

Exxon Valdez Oil Spill
Restoration Project Final Report

Shock Resistance and Observer Classification of Pink Salmon Eggs

Restoration Project 01492
Final Report

John F. Thedinga¹
Mark G. Carls¹
Jacek M. Maselko¹
Stanley D. Rice¹
Ronald A. Heintz¹
Robert E. Thomas²

¹Auke Bay Laboratory
Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
Juneau, Alaska

²California State University, Chico
Department of Biological Sciences
400 West First Street
Chico, CA 95929

February 2003

The Exxon Valdez Oil Spill Trustee Council administers all programs and activities free from discrimination based on race, color, national origin, age, sex, religion, marital status, pregnancy, parenthood, or disability. The Council administers all programs and activities in compliance with Title VI of the Civil Rights Act of 1964, Section 504 of the Rehabilitation Act of 1973, Title II of the Americans with Disabilities Act of 1990, the Age Discrimination Act of 1975, and Title IX of the Education Amendments of 1972. If you believe you have been discriminated against in any program, activity, or facility, or if you desire further information, please write to: EVOS Trustee Council, 441 West 5th Avenue, Suite 500, Anchorage, Alaska 99501-2340; or O.E.O. U.S. Department of the Interior, Washington, D.C. 20240.

Exxon Valdez Oil Spill
Restoration Project Final Report

Shock Resistance and Observer Classification of Pink Salmon Eggs

Restoration Project 01492
Final Report

John F. Thedinga¹
Mark G. Carls¹
Jacek M. Maselko¹
Stanley D. Rice¹
Ronald A. Heintz¹
Robert E. Thomas²

¹Auke Bay Laboratory
Alaska Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
Juneau, Alaska

²California State University, Chico
Department of Biological Sciences
400 West First Street
Chico, CA 95929

February 2003

Shock Resistance and Observer Classification of Pink Salmon Eggs

Restoration Project 01492

Final Report

Study History: Restoration Project 01492 was initiated in FY00 as Project 00492. Three detailed project plans have been submitted. One paper has been published in *Transactions of the American Fisheries Society* and one paper has been accepted for publication in *Alaska Fishery Research Bulletin*.

Detailed Project Plans:

- 1) Project 00492. Were Pink Salmon Embryo Studies in Prince William Sound Biased?
- 2) Project 01492. Were Pink Salmon Embryo Studies in Prince William Sound Biased?
- 3) Project 02492. Were Pink Salmon Embryo Studies in Prince William Sound Biased?

Publications:

- 1) Thedinga, J. F., M. G. Carls, J. M. Maselko, R. A. Heintz, and S. D. Rice. In press. Resistance of Naturally Spawned Pink Salmon Eggs to Mechanical Shock. *AK. Fish. Res. Bull.*
- 2) Carls, M. G., J. F. Thedinga and R. E. Thomas. 2004. Observer Classification of Live, Mechanically Damaged, and Dead Pink Salmon Eggs. *Trans. Am. Fish. Soc.* 133:245-251.

Abstract: After the 1989 *Exxon Valdez* spill, pink salmon (*Oncorhynchus gorbusha*) embryo mortality was greater in oiled streams than in reference streams. The possibility that systematic, sampler-induced errors explain these differences is compared to an alternative hypothesis, oil toxicity. Resistance to hydraulic shock in a naturally-spawned pink salmon egg population increased more slowly than in uniform-aged eggs as a result of mixed embryo ages. The number of eggs that died from natural causes was unrelated to sample time. The most common observer errors were misclassification of shocked eggs as live ($\leq 9 \pm 1\%$) or as dead ($\leq 4.6 \pm 1\%$). When observation times were restricted to 0.5 h, $< 3\%$ of shocked eggs were classified as dead. Combining our observations with modeled estimates of the magnitude of observer error required to eliminate statistical differences between oiled and reference streams, we found these potential errors were insufficient to explain survival differences in Prince William Sound streams. Rather, we conclude that the weight of evidence favors oil toxicity as the cause of elevated embryo mortality: various studies documented oil in sediment around oiled streams, embryo exposure, high toxicity of dissolved oil compounds, and a mechanism to transport dissolved oil from beaches to pink salmon eggs.

Key words: Pink salmon, *Oncorhynchus gorbusha*, egg, hydraulic pumping, mechanical damage, shock, natural mortality, observer discrimination, Prince William Sound, *Exxon Valdez*.

Project data:

Lovers Cove Creek shock resistance data: excel spreadsheets. John Thedinga, National Marine Fisheries Service. 907-789-6025. John.Thedinga@noaa.gov

Egg discrimination data: Lotus spreadsheets. Mark Carls, National Marine Fisheries Service. 907-789-6019. Mark.Carls@noaa.gov

Modeling data: Lotus spreadsheets. Jacek Maselko, National Marine Fisheries Service. 907-789-6067. Jacek.Maselko@noaa.gov

Citation: Thedinga, J. F., M. G. Carls, J. M. Maselko, R. A. Heintz, R. E. Thomas and S. D. Rice. 2003. Shock resistance and observer classification of pink salmon eggs. *Exxon Valdez* Oil Spill Restoration Project Final Report (Restoration Project 01492), National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Auke Bay Laboratory, Juneau, Alaska.

Table of Contents

Executive Summary	v
1. Introduction	1.1
2. Resistance of Naturally Spawmed Pink Salmon Eggs to Mechanical Shock	2.1
Abstract	2.1
Introduction	2.2
Methods	2.3
Results	2.4
Discussion	2.5
Conclusions	2.7
Figures	2.8
3. Observer Classification of Live, Mechanically Damaged, and Dead Pink Salmon Eggs ..	3.1
Abstract	3.1
Introduction	3.2
Methods	3.2
Results & Discussion	3.4
Tables	3.7
Figures	3.8
4. Modeling Prince William Sound Embryo Data	4.1
Introduction	4.1
Methods	4.1
Results & Discussion	4.2
5. Conclusions	5.1
Acknowledgments	6.1
Literature Cited	6.1

Executive Summary

After the *Exxon Valdez* spill in 1989, routine hydraulic sampling of pink salmon eggs (*Oncorhynchus gorbusha*) by Bue et al. (1996; 1998) found higher embryo mortalities in oiled streams compared to non-oiled streams. However, Brannon et al. (2001) suggested these mortality differences were caused by differences in the proportion of immature eggs susceptible to sampler-induced mechanical damage (shock) as a result of systematic differences in spawn timing between oiled and reference streams and were not related to oil exposure.

This study, examined some of the underlying questions surrounding the egg-shock dispute with both field and laboratory components and found no reason to dispute the conclusion that the exposure to oil caused embryo mortality in oiled streams related to oil exposure (Rice et al. 2001). Objectives of the field tests were to determine how shock resistance of naturally spawned pink salmon eggs changes over time, how it relates to spawn timing, and how to provide accurate pre-disturbance estimates of egg mortality. Naturally spawned eggs were hydraulically sampled in Lovers Cove Creek and known-age eggs were sampled periodically at the Little Port Walter Hatchery. In controlled laboratory tests, the ability of observers to accurately discriminate among live, dead, and shocked eggs was tested with six laboratory trials.

Resistance of spawned pink salmon eggs to hydraulic shock increased sigmoidally over time but was more protracted in the stream than in uniform-aged hatchery eggs because maturity varied widely in the naturally spawned egg population. The percentage of shock resistant eggs (embryos) in Lovers Cove Creek increased from about 4% to 98% between September 27 and November 17 (14 days before to 37 days after spawning ended) and was significantly correlated with time ($R^2 = 0.85$; $P < 0.001$). Percentages of maturing eggs, as determined by pigmented eyes, increased in parallel to resistance to shock ($R^2 = 0.84$; $P < 0.001$), but resistance to shock preceded eye pigmentation. The percentage of eggs that died from natural causes in Lovers Cove Creek was not correlated with time ($R^2 = 0.04$; $P = 0.81$) and varied widely by transect (2-59%). Shock resistance in uniform-age eggs increased from 0-94% between 1 and 28 days after fertilization.

Human observers accurately discriminated among live, dead, and shocked eggs, particularly when observation was limited to realistic time intervals. Average discrimination error (mean total percentage of misclassified eggs, averaged across all observation times and observers) ranged from 1-4% in early and late trials, but peaked at 10-12% in the second and third of six independent trials ($P_{ANOVA} < 0.001$). The most common mistake (up to 9%, averaged across all observation times and observers) was misclassification of shocked eggs as live eggs, an error that is irrelevant in field studies designed to determine natural death rate. Misclassification of shocked eggs as dead eggs was the second largest source of error (up to 4.6%) when observation times were unrestricted (≤ 60 minutes); this was reduced to $< 0.5\%$ when observations were limited to ≤ 12 minutes. The third most common error ($\leq 2.3\%$) was dead eggs classified as shocked eggs. Inexperienced observers were easily trained to identify mechanically damaged eggs in one session, thus they can provide accurate data in field settings. Observer learning was evident in the first trial, and error rates declined after each individual understood the egg classification system.

Ideally pink salmon eggs should be allowed to incubate for a month before hydraulic sampling takes place so that most embryos become resistant to mechanical shock. However, our

improved sampling technique provides clear distinction between shocked and naturally dead eggs, thus allowing greater latitude in sample timing and alleviating problems posed by differing egg maturity within and between streams. In all cases, we caution that observers should be very intentional about discriminating between shocked and naturally dead eggs. Combining live and shocked eggs into one “live” category provides an accurate description of pre-sampling conditions, but combining shocked and dead eggs into one “dead” category does not.

We conclude that the weight of evidence favors oil toxicity as the cause of elevated embryo mortality in oiled PWS streams, not sampler-induced damage. The hypothesis by Brannon et al. (2001) requires systematic differences in spawn timing between oiled and reference streams and failure of observers to correctly recognize eggs damaged by sampling methods. Evidence to support meaningful systematic differences in spawn timing is weak. Estimated run-timing differences in all years except 1990 and 1992 were typically small and ranged from 2 to 4 d (Craig et al. 2002). Furthermore, the slow change in shock resistance in natural systems (Thedinga et al., Chapter 2) suggests these small differences in timing are of little consequence. Craig et al. (2002) identified one measure, the difference between spawn timing and sample date, that supports the possibility their observations were affected by sampling date, but they also found that the lack of specific run timing data in most years did not allow conclusive determination of the magnitude of the timing effect. The conclusion that embryo survival was lower in oiled streams than in reference streams was unchanged when sample timing was included in the analysis as a covariate in the one year (1991) when run-timing was monitored with rigor in many PWS streams (Craig et al. 2002).

To estimate the possible importance of observer classification errors in the data of Bue et al. (1996; 1998), the error required to eliminate statistical differences between oiled and reference streams was estimated by modeling (Maselko et al., Chapter 4) and was compared to classification errors observed in our laboratory study (Carls et al., Chapter 3). We conclude that the likelihood was small that potential misclassification errors by Bue et al. (1996; 1998) were sufficient to explain observed survival differences between oiled and reference streams.

