Modelling terrestrial interactions and shelter use in great crested newts (*Triturus cristatus*)

Jan C. Malmgren¹*, Per-Åke Andersson², Simon Ekdahl³

Abstract. The great crested newt (*Triturus cristatus*) is a pond-breeding salamander with a prolonged aquatic phase and a period of activity in the terrestrial environment prior to hibernation. Individuals use ground-covering objects and burrows as shelters from predators and unfavourable conditions. We investigated how interactions and spacing patterns were formed in staged experiments where paired newts were exposed to an arena with two shelters. Shelter use was predominant in all sets of great crested newts, irrespective of sex and size. Males and females differed in their shelter-use response. Females were more prone to reside in shelters when meeting other females than when meeting males, whereas males appeared indifferent to the sex of another individual. The social interaction was uniform for all combinations, with a 2:1 relation between the numbers of pairs sharing shelters vs. positioned in different shelters, instead of a random 1:1 distribution. Thus, the differences in shelter sharing could be attributed to the predominance of shelter use for females when meeting other females. We suggest that this behaviour, when females respond to the presence of another female, is a form of resource protection. Although sex effects were clearly noted, we could not demonstrate that size-related characteristics affected the outcome of any social interactions.

Introduction

Terrestrial spacing patterns and interactions among salamanders are shaped by resource use and competition (Brown and Orians, 1970; Bruce et al., 2000). If resources (e.g., food and shelters) are of limited supply, defendable, and important for individual fitness, territorial behaviour may be expressed – albeit correlated with life cycle complexity (Mathis et al., 1995). Studies on effects from terrestrial resource competition have essentially focused on plethodontid salamanders (Jaeger, 1981; Keen and Reed, 1985; Hairston, 1987; Maerz and Madison, 2000), which are largely terrestrial and establish and defend territories linked to food access, shelter availability and population density. However, salamanders with the ancestral biphasic lifecycle – living in both aquatic and terrestrial habitats – have received much less attention. It has been suggested that because pond-breeding species are less dependent on the terrestrial component of their habitats, they are less likely to defend resources and to display territoriality there (Martin et al., 1986). Competitive interactions, for example agonistic behaviour such as aggression and biting, have nevertheless been observed in the terrestrial phase of ambystomatid salamanders (e.g., Ducey and Rissem, 1988; Ducey, 1989, Smyers et al., 2001, 2002). However, many *Ambystoma* species are explosive pond breeders and only briefly spend time in their aquatic habitat (Petranka, 1989). Thus, the hypothesis that such amphibians are less likely to display territoriality, has apparently yet not been tested in caudates with a prolonged breeding season in relation to the terrestrial phase, and not in any salamandrid species (but see Riberon and Miaud, 2000; Gautier and Miaud, 2003; Gautier et al., 2004).

Great crested newts (*Triturus cristatus*) are pond-breeding Eurasian salamanders (family Salamandridae) with a biphasic lifecycle. Much of the knowledge about their biology (reviewed by Griffiths, 1996; Thiesmeier and Kupfer, 2000; Arntzen, 2003) concerns aquatic characteristics, mainly because newts are readily accessible to study when adults and larvae reside in breeding ponds. Individuals spend a considerable amount of their lifetime in terrestrial habi-
tats, but the ecology of this phase is poorly known. This is particularly challenging since there are conservation concerns associated with the great crested newt throughout its range of distribution, and there is an increasing need for knowledge-based management practices for terrestrial habitats (e.g., Langton et al., 2001; Malmgren, 2001, 2002). Telemetric studies on great crested and closely related newts have demonstrated that adults use burrows of small mammals and crevices under logs, rocks, grass tussocks, and other ground-covering objects as land refuges (Jehle, 2000; Jehle and Arntzen, 2000; Schabetsberger et al., 2004; Malmgren et al., unpubl. data). Such structures are thought to constitute important spatial units for individuals because they provide shelters from predators and unfavourable conditions. Newts are likely to use shelters as sanctuaries to which they return during the day after nightly movements outside burrows (e.g., while feeding), or during periods of drought or cold (i.e., for aestivation and overwintering, respectively). Despite the ability of newts to use olfactory cues for navigating (e.g., Joly and Miaud, 1993; Malmgren, 2002) and identifying conspecifics (Hayward et al., 2000), there is much to learn about the social interactions among newts in relation to density-dependent resource availability, such as shelters.

In south-central Sweden, the aquatic breeding period of adult great crested newts extends from late April to late July or early August, and the active terrestrial period proceeds into late September – after which hibernation takes place (Malmgren, 2001). The importance of this relatively short period of terrestrial activity for foraging and preparing for hibernation is largely unknown. We hypothesised that such a short period does not favour competitive interactions between individuals to defend terrestrial shelters, unless their qualities are essential for foraging or survival to or through hibernation, particularly if territorial behaviour is costly.

In this study, we used an experimental arena in the laboratory to test the hypothesis that paired individuals of the great crested newts interact and influence each others’ shelter use and the choice of sharing shelters. We predicted that shelters are an essential resource to newts and that mutual exclusion or even agonistic behaviour would indicate that shelter sharing may be costly and shelters worth defending. Conversely, non-territorial behaviour (e.g., aggregation) would indicate that shelters are important but not necessarily limiting to the newts or that company of individuals is even preferred. Specifically, in this study we tested whether (A) newts use shelters when provided, (B) individuals co-occupy or exclude each other from shelters, and (C) sex and size of two interacting individuals affect how they position themselves in relation to shelter availability. In order to test this, we used a novel and sample-efficient group-based tournament design evaluated by exact test statistics and Monte Carlo simulations, also involving decision-models which we test our data against. This approach may be useful in future behavioural studies where the availability of individuals to sample is limited.

