

MECHANISMS OF SPERM COMPETITION: TESTING THE FAIR RAFFLE

BRYAN D. NEFF^{1,2} AND LINDI M. WAHL^{3,4}

¹Department of Biology, University of Western Ontario, London, Ontario N6A 5B7, Canada

²E-mail: bneff@uwo.ca

³Department of Applied Mathematics, University of Western Ontario, London, Ontario N6A 5B7, Canada

⁴E-mail: lwahl@uwo.ca

Abstract.—Sperm competition is a major force of sexual selection, but its implications for mating system and life-history evolution are just beginning to be understood. Of particular importance is understanding the mechanisms of sperm competition. Models have been developed to determine if sperm competition operates in a fair raffle process, whereby each sperm from competing males has an equal chance of fertilizing a female's ova, or if it operates in a loaded raffle process, whereby one male's sperm has a fertilization advantage. These models require data on relative sperm and offspring (paternity) numbers of competing males. Here we develop a model based on maximum-likelihood methods for differentiating between the fair and loaded raffle processes. The model calculates the relative competitiveness of two males' sperm (loadings) as well as the economy of scale (nonlinear returns to sperm number). Previous models implicitly assumed that there is no economy of scale, which may not be the case when there is cooperation or interference among sperm from a given male. We demonstrate that our model has superior power—in some instances more than double—than previous models. We apply our model to an example of sperm competition in the guppy (*Poecilia reticulata*) and show that the system may be characterized by a loaded raffle attributable to effects of second male precedence.

Key words.—Guppy, paternity, polyandry, precedence, P_2 .

Received March 2, 2004. Accepted May 4, 2004.

Sperm competition is a major force of sexual selection (Birkhead and Møller 1998; Birkhead and Pizzari 2002). Parker (1998) defines sperm competition as the competition between the sperm from two or more males for the fertilization of a given set of ova. Sperm competition is widespread in nature and because the mechanism of competition varies across mating systems, mating success typically is a poor indicator of reproductive success. Thus, there has been a growing interest in understanding the mechanisms of sperm competition in mating systems (Birkhead and Møller 1998; Wedell et al. 2002).

A fundamental mechanism in sperm competition is called the raffle (Parker 1990, 1998), whereby a male's probability of fertilizing a female's ova is related to the proportion of all sperm provided to the female that is his. A fair raffle occurs when this probability is simply equal to each male's proportion of sperm. A loaded raffle occurs when one male's sperm has a competitive advantage. Such an advantage may come from, for example, sperm displacement, whereby one male's sperm is displaced from the female by a more recent mating partner; sperm ejection, whereby a more recent mating partner's sperm is ejected by the female; or differences in sperm quality (Pizzari and Birkhead 2000; Simmons 2001; Neff et al. 2003).

Parker et al. (1990) develop a linear model designed to differentiate between the processes of the fair and loaded raffles (for two competing males). The underlying assumption of this model is that

$$P_2 = \frac{N_2}{N_1 + N_2} = \frac{rS_2}{S_1 + rS_2}, \quad (1)$$

where P_2 is defined as the paternity of the second male, N_1 and N_2 are the number of offspring sired by male 1 and male 2, respectively, and S_1 and S_2 are the numbers of sperm transferred to the female by the first and second male. Equation

(1) does not allow for the possibility of sperm limitation; that is, ejaculates are assumed to contain enough sperm for complete fertilization in the absence of competition. This assumption has implications from a game theory perspective (Mesterton-Gibbons 1999).

The parameter r in equation (1) is a measure of the loading given to the second male's sperm relative to the first male's sperm. Note that Parker et al. (1990) originally defined r in reference to the first male's sperm relative to the second male's sperm. We elected to redefine r here in reference to the second male because it is a more natural fit with our focus on the paternity of the second male (P_2 ; also see Eggert et al. 2003). When $r > 1$, each sperm from the second male is worth more than that of the first male's. For example, when $r = 2$ each sperm from the second male is twice as likely to fertilize the female's ova than the first male's sperm. This may be true when the second male has competitively superior sperm, possibly attributed to increased swimming speed or longevity, or when the female is twice as likely to eject or otherwise incapacitate a sperm from the first male than the second male. Alternatively, when $r < 1$ the first male has competitively superior sperm. For a fair raffle, $r = 1$.