In contrast to the sampler artifact hypothesis, there is ample support for an alternative hypothesis, that embryos in oiled PWS streams were damaged by exposure to *Exxon Valdez* oil. These studies document 1) oil in pink salmon habitat (e.g., Brannon and Maki 1996; Geiger et al. 1996; Murphy et al. 1999), 2) exposure of pink salmon embryos to oil (Wiedmer et al. 1996), 3) high oil toxicity (e.g., Marty et al. 1997; Heintz et al. 1999), 4) dissolution of toxic oil compounds into surrounding water (e.g., Carls et al. 1999; Heintz et al. 1999), 5) a mechanism to deliver dissolved hydrocarbons to pink salmon eggs (Carls et al. 2003), and 6) detection of *Exxon Valdez* oil in the water of one contaminated stream a decade after the spill (Carls et al. 2004b). These studies support the conclusion by Rice et al. (2001) that exposure to oil increased the incidence of pink salmon embryo mortality.

Our study provides specific guidelines to minimize potential errors in future field studies, clarifies the ability of observers to discriminate among natural and sampler-induced mortality, but does not support the sampler-induced egg damage hypothesis (Brannon and Maki 1996; Brannon et al. 2001).

Chapter 1

Introduction

One ongoing dispute between government and industry researchers concerning the impact of the *Exxon Valdez* oil spill on pink salmon in Prince William Sound (PWS) is the potential of the sampling technique (hydraulic egg pumping) to bias results among streams where spawn run-timing may have been different. Government researchers concluded that pink salmon embryo survival was lower in oiled streams than in non-oiled streams from 1989-1993 and again in 1997 (Bue et al. 1996; 1998; Craig et al. 2002). Industry researchers allege that government sampling in oiled streams was earlier than in reference streams relative to run timing, thus biasing estimates of egg survival, because early egg stages are more susceptible to mechanical damage caused by hydraulic pump sampling than later stages (Brannon and Maki 1996; Brannon et al. 2001). Industry researchers further contend that government observers failed to discriminate between previously dead eggs and those killed by sampling, thereby compounding the problem. The controversy continues after 11 years to cloud estimates of damage, restoration strategies, the impact of long term damage, and the definition of full recovery for this species. This study attempts to shed some light on the controversy.

The original objectives of Restoration Project 00492 were to conduct experimental studies for identifying shocked eggs and determining the effect of time on egg shock resistance in FY01. These studies were successfully carried out in fall 2000 and have resulted in two papers (Chapters 2 and 3). Salient findings from the field study were that 1) resistance of naturally spawned eggs to hydraulic shock increased sigmoidally from < 2% to 98% between late September and mid November, 2) the number of eggs that died from natural causes was unrelated to sample time, 3) shock resistance increased more slowly in natural systems than in reference eggs of uniform age because extended spawning resulted in a population of mixed-age embryos. In the laboratory we determined that 1) mean errors in egg classification were typically $\leq 3.5\%$ and did not exceed 12%, 2) the most common error was misclassification of shocked eggs as live ($\leq 9 \pm 1\%$), 3) the second most common error was classification of shocked eggs as dead ($\leq 4.6 \pm 1\%$) when observation times were unrestricted (≤ 60 minutes), and 4) classification errors were minimized when the time between shock and observation was limited. For example, misclassification of shocked eggs as dead was reduced to < 0.5% when observations were limited to ≤ 12 minutes. We conclude that rapid removal of all eggs from water allows accurate discrimination between shocked and previously dead eggs, an improved method we recommend for future studies, and we caution that combining shocked and dead eggs into one “dead” category does not accurately describe natural mortality. Chapter 2 had been accepted for publication and chapter 3 has been published in a peer reviewed publication.

Further, we modeled the potential for misclassification error (Chapter 4) and re-evaluated the potential for sampling bias to affect the observation of Bue et al. (1996; 1998) that embryo mortality in oiled intertidal streams was greater than in reference streams. We conclude the likelihood that true misclassification errors in the study by Bue et al. (1996; 1998) were as high as the modeled values are small: a misclassification of 9.5% shocked eggs as dead was required to render differences between oiled and reference streams nonsignificant, more than twice the

largest mean misclassification rate observed in one trial of our laboratory study under the worst conditions (unrestricted observation time).

An additional objective was added in FY02, to evaluate the veracity of the original egg categorization by Bue et al. (1996; 1998) by inspecting the preserved egg collection from PWS (1991-1992). Preliminary analysis of the preserved eggs showed that the developmental stage of the eggs could not be accurately determined because of sample deterioration, and we concluded that further inspection of this collection would not help clarify the controversy over potential bias in the egg mortality studies.

At the same time our work was in progress, Craig et al. (2002) completed additional analyses of the original data set by Bue et al. (1996; 1998) to evaluate the possibility that systematic differences in run timing caused the observed differences in egg mortality between oiled and reference streams. Estimated run-timing differences in all years except 1990 and 1992 were typically small; differences ranged from 2 to 4 d (Craig et al. 2002). Craig et al. (2002) found that one measure, the difference between spawn timing and sample date, supports the possibility that their observations were affected by sampling date, but that the lack of specific run timing data in most years did not allow conclusive determination of the magnitude of the timing effect. The conclusion that embryo survival was lower in oiled streams than in reference streams was unchanged when sample timing was included in the analysis as a covariate in the one year (1991) when run-timing was monitored with rigor in many PWS streams (Craig et al. 2002).

Other concurrent studies support the conclusion (Rice et al. 2001) that pink salmon embryos were damaged by exposure to *Exxon Valdez* crude oil. These studies demonstrated 1) *Exxon Valdez* oil in beaches surrounding intertidal pink salmon streams in PWS (e.g., ADFG 1989; Brannon and Maki 1996; Murphy et al. 1999), 2) that low concentrations are toxic to fish embryos (4-18 $\mu\text{g/L}$ total polynuclear aromatic hydrocarbons; Marty et al. 1997; Heintz et al. 1999), 3) that toxic quantities of oil dissolve into pore water from oil-coated sediment (e.g., Marty et al. 1997; Heintz et al. 1999), 4) that groundwater drains from surrounding beaches into water below these streams, thus potentially exposing incubating pink salmon eggs (Carls et al. 2003), and 5) that oil was detected by passive membrane samplers in subsurface water of one PWS stream in 1999 (Carls et al. 2004b). These observations are all consistent with the conclusion that pink salmon embryos were damaged by *Exxon Valdez* oil in PWS.

Chapter 2

Resistance of Naturally Spawned Pink Salmon Eggs to Mechanical Shock

John F. Thedinga, Mark G. Carls, Jacek M. Maselko, Ronald A. Heintz, and Stanley D. Rice

Abstract

Routine hydraulic sampling of pink salmon eggs (*Oncorhynchus gorbuscha*) is the subject of a long-running dispute over impacts of the *Exxon Valdez* oil spill on embryo survival in Prince William Sound, Alaska, because relationships between the time of spawning, sensitivity of eggs to mechanical damage, and sample timing were unclear. Previous laboratory or hatchery studies demonstrate that resistance of eggs to mechanical damage increases with maturity, but applicability to natural populations requires an understanding of embryo age distributions and the ability to discriminate between sampler-induced egg mortality and natural mortality. Resistance of naturally spawned eggs to hydraulic shock, determined six times between late September and mid-November in a southeastern Alaska stream, increased sigmoidally from < 2% to 98%. In contrast, the number of eggs that died from natural causes was unrelated to sample time. Rapid removal of all eggs from the water allowed accurate discrimination between shocked and eggs dead prior to sampling, an improved method we recommend for future studies. The rate of shock resistance increase was slower in naturally spawned eggs than in uniform-age embryos subjected to the same hydraulic shock. We caution that combining shocked and dead eggs into one “dead” category does not accurately describe natural mortality.

Introduction

Hydraulic pumping is a typical way to assess salmon spawning success during the freshwater incubation stage, and pumping data are used in population dynamic models. In this method, a mixture of air and water is violently injected into streambed gravel by a hydraulic pump, dislodging eggs and forcing them to the surface where they are trapped in a net (McNeil 1964a). Live eggs and dead eggs are then counted; the total number of live eggs is typically used in management models to predict potential run strengths when fish return as adults. The challenge is to accurately discriminate between eggs that were previously dead from eggs mechanically damaged (“shocked”) by hydraulic pumping. Mechanical disturbance can cause embryo mortality 0-12 days after fertilization (Jensen and Alderdice 1989; Jensen 1997) and possibly for the first 20 days (Collins et al. 2000) because the delicate viteline membrane serves as the primary barrier between surrounding water and the yolk during this time period. If the membrane is ruptured, water penetrates the yolk, lipoproteins coagulate, and the embryo dies. Thus ideally, streams should be hydraulically sampled after eggs become resistant to shock, but each stream may have a relatively long spawning period (1.5-2 months; Dvinin 1952) in the fall, and advancing winter can limit sampling options.

Hydraulic sampling was used in Prince William Sound following the 1989 *Exxon Valdez* oil spill to assess the potential impact of spilled oil on the hundreds of pink salmon streams that cross oiled beaches; the results are controversial. Bue et al. (1996; 1998) demonstrated higher levels of mortality of pink salmon (*Oncorhynchus gorbuscha*) embryos in the intertidal reaches of streams in Prince William Sound, Alaska that were exposed to *Exxon Valdez* oil; but these observations were challenged by others (Brannon and Maki 1996; Brannon et al. 2001) because relationships between the time of spawning, egg sensitivity, and hydraulic sampling were unclear. Redd superimposition (disruption of redds by later-spawning fish; Dvinin 1952; McNeil 1964b; Heard 1991) is an important consequence of long spawning periods and up to 1/3 of spawned eggs can be displaced or damaged by superimposition (Fukushima et al. 1998). As a result, displaced eggs die and average embryo development is controlled by later-spawning fish. Brannon et al. (2001) argue that the higher embryo mortality in oiled streams was simply an artifact of sample timing and suggest that spawning was frequently later in oiled streams than in non-oiled streams and consequently, these streams had greater proportions of shock sensitive embryos. Craig et al. (2002), however, continued to find oil-induced mortality when the time of spawning was included as a covariate in 1991, the only year in which the timing of spawning for individual streams was accurately monitored (Rice et al. 2001). This controversy stimulated an examination of the resistance of naturally spawned pink salmon egg populations to mechanical damage.

The objectives of this study were to determine how shock resistance of naturally spawned pink salmon eggs changes over time, how it relates to the timing of spawning and how to provide accurate pre-disturbance estimates of egg mortality. While hatchery-related studies have examined shock induced mortalities, Collins et al. (2000) were the first to publish the results of a study on the resistance of naturally spawned pink salmon embryos to shock in the field, but populations were only sampled twice, about the time spawning ended and a month later. Their data established that naturally spawned eggs become resistant to shock within one month, but did not define the rate of shock resistance. Thus, our primary objectives were to determine how

shock resistance in naturally spawned eggs changes over time (requiring periodic sampling during and after spawning) and how it relates to the timing of spawning. Spawning time was determined by periodically counting spawners in the stream. Most importantly, our measurements had to separate naturally dead eggs (dead prior to sampling) from eggs damaged or killed during the sampling process; hence, post-collection egg handling was improved to ensure accurate classifications. In this study, eggs were classified as dead only if they died naturally in the stream prior to sampling, distinct from eggs killed by the sampling process. Ancillary data were collected to compare shock resistance in naturally spawned eggs to resistance in artificially spawned eggs of known age. Our study was conducted on a relatively homogenous reach of stream where cross-transects were sampled once, and the sampling did not influence other transects sampled at later dates.

