

Simulating Selective Mortality on Tadpole Populations in the Lab Yields Improved Estimates of Effect Size in Nature

STEVEN D. MELVIN^{1,2} AND JEFF E. HOULAHAN¹

¹*Department of Biology, University of New Brunswick, Saint John, New Brunswick, Canada E2L 4L5*

ABSTRACT.—Many populations normally experience high levels of mortality throughout larval development, but this is generally overlooked with laboratory experimental protocols. Evidence suggests that mortality is nonrandom in natural tadpole populations, so high survivorship, typical of laboratory populations, may poorly represent populations in nature. We compared survival, growth and development, and population variance of tadpoles in natural ponds with those in the laboratory at low and high densities. In the laboratory, high-density groups were reared with no selection and with selection imposed against different size classes to identify if, and how, mortality influences natural tadpole populations and to investigate whether imposing selection against certain size classes produces responses more consistent with those observed in natural systems. Our results suggest that selective mortality removes smaller individuals in natural populations. We demonstrate that introducing selection against small individuals artificially, in the laboratory, results in individual growth and development, population variance, and statistical power that more closely resembles that observed in natural populations. This is important from an ecological perspective because it demonstrates how selection acts on natural tadpole populations. More importantly, this demonstrates that laboratory experiments can be designed to provide better qualitative estimates for responses of natural populations by considering and simulating natural rates of mortality.

Laboratory experiments are often implemented as surrogates for whole-ecosystem studies (Melvin and Houlihan, 2012) because they allow for the controlled manipulation of environmental variables and provide results that are generally easier to analyze and interpret. While it is acknowledged that studies performed in artificial experimental venues may have limited relevance for responses in natural systems (Cairns, 1983; Schindler, 1987; Carpenter, 1996, 1999; Skelly, 2002, 2005), the ability to control for complex biological interactions and conditions makes laboratory studies extremely useful tools. Therefore, research aimed at increasing the relevance of laboratory studies whilst maintaining a high degree of experimental control could offer a major improvement to controlled laboratory protocols.

The issue of the reduced relevance of laboratory populations has been of particular interest for amphibian community ecology (Hairston, 1989; Jaeger and Walls, 1989; Morin, 1989; Skelly and Kiesecker, 2001; Skelly, 2002, 2005; Chalcraft et al., 2005), but these concerns may be transferrable broadly to other species and systems. Reservations about laboratory studies have focused primarily on the fact that they do not reflect the true complexity experienced by natural populations because they represent very simplified habitat-types, exclude many natural biotic and abiotic interactions, and normally investigate smaller spatial and shorter temporal scales than are of concern in natural systems (Jaeger and Walls, 1989; Chalcraft et al., 2005). However, an additional unexplored issue is that individuals in laboratory populations may not be similar to those in the natural populations they are intended to represent because selective forces that work to structure natural populations are generally excluded from laboratory studies.

Few would disagree that laboratory studies are useful for studying biological interactions under controlled conditions. However, control over experimental conditions is typically realized at the cost of environmental realism (Morin, 2011), and the size and structure of laboratory populations is rarely

representative of natural populations (Reznick and Ghalambor, 2005). Indeed, differences in effect sizes amongst experiments carried out in different settings have been described (Skelly, 2002) and have been broadly attributed to differences in the size and structure of the experimental venues. While this may be true, a more specific explanation is that experiments at small spatial scales with low complexity often have unnaturally high population survival (Melvin and Houlihan, 2012) and, so, populations in these settings may be dissimilar from natural populations that experience much higher mortality. Our recent meta-analysis demonstrates that tadpole mortality is normally very low in laboratory populations (12%) and increases with the scale and complexity of the experimental venue, from mesocosm (34%) to in situ field enclosures (64%) and whole-ecosystem field studies (95%) (Melvin and Houlihan, 2012). In addition, tadpole size in natural populations tends to be very homogeneous compared to laboratory populations, indicating that mortality in natural tadpole populations may be nonrandom and related to size (Melvin and Houlihan, 2012). It has been suggested previously that small individuals suffer greater predation in natural tadpole populations (Smith, 1983), but we are aware of no study that has investigated if the hypothesized selection of small individuals helps to explicitly explain qualitative differences in the magnitude of laboratory and natural population responses.

The first objective of this study was to test the hypothesis that mortality is nonrandom in natural populations of developing tadpoles. We predicted that high mortality in field populations corresponds with tadpole size distributions that are skewed positively and that growth, development, and population variance will be more similar between lab and field populations if size-dependent mortality is incorporated into the laboratory experimental design. The second objective was to investigate whether incorporating artificial selection pressures that are reflective of natural population mortality into a laboratory experiment results in estimates of density-dependence effects that are (qualitatively) more representative of results that we see in natural systems. We predicted that incorporating size selection that is reflective of natural population mortality into the laboratory experimental design would result in more-similar estimates of the effects of density dependence between

²Corresponding author. Present Address: Central Queensland University, School of Medical and Applied Sciences, Queensland, 4680, Australia. E-mail: s.d.melvin@cqu.edu.au or smelvin18@gmail.com

DOI: 10.1670/12-206

laboratory and natural systems. A final objective was to investigate whether the introduction of size selection into laboratory populations influences the statistical power of these protocols to detect small effect sizes.

MATERIALS AND METHODS

Laboratory and field experiments were carried out simultaneously to facilitate comparison of developmental profiles of laboratory populations with those of natural Wood Frog (*Lithobates sylvaticus*) populations. Egg masses for the laboratory experiment were collected from the same experimental ponds used for the field experiment to reduce the potential for genetic variability across the different venue types.