Material and methods

Capture and housing of newts

In July 2002, we collected 34 postbreeding adult great crested newts (17 females and 17 males) using drift fences with pitfall traps during an emigration episode from a breeding pond (Lillsjön, 16 km W Örebro, south-central Sweden: 59°15′N 14°58′E). All newts were sexed and measured (SVL, snout-to-vent length: from tip of snout to posterior end of cloacal lips). They were housed individually in ventilated plastic terrariums (17 × 17 × 11 cm) with moist Sphagnum moss as floor substrate, placed in a shaded outdoor pen for a maximum of 26 days (mean = 22.4). Each newt was used in experiments every second night and returned to its individual terrarium and fed small crickets and earthworms every other day.

Experimental setup

We designed a setup to produce artificial ‘shelter/non shelter’ environments in which pairs of individuals were tested in enclosures. The general idea was to let each individual select a position in relation to shelter availability and the presence of another individual in paired trials. Because of a limited supply of newts, and to evaluate the effect of sex on interactions among individuals, we divided the newts into
Terrestrial interactions in great crested newts

Three separate groups, or trial sets; (1) ten females (female set), (2) ten males (male set) and (3) seven females and seven males (mixed set). Within the female and male sets, each newt was faced against all other individuals of the same sex (producing 45 trials per set), whereas in the mixed set each newt was faced against all individuals of the opposite sex (producing 49 trials). None of the subjects were faced with newts from another set, or with an individual from a prior trial. Thus, each trial represents a unique encounter between two individuals.

Twelve identical arenas were constructed from semi-transparent plastic boxes (40 × 30 × 22 cm), used in simultaneous trials. Each arena was furnished with a substrate of two plastic foam sheet layers (polyurethane, 5 and 20 mm thick, dark grey and off-white, respectively). Two holes (∅ 105 mm) were drilled through the thicker off-white layer, which was placed over the dark grey layer. Each hole was partly covered by a plastic plate (120 × 90 × 3 mm), leaving a 20 mm wide entrance opening. Thus, each arena presented two shaded artificial pit-shelters with entrances facing away from each other (fig. 1). Each trial was prepared by thoroughly cleaning and washing the arena boxes and shelter plates with detergents, water and 70% ethanol. They were then left to dry before refurnishing with entirely new substrate linings. We moistened the substrate with 0.2 litres of dechlorinated (i.e., aerated) tap water. Since the great crested newt is mainly nocturnal, all experiments were conducted during 1900-0800 h in a laboratory room, at around 20°C. The room was held dark during 2200-0600 h, and during 0600-2200 h daylight conditions were simulated by fluorescent tubes (2 × 36 W) placed 1 m above the arenas, to mimic ambient conditions. Trials were started at 1900 h by placing two individuals side by side in the middle of each arena between the shelters, facing opposite long sides. Arenas were then sealed with ventilated semi-transparent lids. At 0800 h the trial ended and the positions of both newts were noted by drawing the positions of each individual.

Figure 1. Schematic overview of the experimental arena, from above. Two layers of polyurethane foam lined the floor. The top foam sheet had two holes (one shown with dashed line), each partly covered by a plastic plate (white rectangles) to produce artificial shelters. The solid black area indicates the entrance opening, which was 20 mm at the widest point (white line). The shelter consisted of a 20 mm deep and 105 mm wide circular pit, lined with a moistened grey thin polyurethane foam layer, and partly covered by the plastic plate.

Statistical analysis

The predicament of small samples is that any approximation, for example by a standard distribution or a standard test statistic, is flawed. As an example, such approximations are implicit in regression models. Thus, we tried to avoid these difficulties by only performing exact tests. This means that if a test quantity, Z, has no well-known, tabulated distribution function, its distribution can be determined by Monte Carlo simulation. By comparing each of 10 000 randomly generated Z-values with the observed (sample) value z0, a P-value can be computed that denotes the probability of getting an equally or more extreme test value in the relevant distribution tail – \( P = P(Z \geq z_0) \) above – or for comparison with the used level of significance \( \alpha = 5\% \) (one-tailed test) or \( \alpha/2 \) (two-tailed test). In this paper we used well-established test statistics (all presented in Appendix) becoming, e.g., asymptotically Normal-distributed as the sample size tends to infinity, but instead of relying on the asymptotic properties exact tests were performed. Further, we applied parametric resampling (see, e.g., Efron and Tibshirani, 1993, pp. 53, 306; Hjorth, 1994, p. 131) by drawing the random numbers without replacement, in order to force the parameter estimates to coincide with that of the observed sample. This is a means for reducing variance, insofar as any expected conditional variance is less than the corresponding unconditional variance. However, if a P-value is computed by simulation, this is an approximation of \( \hat{P} \) as well, but one that is up to our own ambition to improve, simply by running a longer simulation. Moreover, by running 25 repeated simulations we could form an approximate 95%-confidence interval for the P-value in every case of interest. Since our conclusions are based on statistical tests, mostly with one parameter value specified by a null hypothesis, our general resampling procedure is simpler than for example bootstrap methods (e.g., Efron and Tibshirani, 1993; Hjorth, 1994; Manly, 1997), except in our final step of analysis, where several parameters are involved. All statistical analyses were implemented in Matlab\textsuperscript{2}.