Inverting equation (1) and rearranging yields

$$\frac{1}{P_2} = \frac{1}{r} \frac{S_1}{S_2} + 1. \quad (2)$$

Parker et al. (1990) therefore suggest plotting $1/P_2$ (y-variable) versus S_1/S_2 (x-variable), and performing a least-squares linear regression fit to these data. In a fair raffle, the data should be consistent with a slope of one (i.e., $r = 1$) and an intercept of one.

Analysis by Eggert et al. (2003), however, shows that plotting $1/P_2$ and S_1/S_2 results in nonnormally distributed data, which violates the assumption of the least-squares linear regression. This violation can lead to erroneous analysis of sperm loading. From equation (1), Eggert et al. derived

$$\log\left(\frac{N_2}{N_1}\right) = \log\left(\frac{S_2}{S_1}\right) + \log(r) \quad (3)$$

(note that $r = p_1/p_2$ in Eggert et al. 2003). A least-squares linear regression of $\log(N_2/N_1)$ (y-variable) versus $\log(S_2/S_1)$ (x-variable) should yield a slope of one and a y-intercept of zero if the raffle is fair.

By using logarithms, the method of Eggert et al. restores normality in the proportional data (N_2/N_1 and S_2/S_1) and therefore does not violate the assumption of the linear regression. They demonstrated their method with data from a study of the decorated cricket (*Gryllodes sigillatus*) and initially showed that based on equation (2) the fair raffle was rejected, although the data appear skewed (see fig. 1b in Eggert et al. 2003). When equation (3) instead was used, the data appear less skewed and a fair raffle was not rejected (see fig. 1c in Eggert et al. 2003).

A shortcoming of the method of Eggert et al. (2003) is that it requires that N_1 and N_2 be nonzero. However, because the number of ova that two males' sperm compete for are finite, it is possible that N_1 or N_2 can be zero due to chance. This is especially probable when the total number of ova is small—for example, when there is only one ovum either N_1 or N_2 must be zero, or when the relative number of sperm given to a female from the males is highly skewed. When N_1 or N_2 is zero, the brood must be omitted from the analysis because in either case $\log(N_2/N_1)$ is undefined (i.e., either negative or positive infinity).

In this paper, we develop an alternative model based on maximum-likelihood methods. Our model enables all data to be used (i.e., even when N_1 or N_2 is zero) and can be easily generalized to analyze any function relating sperm numbers to fertilization success (P_2). For example, here we expand Parker et al.'s (1990) original equation to include the possibility that there is an economy of scale in sperm number, that is, nonlinear returns to sperm number (e.g., Moore et al. 2002). We show that our model has higher statistical power to detect deviations from the fair raffle than the model of Eggert et al. (2003). We demonstrate our model with a biological example of sperm competition in the guppy (*Poecilia reticulata*). The guppy is a small (about 2–3 cm in length) live-bearing fish with internal fertilization that inhabits tropical freshwater streams (Houde 1997). Females encounter males sequentially and often prefer males that have bright orange coloration, but mate multiply and gain only sperm from males. Females are able to store the sperm of males for several months, but there appears to be last-male precedence in fertilization success (Evans and Magurran 2001; Pitcher et al. 2003).

MATERIALS AND METHODS

The Model

Although the maximum-likelihood method we develop can be used with any equation relating sperm numbers to fertilization success, we illustrate the technique using a variation of Parker et al.'s (1990) raffle equation

$$\frac{N_1}{N_1 + N_2} = \frac{S_1^t}{S_1^t + rS_2^t}, \quad (4)$$

where t is a measure of the economy of scale to sperm number. When $t > 1$ there is an increasing economy of scale and a nonlinear return to sperm number. For example, when $t = 2$ and the first male has placed twice as much sperm in the female than the second male (and assuming $r = 1$), the first male will be four times more likely to fertilize any one of the female's ova and he will on average fertilize 80% of the ova. When $t < 1$ there is a decreasing economy of scale, as might be the case when there is interference or some other form of diminishing returns to sperm number. When $t = 0.5$ the first male would have to place 16 times as much sperm in the female as the second male to be four times more likely to fertilize her ova. When $t = 0$ there is no relationship between sperm number and fertilization success.

