Venue and outcome in ecological experiments: manipulations of larval anurans

David K. Skelly and Joseph M. Kiesecker, School of Forestry and Environmental Studies and Dept of Ecology and Evolutionary Biology, Yale Univ., 370 Prospect Street, New Haven, CT 06520, USA (david.skelly@yale.edu) (present address of JMK: Dept of Biology, Mueller Lab, Pennsylvania State Univ., University Park, PA 16802, USA).

The choice of venue is believed to be a critical decision for ecologists conducting experiments, yet the existence of tradeoffs in design and the impact of venue on outcome have not been quantitatively evaluated. To these ends we reviewed a set of 227 comparisons from 52 studies designed to quantify the effect of density on the growth performance of larval anurans. We predicted that both design and outcome would be related to experimental venue (laboratory, mesocosm, and field). We found that, in contrast to our predictions, laboratory experiments did not include more factors or yield more precise estimates of responses compared with mesocosm and field manipulations. In partial support of our prediction, we found that laboratory experiments did utilize somewhat larger numbers of replicates and included more levels per factor compared to the alternative venues. In addition we found that, as predicted, raising tadpoles at higher densities of conspecifics, or in the presence of heterospecifics tended to decrease growth. This effect, consistent with competition, was strongly dependent on experimental venue. In particular, we found that interspecific effects were much stronger in mesocosm versus field manipulations. This result is particularly relevant because conceptions of amphibian assemblages have included interspecific competition as integral to natural patterns of amphibian distribution. We conclude that hypothesized tradeoffs among experimental venues have not acted in a particularly strong way in shaping the design of tadpole density manipulations. However, venue may mediate strong effects on the outcome of such experiments. Quantification of such methodology based biases will help ecologists more effectively match the goals of their experiments with decisions regarding venue.

Over the last three decades ecology has become broadly experimental. As experimentation has become part of the culture of ecology, ecologists have contemplated strategies for choosing among the various options that confront all experimenters. Important examples include the selection of experimental venue, the choice of experimental design, and the degree of replication. At the same time, the wider use of experiments has led to consideration of the role of experimentation within ecological research programs (Peters 1991, Werner 1998, Wootton and Pfister 1998).

There is a broad perception among ecologists that the choice of experimental venue is a critical decision that can constrain other attributes of an experiment. In particular, a number of ecologists have suggested that field experiments tend to be limited to less complicated designs and lower levels of replication (Mertz and McCauley 1980, Hairston 1989a, Jaeger and Walls 1989, Peters 1991, Morin 1998). In addition, it has been hypothesized that the field environment is more variable leading to less precision among replicates (Hairston 1989a, b, Morin 1998). Finally, exposure to the natural environment is thought to limit the ability of an experimenter to control conditions experienced by the subjects of the manipulation (Mertz and McCauley 1980, Morin 1989, 1998, Peters 1991). These perceived tradeoffs, coincident with rising standards for statistical rigor, have motivated ecologists to experiment within laboratories (e.g. Mertz and McCauley 1980).

Even as the benefits of laboratory manipulation have been repeatedly asserted, critics also have been nearly unanimous in contending that field experiments are more realistic and that only field experiments can tell us about the quantitative importance of factors in natural populations (e.g. Carpenter 1996). Perhaps for this reason, major reviews of ecological experiments often have been limited to field experiments (e.g. Connell 1983, Schoener 1983, Goldberg and Barton 1992, Gurevitch et al. 1992, 2000). In response to their perceptions of these tradeoffs in experimentation, some ecologists turned to a compromise solution in the form of hybrid experiments conducted in mesocosms (Wilbur 1989, 1997, Morin 1998). For experimenters employing this hybrid approach, a major goal has been the develop-

ment of experimental techniques that capture attributes of the natural environment without suffering as acutely from the attendant limitations of field experimentation.

The backdrop for decisions about experimental venues are the goals which motivate ecologists to conduct experiments in the first place (Wootton and Pfister 1998). Thus, experiments conducted to test ecological theory often take place in laboratory settings. It has been reasoned that it is simply too difficult to deal with the challenges of replicating sufficiently against a noisy background to conduct such investigations in the field (e.g. Morin 1998). By contrast, mensurative experiments, designed to quantify the strength of an interaction, often have been conducted in the field (e.g. Paine 1992). In this case, it is believed that the field setting will yield results which are more applicable to understanding the conditions faced by natural populations. In all cases, the decisions made by ecologists appear to reflect choices based on personal assessments of the costs and benefits of conducting an experiment in different venues. However, to date, those reckonings have been made largely in the absence of formal analyses. Neither the degree to which venue mediated tradeoffs affect experimental design, nor the size of venue mediated effects on experimental outcome have been quantified previously. Thus, we may not know how large the advantages of one venue over another really are. In the absence of quantitative data it is possible that prevailing perceptions may distort the advantages and disadvantages of different venues.

In order to quantify venue mediated tradeoffs in design and effects on outcome, we review designs and results from a set of published experiments. To meet these objectives, we have focused on a single taxon, anuran larvae. Anuran larvae have been used extensively in experiments designed to elucidate principles concerning the structure of ecological communities generally (e.g. Morin 1983), and of the role of competition in particular (e.g. Morin and Johnson 1988). Anuran larvae also are closely associated with the development of techniques to exploit the use of mesocosms in order to mitigate venue mediated tradeoffs (e.g. Wilbur 1997). For these reasons, we believe that larval anurans are a particularly appropriate taxon in which to investigate the effects of venue. In addition, a sharp taxonomic focus provides the greatest chance of isolating effects attributable to venue, rather than those attributable to the mechanical distinctions of working with different types of organisms.

Our choice confers a number of additional advantages as well. Ecological experiments on larval anurans began in the 1960s (Dumas 1964) and their development mirrors that of ecology as a whole. Studies are numerous and, owing to the relative tractability of larval amphibians, these experiments represent a wide array of designs and venues. In addition, we have targeted experiments that include manipulations of density. The impact of competition on populations has been a focus of interest during the entire history of experimental ecology (Kingsland 1995). As a result of this research, competition for food resources, particularly between species, continues to be considered an important factor in regulating the abundance and distribution of species (e.g. Rowe and Dunson 1994, Wilbur 1997, Resetarits and Fauth 1998, Alford 1999, Fauth 1999, Petranka and Kennedy 1999).