Laboratory Experiment.—Egg masses were collected from ponds located at the experimental wetlands area of Canadian Forces Base (CFB) Gagetown, New Brunswick, Canada (45°40′53.32″N, 66°29′49.08″W) and transported to the Canadian Rivers Institute wet-labs at the University of New Brunswick, Saint John, New Brunswick in buckets filled with natural pond water. Three clutches were maintained in 20-L aquaria filled with dechlorinated tap water at natural ambient temperatures until the majority had hatched. When the majority of tadpoles reached Gosner stage 26 (Gosner, 1960), individuals were haphazardly allocated to their experimental aquaria (20 L filled with 10 L of dechlorinated water). There were six replicates each of five treatments: 1) low density (1 tadpole/L), 2) high density (5 tadpoles/L), 3) high density with random selection, 4) high density with selection against small individuals, and 5) high density with selection against large animals. Selection was carried out weekly with two tadpoles removed biweekly for the first 6 weeks and one tadpole for the remainder of the experiment, such that 60% of the original population was removed from the selection-pressure treatment groups over the developmental period. This number represents a conservative estimate of population mortality based on our recent meta-analysis (Melvin and Houlihan, 2012).

Tadpoles were fed ad libitum twice weekly with rabbit pellets supplemented with boiled kale, and biweekly water changes were performed to remove feces and food waste. Ten tadpoles were haphazardly removed from each aquarium once per week and placed onto a 0.5-cm grid in a plastic tray where they were photographed for length measurements (snout-vent length [SVL]; full body length) using ImageJ™ image analysis software (National Institutes of Health, USA). Tadpoles were then developmentally staged (Gosner, 1960) and returned to their experimental aquaria. Handling for measurements was performed with care and resulted in no mortality or observable physical damage.

Field Experiment.—Four natural ephemeral ponds located at CFB Gagetown, New Brunswick, Canada (45°40′53.32″N, 66°29′49.08″W) were divided in July 2008 by burying high-density polyethylene sheets vertically into the sediment. Ponds were intermediate in size (average dimensions = 30 m × 30 m; depth = 10–40 cm) and were observed to be breeding sites for Wood Frogs during preliminary observations in the spring of 2008. We manipulated larval Wood Frog density in the sections to achieve low- (~1 tadpole/L) and high-density (~5 tadpoles/L) treatments. We did this by removing or stocking (depending on whether we needed to add or remove egg masses to achieve target densities) freshly laid egg masses and then monitoring tadpole growth and development until metamorphosis. Ponds were checked daily throughout the breeding season to ensure

that egg masses were identified, and moved, within 24 h of being laid. To achieve target densities we estimated the volume of each section using measurements of length and width and average pond depth along transect lines (data not shown) and then added egg masses as required. We estimated the number of embryos per egg mass at 800 (Corn and Livo, 1989; Duellman and Trueb, 1994), resulting in 18–24 masses per section in our ponds. Ten tadpoles from each quadrat were selected randomly for growth measurements and determination of stage of development (Gosner, 1960). We intended to take length measurements (i.e., SVL and full body length) at approximately 2-week intervals, but decreased our sampling efforts to reduce our disturbance of the natural ponds. For measurements of growth, tadpoles were placed on a smooth 0.5-cm grid and photographed for analysis using ImageJ™ (National Institutes of Health, USA) image analysis software.

Ponds were dip-netted every 2–3 weeks throughout the larval developmental period, and the number of tadpoles collected per sweep was used to estimate tadpole density as an indication of survival. The same person sampled all ponds for the duration of the experiment, and care was taken to perform sweeps in a consistent manner. The area of the dip-net was approximately 582 cm² and the average length of a sweep was 94 cm; therefore, the volume of water sampled with each sweep was estimated to be 55 L. From this, we converted the number of tadpoles caught per sweep into estimates of tadpoles per liter. These density estimates are used to approximate abundances between pond halves and over time, which is indicative of population mortality. These are coarse estimates, and we emphasize that the purpose of this study was not to compare intraspecific competition between lab and field venues directly. The purpose of this study was to estimate how the responses of tadpoles in natural populations (with known high mortality) to manipulations of tadpole density compared to the same manipulations in a laboratory setting.

Data Analysis.—Frequency distributions for different tadpole size classes were graphed using Prism 5 for Mac OS X (GraphPad Software, Inc.). Distributions were used to visualize population size structure on days 20–24 of the experiment for the high- and low-density treatment groups because there was evidence that significant mortality had occurred in the natural ponds by day 20 (see Fig. 2). Distributions were analyzed for Skewness using SPSS version 17 (IBM Inc., Chicago, IL) statistical software. The z-table was used to assign statistical significance. Briefly, the measure of Skewness was divided by its standard error (SE) for each population and was deemed to deviate significantly from normal if this value was larger than 1.96 (5% significance threshold; Table 2).