Methods

Study area

Lovers Cove Creek is located on eastern Baranof Island in southeastern Alaska (Fig. 2.1). The stream has three primary channels that enter an extensive intertidal area that extends approximately 458 m (Hanavan and Skud 1954), encompasses about 3.5 ha, and flows into the head of Port Walter. About 60,000 pink salmon spawn in the intertidal portion of the stream each year. Our observations were restricted to a 100 m segment in the eastern branch of Lovers Cove Creek including a straight 83 m intertidal section with fairly uniform gradient and gravel size (Martin 1973); mean channel width was 15 m.

Hourly water temperature of Lovers Cove Creek was measured with a thermograph beginning October 19, 2000. Daily Sashin Creek temperatures recorded at the Little Port Walter hatchery were regressed with Lovers Cove Creek temperatures to estimate Lovers Cove Creek temperatures prior to October 19.

Adults were counted in the study section of the stream about twice a week from September 29 to October 11. Two observers positioned about 20 m from the stream counted the fish at low tide and the mean of their observations was used as the count for that day.

Resistance of naturally spawned eggs to mechanical damage

The 100-m study section of Lovers Cove Creek was divided into 25 transects, each perpendicular to stream flow, and marked on the banks with metal stakes. Three or four randomly selected transects were sampled without replacement during one of six 2-3 d sampling periods between September 27 and November 15, 2000. For each transect sampled, a line was fastened across the stream, and gravel was excavated within 1.5 m upstream of the line with a 1-m long by 3.8 cm diameter stainless steel probe discharging a 170 L/min. air-water mixture. A cylindrical basket (0.1 m²) with a 1 mm mesh collection bag surrounded the probe. Each 0.1 m² area was pumped for 1 minute; dislodged eggs were transferred immediately to a plastic tray. Eggs from consecutive pumpings were combined to form samples of at least 100 eggs; pumping along a transect continued until a minimum of five samples were completed. Eight to 35 pumpings per transect were necessary to obtain sufficient numbers of eggs; a total of 2415-4756 eggs were collected during each sample period. Care was taken to not walk on adjacent transects to avoid damaging incubating eggs.

Eggs were sorted from the gravel and debris, removed from the water, and placed on a screen (1 mm mesh size) for classification and counting. Eggs were classified as live (pink) without visible eye pigmentation, live with pigmented eyes, dead (white), or shocked (changing from pink to white). Removal of eggs from water arrested the characteristic change in shocked eggs from pink to white that occurs as proteins coagulate, and ensured that shocked eggs were not confused with natural egg mortality. (Serendipitous observation in a concurrent experiment, Carls et al. (2004a) demonstrated that removal from water arrests color change.) After initial classification, to ensure that mildly shocked eggs were not misidentified as live eggs, remaining pink-colored eggs were gently placed in water for about 10 minutes to allow continued whitening of mildly shocked eggs. Empty chorions were not counted and were not used for egg classification because of the uncertainty of their origin.

Shock resistance in known-age eggs

To compare the shock resistance of naturally spawned eggs to those of known age, gametes from Lovers Cove Creek pink salmon (3 females, 2 males) were artificially crossed, incubated in the nearby Little Port Walter hatchery, and periodically tested for shock resistance. Egg subsamples were placed in a simulated redd, a 208 L barrel filled with 80 cm of Lovers Cove Creek gravel and pumped with the same equipment and procedures used to collect naturally spawned eggs. On four days (1-28 d after fertilization), about 200 eggs were placed within a 10 cm diameter by 2 cm high aluminum ring in the simulated redd. The ring was covered with eight $0.05 \times 2 \times 10$ cm pieces of plastic to protect the eggs as they were covered with approximately 20-25 cm of additional gravel from Lovers Cove Creek: (20-25 cm redd depths are typical for well-populated spawning grounds; Heard 1991). Before hydraulic sampling, water was added so that the gravel surface was covered with about 20 cm of water.

Data analysis

Shock resistance, percentage of eyed eggs, and percentage of naturally dead eggs with time were described by logistical regression (SAS GENMOD procedure, SAS 1999), corrected for overdispersion (Williams 1982). The times when half (or 90 %) of the embryos became resistant to shock (or became eyed) were estimated from the logistic equations and are reported as time \pm SD.

Results

Run timing

Pink salmon spawning in Lovers Cove Creek began about September 1 and ended 40 days later (October 11). There were two peaks in the run, September 5 (294 fish) and October 5 (393 fish) (Fig. 2.2); a total of 1085 adult pink salmon were counted. Spawning ended on October 11, 2000. Mean daily water temperature in the creek during spawning was 7.2°C.

Resistance of naturally spawned eggs to mechanical damage

Resistance of eggs to hydraulic shock in Lovers Cove Creek increased sigmoidally over time (Fig. 2.3a). The percentage of shock resistant eggs (embryos) increased from about 4% to 98% between September 27 and November 17 (14 days before to 37 days after spawning ended)

and was significantly correlated with time ($R^2 = 0.85$; $P < 0.001$). When spawning ended on October 11, an estimated 23% of the eggs were resistant to shock. Twelve days after spawning ended, the upper end of the sensitivity period predicted by Jensen (1997), just over half (55%) of the eggs were resistant to shock. Twenty days after spawning ended, the end of the sensitivity period predicted by Collins et al. (2000), 76% of the eggs were resistant to shock. Resistance to hydraulic shock did not reach 90% until 28 days after all spawning ended.

Percentages of maturing eggs, as determined by pigmented eyes, increased in parallel to resistance to shock ($R^2 = 0.84$; $P < 0.001$), but resistance to shock preceded eye pigmentation (Fig. 2.3a). For example, half the eggs were resistant to shock on October 21 (10 d after spawning ended), but half of the embryos did not have pigmented eyes until October 31 (20 d after spawning ended).

Naturally dead eggs

The percentage of eggs that died from natural causes in Lovers Cove Creek was not correlated with time ($R^2 = 0.04$; $P = 0.81$) and varied widely by transect (Fig. 2.3b). Mean percent dead eggs varied from 12-59%, as computed on a daily basis. Percentages of dead eggs by transect ranged from 7-95%.

Shock resistance in known-age eggs

Shock resistance in known-age eggs increased more rapidly than in naturally spawned eggs because these eggs were all spawned at the same time (Fig. 2.3c). Percentages of eggs resistant to shock increased from 0-94% between 1 and 28 days after fertilization. Twelve days after fertilization, just 25% of the eggs were resistant to hydraulic shock but 20 d after fertilization, 80% were resistant. Resistance to hydraulic shock reached 90% 22.7 ± 0.4 d after spawning ended.

Discussion

Results of this study demonstrate that resistance of naturally spawned pink salmon eggs to hydraulic shock increases sigmoidally over time, but changes more slowly than in a population of simultaneously fertilized eggs. Pink salmon egg susceptibility to mechanical damage has previously been thoroughly studied with known-age eggs (e.g., Jensen 1997). The application of laboratory results to field research requires an understanding of embryo age distributions in wild populations and testing to ground-truth these estimates. In this discussion, we compare our results to the only other field study published on this topic (Collins et al. 2000) and discuss the need to discriminate between natural egg mortality and mortality caused by the sampling technique. These issues relate to interpretation of pink salmon egg mortality data collected after the *Exxon Valdez* oil spill, a topic that has proven controversial (e.g., Bue et al. 1998; Brannon et al. 2001; Rice et al. 2001) because the relationship between the time of spawning and hydraulic sampling was unknown in most years.

An obvious primary difference between naturally spawned egg populations and those used in the laboratory assessment of shock resistance is that embryo ages will be more variable in the wild than in the laboratory. The expected effect is that average resistance to shock should rise more slowly in eggs from wild populations than in uniform-age eggs because of larger variability

in egg maturity in wild populations, delaying the rate of increase in resistance. This is just what was observed – adults spawned over a 40-d period and the rate of egg resistance to shock increased more slowly in the wild population than in uniform-age embryos subjected to the same hydraulic shock.

Redd superimposition also contributed strongly to delayed resistance to hydraulic shock by skewing the population toward immaturity as older embryos were displaced and killed. For example, 65% of the spawning was completed three weeks prior to October 17, and at observed water temperatures of approximately 7°C those eggs should have matured to eyed stage, yet only 12% were eyed on this date. Half the laboratory embryos were resistant to shock within 16 d, but 16 d after 75% of the spawning was complete, just 32% of wild embryos were resistant. The exact effects of superimposition likely vary spatially and temporally, cannot be predicted with high precision, and will also vary according to population differences in run timing.

Our results are similar to those of Collins et al. (2000) but we sampled more frequently and thus were able to clearly define the rate at which resistance to shock developed. Collins et al. (2000) reported that shock resistance of pink salmon eggs in Prince William Sound increased from about 58% at the end of spawning to 98% about one month later. We observed less shock resistance (23%) at the end of spawning, possibly because of differences in the timing of spawning, superimposition between study sites, or hydraulic energy, but a month after spawning 92% of the eggs were resistant to shock, similar to estimates by Collins et al. (2000). Collins et al. (2000) predicted naturally spawned eggs would become resistant to shock in about 20 days but our data indicate about a quarter of the embryos can be damaged by hydraulic pumping at this time. Most, but not all embryos eventually become immune to shock; 36 d after spawning ended only 2% were shocked, a result reasonably consistent with Collins et al. (2000; 2% shocked 20-29 d after spawning) given the potential differences between streams, salmon behavior, and sampling procedures.

Importance and strategy for discriminating natural mortalities from unnatural shocked mortalities

Discrimination between naturally dead eggs and shocked eggs is crucial in determining the accurate number of pre-sampling mortalities in a stream. For example, after the *Exxon Valdez* oil spill, the question was “did oil affect pink salmon eggs in the oil-exposed intertidal reaches of streams”? Unfortunately, the discrimination issue led to controversy; Bue et al. (1996) found an oil-related effect, but Brannon et al. (2001) alleged that there were differences in timing of spawning between oiled and reference streams, hence differences in percentages of shocked eggs, and that shocking was responsible for recorded survival differences. Bue et al. (1998) and Craig et al. (2002) were able to address some of these concerns in two ways: (1) by incubating eggs from oiled and reference streams together in hatchery incubators without the issue of shocking or run timing and (2) by including the time of spawning as a covariate in the analysis for the only year (1991) with spawn timing data. This controversy clearly illustrates the importance of distinguishing dead from shocked eggs.

Modification of egg handling techniques developed during this study improves the ability of future observers to discriminate between naturally dead eggs and those damaged by the sampling procedure. How eggs are handled after they are collected affects the speed at which shock symptoms appear. If they remain in fresh water, shocked eggs continue to change from

pink to opaque white and within minutes become very difficult to distinguish from previously dead eggs. Depending on the severity, shocked eggs maintained in water can become indistinguishable from dead eggs in less than 10 minutes. Typically, when researchers hydraulically sample eggs, the eggs remain in a water-gravel mixture for extended periods as many eggs are collected, thus natural mortality can be overestimated. Our modification is to limit pumping to short time intervals (one minute), then quickly remove all eggs from water. If the eggs are in air, water can not enter, thus eggs remain pink and clearly distinguishable from naturally dead eggs. If an observer wishes to quantify shocking, eggs can be placed in water after the dead are counted, and shocked eggs will take on water and will continue to whiten in a few minutes.