Finally, a focus on anuran experiments allows us to construct three nonoverlapping venue categories each represented by a substantial number of studies: laboratory, mesocosm, and field. Laboratory experiments include those manipulations conducted indoors in small containers. Mesocosm experiments have been conducted outdoors in replicated, artificial containers. Most commonly, mesocosms used for anuran experiments have been constructed using cattle watering tanks. Field experiments have been dominated by manipulations conducted in mesh enclosures but also include manipulations of entire wetlands (e.g. Smith 1983). Note that while the definitions described here are consistent with usage by amphibian ecologists (e.g. Wilbur 1997, Morin 1998), it is obvious that these terms have different meanings for different ecologists (e.g. mesocosm: Odum 1984, Petersen et al. 1999; field: Gurevitch et al. 1992, 2000).

In this study, we evaluate three predictions regarding attributes of experiments based on widely held perceptions. Relative to field experiments, we expect (1) the design of mesocosm and laboratory experiments to be more complex, (2) replication within mesocosm and laboratory experiments to be higher, and (3) precision within mesocosm and laboratory experiments to be higher. Within each study reviewed, we measured complexity as the number of experimental factors included and the number of levels per factor, replication as the number of replicates of the control treatment (defined below), and precision as the % coefficient of variation (% CV) of the control treatment.

We also evaluated two predictions regarding the outcomes of tadpole density experiments conducted in laboratory, mesocosm, and field venues. These were (1) that tadpoles raised at higher densities of conspecifics, or in the presence of additional heterospecifics would grow more slowly than tadpoles raised at lower density or without heterospecifics, and (2) that the measured impact of density manipulations on growth would be independent of experimental venue. These latter two predictions were evaluated using meta-analyses of ln response ratios of growth responses.

Methods

Literature search

We compiled a preliminary list of studies by searching a set of computerized databases (Aquatic Biology, Agriculture & Fisheries Resources; Biological & Agricultural Index; BIOSIS; Science Citation Index). In each database, we used the search string "(tadpole? or anuran larvae? or larval anuran?) and (competition or density)". The reference lists of all papers identified during the database searches also were searched and potentially relevant papers were also included in the preliminary list. Our sample includes papers published through 1999.

We pared our list to include only those studies that included one or more comparisons meeting five criteria (Appendix): (1) the experimenter directly manipulated densities of at least one larval anuran; (2) a growth related response (growth rate, final size) was measured on a given species for the remainder of the larval period or until the end of a predetermined time period, (3) the published growth response included the mean, some estimate of variance (variance, standard deviation, standard error), and sample sizes for each treatment, (4) the treatments were conducted simultaneously and in the same location; and (5) relevant pairs of density treatments were not confounded by variation in other manipulated factors (e.g. food density, predators, additional competitors, hydroperiod manipulations).

Each study was characterized as including intraspecific manipulations, interspecific manipulations, or both. Experimental venue was designated as laboratory, mesocosm, or field based on the definitions presented in the Introduction.

Attributes of experiments

For each of the 54 experiments (two studies contained experiments in more than one venue) we recorded the number of experimental factors, the number of levels per factor, and the number of replicates used. When the number of replicates varied within the experiment, we used the number from the control treatment. In addition, we calculated the % CV of growth responses of the "control" treatment (designated as the lowest density treatment in intraspecific experiments and the single species treatment in interspecific comparisons). These attributes were analyzed individually using one-way ANOVAs to determine whether they varied among experimental venues.

Meta-analysis of growth responses

We performed a meta-analysis on growth responses of tadpoles subjected to different density treatments.

We have focused on growth responses because growth is a sensitive indicator of performance and a strong predictor of metamorphic traits among anurans (Werner 1986). These metamorphic traits, in turn, have been related to survivorship and reproductive success (e.g. Smith 1987). In addition, growth responses have been reported by virtually all authors of experiments focused on larval anurans, while other measures (e.g. survival, development rate) are reported less frequently.

Following the standard convention for meta-analyses we have focused on comparisons of control and experimental treatments from each study. We designated the lowest density treatment to be the control in intraspecific experiments and the single species treatment to be the control in interspecific comparisons. In both cases, if the experimental treatment mean is lower than the control treatment mean, the trend is consistent with a competitive effect. Most studies included more than one comparison. In some cases a single control treatment (the one with the lowest density) was compared with more than one experimental treatment (following convention of Gurevitch et al. 1992). Experimental treatments were never used more than once.

We gleaned the necessary data from the text and tables of each paper where possible. In some cases, growth responses were presented only in figures. In these instances, we estimated means and errors by digitizing the image and comparing the length of bars (or position of symbols) relative to the scale on the response axis. Stems estimating error were quantified similarly. We verified the effectiveness of this technique by digitizing the relevant figures from our own papers (e.g. Skelly 1995b) and comparing the resulting estimates with the original data.

In meta-analysis the standard deviation is used as the estimate of error about treatment and control means. In cases where authors reported the standard error (SE) instead, we calculated the standard deviation as $((\sqrt{n})\text{SE})$ where n is the sample size for the treatment. In a few cases where no estimate of error was reported, it was possible to estimate the standard deviation by calculating the square root of the within group mean square as reported in an ANOVA table (Gurevitch et al. 1992).

Data from each study were analyzed using MetaWin (Rosenberg et al. 1997) and effect sizes were calculated and analyzed as ln response ratios (Hedges et al. 1999). In all of the comparisons used in this meta-analysis, a positive effect size indicates that growth tended to be reduced at higher density or in the presence of the heterospecific (following the convention of Gurevitch et al. 1992). Comparisons among subgroups were evaluated using the homogeneity statistic, Q_b (Hedges and Olkin 1985).

Results

Our review included data collected from a total of 52 studies incorporating 137 intraspecific and 90 interspecific comparisons (Table 1, Appendix). Studies were conducted on four continents (Australia, Europe, North America, South America) and included anurans from six families and 35 species. Within the sample, manipulations in the laboratory were most numerous (85 comparisons), followed by those conducted in mesocosms (79), and in the field (63). While all experiments shared manipulations of density, additional factors focused on biotic manipulations of predators and pathogens, and abiotic manipulations such as pH and hydroperiod.

We made a set of predictions based on expected tradeoffs in the attributes of experiments conducted in different experimental venues. Contrary to the first prediction we found no evidence that, relative to laboratory and mesocosm experiments, field experiments included fewer experimental factors (Fig. 1a, one-way ANOVA: df = 2.51, $F = 1.61$, $P = 0.21$). In partial support of this prediction, we did find that the number of levels per factor was related to venue (Fig. 1b; one-way ANOVA: df = 2.51, $F = 3.56$, $P < 0.04$). Laboratory experiments tended to include more levels per factor than alternative venues (Tukey HSD test: $P < 0.05$); there was no evidence that mesocosm and field venues differed in the number of levels per factor (Tukey HSD test: $P > 0.05$).

