Pseudoreplication in species comparisons: do individual differences matter?

Wade B. Worthen[1]* & Lauren Hoke[2]

[1] Biology Department, Furman University, Greenville, SC 29613, USA; wade.worthen@furman.edu

[2] 1303 Fairmont Street, Durham, NC 27713, USA; lc.hoke518@gmail.com

* Corresponding author

Abstract. Pseudoreplication occurs in many behavioral studies of odonates, because a few unmarked individuals are sampled repeatedly and are used as estimators of the species' behavior. This can confound individual differences with species differences. Here, we tallied perches of marked and unmarked male libellulids on artificial perches of seven heights (10–120 cm). We estimated the effect of pseudoreplication on species-level contrasts of mean male perch height by comparing the results of four different analyses: 1) a nested ANOVA (analysis of variance) evaluating individual and species effects on perches by marked individuals; 2) a one-way ANOVA comparing species using mean perch heights of the same marked individuals; 3) a one-way ANOVA comparing species without regard to individuals (pseudoreplication of these marked individuals); and 4) a one-way ANOVA using perches by unmarked individuals observed at the same time (pooling across pseudoreplicated individuals in a larger independent data set). Species differences were qualitatively similar across all analyses, and mean perch heights computed on individual means, pooled (pseudoreplicated) data on marked individuals, and data on unmarked individuals were highly correlated. Pseudoreplication altered patterns slightly, but these effects were overcome in the larger data set on unmarked individuals.

Further key words. Dragonfly, Anisoptera, marking, pseudoreplication

Introduction

In many studies of habitat use, perch selection, or foraging behavior by different dragonfly species, the behaviors or abundances of unmarked individuals are tallied (Osborne & Samways 1996; Raab et al. 1996; De Marco & Resende 2002; May & Baird 2002; De Marco & Resende 2004; Worthen & Patrick 2004; Worthen & Jones 2007; Remsburg et al. 2008; Silva et al. 2010; Worthen & Morrow 2016). Each observation is treated as an independent observation, so five observations of one individual are treated the same as single observations of five different individuals. Resampling the same sampling unit – called 'pseudoreplication' – conflates sources of error and confounds treatment effects (Hurlbert 1984). The effect declines as multiple individuals of each species are sampled, but when one species is disproportionately represented by a few individuals, the degree to which the 'mean behavior of a species' is being described is a function of how representative these individuals are.

Because species consist of individuals that vary, differences between species should be measured against this variation within species. The ideal experimental design would be completely balanced, in which a standard number of individuals of each species are sampled a standard number of times (10 individuals within each species, sampled 10 different times each, for example). If these data meet the assumptions of parametric tests, then the correct analysis to use is a 'nested' ANOVA, which evaluates differences between species as a function of the variation between individuals within these species (Sokal & Rohlf 1995).

In order to discriminate between individuals and avoid pseudoreplication, it is common in odonate behavioral studies to uniquely mark individuals – usually by numbering a wing with a felt-tipped marker (Cordero-Rivera & Stoks 2008; but see Moore 1987). However, studying marked individuals may introduce problems, too. Because dragonflies are mobile and relatively short-lived (the median estimate of average longevity in the reproductive period for Anisoptera is 11.5 days; Corbet 1999: 302), 'resighting' rates of marked individuals can be low (10–30%, Forbes et al. 2004; Dolný et al. 2013; Khelifa et al. 2016) – drastically reducing sample sizes and statistical power. In addition, the process of wing-marking may reduce survivorship (Parr & Parr 1979; Grether 1997; Cordero-Rivera et al. 2002), further reducing sample size. Marking can also change behavior. Individuals captured and marked may be more likely to leave the immediate site (Cordero-Rivera & Stoks 2008), potentially changing relative abundances and perceived patterns of habitat use. In addition, wing-marking may inadvertently augment natural wing patterning and affect a variety of important behaviors, from mate choice and reproductive success (Grether 1996) to intra- and interspecific aggression (Anderson & Grether 2010, 2011; Anderson et al. 2011). Indeed, recent studies suggest that marking the abdomen rather than the wing might reduce marking-related mortality and reduce behavioral effects (Anderson et al. 2011).

So, it seems that both methodologies – studying marked or unmarked individuals – have their particular pitfalls. Here, we address the primary statistical pitfall of studying unmarked individuals by asking, "how much do individual differences matter for species-level contrasts?" We marked individual dragonflies, recorded their perch heights, and analyzed these data both with and without respect to individual identity. The results were compared to determine the effects of 'pooling' data across individuals on species-level comparisons, and were compared to perch heights of unmarked individuals perching at the same time.