Percentages of naturally dead eggs in Lovers Cove Creek varied among sampling periods but were not related to time, unlike percentages of shocked eggs that were significantly correlated with time. This difference provides further evidence that our improved sampling technique successfully allowed discrimination between shocked and dead eggs. This inference assumes that natural mortality (e.g., redd superimposition, stream bed scour, disease, inadequate dissolved oxygen, or disturbance by wading animals) was essentially constant. Although daily variance in natural mortality was occasionally high, this was probably due to dead eggs pumped from gravel with poor incubation conditions; that is spatial variability, not temporal variability. When calculated by sampling period (1-3 d), mean percentages of dead eggs (20-36%) in Lovers Cove Creek were similar to those reported by Collins et al. (2000) (21%). In sharp contrast to the unchanging percentage of naturally dead eggs, percentages of eggs described as shocked decreased predictably from nearly all to almost none, also providing evidence that the method provides good data.

Conclusions

Ideally pink salmon eggs should be allowed to incubate for a month before hydraulic sampling takes place so that most embryos become resistant to mechanical shock. However, our improved sampling technique provides clear distinction between shocked and naturally dead eggs, thus allowing greater latitude in sample timing and alleviating problems posed by differing run timing and egg maturity within and between streams. In all cases, we caution that observers should be very intentional about discriminating between shocked and naturally dead eggs. Combining live and shocked eggs into one “live” category provides an accurate description of pre-sampling conditions, but combining shocked and dead eggs into one “dead” category does not.

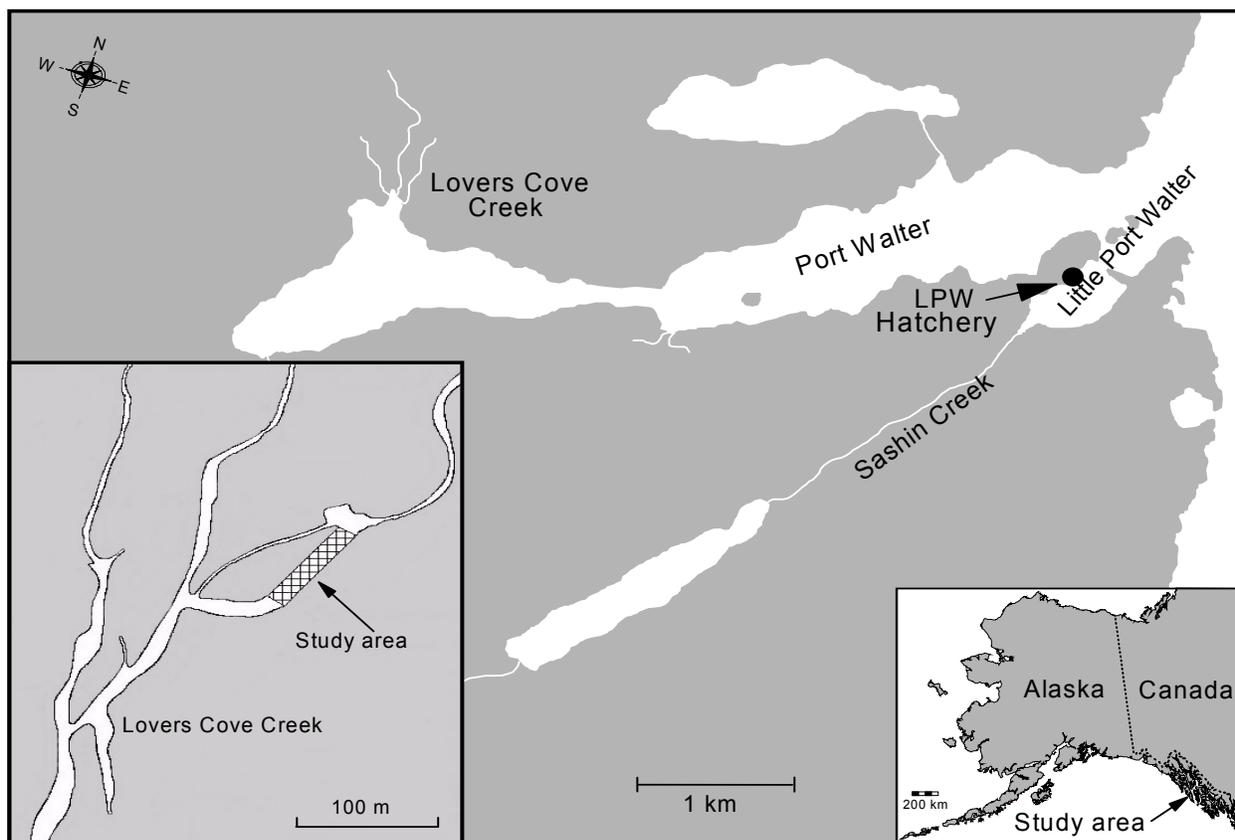


Figure 2.1. Lovers Cove Creek study area in southeastern Alaska.

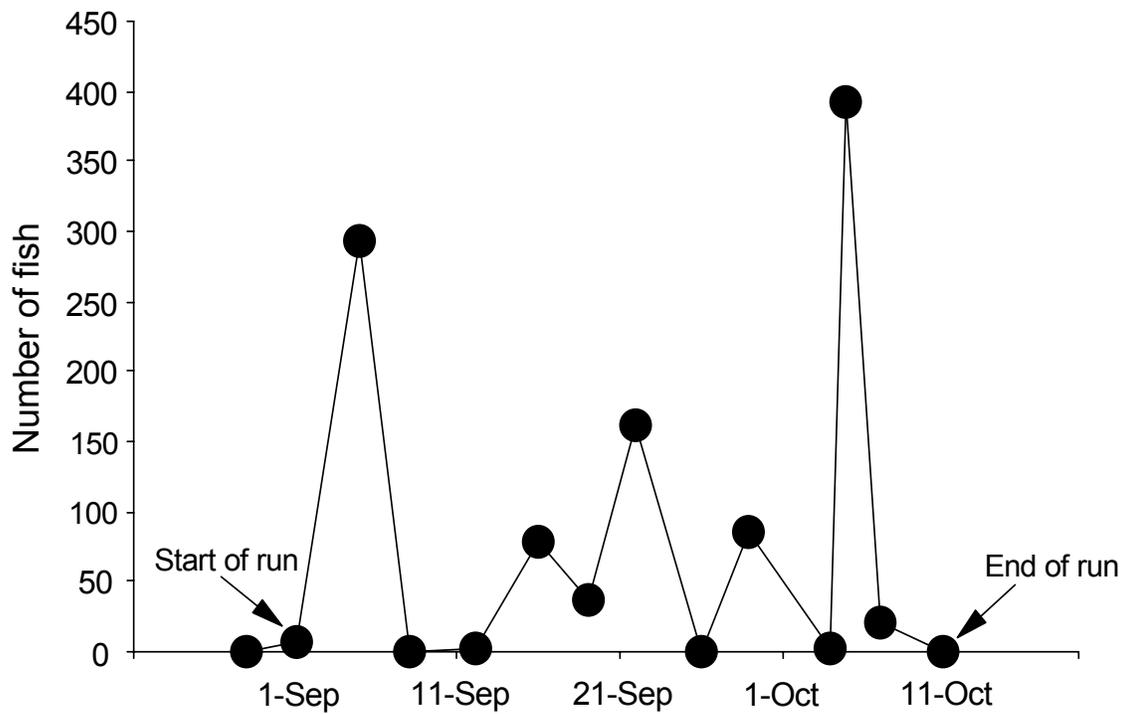


Figure 2.2. Run timing: number of adult pink salmon counted periodically in a 100 m study section of Lovers Cove Creek, Alaska in 2000.

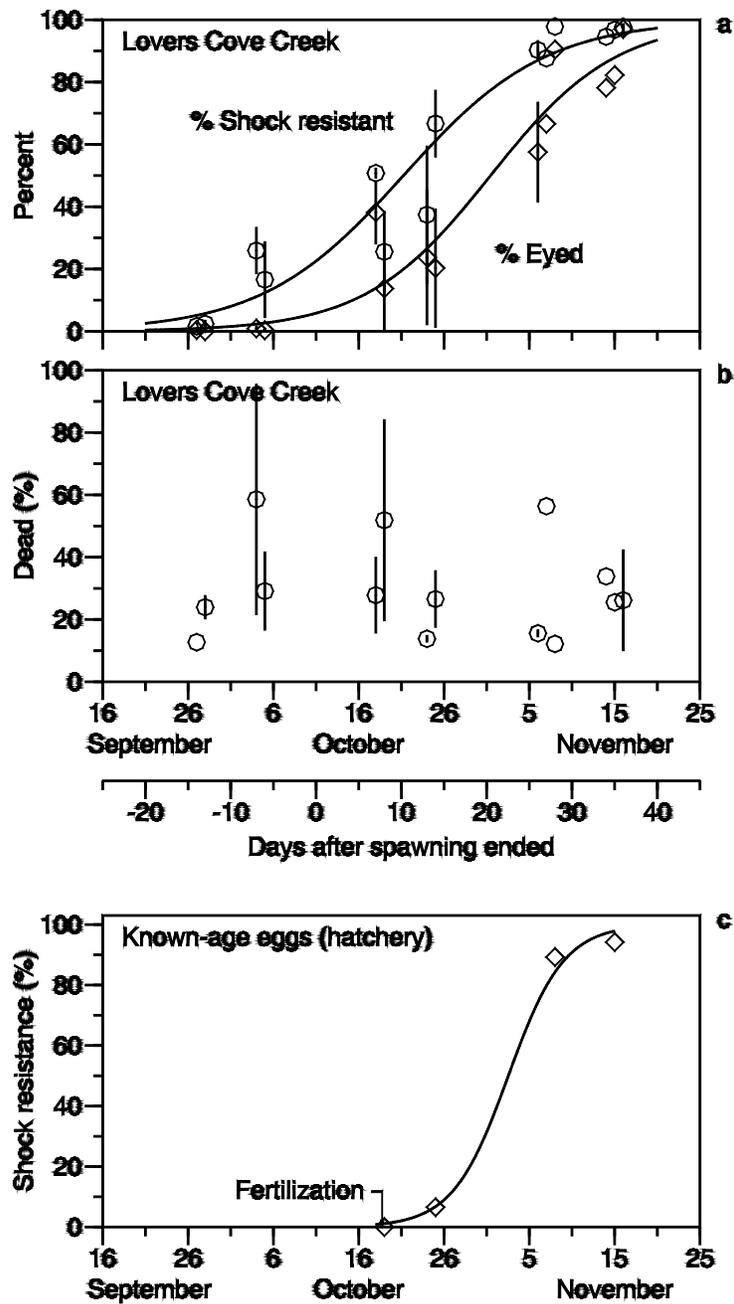


Figure 2.3. (a) Mean percentages of pink salmon eggs from Lovers Creek that survived hydraulic shock and mean percentages of live eggs that were eyed (\pm SE); $n = 1$ where error bars are absent. (b) Mean percentages of pink salmon from Lovers Cove Creek that died of natural causes. (c) Percentages known-age pink salmon eggs that survived hydraulic shock in an artificial redd.

Chapter 3

Observer classification of live, mechanically damaged, and dead pink salmon eggs

M. G. Carls, J. F. Thedinga, and R. E. Thomas

Abstract

Susceptibility of pink salmon (*Oncorhynchus gorbuscha*) eggs to mechanical damage (shock) was studied to test the ability of observers to discriminate among live, dead, and shocked eggs. In a series of six laboratory trials, the mean error rate in discrimination did not exceed 12% and was $\leq 3.5\%$ in 4 of 6 trials. The most common error was misclassification of shocked eggs as live ($\leq 9 \pm 1\%$), an error that is irrelevant in field studies designed to determine natural death rate. The second most common error was shocked eggs classified as dead ($\leq 4.6 \pm 1\%$) when observation times were unrestricted (≤ 60 minutes); this was reduced to $< 0.5\%$ when observations were limited to ≤ 12 minutes. Inexperienced observers were easily trained (within 1 hour) to classify eggs. To accurately describe natural systems before sample disturbance, shocked and dead egg categories should not be combined when reporting data.

