatial scales as penguins are constrained to feeding within approximately 200 km of their breeding site, whereas shearwaters may travel as far as the Polar Frontal Zone (> 1000 km from the breeding site) to obtain food (Weimerskirch & Cherel 1998).

It has been postulated that wild animals respond to humans as they would to potential predators (Frid & Dill 2002; Beale & Monaghan 2004). As such, burrowing seabirds (eg shearwaters and little penguins) are likely to respond aggressively when researchers are working in a burrow where their genetic investment (the chicks) resides. The common activities researchers perform are nest-checking for presence or absence of adults, eggs and chicks, and capture of birds for measurements such as weighing. Disturbances of this nature to nests and birds can result in decreased nest-site fidelity (abandonment of nest sites in subsequent years) in some burrowing species (Blackmer *et al* 2004). Disturbance effects may be exacerbated by secondary effects such as nesting density (de Villiers 2008) which is important as overall declines can result from density dependence rather than direct disturbance (Gill 2007). Moreover, substrates in colonies containing burrowing species are often fragile and unstable as burrowing compromises soil stability; investigator damage can occur by partial or total collapse of a burrow while investigators are navigating through the colony or checking burrows.

Our primary aims in this study were to quantify the effect of investigator disturbance on three key life-history traits — rate of egg laying, survival, and growth of chicks of short-tailed shearwaters and little penguins over multiple years. Specifically, we hypothesise no short-term effects on those life-history traits. However, long-term effects are less certain. In this study, we aim to provide information: i) relevant to future seabird research planning; ii) that can be incorporated into our primary research on the response of key life-history traits in these species to changes in the marine environment; and iii) to contribute to the broader debate on animal welfare in wildlife research.

Materials and methods

Fieldwork was undertaken on Wedge Island in south-east Tasmania, Australia (43° 07’ S, 147° 40’ E) on short-tailed shearwaters and little penguins during October–April 2002–03 (a pilot study with reduced sampling density), 2003–04, 2004–05, 2005–06, 2006–07, 2007–08 and 2008–09 (see Table 1 for schedule and activities). All seasons will subsequently be referred to by their finishing year, eg 2002–03 is referred to as the 2003 season. Marking of burrows was carried out in 2003–04 with no further work undertaken due to researcher illness.

Fieldwork was conducted as part of a long-term study into the effects of environmental variation on life-history characteristics of short-tailed shearwaters and little penguins and was conducted with the permission of the University of Tasmania Animal Ethics Committee and the Nature Conservation Branch of the Tasmanian Parks and Wildlife Service. Four levels of investigator disturbance were tested (two levels of nest checking and two levels of handling) for each species, lower and higher levels of disturbance within each activity are regarded as ‘controls’ and ‘treatments’, respectively, the details of which are expanded upon below.

Researcher activity 1 — Nest checking

The island (1.6 × 0.8 km; length × width at its widest point) was divided into 35 east/west line transects (length 40–450 m). A stratified sampling regime was used such that, transects 1–20, which were in the area of highest little penguin density, were 25 m apart, and transects 21–35, in low penguin density areas, were placed 50 m apart. Circular sampling units (quadrats) were placed at 20 m intervals on transects 1–20 and at 25 m intervals on transects 21–35. All burrows within a 2-m radius of the sampling unit centre were individually marked using flag markers (plastic tabs mounted on wire staves). A total of 674 shearwater burrows were marked in 241 quadrats in this way during the 2004 season; of these, 473 burrows (transects 17–35) were checked twice per season to monitor survival of chicks (once in December to establish laying success and once in April to establish fledging success) and the remaining 201 burrows (transects 1–16) were checked once per month (a total of seven times per season; Table 1). Burrows were checked during daylight hours between sunrise and sunset. Transects 1–16 were chosen for the intensively checked area due to the smaller colony size and stable substrate of this area’s perimeter which allowed for access without frequently traversing unstable areas. Burrow contents were observed without handling animals by using a custom-made 'burrow scope' (a lipstick camera with infra-red [IR] lights attached to a 3-m flexible tube and wired to a 3.5-inch colour monitor, mark 4 model, manufactured by Richard Holmes, Zoology workshop, University of Tasmania, Australia).

Burrows checked seven times over a season constituted the 'test' treatment and burrows checked twice a season the 'control' treatment. While the control burrows were checked twice per season, the individual birds in the burrows were disturbed only once (incubating adults in December and nearly fledged chicks in April). We recognise that a true control would involve no disturbance, however due to the nature of the study and the method of data collection, this was not possible. We also categorised quadrats as either stable or unstable based on the fragility of the burrows and the surrounding substrate. A quadrat was considered unstable if there was soil movement of the burrow roof in greater than two burrows, and otherwise stable. Only low numbers of little penguins were found in these quadrats and their breeding was not synchronised. This precluded their use in this analysis.

Differences between years in the number of shearwater eggs laid were examined to ascertain any ongoing effect on nest-site fidelity, indicated by decreasing numbers of eggs being laid each year, particularly in the intensively checked areas.

Researcher activity 2 — Chick handling

Each season, short-tailed shearwater burrows (n = 50) were randomly assigned to either a control group (n = 25) or a test group (n = 25). Control chicks for each species were weighed, as follows, once on their initial capture and again prior to fledging. Test chicks were weighed an additional two times for a total of four measurements throughout development. Chicks were removed from burrows and weighed in a cloth bag using a 1- or 2-kg Salter® spring balance (Super Samson models, Salter Australia Pty Ltd, Melbourne, Australia) whereas head length (from the tip of the beak to the occipital condyle), beak length (from the tip of the beak to the base where the feathers begin), beak width and beak depth (from just behind the nares) were measured to within ± 0.05 mm with Vernier callipers. Handling time (from extraction from the burrow to return) averaged approximately 3 min and was constant for the duration of the study.