The second prediction focused on replication. In this case we found some support for a relationship between replication and experimental venue (Fig. 2; one-way ANOVA: df = 2.51, $F = 4.97$, $P < 0.01$). Laboratory experiments tended to incorporate a higher degree of replication compared with mesocosm or field experiments (Tukey HSD test: $P < 0.05$). There was no evidence that mesocosm and field experiments differed in their degree of replication (Tukey HSD test: $P > 0.05$).

We also predicted that laboratory and mesocosm experiments would be more precise than field experiments. However, % CV did not vary among venues (Fig. 3; one-way ANOVA: df = 2.149, $F = 1.02$, $P = 0.36$).

We made a second set of predictions focused on the outcome of experiments conducted in different experimental venues. The first of these predictions was that, irrespective of venue, tadpoles of a given species reared
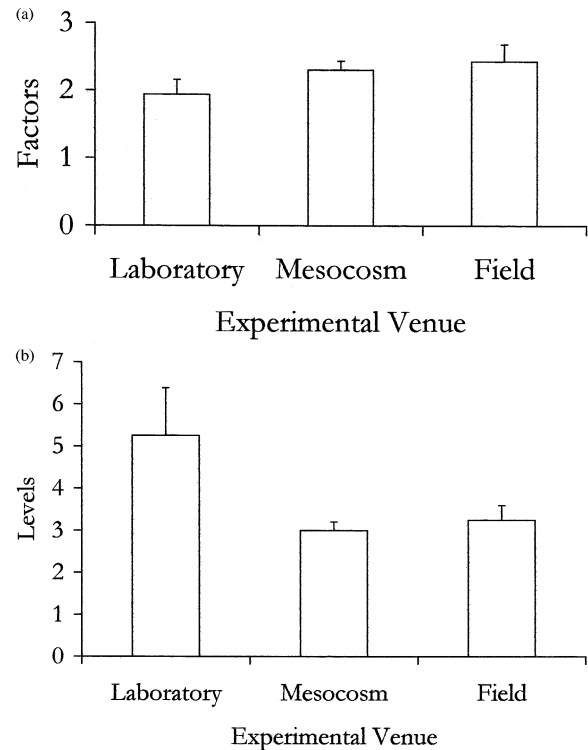

Fig. 1. Design complexity represented as (a) the number of experimental factors employed and (b) the number of levels per factor in 54 experiments categorized as occurring in laboratory, mesocosm, or field venues. Error bars represent 1 SE.

Table 1. Attributes of comparisons included in meta-analysis of competition experiments focusing on larval anurans. Each comparison has been characterized as an additive manipulation of conspecifics (intraspecific) or heterospecifics (interspecific). In addition, manipulations were characterized by the experimental venue: laboratory, mesocosm, or field (see Methods for details). The total number of comparisons considered is included in parentheses following the venue description. Experimental unit volume (liters) is the volume of a single container or enclosure used in the manipulation. Stocking density is the total density (individuals/liter) of tadpoles in the experimental treatment (conspecifics or heterospecifics added). Density factor is the ratio of the experimental density and the control density. An addition of an equal density of heterospecifics or conspecifics would result in a density factor of two. For each measure the mean ± 1 SE is presented.

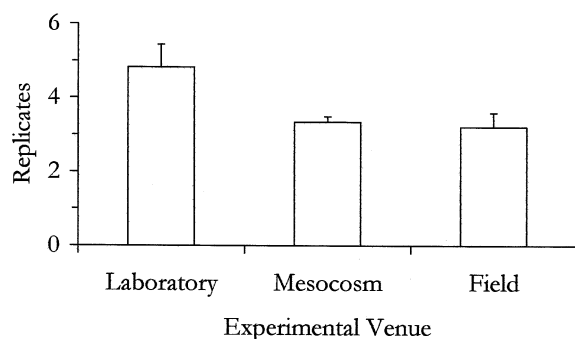| Manipulation/Venue	Experimental unit volume	Stocking density	Density factor
Intraspecific			
Laboratory (73)	5 ± 1	43.8 ± 19.8	7.0 ± 0.9
Mesocosm (34)	712 ± 66	0.3 ± 0.1	2.7 ± 0.2
Field (30)	346 ± 35	1.4 ± 0.3	4.6 ± 0.7
Interspecific			
Laboratory (12)	5 ± 1	5.4 ± 1.6	2.2 ± 0.1
Mesocosm (45)	637 ± 57	0.7 ± 0.1	3.8 ± 0.8
Field (33)	577 ± 39	1.4 ± 0.5	2.0 ± 0.1

Fig. 2. Number of replicates of the control treatment for each of 54 experiments categorized as occurring in laboratory, mesocosm, or field venues. Error bars represent 1 SE.

at higher density or in the presence of a heterospecific, grow more slowly. This prediction was supported for both intraspecific ($n = 137$, ln Response ratio $= 0.11$, $P < 0.05$) and interspecific ($n = 90$, ln Response ratio $= 0.06$, $P < 0.05$) manipulations. Overall, the growth reduction was larger for tadpoles subjected to intraspecific manipulations. While most individual comparisons were consistent with the overall effects described, a total of 33 (15%) showed the opposite trend.

We also predicted that experimental outcome would be independent of venue. This prediction was rejected for both intraspecific (Fig. 4; $Q_b = 108.8$, df $= 2$, $P < 0.001$), and interspecific comparisons (Fig. 4; $Q_b = 1504$, df $= 2$, $P < 0.001$). In both cases, effect sizes tended to be highest among mesocosm manipulations, intermediate among field manipulations, and lowest among laboratory manipulations.