We performed the statistical tests in a step-by-step procedure, similar to for example stepwise multilinear regression, with the initial steps focusing on the relevance of our treatment of the paired meetings as repeated independent statistical trials. First (step 1), we scrutinized the series of experiments for each individual, and secondly (step 2) the series for different collections of newts. After this initial justification of pooling newts into sets, we investigated the group behaviour (steps 3-4). Finally (step 5) we chose a decision model. From an individual newt’s perspective, possible outcomes from our experimental design would be to either occupy a shelter or to remain outside the shelters at the end of a trial, combined with either sharing or not sharing a particular shelter with the other newt. Our primary aim was to characterise the outcomes \( y \) of each trial, in relation to shelter use and sharing, by trying different z-variables.

\[ P(Z \leq 0) = \frac{1}{2} \text{erfc}\left(\frac{z}{\sqrt{2}}\right) \]

\[ P(Z \geq 0) = 1 - P(Z \leq 0) = \frac{1}{2} \text{erfc}\left(\frac{z}{\sqrt{2}}\right) \]
based on newt set, gender and size. We also investigated two types of decision models. Apart from a single-stage model for \( y \) we formulated and tried two-stage models, with separate decision models for stage 1 (pure shelter choice, outcome \( y_1 \)) and stage 2 (pure interaction choice, outcome \( y_2 \)), with combined outcomes \( y = (y_1, y_2) \). Since each combined outcome occurs twice in the observation data – for both individuals in the pair on trial – we consequently generated those Monte Carlo random numbers per trial, not per individual.

Since each individual newt experienced repeated trials, albeit meeting different individuals in each instance, we first investigated (1) learning effects by interpreting learning for an individual as an increased number of repeated choices over the series of trials. As for shelter occupation this should be manifested by the newt almost always settling inside or almost always settling outside a shelter towards the end of the series. Evidence for a learning effect would then be recognised as a significantly changed proportion of in-shelter outcomes between the first and second half of the series. To test this hypothesis we considered the number of in-shelter outcomes in the whole series as given (by the sample) for each individual, whence we could apply a hypergeometrically distributed test statistic \( X_1 \) for the number of in-shelter outcomes in the first half. As for shelter sharing a constant individual probability was not an obligation or even expected, since the meeting characteristics, e.g. difference in size, might have an influence upon the outcome of a trial. We will return to this matter in (5) below.

We controlled for (2) homogeneous populations by testing if there were differences between the individuals in each data set. In the mixed set we also distinguished between the two subsets of male and female newts. Thus the three experimental data sets (with subsets) constituted four different random samples, each representing the local statistical population of newts of a certain sex in a certain test situation: \( F_f \) (Female meeting female), \( M_m \) (Male meeting male), \( F_m \) (Female meeting male) and \( M_f \) (Male meeting female). The null hypothesis was that the proportion of shelter use is the same for the whole population. By considering the overall in-shelter use in each examined set as fixed, the in-shelter vs. out-shelter use across individuals becomes multivariate hypergeometrically distributed, but due to its nasty combinatorial effects on CPU-times, we turned to a standard (asymptotic) Chi-square homogeneity test, for test statistic \( V \).

To test for (3) gender effects, as to shelter use and interactions among newts, we formulated probability models for newt decisions, based on our experimental setup. In the geometric choice model we assumed that a newt’s position would be random according to the total base area of the arena (\( 1200 \text{ cm}^2 \)) and the area of each shelter opening (\( 11.5 \text{ cm}^2 \)). In the single-stage choice model we assumed that each shelter and the outside area of the arena represent three different outcomes – all equally probable. In the two-stage choice model we assumed that an individual in the first decision stage chose whether to be inside or outside a shelter, and that the paired newts in the second decision stage chose whether to share or not to share a shelter. We had two alternative formulations for the 1st stage of the two-stage model – either that the two outcomes \( (y_1 = (I, O)) \) are equally probable (\( p_I = \frac{1}{2} \)), or that the shelter outcome was determined by an unknown probability \( p_I \) that had to be estimated from data. Common to all these models is a random choice sub-model for the 2nd stage, i.e. – given that the two newts use the shelters provided – the probabilities of sharing (I) or not sharing (In) a shelter are equal (\( q_s = \frac{1}{2} \)). We also supplied an interaction sub-model for the 2nd stage, whence the sharing probability \( q_s \) had to be estimated. We performed several tests in order to evaluate the decision models, and to investigate whether the kind of meeting, with another newt of the same or the opposite sex, had an influence on newt behaviour. As for (3a) the pure in/out-shelter choice, we formulated and applied binomial statistics \( X \) to the cases \( p_I = \frac{1}{2} \), \( p_I = \frac{3}{4} \) and asymptotically standard-normal distributed test quantities \( Z \) to all relevant pairs of populations. Accepted decision models can be used to judge whether the paired individuals make independent choices inside vs. outside shelters. We utilised asymptotically Chi-square distributed test statistics \( Y \) (as to the simultaneous number of newts found in-shelter, i.e., 0, 1 or 2, rather than the standard tests of independence). By (3b) the pure interaction choice, we imply whether to share or not to share a shelter, given that both individuals in a pair settle in-shelter. We used an asymptotically normal test statistic \( Z \) for pair-wise comparisons between populations and a binomial test quantity \( X \) for an examination of the pure random choice hypothesis (\( q_s = \frac{1}{2} \)). As for (3c) the combined choice inside vs. outside and sharing vs. not sharing shelter, as described by our different decision models, we tested the null hypothesis that the two individuals in a pair made independent choices and that the two shelters were equally probable. For the combined choice of a pair of newts, four possible outcomes of the trial were identified: both newts outside shelter, one inside and one outside, both inside sharing shelter and both inside in different shelters. For an evaluation of the model hypotheses we applied distribution tests, based on asymptotically Chi-square distributed test quantities \( Y \).