Growth, development, and coefficient of variance (COV) data were analyzed using SPSS version 17 (IBM Inc.) statistical software. Mixed-model repeated-measures analysis of variance (ANOVA) (two-factor within-subjects) was used to investigate overall differences amongst treatment groups using the restricted maximum likelihood method of estimation. Tank averages were analyzed for each dependent variable (i.e., SVL, full body length, Gosner stage, COV) with density–selection treatment as a fixed factor and time as a random factor. Prior to the analysis, all data were tested for normality and homogeneity of variance using the Levene’s test (Levene, 1960) and normal probability plots. To satisfy the requirements of the repeated-measures ANOVA, the assumption of sphericity was tested using Mauchly’s Test (Mauchly, 1940). If violated, the Greenhouse-Geisser correction was applied to adjust the degrees of freedom

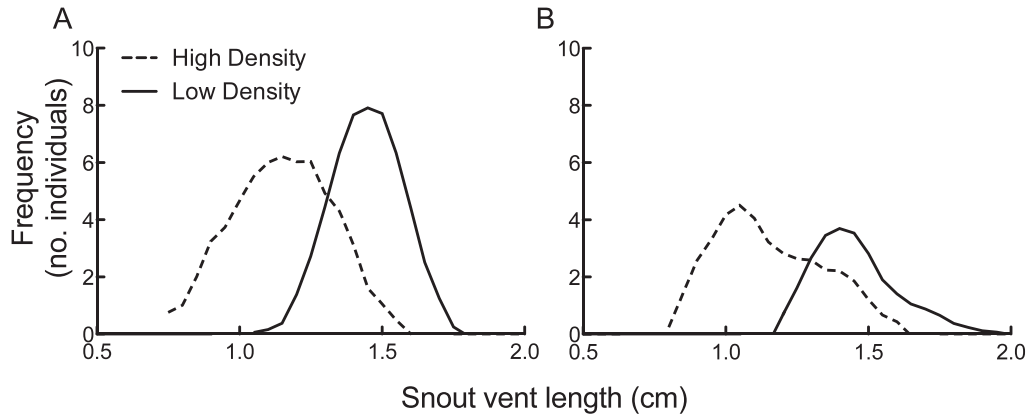


FIG. 1. Frequency distributions (no. individuals) of Wood Frog, *Lithobates sylvaticus*, tadpole size classes at experimental day 20–24 for individuals reared at low (1 tadpole/L) and high (5 tadpoles/L) stocking densities in (A) the laboratory and (B) natural ephemeral ponds.

such that a valid *F*-ratio was obtained, thereby reducing the Type I error rate (false positive). Differences amongst groups were analyzed using a Bonferonni post hoc test. Our primary question was, “Does incorporating artificial selection pressures that are reflective of natural population mortality into a laboratory experiment result in estimates of density dependence effects that are (qualitatively) more representative of results that we see in natural systems?” Therefore, our analysis focused on comparisons between each of the high-density treatments and the low-density treatment in the laboratory (although all groups were compared), and between high-density and low-density treatments in the field. The results of these independent analyses were then visually compared between laboratory and field populations for similarity in the magnitude of response.

We were interested in whether the introduction of size selection on populations in the laboratory influences their ability to detect significant differences in growth (i.e., SVL, full body length) when effect sizes are small. Therefore, we performed a power analysis to assess the probability of Type II error (false negative) for our experiments, using both the laboratory and field populations, to ensure that incorporating size selection in the laboratory does not negatively influence these protocols. For power analyses, we chose critical effect sizes

based on estimates of variability in natural populations (Munkittrick et al., 2009) to reflect potentially ecologically meaningful responses. Differences in tadpole size of approximately 10% were observed amongst the populations of Wood Frogs in our field experiments (Table 1), and evidence suggests differences in tadpole size as low as 20% elicit significant differences in predator success rates in Green Frog (*Rana clamitans*) populations (Formanowicz, 1986). Populations were analyzed for their power to detect changes in tadpole growth of 10%, 15%, and 20% using population means and standard deviations at the final measurement prior to metamorphic climax (Gosner stage 42; Gosner, 1960). We analyzed for power to detect differences between two independent groups (*t*-test; two-tailed) using G*Power™ software (alpha, 0.05). We performed two separate analyses; first with a sample size of four replicates, as this is a common benchmark used in laboratory experimentation and reflects the number of replicates of our field experiment, and second with a sample size of six replicates because this reflects the actual number of replicates used in our laboratory experiment.

RESULTS

Is Mortality Nonrandom in Natural Tadpole Populations?—Frequency distributions for laboratory populations were approximately normal at both high and low density (Fig. 1A), while distributions for field populations were skewed in a manner consistent with selective mortality acting upon individuals at the small end of the size distribution (Fig. 1B). Although both high- and low-density populations in the field had visibly skewed distributions, only the low-density population significantly deviated from normal (Table 2). Mortality was very low in laboratory populations (less than 5%) in all replicates of all treatment groups, while tadpole density estimates indicated high population mortality throughout the developmental period in the field (Fig. 2). Density estimates are coarse because of the avoidance of intense sampling in the natural ponds but are highly consistent with expected mortality based on the literature and with our personal observations during field experimentation.

Does Incorporating Selective Mortality in Laboratory Populations Produce Results that Are More Consistent with that of Natural Populations?—Differences in growth and development between high- and low-density groups in the laboratory were most consistent with field populations when small tadpoles were

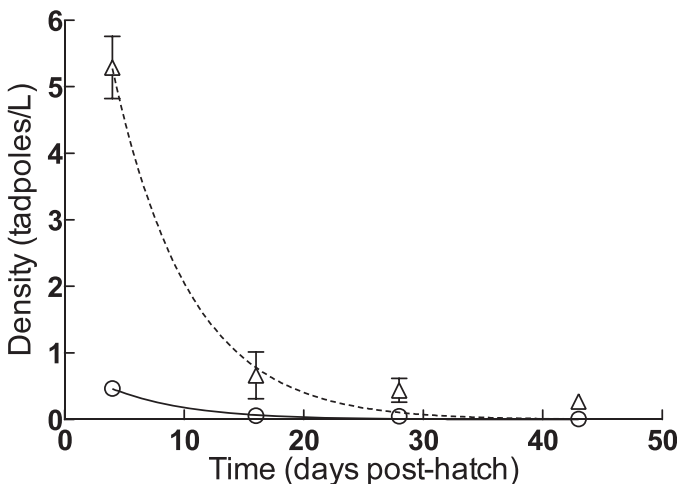


FIG. 2. Tadpole density estimates at low (circles; 1 tadpole/L) and high (triangles; 5 tadpoles/L) stocking density in natural ephemeral ponds throughout the larval developmental period. Bars represent 1 SD.