Discussion

As experimentation has become ingrained in their discipline, ecologists have begun to more closely scrutinize
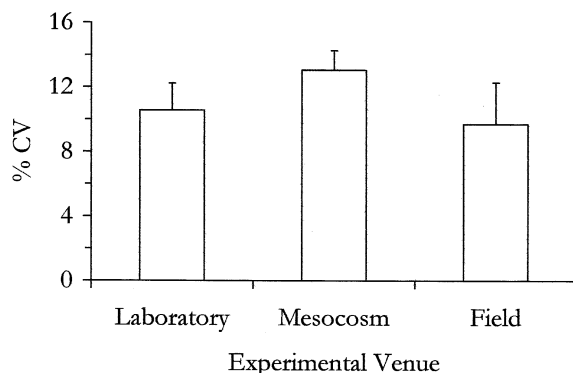

Fig. 3. Per cent coefficient of variation (% CV) of growth responses among control replicates. Responses were collected from 54 experiments categorized as occurring in laboratory, mesocosm, or field venues. Error bars represent 1 SE.
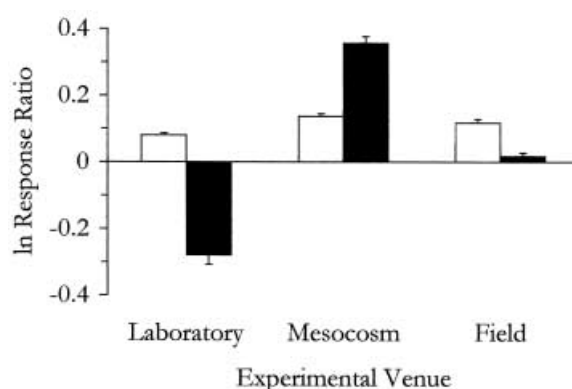

Fig. 4. Cumulative effect sizes (ln response ratios) in tadpole competition experiments. Experiments were categorized as intraspecific (open bars) or interspecific (filled bars) manipulations, and by experimental venue: laboratory, mesocosm, or field. Positive effect sizes indicate that tadpole growth was reduced when density of conspecifics was increased or when heterospecifics were added.

the way decisions are made regarding the design of experiments and of the role experiments play in producing inferences about the natural world (e.g. Morin 1998, Werner 1998). Broadly speaking, ecologists have come to believe that choices made about one aspect of an experimental design may constrain other aspects of the design, and that such choices could influence the outcome of the experiment and the inferences drawn. Of all the decisions made by an experimenter, the choice of experimental venue must rank as one of the most critical in this regard. In this study we set out to quantify patterns in the design and outcome of experiments flowing from the choice of venue.

Our first goal was to evaluate three widely professed beliefs regarding tradeoffs encountered when designing ecological experiments. Much to our surprise, we found little support among the experiments we reviewed. Relative to their mesocosm and field counterparts, laboratory experiments utilized more levels per factor and more replicates, but were equivalent in the number of factors evaluated and in the precision of responses. There was no evidence of venue mediated differences in design between mesocosm and field experiments.

The supposition that more complex experimental designs can be accomplished in the laboratory or in mesocosms is based on the idea that it is difficult to add additional factors or levels, and therefore additional experimental units, to experiments conducted in the field. However, most of the experiments we reviewed included one, two or three factors, and there was no evidence that field experiments contained fewer factors within this range. As has been pointed out, factorial designs explode into large numbers of experimental units when extended beyond three factors (Wilbur 1997). Perhaps the difficulty of conducting (and interpreting) any experiments of this size has moderated the imposition of a complexity based tradeoff.

One of the primary rationales for using mesocosms and laboratory containers has been the relative ease of replication. In practice, it appears that these advantages have been expressed moderately at best. There is no evidence from the structure of the experiments we reviewed that mesocosms are easier to replicate. And while the typical laboratory experimental unit is more than order of magnitude smaller than the typical mesocosm or field enclosure, the relative ease of setting up extra laboratory containers has yielded just a small increase in replicates per treatment. These patterns may imply that the two to five replicates used in most experiments regardless of venue have supplied statistical power adequate to researchers' goals. An alternative explanation is that the degree of replication may be determined by other considerations, (such as limits on the number of mesocosms, shelf space, or availability of field sites), or even rules of thumb.

Among our most puzzling findings is the lack of a gradient in precision. We fully expected to see more "noise" in responses emanating from field versus mesocosm and laboratory experiments. These results deserve additional investigation. Laboratory experiments on tadpoles are done in small, identical containers usually in environments that impose some control over light and thermal regimes and where food is carefully metered out. The fact that responses from this environment are as variable as those from mesocosm and field environments suggests one of three likely mechanisms. First, it may be that the larger number of individuals that comprise experimental populations in mesocosm and field experiments foster an averaging effect on responses not found among smaller numbers of individuals in laboratory containers (see discussion by Petersen et al. 1999). Alternatively, the field environment may be more variable, but important environmental attributes may not vary in a synchronized fashion. Thus, a warmer location fostering more rapid growth also may have a lower dissolved oxygen or a lower abundance of food resources resulting in little net variation in growth among replicates. Finally, experimentalists may be incorrect. Among the environmental attributes that most directly impact the responses we analyzed, the environments provided by alternate venues may not vary the way we pre-supposed. Whatever their origin, our findings suggest that precision has not been a particular problem in field experiments on larval anurans and laboratory and mesocosm environments offer no measurable advantage in this regard.

Overall, our results suggest that the tradeoffs hypothesized to constrain researchers' abilities to conduct complex, powerful, and precise experiments have not acted in an especially strong way to structure experimental designs. Phrased another way, there is little evidence from our analyses that there are strong advantages or disadvantages to using one venue over another based on considerations of design complexity, replica-

tion, and precision. While there is no guarantee that experiments conducted on other taxa or other interactions will produce the same message, our findings strongly suggest that these widely held beliefs should be further evaluated. The advantages of such scrutiny are several. Most notably, if venue does not constrain experimental designs as much as is typically thought, then ecologists can choose experimental venues based on other considerations.

Of all possible considerations, one of the most critical to many ecologists is the notion of realism. This concept is frustratingly difficult to define, but undeniably important to much that ecologists do. When measuring the strength of an interaction, and even when testing ecological theory, many ecologists would like to believe that their experimental results offer a realistic picture of the factors that affect natural populations and communities. Unfortunately, realism is one of the most difficult experimental attributes to calibrate. An ecologist performing a manipulation unavoidably decreases realism in hard to define ways. While it is often implied that experiments performed in the field are more realistic than those performed in mesocosms or laboratory mechanisms, at base, this can only be called an assumption.

A number of ecologists have dealt with this thorny issue by adopting an integrative approach. By rooting their research in a conceptual framework and conducting experiments in different venues, they seek to evaluate hypotheses through consensus in outcomes that can extend across venues (e.g. Sarnelle 1997, Werner 1998). Consequently, one metric of the importance of venue is the degree to which experiments conducted in different venues achieve congruence in outcomes. If experimental outcome is largely independent of venue, then there is some reason to think that resulting conclusions are more likely to reflect a process or pattern of interest rather than the methodological artifacts of our manipulations.