To investigate (4) size effects on the outcomes for paired individuals we used two different approaches. For the choice inside or outside shelters within each set we used asymptotic Chi-square tests of independence, splitting the set into two categories, corresponding to small and large SVL-related x-values. For the choice of sharing or not sharing shelters, given that both newts were inside shelters, we used the same type of test variables \( R \) as well as applied asymptotic logistic regression tests, with test statistics \( G \) and logits \( q_s = a + bx \). Here we used semi-parametric resampling: While keeping the number of sharing outcomes fixed in the simulations we estimated \( a \), \( b \) by maximising the likelihood function. In all these tests \( x \) stands for one difference in SVL, average SVL, minimum SVL, maximum SVL, male SVL in the mixed set and female SVL in the mixed set, for each pair of newts.

To evaluate (5) the choice of decision model, the analysis depended on the results in steps 3-4 above. For example, if possible we wanted to complete the learning and homogeneity tests in steps 1-2 before proceeding. Even further, if more than one decision model remained from the early
steps, we would decide to use approximate Chi-2 goodness-of-fit tests for choosing one model for recommendation. If several distribution parameters proved to be needed in the model, we would be ready to consider using bootstrap technique for evaluation purposes.

Uncommented Matlab® code, for the steps of analysis presented here, is available from the first author upon request.

Results

Between experiments, with only three trials remaining, a male escaped from its individual terrarium after having been involved in trials. Further, in eleven trials an individual had pressed itself down under the foam lining. We omitted these trials from the results, together with a female outlier (see below), thereby lowering the number of outcomes for each set (final set numbers: females \( n = 34 \), males \( n = 36 \), mixed \( n = 46 \)).

1) Learning effects. – A crucial test of our experimental design was whether a newt would show a systematically changed behaviour over a series of trials, through a changed proportion of in-shelter outcomes. We could not detect a significant difference in behaviour between the first and second half of a series for any of the individuals in any of the three sets (all \( P > 0.05 \) from 34 exact tests; the most extreme result \( P = 0.14 \) attained for a male in the male set). Thus we found no evidence of learning effects.

2) Homogeneous populations. – Another crucial test was whether the newts in the sets behaved differently in relation to the experimental setup. No significant differences in the proportion of in-shelter choices were found between the individuals in the male set (\( P = 0.21 \)) and mixed set (\( P = 0.0685 \pm 0.0012 \) (95% = simulated 95%-confidence interval). Hence there was no evidence of inhomogeneous populations. These results, together with the absence of significant learning effects, suggest that the general statistical demand for independent random trials was fulfilled. Thus, our economic group-based tournament-design of reusing a limited number of newts in several trials, instead of using different individuals in every experiment, was validated. The option of testing properties like learning and homogeneity was therefore a bonus from the experimental design, which would have been more or less impossible with alternative designs. Further, through this method we detected that one individual, \( \tilde{N}_2 \), in the female set, clearly deviated behaviourally from the rest, in that she never shared a shelter. In the in-shelter vs. out-shelter homogeneity test the simulations corresponded to \( P < 0.001 \) with female \( \tilde{N}_2 \) included, and \( P = 0.40 \) with her data removed. Hereafter we point out the (few) test results where it makes a significant difference to have female \( \tilde{N}_2 \) included or excluded.

3a) Pure in/out-shelter choice. – In this step we tested the various proposals for the in-shelter probability \( p_I \). In the female set the newts were observed to use shelters in 91% of the cases, whereas the proportion was 67% in the male set. In the mixed set the in-shelter proportion was 61%, with 63% for the female \( Fm \) and 59% for the male \( Mf \) population. The hypothesis of two equally probable inside/outside outcomes \( (p_I = \frac{1}{2}) \) could be rejected in tests of all sets \( (P < 0.025, \text{two-tailed tests}) \). The hypothesis \( p_I = \frac{1}{2} \) was accepted for the male and mixed sets, both when treated separately and pooled \( (P = 0.21) \), but was rejected for the female set \( (P < 0.001) \). The hypothesis of a common, unknown \( p_I \)-value was tested by pair-wise comparisons. The only significant results were between the single-gender sets \( Ff \) and \( Mm \) \( (P < 0.001) \) and between the female populations \( Ff \) and \( Fm \) \( (P < 0.001) \). Thus, in trials where females met females they displayed a different behaviour than compared to other meetings. We also tested if the two newts in a pair made independent choices of ending up inside or outside shelter. For all our accepted probability models, i.e., both \( p_I = \frac{1}{2} \) and an estimated \( \hat{p}_I = 0.634 \) for \( Mm + Mf + Fm \) and an estimated \( \hat{p}_I \) for \( Ff \), the null hypothesis of independent choices could
not be rejected. The independency test result closest to significance was distant: For \( Mm + Mf + Fm \) we registered 36 \( I \setminus I \), 32 \( I \setminus O \) and 14 \( O \setminus O \) outcomes, and received \( P = 0.21 \) with \( p_I = \frac{3}{2} \). Consequently, for characterizing the first model stage of outcomes (i.e., \( y_1 = \{I, O\} \)), one \( x \)-variable was needed indicating a newt’s population, since shelter-choice depends on whether a newt belongs to the population \( Ff \) or not.