Equation (4) is an expectation and therefore does not explicitly deal with the chance variation that occurs in real biological systems. The simplest model of this variation is the binomial distribution, which assumes that offspring are sired independently with the right side of equation (4) giving the probability that a given offspring is sired by male 1. If $\Pr(N_1, N_2 | S_1, S_2, r, t)$ denotes the probability of observing exactly N_1 and N_2 , given that the experimental conditions were S_1, S_2, r , and t , we can write this value as

$$\Pr(N_1, N_2 | S_1, S_2, r, t) = \binom{N_1 + N_2}{N_1} \left(\frac{S_1^t}{S_1^t + rS_2^t}\right)^{N_1} \left(1 - \frac{S_1^t}{S_1^t + rS_2^t}\right)^{N_2}. \quad (5)$$

Given a set of B broods for which N_i and S_i are known, the probability of making this exact set of observations is given by

$$\Pr(\text{obs}) = \prod \Pr(N_1, N_2 | S_1, S_2, r, t), \quad (6)$$

where the product is over all broods. Therefore, if we can find the values of r and t that maximize equation (6), these values will be the maximum-likelihood estimates of the unknown parameters.

We developed an algorithm using the C++ programming language to determine the most likely value of r and t based on equation (6). The algorithm uses the downhill simplex method (Nelder and Mead 1965) presented in Press et al. (1992). This routine tries numerous values of r and t and, by evaluating both the value of $\Pr(\text{obs})$ and its gradient, is able to efficiently search the plane and find where $\Pr(\text{obs})$ reaches its maximum. Our algorithm also determines if r and t are significantly different from one (a fair raffle) and if t is significantly different from zero ($t = 0$ when fertilization success is independent of sperm number) as described more fully in the next section. We have made our algorithm available as an executable file downloadable at <http://publish.uwo.ca/~bneff/links.htm>.

Extending the approach of Eggert et al. (2003) to equation (4), it can be shown that

$$\log\left(\frac{N_2}{N_1}\right) = t \log\left(\frac{S_2}{S_1}\right) + \log(r). \quad (7)$$

Thus, using Eggert et al.'s least-squares linear regression of $\log(N_2/N_1)$ (y-variable) versus $\log(S_2/S_1)$ (x-variable) should yield a slope of t and a y-intercept of $\log(r)$.

Significance Testing and Power Analysis

To compare the statistical power of our algorithm to the model of Eggert et al. (2003), we performed several simulations. First, for our model we established 95% confidence intervals for r and t in the following way. A simulation generated B broods, each with a random number of offspring uniformly distributed between N_{min} and N_{max} . For each brood, sperm number was randomly assigned to each male as a uniformly distributed, real number between 1 and 10. Thus, S_1/S_2 could take on any value between 0.1 and 10 (the subsequent analysis was similar when S_1/S_2 was instead allowed to range between 0.5 and 2; data not shown). Then for each offspring within the brood, a uniformly distributed random number was generated between zero and one and if the number was less than $S_1/(S_1 + rS_2)$ with $r = t = 1$ (the null model) the offspring was assigned to the first male; otherwise the offspring was assigned to the second male. The process was repeated until all offspring were assigned in the B broods. The downhill simplex method was then used to estimate r and t and this was all repeated for a total of 100,000 datasets of B broods. The r and t estimates were then ranked from which the 2.5% and 97.5% values were extracted.

Next, we generated 10,000 pseudo-datasets based on the technique described above, but with a predetermined r and t . From these pseudo-datasets our method and that of Eggert et al. (2003) were used to estimate r and t , and these were compared to the expectation under the null hypothesis of a fair raffle (i.e., $r = t = 1$). We then calculated the proportion of the datasets in which the null hypothesis was rejected. For our model, the 95% confidence interval was used to reject or accept the null hypothesis, while for Eggert et al. (2003), t -tests using the standard errors established by the linear regression were used. In the case that the predetermined r and t were set to one, these proportions represent Type I error rate, otherwise they represent statistical power.

Biological Example

We used our algorithm and that of Eggert et al. (2003) to examine sperm competition in the guppy (for full details of the experiment see Pitcher et al. 2003). Briefly, females were sequentially mated to two males presumably differing in their genetic quality based on a sexual ornament (area of orange coloration). In the first set of trials ($B = 8$), females were first mated to a high-quality male followed by a low-quality male (“showy/plain” trials in Pitcher et al. 2003). In the second set of trials ($B = 9$), females were first mated to a low-quality male followed by a high-quality male (“plain/showy” trials in Pitcher et al. 2003). The sperm number transferred to the female was estimated based on each male’s sperm load, which was calculated by collecting and counting the spermatozeugmata from each male three days after they mated. This estimation method has been used by other researchers (e.g., Pilastro and Bisazza 1999; Evans and Magurran 2001). In each brood, paternity was unambiguously assigned using a Y-linked genetic marker that produced a visible pattern on the offspring. On average 8.5 ± 3.4 (SD; range = 4–16) offspring were analyzed per brood.