TABLE 1. Average growth and developmental stage of Wood Frog, *Lithobates sylvaticus*, tadpoles reared at low density and high density (5 tadpoles/L) with no selection, random selection, or selection against the largest or smallest tadpoles, at the last sample time prior to metamorphic climax. Measurements of tadpoles reared in a natural pond are also shown as well as the corresponding measurements of tadpoles reared at low density (1 tadpole/L). The magnitude of response (effect size %) amongst high- and low-density treatments is shown for field and laboratory populations with no selection pressures and for laboratory populations with imposed selection pressures.

Selection pressure	Snout-vent length (cm)			Full length (cm)			Gosner stage		
	High density	Low density	Effect size (%)	High density	Low density	Effect size (%)	High density	Low density	Effect size (%)
No selection	1.13	1.43	21.01	2.93	4.02	27.01	32.83	38.51	14.75
Large	1.16	1.43	18.85	3.10	4.02	22.87	33.88	38.51	12.03
Random selection	1.23	1.43	13.95	3.23	4.02	19.63	34.36	38.51	10.79
Small	1.26	1.43	11.85	3.35	4.02	16.65	36.73	38.51	4.63
Field	1.44	1.60	9.80	3.60	4.08	11.84	36.80	38.70	4.91

selectively removed (Figs. 3, 4; Table 1). Tadpoles reared at low density in the laboratory had significantly greater SVL (Fig. 3A; $F_{4,75} = 49.19$, $P < 0.001$) and full body length (data not shown in a figure; $F_{4,75} = 58.98$, $P < 0.001$), and developed faster (Fig. 4A; $F_{4,125} = 102.74$, $P < 0.001$) than those reared at high density, regardless of the selection-pressure treatment. In the field, a significant difference in SVL of tadpoles from high- and low-density treatments was also observed (Fig. 3B; $F_{1,12} = 6.37$, $P = 0.045$), and the full body length of those tadpoles from high-density pond-halves was significantly smaller than those from the low-density halves (data not shown; $F_{1,12} = 9.31$, $P = 0.023$). However, there was no difference in developmental stage between tadpoles from high- and low-density pond-halves (Fig. 4B; $F_{1,8} = 2.23$, $P = 0.209$). In the laboratory, there was a significant interaction between time and density for SVL ($F_{12,75} = 4.60$, $P < 0.001$), full body length ($F_{12,75} = 4.84$, $P < 0.001$), and development ($F_{20,125} = 4.44$, $P < 0.001$). Conversely, there was no interaction between time and density for SVL ($F_{2,12} = 0.45$, $P = 0.648$), full body length ($F_{2,12} = 0.58$, $P = 0.575$), or Gosner stage ($F_{2,8} = 0.44$, $P = 0.658$) in the field. Overall variation in population growth (COV) was significantly different among groups (Fig. 5A; $F_{4,75} = 16.04$, $P < 0.001$) in the laboratory, but the selection of small tadpoles reduced COV to closely resemble that of the low-density treatment groups (Tukey HSD; $P = 0.925$) and the COV of field populations at both high and low density (Fig. 5B).

Does Incorporating Selective Mortality in Laboratory Populations Influence Statistical Power?—Power was high for all populations for effect sizes of 15% and 20%. Field populations, laboratory populations at low density, and laboratory populations at high density with selection against small individuals had very high power (>0.95) to detect even small differences in growth (10%). By contrast, laboratory populations at high densities with no selection, and with random selection, had relatively low power to detect differences of the same magnitude (Fig. 6).

DISCUSSION

Is Mortality Nonrandom in Natural Tadpole Populations?—Our data support the hypothesis that small tadpoles are selected against in natural populations (Smith, 1983; McCoy et al., 2011). Visualization of populations following the mortality of numerous individuals in the field identified positively skewed size distributions. A distribution mean greater than the median suggests that a higher number of small individuals were eliminated from the population compared to large individuals. As further evidence, when small individuals were selectively removed from laboratory populations, the resulting magnitude of response to increased density was similar to that observed in the field. Based on our results, it is not possible to explicitly attribute this size-selective mortality to any one cause but, based on high numbers of Odonates observed in the ponds, it is likely that predation is a major factor (Brockelman, 1969; Calef, 1973). Odonates are known to be gape-limited predators (Semlitsch, 1990), and this explains the extreme mortality early in the developmental period that tapers off with time and also the size-selective elimination of smaller individuals. Our preceding work further indicates the broad importance of predatory interactions for influencing survivorship in natural tadpole populations (Melvin and Houlihan, 2012).

The process of natural selection is a pillar of ecological and evolutionary research (Darwin, 1869; Fisher, 1930; Endler, 1986). Nevertheless, laboratory-based experimental protocols tend to overlook the importance of selective processes for shaping natural populations (Collins, 1986). For instance, many anuran amphibians are excellent examples of r-selected species with high fecundity, low parental investment, and high mortality suffered early in the developmental period (MacArthur, 1961, 1962; Reznick et al., 2002). Despite these known life-history characteristics, the effects of high, early-stage mortality are rarely considered in laboratory experimental designs. In fact, the results of laboratory experiments with developing tadpoles are generally not considered unless the mortality of control populations is less than 5–10% (Gutleb et al., 2007; Lutz et al.,

TABLE 2. Summary table of Skewness and Kurtosis for populations of Wood Frog, *Lithobates sylvaticus*, reared in the laboratory and natural ponds at high- (5 tadpoles/L) and low-density (1 tadpole/L) as shown in Figure 1.