Methods

Sampling procedure

We sampled dragonflies from 08-vi- till 08-vii-2010 at Furman Lake on the campus of Furman University in Greenville, SC, USA (34°55'35.99"N, 82°26'27.75"W). Four perching stations were established along the western shoreline of the lake at

10 m intervals. At each station, wooden dowels were placed 10 cm apart, 50 cm from shore, emerging 10, 20, 40, 60, 80, 100, and 120 cm above the water's surface. Within each station, the order of perch heights was randomized each day to equalize accessibility of perches over the course of the study – so short perches weren't always between two taller perches, for example. Dragonflies were collected from these perches and surrounding vegetation by aerial net each morning (approximately 9:00 am to noon, EST), Monday–Friday, as weather permitted. Captured individuals were weighed, marked (numbered on a wing with a Sharpie© marker), photographed, and released; the procedure took approximately three minutes.

Perch heights of marked and unmarked individuals were recorded for a two-hour interval each day, between 1:30 pm–4:00 pm EST (UTC-5), weather permitting. We separated the capture phase from the observation phase to minimize disturbance to perching dragonflies. Although 333 individuals from ten species were marked, only 43 individuals from five species were observed perching in our arrays and were included in the analyses of marked individuals: *Perithemis tenera* (Say, 1839), *Erythemis simplicicollis* (Say, 1839), *Pachydiplax longipennis* (Burmeister, 1839), *Libellula incesta* Hagen, 1861, and *Libellula luctuosa* Burmeister, 1839.

Statistical analyses

To assess the costs and benefits of pooling data from different individuals on species-level comparisons, we analyzed differences in mean perch height between species four ways. First, we conducted a nested ANOVA (analysis of variance) on the perch heights of marked individuals, including both 'species' and 'individual (species)' effects. The data set was not balanced: there were unequal numbers of observations among individuals, and unequal numbers of individuals among species. MacDonald (2014) suggests that a Satterthwaite correction is appropriate in this case, but states that this may increase the probability of a type I error (false positive); we chose not to employ the correction to maintain a more conservative test. Second, we calculated mean perch heights for each marked individual, and compared species using these mean values in a one-way ANOVA. The 'species' effect in this analysis would be the same as in a balanced nested ANOVA, but can be slightly different in an unbalanced data set like ours (MacDonald 2014). Third, we pooled perch heights of marked individuals, and compared mean perch height among species in a one-way ANOVA. This evaluates all perch events as independent, and mimics the situation where unmarked dragonflies are sampled. Typically, the benefit of observing unmarked individuals is that more individuals and more observations are included. So, we also conducted a fourth one-way ANOVA that compared mean perch heights among these species using observations on different, unmarked individuals observed during the same observation periods. The last three ANOVA were followed by post-hoc Tukey mean comparison tests to compare mean perch heights among species. (Both the nested ANOVA and ANOVA using

individual mean values compute species mean comparisons the same way, so only one is needed.) These mean perch heights, generated by different analyses, were compared using Pearson product-moment correlations. All statistical tests were performed using SPSS©, version 21 (IBM Corp. 2012).

Results and discussion

Of the 333 individuals that were marked, 75 were observed again: 43 on the experimental arrays and 32 in nearby vegetation. This resighting rate of 22 % is consistent with previous studies of libellulids, which typically range from ~10–30 % (Forbes et al. 2004; Chin & Taylor 2009; Dolný et al. 2013; Khelifa et al. 2016).

There was a highly significant 'species' effect ($p < 0.0001$) on mean perch height in all three analyses using marked individuals, regardless of whether individual variation was accounted for in the model (Nested ANOVA; Table 1a), averaged per individual (one-way ANOVA on mean perch height by individuals; Table 1b), or ignored (one-way ANOVA, pooled across marked individuals; Table 1c). There was also a highly significant 'species' effect ($p < 0.0001$) on mean perch height of unmarked individuals (Table 1d). In addition, patterns among species in mean perch height were largely consistent across analyses (Table 2). Using mean perch heights of marked individuals as sampling units to compute species means (which approximates the approach of a nested ANOVA), *Perithemis tenera* and *Erythemis simplicicollis* perched significantly lower than the other three species, which did not differ from one another (Table 2). When perch height values of marked individuals were pooled (pseudoreplicating without regard to individual), *P. tenera* and *E. simplicicollis* again perched significantly lower than the other species, and the difference between *Pachydiplax longipennis* and the two *Libellula* species was resolved as significant, as well (Table 2). The mean perch heights of all species were significantly different from one another in the largest data set of unmarked individuals (Table 2).