Introduction

Hydraulic pumping is used as a research and management tool to collect eggs from salmon redds. For example, after the *Exxon Valdez* oil spill, pink salmon eggs were collected by hydraulic pumping to determine if exposure to oil had reduced embryo survival (e.g., Bue et al. 1998). This violent injection of an air-water mixture into streambed gravel can mechanically damage (shock) developing salmon eggs, particularly when embryos are the least mature, obscuring the distinction between natural and sampling mortality. The appearance of shocked eggs changes from pink to white as freshwater penetrates the ruptured vitelline barrier and causes protein coagulation. Resistance to shock increases as embryos mature and reinforce the barrier with epidermal tissue. In this paper, we use the word ‘egg’ to refer to either infertile or fertilized eggs. Early embryonic development is not visible macroscopically, thus the visual cues used by observers to classify condition are based on egg structures, not embryonic tissue. The same cues are used for infertile eggs and for fertilized eggs with visible (more mature) embryonic development.

Although the susceptibility of pink salmon (*Oncorhynchus gorbuscha*) eggs to shocking before pigmentation of embryonic eyes has been adequately described in laboratory (Smirnov 1954, 1975; Jensen and Alderdice 1989; Jensen 1997) and in field studies (Collins et al. 2000; Thedinga et al. In press), observer ability to discriminate among live, dead, and shocked eggs has not been reported. Only if the typical technician can readily discriminate egg condition will field data be accurate. Thus, our goal was to study observer ability to correctly identify damaged eggs rather than to study mechanical egg damage *per se*. We are aware of no other comparable research. To examine observer discrimination, 10 people repeatedly classified unknown mixtures of live, dead, and shocked eggs throughout early embryonic development and until remaining eggs were all resistant to shock. To supplement the primary observations and aid in data interpretation, resistance of eggs to shock was recorded throughout the one month trial period.

Methods

Most pink salmon eggs examined were collected from wild stock, fertilized, and maintained in hatchery conditions. Additional naturally spawned eggs were included in some trials (as explained later). Gametes collected on September 18, 2000 from pink salmon at Sashin Creek (Little Port Walter, Alaska) were kept cool and flown to the Auke Creek hatchery (Juneau, Alaska) where they were crossed. Ova were placed in plastic cups, a few milliliters of milt were added, freshwater was added, and the mixture was gently decanted between two cups three times. Fertilized eggs were placed directly into a Heath-tray incubator¹ with flowing freshwater. To control microbial growth, approximately 100 L of seawater was added to the incubator every 2-3 days with the freshwater flow interrupted for 1 hour.

¹Reference to trade names does not imply endorsement by NOAA Fisheries.

Egg resistance trials

At 3-7 d intervals, groups of eggs were mechanically shocked by dropping them from a height of 1 m onto a hard surface. Shock intensity was intentionally high to ensure shocking and to emulate the vigorous shock potential of hydraulic sampling methods. Numbers of shocked, slightly shocked, and surviving eggs were recorded (determined by color change). Samples of each classification were preserved and later inspected for development. Beginning 36 d after fertilization, infertile eggs were easily distinguishable from developing eggs, and observations were subdivided into developing and infertile groups. The percentage of developing eggs that survived shock was calculated from direct observation in later samples, or estimated from the average fertility rate in preserved early samples.

Discrimination trials

Ten observers were presented random mixtures of live, dead, and shocked pink salmon eggs in 10 petri dishes, with 30-66 eggs/dish. The numbers of eggs in each category were unknown to the observers. Each observer began classifying eggs about 5 minutes after shocking with up to 5 minutes allowed per dish for observation. Observers were allowed to sort eggs, but were required to gently mix them before moving to a new dish and were not allowed to compare results. During each 1-hour test period, observers rotated to a new dish every 5 minutes for the first 40 minutes and every 10 minutes thereafter. Based on their own assessments of prior experience, observers were subdivided into three classes, experienced, inexperienced, and intermediate.

Some variables were adjusted among trials. The total number of eggs per dish was constant (50) in trials 1 and 2, but variable in all other trials. In trial 1, nearly all dead eggs had visible microbial growth. Some live eggs were shocked during the course of this trial due to rough handling by inexperienced observers. In trials 2-4, dead eggs pooled from preceding trials were used as dead eggs, and the percentage of eggs with microbial growth was controlled (0-25% of dead eggs per dish). In trials 5-6, naturally spawned dead eggs were used: these were sampled by hydraulic pumping a week preceding each trial, fixed in 5-10% phosphate-buffered formalin, and soaked 1-2 days in flowing freshwater before observations. Dead eggs in trial 5 were typical dead eggs collected from Lover's Cove, near Little Port Walter, Alaska (October 17, 2000). Dead eggs in trial 6 were a mixture of typical dead, recently dead, and dead eyed eggs chosen at random (0-33% possible for each category).

Potential sources of error were recorded or calculated. All six potential sources of misclassification were calculated (live eggs scored as shocked or dead, shocked scored as live or dead, and dead scored as live or shocked) and expressed as a percentage of the total number of eggs reported. Other sources of error included incorrect egg counts, record-keeping errors, and additional shocking during the course of observation. Obvious record errors were infrequent and corrected. The number of shocked eggs was adjusted for change during the first trial (only) by assuming that most observers had correctly identified all shocked eggs in each dish. Where the true number of shocked eggs was ambiguous, the assessment of experienced observers was favored, and values were adjusted as infrequently as possible. Ambiguities in the first trial were due to additional egg shocking caused by inadvertent rough handling by inexperienced observers.

For any given trial, 80-100% of the original observers were present. Seventy percent of the original observers participated in all trials. One to two observers were added as substitutes in

three trials, and these substitutes each participated in two to three trials. Both substitute observers were inexperienced. Two of the original observers who could not participate in all trials were also inexperienced, and one was intermediate.

Data were analyzed with ANOVA or regression methods as appropriate. Percentages (but not cumulative percentages) were arc-sine transformed (Snedecor & Cochran 1980) prior to ANOVA. Multi-factor ANOVA (total percent eggs misclassified = trial + experience + time + dish nested in trial) was used to describe general observer discrimination across trials: trial = trial number, experience = experience level, and time = time after shock. Observer experience was summarized with two-factor ANOVA (cumulative percent eggs misclassified = trial + experience), where cumulative percentages were determined across all observation times. The importance of time for each of the six possible error types was explored with multi-factor ANOVA (pE = trial + experience + time + dish nested in trial), where pE = percent error. Time effects were further explored with logistic regression. Reported data are means \pm SE.

Results & Discussion

Resistance to mechanical damage

Developing eggs were initially susceptible to mechanical damage (shock), but resistance rose sharply after 15 d incubation, and 95% of viable embryos survived shock at ≥ 26 d (after onset of eye pigmentation; Fig. 3.1a). Infertile eggs were invariably shocked. As embryos matured, slight or partial shocking was observed in a few percent of viable embryos between 15 d (head and trunk differentiation) and 26 d (after onset of eye pigmentation; Table 3.1). Time-dependent increases in resistance of eggs to damage caused by the 1 m drop onto a hard surface in this study were similar to increased resistance to hydraulic pumping (Thedinga et al. In press) (Fig. 3.1a). The offset between shock resistance of eggs in this study and those by Thedinga et al. (In press) were most likely due to thermal differences; resistance to pumping and a 1 m drop were very similar in the latter study (data not shown).

Observer error

Average discrimination error (mean total percentage of misclassified eggs, averaged across all observation times and observers) ranged from 1-4% in early and late trials, but peaked at 10-12% in trials 2 and 3 ($P_{ANOVA} < 0.001$; Fig. 3.1b). The high frequency of microbial growth on dead eggs probably helped with discrimination in trial 1. Reasons for the larger error on intermediate days are not clear but could involve degree of separation difficulty, rapidly changing shock resistance, and observer experience. A population of immature embryos (prior to eyeing), which are susceptible to shock, and infertile eggs is likely the most challenging to classify accurately because there are no initial (pre-shock) color differences. When eyed embryos are frequent, classification may be easier because the color of older eggs, which are less susceptible to shock, is darker than immature and infertile eggs. (Older eggs tend to be covered with algal growth, perhaps because waste excretion supplies nutrients for such growth.) Condition of dead eggs is another factor; recently dead eggs are typically harder to distinguish from shocked eggs than dead eggs that have been discolored or have microbial growth. These factors may explain why observer error rates were highest about the time eggs rapidly became resistant to shock (Fig. 3.1).

Sources of observer error

The largest source of error (up to 9%, averaged across all observation times and observers) was misclassification of shocked eggs as live eggs (Fig. 3.2a). Assuming the underlying objective of typical hydraulic studies in natural systems is to distinguish the number of live and dead eggs in a stream before sampler influence, this type of error is unimportant.

Misclassification of shocked eggs as dead eggs was the second largest source of error (up to 4.6%, averaged across all observation times and observers) (Fig. 3.2b). As shocked eggs in trials 2-3 whitened they became more difficult to discriminate from the dead eggs produced in the hatchery. Discrimination between shocked eggs and dead eggs was relatively easy in the first trial because of the distinctive microbial growth on dead eggs. Dead eggs from a natural stream system were different enough from shocked eggs that shocked eggs were infrequently confused with dead, even when the least different wild dead egg category (recently dead) was included (in trial 6). Misclassification of shocked eggs as dead eggs can cause important errors in the field, but can be minimized by quick removal of eggs from water, and by restricting the time between shock (hydraulic pumping) and observation.

The third largest source of error ($\leq 2.3\%$) was dead eggs classified as shocked eggs (Fig. 3.2c). In trial 4, dead eggs were more frequently misclassified as shocked at the end of the trial, suggesting that this group of dead eggs looked more like advanced shocked eggs than eggs that had been dead longer. This error can also be minimized by quickly removing eggs collected by hydraulic pumping from water; under these conditions white eggs must have been dead before removal from water.

Other sources of error were minor, less than 1%, and generally may have been caused by recording errors rather than misclassification (Fig. 3.2d-f). However, misclassification of live eggs as shocked might have been caused by light refraction through the walls of the petri dishes. An inattentive observer might have confused the slight color change caused by refracted light with early shocking. Misclassification of live eggs as dead ($<0.25\%$) was almost certainly due to recording errors, not actual misclassification.

Observer experience

Inexperienced observers are easily trained to identify mechanically damaged eggs in one session and can provide accurate data in field settings. Observer learning was evident in the first trial, and error rates declined once each individual understood the egg classification system. Although inexperienced observers tended to have greater cumulative classification errors throughout the 1 month test period, there were no statistically significant differences between inexperienced and experienced observers ($P = 0.454$; Fig. 3.3a). One observer accounted for most of the 'inexperienced' variance in the first trial, but after 20 minutes and thereafter most shocked eggs were correctly identified by most observers, including the least accurate observer. Thus, learning to identify mechanically damaged eggs is not difficult.