These data were used to calculate r and t for each set of trials and for all trials combined. For our model, r and t were

TABLE 1. Summary of the power analysis for our model and Eggert et al. (2003). Parameters used for the simulation comprise the actual r and t , the number of broods analyzed, and the number of offspring per brood (range = N_{min} – N_{max}). Power is the proportion of datasets in which the null hypothesis ($r = 1$ or $t = 1$) was rejected and the mean number of usable broods refers to Eggert et al. When $r = t = 1$, “power” refers to Type I error rate.

Broods	Offspring	Power: our model		Power: Eggert et al.		Usable broods
		r	t	r	t	
$r = 1, t = 1$						
5	5–10	.051	.048	.050	.051	4.6
10	5–10	.050	.050	.050	.050	9.2
20	5–10	.051	.049	.053	.060	18.3
40	5–10	.050	.049	.049	.075	36.7
5	10–20	.052	.050	.048	.060	4.9
10	10–20	.051	.055	.049	.063	9.8
20	10–20	.051	.052	.051	.064	19.7
40	10–20	.054	.047	.051	.067	39.4
$r = 2, t = 1$						
5	5–10	.336	.058	.174	.049	4.3
10	5–10	.724	.060	.490	.062	8.7
20	5–10	.965	.059	.849	.090	17.5
40	5–10	1.00	.058	.992	.156	35.0
5	10–20	.631	.062	.338	.051	4.8
10	10–20	.958	.064	.857	.054	9.6
20	10–20	1.00	.064	.996	.059	19.3
40	10–20	1.00	.058	1.00	.051	38.6
$r = 1, t = 2$						
5	5–10	.107	.218	.049	.074	3.7
10	5–10	.091	.568	.051	.136	7.6
20	5–10	.085	.898	.049	.296	15.3
40	5–10	.080	.997	.053	.862	30.6
5	10–20	.089	.411	.047	.203	4.4
10	10–20	.083	.875	.048	.598	8.8
20	10–20	.085	.997	.048	.939	17.6
40	10–20	.083	1.00	.053	1.00	35.2
$r = 2, t = 2$						
5	5–10	.381	.220	.130	.066	3.5
10	5–10	.701	.557	.330	.131	7.4
20	5–10	.944	.901	.672	.271	14.7
40	5–10	.999	.998	.949	.540	29.5
5	10–20	.637	.418	.240	.189	4.3
10	10–20	.937	.873	.689	.552	8.6
20	10–20	.999	.997	.962	.898	17.1
40	10–20	1.00	1.00	1.00	.998	34.4

compared to a 95% confidence interval generated as described above with $r = t = 1$ and with the actual offspring and sperm numbers used in the experiment. Estimates of P -values were established by comparing the calculated r and t to the ranking of the 100,000 values generated for $r = t = 1$. We also compared the calculated t to a second 95% confidence interval generated with $t = 0$ and $r = 1$. When $t = 0$ fertilization success is independent of sperm number and each offspring is assigned randomly to either male (with equal probability). We term this the “sperm independent model.”

RESULTS

Significance Testing and Power Analysis

When r and t were set to one for the pseudodata generation, we were able to assess the Type I error rate (Table 1). Across the parameter sets that we examined, a t -test revealed that

the Type I error rate did not differ from the expected 0.05 for our model (r : $t_7 = 1.34$, $P = 0.22$; t : $t_7 = 1.40$, $P = 0.20$). For Eggert et al. (2003) there was no deviation for r ($t_7 = 0.28$, $P = 0.79$), but t had a marginally higher rate than expected ($t_7 = 3.87$, $P = 0.006$). To investigate further, we considered the case of $r = t = 1$, $B = 20$ and 5–10 offspring per brood in greater detail. For this case, the estimated values of r and t were $r = 1.02 \pm 0.19$ (SD) and $t = 1.03 \pm 0.27$ by our method and for Eggert et al. (2003) they were $r = 1.02 \pm 0.19$ and $t = 0.89 \pm 0.24$. Thus, both methods overestimated r by about 2%, while our method overestimated t by 3% and Eggert et al. underestimated t by 11%.