3b) Pure interaction choice. – We tested two proposed models for the sharing-shelter probability \( q_s \). The observed sharing proportions was 62\% in the female set, 61\% in the male and 72\% in the mixed set. Pair-wise comparisons between the sets showed no significant differences, with the most extreme result for the female set vs. the mixed set \( (P = 0.27, \text{two-tailed} \) test). This implies that the three experimental data sets are indistinguishable as to the sharing vs. non-sharing choice, and could therefore be pooled into one overall set. The observed pooled sharing proportion 0.65 was significantly different from random choice \( q_s = \frac{1}{2} \) \( (P = 0.012, \text{two-tailed test}) \), which suggests that the newts interacted in favour of aggregation, with \( \hat{q}_s = 0.65 \). For characterizing the second model stage of outcomes (i.e., \( y_2 = \{Is\setminus Is, In\setminus In\} \)), there was no need for a further \( x \)-variable. Observe that we used the same \( x \)-data both for \( y_1 \) and \( y_2 \), so our simultaneous level of significance is around \( 2 \alpha = 10\% \).

3c) Combined choice. – The considered outcomes for the pair of newts in a trial are \( Is\setminus Is, In\setminus In, In\setminus O, O\setminus O \). We tested if these outcomes could be simultaneously modelled by some \( p_I \) and \( q_s \). For the pooled male+mixed set both \( p_I = \frac{1}{2} \) and the observed (estimate) \( \hat{p}_I = 0.63 \) remained as proposals from test step 3a; for the female set \( \hat{p}_I = 0.91 \). As for the pooled male+mixed sets, with observed frequencies (24, 12, 32, 14) of outcomes \( Is\setminus Is \), etc., the model \( p_I = \frac{1}{2} \) was accepted \( (P = 0.0662 \pm 0.0009 \ (95\%)) \), while \( \hat{p}_I = 0.63 \) was rejected \( (P = 0.035) \).

As for the female set, with observed frequencies (18, 11, 4, 1), \( \hat{p}_I = 0.91, q_s = \frac{1}{2} \) was accepted \( (P = 0.093) \). Thus, the overall dominance for sharing a shelter, which was confirmed on sharing vs. non-sharing in test step 3b above, was only partially verified by these tests for combined choices. The main reason is purely statistical, since significant differences in the in-shelter probabilities \( p_I \) between population \( Ff \) and the others prohibit a full pooling of the data sets here.

4) Size effects. – The independence tests revealed no significant dependence on size, neither for inside vs. outside shelter use, nor for sharing vs. non-sharing. By using \( x = (\text{absolute}) \text{SVL} \) difference in the female set, partitioned into \( L \) = large and \( S \) = small for sharing (\( Is \)) vs. non-sharing (\( In \)), we observed 8 \( L\ canceled 10 S\ canceled 9 L\ canceled 3 S\ canceled 3 \) outcomes, corresponding to \( (P = 0.14) \); all other \( P \)-values were even higher. As for the exact tests based on logistic regression, for the sharing vs. non-sharing interaction, none of the size-related variables \( x \) that we applied to each pair of newts produced an \( x \)-coefficient \( b \) significantly different from zero. In the mixed set, we observed 13 \( Is\ canceled 5 In\ canceled 14 \) outcomes, and by using \( x = \text{male SVL} \), we obtained \( a + bx = -2.7 + 2.5x \) \( (P = 0.13) \); all other \( P \)-values were higher. Thus we found no evidence of size effects.

5) Choice of decision model. – The findings in steps 3-4 did not lead to a unique decision model and we therefore tested for the best model. Moreover, since the sharing-shelter behaviour was found to be uniform for all newts we should complete the testing for unwanted learning and inhomogeneity. We identified two candidate “mixed” models: (I) \( q_s = \frac{1}{2} \), together with \( \hat{p}_I \) for \( Ff \) and \( p_I = \frac{2}{3} \) for \( Mm + Mf + Fm \), (II) a jointly estimated \( \hat{q}_s \), together with \( \hat{p}_I \) for \( Ff \) and \( p_I = \frac{1}{2} \) for \( Mm + Mf + Fm \). In any case the kind of meeting, here an indicator of membership in data set \( Ff \), was the only \( x \)-variable needed (of those tested here). Learning (cf. step 1) was examined by the number \( X_1 \) of
Is outcomes in the first half of trials. With female \( N_2 \) 9 excluded there was no significant result. With \( N_2 \) 9 included, for female \( N_2 \) 5 we observed the proportions \( \hat{q} = 1 \) in the first and \( \hat{q} = 0 \) in the second half (\( P = 0.0079 \), two-tailed test). For the learning tests as a whole, in step 1 and here, on the 5\% level of significance we expected to find \((34 + 34) \cdot 0.05 = 3.4 \) significant results, and we found just 1 (and only if \( N_2 \) 9 was included). Thus, once again we failed to reject the hypothesis of no learning effects.