Error control

Observer error rates for some classifications were time dependent ($0.0001 < P_{\text{ANOVA}} \leq 0.699$) and were consistent with the color change from pink (living) to white (dead) that occurs when eggs are shocked (Fig. 3.3b). The percentage of shocked eggs mistaken for live eggs decreased as time increased ($P_{\text{regression}} < 0.05$ in all 6 trials), and percentages of shocked eggs

mistaken for dead eggs increased with time ($P_{\text{regression}} < 0.05$ in 2 of 6 trials). No consistent relationship between time and other misclassifications was evident.

Observer error rates can be reduced by limiting the time between sampling (mechanical damage) and egg assessment, a conclusion also reached by Thedinga et al. (In press). Because shocked eggs become increasingly difficult to distinguish from dead eggs when maintained in water, egg condition should be assessed as soon as possible after collection. Thedinga et al. (In press) also found that the color change can be arrested by placing egg samples in air; under these conditions water does not enter damaged eggs and protein does not coagulate. Accordingly, classification errors are reduced, yet ample time is available for assessment. Quick removal of eggs from water is the critical issue. Thedinga et al. (In press) conservatively recommend limiting hydraulic pumping to 1 minute intervals. Results of this study suggest that assessment within 10-12 minutes (when eggs are maintained in water) may be adequate. However, Collins et al. (2000) report that eggs shocked by hydraulic sampling can turn opaque white within minutes, suggesting that collection of eggs in 1 minute intervals is a good strategy.

Not only does time restriction reduce the absolute observer error rate, but the sources of observer error change toward more accurate assessment of conditions before sampling (Fig. 3.2). In particular, the average percentage of shocked eggs erroneously scored as dead was reduced below 0.5% and was usually 0% when observation time was limited to 12 minutes or less. The mean error rate was $\leq 2.2\%$ for all categories except for shocked eggs erroneously scored as live. The more elevated error rate in this category only exceeded the typical rate in two trials (8-13%) for unknown reasons. However, misclassification of shocked eggs as live is irrelevant in typical studies, where the objective is to distinguish live and dead eggs before sampler disturbance, so the tradeoff between less accurate separation of live and shocked eggs for more accurate separation of shocked and dead eggs is desirable.

If precise discrimination between live and shocked eggs is necessary, pink-colored eggs can be placed back into water after other data are collected. For example, shocking provides a way of discriminating between infertile and developing eggs, so extended hydration (an hour or more) can improve this measurement.

Finally, unlike the test circumstances in this study where observers were not allowed to compare samples, routine comparison of data among observers should also reduce overall error rates. Site-specific factors can complicate observation, such as intrusion of salt water, which causes dead and shocked eggs to become translucent orange instead of white. (Eggs were no easier to classify in saltwater than freshwater; unpublished observations.) Communication among observers is undoubtedly helpful in field studies.

In conclusion, we recommend that eggs obtained by hydraulic pumping be classified as live, shocked, and dead. Eggs should be quickly removed from water to arrest color changes, and classification should be prompt to limit overall error rates. The percentage of eggs damaged by mechanical shock may potentially provide valuable insight into run timing and egg superimposition in wild runs, although a record of percent eyed eggs may provide the same information. At a minimum, live eggs and shocked eggs should be combined into one “live” category, distinct from eggs dead before observer disturbance if the objective is to study *in situ* mortality. Combining shocked and dead eggs does not accurately reflect pre-sample conditions.

Table 3.1. Experimental time line and estimated stage of development. Development stage was estimated from Smirnov (1975) by comparing measured thermal degree-day units (Deg. Days). “Eyed” embryos were first recorded on October 24 at 310 degree-days, confirming the prediction accuracy at this time.

Trial	Date	Day	Deg. Days	Stage of development
	09/18/2000	0	0	Fertilization
1	09/26/2000	8	77	Embryonic streak to 2 mm embryos
2	10/03/2000	15	142	About 40 somite segments; head & trunk are differentiated. Pericardial cavity is present; caudal bud forms.
3	10/10/2000	22	200	Blood stream; blood may be pigmented. Probably some tail movement. Eye pigmentation about to begin. Two pairs of gills are either supplied with blood or will be shortly.
4	10/17/2000	29	257	Eyes will soon become gray.
5	10/24/2000	36	310	Eyes large, black, iridescent.
6	10/31/2000	43	359	Pectoral fins become capable of movement about now.

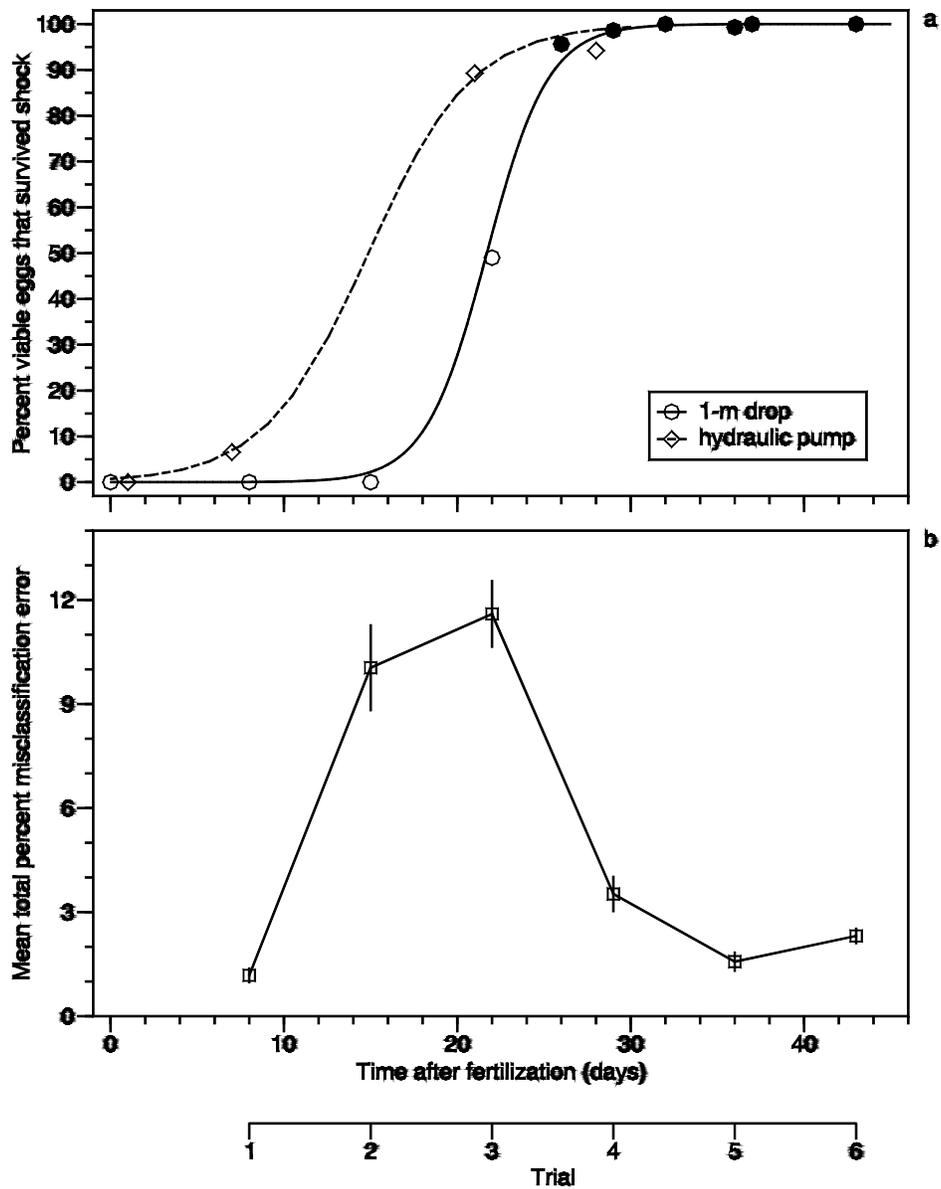


Figure 3.1. Resistance of embryos to mechanical damage (shock) (upper panel) and total percent misclassification as time after fertilization increased (mean \pm SE) (lower panel). In the upper panel, only observations with solid symbols were replicated (error terms were smaller than these symbols); shock resistance of eggs to hydraulic pumping are included for comparison (Thedinga et al. In press).

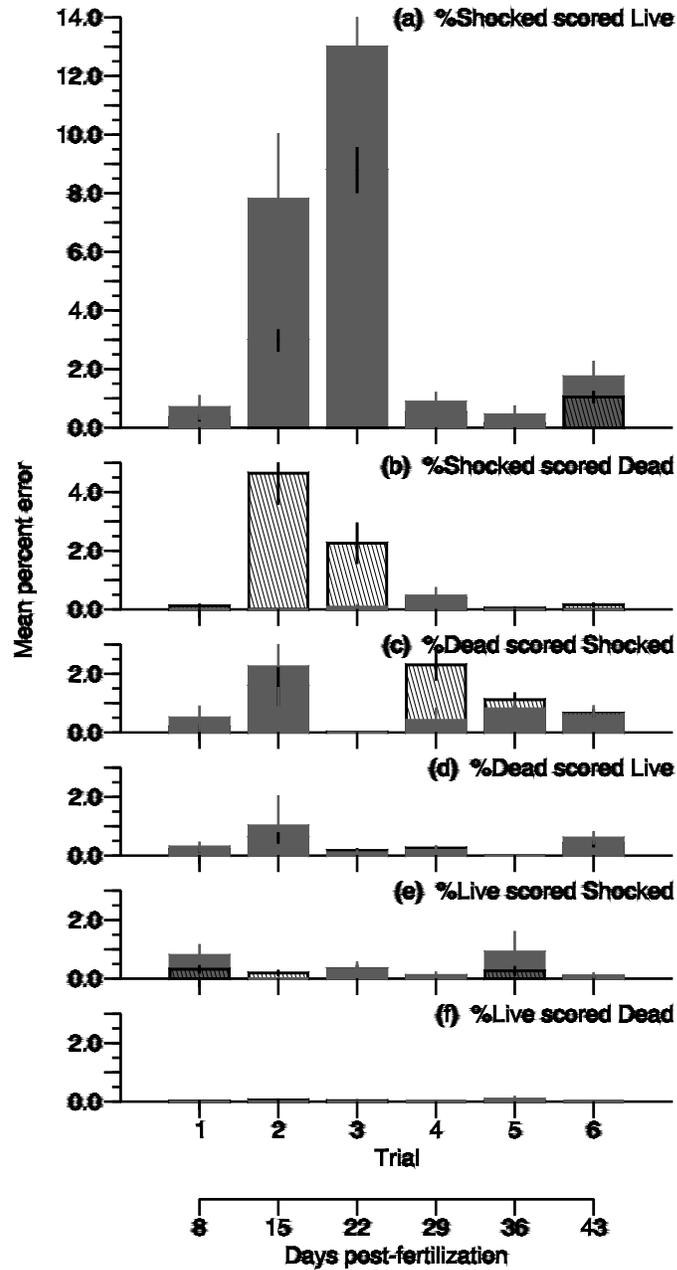


Figure 3.2. Types and frequency of observer classification errors. Hatched bars are means (\pm SE) for all observation times combined, 5-60 minutes. Gray bars are means (\pm SE) where the maximum time between mechanical damage and observation was restricted to 12 minutes or less.