When either r or t was set to 2.0 for the data generation, we were able to assess the statistical power (Table 1). Increasing either the number of offspring analyzed per brood or increasing the number of broods increased power. In all cases, our algorithm outperformed Eggert et al. (2003) leading to significantly higher power (paired t -test r : $t_{15} = 4.63$, $P < 0.001$; t : $t_{15} = 5.22$, $P < 0.001$). This may partly be attributable to the fact that Eggert et al. could not always use all broods (i.e., when either N_1 or N_2 equaled zero). About 8% of broods could not be used in the analysis when each brood contained between 5–10 offspring and about 2% could not be used when each brood contained between 10–20 offspring. Consequently, the greatest difference in power was evident when fewer offspring were analyzed per brood (5–10 vs. 10–20) and when fewer broods were analyzed (5 vs. 10 or 20).

Our model performed least well when $r = 1$ and $t = 2$, suffering from a slightly inflated Type I error rate for the estimate of r . For our model the rate was about 8–10% (instead of the expected 5%). The model of Eggert et al. (2003) did not appear to have an elevated Type I error rate for r (see Table 1).

Figure 1 presents the power analysis for our model when either r or t equals 2.0 for brood sizes between 5–10 and 10–20. As an example, when $r = 2$ and $t = 1$ about 12 broods were required to obtain 80% power when 5–10 offspring were analyzed per brood. Thus, 12 broods would be required to correctly reject the null hypothesis ($r = t = 1$) 80% of the time. Seven broods were required to obtain the same power when 10–20 offspring were analyzed per brood; when $r = 1$ and $t = 2$ about 9 and 16 broods were required for the two respective brood size ranges.

Biological Example

In the guppy, for the showy/plain trials ($B = 8$) our model revealed that $r = 1.41$, which was not significantly different from one ($P = 0.153$), and $t = 0.30$, which was not significantly different from one ($P = 0.085$). However, the sperm independent model ($t = 0$) could not be rejected ($P = 0.284$). For the plain/showy trials ($B = 9$) our model revealed that $r = 3.30$, which was significantly different from one ($P < 0.0001$), and $t = 0.15$, which was significantly different from one ($P = 0.0045$). However, the sperm independent model again could not be rejected ($P = 0.316$).

When the trials were combined and analyzed (to increase power), the model estimated $r = 2.15$, which was significantly different from one ($P = 0.0001$), and $t = 0.34$, which

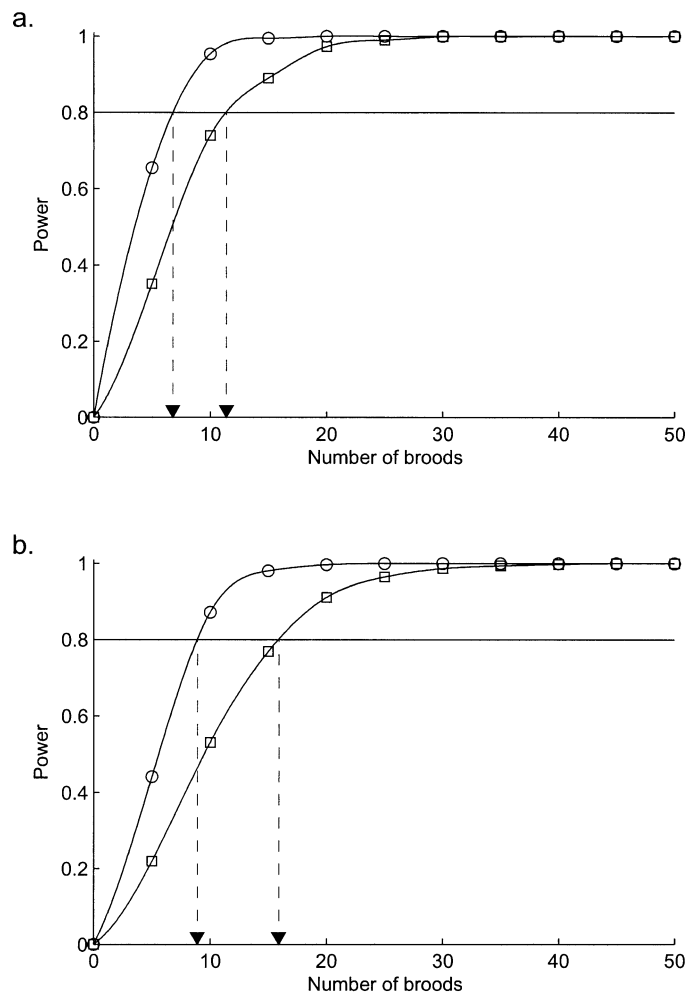


FIG. 1. The relationship between the number of broods and power based on our model for (a) $r = 2$ and $t = 1$ and (b) $r = 1$ and $t = 2$. The lower line (squares) in each panel represents an offspring range per brood of 5–10 and the upper line (circles) represents an offspring range of 10–20. The solid horizontal line denotes 80% power and the arrows denote the number of broods required to obtain this level of power for each line.

was significantly different from one ($P = 0.0052$). In this case the sperm independent model did a poor job explaining the data and the P -value was close to the level required to reject the model ($P = 0.086$).