Venue	Density	N	Skewness	SE	P-value	Kurtosis	SE	P-value
Laboratory	High	60	-0.069	0.309	0.412	-0.778	0.608	0.100
	Low	53	-0.033	0.327	0.379	0.096	0.644	0.441
Field	High	40	0.53	0.374	0.078	-1.046	0.733	0.077
	Low	25	1.231	0.464	0.004 ^a	1.05	0.902	0.122

^a Indicates positively skewed population ($P < 0.05$).

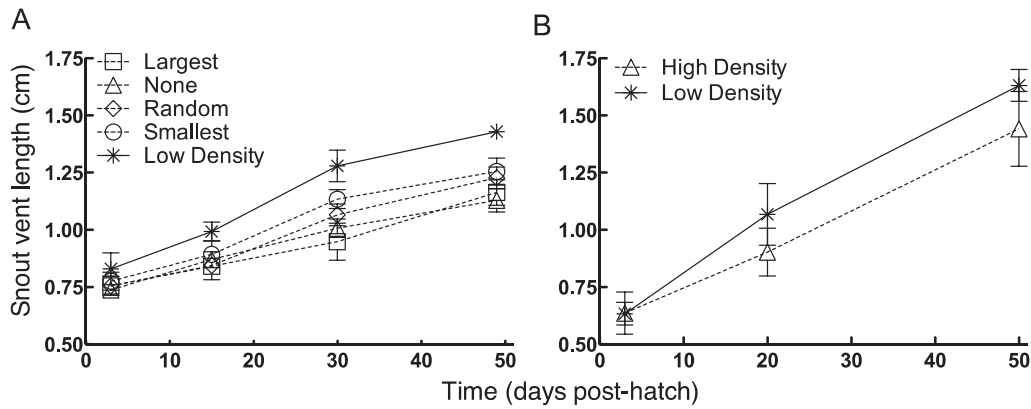


FIG. 3. Snout-vent length (SVL) of Wood Frog, *Lithobates sylvaticus*, tadpoles reared at low (solid line; 1 tadpole/L) and high (dashed lines; 5 tadpoles/L) stocking densities in (A) the laboratory experiencing no selection, random selection, or selection against the largest or smallest tadpoles, and (B) in natural ephemeral ponds. Bars represent 1 SD.

2008; OECD, 2009) because it is otherwise difficult to attribute responses to the treatment. Optimizing survival of control animals provides confidence that mortality experienced in treatment groups is truly the result of the treatment but, as our results illustrate, this increases the possibility that experimental populations are made up of individuals that would almost certainly be selected against in nature (Melvin and Houlihan, 2012).

Does Incorporating Selective Mortality in Laboratory Populations Produce Results that Are More Consistent with that of Natural Populations?—There have been several noteworthy exchanges amongst researchers who favor controlled experimental venues (Hairston, 1989; Wilbur, 1997) and those that assert that true ecological significance can only come from natural ecosystems (Kimball and Levin, 1985; Carpenter, 1996; Schindler, 1998; Skelly, 2002). Realistically, controlled laboratory and whole-ecosystem experiments offer different strengths and weaknesses and, if used properly, studies in both venues will continue to provide insight regarding natural ecological interactions (Jaeger and Walls, 1989; Thompson, 2004; Chalcraft et al., 2005). It is therefore important to question and understand the limitations imposed by different study designs and to work to improve experimental procedures so that they offer the best ecological significance while optimizing environmental control, statistical replication, and repeatability. The results of this study provide a

starting point for developing laboratory protocols that allow us to have increased confidence in inferences to natural systems. Our results demonstrate how introducing selection pressure against small individuals in the laboratory results in individual growth and development, and in population variation (COV), that are better representations of that observed in natural populations where high population mortality is a known occurrence (Melvin and Houlihan, 2012). It is not yet clear to what extent optimizing the representativeness of controlled laboratory experiments can influence results and interpretations of biological interactions. However, intense selection in natural systems is likely to change both the population and the context in which individuals from the population grow and develop, and this creates the potential for misleading inferences from laboratory studies because the population is made up of individuals that are characteristically different from the natural population. Introducing size selection into laboratory study designs can help reduce the potential for misinterpreting responses and, therefore, represents a positive step towards the improvement of laboratory study designs.

Does Incorporating Selective Mortality in Laboratory Populations Influence Statistical Power?—In addition to increasing the ecological relevance of laboratory experiments, incorporating selection into study designs offers the added benefit of increasing statistical power to detect small effect sizes because