As for homogeneity (cf. step 2), the individuals within each of the female and male+mixed sets were tested by the proportion of Is-outcomes for each newt. The null hypothesis of common Is-probabilities was accepted both for \( Ff \) (\( P = 0.41 \)) and \( Mm + Mf + Fm \) (\( P = 0.80 \)). Thus, out of totally 5 homogeneity tests in step 2 and here, we found 1 significant result (for \( I \) vs. \( O \), only with \( N_2 \) 9 included in set \( Ff \)) – which is more than expected \((5 \cdot 0.05 = 0.25) \) but not extreme (\( P = 0.23 \)). We leave the question open whether the results for female \( N_2 \) 9 are purely random or point at true behavioural variation.

As for the best model, the two candidate models (I)-(II) above were judged by a goodness-of-fit test. The non-interaction model (I) was rejected (\( P = 0.012 \)), whereas the interaction model (II) was accepted (\( P = 0.083 \)). In summary, our analysis favours the unified interaction model (II), with an estimated sharing-shelter probability \( \hat{q} = 0.65, 95\%\)-confidence interval \([0.52, 0.76]\). The estimated in-shelter probability for \( Ff \) was \( \hat{\mu}_I = 0.91 \) with female \( N_2 \) 9 excluded (otherwise \( \hat{\mu}_I = 0.84 \)).

Discussion

In their choices between sharing and not sharing a shelter, newts behaved uniformly throughout all experimental trials. When both individuals in a paired trial occupied a shelter, they co-occupied one of the shelters in about two thirds of all cases, and the response was different from a random distribution of individuals. Shelter use in favour of aggregation was thus predominant in all sets of great crested newts tested here. However, males and females differed in their shelter-use response. Females were more prone to reside in shelters when meeting other females than when meeting males, whereas males appeared indifferent to the sex of the other individual. The difference in co-occupancy rates was entirely attributed to the predominance of shelter use among females meeting other females. Although sex effects were clearly noted, we could not demonstrate that size-related characteristics affected the outcome of any social interactions.

The extent of research on the ecology and behaviour of pond-breeding salamanders during the terrestrial phase is scarce despite the high importance of the terrestrial habitat (Dodd and Cade, 1998; Semlitsch, 1998; Malmgren, 2001, 2002; Schabetsberger et al., 2004). However, conducting terrestrial studies on social interactions among newts in the field is difficult since they are secretive, largely burrowing, and nocturnal (e.g., Griffiths, 1996). Results from experimental studies under laboratory conditions can therefore generate hypotheses on how individuals behave in field situations. In the case of the great crested newt, or for almost any other non-plethodontid or ambystomatid salamander, background data on what to expect in the terrestrial phase is practically non-existent.

A notable result from this study is the verification that shelters are important to great crested newts, which is consistent with results from studies on pond-breeding ambystomatid salamanders (Madison, 1997; Regosin et al., 2004). It also confirms observations on the great crested newt from telemetric field studies (e.g., Jehle, 2000; Malmgren et al., unpubl. data), suggesting that the different kinds of behaviours we recorded in the experimental setting are similar to those expressed in the field. Further, our choice of experimental design was validated by the fact that newts involved in experiments did not demonstrate systematically changed behaviour during the course of their respective trial.
series. This, together with the result that individuals in each test set reacted to the experimental setup similarly, indicates that behavioural responses to other individuals override responses to the potential influence from an artificial environment or from past events.

At least theoretically, our results support the hypothesis that shelters have the potential to be resources worth defending for great crested newts. However, if shelter availability per se is not limiting, mutual exclusion may not be a likely outcome, and in our study each arena had two shelters that provided ample space for both individuals to either reside alone or to share a common shelter. Contrary to results obtained in similarly staged experiments on ambystomatid salamanders (e.g., Regosin et al., 2004), we found that great crested newts co-occupied shelters more frequently than expected by chance. For example, Ducey and Ritsema (1988) gave spotted salamanders (Ambystoma maculatum) a choice of multiple shelters, which resulted in lower co-occupancy rates due to the tendency for individuals to avoid or exclude conspecifics. In our study, the relative frequency of co-occupancy in shelters among females was the same as for the other sets of newts, combined with the almost extreme preference for remaining in shelter (91%) when two females met.

The decision model that best fits our data was a two-stage model, in which 1) an individual first chose between residing inside or outside the shelters (F1 mostly in shelters), and then 2) given that both newts had chosen to reside inside a shelter, in two out of three trials newts end up sharing shelters. This suggests that females display a form of passive resource protection when faced with another female in a shelter. We assume that they do not express agonistic behaviour, since this would result in mutual exclusion rather than shelter co-occupancy. This would make sense because females are more likely to be similar to other females than males in their demands on the micro-habitat. Thus, competitive interactions among newts may be context-dependent, as have been suggested for ambystomatid salamanders (Ducey and Heuer, 1991; Regosin et al., 2004). Since newts tend to use natural crevices and holes under rocks or logs, or burrows made by other animals (e.g., Jehle, 2000; Jehle and Arntzen, 2000), they appear to have little ability to create their own shelters. This is consistent with Semlitsch’s (1983) observations of ambystomatid salamanders, and implies that there may be situations when the supply of shelters is limited. However, based on our results, it seems unlikely that great crested newts would respond to such conditions with increasing aggressiveness, such as in many Ambystoma species (Ducey and Heuer, 1991), but this could be a hypothesis to test in future studies (see, e.g., Regosin et al., 2004) using for example direct observation and video recording.