Using Eggert et al. (2003) and all trials combined, t was estimated to be 0.18 (95% CI: -0.76 to 1.12), which was not significantly different from zero or one ($P > 0.05$ for both). The parameter r was estimated to be 1.98 (95% CI: 1.04 to 3.77), which was significantly different from one ($P = 0.040$). Only 12 of the 17 broods could be used in this analysis because either N_1 or N_2 equaled zero for the other five.

DISCUSSION

We developed a model based on maximum-likelihood methods to calculate the loading and economy of scale associated with sperm number when two males compete for fertilizations. Our model provides more precise estimates as

well as superior power for detecting statistical deviations from a fair raffle than previous models. For example, when brood number is low or when few offspring are analyzed per brood, our model can increase power by more than two-fold (see Table 1). The increased power is in part attributed to the fact that our model can use broods in which one of the two males does not fertilize any of the offspring (i.e., when N_1 or N_2 equals zero). Such broods can be influential in the analysis because they represent more extreme datapoints that are especially informative. Indeed, when $r = t = 1$ and 20 broods were analyzed with 5–10 offspring per brood, about 10% of broods could not be used in the Eggert et al. (2003) approach. An important point to note is that the broods omitted from the Eggert et al. method are not uniformly distributed across the range of the measured data, but tend to occur when S_1/S_2 is either small or large. This effect leads to a systematic underestimation of t and this bias can be large (e.g., 11%; see Results).

To ensure that the increased power of our model was not exclusively a result of being able to use all broods, we performed an additional analysis with five broods, 20 offspring per brood, a relative sperm number (S_2/S_1) ranging from 0.5 to 2.0, and $r = 2$, $t = 1$. In this case, all broods were used by both methods and the power for our method still was considerably higher at 0.842 than Eggert et al. (2003) at 0.507. We also considered adding a small value δ to N_1 and N_2 for the Eggert et al. method, so that all broods could always be used. When N_1 and N_2 are large, this small value will have negligible effect. When either N_1 or N_2 is zero, however, this method replaces the undefined point on the plot with a point that will have the original value of $\log(S_2/S_1)$ on the x -axis, but will have an arbitrary value on the y -axis. When $N_2 = 0$ and δ is a very small value, $\log[(N_2 + \delta)/(N_1 + \delta)] = \log(\delta) - \log(N_1 + \delta) \approx \log(\delta)$, a large negative number. Similarly, when $N_1 = 0$, $\log[(N_2 + \delta)/(N_1 + \delta)] \approx -\log(\delta)$, a large positive number. This means that the point containing the zero has not been salvaged, but has instead been replaced by an arbitrary point on the plot. We illustrate this effect for two values of δ in Figure 2.

Our model appeared to have a slightly elevated Type I error rate for r when $t = 2$ (8–10%; see Table 1). However, this was resolved when the null model used to generate the expected 95% confidence interval for r was run with $t = 2$ instead of $t = 1$ (data not shown). Thus, researchers may wish to run the null model with either r or t set to the value calculated from their data in hand. Our executable program allows any r and t to be specified for the null model. Thus, researchers initially can run the program with $r = t = 1$ for the null distribution to generate the estimated r and t based on the data in hand; then they can rerun the program with $r = 1$ and t set to the value calculated from their data to test for $r \neq 1$; and a third time with $t = 1$ and r set to the value calculated from their data to test for $t \neq 1$.

Our model revealed several interesting aspects of sperm competition in the guppy. When all trials were analyzed simultaneously, we found that there was second male precedence ($r = 2.15$) and diminishing returns to sperm number ($t = 0.34$). However, we could not convincingly reject that fertilization success was independent of sperm number altogether ($P = 0.086$). Power analysis assuming $r = 2.15$ and