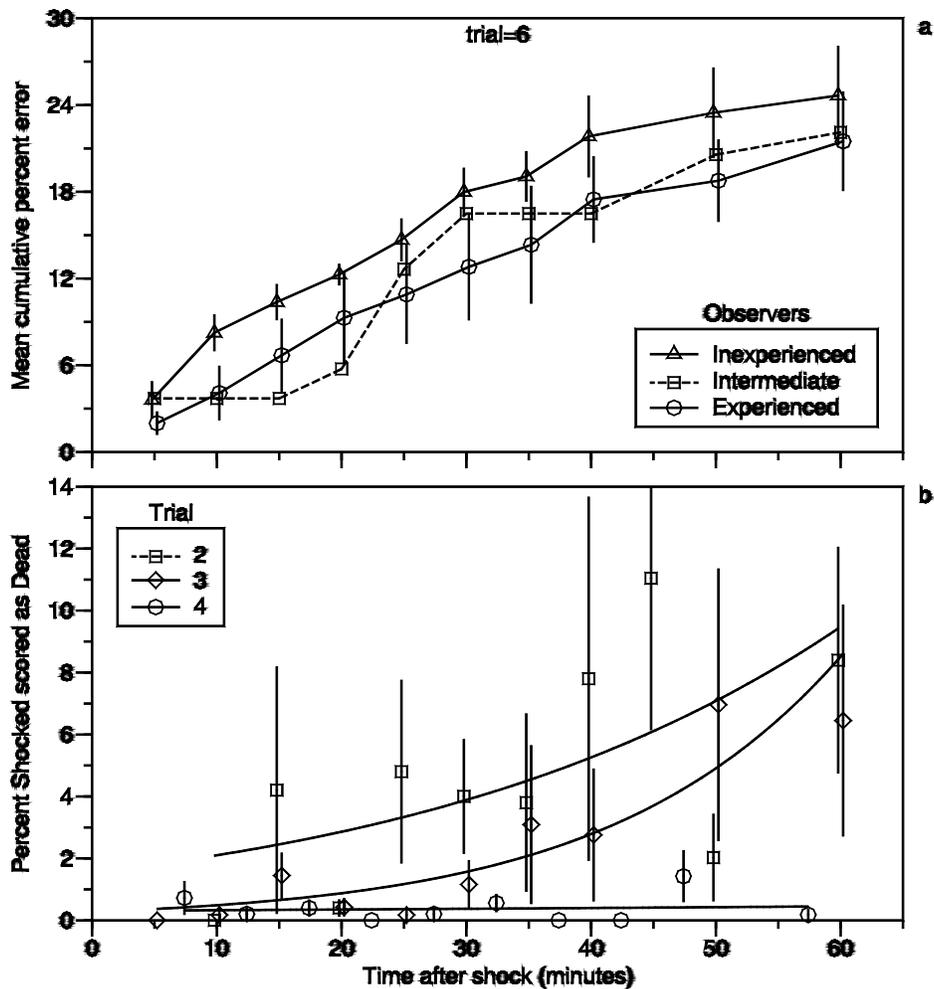


Figure 3.3. Example mean (\pm SE) time related changes in a) cumulative percent classification error by inexperienced, experienced, and intermediate observers and b) percentage of shocked eggs scored as dead. In b), lines are logistic regression fits and results for trials not illustrated were similar to trial 4 results.

Chapter 4

Modeling Prince William Sound Embryo Data

Jacek M. Maselko, Mark G. Carls, and John F. Thedinga

Introduction

The conclusion that pink salmon (*Oncorhynchus gorbuscha*) embryos were damaged by oil after the 1989 *Exxon Valdez* oil spill in Prince William Sound (PWS), Alaska (Rice et al. 2001), has been controversial because some researchers (Brannon and Maki 1996; Brannon et al. 2001) suggest that sampling artifacts or other uncontrolled variables account for the differences in embryo survival between oiled and reference streams. Bue et al. (1996; 1998) and Craig et al (2002) found that embryo mortality was greater in the intertidal reaches of oiled streams than in reference streams from 1989-1993 and again in 1997. One of the alternative reasons Brannon and Maki (1996) and Brannon et al. (2001) suggest was responsible for different embryo mortality between oiled and reference streams was that adult run timing may have varied systematically between stream types, thus embryo susceptibility to the mechanical shock (caused by hydraulic sampling) also varied systematically, and that the technicians who classified the eggs failed to recognize the difference between eggs that died from natural causes and those killed by sampling procedures. Attempts more than a decade after preservation to directly gauge the veracity of the original egg classification by inspecting samples failed because the collection was too degraded to be accurately interpreted.

One way to indirectly discern the accuracy of the original embryo classification by Bue et al. (1996; 1998) is to estimate how much error would be required to eliminate the statistical difference between oiled and reference streams and compare this estimated error with classification errors in contemporary studies designed to study classification errors and the resistance of embryos to mechanical shock in natural systems (Thedinga et al., Chapter 2; Carls et al., Chapter 3). In our model, the misidentification rate between oiled and reference streams was adjusted until mortality differences were not significant, thus providing an estimate of potential bias necessary to explain the original observations. The plausibility of this estimated potential bias was then compared to known misclassification rates.

Methods

We developed a model to test for the misidentification of eggs shocked and killed by the egg pumping procedure used in PWS. By adjusting the misidentification rate in oiled and non-oiled streams so that there was no longer a significant difference between the streams, we

were able to compute the potential field sampler bias. We could then test in the field whether that level of misclassification of eggs shocked and killed by the egg pumping procedure was realistic.

We used a GLM two-factor model based on the intertidal zones where the eggs were collected by Bue et al. (1996; 1998) to account for tidal effect and adjusted the proportion of eggs that were classified as dead or live in all streams. Significant differences between oiled and reference streams in PWS were based on an acceptance level of $P < 0.05$. We adjusted the proportion of misclassified dead eggs until the difference between oiled and referenced streams was insignificant ($P > 0.05$) by the GLM model.

Results & Discussion

Modeled differences in egg mortality between oiled and non-oiled streams became non-significant ($P > 0.05$) when we assumed that 9.5% of eggs in all of the oiled streams were incorrectly counted as dead, but were actually killed by egg pumping and should have been counted as live. Conversely, in the reference streams, 11.3% of dead eggs would have to be incorrectly counted as live before statistical differences between stream types disappeared. A middle-of-the-road model might suggest that a portion of live eggs in oiled streams were misclassified as dead and a portion of dead eggs in reference streams were misclassified as live although the possibility that systematic classification bias was opposite between stream types seems remote.

The likelihood that true misclassification errors in the study by Bue et al. (1996; 1998) were as high as the modeled values is small. In a laboratory study specifically designed to detect and quantify observer bias (Carls et al., Chapter 3), less than 1% of live eggs were misclassified as dead in six individual trials (grand mean across all trials was $0.27 \pm 0.04\%$, $n=599$). This low error rate virtually eliminates the possibility that the results of Bue et al. (1996; 1998) could be explained by such a bias in reference streams. At worst, misclassification of shocked eggs as dead in the study by Carls et al. (Chapter 3) was only about half that required to bias the results of Bue et al. (1996; 1998), and was typically much smaller. The highest mean rate in a trial was $4.6 \pm 1.1\%$; the grand mean was $1.3 \pm 0.2\%$ ($n = 599$). Furthermore, when observation times were restricted to 0.5 h, mean misclassification of shocked eggs as dead was less than 3% in all 6 trials. Although error rates would be expected to be less in a laboratory situation compared to the field, careful and timely handling of eggs in the field can minimize error (Thedinga et al., Chapter 2). The time between collection and classification in the study by Bue et al. (1996; 1998) was typically less than 0.5 h until 1995 when the time was reduced to less than 5 minutes, thus we suggest the likelihood that this type of error forced statistical significance in PWS is small.

Support is weak for the hypothesized systematic differences in embryo sensitivity to shock between oiled and reference streams (Brannon and Maki 1996; Brannon et al. 2001). Estimated run-timing differences in all years except 1990 and 1992 were typically small and ranged from 2 to 4 d (Craig et al. 2002). The slow change in shock resistance in natural systems (Thedinga et al., Chapter 2) suggests these small differences in timing are of little importance. Craig et al. (2002) do identify one measure, the difference between spawn timing and sample date, that supports the possibility their observations were affected by sampling date, but they also found that the lack of specific run timing data in most years did not allow conclusive determination of the magnitude of the timing effect. The conclusion that embryo survival was lower in oiled streams than in

reference streams was unchanged when sample timing was included in the analysis as a covariate in the one year (1991) when run-timing was monitored with rigor in many PWS streams (Craig et al. 2002). Furthermore, as indicated by modeling, the likelihood that shocked eggs were systematically misclassified as dead is also small.

In contrast, there is ample support for an alternative hypothesis, that embryos in oiled PWS streams were damaged by exposure to *Exxon Valdez* oil. These studies document 1) oil in pink salmon habitat, 2) exposure of pink salmon embryos to oil, 3) high oil toxicity, 4) dissolution of toxic oil compounds into surrounding water, 5) cyclic drainage of groundwater from intertidal beach sediment into water below stream channels where pink salmon eggs incubate, and 6) detection of *Exxon Valdez* oil in the water of one contaminated stream a decade after the spill. Both industry-sponsored and government researchers documented that beach substrate surrounding intertidal streams utilized by pink salmon was contaminated by *Exxon Valdez* oil (ADFG 1989; Brannon and Maki 1996; Bue et al. 1996; 1998; Geiger et al. 1996; Murphy et al. 1999; Craig et al. 2002). Although flowing freshwater precluded direct oil deposition in stream channels, cytochrome P4501A was induced in pink salmon embryos incubating in these streams (1989-1991; Wiedmer et al. 1996), an early direct indication of embryo exposure to oil. *Exxon Valdez* crude oil proved to be very toxic; in laboratory studies designed to emulate the composition of polynuclear aromatic hydrocarbons (PAH) observed in PWS, lowest observed effective concentrations were 4-18 µg/L total PAH for pink salmon (Marty et al. 1997; Heintz et al. 1999). This high toxicity results from the PAH present in contaminated water and is consistent with earlier findings that aromatic hydrocarbon toxicity increases with molecular weight and alkyl-substitution (Rice et al. 1977, Neff 1979, Black et al. 1983). Heintz et al. (1999) demonstrated that PAH were transferred from oiled substrate to eggs by water; results for eggs in effluent water and those in direct contact with oiled substrate were not significantly different. Tidal-driven oscillating flows cause advective transfer of chemicals between groundwater, rivers, and oceans (Li et al. 1999) and groundwater moves rapidly from the coarse beaches surrounding pink salmon streams in PWS into the hyporheic zone where pink salmon eggs incubate as a result of tidally driven hydraulic gradients (Carls et al. 2003). Thus, PAH were transferred from oiled beach substrate to incubating pink salmon eggs and in 1999, PAH consistent with *Exxon Valdez* oil were present in hyporheic water of one of six heavily oiled PWS streams, a location where oil was also detected in surrounding beach substrate (Carls et al. 2004b) Additionally, Bue et al. (1998) discounted the possibility that other uncontrolled natural differences among oiled and reference habitat caused differential mortality. These studies support the conclusion by Rice et al. (2001) that exposure to oil increased the incidence of pink salmon embryo mortality.

We conclude that the results of this modeling exercise do not support the differential shock susceptibility hypothesis, but rather are consistent with the hypothesis that oil toxicity increased embryo mortality in PWS streams.