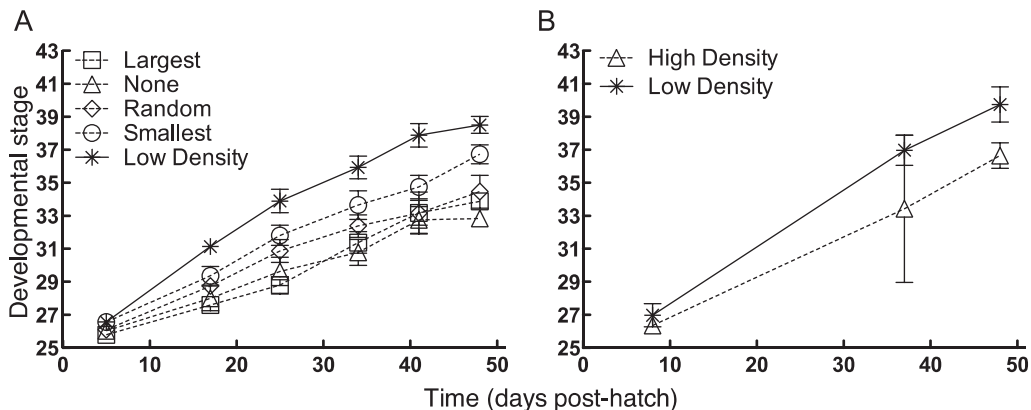


FIG. 4. Developmental stage (Gosner, 1960) of Wood Frog, *Lithobates sylvaticus*, tadpoles reared at low (solid line; 1 tadpole/L) and high (dashed lines; 5 tadpoles/L) stocking densities in (A) the laboratory experiencing no selection, random selection, or selection against the largest or smallest tadpoles, and (B) in natural ephemeral ponds. Bars represent 1 SD.

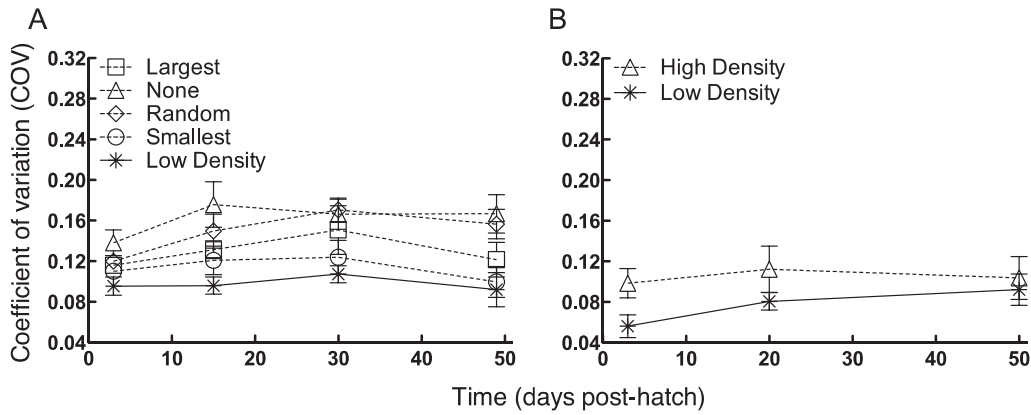


FIG. 5. Coefficient of variation (COV) of Wood Frog, *Lithobates sylvaticus*, tadpoles reared at low (solid line; 1 tadpole/L) and high (dashed lines; 5 tadpoles/L) stocking densities in (A) the laboratory experiencing no selection, random selection, or selection against the largest or smallest tadpoles, and (B) in natural ephemeral ponds. Bars represent 1 SD.

of the resulting reduction in population variation (COV). Power analysis of our data demonstrate that the homogenizing effect of selective mortality on natural tadpole populations causes them to have relatively high statistical power to detect small differences in growth and development (10%). Our results show that controlled laboratory populations at high densities in the absence of selection become very heterogeneous, which lowers their power in comparison to more-homogeneous natural populations. This is somewhat counterintuitive because, in our experience, significant differences in growth and development are regularly observed in the laboratory but less often in the field. However, this is best explained because

sample size and effect size, in addition to population variance, are important for determining power (Murphy et al., 2009), and each of these parameters are generally greater in laboratory experiments. Regardless, inappropriately high densities result in increased variability (higher than usually seen in nature) and increase the risk of making a Type II error (i.e., failure to identify significant differences when they exist).

Conclusion.—The need for caution when extrapolating from controlled populations to the natural world is well understood (Loman, 2001, 2004; Skelly and Kiesecker, 2001; Halverson et al., 2006), but the results of the present study demonstrate that the elimination of natural population mortality from laboratory