Our research permit allowed the use of 50 chicks of each species as opposed to 50 nests, therefore the number of little penguin burrows used for chick growth-rate measurements varied across seasons due to the uneven number of eggs laid by individual parents. Consequently, to weigh 50 penguin chicks, 22, 29 and 27 little penguin burrows (years 2003, 2005 and 2006, respectively) were selected. These chicks were weighed and measured using the same techniques as the shearwater chicks with analogous handling times. For little penguins, years 2007 and 2008 were excluded from the analysis due to extreme breeding asynchrony resulting in only control-level handling for these two years.

Statistical analysis

Researcher activity level 1 — Nest checking (short-tailed shearwaters only)

While it is likely that environmental factors such as year and habitat stability will affect fledging and rates of egg laying, for this study we were primarily interested in the effect of the investigator.

We compared these two levels of disturbance using generalised linear models (GLM) in the R statistical programme (Pinheiro *et al* 2009). The model fixed terms were continuous: 'fledge' (the number of fledged birds per quadrat); 'eggs' (the number of eggs laid per quadrat); and 'year' (2003, 2005, 2006, 2007, 2008, 2009). There were also two binary terms: 'treatment' (1 = control, 2 = treatment, n = 201 and 473, respectively) and 'stability' (1 = unstable, 0 = stable). The stability parameter was included as it occurred, irrespective of transect location and provided a correction factor for any effects resulting from the non-random sampling regime that was employed to reduce any effects of frequently traversing unstable areas (see *Researcher activity 1 — Nest checking* in *Materials and methods*). The number of birds fledged was used as the response variable and models were built to test both additive and interaction terms. Each model was constructed using a Poisson-error distribution and a log-link function as examination of the residuals revealed these to be the most appropriate.

Bayesian Information criteria (BIC and ΔBIC) scores were used to rank models and determine the most parsimonious model as they provide greater strength when the sample size is large and contains a small number of variables with potentially large effects (Burnham & Anderson 2004). Generally, ΔBIC values of: 1) < 2 implying that the models have substantial support; 2) $2 < \Delta BIC < 7$ showing some support; and 3) $\Delta BIC > 7$ showing no support for that model (Burnham & Anderson 2001). As short-tailed shearwaters exhibit a high degree of nest-site fidelity and the same burrows were checked every year (with the exclusion of 2003), it is likely that the same breeding pairs are being disturbed every year. Consequently, if investigator effects are cumulative over time we might expect to detect a decrease in the number of eggs laid per quadrat in the areas with a higher rate of disturbance. Mean number of eggs laid per quadrat were compared using the same GLM approach as for fledging rates.

Researcher activity level 2 — Chick handling (short-tailed shearwaters and little penguins)

Survival (n = 50, 50, 50, 47) and fledging mass of surviving shearwater chicks (n = 25, 28, 36 and 30) were compared using a two-way ANOVA between treatment and year, and little penguin fledging mass was compared across groups using the same GLM as for researcher activity level 1 with corrected Akaike's information criterion (AICc and ΔAICc) scores used to rank models in this case due to the smaller sample sizes. The GLM was used for the little penguin analysis as most breeding pairs of little penguins lay two

Table 2 Model Bayesian Information Criterion (BIC), ΔBIC, wBIC and percentage deviance weights for short-tailed shearwater chicks fledged and eggs laid per quadrat across year, quadrat stability and nest checking.

Model	BIC	ΔBIC	wBIC	Pcdev
fledge~stability + year	1,746	0	0.9549	12.03
fledge~checking + stability + year	1,752.1	6.11	0.0451	12.05
fledge~year	1,771.1	25.05	0	8.03
fledge~year × stability	1,771.7	25.67	0	12.72
fledge~stability	1,776.2	30.18	0	4.19
fledge~checking + year	1,776.7	30.73	0	8.1
egg~stability + year	686.2	0	0.4922	17
egg~checking + stability + year	686.3	0.16	0.4538	19.09
egg~year × stability	691.2	5.07	0.0389	17
egg~year × checking × stability	693.3	7.14	0.0138	24.77
egg~year	698.6	12.41	0.001	9.56
egg~checking + year	701.9	15.72	0.0002	10.31

eggs and often raise two chicks, and therefore comparing fledging weight between chicks using ANOVA would result in pseudo-replication. Survival for handled little penguins was not calculated due to the infrequency of sampling meaning we could not determine with certainty if a chick had died between visits or had fledged prior to us returning and therefore we also interpret the mass comparison results with some caution, recognising that there will be low power associated with this analysis. Fledging mass of little penguins was expressed as mean surviving fledging mass per burrow (n = 14, 16, 13).

Results

Researcher activity level 1 — Nest checking (short-tailed shearwaters)

Egg laying

None of the models were good at explaining egg laying with the best model, accounting for 15.8% of the data variability (percent deviance). The most strongly supported model was egg~stability + year. There was some low-level support for the inclusion of the disturbance term (egg~treatment + year, ΔBIC = 4.14, % deviance = 16.05; Table 2). While this may indicate a disturbance effect, analysis of the model coefficients show that there were only a lower number of eggs laid in the intensively checked area in 2006 (estimate; $P = 0.003$). Fledging success for 2006 in the more intensively checked area was not significantly lower than in any other year and the effect has not been observed since that year. Given the trend for reduced numbers of eggs being laid across the entire colony (Figure 1[b]), it is likely that the lower rate of egg laying in 2006 was not due to the investigator, but is due to an effect not analysed by the current study.

Fledging success

As with egg laying, none of the models explained fledging success (fledge) well, with the best ranked model accounting for only 12.02% of the data variability (percent deviance). The model with the strongest support was fledge~stability + year (Table 2) indicating that the number of chicks fledged was affected most strongly by the stability of the quadrat and the year of the study. This indicates that there were years of relatively poor performance when the effect of the investigator might have been exacerbated (Figure 1[a]). However, there was little support for a disturbance effect during nest checking in any year of the study with the nest-checking parameter actually ranked beneath the null model according to the ΔBIC values (Table 2).