The mean perch heights computed in the two analyses on marked individuals were strongly correlated ($r = 0.992$, $df = 5$, $p < 0.0001$, one-tailed test), but there were some differences. In the pooled analysis that pseudoreplicated individual responses, mean perch height for *P. longipennis* was >10 cm lower, mean perch heights of the two *Libellula* species were 4–7 cm higher, and the rank order of the *Libellula* species was reversed relative to the more appropriate analysis using individual means. So, pseudoreplication introduced some error in the small data set of marked individuals.

The species' mean perch heights of unmarked individuals were perfectly correlated with species means computed on marked individual means ($r = 1.000$, $df = 5$, $p < 0.0001$). In addition, the species' means computed in these two analyses were remarkably similar, differing by less than 2.5 cm for all species.

These comparisons highlight the costs and benefits of observing marked or unmarked individuals. When analyses were limited to marked individuals, species

Table 1. ANOVA describing the variation between species in mean perch height using four different approaches: a) Nested ANOVA, describing 'Species' effects as a function of variation among marked individuals; b) one-way ANOVA, describing 'species' effects using mean perch heights of marked individuals; c) one-way ANOVA, pooling all data across marked individuals; d) one-way ANOVA on unmarked individuals.

Effect	df	MS	F	p
a) Nested ANOVA				
Species	4	112878.71	47.92	0.0001
Ind (Species)	38	2355.32	9.71	0.0001
error	463	242.59		
b) One-way ANOVA, Mean Perch Heights per Individual				
Species	4	8533.20	22.58	0.0001
error	38	377.99		
c) One-way ANOVA, pooled data on marked individuals				
Species	4	112878.70	280.21	0.0001
error	501	402.83		
d) One-way ANOVA, pooled data on unmarked individuals				
Species	4	1280624.06	23137.16	0.0001
error	3546	408.21		

Table 2. Comparisons of mean perch heights of libellulid (Odonata) species, using: mean perch height values of marked individuals (Marked, individual means); all perch height values, pooled across marked individuals (Marked, pooled); and perch heights of unmarked individuals (Unmarked, pooled). Means in a column that are followed by the same letter are not significantly different, $p = 0.05$, Tukey's Multiple Comparison tests).

	Marked, individual means			Marked, pooled			Unmarked, pooled		
Species	N	$\bar{x}$ ±1 sd		N	$\bar{x}$ ±1 sd		N	$\bar{x}$ ±1 sd	
Perithemis tenera	2	10.2±0.3	A	36	10.3±1.7	A	368	11.5±3.8	A
Erythemis simplicicollis	7	16.1±4.3	A	40	15.0±5.1	A	754	17.5±12.2	B
Pachydiplax longipennis	29	73.7±22.4	B	324	62.8±24.3	B	1073	73.5±29.0	C
Libellula luctuosa	2	106.8±9.6	B	39	113.1±9.6	C	819	104.4±18.9	D
Libellula incesta	3	108.6±8.9	B	67	112.3±9.9	C	537	109.8±16.2	E

level comparisons were constrained by the small sample size and some differences between species could not be resolved. And, in this small data set, pseudoreplication had a slight effect on species-level contrasts. Although we could not keep track of unmarked individuals, we know we sampled more unmarked individuals than marked individuals. We only resighted two different marked individuals of *P. tenera*

and *L. luctuosa*, but we often observed 4–5 unmarked individuals of these species at once. By sampling unmarked individuals, we increased the size of the sample, increased the number of individuals sampled, and reversed the negative effects of pseudoreplication: species means more closely matched those in the preferred analysis on marked individuals. So, large samples of unmarked individuals should minimize the effects of individual variation, overcome the effect of pseudoreplication, and may be suitable for species-level comparisons of perch selection. Other behaviors, of course, may show greater individual variation and may be affected by pseudoreplication to a greater degree.

Acknowledgements. This research was supported by an Undergraduate Research Grant from the consortium of South Carolina Independent Colleges and Universities (LH), and a grant from the Research and Professional Growth Committee, Furman University (WW).

References

Anderson C.N., Cordoba-Aguilar A., Drury J.P. & Grether G.F. 2011. An assessment of marking techniques for odonates in the family Calopterygidae. *Entomologia experimentalis et applicata* 141: 258-261

Anderson C.N. & Grether G.F. 2010. Interspecific aggression and character displacement of competitor recognition in *Hetaerina* damselflies. *Proceedings of the Royal Society of London* (B, Biological Sciences) 277: 549-555

Anderson C.N. & Grether G.F. 2011. Multiple routes to reduced interspecific territorial fighting in *Hetaerina* damselflies. *Behavioral Ecology* 22: 527-534

Chin K.S. & Taylor P.D. 2009. Interactive effects of distance and matrix on the movements of a peatland dragonfly. *Ecography* 32: 715-722