Chapter 5

Conclusions

The resistance of pink salmon embryos to mechanical damage (“shock”) increases over the course of a spawning run and as embryos mature. Unlike in eggs of uniform age where shock resistance increases quickly, shock resistance in naturally spawned eggs increases gradually because continued spawning ensures non-uniform embryo ages and, a proportion of developing embryos are continually replaced by newly spawned ones. Therefore, by sampling at different times of a run, embryos of different maturity levels will be encountered resulting in different shock resistances. If samplers are trained to differentiate shocked eggs from those that died previously and sampling procedures are optimized to allow discrimination, the time of sampling should not affect identification of different egg conditions (live, shocked, or dead).

In our study, the rate of misclassification of shocked eggs as dead was $\leq 4.6 \pm 1\%$ when observation times were unrestricted (≤ 60 minutes) and was reduced to $< 0.5\%$ when observations were limited to ≤ 12 minutes. Conversely, $\leq 2.3\%$ of dead eggs were misclassified as shocked eggs. Both experienced and inexperienced observers participated in our study, and they classified a wide range of egg conditions that were probably similar to those observed by ADFG in PWS. Inexperienced observers were easily trained to discriminate egg condition within an hour, thus they can be expected to provide accurate data in field settings. Thus, if there had been a similar error rate in PWS, there would be insufficient bias to change Bue’s (1996) findings that embryos exposed to oil had reduced survival.

In a modeling study designed to retrospectively explore the veracity of the original egg classifications of Bue et al. (1996; 1998), we found the likelihood was small that potential misclassification errors by Bue et al. were sufficient to explain observed survival differences between oiled and reference streams. For example, modeled differences in egg mortality between oiled and non-oiled streams became non-significant ($P > 0.05$) when we assumed that 9.5% of eggs in all of the oiled streams were incorrectly counted as dead. However, the highest mean misclassification rate in laboratory trial was $4.6 \pm 1.1\%$ of shocked eggs scored as dead and were typically smaller ($1.3 \pm 0.2\%$). When observation times were restricted to 0.5 h (the approximate upper limit in Bue’s study), mean misclassification of shocked eggs as dead was less than 3% each of 6 trials. Thus we suggest the likelihood is small that systematic classification bias between oiled and reference streams caused the observed mortality differences in PWS. Rather, we find the hypothesis that oil was directly responsible for mortality compelling; the reasons for high polynuclear aromatic hydrocarbon toxicity are known, and the mechanisms for delivering oil from surrounding substrate to incubating pink salmon embryos have been demonstrated. We agree with Rice et al. (2001) that exposure to oil reduced embryo survival in oiled PWS streams.

Acknowledgments

We thank Brian Woody, Heather Meuret and Robert Bradshaw for counting pink salmon in Lovers Cove Creek. Brian Woody, Talli Leach, Dina Spangenberg, and Mandy Lindeberg assisted with sampling eggs at Lovers Cove Creek and at Little Port Walter hatchery. Our thanks goes to numerous colleagues who were willing to serve as observers in this study. The research described in this paper was supported by the *Exxon Valdez* Oil Spill Trustee Council. However, the findings and conclusions presented by the authors are their own, and do not necessarily reflect the view or position of the Trustee Council.

Literature Cited

- ADFG. 1989. Alaska Department of Fish and Game stream survey maps. Cordova, AK.
- Black, J. A., W.J. Birge, A. G. Westerman, and P. C. Francis. 1983. Comparative aquatic toxicology of aromatic hydrocarbons. *Fund. Appl. Toxicol.* 3:353-358.
- Brannon, E. L., and A. W. Maki. 1996. The *Exxon Valdez* oil spill: Analysis of impacts on the Prince William Sound pink salmon. *Rev. in Fish. Sci.* 4:289-337.
- Brannon, E. L., K. C. M. Collins, L. L. Moulton, and K. R. Parker. 2001. Resolving allegations of oil damage to incubating pink salmon eggs in Prince William Sound. *Can. J. of Fish. and Aquat. Sci.* 58:1070-1076.
- Bue, B. G., S. Sharr, S. D. Moffitt, and A. Craig. 1996. Effects of the *Exxon Valdez* oil spill on pink salmon embryos and preemergent fry. *Am. Fish. Soc. Symp.* 18:619-627.
- Bue, B. G., S. Sharr, and J. E. Seeb. 1998. Evidence of damage to pink salmon populations inhabiting Prince William Sound, Alaska, two generations after the *Exxon Valdez* oil spill. *Trans. Am. Fish. Soc.* 127:35-43.
- Carls, M. G., S. D. Rice, and J. E. Hose. 1999. Sensitivity of fish embryos to weathered crude oil: Part 1. Low level exposure during incubation causes malformations and genetic damage in larval Pacific herring (*Clupea pallasii*). *Envir. Tox. Chem.* 18:481-493.
- Carls, M. G., R. E. Thomas and S. D. Rice. 2003. Mechanism for transport of oil-contaminated water into pink salmon redds. *Mar. Ecol. Prog. Ser.* 248:245-255.
- Carls, M. G., J. F. Thedinga, and R. E. Thomas. 2004a. Observer classification of live, mechanically damaged, and dead pink salmon eggs. *Trans. Am. Fish. Soc.* 133:245-251.
- Carls, M. G., S. D. Rice, G. D. Marty, and D. K. Naydan. 2004b. Pink salmon spawning habitat is recovering a decade after the *Exxon Valdez* oil spill. *Trans. Am. Fish. Soc.* 133:834-844.

- Collins, K. M., E. L. Brannon, L. L. Moulton, M. A. Cronin, and K. R. Parker. 2000. Hydraulic sampling protocol to estimate natural embryo mortality of pink salmon. *Trans. Am. Fish. Soc.* 129:827-834.
- Craig, A. K., M. Willette, D. G. Evans, and B. G. Bue. 2002. Injury to pink salmon embryos in Prince William Sound – Field monitoring. Restoration Report 98191A-1 *Exxon Valdez* Oil Spill Final Report. Alaska Department of Fish and Game, Anchorage.
- Dvinin, P. A. 1952. The salmon of south Sakhalin. *Izv. Tikhookean. Nauchno-Issled. Inst. Rybn. Khoz. Okeanogr.* 37:69-108. (Translated from Russian; Fisheries Research Board of Canada Translation Service 120).
- Fukushima, M., T. J. Quinn, and W. W. Smoker. 1998. Estimation of eggs lost from superimposed pink salmon (*Oncorhynchus gorbuscha*) redds. *Can. J. of Fish. and Aquat. and Sci.* 55:618-625.
- Geiger, H. J., B. G. Bue, S. Sharr, A. C. Wertheimer, and T. M. Willette. 1996. A life history approach to estimating damage to Prince William Sound pink salmon caused by the *Exxon Valdez* oil spill. *Am. Fish. Soc. Symp.* 18:487-498.
- Hanavan, M. G. and B. E. Skud. 1954. Intertidal spawning of pink salmon. *Fish. Bull.* 56:167-185.
- Heard, W. R. 1991. Life history of pink salmon (*Oncorhynchus gorbuscha*). Pages 121-230 in C. Groot and L. Margolis, editors. *Pacific salmon life histories*. Department of Fisheries and Oceans, Biological Sciences Branch, Pacific Biological Station, Nanaimo, British Columbia, Canada. University of British Columbia Press, Vancouver, British Columbia.
- Heintz, R. A., J. W. Short and S. D. Rice. 1999. Sensitivity of fish embryos to weathered crude oil. Part II. Increased mortality of pink salmon (*Oncorhynchus gorbuscha*) embryos incubating downstream from weathered *Exxon Valdez* crude oil. *Environ. Toxic. Chem.* 18(3):494-503.
- Jensen, J. 1997. Mechanical shock sensitivity in salmonid eggs. Pages 1-3 in C. Clark, editor. *Aquaculture Update 78*. Pacific Biological Station, Nanaimo, British Columbia.
- Jensen, J. O. T. and D. F. Alderdice. 1989. Comparison of mechanical shock sensitivity of eggs of five Pacific salmon (*Oncorhynchus*) species and steelhead trout (*Salmo gairdneri*). *Aquaculture.* 78:163-181.
- Li, L., D.A. Barry, F. Stagnitti, and J.-Y. Parlange. 1999. Tidal along-shore groundwater flow in a coastal aquifer. *Environmental Modeling and Assessment* 4:179-188.

- Martin, R. M. 1973. Hydraulic modification of intragravel streambed flow as a means of improving egg-to-fry survival. MS thesis. Humbolt State University.
- Marty, G. D. and seven coauthors. 1997. Ascites, premature emergence, increased gonadal cell apoptosis, and cytochrome P4501A induction in pink salmon larvae continuously exposed to oil-contaminated gravel during development. *Can. J. of Zoo.* 75:989-1007.
- McNeil, W. J. 1964a. A method of measuring mortality of pink salmon eggs and larvae. U. S. Fish and Wildlife Service Fishery Bulletin. 63:575-588.
- McNeil, W. J. 1964b. Redd superimposition and egg capacity of pink salmon spawning beds. *J. Fish. Res. Board Can.* 21:1385-1396.
- Murphy, M. L., R. A. Heintz, J. W. Short, M. L. Larsen, and S. D. Rice. 1999. Recovery of Pink Salmon Spawning Areas after the *Exxon Valdez* Oil Spill. *Trans. Am. Fish. Soc.* 128:909-918.
- Neff, J. M. 1979. Polycyclic aromatic hydrocarbons in the aquatic environment. Sources, fates, and biological effects. Applied Science Publishers LTD, London.
- Rice, S. D., J. W. Short, and J. F. Karinen. 1977. A review of comparative oil toxicity and comparative animal sensitivity. In Wolfe, D. A. (ed) *Fate and Effects of Petroleum Hydrocarbons in Marine Organisms and Ecosystems*. Pergamon Press, New York, p 78-94.
- Rice, S. D., R. E. Thomas, M. G. Carls, R. A. Heintz, A. C. Wertheimer, M. L. Murphy, J. W. Short and A. Moles. 2001. Impacts to pink salmon following the *Exxon Valdez* oil spill: persistence, toxicity, sensitivity, and controversy. *Rev. in Fish. Sci.* 9:165-211.
- SAS. 1999. SAS OnLineDOC Version 8. SAS Institute Inc. Cary, NC.
- Smirnov, A. I. 1954. The effect of mechanical agitation on developing eggs of the pink salmon *Oncorhynchus gorbuscha* (Walbaum). Salmonidae. *Doklady Akademii Nauk SSSR* 97:365-368. (Translated from Russian; Fisheries Research Board of Canada Translation Service 231).
- Smirnov, A. I. 1975. The biology, reproduction and development of the Pacific salmon. *Izdatel'stvo Moskovskogo Universiteta*, Moscow, USSR. (Translated from Russian; Fisheries Research Board of Canada Translation Service 3861).
- Snedecor, G. W., and W. G. Cochran. 1980. *Statistical Methods*, 7th edition. The Iowa State University Press. Ames, Iowa.

- Thedinga, J. F., M. G. Carls, J. M. Maselko, R. A. Heintz, and S. D. Rice. In press. Resistance of naturally spawned pink salmon eggs to mechanical shock. AK. Fish. Res. Bull.
- Wiedmer, M., M. J. Fink, J. J. Stegman, R. Smolowitz, G. D. Marty and D. E. Hinton. 1996. Cytochrome P-450 induction and histopathology in preemergent pink salmon from oiled spawning sites in Prince William Sound. Am. Fish. Soc. Symp. 18:509-517.
- Williams, D. A. 1982. Extra-binomial variation in logistic linear models. Applied Statistics 31:144-148.