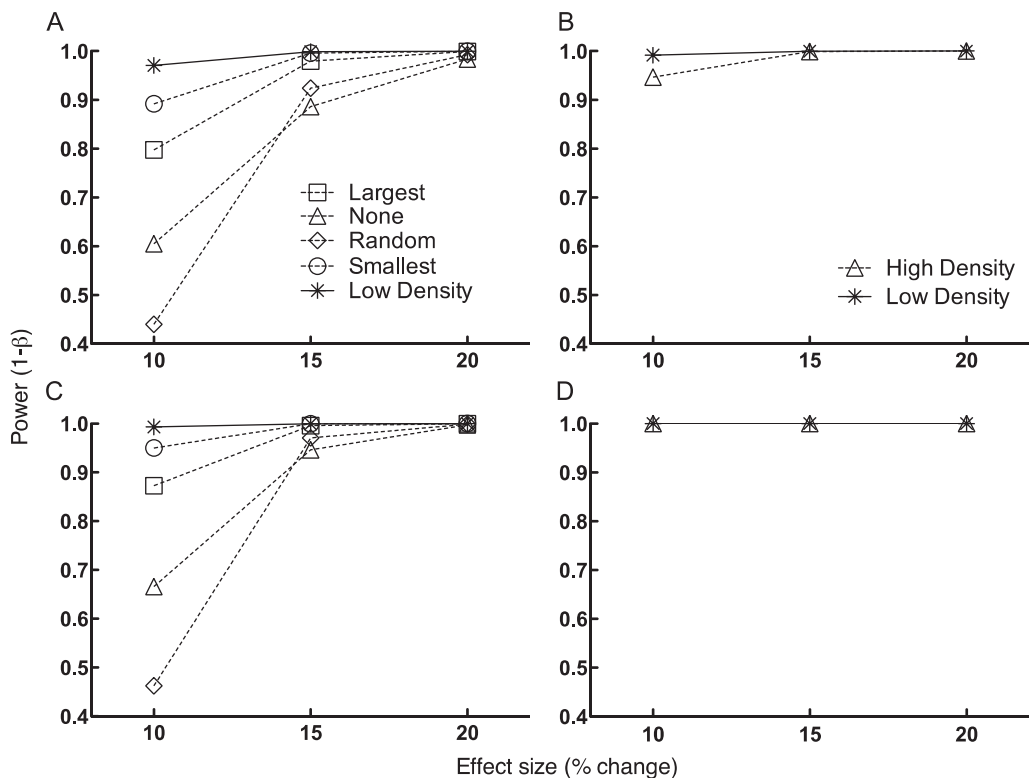


FIG. 6. Statistical power ($1 - \beta$) to detect differences of 10%, 15%, and 20% in tadpole growth for populations of Wood Frogs, *Lithobates sylvaticus*, reared at (A) low (solid line; 1 tadpole/L) and high (dashed lines; 5 tadpoles/L) stocking densities in the laboratory experiencing no selection, random selection, or selection against the largest or smallest tadpoles, and (B) in natural ephemeral ponds with sample size of $N = 4$. The same analysis was performed for (C) laboratory and (D) natural ephemeral ponds with sample size of $N = 6$. For both sample sizes, $\alpha = 0.05$.

experiments could be a major factor leading to inconsistency amongst experimental venues. Based on our findings, there are simple methods for increasing the representativeness of laboratory populations for predicting responses of populations of tadpoles in natural systems. We recommend maximum densities of roughly one tadpole/L be used in controlled laboratory populations or, alternatively, incorporating selection against small individuals when higher densities are required. High densities can influence tadpole growth and development strongly (Brockelman, 1969; Berven and Chadra, 1988; Altwegg, 2003; Belden et al., 2007) but have a negligible influence on survival under controlled conditions (Melvin and Houlihan, 2012). Thus, if laboratory experiments are carried out at such high densities that growth and development are influenced but selection is not incorporated into the study design, the results will reflect the responses of many individuals that would normally perish under natural conditions and therefore have limited ecological relevance. While we advocate the incorporation of artificial selection pressures into laboratory studies, and feel that these methods are generally quite feasible, we recognize that this may prove difficult for certain experimental designs or anuran species. Therefore, research investigating additional methodological improvements that could be developed to enhance the realism of laboratory experiments would be useful. The further refinement and validation of laboratory protocols will help bridge the gap between experimental control and environmental realism while continuing to provide the high level of replication and statistical power that have become the standard for ecological research under laboratory conditions.

Acknowledgments.—We thank all members of the Long-term Experimental Wetlands Area (LEWA) for assistance with experiments and discussions about the representativeness of experiments in artificial venues. Animal use was approved by the University of New Brunswick Saint John Animal Care Committee (Animal Care Protocols 2010-01-01 and 2010-01-02) under the Canadian Council of Animal Care (CCAC). This research was made possible through a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant to J. Houlihan.