Thus, the second goal of our review was to evaluate the outcome of tadpole density manipulations, particularly with respect to the experimental venue employed. As in many model systems, amphibian ecologists have frequently measured the impact of density using experiments. In this system, competition, particularly between species, has been a cornerstone of hypotheses of community structure in amphibian assemblages (Morin 1983, Wilbur 1997, Alford 1999).

The results of our review demonstrate that manipulations of tadpole densities have regularly yielded effects consistent with competition. Overall, larval anurans raised at higher densities of conspecifics, and those raised in the presence of additional heterospecifics, experienced reduced growth. However, it is equally clear that the measured impact of density is highly heterogeneous. As we discuss below, our results suggest that decisions made regarding how to conduct a density

manipulation may have sizable impacts on estimates of competitive effect.

Specifically, we found large differences in estimated effect size among the three experimental venues. Mesocosm experiments were associated with larger declines in growth than either laboratory or field experiments. As seen in Fig. 4, estimates of the impact of interspecific manipulations are highly variable. Why would interspecific effect sizes diverge so widely? A negative estimate (facilitation) for laboratory manipulations is derived from just two studies, both focusing on the same two species (Smith-Gill and Gill 1978, Werner 1992). As in most laboratory studies of tadpole competition (and unlike the mesocosm and field studies reviewed), these researchers provided tadpoles with regularly replenished, per-capita food rations and in both cases it appeared that tadpoles sometimes benefited from increased tadpole density, perhaps because it was associated with a greater absolute amount of food. In addition, laboratory experiments have been conducted in containers averaging just a few liters while both mesocosm and field experiments have been conducted in experimental units of several hundred liters. The impact of scaling, in the absence of other factors, is capable of affecting responses measured by ecologists (Petersen et al. 1997, 1999). These or other differences between laboratory and field experiments could be responsible for an observed switch in competitive dominance between experiments conducted on leopard frogs (*Rana pipiens*) and wood frogs (*R. sylvatica*) (Werner 1998).

While mesocosm and field venues yielded significant effect sizes consistent with competition, the effect size from mesocosm studies was severalfold that for field experiments. In fact, interspecific field manipulations were associated with extremely small effects, equivalent to a 2% change in growth. The outcome of straightforward density manipulations is highly context dependent prompting two important questions: What factors contribute to differences in outcome between mesocosm and field experiments? and, How do results from these experiments relate to the conditions experienced by natural populations?

One potential explanation for the discrepancy in outcomes between mesocosm and field experiments could be differences in stocking densities (Jaeger and Walls 1989). While mesocosm studies have been criticized as having unrealistically high densities, we found that stocking densities within mesocosm experiments were similar to and tending to be even lower than those found in field experiments (Table 1). Experiments conducted in mesocosms and in the field were comparable in other respects as well. The volume of experimental units, the duration of experiments, and the density factor of the manipulations were each broadly similar (Table 1, Appendix).

Given that mesocosms were first implemented as a more tractable, but quasi-natural alternative to field manipulations (Resetarits and Fauth 1998), perhaps it is not surprising that researchers have been successful at mimicking a number of aspects of field experiments. These similarities make it that much more difficult to understand the sharp distinction in experimental outcomes between venues. Nevertheless, there are at least two possibilities that bear further inquiry. First, it is clear that experimenters working in the two venues have tended to focus on different species (Appendix). Perhaps some species which have been examined solely within mesocosms will also compete strongly when evaluated in the field. Second, we know of no studies comparing the composition and abundance of food resources within mesocosms and natural ponds (or field enclosures). If mesocosms contain a lower diversity of food resources, or if resources become depleted during an experiment more rapidly, the measured impact of density on growth could be increased because of heightened resource overlap. Whatever their origin, the existence of a characteristic difference in outcome between mesocosm and field experiments has important implications for amphibian ecologists.

Of even larger relevance is our ability to use any of the experiments to make inferences about conditions experienced by natural populations. Currently, such inferences about anuran communities often appear to be based on the predominant pattern emanating from mesocosm experiments (e.g. Wilbur 1997, Resetarits and Fauth 1998, Petranka and Kennedy 1999). However, if field experiments are assumed to be the most appropriate gauge of the natural condition (e.g. Connell 1983, Schoener 1983, Goldberg and Barton 1992, Gurevitch et al. 1992, Carpenter 1996), then a much different picture emerges. Interspecific competition may occur in natural populations, but it may often be of moderate effect. Patterns, such as segregated distributions, previously attributed to interspecific competition may be due to other factors (Wellborn et al. 1996).

To what extent is the assumption of greater realism valid for the field experiments reviewed here? A particularly relevant concern is that the mechanics of field manipulations may, for whatever reason, tend to mitigate the impacts of increased density. One common criticism is that the mesh walls of enclosures may allow nutrients and food resources to enter thus preventing depletion. There are two counterarguments to this suggestion. First, there are usually tadpoles and other grazers living outside of enclosure walls so it is not clear, a priori, whether the net flow of food resources and nutrients will be into or out of enclosures. Second, single species field experiments conducted under identical conditions have tended to yield sizable impacts of intraspecific density manipulations (Fig. 4). This difference in outcomes might be expected from a theoretical perspective, and demonstrates that field enclosures can

subject larvae to the impacts associated with increased density.

As Morin (1998) has suggested, the continued interest of ecologists in the impact of experimental venue may stem partly from a lack of understanding of the advantages and disadvantages of the available alternatives. However, the picture emerging from this first quantitative analysis of venue mediated tradeoffs does not match the image of inevitably conflicted goals of complexity, precision, and realism typically presented (Hairston 1989a, b, Peters 1991, Morin 1998). This does not mean that one venue should hold primacy over others; however, it does suggest that the role of venue should receive closer scrutiny. The goals of some experiments make decisions regarding experimental venue relatively straightforward. Logistical considerations will necessitate working in the laboratory in some cases or in the field in others. However, in many other cases experimental goals will not, by themselves, make the choice of venue immediately obvious. In those cases, should ecologists choose the most expedient option? Or should they more closely consider the more cryptic costs and benefits of working in one venue over another?

At the very least, the results of our analysis suggest that venue mediated impacts on design and outcome will be worth investigating in other systems. While results on amphibian larvae may not apply directly to other taxa, the findings presented here suggest that we should not be surprised if results from additional analyses do not conform to long held conventions regarding ecological experimentation. Regardless of their outcomes, such analyses will help ecologists make more informed decisions when designing experiments.

References

Alford, R. A. 1999. Ecology: resource use, competition, and predation. – In: McDiarmid, R. W. and Altig, R. (eds), Tadpoles: the biology of Anuran larvae. Chicago Univ. Press, pp. 240–278.