In this paper, we have introduced a fairly novel method of dealing with interactions among individuals, namely by matching each individual against all others in a group using a tournament-approach. Further, we have used statistical methods based on exact tests and Monte Carlo simulations to deal with this type of data, which have enabled us to resolve several aspects not normally available in similar experiments or in standard statistical packages. Although admittedly complex in its structure, the methods proposed here may be useful for future studies in the behavioural sciences where the complexity of interactions are of interest, or where the availability of individuals to sample is limited, for example when there are ethical aspects to consider or when studying threatened species. A sound combination of methods can thus contribute to reveal stages of the life cycle in cryptic species that need to be exposed. Such knowledge may be particularly important for evaluating conservation and management practices where crucial aspects of a species ecology is poorly known.

Acknowledgements. We thank Robert Jehle, and several anonymous referees, for valuable comments on early versions of the manuscript, and Daniel Gustafson, Roger West-
ermark and Cecilia Journath-Pettersson for field assistance. The study was performed at laboratory facilities provided by the Department of Natural Sciences at Örebro University, Sweden. Newts were collected, handled and kept in captivity with permits acquired from the Swedish Environmental Protection Agency, the Swedish Board of Agriculture, and the Swedish Board for Laboratory Animals (Tierp, C 0/84).

References


3a) **Pure in-out-shelter choice**

\[ p_1 := \text{probability of in-shelter outcome for a newt.} \]
\[ n, X := \text{sample size and number of in-shelter outcomes (s.v.), respectively.} \]

Under \( H_0 \): \( p_1 = \frac{1}{2} \) (\( p_1 = \frac{1}{2} \)). Resampling condition: none.

Under \( H_0 \): \( X \in \text{Bin}(n, p_1) \).

\( p_{1,i} := \text{probability of in-shelter outcome for each newt in test group } i = 1, 2 \).

\( n_i := \text{sample size and number of in-shelter outcomes (s.v.), respectively, in group } i = 1, 2 \).

\( n := n_1 + n_2 \) (total sample size), \( x := x_1 + x_2 \) (total number of observed in-shelter outcomes).

Under \( H_0 \): \( p_{1,i} = p_{1,2} \). Resampling condition: \( X_1 + X_2 = x \).

Under \( H_0 \):

\[ p^*_1 := \frac{X_1 + X_2}{n} = \frac{x}{n} \]
\[ Z := \frac{p^*_1 - p^*_2}{\sqrt{p^*_1(1-p^*_1)/(1/n_1 + 1/n_2)}} \in A X(0, 1). \]

\( p_2 := \text{probability of in-shelter outcome for a newt.} \)

\( p_j := \text{probability of experiment outcome } j = 1 (I), j = 2 (O), j = 3 (O) \).

\( n, X_j := \text{trial sample size and number of outcomes } j \) (s.v.), respectively.

\( H_0 \): independent choices, \( p_1 = \frac{1}{2} \) \( (p_1 = \frac{1}{2}) \). Resampling condition: none.

Under \( H_0 \):

\[ p_1 = p^*_1, \quad p_2 = 2p_1(1-p_1), \quad p_3 = (1 - p_1)^2, \]
\[ Y := \sum_{j=1}^{3} \frac{(X_j - np_j)^2}{np_j} \in A X^2(2). \]

\( x_1 := x_1 + \frac{1}{2} x_2 \) (total number of observed in-shelter outcomes).

Under \( H_0 \): independent choices, unknown \( p_1 \). Resampling condition: \( X_j = x_j \).

Under \( H_0 \):

\[ p_1 = p^*_1, \text{ etc..} \]
\[ p^*_j := \frac{X_j}{n}, \]
\[ Y := \sum_{j=1}^{3} \frac{(X_j - np_j)^2}{np_j} \in A X^2(1). \]

3b) **Pure interaction choice**

\( q_s := \text{probability of sharing shelter outcome, given that both newts are in-shelter.} \)
\( n, X := \text{sample size and number of sharing-shelter outcomes (stochastic variables), respectively.} \)

\( H_0 \): \( q_s = \frac{1}{2} \). Resampling condition: none.

Under \( H_0 \): \( X \in \text{Bin}(n, q_s) \).

\( q_{s,i} := \text{prob. of sharing shelter outcome in test group } i = 1, 2 \), given that both newts are in-shelter.

\( n_i, X_i := \text{sample size and number of in-shelter outcomes (s.v.), respectively, in group } i = 1, 2 \).