Researcher activity level 2 — Chick handling

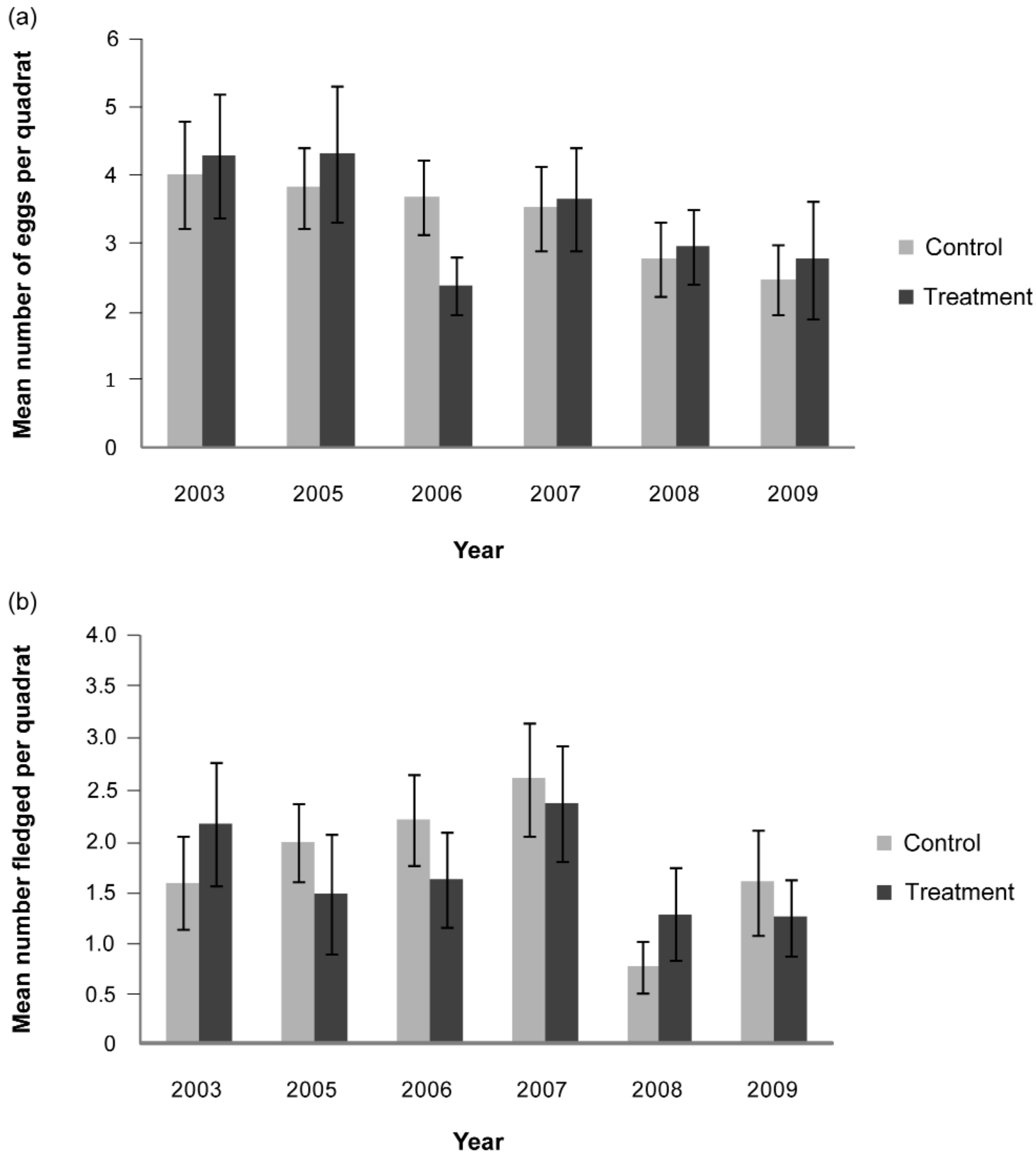
Survival of handled shearwater chicks was related to the level of handling ($F_{3,204} = 2.97$, $P = 0.034$), but this also varied among years (year × handling, $P = 0.03$ and this result is largely driven by the low survival of control chicks (ie those only handled once) in a single year (2003). As handling did not produce a significant result for subsequent years ($F_{3,154} = 1.05$, $P = 0.3$) it seems unlikely that in 2003 handling produced higher survival and this result may be due to an effect not currently analysed in this study (Figure 2).

Fledging mass of short-tailed shearwaters showed no significant difference in either intensity of handling or among years ([$F_{3,114} = 2.255$, $P = 0.086$] year × handling, $P > 0.1$; year, $P > 0.1$; treatment, $P > 0.1$) (Figure 3). For penguin fledging mass, there was virtually no difference in support between the first three models (ΔAICc > 2), mass~year × treatment, mass~year + treatment, mass~year (Table 3) which may suggest an effect of handling on the fledging mass of little penguin chicks in certain years. On examination of the coefficients, only in 2005 did birds handled four times fledge at a lower mass than those handled twice (coefficient estimate = –203.09 [± 89.25], $P = 0.02$), however, in 2003 and 2006, the reverse was true (Figure 3). Given the low power of this analysis, a greater sample size would be needed to ascertain if handling four times during a season is having a negative or positive effect on fledging mass or if this an artefact of the low sample size. The model set showed no support for the effect of treatment on its own (mass~treatment (ΔAICc = 7.08, pcdev = 6.5; Table 3) and overall fledging mass was not lower than average in the year that resulted in a slightly lower fledging mass for handled birds.