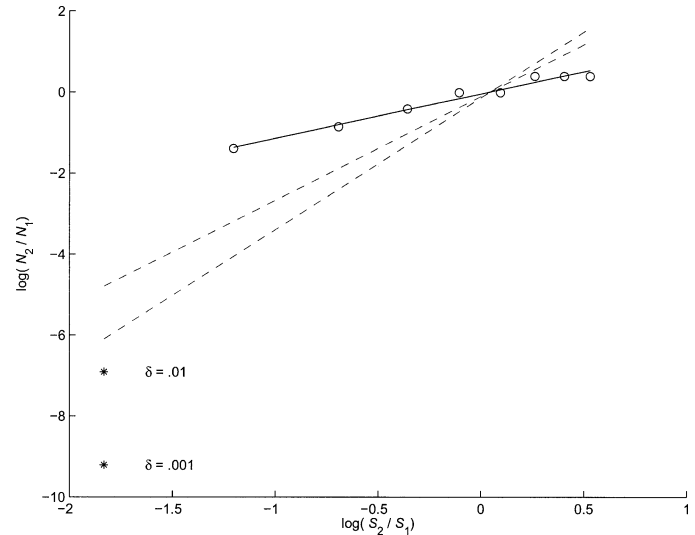


FIG. 2. The effect of adding a small value δ to N_1 and N_2 , to correct for either N_1 or $N_2 = 0$. We simulated nine broods, with 10 offspring analyzed per brood. S_2 varied between 0.8 and 8.5, while $S_1 = 5$. N_2 varied between zero and 6, with the zero occurring when S_2/S_1 was smallest. Circles show the eight broods in which $N_2 > 0$, and the solid line gives the linear regression through these data. By adding either $\delta = 0.01$ or $\delta = 0.001$ to both N_1 and N_2 , a new point (star) is added to the plot. Dotted lines illustrate the regression that would be obtained including all nine broods using these values of δ . Each value of δ adds an arbitrary point to the plot to which the regression may be extremely sensitive.

$t = 0.34$, 17 broods and four to 16 offspring per brood revealed that our analysis had power of only 0.355 to detect a difference from $t = 0$. Thus, we cannot confidently accept the null hypothesis either. If t is in fact greater than zero, then our model indicates that sperm from the second male has about twice the probability of fertilizing a female's ova as sperm from the first male. Furthermore, there are diminishing returns to sperm number (i.e., $t < 1$), suggesting that either male does not benefit much from placing significantly more sperm in the female than his rival.

We were unable to detect an effect of male quality on the sperm competitiveness even though previous work has revealed that higher-quality males have more competitive sperm. Evans et al. (2003) conducted artificial insemination experiments using equal numbers of sperm from pairs of males. When broods were born they used genetic markers to assign paternity to each male, revealing that males with more orange (our measure of quality) had significantly higher fertilization success. Because the females did not get to see or otherwise evaluate the males used to artificially inseminate them, the data suggest that higher-quality males have more competitive sperm. The apparent discrepancy between these results and ours may be attributable to differences between the studies in either statistical power or order effects. When we analyzed the trials separately (i.e., showy/plain vs. plain/showy) our sample was approximately halved. Analysis revealed that this would reduce power to only 0.152 to detect a significant deviation from $t = 0$ (assuming $r = 2.15$, $t = 0.34$, and eight broods with four to 16 offspring per brood). Alternatively, Evans et al. (2003) removed the order effect by first combining equal numbers of sperm from high- and

low-quality males. In our case, the order effect might mask any effect of male quality on competitiveness.

There are two general mechanisms that could explain the male order effect in the guppy. First, females may bias fertilization success in favor of second males. This may be accomplished by increasing the efficiency of sperm transfer through increased receptivity, thereby ensuring a greater proportion of the second male's ejaculate enters the female's gonoduct (Pilastro and Bisazza 1999; Evans and Magurran 2001; Pilastro et al. 2002). Females also may bias the use of sperm that enter their ovaries. Complex structures that store sperm such as epithelial folds have been identified and these structures may be involved in female manipulation of sperm use (Kadow 1954; Kobayashi and Iwamatsu 2002). Furthermore, females may even eject unwanted sperm from the gonoduct (Kadow 1954; also see Pizzari and Birkhead 2000). Second, males that mate last (second male in our experiment) may be able to displace or otherwise incapacitate the sperm of previous sires (first male in our experiment). However, such a mechanism for second male precedence in the guppy has not been described.

In conclusion, we hope that our algorithm will help researchers effectively address sperm competition mechanisms. Specifically, the algorithm enables researchers to estimate both r and t from any number of broods and determine whether these parameters differ from those expected by the fair raffle or the sperm independent model. In addition, our maximum-likelihood approach can be easily generalized to other models of sperm competition. Once the equation for expected paternity has been derived (eq. 4), this relationship can be used directly in equations (5) and (6) to find the maximum-likelihood estimators of any unknown parameters. This obviates the need to derive a means of parameter estimation (such as linear regression of log ratios) for each new model. We hope that this general aspect of our approach will prove useful as the mechanisms underlying sperm competition become better understood and improved modeling approaches are suggested.