LITERATURE CITED

- ALTWEGG, R. 2003. Multistage density dependence in an amphibian. *Oecologia* 136:46–50.
- BELDEN, L. K., M. J. RUBBO, J. C. WINGFIELD, AND J. M. KIESECKER. 2007. Searching for the physiological mechanism of density dependence: does corticosterone regulate tadpole responses to density? *Physiological and Biochemical Zoology* 80:444–451.
- BERVEN, K. A., AND B. G. CHADRA. 1988. The relationship among egg size, density and food level on larval development in the wood frog (*Rana sylvatica*). *Oecologia* 75:67–72.
- BRACKELMAN, W. 1969. An analysis of density effects and predation in *Bufo americanus* tadpoles. *Ecology* 50:632–644.
- CAIRNS, J. 1983. Are single species toxicity tests alone adequate for estimating environmental hazard? *Hydrobiologia* 100:47–57.
- CALEF, G. W. 1973. Natural mortality of tadpoles in a population of *Rana aurora*. *Ecology* 54:741–758.
- CARPENTER, S. R. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. *Ecology* 77:677–680.
- . 1999. Microcosm experiments have limited relevance for community and ecosystem ecology: reply. *Ecology* 80:1085–1088.
- CHALCRAFT, D. R., C. A. BINCKLEY, AND W. J. RESETARITS. 2005. Experimental venue and estimation of interaction strength: Comment. *Ecology* 86:1061–1067.
- COLLINS, J. P. 1986. Evolutionary ecology and the use of natural selection in ecological theory. *Journal of the History of Biology* 19:257–288.
- CORN, P. S., AND L. J. LIVO. 1989. Leopard frog and wood frog reproduction in Colorado and Wyoming. *Northwestern Naturalist* 70:1–9.
- DARWIN, C. 1869. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, 5th ed. John Murray, London, U.K.
- DUCELLMAN, W. E., AND L. TRUEB. 1994. *Biology of Amphibians*. The John Hopkins University Press, Baltimore.
- ENDLER, J. A. 1986. *Natural selection in the wild*. Princeton University Press, USA.
- FISHER, R. 1930. *The genetical theory of natural selection*. Clarendon Press, Oxford, U.K.
- FORMANOWICZ, D. R., JR. 1986. Anuran tadpole/aquatic insect predator-prey interactions: tadpole size and predator capture success. *Herpetologica* 42:367–373.
- GOSNER, K. L. 1960. A simplified table for staging anuran embryos and larvae with notes on identification. *Herpetologica* 16:183–190.
- GUTLEB, A. C., M. SCHRIKS, L. MOSSINK, J. H. J. VAN DEN BERG, AND A. J. MURK. 2007. A synchronized amphibian metamorphosis assay as an improved tool to detect thyroid hormone disturbance by endocrine disruptors and apolar sediment extracts. *Chemosphere* 70:93–100.
- HALVERSON, M. A., D. K. SKELLY, AND A. CACCONI. 2006. Inbreeding linked to amphibian survival in the wild but not in the laboratory. *Journal of Heredity* 97:499–507.
- HAIRSTON, N. G. 1989. Hard choices in ecological experimentation. *Herpetologica* 45:119–122.
- JAEGER, R. G., AND S. C. WALLS. 1989. Review: On salamander guilds and ecological methodology. *Herpetologica* 45:111–119.
- KIMBALL, K. D., AND S. A. LEVIN. 1985. Limitations of laboratory bioassays: the need for ecosystem-level testing. *BioScience* 02:165–171.
- LEVENE, H. 1960. Robust tests for equality of variances. In I. Olkin (ed.), *Contributions to Probability and Statistics*, pp. 278–292. Stanford University Press, USA.
- LOMAN, J. 2001. Intraspecific competition in tadpoles of *Rana arvalis*: does it matter in nature? A field experiment. *Population Ecology* 43:253–263.
- . 2004. Density regulation in tadpoles of *Rana temporaria*: a full pond field experiment. *Ecology* 85:1611–1618.
- LUTZ, I., W. KLOAS, T. A. SPRINGER, L. R. HOLDEN, J. C. WOLF, H. O. KRUEGER, AND A. J. HOSMER. 2008. Development, standardization and refinement of procedures for evaluating effects of endocrine active compounds on development and sexual differentiation of *Xenopus laevis*. *Analytical and Bioanalytical Chemistry* 390:2031–2048.
- MACARTHUR, R. H. 1961. Population effects of natural selection. *American Naturalist* 95:195–199.
- . 1962. Some generalized theorems of natural selection. *Genetics* 48:1893–1897.
- MAUCHLY, J. W. 1940. Significance test for sphericity of a normal n -variate distribution. *Annals of Mathematical Statistics* 11:204–209.
- MCCOY, M. W., B. M. BOLKER, K. M. WARKENTIN, AND J. R. VONESH. 2011. Predicting predation through prey ontogeny using size-dependent functional response models. *American Naturalist* 177:752–766.
- MELVIN, S. D., AND J. E. HOULAHAN. 2012. Tadpole mortality varies across experimental venues: do laboratory populations predict responses in nature? *Oecologia* 169:861–868.
- MORIN, P. J. 1989. New directions in amphibian community ecology. *Herpetologica* 45:124–128.
- . 2011. *Community Ecology*, 2nd ed. John Wiley & Sons, Ltd., U.K.
- MUNKITTRICK, K. R., C. J. ARENS, R. B. LOWELL, AND G. P. KAMINSKI. 2009. A review of potential methods of determining critical effect size for designing environmental monitoring programs. *Environmental Toxicology and Chemistry* 28:1361–1371.
- MURPHY, K. R., B. MYORS, AND A. WOLACH. 2009. *Statistical power analysis: a simple and general model for traditional and modern hypothesis tests*, 3rd ed. Taylor & Francis Group, USA.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). 2009. Amphibian metamorphosis assay. In *Guidelines for the testing of chemicals, section 2: Effects on biotic systems*, pp. 1–33. OECD Publishing, Paris, France.
- REZNICK, D., M. J. BRYANT, AND F. BASHEY. 2002. r- and K-selection revisited: the role of population regulation in life-history evolution. *Ecology* 83:1509–1520.

- REZNICK, D. N., AND C. K. GHALAMBOR. 2005. Selection in nature: experimental manipulations of natural populations. *Integrative and Comparative Biology* 45:456–462.
- SCHINDLER, D. W. 1987. Detecting ecosystem responses to anthropogenic stress. *Canadian Journal of Fisheries and Aquatic Sciences* 44:6–25.
- . 1998. Replication versus realism: the need for ecosystem-scale experimentation. *Ecosystems* 1:323–334.
- SEMLITSCH, R. D. 1990. Effects of body size, sibship, and tail injury on the susceptibility of tadpoles to dragonfly predation. *Canadian Journal of Zoology* 68:1027–1030.
- SKELLY, D. K. 2002. Experimental venue and estimation of interaction strength. *Ecology* 83:2097–2101.
- . 2005. Experimental venue and estimation of interaction strength: reply. *Ecology* 86:1068–1071.
- SKELLY, D. K., AND J. M. KIESECKER. 2001. Venue and outcome in ecological experiments: manipulations of larval anurans. *Oikos* 94:198–208.
- SMITH, D. 1983. Factors controlling tadpole populations of the Chorus frog (*Pseudacris triseriata*) on Isle Royal, Michigan. *Ecology* 64:501–510.
- THOMPSON, D. G. 2004. Potential effects of herbicides on native amphibians: a hierarchical approach to ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry* 23:813–814.
- WILBUR, H. M. 1997. Experimental ecology of food webs: complex systems in temporary ponds. *Ecology* 78:2279–2302.

Accepted: 17 June 2013.

Copyright of Journal of Herpetology is the property of Society for the Study of Amphibians & Reptiles and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.