Discussion

All studies that involve researchers interacting with wild populations of animals should quantify the effects of their activities. Any negative effects need to be considered with respect to the relative importance of the study, for the overall well-being of the population or species, and with respect to the resulting integrity of the data themselves. In our case, we wanted to quantify the effects of environmental stochasticity on life-history parameters but to effectively interpret our results we needed to know firstly of any

Figure 1



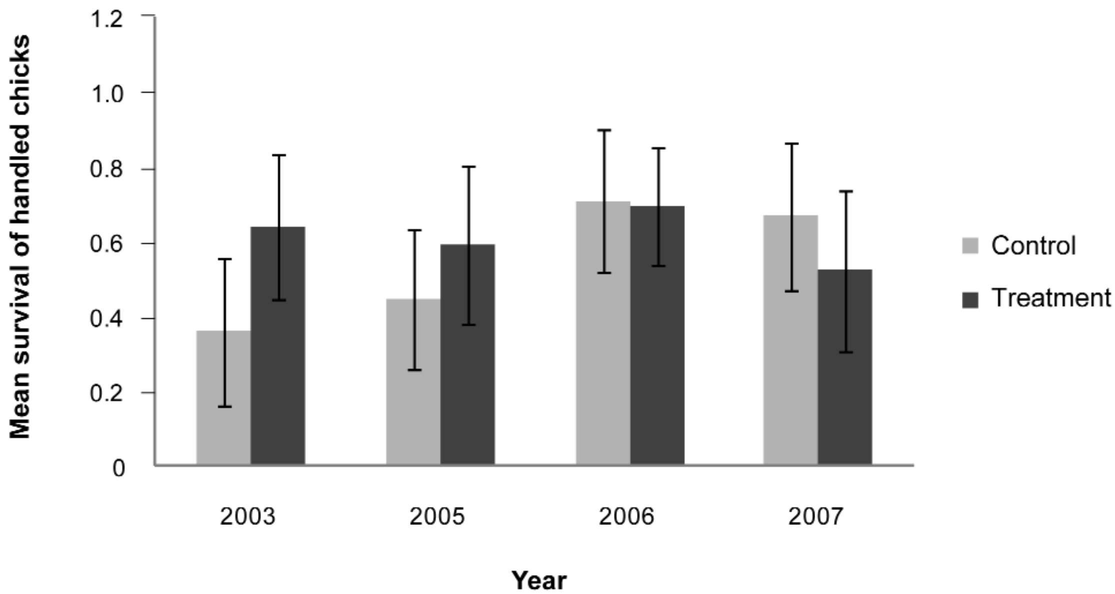
Mean number of (a) short-tailed shearwater eggs laid per quadrat for 2003 and 2009 between twice checked nests (control) and nests checked seven times (treatment) and b) short-tailed shearwater chicks fledged per quadrat between 2003 and 2009 comparing nest checks twice per season (control) and nest checks seven times per season (treatment). The error bars represent the 95% confidence intervals.

confounding effects due to our research activities. Equally important, these effects (or lack thereof) need to be reported, to help produce informed scientific and management decisions for future studies (Wilson & McMahon 2006; Carey 2009).

We have demonstrated that, despite researchers being regularly present in a large seabird colony throughout the breeding season, there were no detectable negative effects of research activities on three key life-history parameters; egg laying, survival of chicks and fledging mass in short-tailed shearwaters although there may be a variable effect on fledging mass in little penguins. However, all observed levels of survival and fledging mass in this study fall within, or exceed the expected norms for both species (short-tailed shearwater survival: 0.4–0.8 [Wooller *et al* 1990], fledging mass 700–800 g [Hamer *et al* 1997], little penguin fledging mass 800–1,000 g [Reilly & Cullen 1981]).

While the levels of disturbance for adult birds were low even at the 'test' level, the observations were consistent over the entire five years of study, which included years of

Figure 2



Mean survival of handled shearwater chicks from 2003–07 (n = 50, n = 50, n = 50, n = 47, respectively) of short-tailed shearwater chicks between birds handled twice (control) and birds handled four times (treatment). The error bars represent the 95% confidence intervals.

reduced egg laying and fledging of chicks. This is an important finding as it shows that even when environmental conditions are poor and researcher effects would be exacerbated, we were unable to observe any negative effects of the research activities on short-tailed shearwaters.

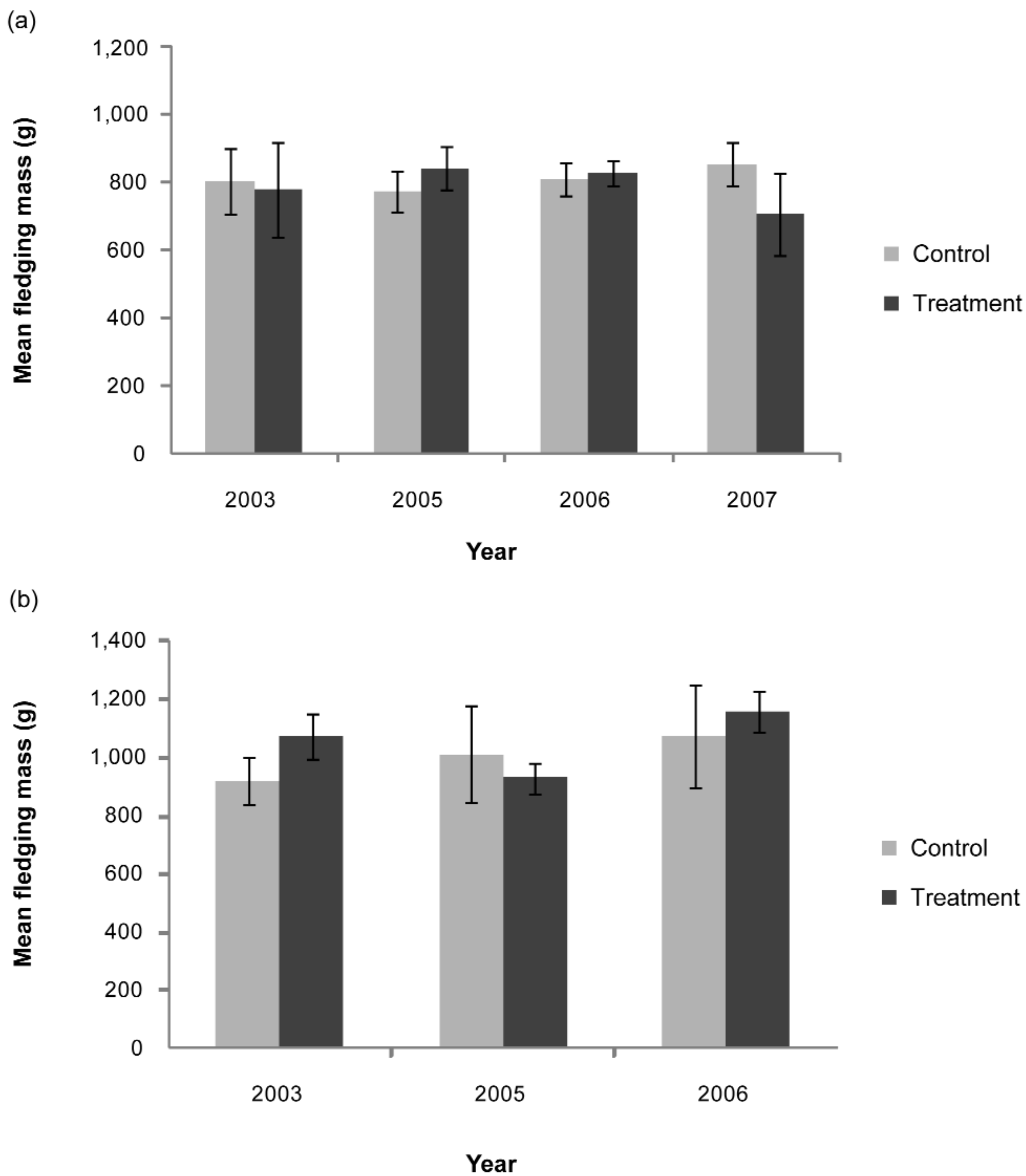
While these results were clear for the short-tailed shearwaters, the results for the little penguins were less so. We detected inconsistent effects in the fledging masses of little penguins that are biologically difficult to interpret. The key reason for this is two-fold. The sample size was not high enough and the rate of sampling was inadequate. For a species such as the short-tailed shearwater, that exhibits synchronous breeding, the sampling regime employed provided enough power to draw conclusions on the effect of investigators. However, the breeding asynchrony of little penguins meant that four-weekly checks were not adequate to establish accurately fledging rates or obtain enough points to calculate growth curves and the sample sizes of 50 chicks (rather than 50 nests) resulted in pseudo-replication of the data for analyses.