Alford, R. A. and Wilbur, H. M. 1985. Priority effects in experimental pond communities: competition between *Bufo* and *Rana*. – Ecology 66: 1097–1105.

Berven, K. A. and Gill, D. E. 1983. Interpreting variation in life-history traits. – Am. Zool. 23: 85–97.

Blaustein, L. and Margalit, J. 1996. Priority effects in temporary pools: nature and outcome of mosquito larvae-toad tadpole interactions depend on order of entrance. – J. Anim. Ecol. 65: 77–84.

Brockelman, W. Y. 1969. An analysis of density effects and predation in *Bufo americanus* tadpoles. – Ecology 50: 632–644.

Brönmark, C., Rundle, S. D. and Erlandsson, A. 1991. Interactions between freshwater snails and tadpoles: competition and facilitation. – Oecologia 87: 8–18.

Carpenter, S. R. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. – Ecology 77: 677–680.

Connell, J. H. 1983. On the prevalence and importance of interspecific competition: evidence from field experiments. – Am. Nat. 122: 661–696.

Cummins, C. P. 1989. Interaction between the effects of pH and density on growth and development in *Rana temporaria* L. tadpoles. – Funct. Ecol. 3: 45–52.

Dash, M. C. and Hota, A. K. 1980. Density effects on the survival, growth rate, and metamorphosis of *Rana tigrina* tadpoles. – Ecology 61: 1025–1028.

DeBenedictis, P. A. 1974. Interspecific competition between tadpoles of *Rana pipiens* and *Rana sylvatica*: an experimental field study. – Ecol. Monogr. 44: 129–151.

Dumas, P. C. 1964. Species pair allopatry in the genera *Rana* and *Phrynosoma*. – Ecology 45: 178–180.

Fauth, J. E. 1999. Identifying potential keystones from field data – an example from temporary ponds. – Ecol. Lett. 2: 36–43.

Gascon, C. 1992. The effects of reproductive phenology on larval performance traits in a three-species assemblage of central Amazonian tadpoles. – Oikos 65: 307–313.

Gascon, C. and Travis, J. 1992. Does the spatial scale of experimentation matter? A test with tadpoles and dragonflies. – Ecology 73: 2237–2243.

Goldberg, D. E. and Barton, A. M. 1992. Patterns and consequences of interspecific competition in natural communities: a review of field experiments with plants. – Am. Nat. 139: 771–801.

Griffiths, R. A. 1991. Competition between the common frog, *Rana temporaria*, and natterjack toad, *Bufo calamita*, tadpoles: the effect of competitor density and interaction level on tadpole development. – Oikos 61: 187–196.

Griffiths, R. A., Edgar, P. W. and Wong, A. L.-C. 1991. Interspecific competition in tadpoles: growth inhibition and growth retrieval in natterjack toads, *Bufo calamita*. – J. Anim. Ecol. 60: 1065–1076.

Griffiths, R. A., Denton, J. and Wong, A. L.-C. 1993. The effect of food level on competition in tadpoles: interference mediated by prototechan algae? – J. Anim. Ecol. 62: 274–279.

Gurevitch, J., Murrow, L. L., Wallace, A. and Walsh, J. S. 1992. A meta-analysis of competition in field experiments. – Am. Nat. 140: 539–572.

Gurevitch, J., Morrison, J. A. and Hedges, L. V. 2000. The interaction between competition and predation: a meta-analysis of field experiments. – Am. Nat. 155: 435–453.

Hairston, N. G. 1989a. Ecological experiments: purpose, design, and execution. – Cambridge Univ. Press.

Hairston, N. 1989b. Hard choices in ecological experimentation. – Herpetologica 45: 119–122.

Hedges, L. V. and Olkin, I. 1985. Statistical methods for meta-analysis. – Academic Press.

Hedges, L. V., Gurevitch, J. and Curtis, P. S. 1999. The meta-analysis of response ratios in experimental ecology. – Ecology 80: 1150–1156.

Hokit, D. G. and Blaustein, A. R. 1994. The effects of kinship on growth and development in tadpoles of *Rana cascadae*. – Evolution 48: 1383–1388.

Jaeger, R. G. and Walls, S. C. 1989. On salamander guilds and ecological methodology. – Herpetologica 45: 111–119.

Kiesecker, J. M. and Blaustein, A. R. 1999. Pathogen reverses competition between larval amphibians. – Ecology 80: 2442–2448.

Kingsland, S. E. 1995. Modeling nature: episodes in the history of population ecology, 2nd ed. – Chicago Univ. Press.

Kupferberg, S. J. 1997. Bullfrog (*Rana catesbeiana*) invasion of a California river: the role of larval competition. – Ecology 78: 1736–1751.

Lawler, S. P. and Morin, P. J. 1993. Temporal overlap, competition, and priority effects in larval anurans. – Ecology 74: 174–182.

Martinez, I. P., Alvarez, R. and Harraez, M. P. 1996. Growth and metamorphosis of *Rana perezi* in culture: effects of larval density. – Aquaculture 142: 163–170.

Mertz, D. B. and McCauley, D. E. 1980. The domain of laboratory ecology. – Synthese 43: 95–110.

Morin, P. J. 1983. Predation, competition, and the composition of larval anuran guilds. – Ecol. Monogr. 119–138.

Morin, P. J. 1986. Interactions between intraspecific competition and predation in an amphibian predator-prey system. – Ecology 67: 713–720.

Morin, P. J. 1987. Predation, breeding asynchrony, and the outcome of competition among treefrog tadpoles. – Ecology 68: 675–683.

Morin, P. J. 1989. New directions in amphibian community ecology. – Herpetologica 45: 124–128.

Morin, P. J. 1998. Realism, precision, and generality in experimental ecology. – In: Resetarits, W. R. and Bernardo, J. (eds), Ecological experiments: issues and perspectives. Oxford Univ. Press, pp. 50–70.

Morin, P. J. and Johnson, E. A. 1988. Experimental studies of asymmetric competition among anurans. – Oikos 53: 398–407.

Morin, P. J., Lawler, S. P. and Johnson, E. A. 1988. Competition between aquatic insects and vertebrates: interaction strength and higher order interactions. – Ecology 69: 1401–1409.

Newman, R. A. 1998. Ecological constraints on amphibian metamorphosis: interactions of temperature and larval density with responses to changing food level. – Oecologia 115: 9–16.

Odum, E. P. 1984. The mesocosm. – Bioscience 34: 558–562.

Paine, R. T. 1992. Food-web analysis through field measurement of per-capita interaction strength. – Nature 355: 73–75.