\( n := n_1 + n_2 \) (total sample size), \( x := x_1 + x_2 \) (total number of observed sharing outcomes).
Terrestrial interactions in great crested newts

\( H_0 : q_s,1 = q_s,2. \) Resampling condition: \( X_1 + X_2 = x. \)
Under \( H_0 \):

\[
q^*_s := \frac{X_1 + X_2}{n} = \frac{x}{n},
\]

\[
Z := \frac{q^*_s - q^*_s,2}{\sqrt{\frac{1}{n_1} - \frac{1}{n_2}}} \in A \chi^2(N(0, 1)).
\]

3c) Combined choice

\( p_j := \text{probability of in-shelter outcome for a newt,} \)
\( p_j := \text{probability of experiment outcome} \)
\( j = 1 (\text{in} / \text{shelter}), \)
\( j = 2 (\text{out} / \text{shelter}), \)
\( j = 3 (\text{in} / \text{shelter}), \)
\( j = 4 (\text{out} / \text{shelter}), \)

\( n : \) trial sample size and number of outcomes \( j \) (s.v.), respectively.
\( H_0 : \) independent choices, \( p_1 = \frac{1}{2} (p_1 = \frac{1}{2}). \) Resampling condition: none.
Under \( H_0 \):

\[
p_1 = p_2 = p_3/2, \quad p_3 = 2p_1(1-p_1),
\]

\[
Y := \sum_{j=1}^{4} (X_j - np_j)^2 \in A \chi^2(2).
\]

\( x_j := x_1 + x_2 + \cdots + x_3 \) (total number of observed in-shelter outcomes).
\( H_0 : \) independent choices, unknown \( p_j. \) Resampling condition: \( X_j = x_j. \)
Under \( H_0 \):

\[
p_1 = p_2 = p_3^2/2, \quad p_3 = 2p_1(1-p_1),
\]

\[
Y := \sum_{j=1}^{4} (X_j - np_j)^2 \in A \chi^2(2).
\]

4) Dependence on size

\( N_{ij} := \text{number of outcomes (s.v.) for} \)
\( i = 1 \) (in) or (sharing),
\( i = 2 \) (out) or (not sharing) and \( j = 1 \) (large difference in SVL-based value), \( j = 2 \) (small difference in SVL-based value),

\( n_{ij} := \text{number of observed outcomes (i,j),} \)
\( H_0 : \) independence of size. Resampling condition:

\[
Y := \sum_{j=1}^{4} (X_j - np_j)^2 \in A \chi^2(2).
\]

\( u_{ij} := \text{given function of SVL for the two newts in experiment} \)
\( k \) of outcomes both in-shelter.
\( p_k := 1/1 + e^{-\mu_0 - \mu_j} = \text{probability of sharing shelter outcome in experiment} \)
\( \mu_0 \) and \( k. \)

\( Y_k := \text{dichotomous stochastic variable in experiment} \)
\( \mu_0 \) and \( k, \)
\( Y_k := 1 (\text{sharing}), \)
\( Y_k := 0 (\text{non-sharing}). \)
\( n := \text{sample size,} \)
\( n := \sum_{k=1}^{n} Y_k \) (number of observed sharing shelter outcomes).
\( D := N_1 \ln (N_1 + (n-N_1)) / n - n \ln n \) (log-likelihood function for model with \( b = 0). \)
\( H_0: b = 0. \) Resampling condition: \( \sum_{k=1}^{n} Y_k = n \).
Under \( H_0 \): ML-estimators \( a^*, b^*, D = d, \)

\[
G := 2 \left( \sum_{k} (Y_k \ln p_k^e + (1 - Y_k) \ln(1 - p_k^e)) - D \right)
\]

\( \in A \chi^2(2). \)

5) Choice of decision model

As for statistics \( X_1 \) and \( V, \) respectively, see steps 1 and 2 above.
\( p_{1,i} := \text{probability of in-shelter outcome for a newt in set} \)
\( i = 1 (F_1), i = 2 (Mn + Mf + Fm), \)
\( p_{j,i} := \text{probability of experiment outcome} \)
\( j = 1 (I_1 / I), \)
\( j = 2 (F_1 / I), \)
\( j = 3 (I_1 / O), \)
\( j = 4 (O_1 / O), \)

\( n_i, X_{j,i} := \text{sample size and number of outcomes (s.v.),} \)

respectively, in new set \( i. \)
\( x_{1,i} := x_1 + x_2 + \cdots + x_3 \) (total number of observed in-shelter outcomes in set \( i). \)
\( H_0 : \) unknown \( p_{1,1} \). \( p_{1,2} = \frac{2}{2}, \) \( q_i = \frac{1}{2}. \) Resampling condition: \( X_{1,1} = X_{1,1}. \)
Under \( H_0 \):

\[
p_{11} = q_1 \cdot p_{1,1}^2, \quad p_{21} = (1 - q_1) \cdot p_{1,1}^2,
\]

\[
Y := \sum_{j=1}^{4} (X_j - np_j)^2 \in A \chi^2(2).
\]

\( H_0 : \) unknown \( p_{1,1} \). \( p_{1,2} = \frac{2}{2}, \) unknown \( q_i. \) Resampling condition: \( X_{1,1} = X_{1,1}. \)
Under \( H_0 \):

\[
p_{11} = q_1 \cdot p_{1,1}^2, \quad p_{21} = (1 - q_1) \cdot p_{1,1}^2,
\]

\[
ML-estimator \ q^*_i := \frac{X_{11} + X_{12} + X_{21} + X_{22}}{n_1}.
\]

\[
Y := \sum_{j=1}^{4} (X_j - [n_1 p_{1,j}^e + n_2 p_{2,j}])^2 \in A \chi^2(2).
\]

Here we applied semi-parametric resampling: While keeping the number of in-shelter outcomes fixed in the simulations, we generated estimates \( q_i \) in the random samples by the same principle as in our observed sample, i.e. as the proportion of \( I / I + O / O \) outcomes out of all \( F / I \) outcomes. The corresponding estimator \( q^*_i \) can be shown to minimize the chi-square sum and to maximise the likelihood function.