We cannot conclude that we had either a positive or negative effect on the fledging mass of little penguins, however, in order to confirm or refute our findings, the sampling rate would need to be higher and more frequent which would potentially impact negatively on the survival or fledging mass of the studied animals. It is perhaps this quandary that results in the under- testing and/or reporting of investigator effects on wild animal populations. Effects can be subtle and difficult to interpret as we have shown here. When any study is being designed it is important to consider the research questions, the breeding biology of the animal, the desired outcome, and then model the study on the minimum impact needed to answer them.

In previous studies (Table 4) that have measured the effects of basic researcher activity (such as nest checking or measuring) on vital rates such as breeding success and growth, results vary according to the species studied, habitat characteristics and the intensity of disturbance. For example, surface-nesting seabirds seem to be far more prone to researcher effects with 7 out of 13 reported procedures resulting in some kind of significant effect on survival and/or growth. These activities ranged in intensity from daily walking of transects (jackass penguin [*Spheniscus demersus*] [Burnham & Anderson 2001]) to daily nest checks (black-legged kittiwake [*Ryssa trydactyla*] [Sandvik & Barrett 2001]) to daily weighing and measuring (black skimmer [*Rynchops niger*], [Safina & Burger 1983]) and attachment of platform transmitter terminals on Adélie penguins (*Pygoscelis adeliae*) (Ballard *et al* 2001). Flipper banding is not discussed here as it has been covered extensively in the literature (see Jackson & Wilson 2002; Saraux *et al* 2011).

However, inter-study comparisons need to be approached and interpreted cautiously due to the varying intensity and nature of the research activity. For example, Hull and Wilson (1996) observed no investigator effect on two species of surface-nesting penguins (royal [*Eudyptes schlegeli*] and rockhopper [*Eudyptes crysocome*]), despite daily visits to the colonies whilst Hockey and Hallinan (1981) observed significant effects of investigators in jack-

Figure 3



The fledging masses of (a) short-tailed shearwaters from 2003–07 (n = 24, n = 29, n = 36 and n = 30, respectively) between birds handled twice (control) and birds handled four times (treatment) and (b) the fledging masses of little penguins from 2003–06 (n = 24, n = 21 and n = 18, respectively) between light handling (control) and more intense handling (treatment). The error bars represent the 95% confidence intervals.

ass penguins with a similar rate of disturbance. Most investigator effects vary within and between species and therefore should be assessed on a case-by-case basis (Casper 2009). Indeed, a recent review on human disturbance (both from tourism and research) in Antarctica by de Villiers (2008) also details species-specific responses to varying intensities of disturbance. However, it could be argued that even those species that have exhibited an investigator effect may find that effect greatly reduced or negated by modifying the sampling regime whilst still maintaining data integrity and usefulness.

While caution is needed when extrapolating from observations such as ours to other species, some generalities are warranted, especially given the consistency of the results over a number of years for short-tailed shearwaters. We would suggest that our sampling regime could provide a

Table 3 Model-corrected Akaike's nformation criterion (AICc), ΔAIC, wAIC and percentage deviance weights for little penguin fledging mass for year and level of handling (twice vs four times per season).

Model	AICc	ΔAICc	wAIC	pcdev
mass~year × handling	804.7	0.00	0.3821	28.67
mass~year + handling	805.1	0.39	0.3137	22.18
mass~year	805.2	0.56	0.2890	18.93
mass~handling	811.7	7.08	0.0111	6.55
mass~1	813.7	9.07	0.0041	0.00

Table 4 Summary of investigator disturbance studies (either published autonomously or as part of a general study) showing species, nesting habit, researcher activities undertaken during the study, recorded effects and whether or not the effect was due to the investigator.