Parris, M. J. and Semlitsch, R. D. 1998. Asymmetric competition in larval amphibian communities: conservation implications for the northern crawfish frog, *Rana areolata circulosa*. – Oecologia 116: 219–226.

Peacor, S. D. and Werner, E. E. 1997. Trait-mediated indirect interactions in a simple aquatic food web. – Ecology 78: 1146–1156.

Pehek, E. L. 1995. Competition, pH, and the ecology of *Hyla andersonii*. – Ecology 76: 1786–1793.

Peters, R. H. 1991. A critique for ecology. – Cambridge Univ. Press.

Petersen, J. E., Chen, C.-C. and Kemp, W. M. 1997. Scaling aquatic primary productivity: experiments under nutrient- and light-limited conditions. – Ecology 78: 2326–2338.

Petersen, J. E., Cornwell, J. C. and Kemp, W. M. 1999. Implicit scaling in the design of experimental aquatic ecosystems. – Oikos 85: 3–18.

Petranka, J. W. and Kennedy, C. A. 1999. Pond tadpoles with generalized morphology: is it time to reconsider their functional roles in aquatic communities? – Oecologia 120: 621–631.

Relyea, R. A. and Werner, E. E. 1999. Quantifying the relation between predator-induced behavioral responses and growth performance in larval anurans. – Ecology 80: 2117–2124.

Reques, R. and Tejedo, M. 1997. Reaction norms for metamorphic traits in natterjack toads to larval density and pond duration. – J. Evol. Biol. 10: 829–851.

Resetarits, W. R. and Fauth, J. E. 1998. From cattle tanks to Carolina bays: the utility of model systems for understanding natural communities. – In: Resetarits, W. R. and Bernardo, J. (eds), Ecological experiments: issues and perspectives. Oxford Univ. Press, pp. 133–151.

Riha, V. F. and Berven, K. A. 1991. An analysis of latitudinal variation in the larval development of the wood frog (*Rana sylvatica*). – Copeia 1991: 209–221.

Rosenberg, M. S., Adams, D. C. and Gurevitch, J. 1997. MetaWin: statistical software for meta-analysis with resampling tests. Ver. 1.0. – Sinauer.

Rowe, C. L. and Dunson, W. A. 1994. The value of simulated pond communities in mesocosms for studies of amphibian ecology and toxicology. – J. Herpetol. 28: 346–356.

Sarnelle, O. 1997. *Daphnia* effects on microzooplankton: comparisons of enclosure and whole-lake responses. – Ecology 78: 913–928.

Schoener, T. W. 1983. Field experiments on interspecific competition. – Am. Nat. 122: 240–285.

Semlitsch, R. D. 1993. Asymmetric competition in mixed populations of tadpoles of the hybridogenetic *Rana esculenta* complex. – Evolution 47: 510–519.

Semlitsch, R. D. and Caldwell, J. P. 1982. Effects of density on growth, metamorphosis and survivorship in tadpoles of *Scaphiopus holbrooki*. – Ecology 63: 905–911.

Semlitsch, R. D. and Reyer, H.-U. 1992. Performance of tadpoles from the hybridogenetic *Rana esculenta* complex: interactions with pond drying and interspecific competition. – Evolution 46: 665–676.

Skelly, D. K. 1995a. A behavioral trade-off and its consequences for the distribution of *Pseudacris* treefrog larvae. – Ecology 76: 150–164.

Skelly, D. K. 1995b. Competition and the distribution of spring peeper larvae. – Oecologia 103: 203–207.

Smith, D. C. 1983. Factors controlling tadpole populations of the chorus frog (*Pseudacris triseriata*) on Isle Royale, Michigan. – Ecology 64: 501–510.

Smith, D. C. 1987. Adult recruitment in chorus frogs: effects of size and date of metamorphosis. – Ecology 68: 344–350.

Smith, D. C. 1990. Population structure and competition among kin in the chorus frog (*Pseudacris triseriata*). – Evolution 44: 1529–1541.

Smith-Gill, S. J. and Gill, D. E. 1978. Curvilinearities in the competition equations: an experiment with ranid tadpoles. – Am. Nat. 112: 557–570.

Sredl, M. J. and Collins, J. P. 1991. The effect of ontogeny on interspecific interactions in larval amphibians. – Ecology 72: 2232–2239.

Sredl, M. J. and Collins, J. P. 1992. The interaction of predation, competition, and habitat complexity in structuring an amphibian community. – Copeia 1992: 607–614.

Steinwascher, K. 1978. Interference and exploitation competition among tadpoles of *Rana utricularia*. – Ecology 59: 1039–1046.

Tejedo, M. and Reques, R. 1992. Effects of egg size and density on metamorphic traits in tadpoles of the natterjack toad (*Bufo calamita*). – J. Herpetol. 26: 146–152.

Tejedo, M. and Reques, R. 1994. Plasticity in metamorphic traits of natterjack tadpoles: the interactive effects of density and pond duration. – Oikos 71: 295–304.

Travis, J., Keen, W. H. and Julianna, J. 1985. The effects of multiple factors on viability selection in *Hyla gratiosa* tadpoles. – Evolution 39: 1087–1099.

Warner, S. C., Dunson, W. A. and Travis, J. 1991. Interaction of pH, density, and priority effects on the survivorship and growth of two species of hylid tadpoles. – Oecologia 88: 331–339.

Warner, S. C., Travis, J. and Dunson, W. A. 1993. Effect of pH variation on interspecific competition between two species of hylid tadpoles. – Ecology 74: 183–194.

Wellborn, G. A., Skelly, D. K. and Werner, E. E. 1996. Mechanisms creating structure across a freshwater habitat gradient. – Annu. Rev. Ecol. Syst. 27: 337–363.

Werner, E. E. 1986. Amphibian metamorphosis: growth rate, predation rate, and the optimal size at transformation. – Am. Nat. 128: 319–341.

Werner, E. E. 1992. Competitive interactions between wood frog and northern leopard frog larvae: the influence of size and activity. – Copeia 1992: 26–35.

Werner, E. E. 1998. Ecological experiments and a research program in community ecology. – In: Resetarits, W. R. and Bernardo, J. (eds), Ecological experiments: issues and perspectives. Oxford Univ. Press, pp. 3–26.

Werner, E. E. and Anholt, B. A. 1996. Predator-induced behavioral indirect effects: consequences to competitive interactions in anuran larvae. – Ecology 77: 157–169.