ACKNOWLEDGMENTS

We thank T. E. Pitcher for the guppy data and M. A. Ball, T. Day, M. Mesterton-Gibbons, and S. K. Sakaluk for comments on the manuscript. This work was supported by NSERC of Canada, Canadian Foundation for Innovation, Ontario Innovation Trust, Compaq Canada, and SHARCNET at UWO.

LITERATURE CITED

Birkhead, T. R., and A. P. Møller, eds. 1998. Sperm competition and sexual selection. Academic Press, San Diego, CA.

- Birkhead, T. R., and T. Pizzari. 2002. Postcopulatory sexual selection. *Nat. Rev. Genet.* 3:262–273.
- Eggert, A.-K., K. Reinhardt, and S. K. Sakaluk. 2003. Linear models for assessing mechanisms of sperm competition: the trouble with transformations. *Evolution* 57:173–176.
- Evans, J. P., and A. E. Magurran. 2001. Patterns of sperm precedence and predictors of paternity in the Trinidadian guppy. *Proc. R. Soc. Lond. B* 268:719–724.
- Evans, J. P., L. Zane, S. Francescato, and A. Pilastro. 2003. Directional postcopulatory sexual selection revealed by artificial insemination. *Nature* 421:360–363.
- Houde, A. E. 1997. Sex, colour, and mate choice in guppies. Princeton Univ. Press, Princeton, NJ.
- Kadow, P. C. 1954. An analysis of sexual behavior and reproductive physiology in the guppy, *Poecilia reticulata* (Peters). Ph.D. diss. New York University, New York.
- Kobayashi, H., and T. Iwamatsu. 2002. Fine structure of the storage micropocket of spermatozoa in the ovary of the guppy *Poecilia reticulata*. *Zool. Sci.* 19:545–555.
- Mesterton-Gibbons, M. 1999. On sperm competition games: incomplete fertilization risk and the equity paradox. *Proc. R. Soc. Lond. B* 266:269–274.
- Moore, H., K. Dvorakova, N. Jenkins, and W. Breed. 2002. Exceptional sperm cooperation in the wood mouse. *Nature* 418:174–177.
- Neff, B. D., P. Fu, and M. R. Gross. 2003. Sperm investment and alternative mating tactics in bluegill sunfish (*Lepomis macrochirus*). *Behav. Ecol.* 14:634–641.
- Nelder, J. A., and R. Mead. 1965. A simplex-method for function minimization. *Comput. J.* 7:308–313.
- Parker, G. A. 1990. Sperm competition games: raffles and roles. *Proc. R. Soc. Lond. B* 242:120–126.
- . 1998. Sperm competition and the evolution of ejaculates: towards a theory base. Pp. 3–54 in T. R. Birkhead and A. P. Møller, eds. Sperm competition and sexual selection. Academic Press, San Diego, CA.
- Parker, G. A., L. W. Simmons, and H. Kirk. 1990. Analyzing sperm competition data: simple models for predicting mechanisms. *Behav. Ecol. Sociobiol.* 27:55–65.
- Pilastro, A., and A. Bisazza. 1999. Insemination efficiency of two alternative male mating tactics in the guppy (*Poecilia reticulata*). *Proc. R. Soc. Lond. B* 266:1887–1891.
- Pilastro, A., J. P. Evans, S. Sartorelli, and A. Bisazza. 2002. Male phenotype predicts insemination success in guppies. *Proc. R. Soc. Lond. B* 269:1325–1330.
- Pizzari, T., and T. R. Birkhead. 2000. Female feral fowl eject sperm of subdominant males. *Nature* 405:787–789.
- Pitcher, T. E., B. D. Neff, F. H. Rodd, and L. Rowe. 2003. Multiple mating and sequential mate choice in guppies: females trade up. *Proc. R. Soc. Lond. B* 270:1623–1629.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical recipes in C: the art of scientific computing. Cambridge Univ. Press, Cambridge, U.K.
- Simmons, L. W. 2001. Sperm competition and its evolutionary consequences in the insects. Princeton Univ. Press, Princeton, NJ.
- Wedell, N., M. J. G. Gage, and G. A. Parker. 2002. Sperm competition, male prudence and sperm-limited females. *Trends Ecol. Evol.* 17:313–320.

Corresponding Editor: T. Day