Species	Nesting habit	Action	Growth or survival effect	Investigator effect
Atlantic puffins (*Fratercula arctica*)[1]	Burrowing	Weighing and measuring once during incubation, every 4 days prior to hatching and up to pre-fledging then every two days until fledging vs late in nesting then every two days until fledging 2 years duration	Yes	Yes
Burrowing owl (*Athene cunicularia*)[2]	Burrowing	Attachment of radio-transmitters and leg bands vs leg bands; testing on natal recruitment and fledging	No	No
Cory's shearwater (*Calonectris diomedea*)[3]	Burrowing	Weighing and measuring at 4-h intervals for three days vs once per day vs once at start and end	No	No
Hutton's shearwater (*Puffinus huttoni*)[4]	Burrowing	Burrows checked through inspection hatch every second day prior to laying, then every day for 4–10 days after laying, chicks weighed and measured every second day until fledging vs burrows checked three times through inspection hatches vs burrows checked three times with burrow-scope	No	No
Leach's storm petrel (*Oceanodrama leucorhoa*)[5]	Burrowing	Both parents weighed, banded (once), measured each day until fate of egg determined vs same procedures once a week vs visited once during incubation and once to determine hatching success. Burrows checked following year to determine ongoing effects	Yes	Yes
North Island little shearwater (*Puffinus assimilis haurakiensis*)[6]	Burrowing	Adults weighed every second day while incubating and chicks weighed after hatching	No	No
Short-tailed shearwater (*Puffinus tenuirostris*)[7,8,21]	Burrowing	Weighing and measuring at 4-h intervals for five nights over three periods and two years (120 measurements) vs weighed once. Weighing twice a day for 84 days, measured every 16 days (136 handles) vs 18–20 handlings over 84 days	No	
		Vs one weigh and measure	Yes	No
		Pre-hatching daily disturbance vs 3 daily vs weekly vs one disturbance	Yes	Yes
Wedge-tailed shearwater (*Puffinus pacificus*)[7]	Burrowing	Weighing twice a day for 84 days, measured every 16 days (136 handles) vs 18–20 handlings over 84 days vs one weigh and measure	No	No
Adelie penguin (*Pygoscelis adeliae*)[9]	Surface nesting	Three seasons, radio-transmitters (RT) attached with signal monitored from a distance vs flipper bands and PIT tags, monitored via weighbridge VS time depth recorder and RT attachment with all undergoing daily nest checks	No	No
		Stomach lavage	No	No
		Platform transmitter terminal (wt 170 g, area 800 mms^2) attachment	Yes	Yes
		TDR attachment (wt 50 g, area 520 mm^2) attachment	No	No

[1] (Rodway *et al* 1996), [2] (Conway & Garcia 2005), [3] (Hamer & Hill 1993), [4] (Cuthbert & Davis 2002), [5] (Blackmer *et al* 2004), [6] (Booth *et al* 2000), [7] (Schultz & Klomp 2000), [8] (Saffer *et al* 2000), [9] (Ballard *et al* 2001), [10] (Cairns 1980), [11] (Sandvik & Barrett 2001), [12] (Safina & Burger 1983), [13] (Verboven & Ens 2001), [14] (Hockey & Hallinan 1981), [15] (Hull & Wilson 1996), [16] (Fraser *et al* 1999), [17] (Davis Jr & Parsons 1991), [18] (Angelier *et al* 2011), [19] (Wheeler *et al* 2009), [20] (Woehler *et al* 2003), [21] (Carey 2011).

Table 4 (cont)

Species	Nesting habit	Action	Growth or survival effect	Investigator effect
Black-browed albatross (*Thalassarche melanophris*)[18]	Surface nesting	Blood sampling, handling time < 5 min, one bird per nest vs no handling or blood sampling	No	No
		Distance marked then nine birds blood sampled prior to foraging vs 7 not sampled	No	No
Black guillemots (*Cephus grille*)[10]	Surface nesting	Nest checks (including diameter, depth, particle size, overall shelter, density and distance from neighbour), daily vs nest checks every 4 days	Yes	Yes
Black-legged kittiwake (*Ryssa trydactyla*)[11]	Surface nesting	Daily nest checks, cross-fostering, weighing of chicks twice and weighing of adults twice vs nests checked once with chicks weighed, measured or counted	Yes (nest attendance) No (chick growth)	Yes
Black skimmer (*Rynchops niger*)[12]	Surface nesting	Weighing and measuring daily vs weekly until hatching and then daily vs weekly	Yes	Yes
Eurasian oystercatcher (*Haematopus ostralegus*)[13]	Surface nesting	Nest checking three times daily for four days vs once every other day	No	No
Jackass penguin (*Spheniscus demersus*)[14]	Surface nesting	One transect walked daily for 7 days vs one transect walked 2 hourly for 2 days vs one transect walked hourly for two days	Yes	Yes
Rockhopper penguin (*Eudyptes crysocome*)[15]	Surface nesting	Traversing transect plus handling of chicks and adults twice weekly vs no handling or traversing	No	No
Royal penguin (*Eudyptes schlegeli*)[15]	Surface nesting	Traversing transect plus handling of chicks and adults twice weekly vs no handling or traversing	No	No
Southern giant petrel (*Macronectus giganteus*)[20]	Surface nesting	Intensive banding programme at three colonies	Yes	Yes
Wandering albatross (*Dimodea exulans*)[19]	Surface nesting	Visual checks twice in three days vs four times in one day vs twice daily for three consecutive days	Yes	Yes
Crested auklets (*Aethia cristatella*)[16]	Crevice nesting	Visual checks two times a day vs every two days vs once a week of nests during incubation	No	No
		Chicks weighed every three days and at least one adult captured within six days of hatching vs chicks weighed every three days vs visual check of nests once a week. Study carried out for two years (1996, 1997)	Yes (in 1997, positive effect with higher disturbance resulting in higher fledging rates)	Yes
Least auklets (*Aethia pusilla*)[16]	Crevice nesting	Nest checks every day up to hatching then every 2–4 days up to fledging with weights and measurements of chicks vs inspected four times after locating eggs, no weights plus disturbance while checking other species vs different location checked twice after locating eggs, no weights, minimal disturbance	Yes (during incubation, no difference in fledging rates)	Yes
Snowy egret (*Egretta thula*)[17]	Tree nesting	Captured and weighed chicks every second day after hatching, banded at 7–10 days for 8–10 days after banding, VS chicks banded at 7–10 days. Both groups with regular nest checks	No	No

baseline for other burrowing seabird species in which sampling effort resulting in minimal disturbance with no discernable impact on growth rates or survival is still contributing useful data.