Werner, E. E. and Glennemeier, K. S. 1999. Influence of forest canopy cover on the breeding pond distributions of several amphibian species. – Copeia 1999: 1–12.

Wilbur, H. M. 1976. Density-dependent aspects of metamorphosis in *Ambystoma* and *Rana sylvatica*. – Ecology 57: 1289–1296.

Wilbur, H. M. 1977a. Density-dependent aspects of growth and metamorphosis in *Bufo americanus*. – Ecology 58: 196–200.

Wilbur, H. M. 1977b. Interactions of food level and population density in *Rana sylvatica*. – Ecology 58: 206–209.

Wilbur, H. M. 1989. In defense of tanks. – Herpetologica 45: 122–123.

Wilbur, H. M. 1997. Experimental ecology of food webs: complex systems in temporary ponds. – Ecology 78: 2279–2302.

Wilbur, H. M. and Alford, R. A. 1985. Priority effects in experimental pond communities: responses of *Hyla* to *Bufo* and *Rana*. – Ecology 66: 1106–1114.

Wiltshire, D. J. and Bull, C. M. 1977. Potential competitive interactions between larvae of *Pseudophryne bibroni* and *P. semimarmorata* (Anura: Leptodactylidae). – Aust. J. Zool. 25: 449–454.

Woodward, B. D. 1987. Interactions between Woodhouse's toad tadpoles (*Bufo woodhousii*) of mixed sizes. – Copeia 1987: 380–386.

Wootton, T. W. and Pfister, C. A. 1998. The motivation and context of experiments in ecology. – In: Resetarits, W. R. and Bernardo, J. (eds), Ecological experiments: issues and perspectives. Oxford Univ. Press, pp. 350–369.

Appendix. Fifty-two studies included in meta-analyses of competition among larval anurans. For each reference, the authors, year of publication, species included (four-letter codes defined at bottom of table), and experimental venue (field, lab, or mesocosm) are followed by the number of intraspecific and interspecific comparisons gleaned from each study.

Author(s)	Year	Species	Venue	No. of comparisons	
				Intraspecific	Interspecific
Alford and Wilbur	1985	BUAM, RASP	M		2
Berven and Gill	1983	RASY	L	15	
Blaustein and Margalit	1996	BUVI	L	1	
Brockelman	1969	BUAM	F	2	
Brönmark et al.	1991	RATE	M	1	
Cummins	1989	RATE	L	4	
Dash and Hota	1980	RATI	L	2	
DeBenedictis	1974	RAPI, RASY	F		8
Gascon	1992	EPFE, OSTA, PHTO	M		6
Gascon and Travis	1992	RAUT	M	2	
Griffiths	1991	BUCA, RATE	M	3	4
Griffiths et al.	1991	BUCA, RATE	M		2
Griffiths et al.	1993	BUCA, RATE	M		2
Hokit and Blaustein	1994	RACS	L	2	
Kiesecker and Blaustein	1999	HYRE, RACS	M	2	2
Kupferberg	1997	HYRE, RABO, RACT	F		2
Lawler and Morin	1993	BUWO, PSCR	M		2
Martinez et al.	1996	RAPE	L	6	
Morin	1986	PSCR	M	2	
Morin	1987	HYVE, PSCR	M		1
Morin et al.	1988	BUWO, HYAN	M		1
Newman	1998	SCCO	L	3	
Parris and Semlitsch	1998	RABL, RASP	M	2	
Peacor and Werner	1997	RACL, RACT	M	1	1
Pehek	1995	HYAN, HYVE, RASP	M		4
Reques and Tejedo	1997	BUCA	L, M	1,1	
Relyea and Werner	1999	RACL, RACT	F		2
Riha and Berven	1991	RASY	L	6	
Semlitsch	1993	RAES, RALE	M	4	6
Semlitsch and Caldwell	1982	SCHO	L	4	
Semlitsch and Reyer	1992	RAES, RALE	M		2
Skelly	1995a	PSCR, PSTR	F		7
Skelly	1995b	PSCR, RACL	F	2	3
Smith	1990	PSTR	F	7	
Smith-Gill and Gill	1978	RAPI, RASY	L	8	8
Sredl and Collins	1991	PSTR	F	1	
Sredl and Collins	1992	HYEX	F	2	
Steinwascher	1978	RAUT	L	4	
Tejedo and Reques	1992	BUCA	L	3	
Tejedo and Reques	1994	BUCA	M	2	
Travis et al.	1985	HYGR	F	4	
Warner et al.	1991	HYFE, HYGR	M	6	
Warner et al.	1993	HYFE, HYGR	M	4	8

Appendix (Continued)

Author(s)	Year	Species	Venue	No. of comparisons	
				Intraspecific	Interspecific
Werner	1992	RAPI, RASY	L	4	4
Werner and Anholt	1996	RACA, RACL	M	4	
Werner and Glennemeier	1999	RAPI, RASY	F		8
Wilbur	1976	RASY	F	5	
Wilbur	1977a	BUAM, RAPA	L, F	5,5	0,3
Wilbur	1977b	RASY	L	3	
Wilbur and Alford	1985	BUAM, HYCR, RASP	M		2
Wiltshire and Bull	1977	PDBI, PDSE	F	2	
Woodward	1987	BUWO	L	2	

Species codes. BUAM *Bufo americanus*, BUCA *Bufo calamita*, BUVI *Bufo viridis*, BUWO *Bufo woodhousii*, EPFE *Epipedobates femoralis*, HYAN *Hyla andersonii*, HYCR *Hyla chrysoscelis*, HYEX *Hyla eximia*, HYFE *Hyla femoralis*, HYGR *Hyla gratiosa*, HYRE *Hyla regilla*, HYVE *Hyla versicolor*, OSTA *Osteocephalus taurinus*, PHTO *Phyllomedusa tomopterna*, PDBI *Pseudophryne bibroni*, PDSE *Pseudophryne semimarmorata*, PSCR *Pseudacris* (= *Hyla*) *crucifer*, PSTR *Pseudacris triseriata*, RABL *Rana blairi*, RABO *Rana boylii*, RACL *Rana clamitans*, RACS *Rana cascadae*, RACT *Rana catesbeiana*, RAES *Rana esculenta*, RALE *Rana lessonae*, RAPA *Rana palustris*, RAPE *Rana perezi*, RAPI *Rana pipiens*, RASP *Rana sphenocephala*, RASY *Rana sylvatica*, RATE *Rana temporaria*, RATI *Rana tigrina*, RAUT *Rana utricularia*, SCCO *Scaphiopus couchi*, SCHO *Scaphiopus holbrooki*.