Corbet P.S. 1999. Dragonflies: behavior and ecology of Odonata. Cornell University Press, Ithaca, New York

Cordero-Rivera A. & Stoks R. 2008. Mark-recapture studies and demography. In: Cordoba-Aguilar A. (Ed.), Dragonflies and damselflies: Model organisms for ecological and evolutionary research: 7-20. Oxford University Press, Oxford

Cordero-Rivera A., Egido-Perez F.J. & Andres J.A. 2002. The effect of handling damage, mobility, body size, and fluctuating asymmetry on lifetime mating success of *Ischnura graellsii* (Rambur) (Zygoptera: Coenagrionidae). *Odonatologica* 31: 117-128

De Marco Jr P. & Resende D.C. 2002. Activity patterns and thermoregulation in a tropical dragonfly assemblage. *Odonatologica* 31: 129-138

De Marco Jr P. & Resende D.C. 2004. Cues for territory choice in two tropical dragonflies. *Neotropical Entomology* 33: 397-401

Dolný A., Mižičová H. & Harabiš F. 2013. Natal philopatry in four European species of dragonflies (Odonata:

Sympetrinae) and possible implications for conservation management. *Journal of Insect Conservation* 17: 821-829

Forbes M.R., Muma K. & Smith B.P. 2004. Recapture of male and female dragonflies in relation to parasitism by mites, time of season, wing length and wing symmetry. *Experimental & applied Acarology* 34: 79-93

Grether G.F. 1996. Sexual selection and survival selection on wing coloration and body size in the rubyspot damselfly *Hetaerina americana*. *Evolution* 50: 1939-1948

Grether G.F. 1997. Survival cost of an intrasexually selected ornament in a damselfly. *Proceedings of the Royal Society of London* (B, Biological Sciences) 264: 207-210

Hurlbert S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54: 187-211

IBM Corp. 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY, IBM Corporation

Kery M. & Juillerat L. 2004. Sex ratio estimation and survival analyses for *Orthetrum coerulescens* (Odonata, Libellulidae). *Canadian Journal of Zoology* 82: 399-406

Khelifa R., Mellal M.K., Zouaimia A., Amari H., Zebsa R., Bensouilah S., Laouar A. & Houhamdi M. 2016. On the restoration of the last relict population of a dragonfly *Urothemis edwardsii* Selys (Libellulidae: Odonata) in the Mediterranean. *Journal of Insect Conservation* 20: 797-805

May M.L. & Baird J.M. 2002. A comparison of foraging behavior in two "percher" dragonflies, *Pachydiplax longipennis* and *Erythemis simplicicollis* (Odonata: Libellulidae). *Journal of Insect Behavior* 15: 765-778

McDonald J.H. 2014. Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, Baltimore, Maryland

Moore A.J. 1987. The behavioral ecology of *Libellula luctuosa* (Burmeister) (Anisoptera: Libellulidae): I. Temporal changes in the population density and the effects on male territorial behavior. *Ethology* 75: 246-254

Osborn R. & Samways M.J. 1996. Determinants of adult dragonfly assemblage patterns at new ponds in South Africa. *Odonatologica* 25: 49-58

Parr M.J. & Parr M. 1979. Some observations on *Ceriagrion tenellum* (De Villers) in Southern England (Zygoptera: Coenagrionidae). *Odonatologica* 8: 171-194

Raab R., Chovanec A. & Wiener A.K. 1996. Aspects of habitat selection by adult dragonflies at a newly created pond in Vienna, Austria. *Odonatologica* 25: 387-390

Remsburg A., Olson A.C. & Samways M.J. 2008. Shade alone reduces adult dragonfly (Odonata: Libellulidae) abundance. *Journal of Insect Behavior* 21: 460-468

Silva D., De Marco P. & Resende D.C. 2010. Adult odonate abundance and community assemblage measures as indicators of stream ecological integrity:

A case study. *Ecological Indicators* 10: 744-752

Sokal R.R. & Rohlf F.J. 1995. Biometry, 3rd ed. W.H. Freeman and Co., New York

Worthen W.B. & Jones C.M. 2007. The effects of wind speed, competition, and body size on perch height selection in a guild of Libellulidae species (Odonata). *International Journal of Odonatology* 10: 257-272

Worthen W.B. & Morrow P.H. 2016. Perch selection by three co-occurring species of *Celithemis* (Odonata: Libellulidae): testing for a competitive hierarchy among similar species. *Psyche* 2016: 1-8

Worthen W.B. & Patrick E.R. 2004. Competitive interactions affect perch-height preferences of three Odonata taxa (Coenagrionidae, Libellulidae). *International Journal of Odonatology* 7: 529-541

Received 5th October 2017