Overall, burrowing seabirds seem to be less susceptible to the effects of research activity than surface-nesting seabirds. Only three of the ten studies on burrowing seabirds reported an effect on growth, reproductive success, or nest-site fidelity (Table 4). In a study of Atlantic puffins (*Fratercula arctica*), which had a control level of disturbance that far exceeded our treatment level, Rodway *et al* (1996) found a 38% reduction in chick productivity which was persistent for one year following the conclusion of the study. In a study of Leach's storm petrel (*Oceanodroma leucorhoa*) (Blackmer *et al* 2004), with disturbance levels similar to those in our study, the researchers recorded a decrease in hatching success in birds checked daily and weekly compared to those checked only once during incubation and once after hatching. A recent study of short-tailed shearwaters (Carey 2011) reported a 100% reduction in hatching success in birds disturbed daily during incubation and a reduction in hatching success in birds disturbed every three days and every week compared to those disturbed once. However, previous work on short-tailed shearwaters (Saffer *et al* 2000), in which birds were weighed twice a day for 84 days (136 handling events), recorded no differences in fledging weights compared to those that were weighed and measured only once. There may be many habitat-related factors other than human disturbance that may affect reproductive performance of seabirds which include the stability of the substrate in which the birds are breeding, densities in colonies, and colony exposure (de Villiers 2008).

Animal welfare implications

As anthropogenic factors, such as climate change, commercial fishing, habitat encroachment and impacts of feral species intensifies, biologists are increasingly required to monitor, understand and quantify the changes these agents bring about on wild animal populations (Minteer & Collins 2008). However, long-term monitoring of wild animal populations often involves procedures that could be perceived as invasive to individual animals, such as banding or tagging (McMahon *et al* 2007). We, and others, would argue that the paucity of substantive research into these effects is a limiting factor in a balanced discussion of these issues (Wilson & McMahon 2006; Carey 2009).

A major impediment to improving the general understanding of investigator disturbance is the tendency for scientists not to publish non-significant results (Koricheva 2003). Consequently, many non-effects of investigator disturbance are mentioned only in passing, or relegated to one or two sentences in a more general paper (eg Hamer & Hill 1993; Booth *et al* 2000; Schultz & Klomp 2000; Cuthbert & Davis 2002). If these results remain unpublished then the *corpus* of knowledge does not develop and decisions and debates remain uninformed. Ideally, all research on wildlife should include a component to assess if the research is likely to have detrimental impacts. The inclusion of a well-designed, rigorous, and easily accessible component to research investigator effects in any new study would contribute greatly to the body of knowledge available to stakeholders when developing proposals for wild animal research.

Acknowledgements

Permission to carry out this research was provided by the University of Tasmania Ethics Committee (A0008138) and under permit (FA 05151) from the Nature Conservation Branch of Tasmanian Parks and Wildlife Service. We thank the many dedicated volunteers for assistance in the field and Adam Stephens, Richard Holmes, Simon Talbot, Anthony Reid and Thomas Sloane for logistical support in getting to and from the island and Mary-Anne Lea, Iain Field, Greg Jackman and two anonymous reviewers for their helpful and insightful comments.

References

Angelier F, Weimerskirch H and Chastel O 2011 Capture and blood sampling do not affect foraging behaviour, breeding success and return rate of a large seabird: the black-browed albatross. *Polar Biology 34*: 353-361

Ballard G, Ainley DG, Ribic CA and Barton KR 2001 Effect of instrument attachment and other factors on foraging trip duration and nesting success of Adélie penguins. *Condor 103(3)*: 481-490

Beale CM and Monaghan P 2004 Human disturbance: people as predation-free predators? *Ecology 41*: 335-343

Blackmer AL, Ackerman JT and Nevitt GA 2004 Effects of investigator disturbance on hatching success and nest-site fidelity in a long-lived seabird, Leach's storm-petrel. *Biological Conservation 116*: 141-148

Booth AM, Minot EO, Imber MJ and Fordham RA 2000 Aspects of the breeding ecology of the North Island little shearwater *Puffinus assimilis* haurakiensis. *New Zealand Journal of Zoology 27*: 335-345

Burnham KP and Anderson DR 2001 Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research 28*: 111-120

Burnham KP and Anderson DR 2004 Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research 33*: 261

Cairns D 1980 Nesting density, habitat structure and human disturbance as factors in Black Guillemot reproduction. *The Wilson Bulletin 92(3)*: 352-361

Carey M 2009 The effects of investigator disturbance on procellariiform seabirds: a review. *New Zealand Journal of Zoology 36*: 367-377

Carey M 2011 Investigator disturbance reduces reproductive success in short-tailed shearwaters *Puffinus tenuirostris*. *Ibis 153*(2): 363-372

Casper R 2009 Guidelines for the instrumentation of wild birds and mammals. *Animal Behaviour 78*: 1477-1483

Conway CJ and Garcia V 2005 Research notes: effects of radiotransmitters on natal recruitment of burrowing owls. *Journal of Wildlife Management 69*: 404-408

Cuthbert R and Davis LS 2002 Adult survival and productivity of Hutton's Shearwaters. *Ibis 144*: 423-432

Dalton R 2005 Animal-rights group sues over 'disturbing'work on sea lions. *Nature 436*: 315

Dalton R 2006 Sea-lion studies come to a halt after court judgement. *Nature 441*: 677

Davis Jr WE and Parsons KC 1991 Effects of investigator disturbance on the survival of snowy egret nestlings. *Journal of Field Ornithology 62(4)*: 432-435

de Villiers M 2008 *Human disturbance to wildlife in the broader Antarctic region: a review of findings*. http://www.scar.org/treaty/atcmxxxi/index.html

Fraser G, Jones IL, Williams JC, Hunter FM and Byrd GV 1999 Breeding biology of crested auklets at Buldir and Kasatochi Islands, Alaska. *The Auk 116(3)*: 690-701

Frid A and Dill LM 2002 Human-caused disturbance stimuli as a form of predation risk. *Conservation Ecology 6*: 11

Gill JA 2007 Approaches to measuring the effects of human disturbance on birds. *Ibis 149*: 9-14

Green JJ and Bradshaw CJA 2004 The 'capacity to reason' in conservation biology and policy: the southern elephant seal branding controversy. *Journal for Nature Conservation 12*: 25-39

Hamer KC and Hill JK 1993 Variation and regulation of meal size and feeding frequency in Cory's shearwater *Calonectris diomedea*. *Journal of Animal Ecology 62*: 441-450

Hamer KC, Nicholson LW, Hill JK, Wooller RD and Bradley JS 1997 Nestling obesity in Procellariiform seabirds: temporal and stochastic variation in provisioning and growth of short-tailed shearwaters *Puffinus tenuirostris. Oecologia 12*(1): 4-11

Hockey PAR and Hallinan J 1981 Effect of human disturbance on the breeding behaviour of jackass penguins *Spheniscus demersus*. *South African Journal of Wildlife Research 11*: 59-62

Hull CL and Wilson J 1996 The effect of investigators on the breeding success of Royal, *Eudyptes schlegeli*, and Rockhopper Penguins, *E. chrysocome*, at Macquarie Island. *Polar Biology 16*: 335-337

Jackson S and Wilson RP 2002 The potential costs of flipper-bands to penguins. *Functional Ecology 16(1)*: 141-148

Koricheva J 2003 Non-significant results in ecology: a burden or a blessing in disguise? *Oikos 102*: 397

Marchant S, Higgins PJ and Ambrose SJ 1990 *Handbook of Australian, New Zealand & Antarctic Birds, Volume 1 Handbook of Australian, New Zealand, and Antarctic Birds Part A. Ratites to Petrels Part B. Australian Pelicans to Ducks.* Oxford University Press: Oxford, UK

McMahon CR, Bradshaw CJA and Hays GC 2006 Branding can be justified in vital conservation research. *Nature 439*: 392

McMahon CR, Bradshaw C and Hays GC 2007 Applying the heat to research techniques for species conservation. *Conservation Biology 21*: 271-273

McMahon CR, Field IC, Bradshaw CJA, White GC and Hindell MA 2008 Tracking and data-logging devices attached to elephant seals do not affect individual mass gain or survival. *Journal of Experimental Marine Biology and Ecology 360*: 71-77

Minteer BA and Collins JP 2008 Why we need an 'ecological ethics'. *Frontiers in Ecology and the Environment 3*: 332-337

Pinheiro J, Bates D, DebRoy S and Sarkar D 2009 R Development Core Team. *Nlme: Linear and Non-linear Mixed Effects Models. R Package Version 3.1-92.* http://CRAN.R-project.org/package=nlme

Reilly PN and Cullen JM 1981 The Little Penguin *Eudyptula minor* in Victoria II: breeding. *Emu 81(1)*: 1-19

Rodway MS, Montevecchi WA and Chardine JW 1996 Effects of investigator disturbance on breeding success of Atlantic puffins *Fratercula arctica*. *Biological Conservation 76*: 311-319

Saffer VM, Bradley JS, Wooller RD and Meathrel CE 2000 The effect of human activity on the growth rates of Short-tailed Shearwater *Puffinus tenuirostris* chicks. *Emu 100*: 49-53

Safina C and Burger J 1983 Effects of human disturbance on reproductive success in the Black Skimmer. *Condor 85*: 164-171

Sandvik H and Barrett RT 2001 Effect of investigator disturbance on the breeding success of the black-legged kittiwake. *Journal of Field Ornithology 72*: 30-42

Saraux C, Le Bohec C, Durant JM, Viblanc VA, Gauthier-Clerc M, Beaune D, Park YH, Yoccoz NG, Stenseth NC and Le Maho Y 2011 Reliability of flipper-banded penguins as indicators of climate change. *Nature 469*: 203-206

Schultz MA and Klomp NI 2000 Chick-provisioning behaviour of two shearwaters breeding in south-eastern Australia. *Austral Ecology 25*: 319-326

Skira I 1991 The short-tailed shearwater: a review of its biology. *Corella 15*: 45-52

Verboven N, Ens BJ and Dechesne S 2001 Effect of investigator disturbance on nest attendance and egg predation in Eurasian Oystercatchers. *The Auk 118(2)*: 503-508

Vertigan C 2010 *The life history of short-tailed shearwaters (*Puffinus tenuirostris*) in response to spatio-temporal environmental variation.* PhD Thesis, University of Tasmania, Australia

Weimerskirch H and Cherel Y 1998 Feeding ecology of short-tailed shearwaters: breeding in Tasmania and foraging in the Antarctic? *Marine Ecology Progress Series 167*: 261-274

Wheeler M, De Villiers M and Majiedt P 2009 The effect of frequency and nature of pedestrian approaches on the behaviour of wandering albatrosses at sub-Antarctic Marion Island. *Polar Biology 32*: 197-205

Wilson R and McMahon CR 2006 Measuring devices on wild animals: what constitutes acceptable practice? *Frontiers in Ecology and the Environment 4*: 147-154

Woehler E, Riddle M and Ribic C 2003 Long-term population trends in southern giant petrels in East Antarctica. In: Huiskes A, Gieskes W, Rozema J, Schorno R, van der Vies S and Wolff W (eds) *Antarctic Biology in a Global Context* pp 290-295. Backhuys Publishers: Leiden, The Netherlands

Wooller R, Bradley J, Skira I and Serventy D 1990 Reproductive success of short-tailed shearwaters *Puffinus tenuirostris* in relation to their age and breeding experience. *Journal of Animal Ecology 59(1)*:167-170